United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
MD-14

EPA-454/R-97-004f
July 1997

EPA

EIIP Volume VI

Quality Assurance Procedures

STAPPA • ALAPCO • EPA
emission inventory
improvement program

U.S. Environmental Protection Agency
Region 5, Library (PL-12J)
77 West Jackson Boulevard, 12th Floor
Chicago, IL 60604-3590

PREFACE

As a result of the more prominent role given to emission inventories in the 1990 Clean Air Act Amendments (CAAA), inventories are receiving heightened priority and resources from the U.S. Environmental Protection Agency (EPA), state/local agencies, and industry. More than accountings of emission sources, inventory data are now providing the prime basis for operating permit fee systems, State Implementation Plan (SIP) development (including attainment strategy demonstrations), regional air quality dispersion modeling assessments, and control strategy development. This new emphasis on the use of emissions data will require significantly increased effort by state/local agencies to provide adequate, accurate, and transferrable information to meet various agency and regional program needs.

Existing emission inventory data collection, calculation, management, and reporting procedures are not sufficient or of high enough quality to meet all of these needs into the next century. To address these concerns, the Emission Inventory Improvement Program (EIIP) was created. The EIIP is a jointly sponsored endeavor of the State and Territorial Air Pollution Program Administrators/Association of Local Air Pollution Control Officials (STAPPA/ALAPCO) and the U.S. EPA, and is an outgrowth of recommendations put forth by the Standing Air Emissions Work Group (SAEWG) of STAPPA/ALAPCO. The EIIP Steering Committee and technical committees are composed of state/local agency, EPA, industry, consultant, and academic representatives. In general, technical committee participation is open to anyone.

The EIIP is defined as a program to develop and use standard procedures to collect, calculate, store, and report emissions data. Its ultimate goal is to provide cost-effective, reliable, and consistent inventories through the achievement of the following objectives:

- Produce a coordinated system of data measurement/calculation methods as a guide for estimating current and future source emissions;
- Produce consistent quality assurance/quality control (QA/QC) procedures applicable to all phases of all inventory programs;
- Improve the EPA/state/local agency/industry system of data collection, reporting, and transfer; and
- Produce an integrated source data reporting procedure that consolidates the many current reporting requirements.

EIIP goals and objectives are being addressed through the production of seven guidance and methodology Volumes. These seven are:

- Volume I: Introduction and Use of EIIP Guidance for Emissions Inventory Development
- Volume II: Point Sources Preferred and Alternative Methods
- Volume III: Area Sources Preferred and Alternative Methods
- Volume IV: Mobile Sources Preferred and Alternative Methods
- Volume V: Biogenics Sources Preferred and Alternative Methods
- Volume VI: Quality Assurance Procedures
- Volume VII: Data Management Procedures

The purpose of each Volume is to evaluate the existing guidance on emissions estimation techniques, and, where applicable, to identify the preferred and alternative emission estimation procedures. Another important objective in each Volume is to identify gaps in existing methods, and to recommend activities necessary to fill the gaps. The preferred and alternative method findings are summarized in clear, consistent procedures so that both experienced and entry level inventory personnel can execute them with a reasonable amount of time and effort. Sufficiently detailed references are provided to enable the reader to identify any supplementary information. Users should note that the number of source categories or topics covered in any Volume is constantly expanding as a function of EIIP implementation and availability of new information.

It is anticipated that the EIIP materials will become the guidance standard for the emission inventory community. For this reason, the production of EIIP Volumes will be a dynamic, iterative process where documents are updated over time as better data and scientific understanding support improved estimation, QA, and data management methods. The number of individual source categories addressed by the guidance will grow as well over time. The EIIP welcomes input and suggestion from all groups and individuals on how the Volumes could be improved.

VOLUME VI

QUALITY ASSURANCE PROCEDURES

January 1997

Prepared by:
Eastern Research Group, Inc.
Radian International, LLC

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

ACKNOWLEDGEMENT

This document was prepared by Radian International, LLC, and Eastern Research Group, Inc., for the Quality Assurance Committee of the Emission Inventory Improvement Program. Members of the Quality Assurance Committee are:

Tom Ballou, Virginia Department of Environmental Quality
Lee Beck, U.S. Environmental Protection Agency, Air Pollution Prevention and Control Division
Denise Bien, Ohio Department of Environmental Services
Victoria Chandler, North Carolina Department of Environment, Health and Natural Resources
Dan Herman, Massachusetts Department of Environmental Protection
Sheila Holman, North Carolina Department of Environment, Health and Natural Resources
William Kuykendal, U.S. Environmental Protection Agency, Emission Factor and Inventory Group
Mena Shah, California Air Resources Board
Dale Shimp, California Air Resources Board
Jim Tomich, Bay Area Air Quality Management District

The other contributor and reviewer was:

Glenn Sappie, North Carolina Department of Environment, Health and Natural Resources

VOLUME VI: CHAPTER 1

INTRODUCTION: THE VALUE OF QA/QC

January 1997

Prepared by:
Eastern Research Group, Inc.

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

DISCLAIMER

This document was furnished to the Emission Inventory Improvement Program (EIIP) and U.S. Environmental Protection Agency by Eastern Research Group, Morrisville, North Carolina. This report is intended to be a final document and has been reviewed and approved for publication. The opinions, findings, and conclusions expressed represent a consensus of the members of the Emission Inventory Improvement Program. Mention of company or product names is not to be considered as an endorsement by the U.S. Environmental Protection Agency.

CONTENTS

Section		Page
1	Background	1.1-1
2	The QA Program and Its Importance	1.2-1
3	Objectives	1.3-1
4	References	1.4-1

FIGURES

Page

1.2-1 Elements of ISO 14001 . . . . . 1.2-5

1

BACKGROUND

The Clean Air Act (CAA) requires state and local air quality agencies to develop complete and accurate inventories as an integral part of their air quality management responsibilities. These air emission inventories are used to evaluate air quality, track emission reduction levels, and set policy on a national and regional scale; however, they are often developed and compiled on a local level by multiple agencies and individuals. Experience with the 1990 State Implementation Plan (SIP) base year inventories brought to light deficiencies and inconsistencies in the inventory development processes now being used. In addition, the current leeway in selecting these processes has resulted in data sets of unknown quality and varying degrees of completeness. More uniform and systematic approaches to collecting and reporting data are needed as well as, standardized procedures and guidance to eliminate variations in interpretation.

To address the problems of the current inventory processes and comply with the CAA, the U.S. Environmental Protection Agency (EPA), in conjunction with State and Territorial Air Pollution Program Administrators/Association of Local Air Pollution Control Officers (STAPPA/ALAPCO), has established the Emission Inventory Improvement Program (EIIP). The EIIP comprises several committees with representatives from state and local agencies, EPA, and industry. Its main goal is to improve the quality of the emissions data collected as well as the manner in which data and information are transferred and shared.

The EIIP Quality Assurance Committee was formed to develop (1) a plan for the EIIP's quality assurance (QA) program, (2) a comprehensive QA source document of methodologies and tools for use in developing emission inventories, and (3) an emission inventory quality rating system. This volume is the EIIP QA source document; it incorporates all products prepared by the EIIP QA Committee including a model QA plan and a quality rating system.

It is important to recognize that good quality assurance/quality control (QA/QC) procedures will only produce results that are as good as the estimation methodology allows. Some emissions estimates are inherently more accurate than others because they are based on well-defined and well-understood processes and/or source-specific data. For example, annual emissions estimates for a boiler with a continuous emissions monitor (CEM) should be of higher quality than estimates based on fuel use and an emission factor. QA/QC procedures are required to ensure confidence in the estimates from both types of methods. However, the QA/QC procedures required for the CEM

data are more detailed and time-consuming than those required for the emission factor approach, just as more effort is expended in acquiring the CEM data.

Because of the different emissions estimating methods that can be used, the EIIP recognized that an inventory quality program had to address both emissions estimation uncertainty and data quality. Uncertainty is largely a function of the estimation methodology. The quality of an estimate is determined partly by the inherent uncertainty of the method as well as by the procedures used to ensure that errors are minimized. Therefore, the EIIP QA Committee worked with the Point, Area, and Mobile Sources Committees to ensure that inherent uncertainties in the emission estimation methods are discussed as fully as possible in each chapter of the appropriate emissions estimation volumes (Volumes II, III, and IV).

In addition, a data attribute rating system (DARS), originally developed by the EPA's Air Pollution Prevention and Control Division (APPCD) as a research tool, was adapted and used to rank the EIIP point and area source methods. The DARS scores provide a means of assessing the relative merits of alternative approaches. The general issues associated with uncertainty were also addressed by the EIIP's QA program. The uses of uncertainty analysis and of rating systems such as DARS are encouraged by the EIIP QA program. These methods can serve as indicators of data quality, be used to identify appropriate estimation methods, and help determine which sources are in need of improvement. Also, Chapter 4 of this volume focuses on the determination and evaluation of uncertainty in emission estimates and the methodology available to do this.

2

THE QA PROGRAM AND ITS IMPORTANCE

QA activities are essential to the development of comprehensive, high-quality emission inventories for any purpose. Furthermore, a well-developed and well-implemented QA program fosters confidence in the inventory and any resulting regulatory and/or control program.

An overall QA program comprises two distinct components. The first component is that of quality control (QC), which is a system of routine technical activities implemented by inventory development personnel to measure and control the quality of the inventory as it is being developed. The QC system is designed to:

- Provide routine and consistent checks and documentation points in the inventory development process to verify data integrity, correctness, and completeness;
- Identify and reduce errors and omissions;
- Maximize consistency within the inventory preparation and documentation process; and
- Facilitate internal and external inventory review processes.

QC activities include technical reviews, accuracy checks, and the use of approved standardized procedures for emission calculations. These activities should be included in inventory development planning, data collection and analysis, emission calculations, and reporting.

The second component of a QA program consists of external QA activities, which include a planned system of review and audit procedures conducted by personnel not actively involved in the inventory development process. The key concept of this component is independent, objective review by a third party to assess the effectiveness of the internal QC program and the quality of the inventory, and to reduce or eliminate any inherent bias in the inventory processes. In addition to promoting the objectives of the QC system, a comprehensive QA review program provides the best available

indication of the inventory's overall quality completeness, accuracy, precision, representativeness, and comparability of data gathered.

A common failure of many inventory development programs is that inadequate resources are devoted to QA/QC activities. A rule of thumb used by many QA professionals is that 10 percent of the in-kind resources of any project should be allocated to QA activities. This does not include the costs of QC, which are assumed to be built into the process.

The actual amount of effort spent in QA/QC of an inventory will vary depending on the desired quality and the complexity of the inventory. However, QA/QC efforts will generally be proportional to the effort expended on emission calculation. For example, estimating an area source's volatile organic compound emissions using simple activity data (such as population) and an emission factor requires relatively little effort for both the calculation and the QA/QC checks. If a survey of local sources is used, resource expenditures for the calculation are increased. Given that more resources are invested in emissions calculations, it is logical to also invest more heavily in ensuring the quality of the data.

It is essential to have a written plan for both the inventory preparation and the QA/QC procedures. Planning includes an assessment of resources and available information. The purpose and end-use of an inventory will dictate the data quality objectives (DQOs). (See Chapter 4 of this volume for more information about DQOs.) The DQOs and available information and resources will determine QA/QC procedures and the scope of the effort.

Simple QA procedures, such as checking calculations and data input, can and should be implemented early and often in the process. More comprehensive (but also more expensive) procedures should target:

- Critical points in the process;
- Critical components of the inventory (e.g., larger or more important sources); and
- Areas or activities where problems are anticipated (e.g., if a complex model is being used for the first time).

Too often, QA activities are concentrated at the end of the inventory process. An effective QA program will include planning, numerous QC checks during inventory development, and QA audits at strategic points in the process. These strategic points need to be identified in the planning stage and will vary somewhat between agencies

and inventories. However, the ideal QA program would include at least one audit conducted after the planning is completed and before the emissions calculations are more than 25 percent completed; another should occur near the end of the process to assure that the final products meet the DQOs. Other audits between these two points are desirable, but the exact scope, timing, and number of audits will depend on the DQOs and resources available as well as the procedures and methods being used to estimate emissions.

Failure to implement and adhere to a QA program will almost certainly lead to undesirable consequences, such as:

- Contamination of subsequent calculations and decisions because of mistakes missed early in the process;
- Increased cost because work has to be redone;
- An incomplete and/or inaccurate inventory even if work is redone;
- Obstruction of the rule-making and enforcement processes;
- Establishment of regulations that are not realistic (emission estimates are based on incorrect emission factors, emissions overestimated or allocated to wrong processes); and
- Embarrassment to all concerned.

Therefore, the EIIP QA program strongly recommends that inventory personnel obtain the commitment of their management to the quality program. This will require the commitment of resources to provide training and proper equipment (e.g., computers) as well as providing sufficient time for the inventory staff to get the work done. This commitment of time and resources will ultimately pay off. In a presentation at a 1994 Air and Waste Management Association conference, Boothe and Chandler (1994), described the careful steps used by the North Carolina Division of Environmental Management (NCDEM) to quality assure emissions data for use in urban airshed modeling. The authors conclude:

> Although the QA process can take significant time and effort . . . [it] will save time ultimately by reducing the processing of invalid emission files. In addition, a thorough QA system ensures confidence in the modeling results . . . [which] provides more confidence in the resulting regulatory decisions.

The critical role played by management in supporting and maintaining quality systems is the core of the emerging environmental management standards under development by the International Standards Organization (ISO). ISO 14000 standards can be categorized into five groups: environmental management systems, environmental audits, environmental performance evaluation, environmental labeling, and life cycle assessment. Because these standards are voluntary, they are not prescriptive. Environmental Management Systems, the first standard (ISO 14001), requires conformance to the elements shown in Figure 1.2-1. The standard states that "top management" must define the organization's environmental policy. Among other things, the policy must ensure continual improvement, must provide a framework for setting and reviewing objectives, and must be documented.

ISO 14000 does not set environmental standards or criteria; rather, it specifies a system for managing environmental quality. As ISO 14000 standards are developed in the future, the EIIP QA guidance should be compatible with the standard. In particular, the EIIP has looked for ways to address the concept of continual improvement in emission inventory quality and to introduce quantitative measures of the quality of emissions estimates so that improvements can be measured.

Environmental Management System Requirements

I. Environmental Policy

II. Planning
- A. Environmental aspects
- B. Legal and other requirements
- C. Objectives and targets
- D. Environmental management programs

III. Implementation and Operation
- A. Structure and responsibility
- B. Training, awareness, and competence
- C. Communication
- D. Environmental management system documentation
- E. Document control
- F. Operational control
- G. Emergency preparedness and response

IV. Checking and Corrective Action
- A. Monitoring and measurement
- B. Nonconformance and corrective and preventative action
- C. Records
- D. Environmental management system audit

V. Management Review

FIGURE 1.2-1. ELEMENTS OF ISO 14001

This page is intentionally left blank.

3

OBJECTIVES

The objectives of the QA source document are to identify, improve, consolidate, and document QA practices and procedures at all steps of the inventory development and review processes. Tools, procedures, and methods useful for inventory QA/QC were identified by surveying inventory experts in government agencies and in the private sector. These procedures are compiled in this document and linked to the appropriate category of sources or stages in the overall inventory process.

This QA source document is intended to be a "living" document that will be updated as needed. For example, several promising and potentially useful techniques for performing sensitivity analyses and statical checks on emissions estimates are presented in this document, but specific details on implementation are not included. The scope of the EIIP QA program does not extend to developing comprehensive "how to" steps for each method. Instead, an overview of methods is provided, with details and references to more detailed studies supplied where available. Future enhancements to this document will depend on the continuation of the EIIP itself, and on feedback from the users of this document.

Although this document focuses on the needs of state and local agencies, it can be used by anyone in government, industry, or research institutions who is concerned about inventory quality (general QA/QC information). For additional specific inventory QA/QC information, refer to the specific QA/QC sections of other EIIP technical documents.

A complete glossary of terms used in the QA/QC process is included in Volume I of this series. The methods and tools presented in this document are designed to reduce the number of procedural and technical errors. A procedural error is caused by the lack of clear and effective management of the QA/QC process including, but not limited to, inadequately trained staff, improper planning, lack of adequate QA, or lack of data tracking and handling protocols. Technical errors are directly related to the methods and technologies used to develop emission estimates. A technical error may result from the incorrect use of spreadsheets or emission inventory software; the use of incorrect data, methodology, and/or assumptions; mathematical miscalculations; or failure to include all emission sources. A good QC program is the best mechanism for minimizing technical errors; QA activities may catch technical errors as well, but less reliably.

Previously published EPA guidance documents have focused primarily on minimizing procedural errors. This volume includes some specific tools addressing technical errors and expands procedural QA tools.

The EIIP quality program is also concerned with providing tools to numerically evaluate emission inventories. One of these tools is the uncertainty analysis, which is an evaluation of the precision and accuracy of an emissions estimate. The most useful uncertainty analysis is quantitative and is based on statistical characteristics of the data such as the variance and bias of an estimate. However, uncertainty can be evaluated qualitatively using expert judgement. Typically, a nonnumerical ranking is used (e.g., "high" or "low" uncertainty).

Another quantitative tool is a sensitivity analysis. The effect of a single variable on the resulting emissions estimate generated by a model (or calculation) is evaluated by varying its value while holding all other variables constant. Sensitivity analyses can help focus QA/QC activities on the data that have the greatest impact on emissions estimates.

The remainder of this source document is organized as follows. Chapter 2, *Planning and Documentation*, discusses the vital role of planning and good QA/QC documentation. The minimum requirements for specific documents and examples are included. A time line showing where specific documents fit in the overall process is provided. Chapter 3, *General QA/QC Methods*, is a compilation of tools, procedures, and methods that can be used for QA/QC that are presented from simplest to most comprehensive. Chapter 4, *Evaluating the Uncertainty of Emission Estimates*, addresses sources of uncertainty and presents methods for evaluating the quality of emission estimates including rating systems and uncertainty analysis. Chapter 5, *Model QA Plan*, demonstrates how some of the information presented in the previous sections can be used in a QA plan. Although this section specifically targets state and local agencies that prepare regional inventories, the plan presented here can easily be adapted to other scales.

4

REFERENCES

Boothe, L., and V. Chandler. 1994. *Quality Assurance of North Carolina Precursors of Ozone Inventories, Emission Preprocessor System and the Urban Airshed Model Output.* Presented at the 87th Air and Waste Management Association Annual Meeting and Exhibition, Cincinnati, Ohio, June 19-24.

This page is intentionally left blank.

VOLUME VI: CHAPTER 2

PLANNING AND DOCUMENTATION

January 1997

Prepared by:
Eastern Research Group

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

DISCLAIMER

This document was furnished to the Emission Inventory Improvement Program (EIIP) and U.S. Environmental Protection Agency by Eastern Research Group, Morrisville, North Carolina. This report is intended to be a final document and has been reviewed and approved for publication. The opinions, findings, and conclusions expressed represent a consensus of the members of the Emission Inventory Improvement Program. Mention of company or product names is not to be considered as an endorsement by the U.S. Environmental Protection Agency.

CONTENTS

Section		Page
1	Overview of QA/QC Planning and Documentation	2.1-1
	1.1 Inventory Categories and Required QA Plan Elements	2.1-4
2	Organization and Staffing	2.2-1
	2.1 Example of External Review of Selected Sources	2.2-4
3	QA Plan	2.3-1
4	Statement of DQOs	2.4-1
	4.1 A Hypothetical Example of DQOs for a SIP Reasonable Further Progress (RFP) Inventory	2.4-3
	4.2 An Alternative Approach to a DQO Statement	2.4-5
	4.3 An Example DQO Approach from the NAPAP Emission Inventory	2.4-7
5	Data Handling	2.5-1
	5.1 Data Gathering	2.5-1
	5.2 Electronic Databases	2.5-4
	5.3 Tracking Data Entry	2.5-4
	5.4 Standardized QA Procedures for Electronic Data Submittals	2.5-5
6	Documentation of Inventory Components	2.6-1
	6.1 Documentation of Calculations	2.6-2
	6.1.1 Documentation of Hand Calculations	2.6-3
	6.1.2 Documentation of Spreadsheet Calculations	2.6-3
	6.1.3 Documentation of Emissions Databases or Models	2.6-4
	6.2 Documentation of QA/QC Procedures	2.6-8
	6.2.1 External QA Review Report of VDEQ Inventory	2.6-8
	6.2.2 QA Review of a State Ozone Precursor Inventory	2.6-8
	6.2.3 QA Review of an Inventory for a Specific Industry	2.6-10
	6.3 Reporting the Inventory Estimates	2.6-13
7	References	2.7-1

FIGURES

Page

2.2-1 Organization Chart for a SIP Inventory Staff 2.2-2

2.3-1 Components of a Comprehensive QA Plan 2.3-2

2.4-1 Inventory Quality Table Used by IPCC/OECD 2.4-6

2.5-1 Example Contact Report 2.5-3

2.5-2 Example Inventory Development Tracking Sheet 2.5-6

2.5-3 Example Data Entry Log 2.5-7

2.5-4 Example Standardized Data Handling Procedure 2.5-8

2.6-1 Documentation of a Spreadsheet Used to Develop Area Source Emissions . 2.6-5

2.6-2 Example Data Reporting Table for IPCC Inventories 2.6-15

TABLES

Page

2.1-1 Inventory Planning, Preparation, and Documentation Steps and Associated QA/QC Activities 2.1-2

2.1-2 Definition of Inventory Levels 2.1-6

2.2-1 QA Coordinator Staffing: Preferred and Alternative Methods 2.2-3

2.2-2 Methods Used to Achieve QA Objectives for VDEQ Inventory Review . . . 2.2-6

2.3-1 Minimum QA Plan and Technical Work Plan Requirements for Inventory Levels 2.3-4

2.3-2 QA Plan: Preferred and Alternative Methods 2.3-5

2.4-1 Preferred and Alternative Methods for DQOs and DQIs 2.4-3

2.4-2 DQO Table for an RFP Inventory 2.4-4

2.6-1 VDEQ Corrective Action Form 2.6-9

2.6-2 Summary of the Survey Data Reviewed for Completeness and Reasonableness in the OCS Production-Related Inventory 2.6-11

This page is intentionally left blank.

1

OVERVIEW OF QA/QC PLANNING AND DOCUMENTATION

Inventory development activities are often limited with respect to time and resources. A key to the planning process is to identify and document these limitations, prioritize inventory-development efforts, and assure that limited resources are effectively budgeted based on priorities. It is vital, therefore, that planning and documentation activities be viewed as integral, not optional. These activities will assure development of the highest quality inventory possible within resource limitations. Planning and documentation are time-consuming in the short term. However, in the long term, good planning and documentation of inventory preparation and quality assurance/quality control (QA/QC) activities will strengthen the reliability and credibility of the inventory.

Planning and documentation are complementary activities, as shown in Table 2.1-1. Documentation of all inventory and QA/QC activities is vital because it provides:

- A record of the planned activities (including QA/QC procedures);
- A statement of the level of quality sought;
- A record of the actual activities;
- Sufficient information to perform the QA/QC activities; and
- A report on the inventory and an assessment of its quality.

Thorough planning helps ensure that the inventory data quality objectives (DQOs) are identified and ultimately met. Inventory planning activities specific to estimating emissions from point, area, mobile, and biogenic sources are discussed in the appropriate volumes of this document series. Volume I of this series provides guidance on planning and documentation for the inventory as a whole. The intent of this chapter is to reinforce the benefits of good inventory QA/QC planning and documentation, while acknowledging that the QA/QC processes selected must be flexible enough to accommodate the agency's needs and goals in developing an inventory within resource limitations. The agency's inventory needs and goals define the DQOs of the inventory.

TABLE 2.1-1

INVENTORY PLANNING, PREPARATION, AND DOCUMENTATION STEPS AND ASSOCIATED QA/QC ACTIVITIES

Inventory Activity	QA/QC Activity
1. Preliminary Planning Activities: • Define purpose and scope of inventory • Define organization and staffing roles	• Define and document DQOs (see Chapter 4 of this volume) • Identify QA coordinator; assign QA/QC responsibilities to inventory staff
2. Prepare Technical Work Plan: • Identify geographical area • Delineate pollutants to inventory • Establish point/area source cutoffs • Prioritize source categories for inclusion in inventory • Prioritize data sources • Delineate emissions estimation procedures	• Prepare QA plan concurrently with or after technical work plan • Document data-gathering methods in QA plan • Choose QA/QC procedures to be used • Determine data quality indicators (DQIs) that will be used to measure quality
3. Prepare Inventory: • Data collection • Data handling • Estimate emissions • Document inventory development activities	• Follow data handling procedures as documented in QA plan • Conduct routine QC activities • Conduct independent QA audits • Document QA/QC steps coinciding with inventory development activities
4. Inventory Reporting: • Document methods, data sources, adjustments • Discuss sources excluded and explain why • Present estimated emissions	• Prepare QA audit reports • Document QC findings and resolution of problems • Discuss QA in final inventory report; prepare separate QA report

QA/QC planning and documentation should not be viewed as optional tasks in preparing inventories; however, the level of effort may vary with the inventory DQOs. In some cases, it may not be necessary to document the planning stages of the inventory in detail. The DQOs and the level of effort required to develop the inventory determine the planning and documentation steps that should be taken. For example, an inventory that is compiled entirely from published data and does not require additional data gathering could be completed with minimal documentation of QA procedures. Such an inventory may be a simple way for an agency to delineate areas for more in-depth research on emissions sources and quantities, for example.

The inventory planning process can be described in detail in a technical work plan in which the inventory preparers typically:

- Specify the geographical area covered and base year;
- Establish and document the inventory DQOs;
- Select pollutants to be included;
- Delineate point/area source cutoffs;
- Prioritize emissions source categories and data needs;
- Identify and prioritize data sources; and
- Describe the inventory methods to be implemented.

Other aspects of a technical work plan address the selection of data handling systems, growth factors that will be used if emissions projections will be needed, and other considerations specific to the inventory's intended use (e.g., modelling).

A technical work plan can be a part of a QA plan, or a separate QA plan can be prepared in addition to the technical work plan. A QA plan contains details on the QA/QC procedures to be implemented throughout the inventory development process. A QA plan also includes a discussion of how the applicability of the data obtained will be assessed and the procedures that will be used to manipulate the data and ensure that QC checks and QA audits are performed throughout the inventory development process. The QA/QC procedures to be implemented vary depending on the ultimate use of the inventory results.

For any type of inventory, documentation is needed to record the information used in data sheets, teleconferences, model inputs, and results. All calculations and

spreadsheets should be clearly documented so they can be reviewed, verified, and easily updated in the future if appropriate. Documentation should be sufficient to allow reconstruction of emissions development activities. Any required reporting that accompanies the inventory should include a compilation of emissions estimates and some type of summary documentation. The inventory report should also document the QA/QC procedures used, even if only to state that the spreadsheets were reviewed by other team members for calculation errors only and to present a discussion of the findings.

1.1 Inventory Categories and Required QA Plan Elements

The intended uses of the inventory drive the documentation needs for the inventory development and QA/QC program. Ultimately, the quality and reliability of an emissions inventory and its documentation are an appraisal of how well it has supported the goals of the program. A detailed QA plan similar to the one presented in Chapter 5 of this volume is not essential for all types of inventories.

Based on their different uses, emissions inventories can be categorized into four general groups. Each of these inventory categories may have slightly different inventory planning, QA/QC, and documentation needs. A report prepared for the United Nations Task Force on Emissions Inventories (Mobley and Saeger, 1994) lists three primary uses of emissions inventories:

- Assessments of air quality problems in an area to identify emissions sources;
- Input for air quality models; and
- Input for regulatory activities associated with policy making.

The above list focuses on the state/local agency perspective. However, other groups develop and use inventories for other reasons. Industrial facilities prepare inventories as part of a permit application or to show compliance with an existing permit. They may also submit annual inventories to be used as the basis for calculating fees. Researchers may also develop inventories to identify sources of pollutants and to use as the basis for modeling or research into mitigation opportunities.

The documentation needs of a research study designed to assess air quality problems are not as stringent as those for the other inventory uses. Top-down inventory development methods are typically used to develop emissions estimates. For an emissions inventory that will serve as input to an air quality model, the methods and activities required to

validate the data are more demanding and require suitable documentation. The baseline inventory data must be source-specific, with detail on the spatial and temporal variability. Emissions inventories used in regulatory activities that define policy options, assess fees, or to demonstrate compliance require the most significant level of documentation. These data could potentially be used in litigation and must therefore stand up to extreme scrutiny.

To date, there have been few attempts to define inventory categories on the basis of quality standards. Table 2.1-2 provides a proposed formal classification of inventories. This classification was derived from the EPA's Air Pollution Prevention and Control Division (APPCD). APPCD delineated four general project categories for field projects and specifies the accompanying QA plan requirements (EPA, 1994). The key point in delineating these categories is that although good QA/QC procedures should be followed in developing any inventory, a detailed record of the planned and implemented activities is not always required. Assigning an inventory to a category level designates what is needed in terms of project staffing, a technical work plan, a QA plan, data handling and tracking, and the level of written documentation needed.

For example, data handling, tracking, and documentation requirements are least stringent for a Level IV inventory, which is usually compiled from previously published emissions data and thus involves no original data gathering. An example of a Level IV inventory is the area source hazardous air pollutant (HAP) emissions inventory that was developed for the Chicago, Illinois area by combining State Implementation Plan (SIP) activity and volatile organic compound (VOC) emissions data with HAP emission factors and speciation profiles (EPA, 1995). The goal of this type of inventory was to quantify HAP emissions in order to evaluate emission reductions from proposed area source VOC regulations. Because the goal was primarily to obtain information and did not directly support rulemaking or compliance, a Level IV inventory was acceptable. No site-specific data were gathered for this effort. In preparing a Level IV inventory, a QA Coordinator should be named, but need not be independent of the inventory staff. A technical work plan should be prepared, but it can be a separate document from the QA plan, or a QA plan may not even be prepared. All calculations should be documented for a Level IV inventory.

A Level III inventory differs from a Level IV inventory because site-specific data of some type are gathered, so more stringent QA and documentation activities are needed. Because the resulting inventory will not be used in direct support of decision making, some flexibility is still available. However, because it may be used to support decision making or to guide future research efforts, a more detailed QA plan is warranted. An example of a Level III inventory is one that is prepared as part of an air pollution control device market potential and performance evaluation. A QA plan that describes the QA/QC procedures to be implemented should be prepared, but it can be separate

TABLE 2.1-2

DEFINITION OF INVENTORY LEVELS

Inventory Levels	Inventory Use	Requirements	Example
I	Inventories supportive of enforcement, compliance, or litigation activities.	Requires the highest degree of defensibility. Generally involves source sampling or mass balance based on site-specific data; performance audits of equipment, traditional QA plan for source sampling activities.	Monitoring for compliance
II	Inventories that provide supportive data for strategic decision making or standard setting.	Site-specific (or region-specific) data are generally required, but not necessarily direct source sampling, performance audits of equipment.	State Implementation Plan (SIP) inventory
III	Inventories developed for general assessments or research that will not be used in direct support of decision making.	May or may not include direct measurement of sources, but often involves site-specific data of some type. QA requirements must be flexible.	Evaluation of effectiveness of alternative controls or mitigation methods; bench-scale or pilot studies
IV	Inventories compiled entirely from previously published data or other inventories; no original data gathering.	Flexible and variable.	Inventory developed for informational purposes; feasibility study; trends tracking

from the technical work plan. It is preferred that a written DQO statement be prepared, and data handling and tracking and calculational procedures should be documented in some fashion.

The minimum QA plan and technical work plan requirements for Level I and II inventories are similar, but less detail is required for a Level II inventory. A SIP inventory is a good example of a Level II inventory; the results of the inventory are used to support decision making, but do not require the same level of defensibility as is needed for a Level I inventory. The primary difference in the QA/QC requirements for Level I and II inventories is that alternative methods regarding staffing and written documentation of DQOs and data quality indicators (DQIs) are acceptable for a Level II inventory but not a Level I inventory.

Level I usually applies to a specific facility or source, and is generally the result of a regulation or litigation. The elements of the QA plan are often specified in the regulation; for example, in the Code of Federal Regulations (CFR), 40 CFR 75 gives specific QA plan requirements for nitrogen oxides (NO_x) and sulfur dioxide continuous emission monitors installed in utility boilers to comply with acid rain provisions of the 1990 Clean Air Act Amendments. Precision and bias determinations are usually required for source sampling data. Other elements of the QA plan include sample custody, instrument calibration, and instrument maintenance requirements. The QA plan requirements are discussed further in Section 3 of this chapter.

This page is intentionally left blank.

2

ORGANIZATION AND STAFFING

An important aspect of QA planning and documentation is the assignment of staff and responsibility. If staff members are clear on their roles and responsibilities, there is less chance of duplication of effort or missed inventory QA/QC steps. Clearly delineating staff roles also allows the inventory director to focus on matching staff capabilities with inventory development needs. The QA staff should have a good understanding of emissions inventory development procedures.

The QA plan identifies key inventory staff and responsibilities. The responsibilities of any outsiders involved in preparing or reviewing the inventory should also be clearly identified. For example, if a state air quality agency is preparing the inventory, consultants, industry personnel, staff from other state agencies, and EPA may be involved in the process.

The overall responsibility for developing the inventory is usually assigned to the agency director. Direct supervision of the inventory preparation process, including making decisions as to the level of effort and funds required to develop the inventory, delineating the DQOs, and evaluating the methods that will be used to create the inventory, is usually the responsibility of an inventory director.

For Level I and II inventories, QA responsibilities are usually assigned to a QA Coordinator. As shown on the example organization chart in Figure 2.2-1, it is preferable to have an independent QA Coordinator who communicates with both the agency and the inventory directors. Ideally, the QA Coordinator should not be a member of the inventory development staff (Table 2.2-1). The QA Coordinator reviews the staff training procedures and conducts QA audits throughout the inventory development process to verify that QC steps are being followed. As discussed in Chapter 1 of this volume, the role of the QA Coordinator is to provide an objective assessment of the effectiveness of the internal QC program and the quality of the inventory, to identify any bias in the inventory process, and to ensure that corrective actions are taken to reduce or eliminate bias.

Flexibility is needed in the QA/QC process so that the availability of staff and resources can be considered. For example, it may not be possible to have an independent QA Coordinator who is not involved in the development of the inventory. This is particularly true for Level III and IV inventories. In these cases, an alternative

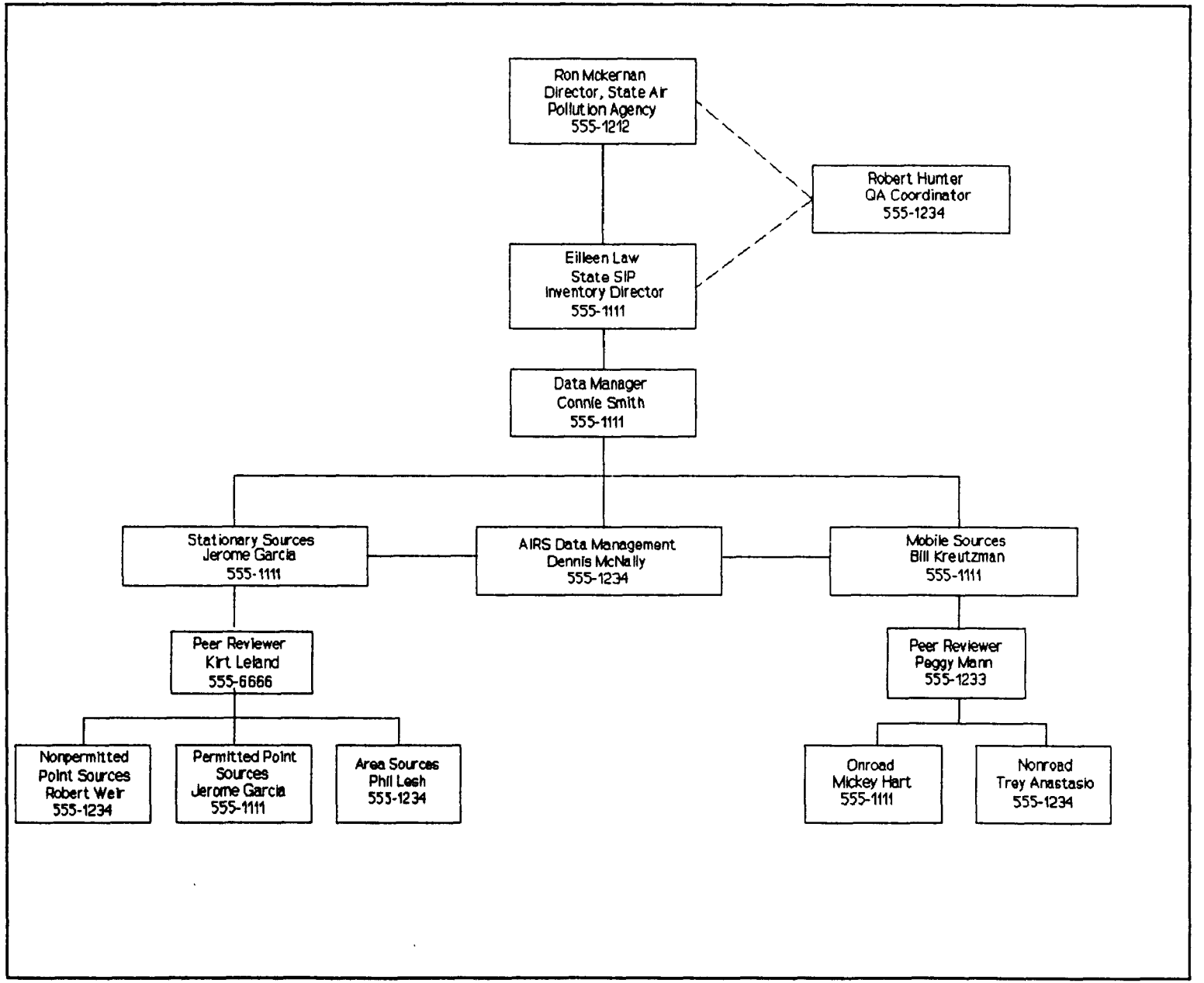

FIGURE 2.2-1. ORGANIZATION CHART FOR A SIP INVENTORY STAFF

TABLE 2.2-1
QA COORDINATOR STAFFING: PREFERRED AND ALTERNATIVE METHODS

Method	Staffing
Preferred	QA Coordinator is independent of inventory staff; conducts periodic QA audits, reviews entire inventory, and prepares QA report. Critical for Level I inventories that may be used in litigation activities. Desirable for Level II inventories.
Alternative 1	QA Coordinator is member of inventory staff; coordinates external review of entire inventory. May prepare some parts of QA report. Is acceptable for Level II, III, and IV inventories.
Alternative 2	QA Coordinator is member of inventory staff; prioritizes selected source categories for external review, coordinates review and incorporation of comments, and prepares QA report. Is acceptable for Level II, III, and IV inventories.

organization can be used. Referring again to Figure 2.2-1, the inventory director, stationary sources lead, mobile sources lead, biogenics lead, or data manager could serve as the QA Coordinator. As such, the QA Coordinator would be responsible for identifying independent peer reviewers for each general inventory category and coordinating an in-depth review. The QA Coordinator is also responsible for ensuring that all peer review comments are addressed satisfactorily.

Because of this reliance on peer review, it is important that QA staff members be chosen for their expertise in a particular area of inventory development. Expert judgement may be needed to determine the quality of the approach if no prescribed method or the required data do not exist, or certain situational factors make the preferred method inappropriate.

If inventory development resources are even more constrained, a second alternative is for the QA Coordinator to prioritize the inventory categories for external review. If this method is chosen, it is very important that the QA reviewer make clear recommendations for corrective actions and that the agency follow up on those recommendations. An external reviewer may not be able to make sure that problems are resolved. Therefore, the inventory director (or a designee) should be made accountable for following through on the recommendations of the external reviewer. A written report should be prepared to document those activities.

Regardless of the distribution of QA responsibilities, it is crucial to impress upon all inventory development and QA staff members the importance of identifying errors throughout the entire inventory development process.

2.1 EXAMPLE OF EXTERNAL REVIEW OF SELECTED SOURCES

The Virginia Department of Environmental Quality (VDEQ) hired an independent consultant to conduct a QA/QC review of portions of its 1990 ozone nonattainment SIP area source emissions inventory. The following objectives were specified for QA activities:

- Ensure that EPA guidance was correctly interpreted and implemented;
- Where EPA guidance was not followed (or was not available), assess the reasonableness of the approach used by VDEQ;
- Ensure the accuracy of input data by verifying data transcriptions from original sources (where appropriate), model inputs, and validity of any assumptions;
- Where appropriate, check the accuracy of spreadsheet calculations by replicating the calculations for at least one county's emissions; and
- If possible and where appropriate, perform some independent comparisons of emissions to other inventories.

The area source categories were selected for review based on the magnitude of the estimated emissions and/or the complexity in estimating emissions. The following area source categories were identified for external review:

- Stationary source solvent use;
- Vehicle refueling and related activities;
- Fuel combustion;
- Incineration and open burning;
- Bioprocesses;

- Waste handling facilities; and
- Leaking underground storage tanks.

In addition, the point source and the mobile source inventories were reviewed. The consultant delineated specific quality objectives that would be covered by the review, as shown in Table 2.2-2. This makes clear the scope of the QA activities, making it easier for subsequent inventory users to identify additional QA procedures that they might want to perform. (Note that this was not a complete QA audit; it is shown as an example of what can be done even when resources are limited.)

All issues encountered during the QA activities were communicated to VDEQ throughout the review. Upon completion of the external review, agency personnel had an idea of the overall quality of the emissions inventory, and could have chosen to have additional source categories reviewed if there had been some question about the data quality. Resolution of the issues was the responsibility of the QA Coordinator on the VDEQ staff.

Other states have used a similar approach, but called on staff from other state agencies to serve as peer reviewers. Regardless of who actually does the review, the key element is that the review be done independently of the actual inventory development.

TABLE 2.2-2

METHODS USED TO ACHIEVE QA OBJECTIVES FOR VDEQ INVENTORY REVIEW[a]

QA Objective	Point	Area	Nonroad	Onroad
1 Ensure correct implementation of EPA guidance.	Peer review of documentation.			
2 Assess reasonableness of VDEQ's approach where EPA guidance not used or unavailable.	• Peer review of documentation.	• Peer review of documentation; • Compared with results from other methods.	• Peer review of documentation.	• Peer review of documentation; • Compared with results from other methods.
3 Ensure accuracy input data	• Not possible since original data forms not provided.	• Checked spreadsheet entries against copies of originals; • Checked accuracy of conversion factors; • Assessed assumptions made to calculate input data (e.g., temperature).	• Checked spreadsheet entries against copies of originals; • Checked accuracy of conversion factors; • Assessed assumption made to calculate input data (e.g., engine matching).	• Used MIDAS software to evaluate MOBILE model input data.
4 Check accuracy of calculations by replicating a sample.	• All calculations were done by SAMS and are presumed to be correct.	• Recalculated emissions by hand.	• Recalculated emissions in a spreadsheet.	• Reran MOBILE input model for selected counties.
5 Perform independent comparisons with other inventories or data sources.	• Compared to TRI data; • Cross-checked SAMS and AIRS-AFS data.			• Modeled VMT compared to HPMS; • Emissions compared to those of four other states.

[a] Key to Acronyms:

AFS: AIRS Facility Subsystem
AIRS: Aerometric Information Retrieval System
HPMS: Highway Performance Monitoring System
MIDAS: MOBILE Input Data Analysis System
SAMS: SIP Air Pollution Management Inventory Subsystem
TRI: Toxic Release Inventory

3

QA PLAN

Prior to initiating work on any emissions inventory, it is imperative that inventory development procedures and QA/QC procedures be agreed upon and documented. Good documentation and involvement of QA personnel during the planning stages enhances the effectiveness of the QA/QC program and decreases the number of quality concerns found during the audits because expectations are clearly outlined in the QA plan and discussed with inventory development personnel prior to starting the work.

Some type of QA plan should be prepared for all inventories, with the level of effort depending on the DQOs of the inventory. The preferred method is to create an integrated technical work plan/QA plan. Chapter 5 contains a model of this type of QA plan for a Level II inventory, and Figure 2.3-1 outlines the key elements to be included. Table 2.3-1 shows which QA plan elements should be documented--at a minimum--for each category of inventory. The use of additional QA plan elements for Levels II, III, and IV inventories is strongly recommended. Furthermore, not all inventories will fit neatly into one of these categories. Modifications to the QA plan elements should be made with the end-use of the inventory in mind.

The purpose of the inventory and the inventory DQOs are defined in the introduction of a combined technical work plan/QA plan. The DQOs should be agreed upon by the agency director and the inventory director (setting DQOs is discussed in Section 4 of this chapter).

The program summary is an executive summary of the QA/QC procedures that will be used to ensure the quality of the inventory. The summary highlights the interaction among functional groups, explains the flow of data through the agency, identifies the points in the inventory procedures where QC is applied, and specifies the frequency of QA audits. Inventory constraints to be acknowledged include limitations on time, resources, data processing capabilities, and availability of personnel. The impacts of any constraints on DQOs should be projected if possible.

The technical work plan discusses staff assignments and responsibilities, including those of inventory development personnel and the QA Coordinator. It also establishes a commitment to the inventory development and QA/QC processes by delineating the resources required to develop the inventory and indicating how they have been (or will

POLICY STATEMENT

INTRODUCTION
- Purpose of inventory
- DQOs

PROGRAM SUMMARY
- Major program components and technical procedures
- Data flow through agency
- Points where QC procedures will be applied and frequency of QA audits
- Inventory constraints

TECHNICAL WORK PLAN
- Organization
  - Organization chart
  - Discussion of roles
- Resources required/how obtained
- Resource allocation
- Personnel training
- Project documentation requirements
  - Guidelines for those supplying data
  - Guidelines for those using data (data handling)
- Schedule

GENERAL QA/QC PROCEDURES
- Data quality ratings
- QA/QC techniques to be used
- QA/QC checkpoints
- Systems audits
- QA/QC review of entire inventory
  - Ensure no double-counting of emissions
  - Compare to other regional inventories for consistency
  - Completeness determination

FIGURE 2.3-1. COMPONENTS OF A COMPREHENSIVE QA PLAN

INVENTORY PREPARATION AND QA/QC ACTIVITIES

- Planning--technical approach
  - Role of state or local agency personnel; minimum QA/QC activities
  - Role of facility personnel; minimum QC requirements
  - Acceptable methods to estimate emissions; preferred method
- Sensitivity analysis to identify key sources and critical data
- Data collection and handling procedures for agency personnel
- Review of estimates
  - Data integrity QC checks
  - Completeness QC checks; ensure all emissions units at a source are included and that all point source facilities in inventory area are included
  - Consistency and reasonableness QC checks
- Data reporting

CORRECTIVE ACTION MECHANISMS

- Identification of problems during QC process
- Identification during QA process
- Documentation of corrective actions

REFERENCES

FIGURE 2.3-1. CONTINUED

TABLE 2.3-1

MINIMUM QA PLAN AND TECHNICAL WORK PLAN REQUIREMENTS FOR INVENTORY LEVELS

Element	I[a]	II[a]	III	IV
Description of project's purpose, scope, and end uses[b]	✓	✓	✓	✓
DQO statement	✓	✓	✓	
Staff organization and responsibilities[b]	✓	✓	✓	✓
Specification of data gathering/ sampling procedures[b]	✓	✓	✓	✓
Specification of estimation methods[b]	✓	✓	✓	✓
Description of internal QC checks	✓	✓	✓	
Specification of QC checkpoints	✓	✓		
Description of QA procedures to be implemented	✓	✓	✓	
Specification of QA checkpoints	✓	✓	✓	
Systems audits[c]	✓	✓		
Calculation or discussion of DQIs	✓	✓	✓	
Corrective action plan	✓	✓		
QA report	✓	✓		

[a] Although all data elements are recommended for both Level I and II, the amount of detail may vary.
[b] If two documents are prepared, these elements may be addressed in a separate technical work plan.
[c] Specific types of audits will vary (see Chapter 3, Section 8 of this volume); performance audits mandatory for all instruments used to collect data.

be) obtained and allocated among the functional groups. The technical work plan also establishes the agency's commitment to personnel training, project documentation, and schedule requirements.

The QA plan presents the general QA/QC procedures that will be implemented including a discussion of how the data will be quality rated, the QA/QC techniques to be used, and the QA/QC checkpoints. The QA plan also establishes when the QA Coordinator will complete the systems audits. Systems audits that evaluate the documentation and procedures associated with the inventory development activities are discussed. The QA plan describes the steps that will be taken as a QA/QC review of the entire inventory is conducted to ensure that no double-counting of emissions occurs. The QA plan explains that the inventory will be compared to other regional inventories for consistency and to verify that all sources of emissions are included. Along with the QA/QC activities, the inventory preparation steps including planning, data collection, review of estimates and reporting, are described in the QA plan. The QA plan also notes what corrective action mechanisms will be implemented throughout the inventory development process.

In addition to this volume, the EPA has published other information that may be helpful in preparing a QA plan (EPA, 1988, 1989).

One alternative to a combined technical work plan/QA plan is that the QA plan and technical work plan be separate documents. However, the use of a combined document is likely to result in a more cohesive and integrated inventory development and quality program. Table 2.3-2 summarizes the preferred and alternative methods.

TABLE 2.3-2

QA PLAN: PREFERRED AND ALTERNATIVE METHODS

Method	Description
Preferred	Prepare an integrated technical work plan/QA plan prior to initiating inventory development.
Alternative	Prepare separate technical work and QA plans prior to initiating inventory development.

This page is intentionally left blank.

4

STATEMENT OF DQOs

The first step in planning any inventory is to define the purpose and intended use of the inventory. This information will, in turn, be used to determine the DQOs for the inventory as well as the QA/QC requirements.

DQOs are statements about the level of uncertainty a decision-maker is willing to accept. Their purpose is to ensure that the final data will be sufficient for its intended use. DQO statements must identify the end use or intended purpose of the data and the level of uncertainty anticipated in the emissions estimates.

It is very important to recognize that DQOs are method-specific; they are based on what is possible for a given methodology and the quality of the data available. The inventory preparers should look at the historical data. What problems have they had in the past that limited inventory quality? Can these problems be overcome for this effort? If this inventory is for a source or region that has never been inventoried, information and experiences from similar efforts should be studied.

DQOs must be realistic and achievable. However, recognition that the inventory quality is less than desirable should be documented in the DQO statement and discussed further in the QA plan (as a constraint). The impacts of weaknesses in methods, data, or other inventory elements should be included in any discussion of uncertainty (see Section 6 of this chapter and Chapter 4 of this volume).

DQOs should be planned in advance and written down. The DQIs that will be used to measure these objectives should also be specified. Specific methods for defining DQOs are discussed in Chapter 4. A complete DQO statement should address:

- Accuracy (or uncertainty) of emission estimates;
- Completeness;
- Representativeness; and
- Comparability.

The issue of *accuracy* has plagued inventory users since the concept of emissions inventories was introduced. Where emissions are measured directly, statistical measures of bias and precision can be used to qualify data accuracy. However, this is rare in a regional inventory. Emissions are usually estimated using factors and surrogate activity data. In some cases, quantitative measures of uncertainty can be made. Also, relative quality ranking systems (such as the Data Attribute Rating System or DARS discussed in Chapter 4) may be used as a quantitative method. At the very least, a qualitative assessment can be employed. For example, a discussion of data strengths and weaknesses, uncertainties, and other qualifiers will set a level of confidence for the inventory user.

The relevance of the other terms is easily shown. *Comparability* can be defined by the intended use of the inventory. For example, emissions trading programs generally require that the emissions of the sources involved be estimated or measured using similar (or identical) methods. In this situation, comparability of the estimates may be the most important DQO.

Representativeness means that the inventory is representative of the region and sources it is meant to cover; for example, if a regional ozone precursor inventory is being prepared, the categories of sources included should represent all of the major sources of VOCs, carbon monoxide (CO), and NO_x in the region. Likewise, the methods and emission factors should be representative of local conditions. The DQOs would assess how representative a national average emission factor is for certain area source categories, and may lead to the decision to use a survey of local sources instead.

Inventory preparers have always been concerned about *completeness*. The DQOs for an inventory must first establish the reference for assessing completeness. It could be a list of businesses in the area, the list of sources from a previous inventory, or some other standard. The DQOs may state that 90 percent completeness is acceptable. Different sections and subsections could have different targets for completeness; 100 percent of the 100-ton sources but only 80 percent of the remaining sources might be required.

Despite the best intentions of inventory preparers, the development effort is often constrained by schedules, resource limitations, and lack of data. The DQOs for the inventory should be realistic and need to account for any factors that will limit inventory quality. Table 2.4-1 lists the preferred and some alternative methods for DQO statements. Other alternative methods are feasible, and can be made very specific to the needs of the inventory. The important thing is that some thought be given in advance to the desired quality of the product.

TABLE 2.4-1

PREFERRED AND ALTERNATIVE METHODS FOR DQOS AND DQIS

Method	Description
Preferred	Written DQO statement addresses accuracy or uncertainty, completeness, representativeness, and comparability. Quantitative methods are used to document inventory quality DQIs; DQIs are specified in the DQO statement and linked directly to DQOs. Critical for Level I inventories; desirable for Levels II and III.
Alternative 1	Written DQO statement addresses specific criteria, but may include less detailed discussions of accuracy or uncertainty, completeness, representativeness, and comparability, or may exclude discussion of one or more of these elements. Qualitative methods are used for DQIs, and each DQO identified is addressed specifically. The DQO statement provides some guidance on the elements to be considered for each DQI. Acceptable for Levels II and III; desirable for Level IV.

Having determined the DQOs, the next, and often more difficult, step is to identify the DQIs that will be used to measure the progress towards each DQO. Quantitative measures (such as confidence limits, numerical ranking systems, or letter grades) are preferable. However, implementing these is also more difficult. An alternative is to use qualitative DQIs, which may simply be a critical discussion of the inventory's strengths and limitations. Specific methods and examples of DQIs are provided in Chapter 4 of this volume. Table 2.4-1 also summarizes the preferred and alternative methods for DQIs.

4.1 A HYPOTHETICAL EXAMPLE OF DQOS FOR A SIP REASONABLE FURTHER PROGRESS (RFP) INVENTORY

The four general DQO categories described above have been informally used by inventory analysts to review inventory quality. Therefore, this explicit use in a more formal DQO statement is fairly easy to envision. Table 2.4-2 presents a hypothetical summary of DQOs and the minimum DQI values that might be set for an update (or RFP inventory) to a SIP inventory.

TABLE 2.4-2

DQO TABLE FOR AN RFP INVENTORY

DQO	Inventory DQI Target Values
Accuracy/Uncertainty	• Achieve DARS score of ≥0.7 for all area sources contributing >10% of total emissions of VOC or NO_x. • Achieve DARS score ≥0.8 for all point sources ≥100 tons per year (tpy). • Quantify variability of all emissions based on source test data or surveys. • Use expert judgement method to estimate uncertainty for all sources >5% of emissions of any pollutant.
Completeness	• 100% of all point sources ≥100 tpy. • 90% of all other point sources. • Top 15 area sources listed in 1990 base year SIP inventory.
Representativeness	• Counties A, B, C, and D. • 1993 daily ozone season.
Comparability	• Results to be compared to 1990 base year inventory.

Presumably, the inventory preparers have considered the sources to be inventoried and the methods/data available. They have chosen to use DARS to set quantitative targets for inventory quality, but also want a quantitative assessment of the variability or uncertainty.

The inventory preparers have also made the determination that as long as they have included all of the ≥100 tpy point sources, they only need to include 90 percent of the remaining point sources. They are also dropping smaller area sources. Not much is made of representativeness in this case. Only the region and relevant year are specified. However, other attributes could be added such as seasonal considerations, conditions that might require adjustments to some emission factors, or other specifications.

Finally, if this inventory must be comparable to the 1990 base year inventory because it will be used to show reductions (or increases) in emissions, the use of comparable methods may be very important. Otherwise, detailed documentation will be needed to demonstrate that differences in emissions are not simply results of the different methodologies, but result from real changes in activity. Additional guidance may be needed on how to ensure comparability while achieving the other DQOs. For example, if meeting the accuracy DQO requires use of a new (and improved) method, the 1990 estimate may have to be recalculated or adjusted to ensure comparability.

A table such as the one shown in Table 2.4-2 may sufficiently describe the DQOs. However, usually some additional details are needed. These may be provided as text in the DQO statement or may be included elsewhere in the QA plan (see Section 3 of this chapter).

4.2 An Alternative Approach to a DQO Statement

The Intergovernmental Panel on Climate Change (IPCC), in collaboration with the OECD and the International Energy Agency (IEA), has led the development of a series of inventory guidance documents for the preparation of greenhouse gas inventories (IPCC/OECD, 1994a, 1994b, 1994c). The guidelines are to be used by individual countries to prepare national inventories. The guidelines must balance the need for well-documented inventories of known quality with the widely varying resources and technology available to participating countries.

The IPCC has essentially one clearly stated DQO, which is to "use comparable methodologies for inventories" because each country is free to use a range of methods at different levels of detail. The IPCC's approach to ensuring comparability is to establish minimum requirements for reporting data that allow for comparison and identification of differences in the methods used. These reporting requirements include "Minimum Data Tables" for various categories, standardized summary and overview tables, and specific reporting elements.

The IPCC has incorporated other DQOs somewhat informally into its guidelines. In a chapter entitled "Reporting the National Inventory," specific QA/QC procedures are listed (the term "verification" is used rather than QA/QC). The last task listed (IPCC/OECD, 1994a, p. 3.5) is to "prepare a brief self-assessment of the quality of the resulting inventory" Furthermore, specific tables are provided for reporting the inventory quality, as shown in Figure 2.4-1. This table indirectly identifies the data attributes of interest without setting specific objectives or targets for quality. Quantification of uncertainties is strongly encouraged as well, and specific methods are described in the IPCC guidance (see also Chapter 4 of this volume).

Table 7A Overview Table for National Greenhouse Gas Inventories

OVERVIEW TABLE

GREENHOUSE GAS SOURCE AND SINK CATEGORIES	CO_2		CH_4		N_2O		Documen-tation	Disaggregation	Footnotes
	Estimate	Quality	Estimate	Quality	Estimate	Quality			
Total National Emissions and Sink									
1 All Energy (Fuel Combustion + Fugitive)									
A Fuel Combustion									
B Fugitive Fuel Emission									
2 Industrial Processes									
3 Solvent and Other Product Use									
4 Agriculture									
A Enteric Fermentation									
B Animal Wastes									
C Rice Cultivation									
D Agricultural Soils									
E Agricultural Waste Burning									
F Savannah Burning									
5 Land Use Change & Forestry									
6 Waste									

KEY

Estimates		Quality		Documentation		Disaggregation	
code	*meaning*	*code*	*meaning*	*code*	*meaning*	*code*	*meaning*
PART	Partial estimate	H	High confidence in estimation	H	High (all background information included)	1	Total emissions estimated
ALL	Full estimate of all possible sources	M	Medium confidence in estimation	M	Medium (some background information included)	2	Sectoral split
NE	Not estimated	L	Low confidence in estimation	L	Low (only emission estimates included)	3	Sub-sectoral split
IE	Estimated but included elsewhere						
NO	Not occurring						
NA	Not applicable						

FIGURE 2.4-1. INVENTORY QUALITY TABLE USED BY IPCC/OECD
(TABLE HAS BEEN ADAPTED FROM IPCC/OECD AND EDITED TO FIT PAGE)

Rather than stating detailed DQOs and specifying targets for the DQIs, the IPCC approach is to make it as easy as possible for inventory users to assess the comparability of the methods. This approach is not the best way to ensure quality, but in some circumstances it is probably a more realistic approach.

4.3 An Example DQO Approach from the NAPAP Emission Inventory

Another example of documented inventory DQOs is from the EPA's 1985 National Acid Precipitation Assessment Program (NAPAP) emission inventory QA/QC plan (EPA, 1986). The objective of the NAPAP inventory was to compile a comprehensive and accurate inventory of emissions and facilities data from natural and anthropogenic sources for the 1985 base year. The EPA developed the DQOs based on the use of the inventory data. For example, one key use of the data was to support atmospheric modeling activities. This required accurate location of emissions sources both geographically and spatially. The EPA also acknowledged the constraints to the inventory because of a tight schedule, budget constraints, and limited availability of emission factors for some source categories (which will affect accuracy). To overcome the problem of scheduling the resources needed to assist with the resolution of questions raised in QA/QC checks, the EPA developed a computerized routine to check for as many of the NAPAP DQOs as possible. The EPA also prioritized sources to be included in the inventory based on pollutants and pollutant quantities emitted, stack heights, and type of industry (combustion sources, petroleum refineries, etc.).

The definition and specification of DQOs in the NAPAP QA/QC plan does not entirely coincide with the approach presented here. However, the basic concept of data accuracy is addressed. The DQOs focused more on identifying critical data elements (such as ensuring accuracy of geographical location), most of which related to representativeness of the inventory.

This page is intentionally left blank.

5

DATA HANDLING

Data handling is an important but often overlooked element of good QA. Information (or data) can come from many different sources, requiring varying degrees of checking, processing, and storage. The key elements of data handling procedures that need to be addressed in the QA program are:

- Tracking data received from different sources in various formats;
- Documenting and managing corrected data; and
- Checking data after conversion to inventory format.

These procedures can be done manually or by use of computerized databases. The methods selected will depend on several factors including the size of the inventory, the inventory level, the number of calculations to be made, and time and budget constraints. In developing most inventories, both hard-copy and electronic data must be dealt with.

5.1 DATA GATHERING

The data handling section of a QA plan discusses how data will be gathered and how subsequent emissions calculations will be affected. The backbone of any data handling system is the project filing system. The organization of the filing system, specifically the names of files and examples of the contents of each, should be specified in the QA plan. The filing system should be set up so that a newcomer could find all relevant data in a logical order if needed. Could an independent QA auditor, for example, trace the sources of data reported in the inventory through the filing system back to the original source?

Regardless of the inventory category level, pertinent information for data obtained from all sources should be kept in a project file. If data are obtained from facility surveys or site visits, the original survey forms and site visit notes and reports should be kept in the project file. Complete copies of all pertinent source test reports should also be kept in the project file. For data obtained from other media, such as electronic bulletin boards or databases, hard-copy printouts of pertinent data should be kept in the project file (if they are not too cumbersome), along with an electronic copy of the original data. For very large databases, a hard-copy summary or description of the database (including date and contents) should be kept with the electronic copy.

Data or information used to develop assumptions or estimation methods can come from several different sources including:

- Published books, documents, reports, or articles;
- Unpublished documents or reports;
- Personal communication via letter, facsimile, or computer e-mail; and
- Personal communication (spoken).

If the data source is published (and presumably available to anyone), a complete citation of the source should be kept in the project file. If feasible, the pages containing specific data should be copied and kept in the file.

Unpublished data sources require that more information be maintained in the file. It is preferable that the entire document, letter, or facsimile be kept in the file. If this is not possible for larger documents, the relevant pages and cover/title pages should be copied and filed. Computer e-mail (or other electronically transmitted information) should be printed and filed.

Any information obtained by telephone, at a meeting, or by other unwritten means should be recorded in a contact report. Standardized forms will help remind staff to record all pertinent information. An example is shown in Figure 2.5-1.

As data are obtained from external sources, it is important to document standardized log-in procedures and verify periodically that the procedures are being followed. Similarly, anyone taking something out of the file for temporary use should sign it out. An early QA audit is a good way to evaluate data handling procedures.

Even when much of the inventory development is done by an electronic database system, some handwritten documentation is usually needed. Each staff member should be assigned a project-specific notebook for recording all calculations and assumptions. If spreadsheets are used for any part of the inventory development, the project file should contain up-to-date electronic versions. The project file should also include copies of completed data entry forms used if the data are combined into a master database. The project file should also contain a discussion of statistical data handling procedures (if relevant), with written documentation of the assumptions made and how outliers in the data were treated.

QC review by inventory team members throughout the emissions estimation process is crucial. It is at this stage in the process that all assumptions, data entry, and

CONTACT REPORT

Date 6/5/95 Originator Joan Brown

CONTACT BY: TELEPHONE X MEETING OTHER _____

NAME, TITLE, & ORGANIZATION
J.P. Morgan, Ozoneville DEQ
ADDRESS & TELEPHONE NUMBER
541-5555
PURPOSE OR SUBJECT (Give project number if appropriate)
XYZ Industries Title V Permit Application

SUMMARY:

J.P. Morgan (permit engineer) was called to determine if the state had any guidance on how to group emissions from a set of related emission units (see contact report for J. O'Conner).

J.P. indicated that the state will allow grouping of emission units as long as each piece of equipment is listed (such as paint booth, drying oven, touch-up). Emissions from each piece of equipment do not need to be specified individually as long as mass balance is used to estimate emissions from the group (paint used - waste = emissions) and all emissions go to one stack. (He also said if hardware controls were added later, we may have to be more specific about emissions).

ACTION: Distribute this information to the XYZ team.

FIGURE 2.5-1. EXAMPLE CONTACT REPORT

calculations are reviewed for technical merit, and transcription and/or data entry errors are also detected. The key here is to encourage all inventory and QA staff members to identify errors throughout the entire inventory development process. Specific QA and QC methods are described in Chapter 3 of this volume and some examples of documentation for specific methods are given. These completed checklists, log sheets, tables, or reports should be kept in the project file.

5.2 ELECTRONIC DATABASES

Data gathered and entered into an electronic database to develop and store emissions estimates should first be validated. A QA/QC program that deals with electronic databases should:

- Check the accuracy of the data input;
- Ensure that emissions are calculated accurately and in a manner consistent with selected methods;
- Ensure that all emissions units are reported and emissions are calculated correctly; and
- Ensure the overall integrity of the database file.

These objectives are met by technically reviewing the input data (QC review), reviewing the emissions estimation methodology, comparing the results of some emissions estimates with estimates calculated by hand, and developing a checklist of emissions sources based on site visits and data gathering efforts and comparing the list to the emissions sources in the database file. Specific methods for accomplishing these tasks are discussed in Chapter 3. Logs should be maintained to track data for each emissions unit as it is entered into the database and its accuracy is checked. Two examples of systems for tracking data entry and flow are provided below. These examples apply most directly to Level I and II inventories. For Level III and IV inventories, less stringent approaches can be used, provided data entry and manipulation are of high quality.

5.3 TRACKING DATA ENTRY

A fairly typical approach for inventory development requires some data manipulation prior to entry into an emissions calculation program or a computer-based data repository. Sometimes the data will change hands several times during this process. It

is therefore very important that the person responsible for each step and the date of the activity be recorded.

In the first example system for tracking data entry presented, data to be used in estimating emissions for a military base were gathered on site. Emissions for some sources are to be estimated using the Air Quality Utility Information System (AQUIS) designed by Argonne National Laboratory for the U.S. Air Force; for others, emissions must be calculated in spreadsheets or by hand, and then entered in AQUIS. Data entry is handled by two people, one of whom is the database administrator. The data were reviewed for accuracy (using sample calculations, peer review, and reality checks as described in Chapter 3 of this volume) prior to entry into AQUIS. After AQUIS data entry, the data were reviewed again to correct any transcription errors. Each step was recorded on a logsheet that was updated electronically and distributed to staff periodically. An example of this tracking sheet is shown in Figure 2.5-2.

In addition, data entry into AQUIS was tracked in more detail. This was necessary because more than one person could enter data; also, the inventory was changing often with updates or corrections. Figure 2.5-3 gives an example of a data entry log.

This level of manual tracking is very tedious, but is necessary--especially when several people are working on the same database. More sophisticated emissions software may have tracking procedures built in that show when an element was last changed and by whom. Depending on the level of detail provided by these computerized "audit trails," some of the manual tracking may not be needed.

5.4 Standardized QA Procedures for Electronic Data Submittals

State agencies must assemble data from many sources, often passing them through different groups within the agency. A second example of a standardized data handling procedure is provided by the California Air Resources Board (ARB), and is depicted in Figure 2.5-4. ARB has standardized procedures to handle emissions inventory data as they are received by the air quality management districts and ARB. The data are reviewed in a hierarchical fashion, beginning with Level 0 data that are provided directly by the source facilities; the district has little direct knowledge of the data quality.

After district personnel have reviewed the facility data by checking for completeness and accuracy and screening for computational errors, the data are referred to as Level 1a data. When the Level 1a data are converted by the district to standard ARB format they termed Level 1b data. This process allows for another completeness check. After ARB inputs the Level 1b data into the electronic database system (EDS), they become

OZONEVILLE ARSENAL INVENTORY REPORTS				County name:	Ozoneville	
Source Category Progress Log				Inventory manager:	A. Griffith	
Last update:	06/06/96					
		Enter date :				
Source Category Description	Source Manager	Through Review	EDI Entry Started	EDI Entry Complete	QA/QC Complete	Write-up Complete
Emergency generators	B. Fife	x	7/20/94	7/28/94	8/04/94	x
Boilers	B. Fife	x	7/20/94	7/28/94	8/04/94	x
Degreasers	N. Bates					
Paint booths	N. Bates	x	7/20/94	7/28/94	8/01/94	
Welding	N. Bates	x	7/08/94	7/09/94		
Cutback asphalt	B. Fife	x	7/20/94	7/21/94		
Abrasive cleaners	B. Fife	x	7/05/94	7/07/94	8/01/94	x
Spills	B. Fife					
Miscellaneous paint use	B. Fife	x	8/06/94			
Wood cyclones	F. Kruger					
Fire fighter training	F. Kruger	x	7/28/94	7/31/94	8/01/94	

FIGURE 2.5-2. EXAMPLE INVENTORY DEVELOPMENT TRACKING SHEET

	EDI Data Entry Log for:						
		Last Update:	06/06/96				
Date	EDI ID	Source Category	Name of data entry person	Date QC'd	QC Name	Problem? (Y/N)	Describe Problem
7/8/94	24000	IFRTS	B. Fife	8/1/94	A. Griffith		
7/8/94	24001	IFRTS	B. Fife	8/1/94	A. Griffith		
7/8/94	24002	IFRTS	B. Fife	8/1/94	A. Griffith		
7/8/94	24003	IFRTS	B. Fife	8/1/94	A. Griffith	Y	Check throughput next visit
7/19/94	16000	Fuel Loading	G. Pyle	7/26/94	H. Crump		
7/19/94	16001	Fuel Loading	G. Pyle	7/26/94	H. Crump		
7/19/94	16002	Fuel Loading	G. Pyle	7/26/94	H. Crump	Y	EDI not speciating emissions
7/19/94	16003	Fuel Loading	G. Pyle	7/26/94	H. Crump		
7/19/94	16004	Fuel Loading	G. Pyle	7/26/94	H. Crump		
7/19/94	16005	Fuel Loading	G. Pyle	7/26/94	H. Crump		
7/20/94	13000	Cutback Asphalt	H. Crump	7/25/96	B. Fife		
7/20/94	14010	Boilers	H. Crump	7/25/96	O. Taylor	Y	EDI calcs different from hand calcs
7/20/94	14020	Boilers	H. Crump	7/25/96	O. Taylor		
7/20/94	14030	Boilers	H. Crump	7/25/96	O. Taylor		
7/27/94	14040	Boilers	H. Crump	7/25/96	O. Taylor		
7/27/94	14050	Boilers	H. Crump	7/25/96	O. Taylor		
7/27/94	20000	Degreasers	J. Beasley	8/19/96	B. Taylor	Y	Check values entered for fuel usage
7/27/94	20001	Degreasers	J. Beasley	8/19/96	B. Taylor		
7/27/94	20002	Degreasers	J. Beasley	8/19/96	B. Taylor		

FIGURE 2.5-3. EXAMPLE DATA ENTRY LOG

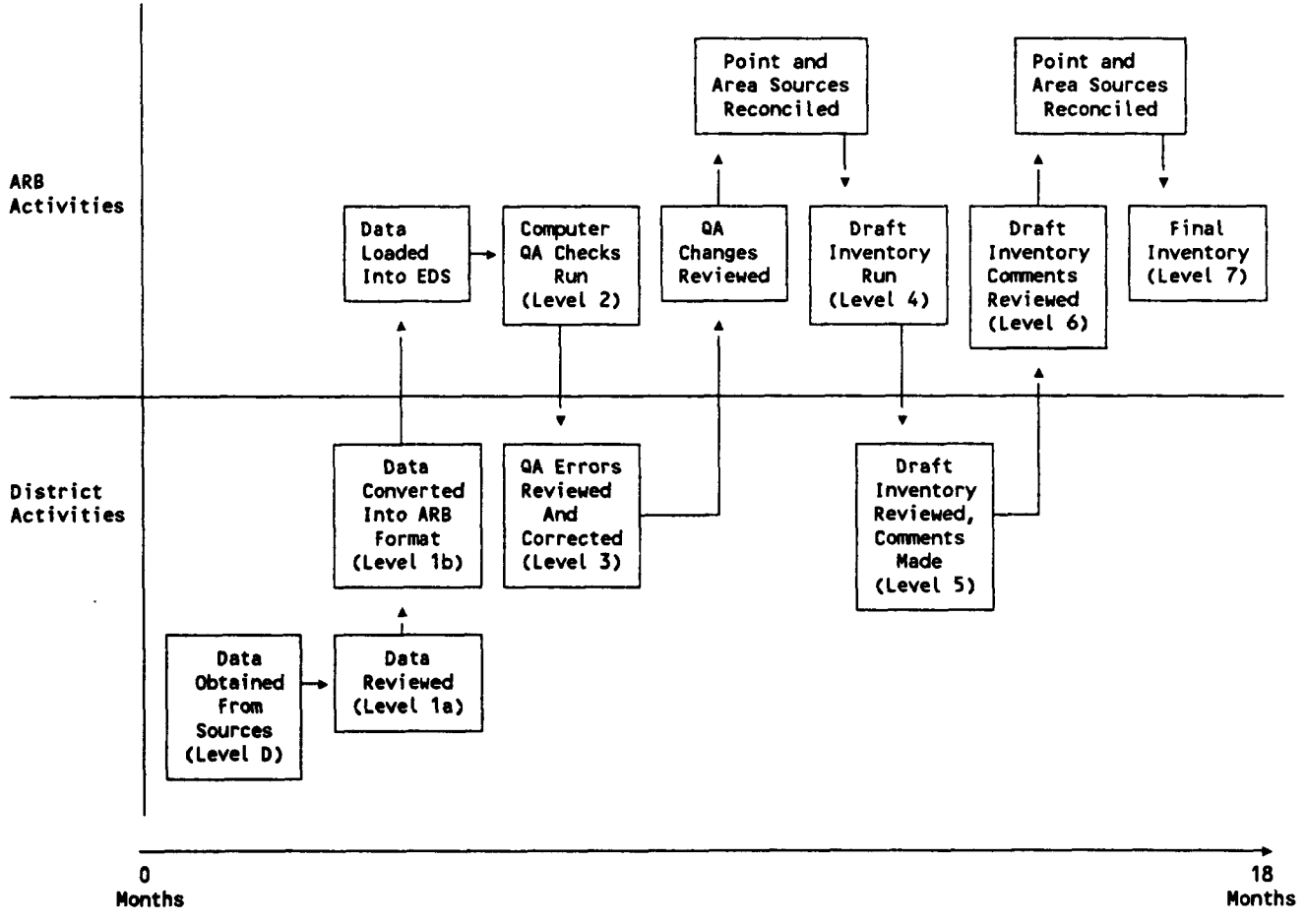

FIGURE 2.5-4. EXAMPLE STANDARDIZED DATA HANDLING PROCEDURE

Level 2 data. Then data are subjected to a variety of computerized QA checks by an independent staff.

Level 2 data are transmitted back to the districts along with a summary of potential errors. There are numerous transmittals to and from ARB (Levels 3-6) as the districts make the necessary corrections and return the data to ARB with recommended changes to the data in EDS. Final inventory data are referred to as Level 7 data.

This page is intentionally left blank.

6

DOCUMENTATION OF INVENTORY COMPONENTS

Previous sections of this chapter have stressed the importance of planning and described the appropriate documentation for each specific procedure. However, written documentation of calculations, assumptions, and all other activities associated with developing the emissions estimates is also a key element of the QA program. Preferred and alternative methods for documenting the planning procedures (technical work plan/ QA plan) and DQOs (DQO statement) are described in Sections 2, 3, and 4 of this chapter.

This section covers documentation of the work that is actually performed during inventory development. The following topics are addressed:

- Documentation of calculations (hand calculations, spreadsheets, databases);
- Documentation of the QA program implementation; and
- Documentation of the results (the inventory report).

As with other elements of inventory preparation, the level of documentation detail required for a specific inventory will vary. The definitions of the Levels I through IV inventory categories given in Section 3 may be used as a guide to the amount of documentation required.

Documenting inventory preparation activities allows the QA Coordinator and others to ensure that the inventory report accurately reflects the data. Examples of topics requiring good documentation in the inventory development process include:

- Point/area source cutoffs to demonstrate that double-counting of emissions does not occur;
- Point source information on survey mailout procedures, tracking and logging of returned surveys, and verification procedures for source test data;

- Adjustments made to source test data to represent longer periods of time, seasonal influences, etc;
- Data obtained from permit and compliance files;
- Adjustments made for applicable rules: control efficiency (CE), rule penetration, and rule effectiveness;
- For area sources in particular, information obtained on emission factors and activity data;
- Data references;
- Adjustments made for local conditions, and assumptions made to adjust for scaling up emissions to account for "nonreported" sources; and
- Mobile source documentation: vehicle miles traveled (VMT), traffic speeds, miles of roadway for each roadway class, hot- and cold-start percentages, vehicle age distribution, etc.

6.1 DOCUMENTATION OF CALCULATIONS

Emissions calculations are generally accomplished using one or a combination of the following methods:

- Handwritten calculations;
- Spreadsheets; and
- Emissions models or databases.

The electronic methods can be very simple or quite complex. However, even when a sophisticated emissions database or estimation program is used, some calculations on the input data may still be needed. At the very least, any assumptions or caveats about the data should be documented. Documentation of calculations should be performed for all inventories, Levels I through IV.

6.1.1 Documentation of Hand Calculations

All calculation sheets should provide the following information:

- The preparer's name;
- Date created (and modified, if applicable);
- Signature of reviewer;
- Date reviewed;
- Citations for all data used;
- List of assumptions; and
- Page number (showing total pages as well).

Calculations should be done in ink (not pencil); any errors should be corrected by drawing one line through the number and writing the correct value above (or nearby). Standardized calculation sheets can be used to prompt staff to remember to provide all of the above information. In addition, written procedures for documentation requirements should be provided, preferably as part of the technical work plan. However, if the agency already has standard guidance or standard operating procedures (SOPs) for documentation, these can be referenced in the technical work plan.

6.1.2 Documentation of Spreadsheet Calculations

Documentation is also important on spreadsheets used to calculate emissions, whether they are part of a formal inventory report or an informal report. The spreadsheet contains all pertinent information used to estimate emissions. The information that should be included on the spreadsheet is similar to that required for handwritten calculations. Because spreadsheet calculations are hidden on the hard copy, additional documentation is needed. The minimum information required is:

- The preparer's name (author);
- Date created (and modified, if applicable);
- Spreadsheet version number;

- Name of spreadsheet reviewer (QC check);
- Date reviewed;
- Citations of all references from which data were obtained (the project file will contain copies of all reference materials);
- All constants, factors, or other data (i.e., no hidden data);
- All calculation documentation (as footnotes or in some other manner); and
- Page number.

A number of alternative spreadsheet designs will fulfill the above requirements. The simplest approach is to show all values used in a calculation as columns (or rows). However, if there are a large number of repetitive values, more concise layouts are better. Figure 2.6-1 gives an example of selected pages from a spreadsheet used to develop area source emissions from gasoline distribution (the entire spreadsheet is not shown). The information at the top serves to document assumptions and values used in the equations; they are also referenced by the appropriate equation in the main table of the spreadsheet. The footnotes at the bottom of the last page also document assumptions and equations used.

6.1.3 DOCUMENTATION OF EMISSIONS DATABASES OR MODELS

Increasingly, emissions inventories are being developed and/or compiled using computerized emissions databases or models. Presumably, the methods, assumptions, and any data included with the software are documented in a user's manual or a technical manual. If not, the user should conduct extensive and careful QA of the model (see Chapter 3, Section 4, "Calculation Checks") or find a better documented system.

Even if the system is well documented, the user will need to provide information about the input data. Comment fields, if available and sufficiently large, can be used to record assumptions, data references, and any other pertinent information. Alternatively, this information can be recorded in a separate document, electronically or otherwise. If at all possible, the electronic database should record a cross-reference to that document. This cross-reference could be a file name (and directory or disk number), a notebook identification number or code, or other document.

	Project number:	123-456-78-99			Special Inputs						
	Project name:	ODE Emissions inventory			1. Dist of OZ Gasoline (Reference 1)						
	Date Created:	08/06/88			County	VMT	Percentage				
	Last modified:	06/06/96			A	8852	17.20%				
	Author(s):	Joseph Smith			B	12804	24.89%				
	File Name:	OZ Gas Distribution			C	7954	15.46%				
	Checked:	Maryanne Jones			D	17471	33.96%				
	Version Number:	4.00			E	4371	8.50%				
	Purpose:	To calculate the VOC emissions for the			OZ Total	51452	100.00%				
		distribution of gasoline in Ozone State.									
			2. Base Emission Factors (Reference 2)								
			2.1 Transit losses (loaded w/ product)				0.005	lb VOC/1000 gal of gas			
			2.2 Transit losses (return w/ vapor)				0.055	lb VOC/1000 gal of gas			
			2.3 Splash underground tank filling				12.18	lb VOC/1000 gal of gas			
			2.4 Submerged underground tank filling				5.04	lb VOC/1000 gal of gas			
			2.5 Vapor balance underground tank filling				0.3	lb VOC/1000 gal of gas			
			2.6 Underground tank breathing				1.0	lb VOC/1000 gal of gas			
Ozoneville Metropolitan AQMA O3 Season: Summary of Emissions From Automobile Gasoline Distribution, SCC 25-01-060- , and Transit Losses, SCC 25-05-030-120											
	(b) OZ State County	(c) 1990 County Gasoline Use (Reference 3) (1000 gal/yr)	(d) Transit Loss (loaded tank truck) (tons/yr)	(e) Transit Loss (empty tank truck) (tons/yr)	(f) Total Transit Losses (tons/yr)	(g) Total Transit Losses (lb/day)	(h) Station Loading Losses (tons/yr)	(i) Station Loading Losses (lb/day)	(j) Station Breathing Losses (tons/yr)	(k) Station Breathing Losses (lb/day)	(Ref 6) (l) Vehicle Refueling (tons/yr)
1. Non-Attainment Areas Within OZ State											
1.1 Ozoneville Metropolitan Area											
	County1	213,097	0.7	7.3	8.0	51.2	149.1	956.0	106.5	585.4	286.6
	County2	308,235	1.0	10.6	11.6	74.1	215.7	1,382.8	154.1	846.8	414.5
	County3	630,114	2.0	21.7	23.6	151.5	405.1	2,597.0	315.1	1,731.1	805.7
	County4	191,479	0.6	6.6	7.2	46.0	134.0	859.0	95.7	526.0	257.5
	County5	39,208	0.1	1.3	1.5	9.4	112.8	723.1	19.6	107.7	169.4
	County6	420,585	1.3	14.5	15.8	101.1	294.3	1,886.8	210.3	1,155.5	565.6

FIGURE 2.6-1. DOCUMENTATION OF A SPREADSHEET USED TO DEVELOP AREA SOURCE EMISSIONS

	(b) OZ State County	(c) 1990 County Gasoline Use (Reference 3) (1000 gal/yr)	(d) Transit Loss (loaded tank truck) (tons/yr)	(e) Transit Loss (empty tank truck) (tons/yr)	(f) Total Transit Losses (tons/yr)	(g) Total Transit Losses (lb/day)	(h) Station Loading Losses (tons/yr)	(i) Station Loading Losses (lb/day)	(j) Station Breathing Losses (tons/yr)	(k) Station Breathing Losses (lb/day)	(Ref 6) (l) Vehicle Refueling (tons/yr)
	County7	105,224	0.3	3.6	3.9	25.3	73.6	472.1	52.6	289.1	141.5
	County8	50,619	0.2	1.7	1.9	12.2	33.5	214.8	25.3	139.1	65.3
	County9	641,903	2.0	22.1	24.1	154 3	412.7	2,645.6	321.0	1,763.5	820.8
	County10	234,536	0.7	8.1	8.8	56.4	155.2	995.2	117.3	644.3	302.5
	Subtotal	2,835,000	8 9	97.5	106.3	681.5	1,986.2	12,732.3	1,417.5	7,788.5	3829.4

2. Spreadsheet Calculation Reference and Examples

(d) Transit Losses, (loaded tank truck) (tons/yr)

=1990 County Gasoline Use (1000 gal/yr) x Transit losses (loaded w/ product, 0.005 lb VOC/1000 gal of gas), D43 =$C43*$H$12/2000

(e) Transit Losses, (empty tank truck) (tons/yr)

=1990 County Gasoline Use (1000 gal/yr) x Transit losses (return w/ vapor, 0.055 lb VOC/1000 gal of gas), E43 =$C43*$H$13/2000

(f) Total Transit Losses, (tons/yr)

= Sum of Transit Losses, (loaded tank truck) and Transit Losses, (empty tank truck), F43 = D43 + E43.

(g) Total Transit Losses, (lb/day)

= [Total Annual Emissions (tons/yr) x 2000 (lb/ton) x ozone seasonal adj.]/[52 (wk/yr) x activity level (days/wk)], G43 = [F43 x 2000 x 1]/[52 x 6].

(h) Station Loading Losses (tons/yr)

Counties where Stage 1 Controls Required

= {[1990 County Gasoline Use (1000 gal/yr) x submerged tank filling (5.2 lb VOC/1000 gal of gas) x

[1-(rule effectiveness, 80%) x (rule penetration for submerged filling)x(control efficiency for vapor balance and submerged fill)]}/2000

H43=(($C43*$H$15)*(1-($T$5*$P12*(($H$15-$H$16)/$H$15))))/2000

Counties where Stage 1 Controls Not Required

= {[1990 County Gasoline Use (1000 gal/yr) x submerged tank filling (5.2 lb VOC/1000 gal of gas)] x [1-(effectiveness of submerged filling, 90%) x (control efficiency)]}/2000

H57=(((($C57*$H$14)*(1-($T$7*(($H$14-$H$15)/$H$14))))/2000

(i) Total Loading Losses, (lb/day)

= [Total Annual Emissions (tons/yr) x 2000 (lb/ton) x ozone seasonal adj.]/[52 (wk/yr) x activity level (days/wk)], I43 = [H43 x 2000 x 1]/[52 x 6].

(j) Station Breathing Losses, (tons/yr)

= 1990 County Gasoline Use (1000 gal/yr) x Station Breathing Losses (1.0 lb VOC/1000 gal of gas), J43 =($C43*$H$18)/2000.

(k) Total Breathing Losses, (lb/day)

= [Total Annual Emissions (tons/yr) x 2000 (lb/ton) x ozone seasonal adj.]/[52 (wk/yr) x activity level (days/wk)], K43 = [J43 x 2000 x 1]/[52 x 6].

FIGURE 2.6-1. CONTINUED

(l) Vehicle Refueling, (tons/yr)
= 1990 County Gasoline Use (1000 gal/yr) x Refueling emission factors from Mike Keenan, NYSDEC, [L43 = C43 * Emission factor from M. Keenan].
(m) Vehicle Refueling, (lb/day)
= [Total Annual Emissions (tons/yr) x 2000 (lb/ton) x ozone seasonal adj.]/[52 (wk/yr) x activity level (days/wk)] [M43 = (L43 x 2,000 x 1.0)/(52 x 7)]
(n) TOTAL VOC Annual Emissions (tons/yr)
= Sum of annual transportation emissions + annual station loading emissions + annual tank breathing emissions + annual refueling emissions, N43 = $F43 + $H43 + $J43 + $L43.
(o) TOTAL DAILY VOC Emissions (lbs/yr)
= Sum of daily transportation emissions + daily station loading emissions + daily tank breathing emissions + daily refueling emissions, O43 = $G43 + $I43 + $K43 + $M43.
3. References
1. Fax. OZ State Dept. of Transportation, Coffee, G., to Environment Corporation, Pacino, A. July 27, 1992. Distribution of Ozone City gasoline use is based on motor vehicle miles traveled within county, p. 18.
2. Compilation of Air Pollutant Emission Factors, Volume 1: Stationary Point and Area Sources, Fourth Edition, EPA AP-42, U.S. Environmental Protection Agency, Research Triangle Park, NC, September 1985 (AP-42), Section 4.4 Transportation and Marketing of Petroleum Liquids, Table 4.4-7, Evaporative Emissions from Gasoline Service Station Operations.

FIGURE 2.6-1. CONTINUED

6.2 DOCUMENTATION OF QA/QC PROCEDURES

QA/QC activities and results should also be documented, either as a part of the inventory report or as a stand-alone QA report. The procedures used to meet the QA/QC objectives, the technical approach used to implement the QA plan, and the results of the QA audits should be documented.

The QA report should summarize the results of all QA activities including key problems found, corrective actions, and any further recommendations. The QA report should also discuss the inventory quality, preferably including quantitative DQIs. If no quantitative measures of quality were planned, then a qualitative assessment of the inventory's strengths, weaknesses, and uncertainties should be provided. More than one report or document may be generated as the result of QA/QC activities. In particular, a report should be prepared for each QA audit. Also, any peer review reports, checklists, forms, or QA/QC tables (or electronic database reports) constitute part of the written records of QA/QC program implementation.

For Level III and IV inventories that may not include a QA report, QA/QC activities should be documented informally (i.e., handwritten notes, comments) and kept as part of the project file. It is important that some written documentation be kept in the event that data quality is questioned.

6.2.1 EXTERNAL QA REVIEW REPORT OF VDEQ INVENTORY

As described in previous sections, an outside consultant was used by VDEQ to review specific elements of the 1990 base year SIP inventory for Virginia.

The VDEQ corrective action form shown in Table 2.6-1 facilitates documentation of QA comments and resolution of issues. This summary of major technical issues found during QA review would also include the name of the VDEQ staff member responsible for resolving each issue, his/her action plan, and proposed date of resolution should be recorded on the form. When the corrective action plan has been completed, the appropriate sections of the inventory should be reviewed again to verify that the emissions estimates are correct. The name of the QA reviewer and the date of the review should be recorded on the form.

6.2.2 QA REVIEW OF A STATE OZONE PRECURSOR INVENTORY

The results of QA checks of North Carolina's ozone precursor inventories were described in a paper presented at a meeting of the Air and Waste Management Association in June 1994 (Boothe and Chandler, 1994). For onroad mobile sources, QA

TABLE 2.6-1

VDEQ CORRECTIVE ACTION FORM

Source Category	Issue	VDEQ Person Responsible	Proposed Date of Resolution	Action Plan	Revised Inventory Reviewed and Approved (signature, date)
AREA SOURCES					
Surface Cleaning	• Use adjusted employment (total-point) to calculate area emissions.				
	• Apply effects of regulations or state clearly that there are none.				
Asphalt Paving	• Use 42 gal/bbl for conversion.				
	• Recalculate emissions without CE.				
Gasoline Tank Truck Unloading (Stage 1)	• Do not apply in counties outside VOC control areas.				
	• Recalculate controlled emissions.				
Wood Consumption	Residential: • Use correct heating value; • Use general wood stove factor rather than catalytic.				
	Commercial/Institutional: • Use correct heating value.				
	Industrial: • Use correct heating value; • Use SAF of 1 (or justify if less).				
Prescribed/Slash Burning	• Recalculate emissions after correcting error in input data.				
Municipal Solid Waste Landfills	• Use emission factor consistent with EPA landfill model.				
Pesticides/Commercial-Consumer Solvent Use	• Correct for double-counting.				

checks of VMT applications and projections included consistency checks of VMT data, evaluation of linear regression analysis determining total rural and urban VMT for all counties in the domain, and review of the disaggregation process of projected rural and urban VMT. QA checks applied to nonroad mobile source spreadsheets verified that the correct EPA-supplied spreadsheet was used, the correct county populations were used, and that the projected NO_x emission factors were adjusted to reflect future NO_x standards.

For area source spreadsheets, QA checks determined if point source adjustments were correctly made and evaluated formulas used to estimate emissions. QA checklists were developed to track and ensure that all sources were reviewed and to document any errors that were found. To QA point source emissions estimates, the EPA SIP Air Pollutant Inventory Management Systems (SAMS) internal QA utilities were used to verify that required fields had proper parameter data entries. When the point source data were in EPA batch transaction format, additional QA checks were performed.

Based on the results of the QA checks implemented by the NC Department of Environmental, Health and Natural Resources (DEHNR) and described briefly above, the authors concluded that although the QA process can take significant time and effort, implementation of a rigorous QA system throughout the entire inventory development process ultimately saves time by reducing the processing of invalid emissions files. A thorough QA system that is well documented will also ensure greater confidence in the modeling results (Boothe and Chandler, 1994).

6.2.3 QA Review of an Inventory for a Specific Industry

Another example of documentation of QA/QC activities is shown in Table 2.6-2. The completeness and reasonableness checks listed in Table 2.6-2 were delineated in a project undertaken to assess air quality impacts associated with the development of outer continental shelf (OCS) petroleum reserves using photochemical modeling (Steiner et al., 1994). Emissions were estimated by soliciting activity and operating data with survey forms. The survey data were entered into Paradox® database tables and input into an inventory system designed in the format required by the EPA's Urban Airshed Model (UAM) Emissions Preprocessor System (EPS). Emissions were calculated using EPA emission factors.

To ensure that the quality of the inventory was technically acceptable, a series of automated checks were developed to review the platform data for completeness and reasonableness. For example, reported latitude and longitude were verified, and data were plotted to identify sources outside the inventory domain. Equipment ratings, annual fuel usage, and annual equipment usage were verified by calculating theoretical

TABLE 2.6-2
SUMMARY OF THE SURVEY DATA REVIEWED FOR COMPLETENESS AND REASONABLENESS IN THE OCS PRODUCTION-RELATED INVENTORY[a]

Field	Completeness Check	Reasonableness Check
Platform		
Longitude & latitude	●	●
Crude/condensate & gas throughput	N/A	●
Percent of production in summer of crude/condensate and gas	●	●
Flare and vent stack height	●	●
Vent stack velocity	●	●
Vent diameter	●	●
Number of liquid fuel oil storage tanks per platform & total liquid fuel oil storage capacity per platform	●	●
Equipment		
Equipment type	●	●
Engine type	●	●
Fuel type	●	●
Annual usage	●	●
Annual fuel usage	●	●
Equipment stack velocity	N/A	●
Equipment stack height	N/A	●
Equipment stack diameter	N/A	●
Annual average load = (annual fuel use * fuel consumption rate)/(annual use)	N/A	●

TABLE 2.6-2

CONTINUED

Field	Completeness Check	Reasonableness Check
Crude Tanks		
Capacity	●	●
Average daily throughput	●	●
Dimensions	●	●
Tank color	●	N/A
Crew/Supply Helicopters		
Longitude & latitude of home airport	●	●
Average number of landings & takeoffs per month	●	●
Average cruising speeds	●	●
Geographical area served	●	●
Monthly hours of operation	●	●
Rated capacity	●	N/A
Fuel usage at rated capacity	●	●
Annual hours in operation	●	●
Crew/Supply Vessels		
Monthly hours of operation	●	●
Time at idle at platforms during hours of operation	●	●
Monthly fuel usage	●	●
Fuel type	●	N/A
Geographic area served	●	●
Annual hours of operation	●	●

[a] ● = data checked; N/A = not applicable or necessary.
Source: Steiner et al., 1994.

annual average loads, and flagging any equipment ratings greater than 200 percent or less than 20 percent of the calculated annual average loads.

Tables were generated of data that failed the completeness and reasonableness checks. Data in the tables were then checked against the original survey form data, and any changes made to the data were documented in the project file. Data corrections were entered into a data correction module within the inventory system. Corrections made typically consisted of unit conversions, which significantly affected the emissions estimates.

This example shows how QA procedures should be tailored to fit the inventory. The list of data reviewed is very specific to the industry. The level of detail gives the user some confidence in the quality of the QA program itself. (This paper also provides a good qualitative summary of uncertainty, which is discussed in Chapter 4 of this volume.)

The QA/QC efforts applied to this inventory resulted in corrections being made to 6,000 records, identification of three omitted platforms, and a decrease of more than two orders of magnitude in total estimated emissions.

6.3 Reporting the Inventory Estimates

Some sort of final report is required, if only to convey the results to interested parties. The simplest report might be a table (or set of tables). This may be all that some end-users want or need.

In most cases, however, detailed inventory and QA/QC reports are crucial. An example of a formal inventory report is one prepared as a SIP submission. This type of report includes summary tables, raw listings of equipment, activity levels and emissions from individual sources, and a QA report. A detailed inventory report allows comparison of baseline inventories from one area to another, the evaluation of the impact of control strategies, and facilitates updates to the inventory and development of projection inventories.

The 1992 EPA report *Example Documentation Report for 1990 Base Year Ozone and Carbon Monoxide State Implementation Plan Emission Inventories* provides examples of formal inventory reporting (EPA, 1992). This document provides guidance for presenting and documenting SIP emissions inventories, and contains examples of how state and local agencies should present and verify inventory development efforts. Example inventory documentation is presented for point, area, mobile, and biogenic sources. The QA documentation section of a formal emissions inventory report should provide enough information to enable comparison of the QA/QC steps completed with

those described in the QA plan. Examples are also presented that illustrate how the procedures used to implement the QA plan should be documented, as well as the results of the QA procedures.

Another example of very specific reporting requirements can be found in the first volume of the *IPCC Draft Guidelines for National Greenhouse Gas Inventories* (IPCC/OECD, 1994a) that specifies the following four documentation standards:

1. "National inventory reports should provide minimum information to enable the results to be reconstructed, and to justify the choice of methodology and data used. This means, for example, that to the extent possible, activity data should be provided at the level of detail at which the emissions are estimated.

2. Documentation should contain enough information to explain differences between national methods and data, and the IPCC default methods and assumptions. Reasons for the differences should be explained and sources of emission factors and other national data should also be clearly cited. Minimum requirements include: emission factors, activity data, and a list of references documenting any differences from IPCC recommendations.

3. Measurement studies containing new values should be referenced, and made available upon request. It is preferable that new emission factor data be contained in published sources.

4. Documentation should be kept for future years (by the country and by the IPCC) and countries are encouraged to publish the documentation of their inventories. This extensive record keeping will facilitate the recalculation of historical inventory estimates when changes in national methods or assumptions occur."

The IPCC reporting guidelines (OECD, 1994a) provide more than 30 tables to be completed and submitted as part of the inventory report. An example inventory quality table was shown in Figure 2.4-1. Another example with one of the data tables is shown in Figure 2.6-2.

5 D I Managed Forests: Annual Growth Increment

Source and Sink Categories			Activity Data	Emissions/Removals Estimates	Aggregate Removals Factors
Forest Type			A Area of Managed Forest (k ha)	B Carbon Removal (Gg C)	C Carbon Removal Factor (Mg C)
					C=B/A
Tropical	Plantations (specify type)				
	Logged	Closed Broadleaf			
		Closed Coniferous			
		Open			
	Other				
Temperate	Plantations (specify type)				
	Commercial	Evergreen			
		Deciduous			
	Other				
Boreal					
			Number of Trees 1000	Carbon Removal (Gg C)	Carbon Removal Factor (Mg C)
					C=B/A
Afforestation Programs					
Village & Farm Trees					

FIGURE 2.6-2. EXAMPLE DATA REPORTING TABLE FOR IPCC INVENTORIES

This page is intentionally left blank.

7

REFERENCES

Boothe, L. and V. Chandler. 1994. *Quality Assurance of North Carolina Precursors of Ozone Inventories, Emission Preprocessor System and the Urban Airshed Model Output.* Presented at the 87th Air and Waste Management Association Annual Meeting and Exhibition, Cincinnati, Ohio, June 19-24.

IPCC/OECD (1994a). Greenhouse Gas Inventory Reporting Instructions. IPCC Draft Guidelines for National Greenhouse Gas Inventories, Volume 1.

IPCC/OECD (1994b). Greenhouse Gas Inventory Workbook. IPCC Draft Guidelines for National Greenhouse Gas Inventories, Volume 2.

IPCC/OECD (1994c). Greenhouse Gas Inventory Reference Manual. IPCC Draft Guidelines for National Greenhouse Gas Inventories, Volume 3.

Mobley, J.D. and M. Saeger. 1994. Procedures for Verification of Emissions Inventories: Report of the Expert Panel on Emissions Verification. Prepared for The United Nations Task Force on Emissions Inventories, Economic Commission for Europe.

Steiner, C.K.R, L. Gardner, M.C. Causley, M.A. Yocke, and W.L. Steorts, 1994. Inventory Quality Issues Associated with the Development of an Emissions Inventory for the Minerals Management Service Gulf of Mexico Air Quality Study. In: *Emissions Inventory: Perception and Reality*, Proceedings of an International Specialty Conference, VIP-38. Air and Waste Management Association, Pittsburgh, Pennsylvania. pp. 327-338.

U.S. EPA. 1995. *Development of Area Source Hazardous Air Pollutant Inventories. Volume 1: Air Toxic Emission Inventories for the Chicago Area.* U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory, Draft Report. Research Triangle Park, North Carolina.

U.S. EPA. 1994. *AEERL Quality Assurance Procedures Manual for Contractors and Financial Assistance Recipients.* U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory. Research Triangle Park, North Carolina.

U.S. EPA. 1992. *Example Documentation Report for 1990 Base Year Ozone and Carbon Monoxide State Implementation Plan Emission Inventories.* U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-92-007. Research Triangle Park, North Carolina.

U.S. EPA. 1989. *Quality Assurance Program for Post-1987 Ozone and Carbon Monoxide State Implementation Plan Emission Inventories.* U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-89-004. Research Triangle Park, North Carolina.

U.S. EPA. 1988. *Guidance for the Preparation of Quality Assurance Plans for O_3/CO SIP Emission Inventories.* U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-88-023. Research Triangle Park, North Carolina.

U.S. EPA. 1986. *Quality Assurance and Quality Control Plan for the NAPAP 1985 Emission Inventory.* U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory, EPA-600/8-86-025. Research Triangle Park, North Carolina.

VOLUME VI: CHAPTER 3

GENERAL QA/QC METHODS

June 1997

Prepared by:
Radian Corporation

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

DISCLAIMER

This document was furnished to the Emission Inventory Improvement Program (EIIP) and U.S. Environmental Protection Agency by Radian Corporation, Research Triangle Park, North Carolina. This report is intended to be a final document and has been reviewed and approved for publication. The opinions, findings, and conclusions expressed represent a consensus of the members of the Emission Inventory Improvement Program. Mention of company or product names is not to be considered as an endorsement by the U.S. Environmental Protection Agency.

CONTENTS

Section		Page
1	Introduction to Methods	3.1-1
2	Reality Checks	3.2-1
	2.1 Overview of Method	3.2-2
	2.2 Comparison to Reference Value	3.2-3
	2.3 Comparison to Other Regions	3.2-4
	2.4 Comparison to Other Categories	3.2-5
3	Peer Review	3.3-1
	3.1 Overview of Method	3.3-2
	3.2 Checklists	3.3-2
	3.2.1 EPA's Level I and Level II Checklists	3.3-3
	3.2.2 Chicago Area Source Toxics Inventory	3.3-3
	3.3 Peer Reviewer Reports	3.3-3
4	Sample Calculations	3.4-1
	4.1 Overview of Method	3.4-2
	4.2 Replication of Spreadsheet or Hand Calculations	3.4-3
	4.3 Replication of Complex Calculations	3.4-3
5	Computerized Checks	3.5-1
	5.1 Overview of Method	3.5-2
	5.1.1 Range and Character Checks	3.5-5
	5.1.2 Look-up Tables	3.5-5
	5.1.3 Pull-Down Menus	3.5-6
	5.2 Internal Checks and Reports: CEIDARS, RAPIDS, Standardized Spreadsheets	3.5-6
	5.2.1 CEIDARS	3.5-7
	5.2.2 RAPIDS	3.5-8
	5.2.3 Standardized Spreadsheets/Databases	3.5-9
	5.2.4 Spreadsheet Auditing Functions	3.5-14
	5.3 Stand-Alone Programs	3.5-14
	5.3.1 EIQA	3.5-14
	5.3.2 Mobile Input Data Analysis System (MIDAS)	3.5-24
	5.3.3 Visual Auditing Programs	3.5-25
	5.4 User-Defined Automated QA Checks	3.5-27
6	Sensitivity Analysis	3.6-1
	6.1 Overview of Method	3.6-2
	6.2 Sensitivity Analysis for an Area Source Inventory	3.6-5

CONTENTS (CONTINUED)

Section		Page
	6.3 Sensitivity Analysis of EPA's Motor Vehicle Emission Factor Model	3.6-9
	6.4 Sensitivity of Air Quality Models to Inventory Input Data	3.6-10
7	Statistical Checks	3.7-1
	7.1 Overview of Methods	3.7-2
	7.2 Descriptive Statistics	3.7-3
	7.2.1 Testing for Normality	3.7-12
	7.3 Statistical Quality Control	3.7-19
	7.3.1 An Example of Acceptance Sampling by Attributes	3.7-21
	7.3.2 Specifying a Sample Plan	3.7-22
	7.3.3 Outlier Analysis: Acceptance Sampling by Variables	3.7-25
	7.4 Statistical Procedures for Comparability Analysis	3.7-28
	7.5 Example of a Statistically Based Comparability Analysis: Regression Model using Dummy Variables	3.7-30
8	Independent Audits	3.8-1
	8.1 Overview of Method	3.8-2
	8.1.1 Adequate Training	3.8-3
	8.1.2 Managerial Support	3.8-4
	8.1.3 Good Planning	3.8-4
	8.1.4 Effective Audit Procedures	3.8-11
	8.1.5 Audit Report	3.8-15
	8.2 Types of Audits	3.8-18
	8.2.1 Management Systems Audits	3.8-19
	8.2.2 Technical Systems Audits	3.8-20
	8.2.3 Performance Evaluation Audits	3.8-22
	8.2.4 Data Quality Audits	3.8-23
	8.2.5 Data/Report Audits	3.8-25
	8.3 Preferred and Alternative Methods	3.8-26
	8.3.1 Independent/External Audit	3.8-27
	8.3.2 Internal Audit	3.8-28
	8.3.3 Internal Data Checking	3.8-29
9	Emissions Estimation Validation	3.9-1
	9.1 Overview of Method	3.9-2
	9.2 Direct and Indirect Measurements	3.9-3
	9.3 Receptor Modeling	3.9-7
	9.4 Inverse Air Quality Modeling	3.9-8
10	References	3.10-1

FIGURES

Page

3.2-1 Comparison of Per Capita Daily VOC Emissions 3.2-6

3.3-1 Example of a Peer Review Checklist 3.3-4

3.3-2 Example of Written Peer Review Report 3.3-5

3.4-1A Example Spreadsheet Calculation 3.4-4

3.4-1B Example of Hand-calculated Check of Spreadsheet Result 3.4-5

3.4-2 Example Hand-calculated Check of Database Algorithm 3.4-7

3.5-1 Example of Standardized Spreadsheet for an Area Source 3.5-11

3.5-2 Example of Standardized Spreadsheet for Point Source Actual Emissions 3.5-13

3.5-3 Point Source Spreadsheet Toggled to Produce Potential Emissions 3.5-15

3.5-4 Example of Visual Tracing of Cells Using a Spreadsheet Audit Function . 3.5-16

3.5-5 Example of Cell Formulas Listing Using a Spreadsheet Audit Function . . 3.5-17

3.5-6 EIQA Point Sources Completeness Report 3.5-20

3.5-7 EIQA Area Sources SIP Rule in Place Consistency Check 3.5-23

3.5-8 Summary of Results from MIDAS Review of the MOBILE5a Input Files 3.5-26

3.5-9 Example of Visual Auditing Tool 3.5-28

3.5-10 Examples of QA Scatter Plots 3.5-30

3.6-1 Percent VOC by Estimation Methodology 3.6-8

3.6-2 Effect of Population, Employment on Area Source VOCs 3.6-9

3.6-3 Sensitivity of NMHC, NO_x Emissions 3.6-12

3.7-1 Histogram Plot of TOG Emissions for Los Angeles, No Data Transformation 3.7-14

FIGURES (CONTINUED)

		Page
3.7-2	Histogram of TOG Emissions from Los Angeles Point Sources, Log Transformation	3.7-15
3.7-3	Example Expected Normal Plot for Los Angeles Point Source Emissions	3.7-16
3.7-4	Expected Normal Plot of Los Angeles Point Source NO_x Emissions without a Log Transformation	3.7-18
3.7-5	Histogram of Los Angeles Log-Transformed NO_x Emissions Data Illustrating a Bimodal Distribution	3.7-20
3.7-6	Probability of Accepting a Dataset as Error Free for Three Different Sample Sizes	3.7-23
3.7-7	Probability of Accepting a Dataset as Error Free when 10% Sample Size is Used	3.7-24
3.7-8	Graphical Comparison of Reported and Predicted VMT Values	3.7-31
3.8-1	Example SIP Emission Inventory Organization Chart	3.8-5
3.8-2	Key Elements of Database Use and Operation	3.8-8
3.8-3	Example Information Flow Chart	3.8-10
3.8-4	Example Data Audit Checklist	3.8-14
3.8-5	Recommendation for Corrective Action	3.8-17
3.9-1	Single Adjustment Factor for All CO Sources	3.9-10

TABLES

Page

3.1-1 Primary QA/QC Functions of General Types of Methods 3.1-2

3.1-2 Minimum Recommended QA/QC Activities by Inventory Category 3.1-4

3.2-1 Reality Checks: Preferred and Alternative Methods 3.2-2

3.2-2 Standard References for Commonly Used Data (Annual Updates) 3.2-3

3.3-1 Peer Review: Preferred and Alternative Methods 3.3-2

3.4-1 Calculation Checks: Preferred and Alternative Methods 3.4-3

3.5-1 Summary of Common Automated Checks 3.5-4

3.6-1 Key Variables for Some Emissions Estimation Models 3.6-3

3.6-2 Key Inventory Variables for Some Air Quality Models 3.6-4

3.6-3 Sensitivity Analyses: Preferred and Alternative Methods 3.6-5

3.6-4 Daily VOC Inventory for a State 3.6-6

3.6-5 Model Input Values Used for MOBILE4.1 Sensitivity Analysis 3.6-11

3.7-1 Commonly Used Statistical Descriptors 3.7-5

3.7-2 Probability Distributions Relevant to Emissions Inventories 3.7-11

3.7-3 Data Transformations Used to "Normalize" Data 3.7-12

3.7-4 Descriptive Statistics for Los Angeles Point Source Inventory 3.7-13

3.7-5 Results of Dummy Variable Regression Analysis of County VMT Data . . 3.7-32

3.8-1 Objectives of Different Audit Types 3.8-12

3.8-2 Preferred and Alternative Methods: Audits 3.8-27

TABLES (CONTINUED)

Page

3.9-1 Comparison of Ambient and Inventory Ratios of CO/NO_x and $NMOC/NO_x$ for Los Angeles for 1987 . . . . . 3.9-5

3.9-2 Validation of Landfill Methane Emissions Models 3.9-6

1

INTRODUCTION TO METHODS

This chapter is a compilation of methods that can be used to implement a quality assurance/quality control (QA/QC) program, and should be consulted when preparing the quality assurance plan (QAP) at the beginning of inventory development (see Chapter 2 of this volume). The QA/QC methods to be employed should be determined and specified in the QAP. The data quality objectives (DQOs) and the inventory level will determine which methods and the number of different methods that are adequate for a given inventory. Such methods need not be limited to those described in this document; in fact, the inventory preparer is encouraged to develop or identify new tools or methods as needed to meet the DQOs set for the inventory.

The QA/QC methods described in this chapter are classified into the following general categories based on the type of activity involved:

- Reality checks;
- Peer review;
- Sample calculations;
- Computerized checks;
- Sensitivity analysis;
- Statistical checks;
- Independent audits; and
- Emissions estimation validation.

These methods are presented in order of increasing complexity. Each method has specific functions, although methods can be combined to cover a wider range of functions. The more complex and/or labor-intensive methods fulfill more than one function. Table 3.1-1 shows the primary function(s) of each type of method.

Any one of these methods by itself does not constitute a complete QA/QC program. Table 3.1-2 shows the minimum combination of methods recommended by the Emission Inventory Improvement Program (EIIP) for inventories at each level (see Chapter 2,

TABLE 3.1-1

PRIMARY QA/QC FUNCTIONS OF GENERAL TYPES OF METHODS

Method	Ensure Reasonableness of Emissions, Data	Ensure Validity of Assumptions, Methods	Ensure Mathematical Correctness	Ensure Valid Data Were Used	Optimize QA/QC Efforts	Ensure Proper Implementation of QA/QC Program	Assess Accuracy of Estimates
Reality checks	X						
Peer review	X	X					
Sample calculations			X	X			
Computerized checks	X		X	X			
Sensitivity analysis		X			X		
Statistical checks	X		X				
Independent audits	X	X	X	X	X	X	
Emissions estimation validation	X	X					X

Section 1 of this volume). Two of the methods--computerized checks and statistical checks--are not included in Table 3.1-2 because their function can largely be achieved using other methods. Moreover, these two general types of methods require specialized tools and/or expertise that may not be available to everyone. However, computerized and statistical checks can and should be used to the extent possible. They usually allow the processing of greater amounts of data than manual methods, and also reduce the chance of human error (on the part of the reviewer). Emissions estimation validation is also not listed in Table 3.1-2 because it requires resources and expertise that are beyond the capabilities of many organizations; however, these methods are very valuable as QA/QC tools and can be used to enhance the value of a QA program.

The methods presented in this chapter are designed to minimize the errors that can occur when developing an emission inventory. These errors can be of two types: procedural and technical.

Procedural errors result from unclear and ineffective management, inadequately trained staff, improper planning, lack of adequate QA, lack of data tracking and handling protocols, and other problems related to how the work gets done.

Technical errors are directly related to the methods and technologies used to develop emission estimates. They may result from:

- *Incorrect use of spreadsheets and databases*: Certain types of errors are solely a consequence of the mechanics of databases and spreadsheets. For example, an inadvertent reference to the wrong cell in a spreadsheet or retrieval of information from the wrong storage location in a database is very difficult to detect if the information in the erroneous data source and in the desired data source are similar.

- *Mathematical errors*: These types of errors can occur in hand calculations or in spreadsheets. They include the use of incorrect conversion factors, mismatched units in the emission factor and activity parameters, incorrect constants, and arithmetic errors.

- *Incorrect use of emission inventory software*: Software for calculating emissions is generally designed to minimize errors; however, it is possible to incorrectly calculate emissions if the user does not follow the instructions.

- *Use of incorrect data*: This covers a wide range of errors including the use of out-of-date data, inappropriate surrogate data, and data that do not

TABLE 3.1-2

MINIMUM RECOMMENDED QA/QC ACTIVITIES BY INVENTORY CATEGORY

Method	Level 1	Level 2	Level 3	Level 4
Reality checks	✓	✓	✓	✓
Peer review	✓	✓	✓	✓
Sample calculations	✓	✓	✓	
Sensitivity analysis	✓	✓[a]		
Independent audits	✓	✓[a]		

[a] Level 2 may include fewer audits and/or different types of sensitivity analyses and audits than Level 1.

correctly match the source category definition. This applies to both activity data and emission factors.

- *Use of incorrect methodology and/or assumptions*: Methodology is defined to include both computer models (e.g., MOBILE5b) and prescribed methodology (e.g., the use of gasoline throughput in a county and an emission factor to calculate evaporative emissions from gasoline dispensing is a type of model). This category covers just about any other errors not included in the first four error types.

- *Failure to include all sources*: This addresses inventory completeness. For a point source inventory, the omission of some emission units makes an inventory incomplete. For a regional inventory, this also covers failure to include entire source categories.

- *Double-counting emissions*: Double-counting of emissions between area and point source inventories is a technical error that results in an overestimate of the emissions reported.

This chapter primarily covers QA/QC methods that address technical errors, although the section on independent audits includes methods for reducing procedural errors. Other chapters of this volume (*Planning and Documentation, Model QA Plan*) as well as

U.S. Environmental Protection Agency (EPA) guidance (EPA, 1988; EPA, 1989) address procedural QA activities. In general, avoiding or minimizing procedural errors will reduce the likelihood of technical errors.

This page is intentionally left blank.

2

REALITY CHECKS

QUICK REVIEW

Definition:	An assessment of an emission estimate or data designed to answer the question, "Does this make sense?"
Objective:	Ensure that the value being checked is not wildly improbable, or that it is within a specified range or limit that is considered reasonable.
Optimal usage:	Anywhere and everywhere throughout the inventory development process. This method provides the quickest and easiest means of spotting large mistakes before the process has advanced very far.
Minimum expertise required:	Moderate. Although relatively inexperienced inventory developers can be trained to use this method (provided they are given some guidance and documents to use as a reference), personnel with some experience with emission inventories will be the most efficient and effective.
Advantages:	Quick, easy.
Limitations:	(1) It is possible that a value is "reasonable" but is still wrong, (2) it is possible that a value appears unreasonable but is still right, (3) the method does not give any insight into the source of an error, and (4) the method is only reliable if a valid point of reference is available.

2.1 OVERVIEW OF METHOD

The reality check is probably the most commonly used QA/QC method. It can be particularly useful for catching large errors early in the process. This check is a form of assessment to answer the questions "Is this number reasonable? Does it make sense?" However, it is important to remember that an estimate can appear to be wrong (that is, higher or lower than expected) but still be right. The opposite is also true--a value can look reasonable when it is actually incorrect. For these reasons, the reality check is never used as the sole criterion of quality.

To be able to answer the above questions, some notion of a reasonable value must be known. If the questioner is knowledgeable about emission inventories, he or she may not need an external reference to answer the question. However, if a standard reference value is available, the preferred method for a reality check is to use an external reference value.

Other types of reality checks include comparing the value to its equivalent value from a previous or alternative inventory or database. Another check might compare the estimated emissions to those calculated for the same category in a nearby or similar region. Table 3.2-1 summarizes the preferred and alternative methods for performing reality checks.

TABLE 3.2-1

REALITY CHECKS: PREFERRED AND ALTERNATIVE METHODS

Method	Procedure
Preferred	Compare data or estimate to a standard reference value.
Alternative 1	Compare data or estimate to a value from a previous or alternative inventory (or database) for the same region.
Alternative 2	Compare data to values used for other regions.
Alternative 3	Use expert or engineering judgment to assess the reasonableness of the values.
Alternative 4	Compare estimates for similar categories within the same inventory.

2.2 COMPARISON TO REFERENCE VALUE

This method can be used to assess the reasonableness of the emission estimates or of the data used to calculate those estimates. For a Level 1 or 2 inventory, this method should be used as much as possible. For example, emissions for many area source categories are estimated using population, employment, or other readily available surrogates for emissive activities. Generally, the preferred area source estimation method is to use local survey data. A good reality check for these locally obtained values is to compare them to some of the statistics generated by federal (or other) agencies.

Standard references for commonly used data are shown in Table 3.2-2. These data sources are often used as a source of activity data for area source calculations; in these cases, an alternative reference for reality checks is needed.

TABLE 3.2-2

STANDARD REFERENCES FOR COMMONLY USED DATA (ANNUAL UPDATES)

Reference	Types of Data	Resolution
U.S. Dept. of Commerce, Bureau of the Census (year). *19__ Census of Population and Housing, Summary Population and Housing Characteristics*, (state, year) CPH-1-34. Washington, DC.[a]	Population, number of housing units by type	Townships, other sub-county units
U.S. Department of Commerce, Bureau of the Census (year). *County Business Patterns* - (state, year). Washington, DC.[a]	Employment by Standard Industrial Classification (SIC) code and size of establishments	County level
U.S. Dept. of Commerce, Bureau of the Census (year). *19__ Census of Manufacturers*. Geographic Series MC87-A1 to -51. Washington, DC.[a]	Employment, hours worked, value of shipments and other data by SIC code group	County and state level
U.S. Department of Commerce (year). *Statistical Abstract of the United States*, (year). Washington, DC.[a]	Wide variety of useful data: economics, population characteristics, energy use, industrial production	National, regional; some state level data

TABLE 3.2-2

CONTINUED

Reference	Types of Data	Resolution
U.S. Department of Energy, (year). *State Energy Data Report Consumption Estimates, 1960-year*, Energy Information Administration, SOW/EIA-0124(xx). Washington, DC.	Fossil fuel consumption by sector	State level
U.S. Department of Transportation, (year). *Highway Statistics* (year). Federal Highway Administration, Washington, DC.	Vehicle miles travelled, onroad and offroad fuel consumption	State level
U.S. Environmental Protection Agency, 1993. *Regional Interim Emission Inventories (1987-1991), Volume 1: Development Methodologies*, EPA-454/R-93-021a. Research Triangle Park, NC.	Emissions of criteria pollutants	County level
U.S. Environmental Protection Agency, 1994. *National Air Pollutant Emissions Trends, 1990-1993*. EPA-454/R-94-027. Research Triangle Park, NC.	Emissions of criteria pollutants	National

[a] Data based on most recent census; therefore, accuracy decreases each year beyond the most recent census.

2.3 COMPARISON TO OTHER REGIONS

Another type of reality check is to compare results to those from different regions. A region that is very similar may be chosen with the expectation of similar emission estimates. An alternative approach is to contrast regions known to be different with the expectation that one region will have higher emissions than the other.

When using this approach, care should be taken to consider (and possibly remove) any effects that make comparisons unequal. For example, two counties may be similar in terms of type of industry and community characteristics, but very different in size.

Therefore, comparing emissions on a per capita, per employment, or other consistent basis should be considered.

Figure 3.2-1 shows an example of a graphical comparison of daily volatile organic compound (VOC) emissions on a per-capita basis in three neighboring regions: the portion (multiple counties) of a large Metropolitan Statistical Area (MSA) in State 1, the portion of the MSA in State 2, and an adjacent county in State 1 that is more similar to the State 2 MSA region with respect to population density, road types, and industrial development. The expectation was that the State 1 adjacent county and State 2 MSA region would have more similar per capita emissions than the densely populated State 1 MSA region. The onroad emissions are as expected: vehicle miles travelled (VMT) per capita is lower on average in densely populated urban areas. However, nonroad emissions for these three areas should be more similar on a per-capita basis. The discrepancy in area source emissions was also cause for concern because both inventories were known to be complete in their coverage of sources. This reality check suggested that the data and methods used for nonroad and area source emissions need further investigating (this example came from early draft inventories of an actual urban area; the difference in nonroad emissions were found to be the result of inventory developers from the various regions using different technical approaches to calculate emissions. The area source differences were found to be partly due to differences in the types of sources and to differences in the level of controls in place).

2.4 Comparison to Other Categories

Comparing categories within an inventory to other categories within the same inventory is another type of reality check. For example, nitrogen oxides (NO_x) emissions from utility boilers could be compared to those from industrial space heaters (point sources). Emissions from solvent usage from different area source categories could also be compared relative to the anticipated level of solvent usage or prevalence of controls. In both examples one may expect to find dissimilar emission estimates; the degree of dissimilarity to expect serves as the reality check and is dependent on the experience of the inventory developer.

Global checking of this type could help to quickly identify emissions estimates that may be erroneous. Additional checking is needed when data are suspected to be unreasonable. One must remember that reality checking is merely a screening process; therefore, it should never be used solely as a means of determining data quality.

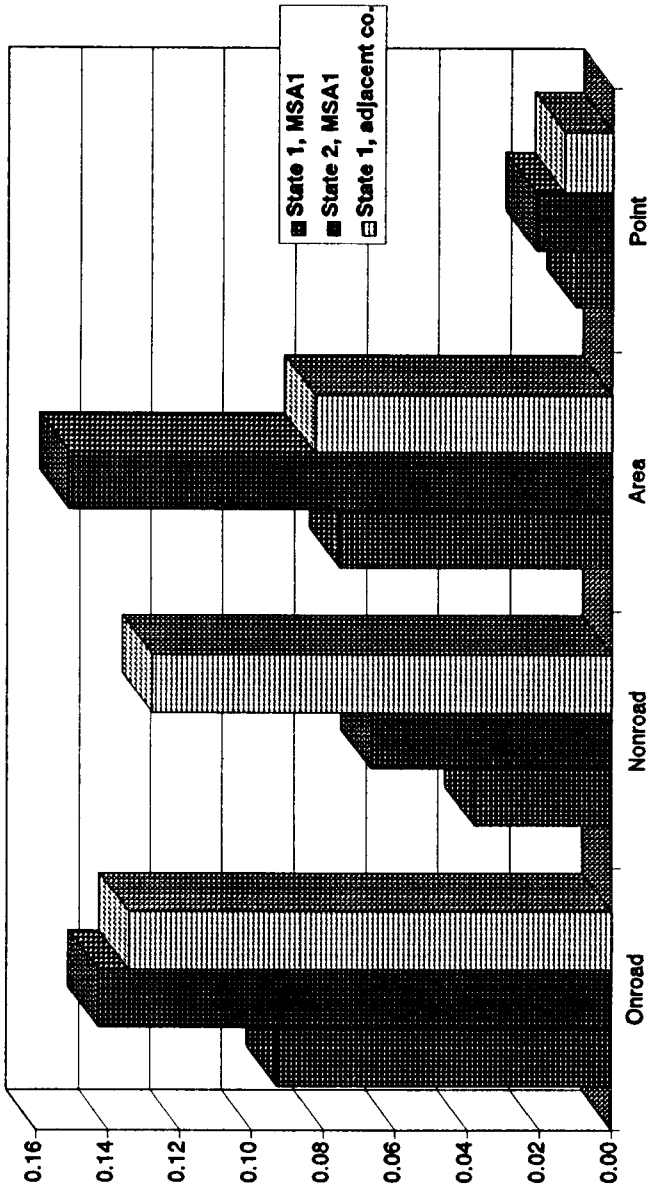

FIGURE 3.2-1. COMPARISON OF PER CAPITA DAILY VOC EMISSIONS

3

PEER REVIEW

QUICK REVIEW

Definition:	An independent review of calculations, assumptions, and/or documentation by a knowledgeable expert in the technical field. This is generally accomplished by reading or reviewing documentation, but may not include rigorous certification of data or references such as might be done in a technical audit (see Section 8 of this chapter, "Independent Audits").
Objective:	Ensure that assumptions and procedures used are reasonable and meet expectations as judged by persons knowledgeable in a specific field.
Optimal usage:	At critical points where: (1) information is organized in a manner amenable to review, (2) before critical decisions or commitment of large amounts of resources must be made, and (3) when the document, inventory, or database is completed, but prior to its release.
Minimum expertise required:	Moderate to high. Although the term "peer" implies equality, the most effective peer review is often done by someone with a high level of expertise. At a minimum, the expertise of the reviewer should be equivalent to the expertise required to manage the aspect of the inventory development activities from which the data to be reviewed were generated.
Advantages:	Usually not a labor-intensive effort; very flexible.
Limitations:	(1) Peer review is a form of reality check (and therefore has similar limitations), (2) the review is often general and does not check accuracy or quality of original data, and (3) for a large and/or complex inventory, it is very easy for a peer reviewer to miss errors.

3.1 OVERVIEW OF METHOD

The peer review process is very flexible and generally does not require specific tools; however, the use of checklists or review forms, which also provide a means of documenting QA/QC procedures, is strongly encouraged. If checklists are not used, the peer reviewer should provide feedback in the form of written comments. In some cases, the peer reviewer's comments may be transmitted verbally, however, the inventory developer should document when the review took place, the major points or items of discussion, and steps that will be taken to address the comments. Table 3.3-1 summarizes the preferred and alternative methods.

TABLE 3.3-1

PEER REVIEW: PREFERRED AND ALTERNATIVE METHODS

Method	Procedure
Preferred	Use of a checklist showing elements to be covered by the review. Provides a guide for the peer reviewer and can be tailored to fit specific situation.
Alternative 1	Written comments by reviewer identifying issues noted.
Alternative 2	Written notes summarizing reviewer's comments identifying issues noted by reviewer as told to author of notes.

At a minimum (i.e., for a Level 4 inventory review), a peer reviewer will perform some reality checks. A more in-depth review will include some checks of the calculations and verification that correct data and emission factors were used. Access to original data and reports is required to conduct in-depth reviews. The extent of the review required will depend on other QA/QC activities, the DQOs, and the inventory category (Levels 1-4; see Chapter 2 of this volume). If adequate calculation checks and/or data verification are being conducted as part of the QA program, the peer reviewer may be asked to concentrate on the validity of the methods and assumptions.

3.2 CHECKLISTS

The peer reviewer should have a clear understanding of what he or she is expected to do. A checklist is the most expedient way to accomplish this. The completed checklist also provides documentation that the specified QA program procedures were followed.

The peer review checklist is usually prepared by the Inventory Development Manager and the task leaders, with the assistance of the QA Coordinator and peer reviewer. Examples of checklists that can be used for peer review are discussed below.

3.2.1 EPA'S LEVEL I AND LEVEL II CHECKLISTS

The EPA's Level I/Level II checklists were used to review the 1990 base year State Implementation Plan (SIP) inventories. Many of the items are specific to SIP inventories and may not apply to other types of inventories. An important limitation of the Level I checklist is that it only deals with procedures or methods used to calculate emissions--it does not address the validity of the assumptions or data used in the calculations. The Level II checklist is more detailed, but also fails to fully address some of these important issues.

A complete copy of the Level I and Level II checklists and guidance for their use is found in *Quality Review Guidelines for 1990 Base Year Emission Inventories* (EPA-450/4-91-022)(EPA, 1991a). An automated version of the checklists is available through the Clearinghouse for Inventories and Emission Factors (CHIEF) bulletin board.

3.2.2 CHICAGO AREA SOURCE TOXICS INVENTORY

The example checklist shown in Figure 3.3-1 was designed for a detailed peer review of individual area source categories. The reviewer is asked to answer nine specific questions. The QA program for this inventory included extensive QC checks of the calculations and a procedural audit, but no technical audit was planned. Therefore, a thorough peer review that included some data verification and calculation checks was used (see Section 4, "Sample Calculations," of this chapter).

This checklist includes a place ("QC review") for the peer reviewer to sign. The QA review signature space is provided for the procedural audit. The QA auditor signs to show that he or she reviewed the form as part of the audit (see Section 8, "Independent Audits," of this chapter). The auditor's signature indicates a check was performed to determine whether approved procedures were followed and whether problems identified were resolved.

3.3 PEER REVIEWER REPORTS

These reports may take the form of a memorandum or other less formal report. Providing a form for the reviewer's comments will facilitate the production of this report. The example shown as Figure 3.3-2 is a reviewer's comments on calculations of VOC losses from gasoline distribution. It is a good example of a peer review because

CHICAGO AREA SOURCE TOXICS INVENTORY
Source Category Review

CATEGORY DESCRIPTION:

INVENTORY REGIONS COVERED:

NAME (person responsible for calculations): _____

Signature (QC Review) _____ Date _____

Signature (QA Review) _____ Date _____

Checklist	YES	NO	Could Not Be Determined
1. Were all speciation profiles clearly referenced?			
2. Were correct emission factors used?			
3. Were sources of activity data clearly referenced?			
4. Were correct activity data used?			
5. Were annual emissions correctly adjusted from daily estimates?			
6. Were all calculations documented?			
7. Are calculations correct?			
8. Were applicable regulations cited?			
If so, were RE and RP correctly applied?			
9. Were adjustments made to account for any excluded VOCs?			

The following problems need to be corrected:

_____

_____

_____

_____

ACTION TAKEN:

_____

_____

FIGURE 3.3-1. EXAMPLE OF A PEER REVIEW CHECKLIST

QC-1 Comments on the NY SIP Inventory Gasoline Marketing Area Sources

9/16/92

GDRives

1. Emission factors are in general agreement with AP-42. Include tank truck in transit loss emission factors in Table 2 of the documentation. Emission factors could be refined a little bit to reflect typical gasoline RVP values and summer temperatures. Note that the state has a gasoline RVP restriction (<9 psia) during the ozone season (NY Rules Subpart 225-3, page 461:0593 BNR). Normal daily mean temperature for NY is probably closer to 70° F for the ozone season than 60. If you plan to refine emission factors any, I would probably select the midpoint of the range for transit losses since they have a tank truck tightness testing and certification program.

2. Assumptions on filling methods are outlined nicely for the NY metro area, what filling assumptions were used for the other areas? I recommend expanding Table 1 to address all areas.

3. Table 1 and Table 3 are not parallel which makes it difficult to review. I'd recommend presenting all information in the same sequence (i.e., county and area orders)

4. Table 3. Don't include NOx and CO for this source since they will just remain blank. Instead, you could present annual gasoline throughput by county, and assumptions on penetration of Stage I and II controls. This would provide all the activity data necessary to independently perform calculations.

5. Table 3. A little bit of clarification on Table 3 would help the review. At first glance, the columns don't add up correctly. With a more detailed look, you can figure it out but it leads to confusion. For example, NY Metro Area appears to have 11 components for a total of 13,192 tpy. However, "NYC Sum" appears to be a subtotal of the same group and should not be included in the totals, by the indention it appears to be just another county. Why don't we put NYC sum under the subtotal row?

6. Calculations look real good for NY area including the application of rule effectiveness for loading operations. I've checked a couple of the counties, and they are spot on. I can't independently verify the calcs for other areas, why did we use 90% rule effectiveness for loading?

7. I would have probably applied rule effectiveness when calculating emissions from submerged filling as well as vapor balance. I understand that you have already discussed with EPA and have reached consensus.

FIGURE 3.3-2. EXAMPLE OF WRITTEN PEER REVIEW REPORT

8. I reviewed all of the county-specific gasoline throughput values to ensure that the data were correctly transcribed from the hard-copy supplied by the state energy office to our spreadsheets. With the exception of Tiago County, all matched correctly. Note that Tiago County should be 25460 instead of 25450.

9. Have we performed any sanity checks on these gasoline consumption data? A couple of checks would be to calculate gal/person, gal/VMT, and gal/registered vehicle for each of the NY counties. Then calculate average and standard deviation for each. Any counties outside of 2 std deviations from the mean deserve to be investigated.

10. I like the inclusion of supporting materials with the writeup. I hope that the plan is to submit this inventory with the documentation.

11. Add dates and document numbers to the references, where appropriate.

FIGURE 3.3-2. CONTINUED

the strengths as well as the weaknesses are described. The reviewer is also very clear about what he did and did not check. This was a very thorough peer review; note the suggestions for improving the presentation of the data (items 3 and 5). This emphasizes the importance of clear and thorough documentation for QA/QC. Finally, the reviewer makes recommendations on how to improve the estimates.

4

SAMPLE CALCULATIONS

QUICK REVIEW

Definition:	Verification of values calculated using spreadsheets, models, or hand calculations by *replicating* the calculations by hand or electronically; a subset or sample of each type of calculation is checked.
Objective:	Ensure that calculations are done correctly, minimizing mathematical and spreadsheet errors.
Optimal usage:	Should be used extensively and frequently as a QC tool; also used as part of a technical audit.
Minimum expertise required:	Low to moderate. Replication of calculations can be done by anyone capable of doing the work (i.e., some expertise in inventory development recommended).
Advantages:	(1) Most reliable way of detecting computational errors, (2) can be done by anyone involved in inventory development, and (3) does not require expensive or complicated tools (only pen and paper).
Limitations:	(1) Does not ensure that approach and assumptions used are correct, and (2) is labor-intensive.

4.1 Overview of Method

Replication of calculations is one of the most basic quantitative QA/QC methods available. It should be used by the author of the original calculations as a self-check, by the person performing QC checks, and as part of any QA audit.

Self-checking using manual or electronic calculations is strongly encouraged when spreadsheets are used to calculate values. Each author should be encouraged to include an example calculation replicating one complete set of calculations on his spreadsheet or as an attachment to help facilitate the review of the technical approach and identify errors. If the original calculations are done by hand, replication by the author may still be useful.

The most common usage of this method is by the person conducting the QC check. The preferred QC method is to replicate one complete set of calculations from beginning to end. In addition to verifying the mathematical correctness of the calculation, the values used should also be verified. For example, use of the correct emission factor should be checked. If the original calculations are in a spreadsheet, spot checks of the remaining values (assuming the same equations are being used repetitively) is sufficient. The percentage of calculations that should be checked will depend on several factors, including:

- The complexity of the calculations;
- The DQOs for the inventory; and
- The error rate encountered.

The error rate is determined during the QA/QC process. If a large number of errors is being detected, the number of checks should be increased, although not necessarily on a given version of the spreadsheet. For example, during the initial check, several errors may be found. The person conducting the QC check may stop after working through one set of calculations and send it back to the author for revision. In the second review, the original calculation is rechecked and found to be correct. The person conducting the QC check then proceeds to check another set of calculations; if no errors are found, he or she may spot check a few more values before approving the calculations (see Section 7, "Statistical Checks," for more discussion on statistical QC procedures).

The procedures for determining the number of calculations to be checked and the number of iterations required before the calculations are approved are a function of the DQOs and should be specified in the QA plan. Ensuring that the specified QC procedures are followed is the responsibility of the QA Coordinator.

In addition to checking the arithmetic used, the proper conversion of all units should be verified. Units analysis may be done independently of the actual values.

Replication of calculations by hand (e.g., using pen, paper, and calculator) is the preferred method for checking calculations originally performed by hand or in a spreadsheet. Spreadsheets can be used as an alternative; however, they are not preferred as a QA technique because they introduce the possibility of other errors unique to spreadsheets. In certain situations, it may not be feasible to replicate calculations; for example, if emissions are generated by a model, the algorithm may either be unknown to the user or too complex to repeat by hand. A check would be to attempt to approximate the result using an alternative method; this is a more complex type of a reality check. Table 3.4-1 summarizes the preferred and alternative methods.

TABLE 3.4-1

CALCULATION CHECKS: PREFERRED AND ALTERNATIVE METHODS

Method	Procedure
Preferred	Hand replication of one complete set of calculations.
Alternative 1	Hand replication of most complex calculations.
Alternative 2	Hand calculation using a different method, attempting to approximate the result.

4.2 REPLICATION OF SPREADSHEET OR HAND CALCULATIONS

Figure 3.4-1A shows spreadsheet calculations of structure fire emissions for several counties. Figure 3.4-1B shows the original hand-calculated QC checks of one county's emissions. Note that the reviewer did *not* verify emissions of all counties; a reality check was used to check the others. Depending on the DQOs for the inventory, this level of checking may be adequate, particularly if combined with a spreadsheet audit (see Section 5.2.4, "Spreadsheet Auditing Functions").

4.3 REPLICATION OF COMPLEX CALCULATIONS

When a model or other software program is used to calculate emissions, hand replication of each type of calculation performed is the preferred method. However, if

STRUCTURAL FIRE EMISSIONS FOR OZONE SEASON

AUTHOR: M. Fauls
PREPARED: 9/10/92
PURPOSE: Prepare ozone season emission estimates of NOx, CO, and VOC from structure fires in New York Counties

CHECKED BY: R. Peer
DATE: 9/15/92

New York Metropolitan Nonattainment Area O3 Season: Summary of Emissions From Structural Fires SCC 28-10-030-000

New York Metropolitan Area	(a) Number of Fires (yearly)	(a) % of Fires in Ozone Season	(b) Tons Burned (yearly)	(c) Emission Factors (lb/ton) NOx	(c) CO	(c) VOC	(d) Activity (days/wk)	(f) NOx Emissions Annual (ton/yr)	(f) NOx Emissions Ozone Season (lbs/day)	(e) CO Emissions Annual (ton/yr)	(f) CO Emissions Ozone Season (lbs/day)	(e) VOC Emissions Annual (ton/yr)	(f) VOC Emissions Ozone Season (lbs/day)
Bronx	6775	24.3%	46070.00	1.4	60	11	7	32.2	170.4	1382.1	7301.1	253.4	1338.5
Kings	9245	24.3%	62866.00	1.4	60	11	7	44.0	232.5	1886.0	9962.9	345.8	1826.5
New York	7787	24.3%	52951.60	1.4	60	11	7	37.1	195.8	1588.5	8391.7	291.2	1538.5
Queens	4367	24.3%	29695.60	1.4	60	11	7	20.8	109.8	890.9	4706.1	163.3	862.8
Richmond	934	24.3%	6351.20	1.4	60	11	7	4.4	23.5	190.5	1006.5	34.9	184.5
Nassau	1094	25.6%	7439.20	1.4	60	11	7	5.2	29.0	223.2	1242.0	40.9	227.7
Suffolk	1313	26.2%	8928.40	1.4	60	11	7	6.2	35.6	267.9	1525.6	49.1	279.7
Westchester	2049	27.5%	13933.20	1.4	60	11	7	9.8	58.3	418.0	2498.9	76.6	458.1
Rockland	193	26.4%	1312.40	1.4	60	11	7	0.9	5.3	39.4	226.0	7.2	41.4
Orange	172	26.8%	1168.12	1.4	60	11	7	0.8	5.1	35.0	219.4	6.4	40.2
			Total Structural Fire Emissions =					161.5	865.2	6921.5	37080.2	1268.9	6798.0

ok

NOTES:

(a) Information supplied by NYSDEC Office of Fire Prevention and Control except for the five New York City boroughs. This information came fro the Fire Department of the City Of New York. FDNY Annual Report and Statistics-1991. Brooklyn, NY, p.65. Percentage of Fires in the Ozone Season for the five New York City boroughs was calculated using data provided by the Fire Department of the City of New York. FDNY Annual Report and Statistics-1990. Brooklyn, NY, p.58. Orange county apportionment based on the number of structure, ref. Orange County Memo, September 3,1992. Tompkins County information is derived from the county population times the statewide average number of fires per 1000 people, New York Department of State, Office of Fire Prevention and Control. Fire in New York 1990. Albany, NY, p.4.

(b) Tons Burned = (Yearly Number of Fires)*(Fuel Loading Factor)
The fuel loading factor is taken from the EPA Procedures Document, Section 4.8.4.
The value used in this inventory is 6.8 tons of material per fire.

(c) Emission factors are taken from the EPA Procedures Document, Section 4.8.4.

(d) Activity is taken from the EPA Procedures Document, Table 5.8-1.

(e) (Annual Emissions t/yr) = (Tons Burned) * (Emission Factor lbs/ton)/ (2000 lbs/ton)

(f) (Ozone Season Emissions lbs/day) = (Tons Burned) * (% of Fires in Ozone Season) * (Emission Factor)/(# of Days in the Ozone Season)
100

FIGURE 3.4-1A. EXAMPLE SPREADSHEET CALCULATION

SIGNATURE R. Peer DATE 9/15/95 CHECKED WA DATE

PROJECT Ozone inventory JOB NO.

SUBJECT QC checks of "Structure Fires" SHEET 1 OF 1 SHEETS

Objective: Check calculations in STRUCTO3.XLS.

Largest (in terms of emissions) county was chosen - Kings.

(1) Annual emissions (tons/yr) = $\frac{\text{Tons burned/yr} * \text{Factor (lb/ton)}}{2000 \text{ lb/ton}}$

Annual emissions (t/yr) = $\frac{62{,}866 \text{ tons/yr} * 60 \text{ lbs CO/ton}}{2000 \text{ lb/ton}} = 1{,}885.98 \text{ t/yr}$

(2) Ozone season emissions (lbs/day) =

$$\frac{(\text{tons burned/yr}) * \left(\frac{\text{\% Fires in } O_3 \text{ Season}}{100}\right) * \text{Factor (lb/ton)}}{\text{\# days in ozone season}}$$

$= \frac{62{,}866 \text{ tons/yr} * 0.243 * 60}{92} = 9{,}962.89$ lb/day

Conclusions:

- Kings Co. calculations verified by replication of CO emissions shown above. NOx and VOC were verified (but not shown).
- Other counties appear correct based on visual comparison of values.
- Documentation adequate, except number of days in ozone season not shown. Equation (f) should show division by 100.

FIGURE 3.4-1B. EXAMPLE OF HAND-CALCULATED CHECK OF SPREADSHEET RESULT

the calculations are very complex or cannot be duplicated easily, an alternative approach is to attempt to duplicate the result using an alternative calculation method.

The following example shows one approach to checking the calculations used in an emissions estimation software program. The Air Quality Utility Information System (AQUIS) (U.S. Air Force, 1994; Smith et al., 1995) is used by the U.S. Air Force, Air Combat Command, to estimate air emissions at a single site.

Note that this type of check only confirms that the mathematical calculations are performed as expected. Any data that are supplied by the software (e.g., conversion factors, emission factors) should also be verified and input data need to be checked (see Section 3, "Peer Review," Section 5, "Computerized Checks," and Section 8, "Independent Audits").

Figure 3.4-2 shows a hand-calculated check of the AQUIS algorithm for industrial reciprocating engine emissions (The original calculations were subsequently typed to give a neater presentation for a report). The input data used in both the program and the hand calculation are shown in Section I. The equation used by AQUIS (as reported in the AQUIS Algorithm Manual) is shown in Section II. Sample calculations are shown for three hazardous air pollutants and the results compared to the program output (The third page, not shown, contains a list of all pollutants calculated by AQUIS).

Although this level of checking for a program can require a significant amount of time, it is necessary. Furthermore, given that these programs are generally used many times over, the effort required to check the algorithms is relatively small.

CALCULATION SHEET

CALC. NO. 2

SIGNATURE J. Scientist DATE 6/30/94 CHECKED R. Peer DATE 7/13/94

PROJECT Inventory Database QA JOB NO.

SUBJECT Review of AQUIS Methodology SHEET 1 OF 3 SHEETS

INTERNAL COMBUSTION GENERATORS

I. Input Data:

Type of Stationary Source = Industrial reciprocating engine
Fuel = Diesel
Rated Power = 201.2 hp
Hourly Max Fuel Usage = 10.13 gal/hr
Annual Average Fuel Usage = 1.6694×10^3 gal/yr
Fuel Unit of Measure (fu) = gal

II. Annual Emissions Calculation:

AAUE = NPIE * HV * EF(P) * AU

where:

AAUE	=	Annual emissions from generators (lb/yr)
NPIE	=	Number of identical pieces of equipment = 1
HV	=	Heating value of fuel = 137 (10^3 Btu/fu)
	=	137×10^3 Btu/gal
EF(P)	=	Emission factor for pollutant P (lb/10^6 Btu) for Type 2 (from Table 3.2.1-4)
AU	=	Annual average quantity of fuel (10^3 fu/yr)
	=	1.6694×10^3 gal/yr

This is the Type 2 equation (one of three types), where type is a function of engine design and pollutant.

FIGURE 3.4-2. EXAMPLE HAND-CALCULATED CHECK OF DATABASE ALGORITHM

CALCULATION SHEET

CALC. NO. 2

SIGNATURE J. Scientist **DATE** 6/30/94 **CHECKED** R. Peer **DATE** 7/13/94

PROJECT Inventory Database QA **JOB NO.**

SUBJECT Review of AQUIS Methodology **SHEET** 1 **OF** 3 **SHEETS**

A. For 1,3-Butadiene:

AAUE = 1 * 137 x 10^3 Btu/gal * 3.91 x 10^{-5} lb/10^6 Btu * 1.6694 x 10^3 gal/yr
= 8.94 x 10^{-3} lb/yr 1,3-Butadiene

AQUIS Printout: AAUE = 0.009 lb/yr 1,3-Butadiene

B. For Benzene:

AAUE = 1 * 137 x 10^3 Btu/gal * 9.33 x 10^{-4} lb/10^6 Btu * 1.6694 x 10^3 gal/yr
= 0.213 lb/yr Benzene

AQUIS Printout: AAUE = 2.213 lb/yr Benzene

C. For Naphthalene:

AAUE = 1 * 137 x 10^3 Btu/gal * 8.48 x 10^{-5} lb/10^6 Btu * 1.6694 x 10^3 gal/yr
= 0.0194 lb/yr Naphthalene

AQUIS Printout: AAUE = 0.019 lb/yr Naphthalene

FIGURE 3.4-2. CONTINUED

5

COMPUTERIZED CHECKS

QUICK REVIEW

Definition:	Electronic methods of checking or verifying data or emissions; they may be a built-in function of an emissions database, model, or spreadsheet, or they may be stand-alone programs.
Objective:	Ensure the accuracy or correctness of a value.
Optimal usage:	Generally used for QC at the point where data are being entered into a computer software program for further processing and/or after emissions have been calculated. Frequency and timing of use depend mostly on the specific technology.
Minimum expertise required:	Low to moderate, depending on the specific tool/method.
Advantages:	(1) Built-in checks can be done with little additional effort and the need for more laborious manual reviews may be eliminated, (2) they allow efficient checking of large amounts of data, and (3) certain software features prevent or minimize some types of errors.
Limitations:	(1) Requires some level of human decision-making (such as determining whether a variable that is outside a range is really an error), (2) some programs--particularly stand-alone programs--require additional training or expertise to use properly, (3) all require a computer, (4) too much reliance may be placed on the computer's ability to detect errors, resulting in too little QA/QC by humans, and (5) cannot catch erroneous values that pass range/limits checks.

5.1 OVERVIEW OF METHOD

Automated checks cover a wide array of computer-based QA/QC functions. They include: (1) common features found in most database programs or models that ensure against entering characters where numbers are expected, (2) range checks that give an expected minimum and maximum for a specific variable, and (3) look-up tables that define permissible entries, and The user is given a message or printed report to document the potential problem. Some computer programs provide users with visual written QA/QC tools. Spreadsheet audit functions, for example, can map cell references and formulas visually in report format. Other computer programs are designed to display data visually through maps, charts, or graphics. These features vary in their implementation between computer programs; it is important that the user understand how they work and then use these features correctly.

Unlike the methods described in previous sections of this chapter, the inventory preparers may have very little choice in selecting automated checks because the emissions estimation system may already have been chosen (or may be outside their control). However, the automated checks described in this section do not provide any new functions; they are automated, efficient ways to screen, summarize, and/or present data for "Peer Review" (Section 3) or "Independent Audits" (Section 8). This section describes some of the general QA/QC functions used in emissions estimation software, such as:

- Features of spreadsheets that can be used to facilitate QA/QC; and
- Some stand-alone software programs that can be used in QA/QC tools.

The computer-literate inventory preparer is encouraged to develop his/her own automated QA/QC programs using the ideas presented here.

Automated QA/QC functions can be used to facilitate peer review or, in some cases, replace manual reality checks. Because large amounts of data can be processed very quickly, these computer-based methods can significantly reduce the amount of time required both to develop and quality assure an inventory. In addition, some of the data entry features used in computer-based estimation software that were designed primarily to simplify data entry also help reduce the number of potential errors. By reducing the amount of information that has to be entered manually, the number of typographical errors is reduced. Furthermore, some programs select the correct emission factor (or other defaults needed to estimate emissions) based on one or more parameters provided by the user. Although they may not eliminate errors, these automated procedures can help reduce the use of incorrect data.

It is important to note that these reports do not actually perform QA functions. Reports that are designed to highlight outliers or suspect data do just that--highlight *suspect* data. Similarly, reports that display information only display information as requested. Human reasoning and judgment must still be applied when reviewing the reports and evaluating them for errors. The preferred method for emissions estimation and performance of many types of QA/QC checks is to use emissions-estimation software with built-in QA/QC functions and data entry procedures that minimize errors. If that is not feasible, the second-best method is to develop standardized spreadsheets or databases using standard software packages, building in as many QA features as possible. Most popular spreadsheet and database packages support the use of look-up tables and provide auditing tools that can be used for QA/QC. Portions of the spreadsheet or database file can be locked or protected to prevent inadvertent changes.

Because the programmer is often not the end user, it is highly recommended that the end user be involved in developing the program and its output features. Involving the user will help maximize the usefulness of the program and make it more user friendly.

Emissions-estimation software or spreadsheet/database programs can be supplemented using stand-alone programs to check or review the data. Some programs provide range checks on data, generating reports that flag potential problems. Emissions estimates can be further verified by using data visualization programs. These software programs are capable of transferring emissions data from spreadsheet/database format to visual tools such as points on a map, charts, and tables.

Table 3.5-1 summarizes the methods and functions described in this section. Preferred and alternative methods were not defined for this group of tools because the technology available to different agencies varies.

As valuable as these automated tools are, they do have some limitations, such as creating a false sense of confidence in the data. It is usually necessary for a reviewer to make a judgment about the data quality by reviewing the computer-generated reports or error messages and deciding which potential errors need to be corrected. Also, even if the value for a particular parameter is within a reasonable range, it can still be wrong. Therefore, these automated tools are not a substitute for peer reviews or audits; rather, they should be used to facilitate other QA/QC activities.

Most emissions-estimation software offers built-in functions designed to prevent the introduction of faulty data into the emissions calculation. Functions such as range and data-type checks are QA functions that flag potential errors. Other features are indirect forms of QA in that they reduce the amount of data that must be entered or limit the options. These features include look-up tables and pull-down menus where the user

TABLE 3.5-1

SUMMARY OF COMMON AUTOMATED CHECKS

Type of Automated Check	Description	Examples	Strengths/Limitations
Variable type check	Alerts user if wrong data type or inappropriate value is entered.	Numeric value is expected, character string entered: a warning is issued immediately or field is flagged in subsequent report.	Reduces errors early in process, especially if warning issued interactively and/or if incorrect data entry prohibited. Report in which error is flagged is easily ignored.
Range (value) checks	Checks value entered to determine if it is within an expected or acceptable range.	Range of stack heights is used to flag a stack height that is too high or low.	Flags suspicious data for further review. Does not eliminate possibility that wrong value entered is within range, or that value outside range could still be correct.
Look-up table	Uses a parameter (such as user-supplied input variable) to select other appropriate parameters from a table.	User enters a source category code (SCC) and program supplies appropriate emission factor.	Eliminates some types of data entry errors; assures data consistency. If wrong value added (i.e., incorrect SCC), all dependent values will also be wrong.
Pull-down menu, pop-up window	Presents selection of possible values for a particular field.	List of possible fuel types is presented to user when entering data to calculate boiler emissions.	Eliminates transcription errors, reduces chance of using wrong value due to user not understanding what is wanted. Does not eliminate possibility that wrong choice will be made by user.
Completeness/ Consistency checks	These two terms are often used to describe similar operations; include a wide array of checks and/or comparisons.	Checks verify that some specified amount of data for certain fields has been entered; or, if a certain field has data, verifies that other required fields also have data. Assure that units, equipment types, IDs, and other parameters are consistent.	Completeness is often difficult to quantify; in practice, a minimum expected value is used to determine completeness. Does not assure that data are correct. Impossible to completely automate these types of checks, some expert judgment usually required. If too much consistency automated into process, inflexibility may result.

must choose from a limited but inclusive list of options. Although not necessarily meant as a QA tool, these features can help minimize several types of data-entry errors.

5.1.1 RANGE AND CHARACTER CHECKS

Data-type checks prohibit the user from entering text strings when numerical values are expected and vice versa. In most cases, data-type and range checks are no more than an automated reality check. Range checks provide an indication that the value is (or is not) within an expected range. However, these checks do not necessarily ensure that the value is correct. Also, range checks may warn the user of a potential problem, but do not usually prevent entry of incorrect data.

Numerical range checks generally fall into one of two categories: (1) the upper and/or lower limit is specified, or (2) the statistical mean or median is used as a point of reference. In the first case, as long as the input value falls within (or above or below) the limit, no error is flagged. Setting this limit is difficult; assuming an upper bound is used, if it is set too high, no potential errors are flagged. If set too low, too many data points are flagged. The use of statistically derived reference values raises similar questions (see Section 7, "Statistical Checks," of this chapter).

A common frustration that many inventory preparers have with automated QA checks that use an average (or default) value as a reference point is that correct, site-specific data are often either higher or lower than a statistical average or default. This flagging of correct but atypical values should not be viewed as a problem with the system; rather, it provides the opportunity to review and explain the data. The *perceived* quality of the inventory to the user or reviewer is increased by this sort of analysis, even though the *real* quality is unchanged. In other words, confidence in the emission estimates has increased.

5.1.2 LOOK-UP TABLES

Look-up tables are another feature of many computer programs. They help ensure data quality by restricting the number of choices a user has to make. A look-up table uses one variable (entered by the user) to look up the correct value for another variable.

For example, some computerized emissions calculation programs supply the correct emission factor by matching it to a source category code (SCC). One approach is to retrieve the correct emission factor based on an SCC entered by the user. Although this ensures a correct match between the code and the factor, if the user enters the incorrect code, the emissions will be calculated with the wrong number. Therefore, look-up tables are not always foolproof.

Other systems minimize this type of error by building more safeguards into the system. The AQUIS used by the U.S. Air Force, Air Combat Command, has separate modules for each source type. To calculate boiler emissions, for example, the user enters the boiler characteristics, fuel use, and fuel type in the "External Combustion" module. The program uses this information to select the correct emission factor. The possibility of error still exists because incorrect boiler data could result in the selection of an incorrect emission factor; however, it is less likely to occur than when the additional step of supplying the correct SCC is required.

Still other systems allow the user to set up the relationships in the look-up table. This approach is used in some commercial emissions database software. Access to these tables is limited to a system administrator who can assign codes, emission factors, and specify other interrelationships between data to be entered in the system. Usually these software packages allow the user to enter alternative emission factors, conversion factors, or constants. The inventory developer using this type of software should take advantage of these features to build in QA/QC features.

5.1.3 Pull-down Menus and Pop-up Windows

Another error-limiting feature is the pull-down menu or pop-up window. This feature is activated when the user selects a field to enter data. The menu/window will appear and offer a limited set of allowable entries.

5.2 Internal Checks and Reports: CEIDARS, RAPIDS, Standardized Spreadsheets

After the data for performing the emissions calculations have been entered, many programs have internal QC programs designed to highlight suspect information for further evaluation. These programs are a form of "reality check," but are capable of processing larger volumes of data than a human reviewer, automatically flagging input/output that falls outside the expected parameters. However, the flagged data must be reviewed by a human to determine whether an error has really occurred. Software packages, such as the Regional Air Pollutant Inventory Development System (RAPIDS) and California Emissions Inventory Database Acquisition Retrieval System (CEIDARS) contain built-in checks and reports designed to evaluate data and present potential errors or highlight important information. Other software packages may allow the user to design a QC report that prints key variables. Depending on the flexibility of the software, the user may be able to design QA reports that perform some of the same functions as described below for CEIDARS and RAPIDS.

5.2.1 CEIDARS

The California Air Resources Board (ARB) uses an Oracle®-based relational database system to compile and maintain toxic and criteria air pollutant inventories. This system, CEIDARS, also provides many built-in QA functions that are used by the state agency to review data entered at the district level (California is divided into multiple air districts each with its own staff responsible for developing inventories and performing other functions). These data have presumably undergone QA/QC checks prior to submittal to the district, and reviewed at the district level prior to entry into CEIDARS (see Chapter 2 of this volume, *Planning and Documentation* for a description of the ARB's QA procedures designed to use these reports).

Eleven different QA reports are provided in CEIDARS; the differences between them are defined partly by the QA features and partly by the category type or level (e.g., area, point, stack).

To ensure the integrity and completeness of the criteria pollutant data, the CEIDARS QA reports are designed to help identify errors, inconsistencies, or omissions in the criteria pollutant inventory data input to the database. The QA reports verify that all critical data fields for each facility are entered into the database, and that various data consistency and range-checking criteria are met. The QA reports are not for general data (emissions) reporting.

Some CEIDARS reports perform integrity checks on data entered by the user (QAFACC, QASTKC, QACEVE, QAPROC). Other reports check key field data (QAKEYS), or the consistency of emissions data with reported process rates and emission factors (QAKEMS). QA reports also perform validity and range checking of various data including stack (QASTACK) and temporal (QATEMP) information.

Completeness Checks

Four different reports perform some completeness checks at the facility, stack, device, and process levels. The facility-level report verifies that at least one emissions record is reported for each facility. If a record is flagged, the reason for the lack of emissions is investigated. If emissions were reported but not entered, the district must resolve the error. However, if no facility emissions were reported (by the facility operator), the record may either be dropped or retained (for tracking). Similar reports flag stacks, devices, and processes that do not have any reported emissions. As with the facility QA report, the reviewer must determine if the omission is truly an error.

Consistency/Validity Checks

Several QA reports are designed to flag inconsistent, invalid, or missing data. One report checks to ensure that key data are present and consistent. For example, if a device has not been defined for a process, the error message "DEVICE NUMBER IS MISSING OR INCONSISTENT" is given. The items checked include county identification (ID) and name; device ID; facility ID, name, and Standard Industrial Classification (SIC) code; pollutant ID and name; process ID, SCC, and SIC codes; and stack ID.

A similar type of data check is performed on process and emissions data to verify that they are valid and consistent. For example, an error is flagged if an invalid emissions estimation code is used. If emissions are reported when the process rate is zero, an error is flagged. The validity of the reported emissions is checked by comparing it to a calculated value (i.e., process rate times an emission factor); if the two values are not within 10 percent of each other, the record is flagged. The items checked include emissions, fraction particulate matter (PM), district-specified reactive organic gas (ROG), PM less than or equal to 10 μm (PM_{10}), fraction ROG, process rate, and sulfur content.

Stack data validity is checked and verified to be within expected ranges by another set of QA checks. For example, if the stack temperature is less than 50°F or greater than 2000°F, an error is flagged. If the flow rate reported for the stack does not equal a computed theoretical flow rate ($\pm$10 percent), an error is flagged. Stack parameters checked include gas flow, gas temperature, gas velocity, stack diameter, stack height, and universal transverse mercator (UTM) coordinates.

Other reports verify the spatial and temporal data, device and process data, and control device data. The spatial and temporal parameters checked include days per week/weekly operating cycle, hours per day/daily operating cycle, and operating weeks per year. The device and process data reported are device ID, equipment size, equipment type code, equipment size unit code, maximum design rate, process rate, and process description.

5.2.2 RAPIDS

RAPIDS is a database management system developed specifically for compiling and maintaining state toxic emissions data into a central regional repository. States using this system have available many QA/QC functions in RAPIDS that should be used prior to submitting their toxic emissions data and estimates into the regional repository. The objective of these QC checks is to ensure that the regional database of emissions data

and estimates is as complete as practicable and is properly QA checked. Below is a list of the QC checks performed by RAPIDS:

- *Data integrity checks*: Ensure that no invalid relationships among data elements exist (e.g., no process without a device, device without a facility, material activity without a material).

- *Field-specific checks*: Ensure that obviously invalid data are not entered into the database [e.g., SIC code, SCC/Area and Mobile Source Subsystem (AMS) Codes, speciation factor (>0.0 to <1.0), data relationships, etc.].

- *Record-specific checks*: Identify a valid relationship between two values in different fields in the same record (e.g., installation date is earlier than the dismantling date of a facility or device, start date/time in an activity record is earlier than the current date).

- *Cross-entity checks*: Identify invalid relationships between data in two different entities (i.e., data elements).

- *Completeness checks*: Refer to the extent to which all sources/devices/ processes and emittants are accounted for in the inventory database.

- *Consistency and reasonableness checks*: Consistency refers to the extent that emissions data and estimates do not significantly vary among similar occurrences of the same area, source, device, process, stream, etc. Reasonableness refers to the extent that emissions data meet certain expectations based on experience.

The last two types of QC checks identify *possible* errors and require manual investigation to determine if an error really exists and how to correct it. Completeness or consistency and reasonableness QC checks do not lend themselves to inclusion as fully automated checks; rather, they require the user to generate reports or other outputs from general purpose software to identify possible errors in the data.

5.2.3 STANDARDIZED SPREADSHEETS/DATABASES

If emissions are to be calculated using spreadsheet or database software, many of the QA/QC features available in emissions-estimation software can be built into standardized spreadsheets or simple database programs. The additional effort required should pay off if the inventory is large, if it may be repeated or updated periodically, or if data entry may be performed by more than one person.

Area Source Inventory Spreadsheet Example

Standardized spreadsheets were developed and used to prepare an area source Level 4 inventory of air toxics in the Puget Sound region. The inventory preparation plan specifies the methods, factors, and speciation profiles used on the spreadsheet. All inventory personnel had a copy of the plan.

Figure 3.5-1 shows the spreadsheet developed for calculating graphics arts hazardous air pollutants (HAPs). It is one of many spreadsheets used to calculate area source HAP emissions for an urban area. The basic spreadsheets were developed by one person; the cells for activity data and any other user-defined parameters were left blank (shown in italics in Figure 3.5-1). The source-specific VOC emission factors and HAP speciation profiles were included in the spreadsheet. The user (i.e., the person calculating emissions for a category) entered the needed data and emissions were automatically calculated. A look-up function uses the HAP code to find the weight percentage (shown in the last table in the spreadsheet) used to estimate the HAPs value.

If spreadsheets must be used for inventory estimates, developing standardized spreadsheets has several advantages for QC. The spreadsheets (without activity data) can undergo QA/QC checks before emissions are calculated and all cells except those needed for data input can be locked (or protected). This will ensure that changes to the data are not made inadvertently. Final QA of the estimates is simplified because only activity data (and any other user-supplied data) need to be reviewed. This method may help eliminate a "QA bottleneck" at the end of the project by dividing the review into several segments.

Another advantage of this system is that the emissions were easily imported into a database for summation and reporting. The format of the spreadsheets was designed so that the emissions were always in a block and in the same pollutant (rows) and county (columns) order. This makes import of the results into a summary spreadsheet or a database program relatively simple. Also, it minimizes the chance of errors in data transfer and compilation.

Point Source Inventory Spreadsheet Example

The North Carolina Department of Environment, Health, and Natural Resources (NC DEHNR) has developed standardized spreadsheets for some types of emissions sources that are currently permitted. These spreadsheets are available to facilities that are preparing permit applications or annual emissions inventories (for fee assessments). An example of one of these spreadsheets is shown in Figure 3.5-2. This spreadsheet uses cell protection so that the user cannot change emission factors or other crucial data. Look-up functions are used to select the correct emission factor and to calculate

Created by: M. Adams	Data entry by:	QC by:
Date: 11/03/94	Date:	Date:
Purpose: Calculate HAPs from graphic arts in Ozoneville using EPA VOC per capita methods and speciation profiles.		
SOURCE CATEGORY:	Graphic Arts	
AMS Code	2425000000	

County	County A	County B	County C
Zone 1 population	1,364,576	510,524	368,416
Zone 2 population	142,743	75,679	97,226
Total population, Zones 1 & 2	1,507,319	586,203	465,642
VOC emission factor, lb/capita	1.3	1.3	1.3
Adjustments (if none, enter 1)	1	1	1
VOC emissions (pounds)	1,959,515	762,064	605,335

HAP NAME	HAP CODE	HAP Emissions (pounds/year)		
Xylene	1	189,955	73,874	58,681
Toluene	2	229,812	89,375	70,994
Glycol ethers	3	346,834	134,885	107,144
Methyl ethyl ketone	4	70,543	27,434	21,792
Ethylbenzene	5	12,169	4,732	3,759
Ethylene glycol butyl ether	6	109,733	42,676	33,899
Diethylene glycol butyl ether	7	92,097	35,817	28,451
Other diethylene glycol ethers	8	74,462	28,958	23,003
Propylene glycol methyl ether acetate	9	21,555	8,383	6,659
Ethylene glycol ethyl ether acetate	10	13,717	5,334	4,237
Other glycol ethers and esters	11	35,271	13,717	10,896
Glycols (unsubstituted)	12	35,271	13,717	10,896
County Total		1,231,418	478,904	380,410

FIGURE 3.5-1. EXAMPLE OF STANDARDIZED SPREADSHEET FOR AN AREA SOURCE

SPECIATION PROFILES

HAP	HAP CODE	WEIGHT PERCENT OF VOC VAPORS Graphic Arts
Xylene	1	9.69%
Toluene	2	11.73%
Glycol ethers	3	17.70%
Methyl ethyl ketone	4	3.60%
Ethylbenzene	5	0.62%
Ethylene glycol butyl ether	6	5.60%
Diethylene glycol butyl ether	7	4.70%
Other diethylene glycol ethers	8	3.80%
Propylene glycol methyl ether acetate	9	1.10%
Ethylene glycol ethyl ether acetate	10	0.70%
Other glycol ethers and esters	11	1.80%
Glycols (unsubstituted)	12	1.80%

FIGURE 3.5-1. CONTINUED

Natural Gas Combustion: Criteria and Toxic Air Pollutant Emissions — Central Office Version

NAME OF FACILITY:	No Name Co.	APP	
CITY:	Rocky Mount	Review Engineer	R. Person
COUNTY:	Nash		

Boiler Input Data

BOILER TYPE >	2			
1 - LARGE BOILER > 100 mmBtu/hr				
2 - SMALL BOILER 10 - 100 mmBtu/hr				
3 - COMMERCIAL < 10 mmBtu/hr				
ACTUAL NG BURNED :	35.00	mm cu. ft/yr		
			.0503 PM Allowable	
FOR POTENTIAL :			0.47	lb/mmBtu
ENTER BOILER SIZE :	25	mmBtu/hr	11.9	lb/hr @ max. input
HOURS OF OPERATION :	8,760	hr/yr	.0516 SO2 Allowable	
ENTER % CONTROL EFF. :	0.00	%	58	lb/hr @ max. input

ACTUAL CRITERIA EMISSIONS

Pollutant	Factor (lb poll./1E6 cu. ft burned)	Emission Rates (lb/yr)	(lb/hr)	(tpy)
TSP	13.7	480	0.05	0.24
PM-10	13.7	480	0.05	0.24
SOx	0.6	21	0.00	0.01
NOx	140.0	4,900	0.56	2.45
VOCs	2.8	97	0.01	0.05
CO	35.0	1,225	0.14	0.61

ACTUAL TOXIC EMISSIONS

Pollutant	Factor (lb poll./1E12 Btu)	Emission Rates (lb/yr)	(lb/hr)	(tpy)
FORMALDEHYDE	220.3	7.9	9.07E-04	0.004

Potential NG usage :	214	mm cu. ft/yr	24,379	cu ft/hr
Heating value :	1,030	Btu/cu. ft		

NOTES:
FACTOR #1 - AP-42, SUPPLEMENT F (7/93), SEC 1.3
FACTOR #2 - EPA 450/2-90-011

ASSUMPTIONS:
- TSP, PM10, AND POM ARE AFFECTED BY CONTROL
- HEAT CONTENT OF NG 1030 Btu/cu.ft
VOCS = TOC*(1-%METHANE)

Disclaimer
"This spreadsheet is for your use only and should be used with caution. DEHNR does not guarantee the accuracy of the information contained. This is a draft spreadsheet and is continually revised and updated. It is your responsibility to be aware of the most current information available. DEHNR is not responsible for errors or omissions that may be contained herein."

FIGURE 3.5-2. EXAMPLE OF STANDARDIZED SPREADSHEET FOR POINT SOURCE ACTUAL EMISSIONS

emissions. If the user wishes to calculate potential emissions, the "ACTUAL NG BURNED" box is left blank; the spreadsheet then uses the data shown under "FOR POTENTIAL" to calculate potential emissions. An example is shown in Figure 3.5-3.

This spreadsheet illustrates some other good QA features. The sources of the factors and the assumptions used are shown in the "NOTES" near the bottom of the page. This spreadsheet does not show a version number (or date created/revised), but some other DEHNR spreadsheets do. Some sort of document control system that uses identifiers, dates, and/or authorship to track different versions of spreadsheets is a good idea, especially when they are distributed widely.

5.2.4 SPREADSHEET AUDITING FUNCTIONS

For emissions that are calculated by spreadsheet applications, audit functions built into most spreadsheet software provide the user and checker with a helpful tool for verifying formula calculations and cell references. Although many spreadsheets have their own specific auditing functions, some of the more common audit options include cell dependency, cell precedence, and error identification. Although these are three different functions, they all share one important characteristic: they help the user identify which cells are referenced in spreadsheet formulas.

Figure 3.5-4 is an example of Excel tracing cell precedents. By highlighting a cell with a formula in it and selecting the *tools|audit|cell precedents* function, Excel displays on-screen arrows that point to the cells used in the calculation, allowing an auditor to quickly and easily verify formulas and calculations. The Lotus audit function operates in a similar fashion by highlighting all cells used in formula calculations.

Both Excel and Lotus provide the user with the ability to generate a description of cells involved in a calculation. Figure 3.5-5 illustrates the "report" generated by Lotus when a cell is selected and the *audit|trace dependents* function is activated.

5.3 STAND-ALONE PROGRAMS

Additional QA/QC functions are available through the use of stand-alone programs developed primarily for this purpose. These include programs with the same types of features as those described above, as well as graphical and mapping systems that can be used to perform visual auditing of data.

5.3.1 EIQA

Emission Inventory Quality Assurance (EIQA) is a software program specifically designed to provide QA/QC capabilities. EIQA provides an automated means of

Natural Gas Combustion: Criteria and Toxic Air Pollutant Emissions — Central Office Version

NAME OF FACILITY: No Name Co. — APP

CITY: Rocky Mount — Review Engineer R. Person

COUNTY: Nash

Boiler Input Data

BOILER TYPE >	2			
1 - LARGE BOILER > 100 mmBtu/hr				
2 - SMALL BOILER 10 - 100 mmBtu/hr				
3 - COMMERCIAL < 10 mmBtu/hr				
ACTUAL NG BURNED :		mm cu. ft/yr		
			.0503 PM Allowable	
FOR POTENTIAL :			0.47	lb/mmBtu
ENTER BOILER SIZE :	25	mmBtu/hr	11.9	lb/hr @ max. input
HOURS OF OPERATION :	8,760	hr/yr	.0516 SO2 Allowable	
ENTER % CONTROL EFF. :	0.00	%	58	lb/hr @ max. input

POTENTIAL CRITERIA EMISSIONS

Pollutant	Factor (lb poll./1E6 cu. ft burned)	Emission Rates (lb/yr)	(lb/hr)	(tpy)
TSP	13.7	2,926	0.33	1.46
PM-10	13.7	2,926	0.33	1.46
SOx	0.6	128	0.01	0.06
NOx	140.0	29,898	3.41	14.95
VOCs	2.8	595	0.07	0.30
CO	35.0	7,474	0.85	3.74

POTENTIAL TOXIC EMISSIONS

Pollutant	Factor (lb poll./1E12 Btu)	Emission Rates (lb/yr)	(lb/hr)	(tpy)
FORMALDEHYDE	220.3	48.5	5.53E-03	0.024

Potential NG usage :	214	mm cu. ft/yr	24,379	cu. ft/hr
Heating value :	1,030	Btu/cu. ft		

NOTES:
FACTOR #1 - AP-42, SUPPLEMENT F (7/93), SEC 1.3
FACTOR #2 - EPA 450/2-90-011

ASSUMPTIONS:
- TSP, PM10, AND POM ARE AFFECTED BY CONTROL
- HEAT CONTENT OF NG 1030 Btu/cu.ft
VOCS = TOC*(1-%METHANE)

Disclaimer

"This spreadsheet is for your use only and should be used with caution. DEHNR does not guarantee the accuracy of the information contained. This is a draft spreadsheet and is continually revised and updated. It is your responsibility to be aware of the most current information available. DEHNR is not responsible for errors or omissions that may be contained herein."

FIGURE 3.5-3. POINT SOURCE SPREADSHEET TOGGLED TO PRODUCE POTENTIAL EMISSIONS

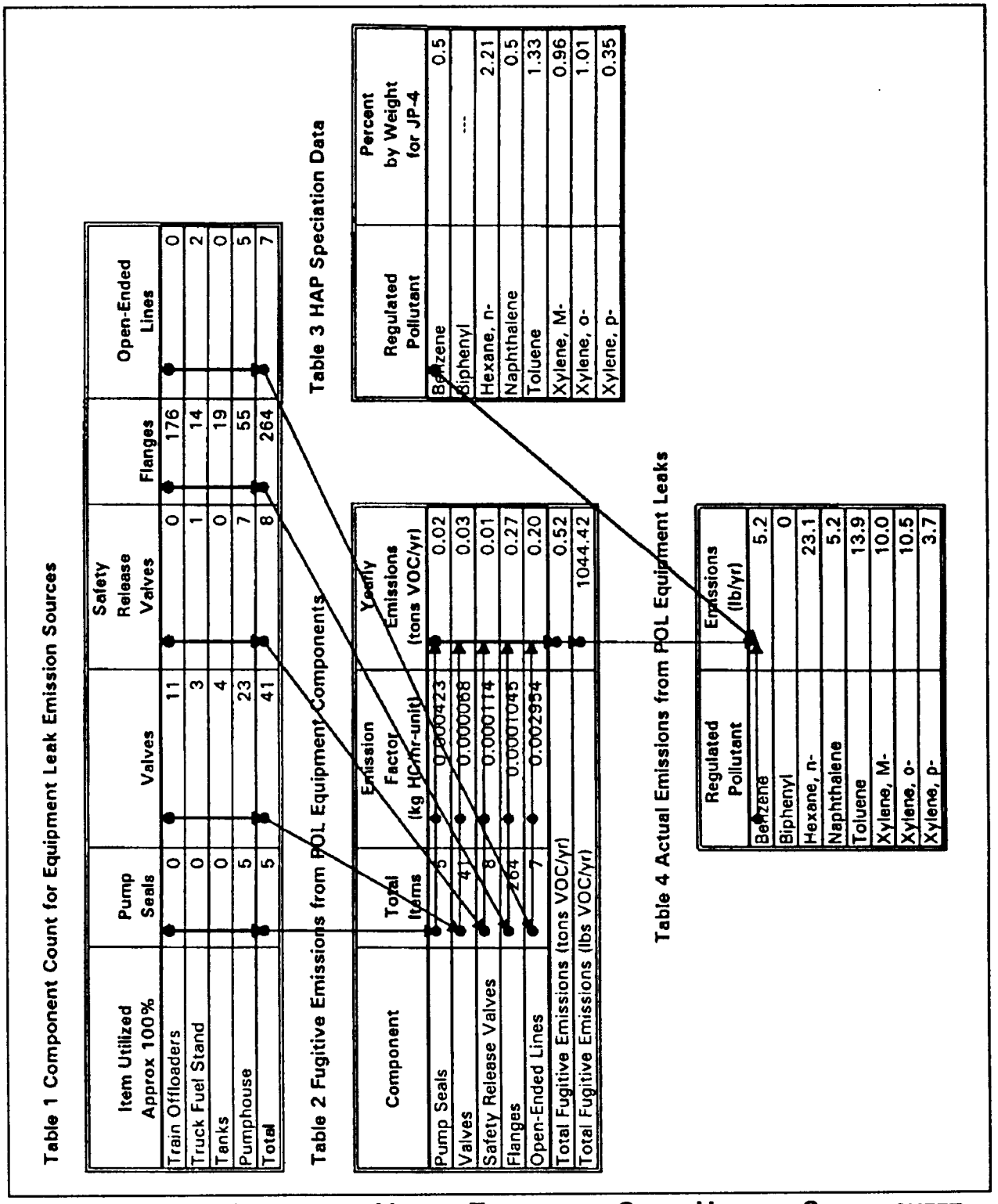

FIGURE 3.5-4. EXAMPLE OF VISUAL TRACING OF CELLS USING A SPREADSHEET AUDIT FUNCTION

Table 1 Component Count for Equipment Leak Emission Sources

Item Utilized Approx 100%	Pump Seals	Valves	Safety Release Valves	Flanges	Open-Ended Lines
Train Offloaders	0	11	0	176	0
Truck Fuel Stand	0	3	1	14	2
Tanks	0	4	0	19	0
Pumphouse	5	23	7	55	5
Total	5	41	8	264	7

Table 2 Fugitive Emissions from POL Equipment Components

Component	Total Items	Emission Factor (kg HC/hr-unit)	Yearly Emissions (tons VOC/yr)
Pump Seals	5	0.000423	0.02
Valves	41	0.000068	0.03
Safety Release Valves	8	0.000114	0.01
Flanges	264	0.0001045	0.27
Open-Ended Lines	7	0.002954	0.20
Total Fugitive Emissions (tons VOC/yr)			0.52
Total Fugitive Emissions (lbs VOC/yr)			1044.42

Table 3 HAP Speciation Data

Regulated Pollutant	Percent by Weight for JP-4
Benzene	0.5
Biphenyl	---
Hexane, n-	2.21
Naphthalene	0.5
Toluene	1.33
Xylene, M-	0.96
Xylene, o-	1.01
Xylene, p-	0.35

FIGURE 3.5-5. EXAMPLE OF CELL FORMULAS LISTING USING A SPREADSHEET AUDIT FUNCTION

Table 4 Actual Emissions from POL Equipment Leaks

Regulated Pollutant	Emissions (lb/yr)
Benzene	5.2
Biphenyl	0
Hexane, n-	23.1
Naphthalene	5.2
Toluene	13.9
Xylene, M-	10.0
Xylene, o-	10.5
Xylene, p-	3.7

Precedents of Cell A:D38
Current file
A:C38: 'Benzene
A:F22: 'Benzene
A:F23: 'Biphenyl
A:F24: 'Hexane, n-
A:F25: 'Naphthalene
A:F26: 'Toluene
A:F27: 'Xylene, M-
A:F28: 'Xylene, o-
A:F29: 'Xylene, p-
A:G22: 0.5
A:G23: ^--
A:G24: 2.21
A:G25: 0.5
A:G26: 1.33
A:G27: 0.96
A:G28: 1.01

A:G29: 0.35
A:D28: +D27*2000
A:D27: @SUM(D22..D26)
A:D22: +B22*C22*8760/907.
A:D23: +B23*C23*8760/907.
A:D24: +B24*C24*8760/907.
A:D25: +B25*C25*8760/907.
A:D26: +B26*C26*8760/907.
A:B22: +B11
A:C22: 0.000423
A:B23: +C11
A:C23: 6.8E-05
A:B24: +D11
A:C24: 0.000114
A:B25: +E11
A:C25: 0.0001045
A:B26: +F11
A:C26: 0.002954

A:B11: @SUM(B7..B10)
A:C11: @SUM(C7..C10
A:D11: @SUM(D7..D10
A:E11: @SUM(E7..E10)
A:F11: @SUM(F7..F10)
A:B7: 0
A:B8: 0
A:B9: 0
A:B10: 5
A:C7: 11
A:C8: 3
A:C9: 4
A:C10: 23
A:D7: 0
A:D8: 1
A:D9: 0
A:D10: 7
A:E7: 176

A:E8: 14
A:E9: 19
A:E10: 55
A:F7: 0
A:F8: 2
A:F9: 0
A:F10: 5

FIGURE 3.5-5. CONTINUED

performing QA edit checks on critical data elements in the emissions inventories submitted by the states to the EPA's Aerometric Information Retrieval System (AIRS); it has also been adapted for use on non-AIRS data. The EIQA program is available for general use on EPA's mainframe computer; however, user support is not provided by the EPA. Reports can be generated that examine the data for point, area, and mobile sources, and MOBILE model input data. By providing the user with an automated checking and reporting system, the net result should be significantly improved quality of the emissions inventory data stored in AIRS. A sample of QA reports generated by EIQA for point, area, and mobile source data are listed below.

Point source data QA reports include:

- Range Checks [compares inputs on an SCC-basis to a range of expected values from 1985 National Acid Precipitation Assessment Program (NAPAP) inventory];
- Consistency Checks (ensures that certain data elements are correct);
- Report Emission by Miscellaneous SCC Summary [shows emission totals by the SCC that are "miscellaneous" source types of VOCs, NO_x, and carbon monoxide (CO)];
- Report Emission by Method Code Summary (reflects the percentage of the total emissions for a pollutant that was obtained by a certain method);
- Report Emission by Source Magnitude Summary (indicates the magnitude of source in the inventory region);
- Report Source Category Completeness Summary (highlights where "expected" data are missing);
- Report Rule Effectiveness Category Summary (reports on the existence of rule-effectiveness estimates in the AIRS Facility Subsystem (AFS) for selected SIC/SCC combinations that are expected to have rule-effectiveness estimates); and
- Report Plants Exceeding 500 Tons Per Year (tpy) Summary [indicates the plants exceeding 500 tpy per pollutant listed (classified by pollutant)].

Figure 3.5-6 shows an example of a source category completeness summary.

EMISSION INVENTORY QUALITY ASSURANCE SYSTEM
POINT SOURCES
SOURCE CATEGORY COMPLETENESS SUMMARY REPORT F

BASE YEAR: 1990 SIP PROGRAM: OZONE STATUS AREA CODE: N0200 DATE REPORT GENERATED: 07JUL94

DEFINITION OF REPORT F:
These source categories and ranges are expected in most nonattainment areas as listed in EPA's Quality Review Guidelines.
An "N" in the "PRESENT" column may indicate and incomplete inventory.

PRIORITY RANKING SCHEME:
1 = High Priority - Resolution/Determination Required
2 = Warning/Advisory - No Formal Action Required
3 = Informative - See Text for Further Description

EMISSION INVENTORY QUALITY ASSURANCE SYSTEM
POINT SOURCES
SOURCE CATEGORY COMPLETENESS SUMMARY REPORT F

BASE YEAR: 1990 SIP PROGRAM: OZONE STATUS AREA CODE: N0200 DATE REPORT GENERATED: 07JUL94
Priority 1: High Priority

Exception Resolution	SCC	Source Category Name	Emission Magnitude Range	Pollutant	Present
____	4-05-xxx-xx	Graphic Arts	10-25 tpy	VOC	N
____	1-03-xxx-xx	Commercial Boilers	10-25 tpy	NO_x	N
____	1-03-xxx-xx	Commercial Boilers	10-25 tpy	CO	Y
____	4-04-xxx-xx	Petroleum Storage	>25 tpy	VOC	N
____	4-06-xxx-xx	Petroleum Market & Transport	>25 tpy	VOC	N
____	3-xx-xxx-xx	Industrial Processes	all ranges	all	Y
____	1-02-xxx-xx	Industrial Boilers	all ranges	NO_x	Y
____	1-02-xxx-xx	Industrial Boilers	all ranges	CO	Y
____	1-01-xxx-xx	Utility Boilers	all ranges	NO_x	N
____	1-01-xxx-xx	Utility Boilers	all ranges	CO	N

FIGURE 3.5-6. EIQA POINT SOURCES COMPLETENESS REPORT

Area source data QA reports include:

- Activity Level Unit Range Check Exceptions (the AMS database includes acceptable SCC-specific activity levels; these data are used to check the reported activity level units in the inventory);

- Fuel Loading Factor Range Check Exceptions (the minimum and maximum fuel loading factors, as found in the EPA document, *Compilation of Air Pollutant Emission Factors Volume I: Stationary Point and Area Sources*, commonly referred to as AP-42, are used to identify the inventory's outliers);

- Base Year Factor Range Check Exceptions (reports records in which the base year does not equal the designated year, 1990);

- Annual Emissions Range Check Exceptions (the EPA document, *Quality Review Guidelines for 1990 Base Year Emission Inventories*, defines per capita range checks for area source emission categories; the program uses these ranges to identify SCCs that are apparently under- or over-reported);

- SIP Rule in Place Consistency Check (identifies area sources that may be covered under SIP rules);

- Pollutant Consistency Check Exceptions (checks for pollutants not reported for certain SCC codes);

- Gasoline Consumption Consistency Check Exceptions [compares gasoline consumption to population estimates and expected state-level Department of Energy (DOE) gasoline consumption estimates];

- Population Consistency Check Exceptions (highlights large population changes);

- Emission Factor Units Consistency Check Exceptions (highlights where factor units are not what were expected);

- SCC Consistency Check Exceptions (highlights potentially missing SCCs); and

- Annual Emission Consistency Check Exceptions (highlights large changes in emissions from prior levels).

An example of an area source report (SIP Rule in Place Consistency Check) is shown in Figure 3.5-7.

Mobile source data QA reports include:

- Range Check Exception Report (expected values for specified vehicles);
- Period Consistency Check Exception Report (checks are designed to ensure internal consistency of source data);
- VMT Consistency Check Exception Report (checks reported VMTs against expected values);
- Data Item Completeness Exception Report (checks for amount of road type, expected pollutant levels, etc.);
- SCC, VMT, and Emissions Summary Report (assists user in performing manual examination of inventory data);
- VMT and Emissions Summary Report (assists user in performing manual examination of inventory data); and
- Emission Factors Summary Report (provides summary tables for total or time-of-day emission factors).

MOBILE model input data QA reports include:

- Geographic Consistency Exception Report (checks status of Inspection and Maintenance (I/M) programs, Anti-tampering Programs (ATP), oxygenated fuels in input files);
- I/M and ATP Control Program Consistency Exception Report (checks I/M and anti-tampering flag parameters);
- I/M and ATP Program Characteristic Range Check Exception (checks I/M, anti-tampering and fuel-related characteristics against expected values);
- Scenario Range Check Exception Report (checks model scenario inputs against expected values);

EMISSION INVENTORY QUALITY ASSURANCE SYSTEM
AREA SOURCES
SIP RULE IN PLACE CONSISTENCY CHECK EXCEPTION REPORT E

BASE YEAR: 1990 FILENAME: AREADAT DATE REPORT GENERATED: 07JUL94

DEFINITION OF REPORT E:
Area sources may be covered under SIP rules in some areas.
Check that this report is consistent with local rules.

PRIORITY RANKING SCHEME:
1 = High Priority - Resolution/Determination Required
2 = Warning/Advisory - No Formal Action Required
3 = Informative - See Text for Further Description

EMISSION INVENTORY QUALITY ASSURANCE SYSTEM
AREA SOURCES
SIP RULE IN PLACE CONSISTENCY CHECK EXCEPTION REPORT E

BASE YEAR: 1990 FILENAME: AREADAT DATE REPORT GENERATED: 07JUL94
Priority 1: High Priority

Exception Resolution	ST	CNTY	CITY	ZN	SCC	SCC Description	Pollutant	SIP Rule in Place Reported	SIP Rule in Place Expected	Rule Effective-ness	Rule Penetra-tion
___	18	005			24-15-300-000	Degreasing	VOC	Y	N	80%	60%
___	18	005			24-20-000-000	Drycleaning	VOC	N	Y	0%	0%

FIGURE 3.5-7. EIQA AREA SOURCES SIP RULE IN PLACE CONSISTENCY CHECK

- Vehicle Speed Summary Report (provides overview of speeds used in model run; and

- Vehicle Registration Distribution Report (provides an overview of model vehicle registration distributions).

5.3.2 MOBILE INPUT DATA ANALYSIS SYSTEM (MIDAS)

MIDAS is a data processing system to facilitate the storage, retrieval, and analysis of input file variables used in mobile source modeling for SIP emissions inventories. It is designed to support evaluating the contents of MOBILE input files, which are intended to reflect all of the possible factors that could affect motor vehicle emissions within a nonattainment area.

The reports generated by MIDAS provide a convenient means of summarizing the input data and flagging potential problems. Because these data are often provided to an inventory preparer by another agency (e.g., the state's transportation department), the inventory preparer may wish to pass these reports back to the data originator for review and comment. Used in this way, MIDAS can be a useful tool to facilitate dialogue between the two agencies. Software-specific reports that MIDAS can generate are listed below:

Input File Reports

These reports provide information about the MOBILE input variables stored in the MIDAS database. Each report summarizes the information for a single database record that contains the variables for a single input file:

- Control Flag Summary (provides a synopsis of the MOBILE flag settings);

- Detailed Optional Input (provides an in-depth view of some of the optional inputs that can be modeled);

- Scenario Summary [allows user to look at contents of each individual scenario record and local area parameter (LAP) record];

- Input File Comments (provides an analysis of an input file as to its likely acceptability when used for a SIP emission inventory); and

- Input File Summary (summarizes the information contained in the previous four reports in order to provide a one-page synopsis of the input file variables and their likely acceptability for SIP modeling).

Area Reports

These reports provide information about the MOBILE input variables used to model an entire nonattainment area:

- Area Control Flag Summary (similar to Control Flag Summary);
- Area Variable Summary (presents selected parameters used to model various inventories for a nonattainment area); and
- Area Registration Summary (provides user-supplied registration distribution for each input file within the selected area).

Miscellaneous Reports

MIDAS also provides the following miscellaneous reports:

- Nonattainment Area Summary (provides a detailed description of an ozone or CO nonattainment area); and
- Database Summary (provides a listing of all the input file records that are currently in the database).

Figure 3.5-8 illustrates the use of MIDAS for a QA review of Virginia's 1990 SIP on-road mobile source inventory. The error messages summarized in the table focused the reviewer's attention on inventory data that departed from the expected (in this case, the MOBILE5a defaults). The discussion of the review provides several important pieces of information to the inventory preparer:

- The specific cause of the error flag is described;
- The relative importance of this possible error is assessed; and
- A recommended action is stated.

It should be noted that the agency did follow up on these recommendations and no errors were found.

5.3.3 VISUAL AUDITING PROGRAMS

Data visualization programs and methods can also be used as part of a QA/QC program to review data and emissions. These programs are capable of displaying the

Error Message from MIDAS Input File Comments Report	Counties/Cities Whose Input File Showed Error Message		Item Requiring Further Review	Noncritical Item
For vehicle types ____, . . . , the percentage of age 1 vehicles is high relative to the percentage of age 2 vehicles.	Charles City County Chesterfield County Chesapeake City Hampton City Hanover County Henrico County Newport City	Norfolk City Poquoson City Portsmouth City Richmond City Suffolk City Virginia Beach City York County	✓ (Norfolk City and Richmond City only)	✓ (all other cities/counties)
For vehicle types ____, . . ., the age 25 distribution is too low.	Charles City County Chesapeake City Hampton City Hanover County Henrico County Newport City Norfolk City	Poquoson City Portsmouth City Richmond City Suffolk City Virginia Beach City York County	✓	
The fraction of LDDT is too high.	Chesterfield County Colonial Heights City Hanover County	Henrico County Hopewell City Richmond City	✓	
Spikes occur in the registration distribution.	Charles City County Chesterfield County Chesapeake City Hampton City Hanover County Henrico County James City County	Newport City Norfolk City Poquoson City Portsmouth City Richmond City Suffolk City York County		✓

FIGURE 3.5-8. SUMMARY OF RESULTS FROM MIDAS REVIEW OF THE MOBILE5A INPUT FILES

data in a more visual format (i.e., maps, timelines, scatter plots). Examples of such programs are Voyager, ARC View, and Map INFO. These software packages allow the user to browse quickly and easily through large spatial and temporal databases. Because the data are always displayed graphically, the user can see relationships that might otherwise have remained obscure or invisible. All of these packages allow the user to import data files in common formats (usually "DBF"). A linked version of Voyager, SIPView, was developed and released by the EPA specifically for use with SIP inventories.

Figure 3.5-9 shows an example from Voyager (Misenheimer, 1996). A point source emissions inventory for Georgia is plotted on a map and the relative magnitudes of the emissions are shown by the vertical bars. The circled items show sources with incorrect UTM coordinates (that is, the points are outside of state boundaries). The reviewer could check each point by displaying the data associated with each point.

If a data visualization software package is to be used as part of a QA/QC program, the following features are desirable:

- Ability to correct the data interactively will make this tool more useful for QC; that is, the user should be able to select the point, display the data, update or correct if necessary, and redisplay it.

- Ability to print reports will allow easy documentation of QA/QC activities.

- Ability to print graphics will also allow documentation of QA/QC activities. Packages that only plot points are less useful than those that display values for variables associated with these points (e.g., bars representing emissions).

5.4 User-defined Automated QA Checks

Inventory preparers with access to computer software and the skills to fully exploit that software may wish to create their own automated QA systems. Some emissions-estimation software programs provide the user with the ability to "create" reports. This allows the user to specify what information is to be printed in a report. Alternatively, database software can be used to develop a program that will read the emissions data and generate reports. QA/QC activities can be facilitated by designing and using reports that present key variables for auditing.

A relatively sophisticated system of automated QA programs has been developed by air quality modelers in Texas (Neece and Smith, 1994). They had previously identified point source emissions as a critical component of ozone modeling and developed a

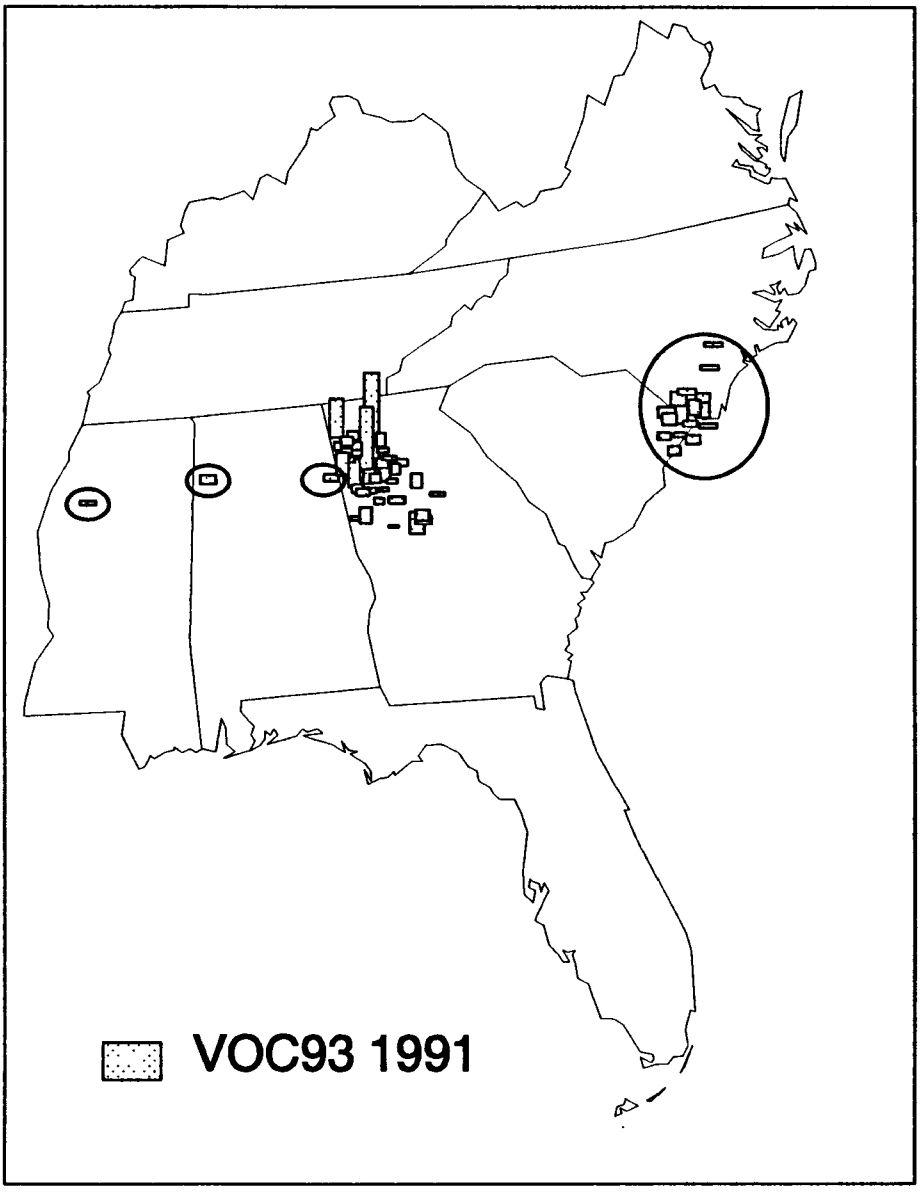

FIGURE 3.5-9. EXAMPLE OF VISUAL AUDITING TOOL

series of checks, many of which are specifically designed to catch errors that will most impact modeling results (see also Section 6, "Sensitivity Analysis," of this chapter).

Two types of procedures are involved. The first is a series of database checks. One of particular interest is a ratio test. The ratio of VOC/NO_x emissions from an emissions unit is compared to the ratio of VOC/NO_x factors for that unit. Emission factors are not always used to estimate emissions, so these ratios may not always agree exactly. However, extreme disparities in the ratios are investigated (See Neece and Smith, 1994, for descriptions and examples of all tests).

The second type of procedure uses graphical analyses to plot histograms (bar charts) showing the distributions of key data. Scatter plots are also used to identify outliers (examples are shown in Figure 3.5-10). (The use of descriptive statistical methods such as these is discussed more fully in Section 7 of this chapter.) Spatial plots of the data are used to check that plant coordinates are correct.

The system of automated QA programs developed by Neece and Smith (1994) is an excellent example of an inventory QA program designed to fulfill a specific need: that of ensuring the quality of results of an air quality model that addresses issues critical to the particular modeling exercise. The developers of an inventory do not always know how their inventory will be used and, even when they know, they may not be aware of any special concerns. It is important that the *users* of an inventory identify those data quality issues and perform their own QA checks.

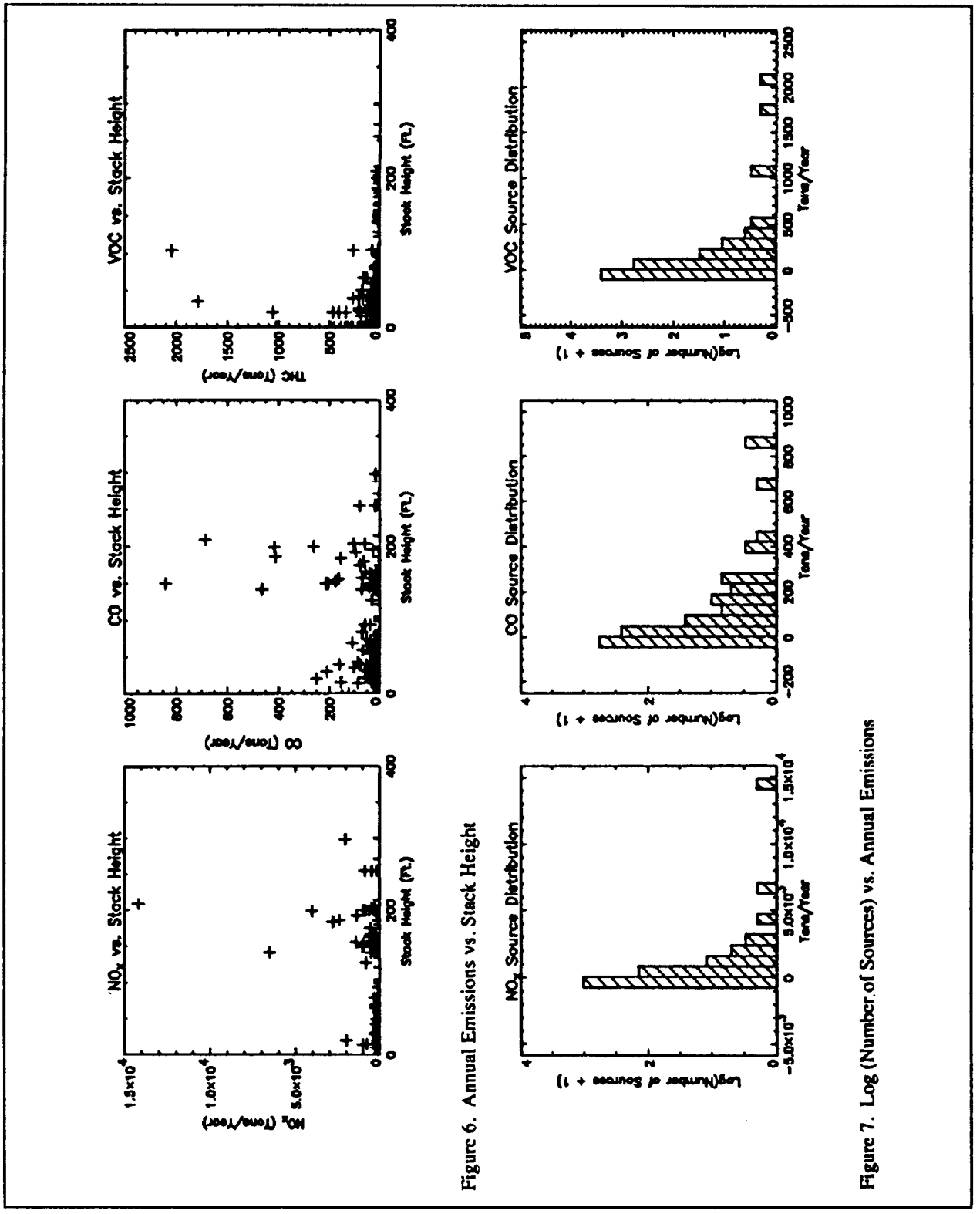

Figure 6. Annual Emissions vs. Stack Height

Figure 7. Log (Number of Sources) vs. Annual Emissions

FIGURE 3.5-10. EXAMPLES OF QA SCATTER PLOTS

6

SENSITIVITY ANALYSIS

QUICK REVIEW

Definition:	A systematic study of how changes in model (or inventory) input parameters affect the output.
Objective:	To identify the parameters that have the greatest effect on the results.
Optimal usage:	When planning both the inventory and the QA program, sensitivity analysis can be used to prioritize the efforts. The user of an inventory also needs to understand the inventory sensitivities with respect to how the inventory is to be used.
Minimum expertise required:	Moderate to high, depending on model complexity and/or availability of sensitivity analysis results.
Advantages:	Allows the most efficient utilization of limited resources by focusing efforts where they will do the most good.
Limitations:	(1) Not a substitute for QA/QC procedures to verify data and calculations, and (2) can be very time-consuming if the model or inventory is very complex.

6.1 Overview of Method

A sensitivity analysis is a process for identifying the magnitude, direction, and form (e.g., linear or nonlinear) of the effect of an individual parameter on the model's result. It is usually done by repeatedly running the model and changing the value of one variable while holding the others constant. If a model has a large number of variables, conducting a thorough sensitivity analysis can require a great deal of effort.

With respect to emissions inventories, sensitivity analyses may be used to: (1) evaluate the overall inventory, (2) identify sensitivities of specific emissions models, or (3) identify key inventory data that, when used as input to air quality models, have the potential to significantly affect the results. These types of analyses should be used when planning an inventory in order to focus both development resources and QA/QC efforts on the critical elements of the database. Sensitivity analyses will also be important to the user of the final inventory.

Sensitivity analyses of inventories can be accomplished through several steps, and even the simplest approach can yield useful results. For example, individual source categories can be ranked in order of their contribution to total emissions for a given pollutant. A variation is to rank the categories and then calculate the cumulative percent of emissions. These ranks can then be used to prioritize categories and target the QA efforts.

Somewhat more complicated, but also more useful, is identifying the impacts of underlying variables on emissions. For example, population is a commonly used surrogate for calculating area source emissions of VOCs. If area sources are a large component of a VOC inventory, small errors in the population value can have a significant impact on the emissions estimate. Knowledge of this fact might prompt an agency to conduct a survey or to use methods other than per capita factors.

Models are often used to estimate emissions for specific source categories. Table 3.6-1 lists many of these models and provides references for some publicly-available sensitivity analyses. Unfortunately, complete sensitivity analyses have not been published for many of these models.

It is ultimately up to the inventory developers to determine the significance of each source category in their inventories and direct QA efforts accordingly. If a source category contributes a relatively small proportion of the emissions, conducting a sensitivity analysis may not be worthwhile.

TABLE 3.6-1

KEY VARIABLES FOR SOME EMISSIONS ESTIMATION MODELS

Model	Description	Key Variables
MOBILE5b[a]	Mobile source emission factors	I/M, VMT mix, speed, mileage accumulation, temperature
BEIS, BEIS-2.2[b]	Biogenic emissions	Canopy effects, temperature, sunlight, biomass/land-use
Landfill Air Emissions Estimation Model[c,d]	Landfill VOC and methane emissions	Decay rate (k), methane potential of waste (L_0), nonmethane organic compound (NMOC) concentration, length of time refuse has been in place
SIMS[d]	VOC emissions from wastewater treatment	Aeration, biodegradation, surface area and wind speed, retention time
TANKS[d]	Emissions of organic liquids from fixed-roof or floating-roof tanks	Location, tank type, tank contents

[a] From data supplied by Office of Mobile Sources; for more discussion, see the example in this section.
[b] Pierce et al., 1990; Birth, 1995; see also Lamb et al., 1993; Gaudioso et al., 1994.
[c] Peer et al., 1992.
[d] From Session 10 of "Emissions Inventory Satellite Workshop Training Manual/Presentation Material."

Finally, if the inventory is being prepared as input for a model, a sensitivity analysis of that model can be very useful for inventory preparation and QA. Table 3.6-2 lists several air quality models that require emissions inventories as input. Although ensuring the accuracy of emissions is obviously important for all these models, the importance of other inventory parameters varies from model to model. The relevant temporal scale, for example, may be annual, daily, or hourly, and the maximum or the average emissions may be required. Spatially allocating emissions presents an additional complication. Some models for some pollutants may be very sensitive to spatial scales; for example, models of tropospheric ozone formation may be very sensitive to the spatial distribution of short-lived pollutants, but relatively insensitive to those with longer half-lives.

TABLE 3.6-2

KEY INVENTORY VARIABLES FOR SOME AIR QUALITY MODELS

Model	Scale[a]	Key Inventory Variables
Industrial Source Complex Long Term (ISCLT)	Local (up to 20 km), annual concentrations	• Emission rates • Point source plume rise[b] • Location of source(s) • Dimensions of buildings adjacent to sources
Industrial Source Complex Short Term (ISCST)	Local (up to 20 km), hourly, 8-hour, 24-hour	
Urban Airshed Model (UAM)	Urban (20 to 200 km), hourly	• Location of point sources • Spatial/temporal allocation of area and mobile source emissions • Emission totals per grid cell • Point source plume rise[b] • Speciation of VOC compounds • Temperatures (used to estimate biogenic emissions)
Regional Acid Deposition Model (RADM) and Regional Oxidant Model (ROM)	Regional (200 to 2000 km), hourly	• Location of point sources • Point source plume rise[b] • Speciation of VOC compounds

[a] Spatial scale conventions from NAPAP (1991) are used here.
[b] Inventory variables used to calculate plume rise are exit velocity, exit temperature, and stack diameter.

Table 3.6-3 summarizes the preferred and alternative methods for using sensitivity analyses in the development of emissions inventories.

TABLE 3.6-3

SENSITIVITY ANALYSES: PREFERRED AND ALTERNATIVE METHODS

Method	Procedure
Preferred	When preparing the Inventory Preparation Plan (IPP/QAP) for an emissions inventory, prioritize data gathering and QA/QC efforts by conducting sensitivity analyses to focus on key variables. The analyses should include: (1) previous inventories, (2) emissions models to be used in preparing the inventory, and (3) models for which the inventory will be used as input data.
Alternative 1	Use existing sensitivity analyses of emissions models and of previous inventories to identify key variables and potential trouble spots.

6.2 SENSITIVITY ANALYSIS FOR AN AREA SOURCE INVENTORY

Area source emissions are often estimated using an emission factor and a surrogate for the emissions activity. Population and employment by SIC code category are commonly used in this approach. As discussed in Chapter 1 of Volume III (Preferred and Alternative Methods for Area Sources), the use of these surrogates is a significant source of uncertainty of the emissions.

Table 3.6-4 shows the percent of VOC emissions attributable to specific area sources for a state and the methods used to estimate the emissions. A large percentage of the emissions were estimated using per capita or per employee factors. To determine just how important population and other surrogates are to the overall estimate, a simple sensitivity analysis was conducted.

The sources were assigned to six general methods groups: (1) per capita, (2) per employee, (3) throughput (this includes any category for which emissions are based on fuel use, production, or consumption of a material), (4) statistical models, (5) surveys,

TABLE 3.6-4
DAILY VOC INVENTORY FOR A STATE

Area Source Category	VOCs, lb/day	% Total	Estimation Procedure
Commercial/Consumer Solvent Use	302,854	18	per capita
Architectural Surface Coating	295,558	17	per capita
Gasoline Distribution	246,175	14	gasoline consumption
Industrial Surface Coating	240,237	14	per employee, per capita
Surface Cleaning Operations	172,429	10	per employee
Dry Cleaning Operations	73,904	4	per employee
Petro. Vessel Loading/Unloading	66,673	4	petro products loaded/unloaded
Automobile Refinishing	51,931	3	per capita
Graphic Arts Facilities	48,037	2	per capita
Asphalt Paving	32,450	1	consumption
Traffic Paints	28,976	1	per capita
Agricultural Pesticides Application	22,421	1	application rate, acres crops
Commercial Bakeries	20,654	1	per capita
Structure Fires	18,728	1	per fire
Municipal Landfills	13,496	0	statistical models
Residential Fuel Combustion	4,571	0	fuel use
Industrial Fuel Combustion	3,775	0	fuel use
Aircraft Refueling	2,111	0	aviation fuel consumption
Apartment Incinerators	563	0	engineering judgment
Wastewater Treatment	462	0	survey
Hospital Sterilizers	299	0	per hospital bed
Forest Fires	292	0	acres burned
Commercial/Institutional Fuel Combustion	266	0	fuel use
Breweries	192	0	state beer production
Barge, Tank Car, Rail Car, Drum Cleaning	119	0	survey, engineering judgment
Medical Waste Incinerators	11	0	engineering judgment
Asphalt Roofing Kettles	1.5	0	per square paper
Orchard Heaters	0	0	survey
Distilleries	0	0	survey
Agricultural/Slash Burning	0	0	survey
Wineries	0	0	state wine production
Open Burning	0	0	survey

and (6) engineering judgment. Based on expert judgment, the ranking of the estimates from highest to poorest quality was:

1. Survey
2. Throughput
3. Statistical models
4. Per employee
5. Per capita
6. Engineering judgment

(For more discussion on emissions quality ranking and/or uncertainty, see Chapter 4 of this volume.)

Figure 3.6-1 is a histogram of percent VOC (from Table 3.6-4) estimated by each of the above methods. An important finding of this simple analysis is that over 70 percent of the area source VOC emissions were estimated using two mid- to low-quality methods. The state air pollution control agency is seeking ways to reduce emissions, and viable control methods have been identified for some coatings. The Agency now is concerned that more accurate methods may be needed to evaluate the effects of controls for these population-based estimates.

Because emissions were estimated using mostly linear models (e.g., emission factor times activity), the relationship between the emissions estimate and population and/or employment is easy to understand. Figure 3.6-2 demonstrates how percentage changes in population or employment affect VOC emissions. The graph shows that a 25 percent change in population will result in a 10 percent change in *total* area source VOCs (that is, VOCs from all sources). This suggests that small errors in the population variable do not have a very large impact overall. If the inventory preparer has alternative estimates of population (or other variable used as activity data), an analysis such as this can help determine whether the choice of alternative values is a significant one.

This type of analysis can be used to target area source categories for improvement. In this example, the state agency's plans for the next inventory might include:

- Dropping the "Open Burning" and "Barge Cleaning" categories because surveys showed these were not important sources of emissions;

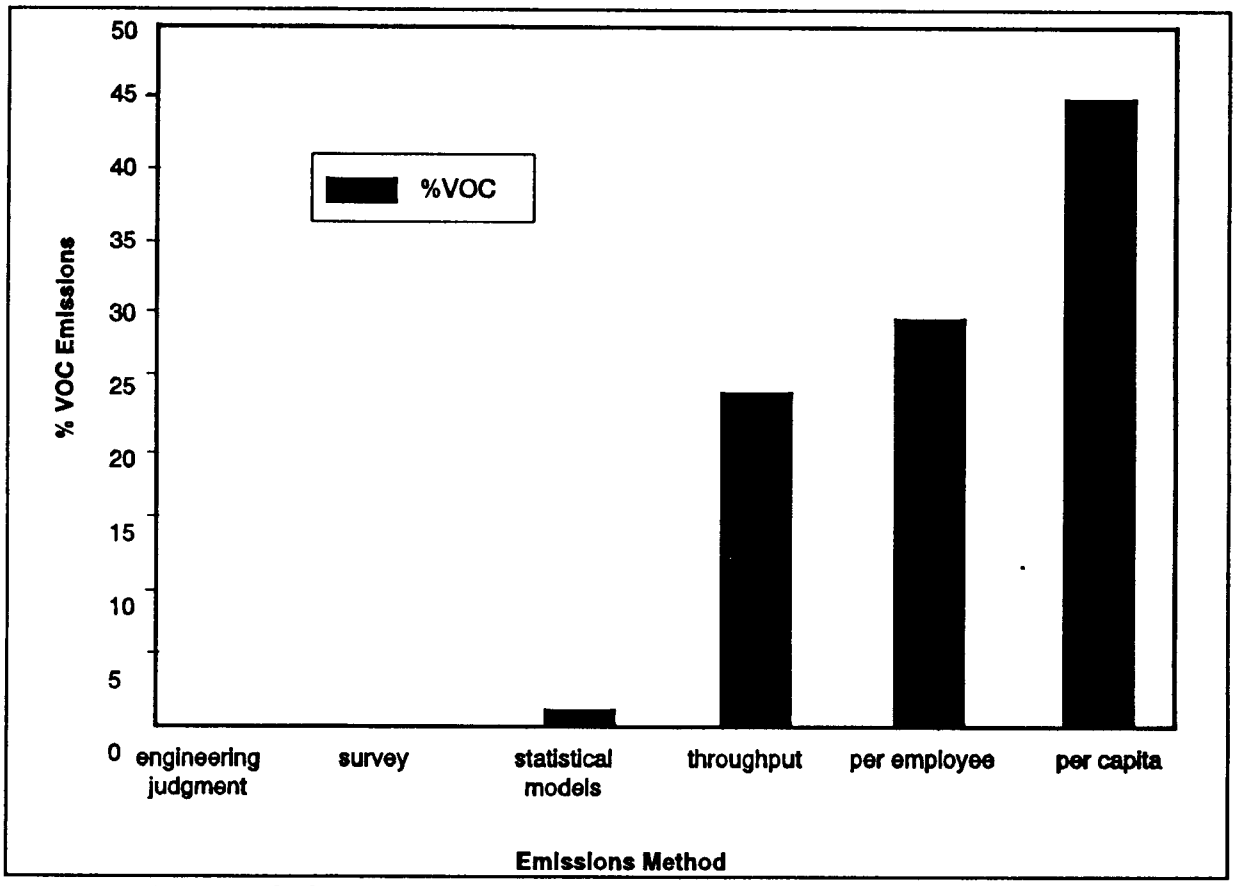

FIGURE 3.6-1. PERCENT VOC BY ESTIMATION METHODOLOGY

- Conducting a survey of paint distributors to develop more reliable estimates for "Architectural Surface Coating;" and

- Intensifying point source surveys to cover more of the industrial surface coating emissions in the point source inventory and using data from smaller sources to develop local per employee factors.

This is a very simple example of a sensitivity analysis. Its simplicity is a result of the methods used to estimate emissions. As more complicated methods are used, the complexity of the analysis increases. Also, specific results from this analysis cannot necessarily be extrapolated to other inventories. Each inventory has unique features and sensitivity analyses need to be conducted to meet specific objectives.

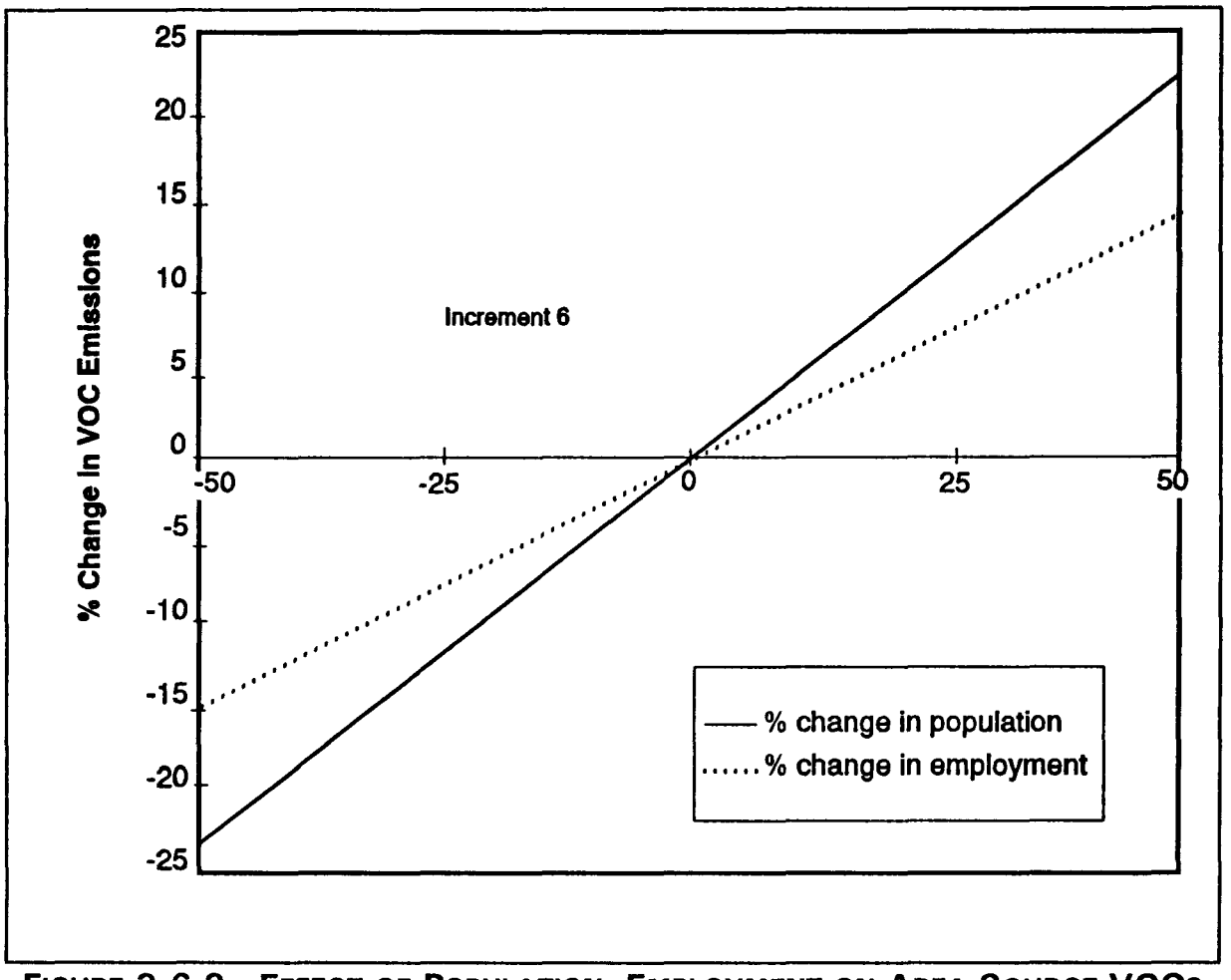

FIGURE 3.6-2. EFFECT OF POPULATION, EMPLOYMENT ON AREA SOURCE VOCS

6.3 SENSITIVITY ANALYSIS OF EPA'S MOTOR VEHICLE EMISSION FACTOR MODEL

The EPA's MOBILE model is the most commonly used method for estimating onroad mobile source emission factors. At the time this volume was written (1997), version 5b of the model was in use. A fairly detailed sensitivity analysis of version 4.1 was conducted in 1992 and the results distributed by the EPA's Office of Mobile Sources (OMS). One example from that analysis is presented here.

The MOBILE model is one of the most complex models commonly used for emissions estimation. Default data are provided for many of the input variables, but the user is encouraged to develop region-specific input data.

This sensitivity analysis compared three test cases. The conditions used for each case are summarized in Table 3.6-5. The results of the analysis for nonmethane hydrocarbon (NMHC) emissions are summarized in Figure 3.6-3a. Not surprisingly, VMT has a significant linear effect on emissions. Just as in the area source example described previously, emissions are estimated by multiplying a factor by an activity value (i.e., VMT). The other variables, however, directly affect the emission factors produced by the MOBILE model. The model was relatively insensitive to changes in Vehicle Age Distribution and VMT mix (less than 5 percent difference in emissions). Other variables had a more profound effect; for example, underestimating the average speed by as little as 5 miles per hour (mph) resulted in greater than a 20 percent increase in emissions.

However, as Figure 3.6-3b shows, model sensitivities are different for other pollutants. VMT mix is very important for summertime NO_x emissions, but temperature is not (These results are based on the same set of conditions as those in Figure 3.6-3a).

These sensitivity studies of the MOBILE model illustrate the importance of analyzing the model for all potential scenarios of interest. Pollutants may be affected differently by changes in variables; also, the time of the year may be an important factor.

6.4 Sensitivity of Air Quality Models to Inventory Input Data

The users of inventories often have little or no involvement in developing the inventory. Therefore, they may subject the inventory to intensive review and specific QA procedures designed to identify problems that could impact the outcome of their particular use of the inventory. One such group is composed of air quality modelers.

Air quality modelers typically conduct sensitivity studies of the air quality model response to changes in emissions data input to the model. These sensitivity studies many times are simple and involve across the board increases or decreases in total emissions. For example, a common sensitivity simulation usually performed for most photochemical modeling studies is the removal of all emissions from the simulation. These bounding simulations, while not realistic, provide the modeler with information as to the overall response of the photochemical model to changes in emissions.

While very useful for interpreting modeling results for a given modeling domain, emission sensitivity simulations *cannot* be generalized to all photochemical modeling studies. Every modeling study will have its own peculiarities and results of sensitivity simulations for the region cannot be extrapolated to other areas. The meteorology for a region, the magnitude and relative mix of biogenic, area, and point source emissions, the reactivity of the emissions, the initial and boundary conditions in the region, and the

TABLE 3.6-5

MODEL INPUT VALUES USED FOR MOBILE4.1 SENSITIVITY ANALYSIS

Variables	Range of Conditions			Notes
	Low	Base	High	
Inspection/ Maintenance (I/M)	Basic	None	High option enhanced	Indicates the effects of I/M options.
Travel (VMT)	75,000	100,000	125,000	A percent change in VMT causes the same percent change in emissions.
Speed (mph)	14.7	19.6 FTP speed	24.5	Indicates the effect of inputting an incorrect speed.
Temperature (°F)	65-90	70-95	75-100	Indicates the effect of a 5°F error in temperature.
Hot/Cold Distribution PCCN/PCHC/ PCCC	5.0/5.0/5.0	20.6/27.3/20.6 FTP cycle	5.0/55.0/5.0	The low condition represents fewer hot or cold starts than the FTP cycle. The high condition represents an increase in hot starts, but few cold starts.
VMT Mix	Fairbanks	MOBILE4.1 default	California	In Fairbanks more VMT is distributed to heavy duty gasoline and diesel trucks than in the MOBILE4.1 default. California has more VMT distributed to cars and heavy duty gasoline trucks.
Mileage Accumulation	Fairbanks	MOBILE4.1 default	California	The low mileage accumulation in Fairbanks on light duty autos was substituted for the MOBILE4.1 default for the low condition. California mileage accumulation is greater than the MOBILE4.1 default.
Vehicle Age Distribution	Fairbanks	MOBILE4.1 default	Phoenix	The Fairbanks vehicle fleet has a smaller percentage of older cars while Phoenix has a larger percentage than the MOBILE4.1 default.

FTP = Federal test procedure
PCCN = % of VMT accumulated by non-catalyst-equipped vehicles in cold-start mode.
PCHC = % of VMT accumulated by catalyst-equipped vehicles in hot-start mode.
PCCC = % of VMT accumulated by catalyst-equipped vehicles in cold-start mode.

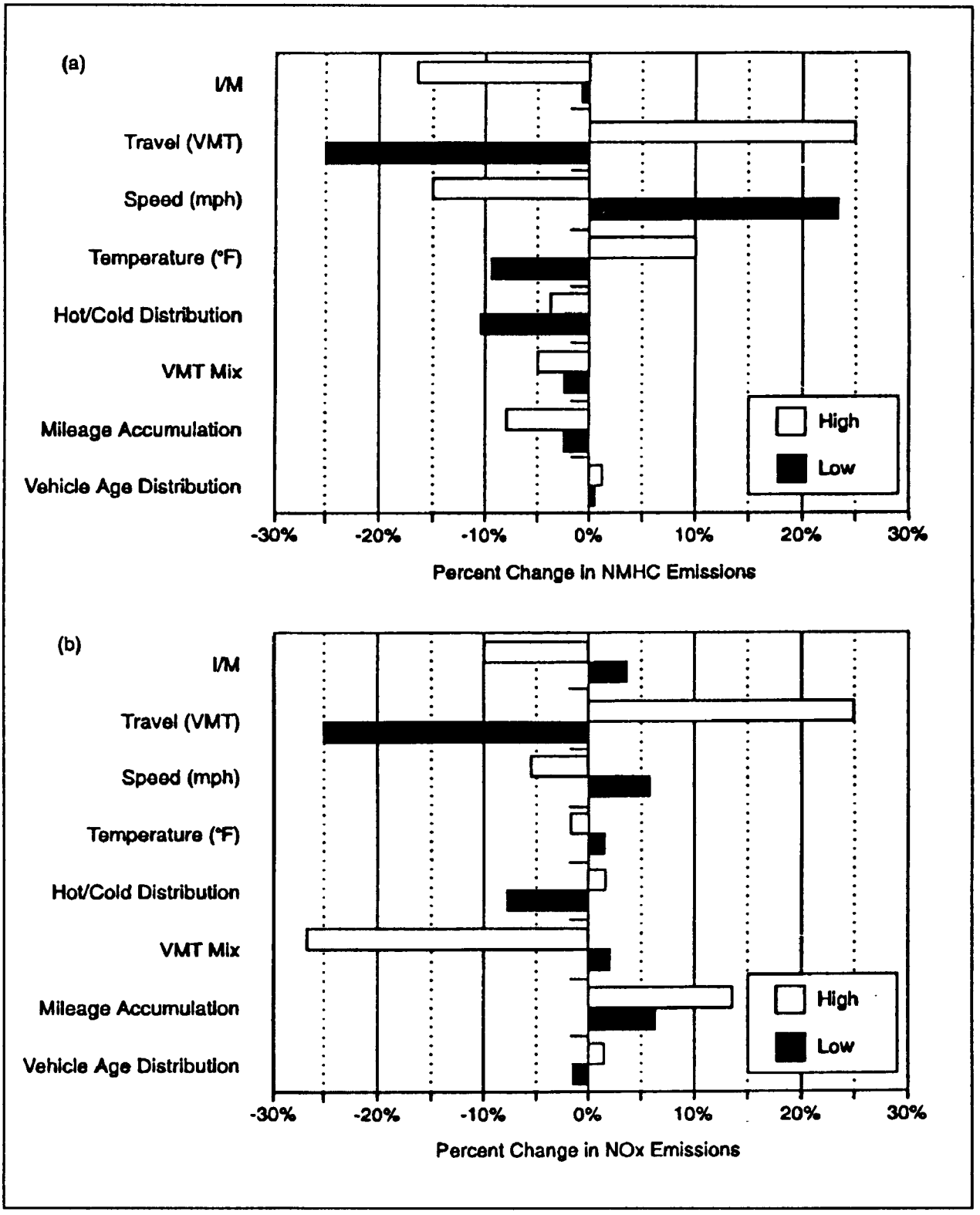

FIGURE 3.6-3(A AND B). SENSITIVITY OF NMHC, NO_x EMISSIONS

occurrence or non-occurrence of transport of precursor pollutants from outside the domain all combine to create a set of unique responses to changes in emissions in a given modeling domain.

With the above caveat in mind, it is still useful to examine the type of information that can be obtained from a properly formulated sensitivity study. Below are three examples of the type of information feedback to the inventory specialist that can result from sensitivity studies.

SOUTHERN CALIFORNIA

Numerous investigators have examined the sensitivity of ozone formation to emissions in the South Coast Air Basin (Los Angeles metropolitan area) in Southern California. A key driver for these emission sensitivity studies was the results of the tunnel study of motor vehicle emissions conducted during the Southern California Air Quality Study (SQAS) (Pierson et al., 1990). The results of this study indicated that onroad motor vehicle hydrocarbon emissions were significantly underestimated in the South Coast Air Quality Management District (SCAQMD) inventory. The SCAQMD addressed this issue in its 1994 Air Quality Management Plan (AQMP) by conducting photochemical model sensitivity studies in which hydrocarbon emissions from all onroad vehicles were doubled (SCAQMD, 1994). The SCAQMD modeling was performed using the Urban Airshed Model (UAM). For all episodes modeled, model performance for the doubled vehicle hydrocarbon emissions scenario was significantly improved above the base case.

Harley et al. (1993), using the California Institute of Technology (CIT) photochemical model for the same Los Angeles episodes, also adjusted upward motor vehicle emissions as indicated from the SCAQS tunnel study results. Again, there was a significant improvement in model performance using the upward adjusted motor vehicle emissions inventory. Results from these two modeling studies using different models and methodology provide significant insight into potential biases in the 1994 AQMP motor vehicle emission inventory and demonstrate the value of emission sensitivity studies.

TEXAS GULF COAST

A different type of emissions sensitivity analysis was performed by Neece and Smith (1994) using UAM simulations for the Texas Gulf Coast. The automated QA program developed by Neece and Smith has been previously discussed in Section 5 of this chapter. In their sensitivity analyses, they examined the sensitivity of peak ozone concentrations to individual components of the emission inventory. Based upon these analyses, they concluded:

- Peak predicted ozone concentration is more sensitive to point source emissions than to area, mobile, or biogenic source emissions.

- "Hot spots" in the grid occur where NO_x and VOC emissions are mixed; these grid cells need to be checked carefully to ensure that point source locations are correct.

- The emissions from large sources outside the boundary of the modeling domain are not detected by UAM; therefore, a test was designed to identify large sources just outside the upwind boundary.

As a consequence of their concern over these model sensitivities, Neece and Smith developed their own automated QA procedures to review point source inventories.

ATLANTA

Biogenic emissions are a very significant source of hydrocarbon emissions in the Atlanta, as first reported by Chameides et al. (1988). In a subsequent analysis of biogenic emissions in the Atlanta area, Cardelino (1995) analyzed the impact on ozone precursor emissions and ozone concentrations of changes in land use patterns, alternative estimates of isoprene emissions from biogenic sources, and variation in temperature reflecting different land use patterns. Based on this analysis, Cardelino concluded:

- There remains uncertainty in the most appropriate algorithm for estimating biogenic emissions of isoprene, a highly reactive component of biogenic emissions.

- Satellite imagery can potentially provide improved representation of temperature fields, leading to improvements in the temporal distribution of emission inventories and to potential improvements in modeled ozone concentrations.

- Future population changes and urbanization can significantly change land use with resulting changes in ambient surface temperatures. These changes should be taken into account when developing future year emissions estimated and when performing photochemical modeling simulations.

The above studies reflect a sample of the information feedback available to the inventory specialist from the performance of photochemical modeling emission sensitivity studies. Again, each sensitivity result cannot be reliably extrapolated outside the specific modeling domain for which the results were obtained. However, the

inventory specialist is encouraged to work with the end-users of inventories to design the appropriate sensitivity studies for a given area and inventory. By gaining an appreciation of the ultimate uses to which an inventory is subjected, end-use concerns can be addressed when planning the inventory, preparing the QAP, and in subsequent inventory development and QA/QC activities. The active involvement of the inventory specialist with the end-use of the inventory will tend to minimize additional analyses and QA that modelers and other end-users perform to better understand the inventory.

This page is intentionally left blank.

7

STATISTICAL CHECKS

QUICK REVIEW

Definition:	The use of descriptive statistics, or statistical tests or procedures to evaluate or control data quality.
Objective:	Descriptive statistics are most commonly used to compare results, and to flag unusual or unlikely values. More formal statistical procedures may use specified sampling methods and performance specifications as a quality control procedure.
Optimal usage:	Descriptive statistics are useful for QA as part of a Peer Review or Independent Audit of a completed or partially completed inventory; particularly useful for comparing different inventories. Statistical QC is potentially a very useful method for inventory QC.
Minimum expertise required:	Moderate to high, depending on how much the system is automated, and the complexity of the analysis.
Advantages:	(1) Can process large amounts of data, (2) can reduce subjectivity of informal reality checks, and (3) contributes to the continuous improvement of inventories.
Limitations:	(1) More sophisticated methods require some understanding of statistical theory, (2) common statistical methods are based on assumptions of normality (which is often not the case for emissions data), and (3) not a replacement for human evaluation and judgment.

7.1 OVERVIEW OF METHODS

Statistical procedures offer a wide array of tools that perform many of the same QA/QC functions described in previous sections. As with the automated checks discussed in Section 5, statistical checks are generally used as a tool to facilitate reality checks (Section 2), peer reviews (Section 3), or independent audits (Section 8). They are also used in some emissions validation procedures (Section 9). Many of the statistical methods described in this system can be automated, and probably should be if the data sets are very large or if the procedure is to be repeated many times.

There are two basic methods used to draw conclusions about data quality; the most commonly used does not require statistical analyses. For example, the usefulness of reality checks stems from the reviewer's ability to examine or compare some values, and to draw a conclusion from them. The reviewer is using *logical inference* to draw that conclusion; that is, the facts are weighed against that reviewer's knowledge of the subject and a conclusion about the validity (this process may also be called engineering judgment or expert judgment). *Statistical inference* is an alternative method for drawing conclusions about data sets. For some types of decisions, statistical inference is more trusted and more reliable than logical inference. However, this is not necessarily true in all situations. Sometimes the answer is so obvious that statistics are not needed. Also, often too little is known about the data distribution to use statistical procedures reliably.

If statistical checks are only alternative ways of performing the other procedures described in this document, it is reasonable to ask why statistics should be used at all. The non-statistical approach to QA/QC requires the reviewers to base their decisions on their experiences with inventory preparation (which includes knowledge of past quality history) and on their understanding of the weaknesses in the data, methods, and process. This approach is sound in principle, but is also subject to several potential problems. The experts' memories may be inaccurate or short; furthermore, key experts may leave the agency, taking this "institutional knowledge" with them. Reliance on historical problems may lead to a failure to identify new problems in unexpected areas. Finally, experts are sometimes wrong; the uncritical acceptance of an expert's value can lead to persistent errors. While it is impossible to fully automate a QA/QC program and not advisable to rely entirely on statistical methods, the use of statistical descriptors, analytical procedures, and random sampling should be used to decrease reliance on key experts. Also, some institutional knowledge can be built into the statistical procedures.

The term *statistical checks* as used in this section describes a miscellaneous set of tools/procedures that are mostly useful for QA but can sometimes be applied to QC as well. However, many of these procedures fit the concept of "statistical quality control." This is the label attached to a formal system of process QC founded by W.A. Shewhart in the 1930s and popularized by W.E. Deming in the second half of the 20th century.

Although originally applied to manufacturing processes, the concept has been expanded to cover services as well. While formal statistical QC procedures have not been widely applied to emissions inventories, QA/QC procedures such as the range checks described under "Computerized Checks" (Section 5) are very similar to the acceptance sampling procedures used in statistical QC. The procedure could be extended to emissions inventories in the future, and some useful concepts from statistical quality control are included in this section.

Commonly used statistical methods for inventory QA/QC are:

1. Descriptive statistics (mean, standard deviation, frequency distributions, etc.) are often used for data presentation and to facilitate peer review (see Section 7.2).

2. Statistical procedures to detect outliers can be used for range check procedures (see Section 7.3).

3. Statistical tests (such as t-tests or similar analyses) can be used for comparability checks (to compare two or more inventories) and for data validation, such as comparing inventory data to ambient data, or evaluating the relationships between parameters used in an inventory (see Section 7.4). The latter is discussed in more detail in Section 9.

The first two types of methods are used fairly commonly; the third method has not been used very much but probably should be.

Because statistical analysis requires specialized expertise and also because essential QA/QC functions can be met using non-statistical methods, the EIIP has not designated any preferred or alternative statistical methods. The information in this section is provided to assist those who wish to incorporate statistical methods in their inventory QA/QC program. Also, if surveys and/or statistical methods are used to develop emission inventories, this section will be useful to peer reviewers or QA personnel who need to assess the validity of the statistical methods used. However, only an overview of important statistical terms and concepts is provided; for more detailed information, the reader should consult the references cited in this volume.

7.2 Descriptive Statistics

The most common use of statistics in emission inventories is to summarize and report both emissions and the data used in calculations. However, statistics are useful at several stages in the inventory process. They can be used to make informed decisions about the appropriate data or model to use, and to document the reasons for that

choice. They are used to screen large datasets or emissions calculations to identify potential errors. Statistical descriptors are also useful for assuring consistency within an inventory. For example, when comparing emissions from different counties, a greater than expected range of emissions for one county may be shown. These would be flagged for further checking of the data and methods.

Table 3.7-1 summarizes descriptive statistics that are useful for these purposes. QA reports (such as some of those shown in Section 5) can and do make use of summary statistics to facilitate peer review of the database. Statistics used in this way are primarily useful as a screening method much like a reality check. They should also be used to report and document the data (as discussed in Chapter 2, reporting and documentation play an important role in QA). While it may be difficult to spot an unusually large value in a vast data array, that number should stand out if the proper summary statistics are displayed.

One major limitation in advancing the use of statistical methods for QA/QC of inventories is that there has been little published analysis of the statistical characteristics of emission inventories. There is a general recognition that the frequency distributions of emissions are often not normal (or Gaussian). The lognormal distribution is held to be more representative of emissions data. However, the emission source, the type of pollutant, and the basis on which emissions are being presented (e.g., per Btu, per unit fuel, per capita, etc.) will determine the resulting distribution. If statistical methods (even the simple methods are described here) are to be used effectively, the expected distribution (i.e., the one that should occur if everything is correct) needs to be known. This will become the reference distribution that can be used as the basis for comparison.

TABLE 3.7-1

COMMONLY USED STATISTICAL DESCRIPTORS

Term	Formula	Explanation
Observation	x_i	The variable or attribute being measured.
Sample size	n	The number of values (or x_i's) in a set.
Arithmetic mean	$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$	The mean, or average, is the most commonly used estimate of central tendency. However, it is very sensitive to extreme values and is not a good measure if there is a very skewed distribution or very large outliers.
Range	$R = x_{max} - x_{min}$	The span of a set of numbers where x_{max} is largest number and x_{min} is smallest number in set. The range is most useful when emissions data are highly uncertain and an expert is using judgment to place "bounds" on the expected values (e.g., "fugitive dust emissions from unpaved roads are between a range of 0.25 and 3.0 times the value estimated using method x").
Variance of a sample	$s^2 = V = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$	A measure of the variability in a sample.

TABLE 3.7-1

CONTINUED

Term	Formula	Explanation
Standard deviation of a sample	$s = \sqrt{V}$	A measure of dispersion in the original units (e.g., tons per year), The standard deviation, along with the mean, are the principal measures used to define a value such as an emission rate and the associated uncertainty in the value. For normally distributed data, approximately 95 percent of values lie within ±2 standard deviations of the mean.
Estimated sum of population	$N\bar{x}$	The estimated sum of x_i's in the population of size N; this formula should be used to estimate area source emissions from a survey of a representative sample of sources.
Standard error of the estimated sum	$s_{N\bar{x}} = \frac{Ns}{\sqrt{n}} \left[\sqrt{1 - \frac{n}{N}} \right]$	The estimated error associated with the sample used to estimate the population sum. The value of s is the sample standard deviation as defined above. Note that as n approaches N, the error of the estimated sum approaches zero.
Median	The value x_i such that there is an equal number of values larger and smaller.	The median is the value from a sample with a cumulative frequency of 50 percent. It is relatively insensitive to extreme values so is a good measure of central tendency when there are extreme values such as is the case for most ambient monitoring data.

TABLE 3.7-1

CONTINUED

Term	Formula	Explanation
Percentile	---	The value on a scale of 100 that indicates the percent of a distribution that is equal to or below it.
Geometric mean	$\bar{y} = \exp\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right]$ where $y_i = \ln x_i$	Many environmental parameters such as air quality monitoring data tend to obey a lognormal distribution rather than a normal distribution. Consequently, the logarithmic transformation of $y_i = \ln x_i$ allows the data to be treated as if coming from a normal distribution.
Confidence interval		Values used to estimate a statistical parameter and that tend to include the true value of the parameter a predetermined proportion of the time if the process of finding the group of values is repeated a number of times. A confidence interval provides an estimate of the upper and lower bounds of a value at a given level of confidence (e.g., there is a 95 percent certainty that another sample would produce an estimate between the upper and lower bounds).
of sample mean	$\bar{x} - t_{\alpha/2,\, n-1}\, s/\sqrt{n}$ $\bar{x} + t_{\alpha/2,\, n-1}\, s/\sqrt{n}$	
of sample proportion (p)	$p - t_{\alpha/2,\, n-1}\sqrt{\frac{p(1-p)}{n}}$ $p + t_{\alpha/2,\, n-1}\sqrt{\frac{p(1-p)}{n}}$	

TABLE 3.7-1

CONTINUED

Term	Formula	Explanation
Absolute bias	$B_{abs} = \bar{x} - Obs$	The bias in an emissions estimate or parameter is the difference between the estimated value and the "true" value that is "known" through measurement or observation. The absolute bias is appropriate when the bias is relatively constant across all estimates of the true value (e.g., an incorrect offset voltage in a data logger produces a constant offset in the measured temperature).
Relative bias	$B_{rel} = \frac{\bar{x} - Obs}{\bar{x}} \times 100$ $B_{rel} = \frac{\bar{x} - Obs}{Obs} \times 100$	The relative bias is the absolute bias expressed as a percentage of either the "true" value or the estimate. It is more appropriate than the absolute bias when the magnitude of the bias is a function of the value itself (e.g., the bias in estimates of VOC emissions from motor vehicles is a function of estimated vehicle miles travelled).
Coefficient of variation	$CV = \frac{s}{\bar{x}} \times 100$	The coefficient of variation expresses the standard deviation as a percentage of the mean value. For lognormally distributed data such as most ambient monitoring data, the coefficient of variation is used analogously to the standard deviation.

TABLE 3.7-1

CONTINUED

Term	Formula	Explanation
Root of the mean squared error	$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(x_i - Obs)^2}{n}}$	The root mean squared error expresses the variability in the bias in a form analogously to the sample standard deviation. The RMSE can be used analogously (but not equivalently) to the standard deviation or coefficient of variation to express uncertainty.

where:

t = Student's t distribution;
α = significance level; and
Obs = observed emission value of the "true" value of the parameter.

Most of the statistical measures shown in Table 3.7-1 apply to data that obey a normal distribution. However, emissions data tend to not be normally distributed for a number of reasons including:

- Emission inventories can be dominated by a few very large sources (i.e., utility NO_x and sulfur oxides (SO_2) emissions);
- The underlying activity data used to estimate emissions are not typically normally distributed (e.g., vehicle mile traveled data, population density in an urban area); and
- The processes that lead to elevated emissions tend to be driven by extreme events that are not normally distributed (e.g., high load and acceleration events for motor vehicle, out of specification sulfur content for fuel oil, high wind speeds for wind blown dust).

Consequently, an inventory specialist should evaluate the emissions data for normality prior to any detailed statistical analysis of the emissions data or its uncertainty. The objective of this section is to discuss methodology available for assessing the distributional nature of emissions data.

Emissions from point sources are also often lognormal when expressed on a per facility basis. The distribution of businesses based on size (where size is measured by number of employees) is a well-documented example of a lognormally distributed variable (Lawrence, 1988). For those industries where emissions are correlated with number of employees, the distribution of emissions is at least partly explained by this fact. Furthermore, for area sources whose emissions are estimated using per employee factors, a lognormal (or other non-symmetrical) distribution of emissions is to be expected.

Some data distributions relevant to emissions inventories are shown in Table 3.7-2. While it is not necessary to understand the underlying mathematics, it is important to recognize that many common statistical procedures are based on the assumption of normality. However, real world data are often *not* normal. Sometimes it is useful to "straighten" or normalize the data so that a certain statistical method can be used. Table 3.7-3 gives several common data transformations and guidance on when to use them. The lognormal distribution is very commonly found in emission inventories as discussed above (an example is given in the following section); the lognormal transformation is a method that is often used on right-skewed data to create a more normally distributed variable. A transformation might be used if the intention is to use the data for further statistical analysis; such as regression modeling. However, transformation back to the original data should be undertaken with caution, and is not always appropriate. For example, the arithmetic mean of a log-transformed variable can be back-transformed (by exponentiating), but this new value is the geometric mean. For lognormally distributed data, the geometric mean is always less than the arithmetic mean (of the untransformed data). The geometric mean is often used to estimate the lognormal mean and the median, but it is a biased estimate of both (Gilbert, 1987).

The arcsine transform is used when the data are proportions. Examples of these data are also common in inventories. The relative contribution of sources to total emissions, DARS scores (see Chapter 4), or control efficiencies are all expressed as values between 0 and 1.0. If an inventory analyst wanted to know if there were differences in the relative contributions of area sources between two counties (expressed as a fraction of total emissions), for example, a t-test could be used after applying an arcsine transformation to the data. The mean and standard deviations of the raw (untransformed data) should be reported along with the results of the statistical test.

Finally, the square root transform is useful for data that follows a Poisson distribution. This type of distribution often occurs when the data are counts of independent items or events occurring in space or time. Typically, values of zero are common. This distribution is most likely to be encountered when collecting certain types of data that are used to develop factors or activity data. For example, to temporally allocate mobile source emissions, a field study to collect traffic counts might be implemented. The data

TABLE 3.7-2

PROBABILITY DISTRIBUTIONS RELEVANT TO EMISSIONS INVENTORIES

Distribution	Sample Mean	Sample Variance	Examples
Binomial	np, where n = number in sample or population, p = proportion	np(1-p)	Proportion of facilities in a SIC group belonging to emissions source category. Rule penetration (i.e., proportion of area source affected by a rule).
Lognormal	If the distribution is not highly skewed, simplest approach is to use equations shown for normal distribution[a]		See text.
Normal			Rarely found in environmental data but most standard statistical formulas assume normal distribution.
Uniform			Often used as default distribution when true distribution unknown.

[a]For a more detailed discussion of the methods used for estimating the mean, standard deviation and median of lognormally distributed data, see Gilbert (1987).

TABLE 3.7-3

DATA TRANSFORMATIONS USED TO "NORMALIZE" DATA

Transform name	Calculation of transformed data	Use when:
log[a]	$\chi' = \ln(\chi)$	Raw data approximate or are known to follow lognormal distribution.
arcsine (or angular)	$\theta = \arcsin \sqrt{p}$, where p is a proportion	Raw data are proportions between 0 and 1.0; if most data points fall between 0.3 and 0.7, this transformation generally not needed.
square root[a]	$\chi' = \sqrt{\chi}$	Raw data are counts of discrete events (distribution is Poisson, i.e., variance increases with the mean).

[a]If any of original data are zero, a very small number is generally added to all of the original χ values before transformation.

are collected in discrete intervals, depending on the location, time of day, and length of interval. Some of the values may be zero while others are quite high. These data are likely to fit a Poisson distribution.

For more discussion of the proper use of data transformations, see Sokal and Rohlf (1969). An example of the use of normality tests and data transformation is given next.

7.2.1 TESTING FOR NORMALITY

The first step is to generate basic statistical summaries of the data. These summaries include the minimum, maximum, mean, median, quartiles, and standard deviation of the data. For normally distributed data, the mean and median should be equal and the minimum and maximum values should be equally spaced below and above the mean. Most likely, this simple test will indicate significant departure from normality. For example, emissions data tend to have a few sources with very high emission rates that produce a highly skewed distribution with a long tail, resulting in the mean being significantly higher than the median.

TABLE 3.7-4

DESCRIPTIVE STATISTICS FOR LOS ANGELES POINT SOURCE INVENTORYa

Pollutant	Number of Sources	Mean (TPD)	Median (TPD)	Upper Quartile (TPD)	Min (TPD)	Max (TPD)	Standard Deviation (TPD)
TOG	24,432	0.012	0.0025	0.0072	~0.0	3.9	0.067
NO_x	6,980	1.2	0.00088	0.0052	~0.0	380	10
CO	6,004	0.025	0.0011	0.0049	~0.0	10	0.21
SO_2	1,033	3.0	0.00095	0.0090	~0.0	370	22
PM_{10}	8,225	0.0042	0.00052	0.0018	0.0	1.1	0.023

[a]All emissions shown as tons per day (tpd).

Table 3.7-4 presents a summary of the basic descriptive statistics for the Los Angeles data set used as an example in the remainder of this section. This data set represents estimated emissions for South Coast Air Basin for the year 2007 for point sources. For each of the pollutants, the descriptive statistics imply a highly skewed distribution. The mean is a factor of 10 greater than the median for all pollutants except for PM_{10}. In fact, the mean is greater than the upper quartile (75th percentile) for all five pollutants.

The second step should be preparation of a histogram of the emissions data. After the initial review of the data presented in Table 3.7-4, it is reasonable to assume that the data must be transformed if some semblance of normality is to be obtained. The most common data transformation is a logarithmic transformation, whereby the emissions estimates are replaced by the natural logarithm (logarithm to the base e) of the emissions data. When this transformations complete, the resultant data many times can be represented by a log-normal distribution.

Figures 3.7-1 and 3.7-2 present two histogram plots of total organic gases (TOG) emissions from the South Coast Air Basin. Figure 3.7-1 gives a histogram of the untransformed emissions data while Figure 3.7-2 presents the data after application of the log transformation. Based upon visual inspection of Figure 3.7-2, the transformation is highly successful in producing a log-normal distribution. The histogram is closely matched by the fitted normal distribution.

Another graphical method for examining the fit of a normal distribution is the expected normal plot. Figure 3.7-3 presents expected normal plots for TOG and NO_x from point

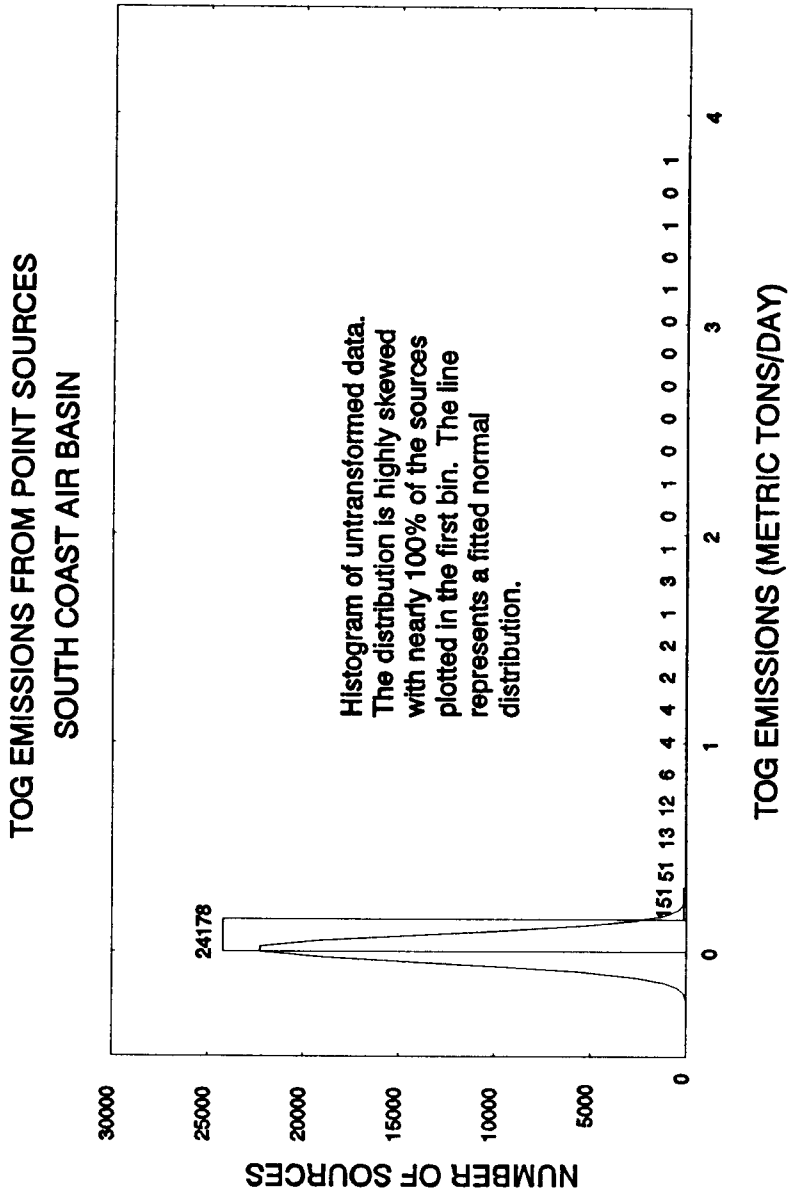

FIGURE 3.7-1. HISTOGRAM PLOT OF TOG EMISSIONS FOR LOS ANGELES, NO DATA TRANSFORMATION

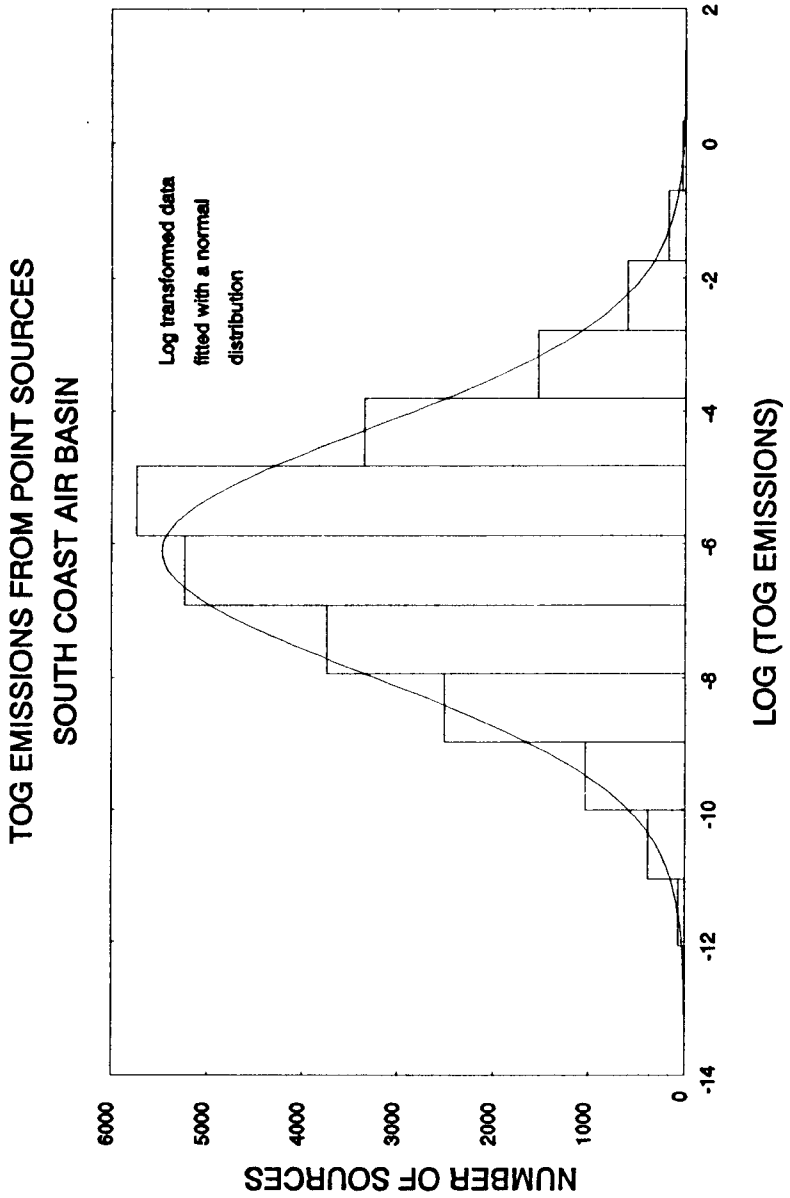

FIGURE 3.7-2. HISTOGRAM OF TOG EMISSIONS FROM LOS ANGELES POINT SOURCES, LOG TRANSFORMATION

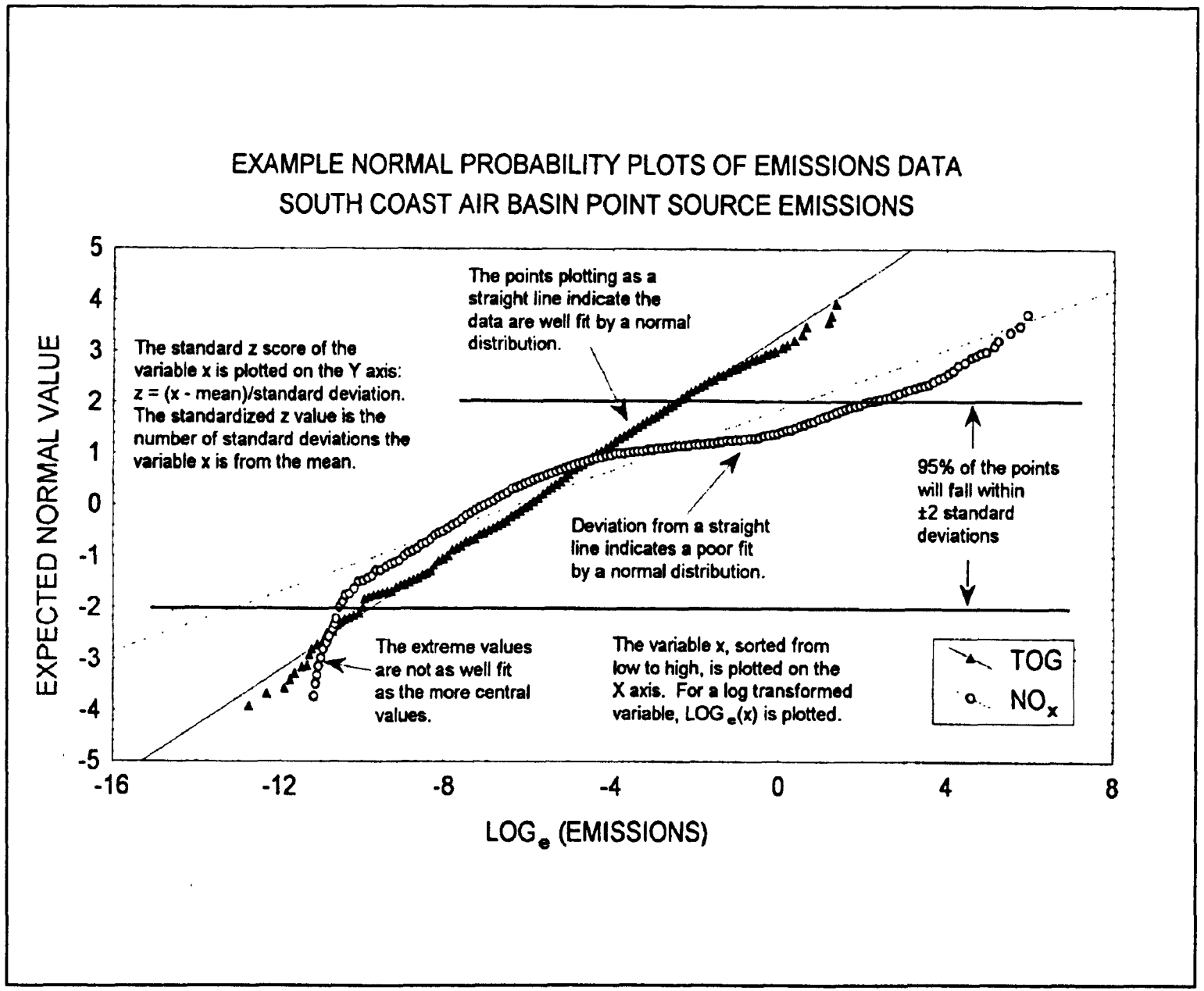

FIGURE 3.7-3. EXAMPLE EXPECTED NORMAL PLOT FOR LOS ANGELES POINT SOURCE EMISSIONS

sources in Los Angeles. In this plot, the observed emission values (log transformed values for this data set) are plotted on the X axis. On the Y axis, the standardized value of the emissions in plotted. The z value standardizes the distribution to have a mean value of zero and a standard deviation of 1.0. The z value is standard statistical transformation and is computed as follows:

$$z = \frac{(x - \bar{x})}{\sigma_s}$$

where:

- z = standard score;
- x = emission value;
- $\bar{x}$ = sample mean; and
- σ_s = sample standard deviation.

For the purposes of the expected normal plot, the distribution being tested is assumed to follow a normal distribution and thus $\bar{x}$ and σ_s estimated from the data set are assumed to be unbiased estimates. For a non-normal distribution, this assumption is violated but the plot itself will identify the lack of fit of the normal distribution.

When the standard score is plotted against the respective value on a linear plot, data from a normal distribution will plot as a straight line on an expected normal plot. As can be seen from Figure 3.7-3, the TOG data for Los Angeles plots as a relatively straight line over most of its range. By comparison, there is a significant deviation from a straight line (and hence log-normality) for the NO_x emissions.

Included on Figure 3.7-3 are elements of the plot that provide insight as to the normality or non-normality of the plotted data. First, if the data are normal, approximately 95 percent will fall within ±2 standard deviations of the mean (between -2 and +2 on the Y axis). From inspection, the TOG data follows a straight line between ±3 standard deviations (approximately 99.7 percent of the data). Both the NO_x and TOG distribution deviate from the straight line at both extremes. This is a common occurrence because the hardest portions of a distribution to fit are the two extremes. For the purposes of emissions uncertainty, the most important factor is the semblance of normality for the bulk of the data. Thus, deviation from normality at the very extremes of the distribution are generally unavoidable but expected.

It is obvious from an expected normal plot when the data clearly do not approximate a normal distribution. Figure 3.7-4 presents an expected normal plot of untransformed Los Angeles point source NO_x emissions. The observed "angle" in the data is very characteristic of a data set that is highly skewed and in need of a logarithmic transformation.

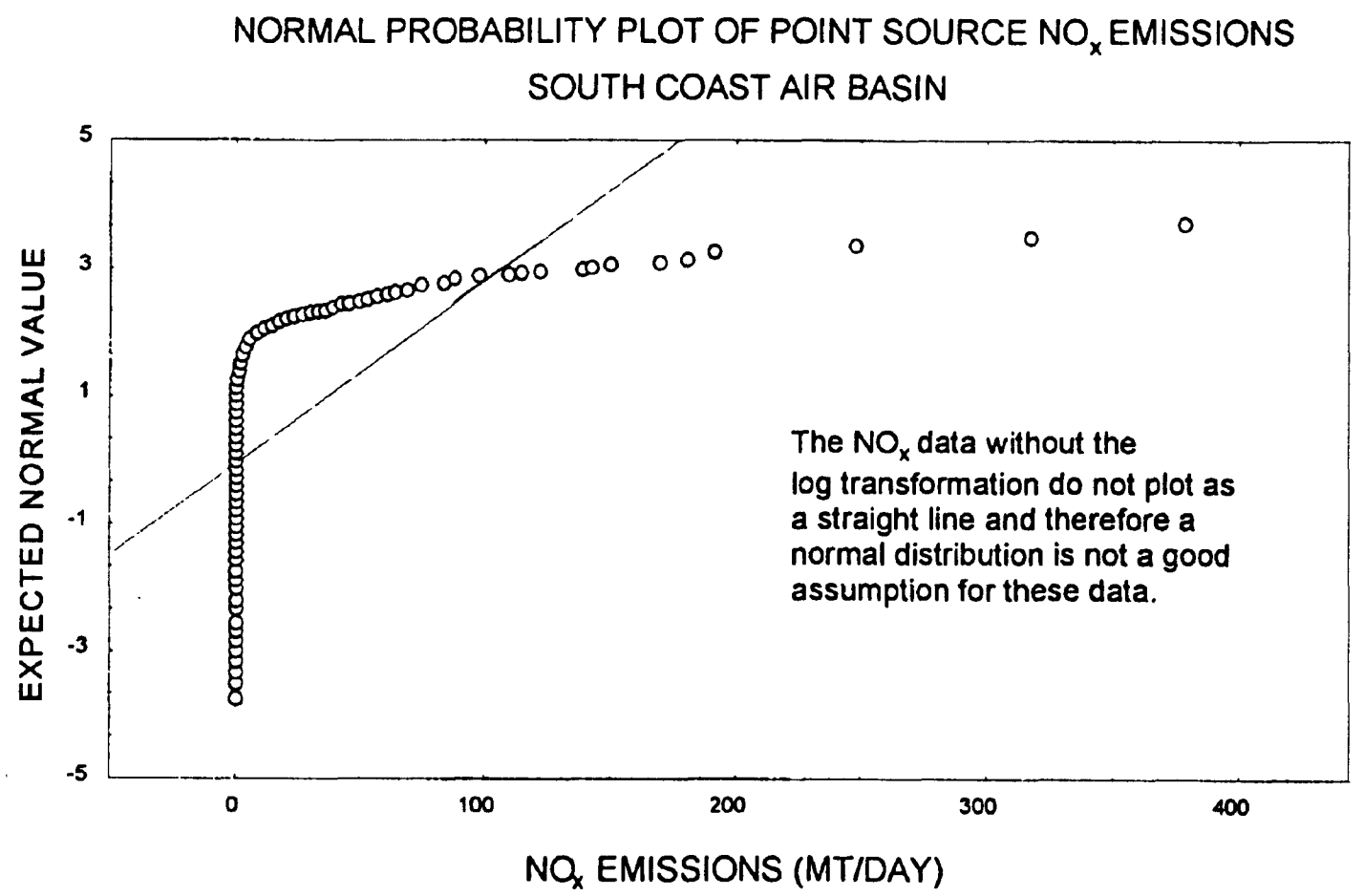

FIGURE 3.7-4. EXPECTED NORMAL PLOT OF LOS ANGELES POINT SOURCE NO_x EMISSIONS WITHOUT A LOG TRANSFORMATION

One factor that leads to non-normality in emissions data is illustrated in Figure 3.7-5. This figure presents a histogram of log-transformed NO_x emissions in Los Angeles that demonstrates a bimodal distribution of emissions. Bimodal distributions are common where there are significant differences in the types of sources producing emissions. For NO_x emissions, small industrial boilers and large utility boilers can potentially contribute to the bimodal nature of NO_x.

Most statistical packages include tests for determining the normality of data. However, care should be taken when judging normality of emissions data on the basis of standard statistical tests such as the Shapiro-Wilk normality test (Shapiro and Wilk, 1965). Emission data sets such as the Los Angeles set used here are very large (over 24,000 sources of TOG in the point source inventory). Because of the large sample size, very small deviations from normality in the distribution are judged to be statistically significant. The above transformed TOG data set, which appears to have a very good visual log-normal fit, fails the Shapiro-Wilk normality test because of the very large sample size. However, for most emissions analysis purposes, the strict deviation of the transformed data set from normality is not important.

7.3 STATISTICAL QUALITY CONTROL

Statistical quality control is a proactive set of procedures in which the assignable causes of variation have been removed from the process (or at least that is the intention) and only random causes are present. The expected error associated with those random causes of variation is characterized by a probability distribution. A tool is then developed based on this probability distribution that can be used to monitor the occurrence of an assignable (and presumably correctable) cause (Woodall and Adams, 1990). This tool may focus on *attributes* or on *variables*. Data that arise from the classification of items into categories are attributes. They are generally measured as counts (or proportions) of the number of items in a given category. For example, in an inventory the QC tool could focus on data entry errors that occur because of incorrect keying of data; the attribute would be a simple "correct/incorrect." When the data are actual measurements, estimates, or calculations, the quality tool is based on variables. If variables are used, the process might be directed towards the value of input data such as stack heights, emission rates, or the emissions themselves.

The type of tool used for statistical quality control is generally some type of control chart or acceptance sampling. The control chart is probably more suitable for continuous processes (although data entry errors for large databases could be monitored using control charts). The concept of acceptance sampling, however, is very similar to the range check concept that is commonly employed for database QC. The difference is that statistical quality control goes well beyond what is normally done for inventory

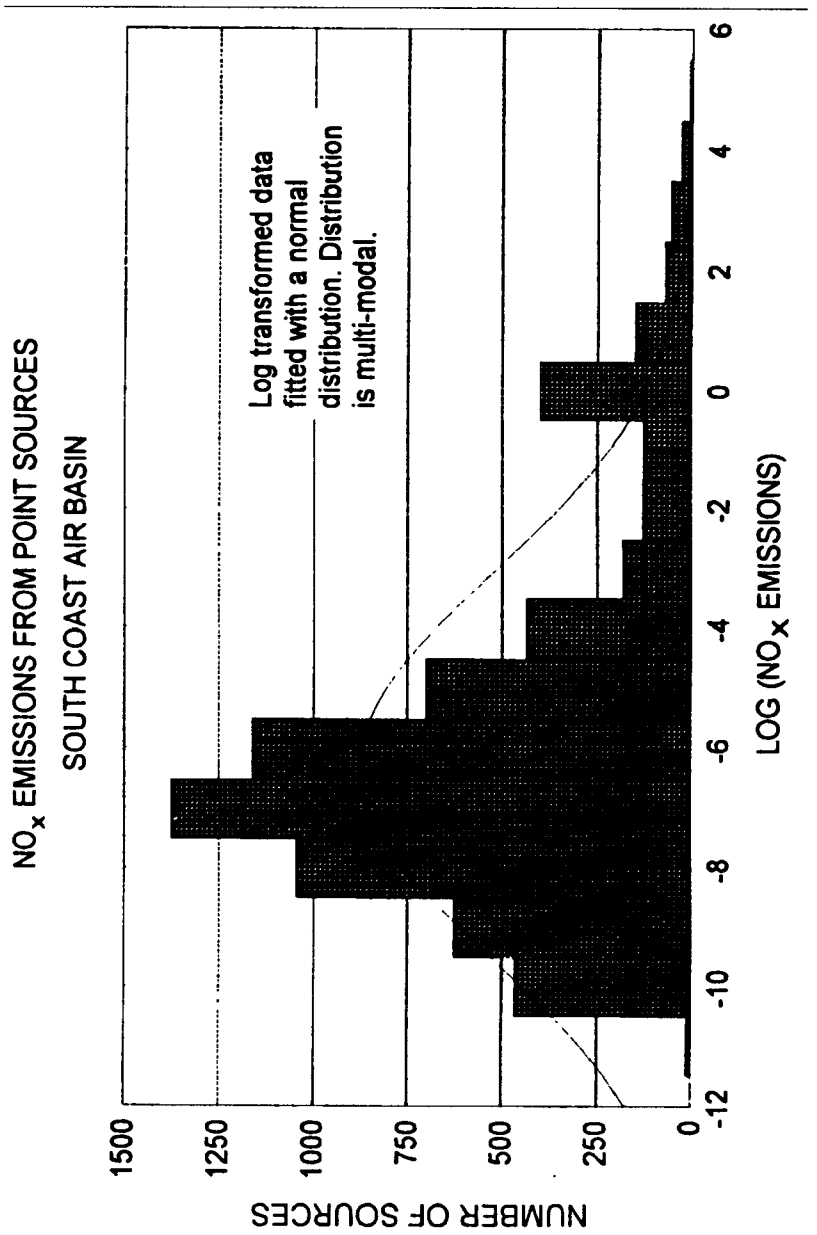

FIGURE 3.7-5. HISTOGRAM OF LOS ANGELES LOG-TRANSFORMED NO_x EMISSIONS DATA ILLUSTRATING A BIMODAL DISTRIBUTION

range checks in that a target quality level is specified, a statistically valid sampling plan for the data set is specified in advance, and procedures for dealing with errors are established ahead of time.

7.3.1 An Example of Acceptance Sampling by Attributes

One of the most common QC activities involves verification that data keyed into a computer database are correct. A state inventory can be very large so that 100 percent inspection (i.e., checking the data to verify correctness) is impractical. The solution is to sample the database, and only check a specified percentage of the data; however, determining the sample design and sample size is not always straightforward. Generally, a stratified random sample is recommended.

The first step in designing a statistical random sample is to divide the data into subcategories. A point source inventory, for example, may be divided into general data (company name, address, owner's name, etc.), location data (UTM coordinates, county, etc.), stack data, process data, or other categories. The importance of these categories will vary. Errors in the general data may be of minor consequence, process data may be critical. The agency should prioritize the data categories as to importance. Sample sizes will depend partly on the importance of the category.

The next step is to choose the acceptable level of error. Several statistical QC symbols and their meaning are defined below (derived from Grant and Leavenworth, 1988):

N	=	number of items in a given lot or batch.
n	=	number of items in a sample.
D	=	number of incorrect items (i.e., data elements not conforming to specifications) in a given database of size N.
r	=	number of incorrect items (i.e., data elements not conforming to specifications) in a given sample of size n.
c	=	acceptance number, the maximum allowable number of incorrect items in a sample of size n.
p	=	fraction incorrect. In a given database, this is D/N; in a given sample, it is r/n.

In general, the inventory preparers will want to minimize the number of errors, so c=0. That is, if any incorrect data elements are found in the batch that is inspected, the data set is unacceptable. In practice, this means that either: (1) the entire data set is inspected and incorrect values are corrected, or (2) the errors in the sample are corrected and the sampling/corrected process is repeated until no errors are found in a sample. The second approach may still allow some errors to remain, but it is much more cost-effective for very large data sets.

Figure 3.7-6 shows the probability that no errors will be found (i.e., P_a or the probability of acceptance) for a given sample size (n) from a database of size N = 5,000. The actual number of incorrect data (D) in the database is shown on the X axis. The graph shows that if 100 points are sampled (n=100) from this data set, there is a 5 percent (P_a = 0.05) chance of finding no errors when the true error rate is about 160/5,000 or 3.2 percent (p = 0.032). A larger sample size decreases the probability of accepting a defective data set. For example, if n = 500, the probability of accepting a dataset with p = 0.032 is near zero. Note that this example assumes only one sample will be taken. Better quality can be attained if the database is repeatedly sampled, inspected, and corrected until the desired level of quality is attained (discussed further in the next section).

7.3.2 Specifying a Sample Plan

Inventory QA plans often specify a percentage of data elements to be sampled, incorrectly assuming that the ratio of sample size to data set size is important. Figure 3.7-7 shows the fallacy of this idea. The curves show that from a quality control perspective the absolute size of the sample is more important than its size relative to that of the data set. Figure 3.7-7 shows four curves for a sampling plan that specifies 10 percent of the inventory be checked. Note that for a given percent incorrect, the smaller N is, the less protection is afforded by using a 10 percent sampling scheme. For example, with only 2 percent incorrect, the probability of accepting the data set of N=1,000 using the 10 percent sampling scheme is around 0.12. The risk of accepting a "bad" dataset is even higher for smaller data sets. This suggests that setting a threshold sample size may be a better strategy. For example, the sample design could specify inspection of data elements for all sources above a certain size, and then sampling a percentage of the remainder. Or, the sample size could be preset at 500 where $N \geq 500$, and 100 percent sampling if $N < 500$.

The sampling plan design should also specify the criteria for determining an adequate number of samples. This can be done *a priori* by specifying the exact number of samples. Or, an iterative approach can be used which requires meeting a target level of quality (or maximum error rate that will be acceptable). For example, the error rate found in one sample may be used to estimate the overall error rate (the average outgoing quality or AOQ) as follows:

$$AOQ = 100p * P_a$$

where P_a is the probability of acceptance as shown in Figures 3.7-6 and 3.7-7, and p = r/n. For example, if p = 0.04 and P_a = 0.012, the AOQ = 0.048. This says that the estimated error rate in the database is 4.8 percent (or, 4.8 percent of the data

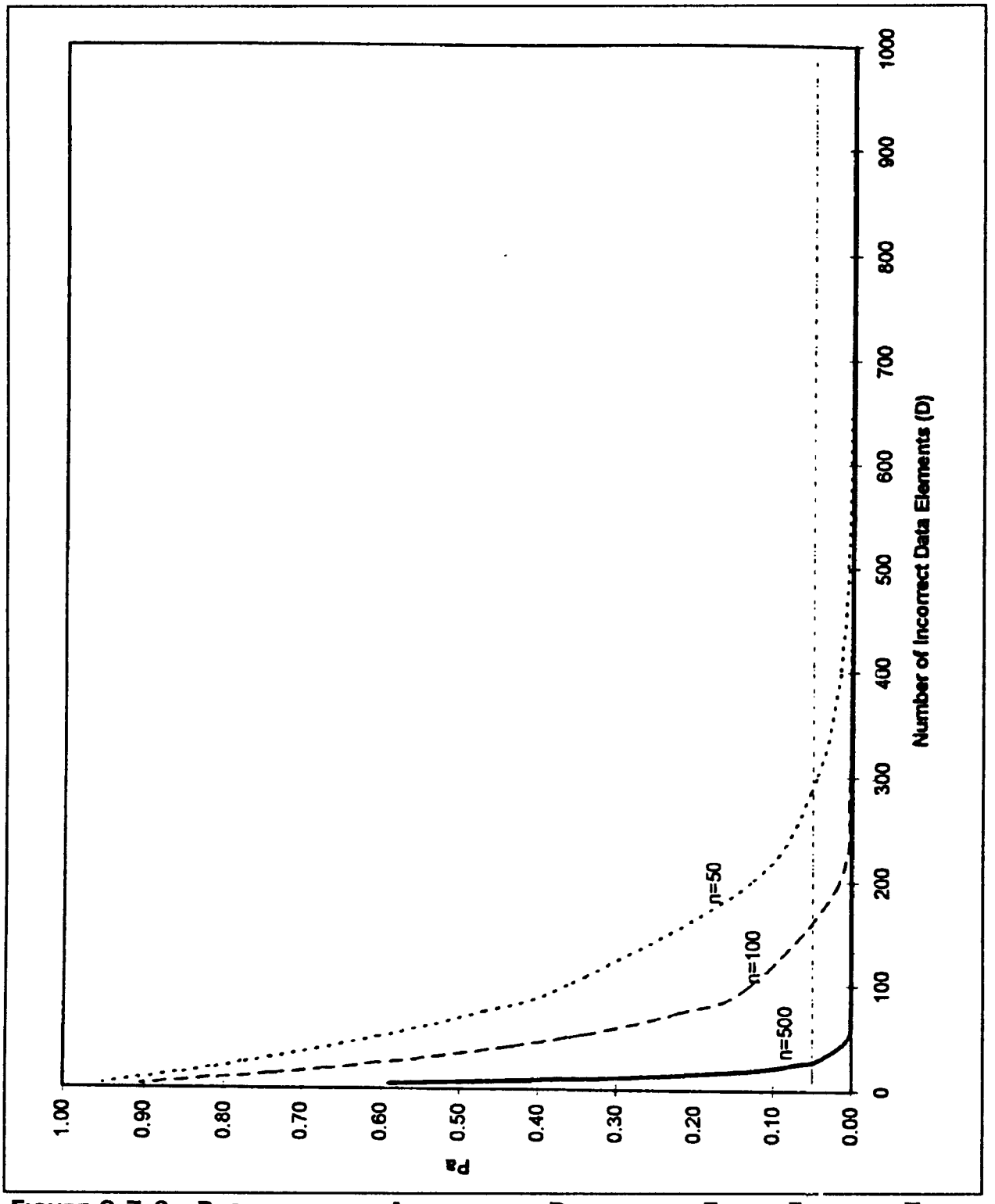

FIGURE 3.7-6. PROBABILITY OF ACCEPTING A DATASET AS ERROR FREE FOR THREE DIFFERENT SAMPLE SIZES

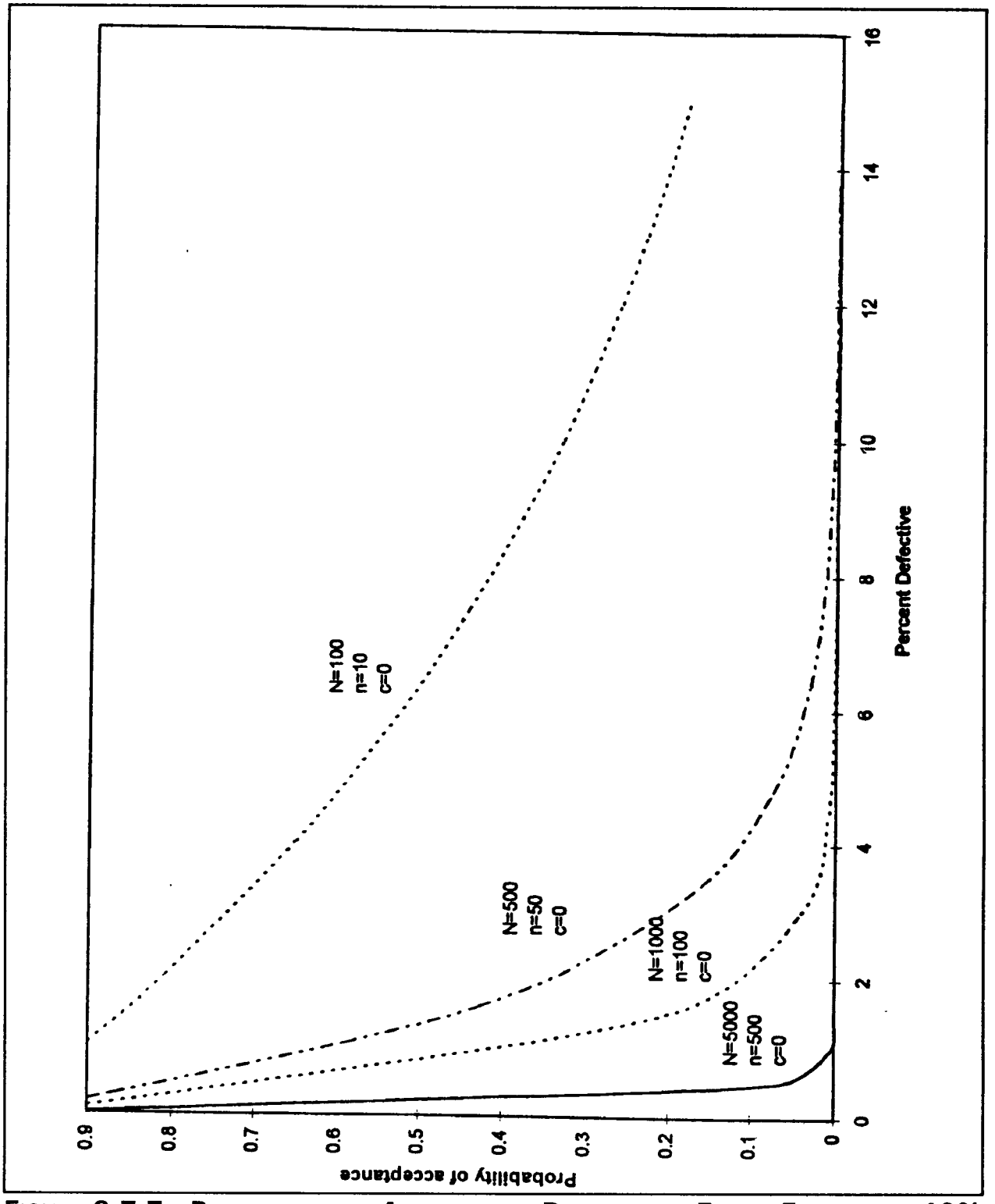

FIGURE 3.7-7. PROBABILITY OF ACCEPTING A DATASET AS ERROR FREE WHEN 10% SAMPLE SIZE IS USED

elements are likely to be incorrect). For a more detailed discussion of AOQ and the theory behind it, see Grant and Leavenworth, 1988.

If the AOQ is too high, another sample may be chosen. Again, any errors found are corrected, and p, P_a, and AOQ are calculated. This process is repeated until the AOQ reaches an acceptable level. Note that this assumes sampling with replacement; that is, each sample is randomly chosen from the entire data set and may include items previously sampled. An alternative approach is to not include previously sampled data elements in subsequent samples (i.e., sampling without replacement). If this approach is used, AOQ is calculated by adjusting for the previously checked and corrected data as shown:

$$AOQ = (100\text{-}f)p * P_a$$

where f = percent inspected and corrected.

The QA program plan should specify the desired AOQs and sample size (n) for each subset of the data. The procedures for sampling should be described in as much detail as possible. Calculation of P_a and AOQ values can be easily performed using the SAS QC® function PROBHYPR; other statistical QC packages are also available that can calculate these terms.

7.3.3 Outlier Analysis: Acceptance Sampling by Variables

Statistically based outlier analysis can be used to check the ranges of specific data and flag values that appear to be in error. This use of statistics to flag outliers is similar to the concept of acceptance sampling by variables used in statistical quality control programs. A common approach is to use the mean and two (2sd) or three (3sd) standard deviations as the basis for the acceptable range. Although three standard deviations is generally used to screen data sets and reject extreme values, for the purposes of inventory QC, two standard deviations may be more appropriate. A large proportion of the data flagged using the 2sd test is likely to be correct, but use of the 3sd test will probably not detect any errors except the very extreme ones.

It is important that the expected distribution of the variable being checked is considered when setting the acceptance criteria (see discussion above under "Descriptive Statistics"). If the underlying distribution is highly skewed to the right (e.g., lognormal), using the arithmetic mean and standard deviation may not be appropriate. In fact, a random sample from a lognormal distribution will tend to result in a large standard deviation, often exceeding the mean. Adding 2sd or even 3sd to the mean may still fall short of the possible range of values at the high end; subtracting 2sd may result in a negative

number. Thus, the range check will not detect any values too small (assuming the real data set is always positive), and will flag too many "errors" at the high end.

Four different methods can be used to deal with this problem. The first is simply not to rely on statistical procedures for range checks. Engineering judgment can be used to set an appropriate upper and lower bound for each variable. This method will work reasonably well assuming that all the variables in the data set are sufficiently understood, and that the "true" range and distribution of the data do not change appreciably between years or geographic regions.

The second approach is to calculate the mean or other statistics using the method that is appropriate for the specific distribution. The greatest advantage of this approach is that it is more likely that errors will be detected, and it reduces the likelihood that correct values will be flagged. The problem with this method is that it requires more information and work to set up the appropriate criteria; it also assumes that the correct underlying distribution can be identified which is often not the case.

The third approach is to use the arithmetic mean and standard deviation of the transformed (and normalized) data. For example, the natural log of the data may more closely follow a normal distribution. If this method is used, the transformed data should be tested using a test for normality to determine if the transformation has had the desired effect. Data must be transformed back to their original form after the analysis is complete (Table 3.7-3 gave examples of transformations).

Assuming that the data are reasonably normal, the outlier test may be made with or without reference to a pre-determined "correct" value. An example of the latter was given in the first paragraph of this section. If the data are normally distributed, 68 percent of the data will fall within 1sd of the mean; 95 percent, 2sd; and 99 percent, 3sd. If more than 1 percent of the data are greater than 3sd away from the mean, these data are suspect.

Of course, the data may be internally consistent (and normally distributed) but still incorrect. Comparison to a reference value should be used as a check. The reference value may be based on a different data set; for example, the distribution of stack heights may be taken from a national inventory and used as the reference distribution for a county- or state-level dataset. Or, the reference value may be based on previous inventories in the region, or based on engineering judgment. Typically, the mean value of the variable is compared to an expected mean using a t-test.

The t-statistic is calculated as:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

where:

t	=	the calculated test statistic;
$\bar{x}$	=	the mean of the test data set;
μ	=	the preestablished mean (or reference value);
s	=	the test data set standard deviation; and
n	=	the number of data values used to calculate the test data set mean and standard deviation.

Assuming that the data are normally distributed, the distribution of the test statistic follows the *Student's t distribution*. The t value can be compared to a statistical table of values generated for the Student's t distribution for various sample sizes and levels of significance. The various sample sizes within a Student's t distribution table are referred to as the *degrees of freedom*. The number of degrees of freedom is equal to the sample size (in this case, the number of values in the data set) minus one. If the absolute value of the test statistic is greater than the value in the Student's t distribution table for a specified level of significance and degrees of freedom, the conclusion would be that the calculated average value is significantly different than the number entered into the average value field. If this is the case, the data must be examined in greater detail to determine if the data are in error, or if the preestablished value entered into the average value field should be changed.

The fourth approach is to use nonparametric methods to evaluate the data. Nonparametric methods make no assumptions regarding the distribution of the data; the sample median, rather than the mean, is used as the basis for testing. A *sign test* can be used to test whether the median of the data is significantly different than a preestablished median. The preestablished median is first subtracted from the values of every data point in the sample, and the total number of positive values (a) and negative values (b) are recorded. Any zero differences are omitted from the test. The test statistic is:

$$\chi^2 = \frac{(|a-b| - 1)^2}{n}$$

where:

χ^2	=	the test statistic;
a	=	the total number of positive values;
b	=	the total number of negative values; and
n	=	a + b.

The test statistic follows a *Chi-square distribution* with one degree of freedom. The value of the test statistic can be compared to a table of values for the Chi-square distribution at various significance levels. If the test statistic is larger than the tabular value, the conclusion is that the sample median is significantly different from the preestablished median.

If the total number of positive and negative differences is larger than 25, a normal approximation to the binomial distribution must be used:

$$Z = \frac{a - 0.5n}{0.5\sqrt{n}}$$

This test statistic (Z) can be compared with the value in a normal distribution table for a selected level of significance.

7.4 Statistical Procedures for Comparability Analysis

Summary statistics are most useful for comparing inventories. Although statistical measures are often used as the basis for comparisons, any decisions made or conclusions drawn from these statistics are usually subjective. That is, the reviewer will compare means, plot emissions versus a variable thought to be correlated with emissions, or evaluate correlations between variables. The reviewer then decides (using his or her own unwritten criteria) that the data are "different" or "reasonably similar." No published references were found that indicate hypothesis testing has been used to determine if two data sets (or inventories) are significantly different, although this method may be appropriate for certain types of analyses.

The types of comparisons most commonly used include:

- Comparisons between inventories;
  - -- between years (same region)
  - -- between regions (same year)
- Comparisons to a reference distribution;

  - between a region and a larger geographic area (such as state to national)

- Comparisons of emissions to a surrogate with an expected predicted relationship (e.g., plot of on-road emissions vs. VMT).

All of the above are useful for the evaluation of inventories; however, their uses strictly for QA/QC need to be considered very carefully. As with other forms of reality checks, differences between two statistical measures do not necessarily mean an error has occurred; similarly, errors will not always be obvious when using this method.

In most cases, an informal comparison of the statistics is sufficient. Often the differences are so obvious as to need to no further statistical analysis; emissions from one inventory may be consistently higher or lower than the other (i.e., biased) and the reasons are easily traced to differences in methods or assumptions used. Biases in inventory data may occur if different methods, factors, activity data, or assumptions are used; within a state (or region), consistent approaches will minimize the occurrence of this type of bias (see Chapter 2, *Planning and Documentation*). However, if inventories prepared by different people and/or for different geographic entities are compared, consistency in approaches cannot be assumed (and is not very likely).

Differences between inventories cannot be assumed to be the result of errors; for example, suppose that evaporative emissions from gasoline marketing are consistently higher (when expressed on a per gallon basis) in two inventories. The QA reviewer should first determine if there is a reasonable explanation for the apparent bias. One state may have implemented Stage I and/or Stage II controls resulting in lower emissions. It is best when conducting a statistical comparison for QA purposes to first remove all the assignable causes of variation. For example, comparisons of uncontrolled emissions might be a better way of detecting possible errors for the gasoline marketing example. Any differences in emissions are more likely to be due to inconsistences or errors in the underlying calculations.

When visual inspections do not reveal any obvious differences, it is generally harder to conclude that the data are correct. Most inventory analysts feel that the uncertainty of emission estimates is too high to allow conclusions to be drawn when small differences between emissions are seen. In this case, the use of statistical tests to detect differences can be of use. The t-test or sign test described in the previous section can be used to determine if there are statistical differences between inventory data if two data sets are being compared. If more than two data sets are being tested, an analysis of variance (ANOVA) can be used for normally-distributed data, or the nonparametric Kruskal-Wallis test should be used if assumptions of normality can be met.

7.5 EXAMPLE OF A STATISTICALLY BASED COMPARABILITY ANALYSIS: REGRESSION MODEL USING DUMMY VARIABLES

When preparing an inventory, it is sometimes necessary to evaluate data prepared by another expert or agency. A common example is found in the development of onroad mobile source emissions; VMT estimates are often done by experts outside the agency preparing the inventory. In this example from North Carolina, the state's department of transportation (DOT) supplied a database of VMT by county over an eight-year period. The inventory analyst at the NC DEHNR had several objectives for this analysis: (1) assess the reasonableness of the estimates, (2) check for any errors in the data, and (3) develop a model to be used for projections of VMT in the future. The method used was to develop a regression model using dummy variables. Dummy variables, sometimes called zero-one variables, have a value of one when the characteristic being described is present in an observation or is set equal to zero otherwise.

Seven North Carolina counties, within three ozone maintenance areas, were chosen for this analysis to illustrate one method that may be used to separate the distinctive intercept and slopes associated with reported VMT. A dummy variable was constructed for each county which has the effect of shifting the intercept term. In addition, groups of counties in the TRIANGLE (Durham and Wake), or the TRIAD (Davidson, Forsyth and Guilford) were coded to distinguish those observations from the remaining base group around CHARLOTTE (Gaston and Mecklenburg). These TRIANGLE and TRIAD dummy variables were then multiplied by the trend variable TIME to create a slope shifting effect between these areas.

The basic method for projecting VMT involves linear trend extrapolation using each separate county VMT over the eight year period. Visually checking this combined data for possible errors would be difficult, but there may be advantages to looking at all the counties together or "pooled," for the purpose of comparing growth rates between the three specific areas of the state. If the model, as constructed, can account for a large percentage in the variability, then the residual or difference between the reported VMT value and the predicted value from the model can be used to indicate possible errors.

Figure 3.7-8 shows a graph of all reported VMT values versus the predicted values from this model. Visual inspection of the plot revealed a possible keying error from Durham County in 1988 (open circle on the figure). The originator of the data set was contacted, and confirmed that this value was incorrect. The value was corrected and the regression was recalculated. This graph can also be used to perform other types of QA checks. For example, the relative ranks and slopes of the county-level data can be compared as a type of reality check (see Section 1). The reviewer might look at the population of each county, the miles of different road types, and other factors that presumably affect VMT to determine if they are consistent with the pattern seen in the data.

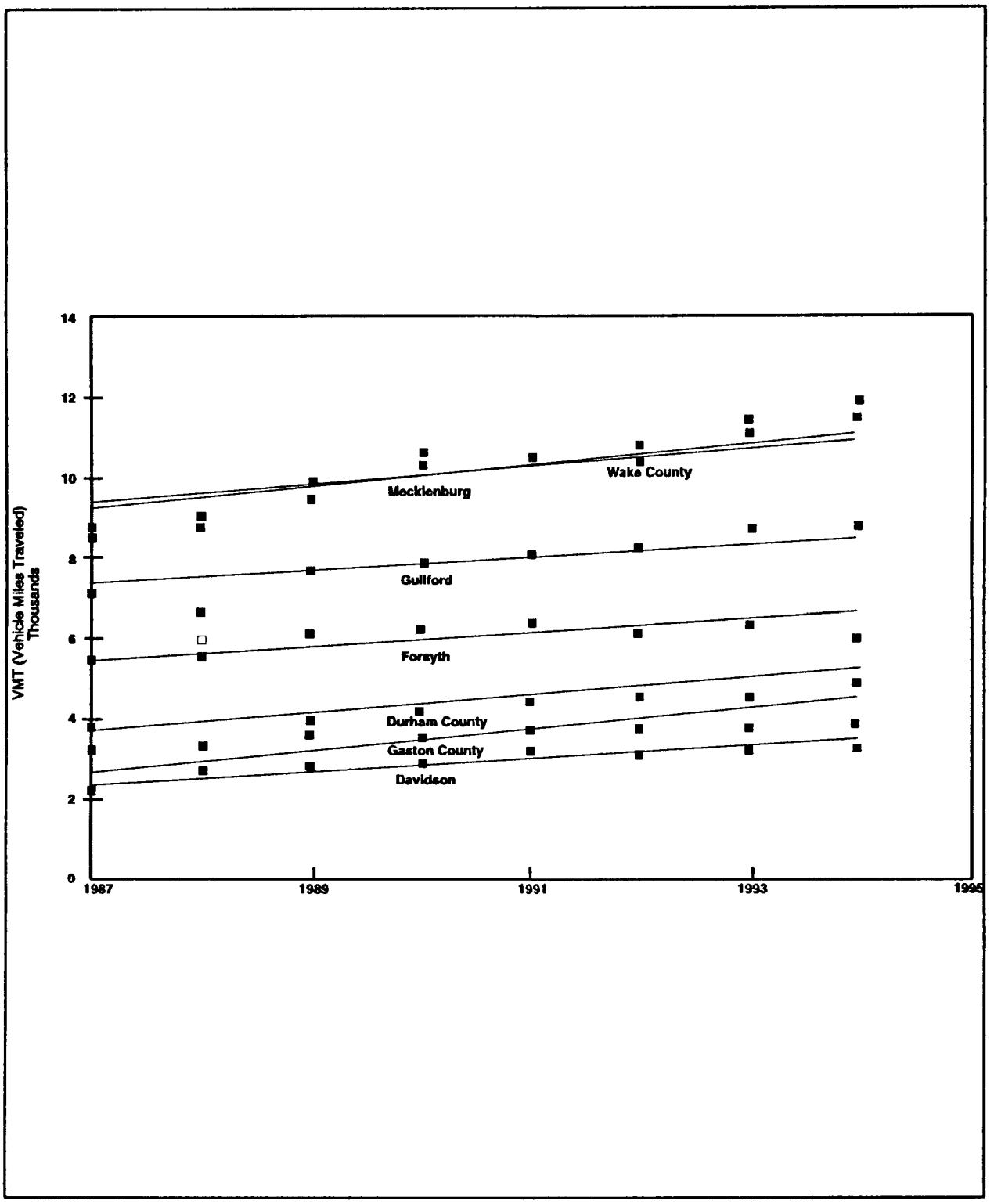

FIGURE 3.7-8. GRAPHICAL COMPARISON OF REPORTED AND PREDICTED VMT VALUES

TABLE 3.7-5

RESULTS OF DUMMY VARIABLE REGRESSION ANALYSIS OF COUNTY VMT DATA

X Variable Name	X Coefficient(s)	Standard Error of Coefficient	"t" Statistic
Time	273.3	52.54	5.20
Davidson	2388.6	277.00	10.52
Durham	3839.8	250.60	15.32
Forsyth	5710.5	227.00	25.16
Gaston	2744.9	250.60	10.95
Guilford	7683.4	227.00	33.85
Mecklenburg	9677.4	250.60	38.62
Wake	9772.8	250.60	39.00
Time*Triad	-100.1	67.83	-1.48
Time*Triangle	-38.2	74.30	-0.51

The regression results can be interpreted around the base case of what is occurring in the Charlotte area. The t-statistic (shown in Table 3.7-5) can be used to determine the significance of including each trend or dummy variable in the model. A general rule for evaluating the t-test is that an absolute value greater than two is significant at the 0.95 confidence level. The model suppresses the normal calculation of an intercept term by assigning each intercept a shifting county dummy variable. These were all highly significant, which indicates that the initial VMT levels between counties are different. The CHARLOTTE area growth trend variable coefficient (time) of 273.3 equals the VMT increases each year and was quite significant for the base case. The slope shift trend for the TRIAD was -100.1, which means that for those three counties the annual growth in VMT was only 173.2, with a t-test value of -1.48 indicating a mildly significant shift. In comparison, the slope shift value of -38.2 for the TRIANGLE counties had a low t-statistic value of only -0.51. This result may be interpreted to mean that while the annual growth of 235.1 in VMT for the TRIANGLE is smaller than the Charlotte area, it should not be considered as significantly different.

This type of analysis may be used by the inventory developer and cited as the basis for any assumptions subsequently used to develop the inventory. For example, the inventory developer may argue that one growth factor can be applied to all counties except those in the TRIAD, and cite this analysis as the basis for that assertion. Alternatively, the peer reviewer (or technical auditor) may perform this type of analysis to test the validity of an assumption used to develop an inventory. If, for example, the project inventory had been prepared using one growth factor for all counties, the peer reviewer might challenge that based on these results.

This page is intentionally left blank.

8

INDEPENDENT AUDITS

QUICK REVIEW

Definition:	A systematic evaluation to determine the quality of a function or activity.
Objective:	Ensure that adequate quality control measures are planned and implemented to help develop a complete and accurate inventory.
Optimal usage:	At critical steps in the inventory development process that could affect the accuracy and completeness of the emissions data.
Minimum expertise required:	Moderate to high technical and auditing experience. Auditors should have audit experience and be able to comprehend the technical process and assess whether approved procedures are followed. An audit team of technical and quality professionals may be required for more complicated inventories.
Advantages:	(1) Help provide confidence in the accuracy and completeness of the emissions results, (2) provide an objective assessment of the effectiveness of the quality control (QC) measures implemented by the emissions inventory development team, and (3) foster continuous improvement in the inventory development process.
Limitations:	(1) Additional staff and funds may be needed to conduct independent audits, and (2) may add to the time required to complete the inventory.

8.1 OVERVIEW OF METHOD

Audits are managerial tools used to evaluate how effectively the emission inventory team complies with predetermined specifications for developing an accurate and complete inventory. As discussed in Chapter 1 of this volume, the EIIP's mission is to develop guidance to improve the emission inventory development process. The development of quality programs, as discussed in Chapter 2, is encouraged to establish effective procedures and standards to increase data quality. After implementing these procedures, audits are conducted to determine whether the procedures are effective and whether additional QC is needed to improve the (procedures and hopefully cut the costs associated with inventory development).

Specifically, audits are managerial tools used to:

- Identify staffing problems, particularly areas where understaffing or inadequately trained staff may adversely affect quality;
- Evaluate the effectiveness of the technical and quality procedures used to develop the emissions data;
- Help ensure the completeness and accuracy of the emissions data;
- Determine whether data quality objectives are being met;
- Determine the need for additional QC measures; and
- Streamline the costs associated with the inventory development activities.

The overall goal of the audit program is continuous improvement in the performance of people, processes, and systems. When continuous improvement becomes a reality, more efficient methods are incorporated, data quality improves, and costs are reduced.

A good audit program is one of the key elements of an effective quality program because it helps to identify the sources of errors and ensures that corrective actions are taken to eliminate them. Because of the importance of the audit to continuous improvement of emissions data, audit planning must be conducted along with the technical plans for inventory development. Quality personnel must be familiar with the critical elements of a good audit program and the types of audits that can be conducted to help meet the data quality objectives established for the emissions data.

The purpose of this section of Chapter 3 is to provide information about:

- The elements of a good audit program;
- The types of audits that can be conducted; and
- Preferred and alternative audit methods that can be implemented.

This information should then be used to design an audit program that ensures the development of continuous improvement in your emissions inventory data.

8.1.1 Adequate Training

Auditors must be adequately trained, have good communication skills, and be able to foster good team relationships with the inventory development team members to develop a successful audit program. They must understand the technical aspects of the work and have auditing experience in order to identify sources of quality concerns and make reasonable recommendations for corrective actions (RCAs). When the technical work is very complex, an audit team of quality professionals and technical experts may be needed. Whatever the make up of the audit team, the members of the audit team must have sufficient training to understand the data and inventory procedures used to estimate emissions.

In order to intelligently discuss the inventory procedures, evaluate the appropriateness of the procedures, and develop a written document that clearly identifies the findings and RCAs, good communication skills as well as technical experience are needed. The auditor must be able to objectively identify the deficiencies and communicate the RCAs in a "constructive" manner, while maintaining good team relationships. This is not an easy task, but it should always be the auditor's goal. If good team relationships are not built, communication of concerns that arise during the development of the emissions data are often not shared with the auditors and a true picture of how well or poorly the work is progressing may never be attained.

The auditor should never be viewed as an inspector who legislates rules and regulations for doing work; instead the auditor must be viewed as an inventory development team member who wants to help save time, improve procedures, communicate the need for additional support and resources to management, as well as improve the effectiveness of the systems and quality of the data. The inventory development team must feel that the goal of the auditor is to "help" rather than "hinder" their efforts to achieve the common goal of developing a complete and accurate inventory.

8.1.2 MANAGERIAL SUPPORT

Without senior managerial support, the audit program is doomed! Who will provide the equipment and people required to do the work? Who will see to it that the actions requested to address the quality concerns found during the audits are taken? Who will approve the audit program?

Before beginning any quality program, senior management must agree that the audit program is needed, support the efforts of the auditor in developing the program, and approve the program prior to its implementation. If these steps are not followed, the auditor will have no recourse if a team member or inventory development manager decides to not act upon the RCAs. Therefore, the auditor works closely with senior management to develop quality procedures and programs that are sufficiently funded and include measures that are agreed upon to be needed to achieve an accurate and complete emissions inventory. The position of senior management (see Ron McKernan) in the organization could be as presented in the example organization chart included as Figure 3.8-1.

8.1.3 GOOD PLANNING

The auditor must adequately plan the audit activities in order for them to be successful. This planning should at least include the following:

- Reviewing the inventory work plan and QAP;
- Reviewing technical procedures that are critical to the technical soundness, completeness, and accuracy of the emissions inventory;
- Identifying the key steps in the inventory development process that will affect its completeness and accuracy;
- Reviewing the results from previous audits (if any);
- Developing an audit checklist that will help to provide a thorough review of applicable systems and data; and
- Developing an audit schedule that includes auditing as many of the critical phases of the inventory development process as possible.

The first documents to review when preparing for the audit are the work plan and QAP. These documents provide the scope of work and identify the technical procedures to be followed. In addition, the QAP describes the approved QA/QC procedures to be used.

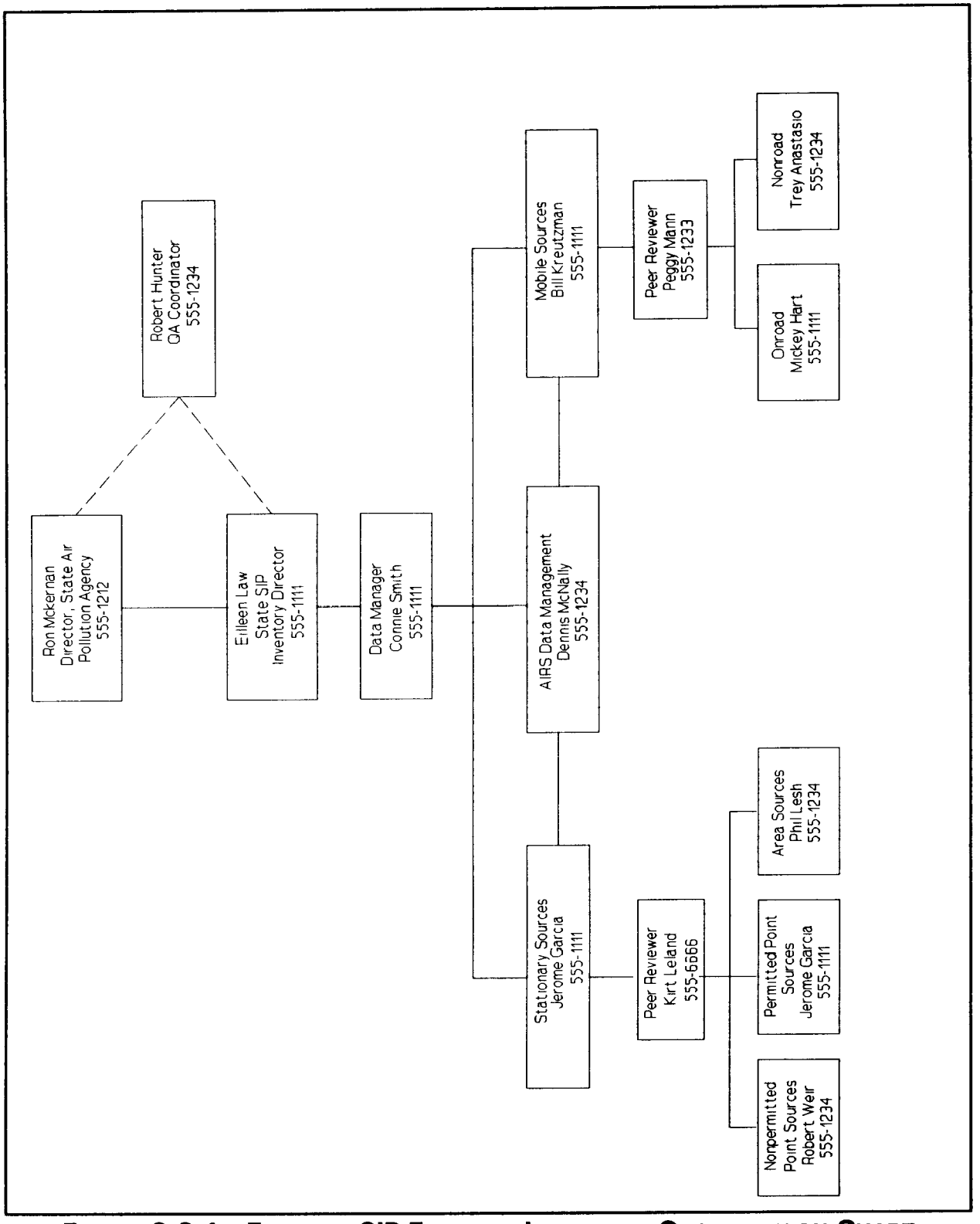

FIGURE 3.8-1. EXAMPLE SIP EMISSION INVENTORY ORGANIZATION CHART

The QAP is usually developed with the assistance of the auditor to ensure that there is agreement concerning the level of QA/QC required to achieve the technical and data quality objectives.

As a result of previous experience auditing similar inventory development activities, an experienced auditor may help identify steps in the inventory development and reporting processes where additional QA/QC may be helpful. He or she may help define more realistic data quality objectives because trends may have been revealed during previous reviews of historical data. In addition, if data completeness and accuracy were the primary concerns during reviews of previous data, additional QA/QC may be recommended during the development of the data collection, management, and technical review procedures.

During the review of the work plan and QAP, the auditor should answer the questions listed below:

- Were the work plan and QAP approved by the appropriate personnel?
- Were the EPA requirements for developing the work plan and QAP met?
- Are there a sufficient number of experienced personnel to develop the inventory within the agreed upon time frame?
- Are the responsibilities of key personnel identified?
- Will the work be adequately supervised?
- Is the objective of the work clearly stated along with the objectives for data quality?
- Is the technical approach sound?
- Are the procedures for documenting, maintaining, and storing data described?
- Are the federal/state requirements for developing the inventory discussed along with the procedures to be followed to achieve them?
- Are the reporting requirements described?
- Are milestones identified to help meet the deadline established for completing the inventory?

- Is the budget mentioned and is it adequate for the work projected?
- Are adequate QA/QC measures described (internal and external)?

Deficiencies found during the review of these documents are documented and brought to the attention of the manager of the inventory development activities as soon as possible.

In addition to the planning documents, the auditor may also review technical standard operating procedures (SOPs) and user's manuals which describe how the data will be gathered, the technical procedures for calculating emissions, and the reporting guidelines (an example is included in Figure 3.8-2). The auditor should look for completeness goals when determining the sources of emissions and a sound technical approach to calculating emissions results. Good SOPs are needed to provide consistency in the approach to gathering data, reviewing it, and developing emissions results. Standard operating procedures are more commonly used in emissions measurement programs; however, written procedures for gathering and using data for emissions inventories are also recommended (see Chapter 2).

The auditor should consider personnel training and supervision of inventory development activities as two of the most important steps in the development of a complete, accurate, and technically sound inventory. Therefore, the organization of the inventory development team and experience of the team members should be reviewed after a basic understanding of the scope of the work and technical expertise required to prepare the inventory is developed.

Adequate training can also be evaluated during the personnel interviews or audits. This evaluation could include the review of personnel training records for evidence of related experience. If deficiencies are found, plans should be made to provide training prior to or during the inventory development process. There should also be evidence of sufficient senior personnel supervision of employees who are less experienced and are being trained to perform inventory development activities.

If previous audits were conducted, the results from these audits are reviewed to determine the quality issues that were identified and actions recommended to eliminate them. The corrective action plan should be a topic of discussion during the audit and, if needed, follow-up activities should be conducted.

After reviewing all of the documents that will help provide a thorough overview of the inventory process, the auditor determines which systems and procedures are key to the successful development of a complete and accurate inventory. These critical procedures and systems will at least include an assessment of the management/supervision of the

KEY ELEMENTS OF DATABASE USE AND OPERATION

1. The system sets up the required data screens for the type of inventory to be developed.

2. The initials of the operator and the date are attached to each added or updated field.

3. The desired area or geographical locations can be selected.

4. For point sources, detailed source information module is used rather than the general module to allow the inclusion of small plant VOC emissions values during summary calculations on a county or grid level. Emissions for remaining point source categories can be added manually.

5. A calculator is available on line.

6. The system screens each data field (length, type of characters) prior to accepting the data added or edited.

7. Each field has a help screen or look-up table (SC Codes, SIC Codes, emissions estimate methods, pollutants, etc.).

8. When data are deleted, all the relative data associated with that key field are deleted to avoid the existence of orphan data. The data deleted are retained in the systems utility area until a request for removal is received. The data held in the utilities part of the system are not used for summaries or calculations.

9. Nonbanked annual emissions, ozone seasonal daily emissions, and carbon monoxide daily emissions can be calculated. If the required data are included, carbon monoxide seasonal calculations can be adjusted to daily rather than 8 hour periods.

10. When data are ready for transportation, transaction files are created. Data are edited prior to transport. If the required data are not available or in the correct format, error codes or messages will be printed to the screen. All corrections must be eliminated prior to transport.

11. The screening or validation conducted during data editing includes 1) checking the field length and data type (numeric or alpha-numeric) and 2) comparing entries against look-up tables.

FIGURE 3.8-2. KEY ELEMENTS OF DATABASE USE AND OPERATION

work, inventory development plans, data gathering activities, completeness checking, data documentation procedures, procedures used to identify emission sources, methods used to calculate emissions, QC measures, and reporting requirements. Questions concerning these critical phases in the inventory development procedures are then included in an audit checklist. Critical phases in the inventory development process may be determined by examining the example Information Flow Chart, shown in Figure 3.8-3. This chart shows data gathering, information flow, and QC activities for a typical inventory development process. Any place where QC activities are indicated are also appropriate places for QA audits. In addition, a total systems audit would cover all activities shown as well as planning activities (not shown). An audit of the software and hardware to be used should be done prior to input of significant amounts of data. Other audits might be done after the report is prepared to validate the emissions estimate generated.

The checklist is the primary audit tool used by the auditor to help prioritize the audit activities so that the most important phases of the inventory process are evaluated first. Space should be provided on the checklist to record the findings during the audit; however, additional notes may be taken. The checklist is retained and becomes part of the records maintained by the auditor. Because the notes are used to generate the audit report, the auditor should record the findings as legibly and thoroughly as possible.

Four checklists are included as appendices to this volume to provide examples of the approaches an auditor could take when evaluating critical phases of emissions inventories. The questions to consider during the evaluation of the effectiveness of the QC program associated with these critical phases also provide examples of the level of detail in each audit. These questions are based on the procedural and quality requirements included in the inventory work plan and QAP.

After developing the checklist(s) that include(s) questions relative to each critical phase of the inventory process to be audited, the auditor determines the frequency of the audits and when they will occur. The available resources often dictate the frequency of these audits; therefore, the audit schedule should be discussed with the inventory development manager prior to establishing the budget. The auditor should request sufficient funds to provide the oversight needed to achieve the data quality objectives established for the work.

Audits should be conducted throughout the development of the inventory so that deficiencies in the QC program are identified and corrected as soon as possible. Planning the audits in this manner will also foster continuous improvement of the accuracy of the data as the inventory is developed. It is recommended that the audits be conducted at least during the planning stages, during the collection of the data, during the development of the emissions estimates, and after the draft of the inventory

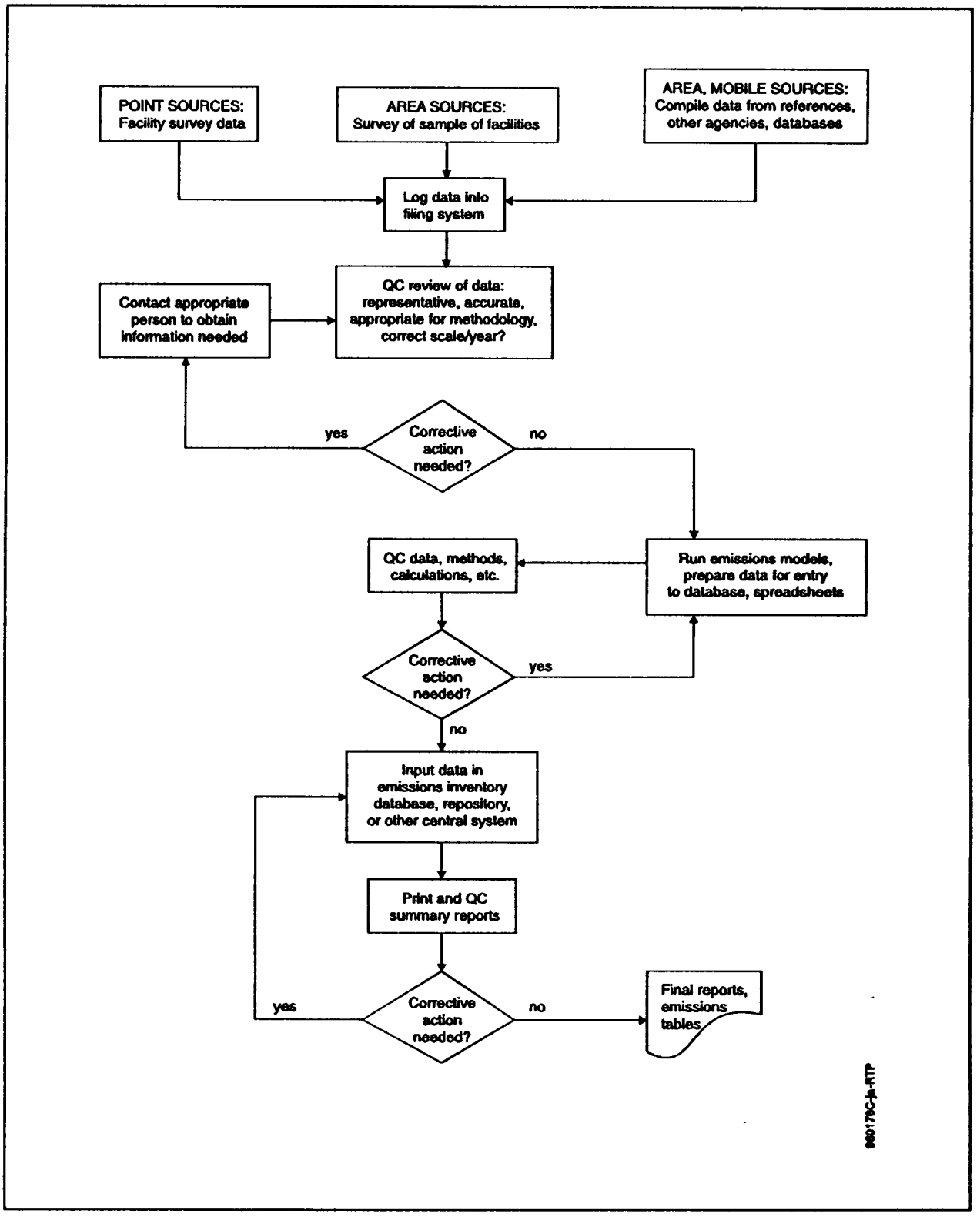

FIGURE 3.8-3. EXAMPLE INFORMATION FLOW CHART

is generated. This schedule will provide an assessment of the adequacy of the development activities and QC procedures associated with the most critical phases of the work.

As the schedule develops, the auditor must also determine what type of audit is appropriate to conduct. The auditor may conduct a management systems audit (MSA), technical systems audit (TSA), performance evaluation audit (PEA), audit of data quality objectives (DQA), or a data/report audit (DRA). The audit type selected depends on the objective of the assessment (see Table 3.8-1).

8.1.4 EFFECTIVE AUDIT PROCEDURES

In order to efficiently use the auditor's time and the time of the inventory development personnel, the auditor should communicate his/her intentions to conduct the audit and hold meetings prior to and after the audit to discuss the audit activities and findings. During each of these steps, notes should be taken to record the information gathered so it can be used to generate the audit report.

The following guidance on effective audit procedures is applicable to all types of audits. However, the relative importance of each step discussed below may vary depending on the type of audit.

Notice of Intent

Approximately one week prior to conducting a systems audit, the auditor should forward a notice of intent to conduct the audit to the inventory development manager. The notice should at least include the following:

- Objective(s) of the audit;
- Proposed date of the audit;
- System(s) or data to be audited;
- Proposed schedule of audit activities; and
- Names of the persons to be interviewed.

This notice helps to ensure that the personnel involved in the critical phases of the inventory process will be available. The schedule of activities will allow the inventory development personnel to plan their activities so that the audit will not interfere with the progression of inventory development activities any more than is absolutely

TABLE 3.8-1

OBJECTIVES OF DIFFERENT AUDIT TYPES[a]

Audit Type	Objective
Management Systems	Determine the appropriateness of the management and supervision of inventory development activities and training of inventory developers.
Technical Systems	Determine the technical soundness, effectiveness, and efficiency of the procedures used to gather data and calculate emission results.
Performance Evaluation	Determine whether the equipment used to collect measurement data operates within acceptable limits.
Data/Report	Determine whether the results reported accurately reflect the emission results calculated and recorded in the supportive data.
Data Quality	Determine the accuracy and completeness of the data used to develop the emission results.

[a]See Section 8.2 for more detailed information about these types of audits.

necessary. Identifying the data that will be reviewed also helps save time by allowing the inventory personnel to gather the data and have it available at the time of the audit.

Opening Meeting

At the start of the audit, the auditor holds an opening meeting to discuss the audit plans. During the meeting, the inventory development personnel are introduced to the auditor by the inventory development manager. One week prior to the audit, the inventory development manager may also be asked to provide a brief summary of the inventory development activities so that the auditor knows the status of the development of the inventory prior to conducting the audit. The status of the work may be updated during the opening meeting. The auditor briefly summarizes the intent of the audit and audit schedule. If needed, changes in the audit schedule may occur to ensure that the auditor is able to interview as many of the key personnel as possible. The auditor also informs the personnel attending the meeting of the proposed date/time of the wrap-up meeting, which is held after the audit.

Conduct of the Audit

After the opening meeting, the audit begins. The most effective approach is to follow the natural progression of inventory development: data collection, data evaluation, emission calculation/estimation, and emission results generation. This sequence of events should take the auditor through the critical steps of the inventory process that were identified when reviewing the planning documents and approved procedures. At each step, the auditor discusses the work with the personnel actually conducting it.

During the interviews or discussions with the inventory development personnel, the auditor asks questions about the systems and data evaluated and compares the answers to the information included in the procedures reviewed during the planning stage of the audit. If approved procedures are not followed, the auditor identifies these occurrences as quality concerns or findings. If responses to procedural questions suggest that personnel training is insufficient, training records may be evaluated to determine if personnel training is a quality concern that may impact data quality.

Likewise, the auditor must consider the operability and ease of use of all instruments used to conduct any of the inventory development activities. This is usually not a concern for inventories developed by an agency using primarily facility data and the emission factor approaches. However, a facility's inventory could involve the use of measured emissions. For example, if the audit includes an evaluation of continuous emissions monitoring of stack emissions, the operation, calibration, and maintenance of the equipment should be evaluated. Equipment maintenance logs should be reviewed to determine if there are recurring problems that could adversely affect the results. After reviewing these records, interviewing the operators, and determining that the operability of the measuring equipment or computers used to calculate emissions is questionable, it should be noted as a finding.

The auditor should plan to spend as much time as possible tracking the receipt, maintenance, and retention of the data gathered and used to estimate/calculate emissions. Some of the major quality concerns found when evaluating emissions data are the lack of sufficient information to verify the accuracy of the emissions results, poor documentation of the supporting data, and the absence of good data management practices. Key questions to ask during the audit to determine the appropriateness of these practices are included in Figure 3.8-4.

As the audit is conducted, the auditor documents the findings on the checklist or in a bound notebook. The notes should be legible and written in sufficient detail to facilitate their use during the development of the audit report, indicate the date the notes were recorded, and include a signature line. The auditor looks for opportunities to praise systems that are working well and personnel who appear to be taking extra

Concerns	Yes	No	Comments
Are the data from the appropriate year?			
Are all of the emission sources identified?			
Is all of the information needed to calculate the emissions results available?			
Are data documented in a manner that will not lead to misinterpretation of the information (e.g., crossing out and obscuring data when making corrections?)			
Are data retained in an easily retrieved manner and controlled so that its location is always known?			
Is black ink used to record data to help reduce the possibility of data loss when there is a need for reproductions?			
Are the procedures used to electronically calculate values clearly explained and are the results reproducible?			
Are the calculated values verified by someone other than the generator?			
Are unit conversions accurate?			
Are the emissions results in the specified units?			
Is the technical approach as indicated in the work plan?			
Auditor:			
Date:			

FIGURE 3.8-4. EXAMPLE DATA AUDIT CHECKLIST

steps to control the quality and accuracy of the data since positive feed-back as well as RCAs should be included in the audit report.

Wrap-up Meeting

Wrap-up Meetings are held after the audit to discuss the findings and any other information that may be pertinent for the auditor to know prior to writing the audit report. This preliminary discussion of the audit findings allows immediate communication of findings that are considered adverse to data quality so that actions can be taken immediately to eliminate them.

The wrap-up meeting is often attended by more personnel than are usually present at the opening meeting. Those personnel interviewed or involved in inventory development activities during the audit may attend this meeting to get immediate feed-back on the audit findings. Senior management may also attend this meeting to get first hand information about the requests for corrective actions and effectiveness of the QC program implemented by the inventory development team.

The auditor should request sufficient time before the meeting to review the notes taken and summarize the audit findings. The goal of the auditor during this meeting is to ensure that the inventory personnel agree that the findings are legitimate concerns that require resolution or corrective actions. The audit findings should be clearly stated along with the impact they may have on the development of a complete and accurate emissions inventory. If additional information can be provided to clarify the auditor's perspective on an issue, it is provided at this time or arrangements are made to provide it soon after the meeting.

8.1.5 AUDIT REPORT

The audit report should be clearly written and distributed to the supervisor/manager of the inventory development activities and the key personnel involved in the audit as soon as possible, but no later than two weeks after the audit. Examples should be given of data, procedures, or systems that are found to be deficient, incomplete, or inaccurate. Examples should also be given of systems that work well and personnel that are doing an exceptional job.

The following information is included in the audit report:

- Identification of the auditor;
- Date of the audit;
- Facility or location of the audit;
- Objective of the audit;
- Descriptions of the system(s), data, procedures audited;
- Names of the persons interviewed;
- Findings (positive feed-back and RCAs); and
- Discussion of follow-up activities that relate to previous audits.

An example audit report is included as Appendix A.

The auditor's goal when reporting an issue found during an audit is to clearly state the concern, the potential impact on the emissions inventory results, and propose a solution or RCA to eliminate the concern. The RCAs should be accompanied by a description of the situation that led to them along with a possible solution to eliminate the quality concern.

For example, if the data for one source was missing, the auditor would indicate that the absence of this information would lead to an incomplete inventory or failure to meet the data quality objective for completeness for a given source. The recommendation could be to use engineering judgment or previous data to estimate the omitted data and flag the results as estimations. The inventory report should include an explanation for reporting estimations rather than results that are based on actual test data.

On the other hand, if an instrument was found to not be appropriately calibrated or tested, or an emissions estimation software system has not been adequately QA'd, the auditor could state that there is a potential for the emissions results developed from this data to be inaccurate. The recommendation could be to qualify or flag the results based on these data to show that they are estimated values because the instrument was not properly tested prior to gathering the supportive data. The auditor could ask the team to evaluate the possible degree of inaccuracy that would be associated with both sets of data given in the examples.

The example included as Figure 3.8-5 could be used to record the RCA and maintain records of the actions taken in response to it. Note that the resolution includes a planned corrective action, identification of the person proposing the action, proposed date of the action, and a description of the action taken.

Follow-up Activities

The auditor must assure that actions are taken in response to the audit findings and that these actions adequately address the quality issues cited in the audit report. Conducting these follow-up activities and determining the recurrence of quality issues are the basis for continuous improvement.

Categorizing the findings helps to identify trends. Trends in the occurrence of quality issues can be used to identify systems that may required additional QA/QC. To promote continuous improvement, the auditor may conduct more system specific audits to determine why the applicable procedures are not producing the desired results. Recommendations can then be made to improve the effectiveness and efficiency of the procedures.

RECOMMENDATION FOR CORRECTIVE ACTION

A. Initial Information

RCA NO:	DATE:	URGENCY LEVEL ☐
ORIGINATOR:	APPROVED BY:	1. Potential for major data loss or information. 2. Potential for failure to achieve data quality objective. 3. Suggested improvement.
ORGANIZATION/INDIVIDUAL RESPONSIBLE FOR ACTION:		

B. Problem Identification

SITE/LAB:	SYSTEM:	DATE PROBLEM IDENTIFIED:
DESCRIPTION OF PROBLEM:		

C. Recommended Corrective Action

DESCRIPTION:	IMPLEMENT BY:

D. Problem Resolution

PLANNED CORRECTIVE ACTION:	PROPOSED BY:	DATE PROPOSED:	SCHEDULE IMPLEMENTATION:
IMPLEMENTED CORRECTIVE ACTIONS:			DATE IMPLEMENTED:

E. QA Verification

VERIFIED BY:	DATE:	COMMENTS:

FIGURE 3.8-5. RECOMMENDATION FOR CORRECTIVE ACTION

The responses from the inventory development manager concerning the RCAs should include a description of the actions to be taken, identification of the responsible person, and date of implementation. The auditor keeps track of the responses to the audit findings to ensure that they are adequate to eliminate the concerns identified and are implemented within a reasonable time. The time between acknowledgement of the concern and implementation of the action should be based on the severity of the impact on data quality. If there is a possibility that the objective of the inventory will not be met or other data are impacted, actions should be taken immediately to eliminate the concern. If new procedures are to be developed or systems revised, but the concern only has the "potential" to adversely impact data quality, the auditor may agree to a longer period between the reporting of the RCA and implementation actions taken in response to it. Regardless of the time interval between the request and implementation of corrective actions, the auditor's job is not done until the quality concern is eliminated. This closure of the issue and frequent audits are the keys to the development of an accurate and complete emissions inventory.

Record Keeping

The auditor should keep all notes taken and information gathered during the audit to provide the supportive data needed to verify the information included in the audit report. A data file should be developed that at least includes the completed audit checklist, notes taken during the audit, technical planning documents that described the objectives of the work and approved procedures, the audit report, and descriptions of actions taken in response to the RCAs.

The records maintained by the auditor are used to monitor the quality of the work during the development of the inventory. Categorizing the findings helps to identify weaknesses in the QC program. This information is also used to plan other audits of systems or data that will help monitor the activities until the number of deficiencies or concerns found during the audit cease to exist or decrease to a reasonable level.

8.2 Types of Audits

The preceding description of an audit assumes the conduct of a complete systems audit that evaluates the effectiveness of all systems associated with the development of an emissions inventory. However, other types of audits may be conducted to allow the auditor to concentrate on specific systems or procedures that are critical to the successful outcome of the inventory development activities.

Because most technical studies involve numerous systems that could result in the collection of volumes of data, other types of audits may be needed to provide a more

thorough assessment of the adequacy of specific systems or the effectiveness of certain procedures. To accommodate the need for a more thorough assessment of phases of the inventory development process, the auditor may conduct MSAs, TSAs, PEAs, DQAs, and DRAs.

The objective of the audit is used to determine which type of audit is appropriate to conduct. When the auditor needs a qualitative assessment of all or parts of the inventory development process, TSAs may be conducted. During these audits, the auditor can focus on one, several, or all of the procedures and systems. If the auditor is concerned about the adequacy of the devices used to take measurements, a quantitative assessment of the measuring device is needed; therefore, a PEA is conducted. If the auditor is concerned about the manner in which data are recorded and whether it is complete, a DQA may be conducted. Moreover, the auditor conducts a DRA if there are concerns about the accuracy and completeness of the emissions results and format of the report. Regardless of the type of audit conducted, the auditor plans to conduct the audits throughout the inventory development process to help improve the procedures and data as the emissions results are developed.

8.2.1 MANAGEMENT SYSTEMS AUDITS

Management systems audits are important because the findings can be used to verify the existence and evaluate the adequacy of the internal management systems that are needed to develop a complete and accurate inventory. Generally MSAs are on-site evaluations of personnel management, project management, resources, training, and scheduling. The objective of the MSA is to determine if the work is adequately managed and supervised to meet the objectives of the work. Therefore, the commitment of management to establish and implement adequate QA and QC programs is evaluated. It is important that the evaluation of the QA program is conducted by an external auditor to provide an objective assessment.

Some of the questions that may be included on the checklist used during the conduct of a MSA are as follows:

- Is there an approved QA program that includes adequately trained QA personnel?
- Is the QA program supported and approved by senior management to ensure that actions are taken beyond the reporting level of the inventory development manager to get RCAs implemented?

- Is there a mechanism that allows the QA personnel to keep abreast of the inventory development activities so that audits can be planned at critical phases in the inventory development process?

- Is there evidence of adequate supervision of inventory development activities?

- Does management provide adequate supplies and personnel to get the inventory developed within the agreed upon schedule?

- Have technical work plans and quality plans been developed and distributed to the inventory development staff and QA staff to ensure that the objectives and approved procedures for conducting the work are clearly understood?

- Are routine meetings held throughout the development of the inventory to provide an opportunity to discuss the successes and short-comings that may have arisen during the conduct of the work?

- Is there an effective corrective action mechanism that ensures the elimination of activities that adversely affect the technical and data quality objectives?

During the MSA, the auditor tries to determine if there are adequate management support and appropriate quality programs to ensure that the quality of the work is controlled. Provisions should be made to implement the QA/QC programs as the inventory is developed to ensure continuous improvement and the achievement of the objective of the work.

8.2.2 TECHNICAL SYSTEMS AUDITS

The TSA is a qualitative assessment to determine if the approved procedures applicable to the work are followed. Because these procedures should be described in the planning documents, this audit is conducted to determine if the approved procedures are being followed and how effective they are.

Example procedures and systems that could be evaluated during a TSA and important questions to answer during the audit are included below:

Technical Planning

- Project Organization and Personnel
  - -- Is there a sufficient number of adequately trained personnel to conduct the work and meet the deadline for reporting the results?
  - -- Does the organizational chart identify the QA Coordinator?
- Are there adequate funds for the resources needed?
- Are the constraints for doing the work outlined in the work plan or QAP?
- Are there procedures and goals for achieving data quality?
- Are technically sound procedures identified and explained in sufficient detail?
- Is there a reasonable inventory development schedule?

Data Collection and Analysis

- Are emission source categories and data elements prioritized?
- Have all emissions sources been identified? How do you know?
- Are data collection procedures adequate?
- Are results calculated in a consistent manner?
- Are instruments used to collect real-time data adequately tested?
- Are the computer software and hardware adequate for their use in managing data or calculating results?
- Are the computer systems validated to be functioning as required prior to each use?
- What procedures are used to verify the accuracy of the calculated results?

Data Handling

- Are data receipt records maintained and tracking numbers assigned to help monitor the completeness of the master data file?
- Is access to the data controlled to help maintain a complete file of all of the supportive data needed to verify the emissions results?
- Are there procedures for handling missing data?

Data Documentation

- Is the information recorded in a manner that will help avoid misinterpretations?
- Are the data documented in a manner that facilitates reconstruction of the development of emission results?
- Are there data correction procedures that help to reconstruct the reason for the correction and ensure the use of the accurate information?

8.2.3 PERFORMANCE EVALUATION AUDITS

Performance evaluation audits are conducted when there is a need to quantitatively assess the accuracy of a measurement system. These audits are infrequently conducted unless there is collection of real-time data to support the emissions results. Sometimes the appropriate corrective actions to take if the operation of the system is questionable are not always apparent; therefore, PEAs are usually conducted in conjunction with TSAs to evaluate the impact that associated systems may have on the operation and maintenance of the measuring device.

The measurement system evaluated during a PEA could be a continuous emissions monitor (CEM) that is used to determine the composition and concentration of the emissions from a stack located at a chemical plant. During the audit, the auditor determines how close the results from the measuring devices are to the known concentration and composition of the reference material. Sometimes the identity of the audit or reference material is unknown to the operator to help provide an objective assessment of the results.

Regardless of the measurement system, the objective is to determine the:

- The accuracy and precision of the measuring device;
- The acceptability of the QC data collected during routine operation of the measuring device;
- If the device is operating within the established control limit established in the QAP; and
- Changes in the quality of the data between audits or over a specified period of use.

The results are used to evaluate the usefulness of the measurement data in determining emissions results. Measurement data naturally result in the development of a more accurate emissions inventory because actual data are used to perform the calculations.

8.2.4 Data Quality Audits

Audits of data quality are conducted to determine the acceptability of the data gathered or developed and used to determine emissions results. During the audit, the auditor assesses the adequacy of the data documentation, collection, review, and management activities. The use of the data to calculate emissions is also evaluated to determine if the technical approach is sound and the conversion of units is accurate.

The auditor should review the results from the quality assessment conducted by the generators of the original data to evaluate the acceptability of data developed externally and forwarded to the inventory developers. If this evaluation is not possible, the results should be qualified by acknowledging that the results were determined from data of unknown quality.

If measurement data are used to calculate emissions, the auditor should review the QA/QC data and determine whether the data are valid for use. The QA/QC data provide important information about sample integrity prior to analysis, sampling technique, sample contamination, and the acceptability of the analytical system. The auditor should review the manner in which the data were flagged to show bias and determine whether the values used to calculate emissions were valid.

Some of the questions that may be asked during this type of audit are listed below.

Data Management

- Is the procedure for logging, filing, retrieving, and storing the data included in a standard operating procedure?
- Are the data logged upon receipt and their use tracked to maintain a complete file of all of the data relevant to the emissions results?
- Are the data retained in one location and access to them controlled (helps to ensure completeness)?
- Are data in the file complete (trace from result back to actual data to see if results are reproducible and supportive data available)?
- Is one person responsible for maintaining the data file?

Data Documentation

- Are data recorded and corrected in a manner that helps to reduce misinterpretation of the information?
- Are data documented in a manner that will render them easy to reproduce, when reproduction is necessary, without risking the chance of losing data because of poor copies?
- Are the name of the data generator and date the data are generated included on each data sheet?

Data Quality Indicators

- Were the data quality goals met?
- How are data gaps eliminated?

Data Generation

- Are calculations explained in sufficient detail to reproduce the results?
- Are units accurately converted?

- Are the results in the required units?
- Are procedures for identifying outliers identified and followed?
- Are standard operating procedures used to round numbers?
- Do results represent the correct year?

Data Processing

- Are computer programs documented and validated?
- Are programs or spreadsheets tested prior to each use?
- Is the accuracy of data transcriptions checked?

Examples of completed work sheets and data audit checklists are included at the end of this section.

8.2.5 DATA/REPORT AUDITS

The data/report audit is a type of systems audit that is conducted to determine the accuracy of the emissions results reported and determine whether there are sufficient data to verify the results. This audit is conducted after the data and report have been checked internally and just before it is forwarded to the requestor. The primary goals of the DRA are to determine whether the report accurately reflects the data and whether the objectives for the work were met by using the approved procedures included in the work plan and QAP. An example DRA checklist is included in the appendix.

Because it is impossible to check every data point reported, the auditor evaluates the complete report and selects key data points that could help assess the accuracy of the remainder of the data. A good understanding of the procedures followed by the data submitters, inventory development team, and reporter helps to conduct an audit that makes efficient use of time and resources.

During the DRA, the auditor attempts to verify the accuracy of all of the concluding statements and raw data included in the report by comparing the data reported to that found in the data file. All of the basic equations used to calculate emissions and the technical approach are confirmed to be as indicated in the approved work plan and QAP. Some of the values are re-calculated to verify the relative accuracy of the emissions data.

In addition, the auditor also determines whether the reporting requirements were actually met.

- Are the results in the required units?
- Is all of the required supportive information included in the report?
- Does the report meet the formatting requirements?
- Is the report dated?
- Are the results from the specified year required to be evaluated?

The questions included in the DQA section are also used to assess the quality of the report and data. The peer review checklist shown in Figure 3.3-1 of Section 3 of this chapter could be used for a DRA.

The DRAs are the last audit to be conducted during the development of the emissions inventory. The information provided by the auditor is used to assess how effectively the QA and QC programs were implemented. The RCAs that require correcting reported values or information are addressed immediately and the auditor remains involved in the corrective action process to ensure that the emissions data are accurately reported. As with all of the audits, information from the audit is used to plan the audit program for the next emissions inventory and help improve the procedures used to control the quality of the inventory data.

8.3 Preferred and Alternative Methods

There are several approaches or methods to follow when auditing emissions inventory development activities. The advantages and disadvantages of these approaches or methods, and the objectives for completing the inventory should be the determining factors when deciding which approach is most appropriate.

For evaluations of inventories developed by state agencies and industry, alternative methods are needed because limited resources or other constraints are inherent to many agencies and facilities developing the inventories. However, before selecting an alternative audit method, one must accept the limited degree of quality that the method will provide. Therefore, alternative methods for conducting audits should only be implemented when limited resources and other constraints do not permit the implementation of the preferred method or approach.

The preferred and alternative audit methods are included in Table 3.8-2.

TABLE 3.8-2

PREFERRED AND ALTERNATIVE METHODS: AUDITS

Method	Identification	Description
Preferred	Independent/External Audit	Audit(s) conducted by an auditor or audit team of experienced quality professionals and technical personnel who do not report to the inventory development team manager or assist in the development of the inventory.
Alternative 1	Internal Audit	Audits conducted by auditors who are also involved in the inventory development process.
Alternative 2	Internal Data Checking	QC checking only by the inventory development team members.

8.3.1 INDEPENDENT/EXTERNAL AUDIT

An independent or external audit is conducted by an experienced auditor who is not involved in the emissions inventory development process. The independence of the auditor is signified by the reporting lines to management. In order for the auditor to be independent of the inventory development team and thus be able to provide an objective assessment of the systems and data evaluated, he/she must not report directly to the manager of the inventory development team for performance evaluations. However, it is not necessary that the audit be conducted by another agency or contractor; if the auditor was not involved in the development of the inventory, the audit may still be independent.

During the conduct of independent audits, the auditor must be able to objectively assess the procedures, systems, and data without fear of being reprimanded if the inventory development manager is not in agreement with the findings or RCAs. For this reason, in an organization where quality programs are well-established, the auditor reports to a QA Director or senior management.

Independent experienced auditors will develop an audit program that is of a higher quality than those conducted by inventory development personnel because of their knowledge of effective quality procedures. Successful auditors will also foster good

working relationships with the inventory development team. This team building approach is usually encouraged by the auditor to help learn more about the procedures and concerns that the team members may have about data quality. Because of the experience and independence of the auditor, ineffective procedures are more readily identified and more effective solutions are recommended when corrective actions are required. In addition, the independent auditor is better equipped to provide continuous improvement throughout the development of the inventory.

One short-coming to this preferred approach to conducting audits is the need for independence. There could be additional costs if there is a need to expand the inventory team to include an auditor or audit team. However, experienced auditors could help eliminate the costs associated with repeating work because of unacceptable data quality or the incorporation of ineffective QC measures.

Independent auditors could be personnel from other groups or agencies within an organization; therefore, selecting personnel from other groups is highly recommended as a means of providing an objective assessment of the quality of the inventory data and effectiveness of the inventory development procedures. Hiring a contractor with inventory QA experience is another option. Regardless of the source of the auditors, technical as well as QA auditing experience are the keys to implementing a successful independent audit.

8.3.2 INTERNAL AUDIT

Internal audits are conducted by personnel who are engaged in the development of the emissions inventory. Some degree of independence is attained by ensuring that personnel do not audit their own work. For example, a point source inventory developer might audit the area source inventory process. However, because the auditor reports directly to the supervisor or manager, he/she may not be able to provide as objective an assessment of the work as an external auditor would. The reporting lines of authority may also prevent the auditor from resolving issues that may not be favorable to the inventory development manager.

Because the internal auditor is also involved in developing the inventory, he/she may not be able to provide an unbiased assessment of the effectiveness of the procedures or systems evaluated. For example, it is difficult for a writer to edit his work because he often overlooks mistakes that are readily seen by someone who has never seen the document or worked on it. The effectiveness of procedures may also be overlooked because the auditor is so accustomed to seeing the work conducted in this manner that he/she never considers revising them. Therefore, an internal audit could lead to a higher incident of error because some mistakes or quality issues may be overlooked.

One advantage of the internal audit is that the cost is probably at least partially included in the budget because the auditor is performing technical duties as well as QA duties. Another advantage could be that the auditor is closer to the work so he/she may understand the technical approach and systems better than an auditor who has not been involved in the development of the inventory. The proximity to the work and working relationship with the personnel can also be an advantage that would help to plan and implement the quality program in less time than it may take an auditor who is only involved in the QA aspects of the inventory. However, the decrease in cost associated with these advantages may not be off-set by the probability of a higher degree of errors.

8.3.3 INTERNAL DATA CHECKING

Internal data checking is the most frequent procedure followed to help ensure the accuracy of the inventories; in this procedure, the only "audit" is a review of calculations and procedures by the personnel who originally did the work (or by their colleagues). This method has the highest probable error rate of the three. The error rate is higher because there is no formal quality program that requires assessments of the effectiveness of all of the systems and procedures used to develop the inventory. Examples of the systems that are critical to data accuracy and completeness are management of inventory development activities, personnel training, data documentation, data collection, data retention, and formal corrective action mechanisms. All of these systems or procedures have a tremendous impact on the quality of an inventory.

Internal data checking primarily involves checking the accuracy of calculated results and data transcriptions using the methods described in previous sections of this chapter. It could also include the use of a peer reviewer to evaluate the soundness of the technical approach. This form of QC often does not lead to continuous improvement unless management establishes a mechanism that provides constant feedback of quality issues and corrective actions to the inventory development team.

Continuous improvement also requires the implementation of measures that ensure frequent evaluations of the effectiveness of the inventory development procedures. The results from these evaluations must be documented and reviewed constantly to determine when revisions of the inventory development procedures are required to avoid data loss and inaccurate emissions results. Therefore, the major disadvantages of only incorporating internal data checking as a means of QC are:

- The absence of a formal quality review program with follow-up activities to help foster continuous improvement; and

- The absence of a formal means of reviewing the effectiveness of technical procedures and evaluating the negative impact that these procedures may have on the accuracy and completeness of the inventory.

If internal data checking and the development of the technical approach to the work are well planned, if there is supervision of the work, and if the inventory development team is very experienced, the quality of the inventories developed by incorporating this method of QC could be increased. The planning would have to include supervision of work practices and review of the data as the inventory is developed. Communication between managers, supervisors, and team members is critical and QC procedures would have to be implemented to determine when there are activities that adversely impact data quality. Status meetings would also have to be held to keep abreast of the quality issues and actions taken to implement corrective actions.

Incorporating these QC measures and using an experienced inventory development team would make the internal data checking method an acceptable approach to controlling the accuracy and completeness of emissions inventories.

9

EMISSIONS ESTIMATION VALIDATION

QUICK REVIEW

Definition:	Comparison of the estimated (model) emissions to real-world measurements (or surrogates).
Objective:	Prove the correctness of the model, assumptions, and data used to estimate emissions.
Optimal usage:	(1) After model development is complete, but before it is used on a wide scale, and (2) after inventory is completed, but before it is used as basis for decision making.
Minimum expertise required:	High. May require expertise in areas outside of inventory development (such as modelling, sampling, ambient data analysis).
Advantages:	(1) Results in an independent test of an emissions model or inventory, (2) may give insights into areas needing more research, and (3) some methods can be used to identify missing or inadequately represented sources.
Limitations:	(1) Can be labor (and resource) intensive, (2) most methods depend on availability of additional data or models of high quality, and (3) some methods are relatively new, and have their own built-in uncertainties.

9.1 Overview of Method

The validation of emission estimates is the ultimate tool for assuring the quality and, more importantly, determining the accuracy of emission estimates. Validation can be used to "prove" (ideally using statistical methods) the accuracy of an emissions model for a specific category. Also, some methods can be used to validate entire inventories.

All validation methods rely on independent measurements, but not necessarily direct measurements of emissions. Surrogate variables known to be correlated with the emission activity may be measured, or ambient concentrations may be measured. In the latter case, additional modelling is generally used in the analysis; these models have some uncertainty associated with them, however.

Furthermore, field measurement and receptor modeling methods offer completely independent checks on emission inventories since the methods rely on monitoring data alone to estimate either the direct or relative contribution of emissions from individual sources (or source groups). The inverse modeling techniques are not totally independent because they use the emission estimate as the starting point for the air quality modeling. However, output of the model is the best-fit adjustment required to the emissions inventory to allow the modeled concentration data to match observed ambient concentrations. Thus, each method provides the emissions inventory analyst the tools to verify the relative or absolute emissions estimated through the emission inventory process and identify potential problem areas, weaknesses, and strengths in current inventory methodology.

Validation methods should only be used to evaluate and/or calibrate a thoroughly QA/QC'd model. They are not substitutes for the standard QA and QC procedures described in previous sections of this chapter. Moreover, because of the potentially high cost of validation procedures, attempting to validate an inventory that has not been thoroughly QA'd can result in false results and a waste of resources.

Because the procedures described here are relatively expensive, require expert skill, and, in some cases, are fairly new and not entirely accepted by the inventory community, the EIIP considers these an optional component of an emissions inventory QA program. At the same time, the EIIP encourages the continual development and use of validation methods. Studies such as the ones described below are ultimately the best way to determine the accuracy (and in particular, the bias, see Chapter 4) of emission inventories.

9.2 DIRECT AND INDIRECT MEASUREMENTS

Direct and indirect measurement methods involve field measurement of actual emissions or surrogate parameters highly correlated with the emitted pollutant. These measurements can be used for two purposes. First, the data can be used to verify and evaluate emission models. Second, the field data can be processed to produce direct estimates of emissions and emissions uncertainty for the sources measured (See Chapter 4). Because of their ability to provide unique sources of information on emissions, field studies have the potential to significantly enhance the understanding of the processes producing emissions and the uncertainty in these emissions.

Prior to using field data to calculate emissions, develop emission factors, or model the data to enhance an understanding of the emission process, one must evaluate the accuracy of the raw data used to generate the field data report. This evaluation should include reconstruction of the work and verification that the field/laboratory work was technically sound and the data points are accurate. The QA/QC information provided with the reported results may also be reviewed to assure that the results are of sufficient quality for their intended uses.

Evaluation of original field data often allows the reviewer to evaluate the bias, if any, associated with each data point. Bias may be related to field sampling techniques, sample contamination, or an unstable measurement system. To assess the validity of the data points, QC sample results or results from the analysis of blanks, matrix spikes, duplicate samples, replicate sample, or instrument calibration check samples can be evaluated. The intended use of the data should be used to determine the appropriate level of review needed. See the levels of inventories based on their intended use in Table 2.1-2 of Chapter 2 of this document.

A well-known field study is the "tunnel study" of Pierson et al. (1990) in which motor vehicle emissions were estimated from air pollutant concentrations measured within the Van Nuys highway tunnel in Los Angeles. The results of this study seemed to indicate that the available emission factors model in use at the time significantly underestimate VOC emissions from motor vehicles while NO_x emissions seem to be estimated satisfactorily. However, the interpretation of these data have been questioned by EPA and others and the results are subject to debate.

Remote sensing has significant potential for use in estimating both point source and area source emissions and emissions uncertainty in industrial settings. For example, Spellicy et al. (1992) used Fourier transform infrared (FTIR) spectrophotometry to estimate speciated VOC concentrations due to emissions from area sources within a petrochemical complex. When these data are correlated with process information and meteorological data, estimates of emissions and emissions uncertainty can be developed.

Indirect measures also provide significant information that can address emission uncertainty issues. For example, Claiborn et al. (1995) released a surrogate tracer species (SF_6) to measure the relative dispersion of the tracer and PM_{10} from paved and unpaved roads. The ratio of the measured PM_{10} and tracer concentrations and the tracer release rate was then used to estimate emission factors for PM_{10} emissions from the roadways. Estimates of uncertainty can then be determined directly from the field measurements. These tracer methods make the fundamental assumption that the tracer is emitted and disperses in the same manner as the pollutant of interest. This assumption may not hold in some situations so care must be taken when performing tracer studies and when analyzing the study results to ensure that this key assumption is met.

Fujita et al. (1992) estimated ratios of various pollutants developed from emission inventories and ambient monitoring data from the South Coast Air Quality Study (SCAQS). Table 3.9-1 presents results from Fujita et al. in which they compared the ratio of ambient and emission inventory estimated ratios of CO/NO_x and $NMOC/NO_x$. Included in Table 3.9-1 are the adjustment factors to the exhaust portion of the emission inventory that would be required to bring the emission inventory ratios to equality with the observed ambient ratios (i.e., remove a potential bias). Based upon this analysis, they suggested that VOC from motor vehicles, relative to NO_x, appears to be underestimated in motor vehicle emission inventories, the same conclusion reached by Pierson et al. discussed above.

Janssen and Koerber (1995) have also compared computed to observed VOC to NO_x ratios as part of the Lake Michigan Ozone Study (LMOS) in an effort to understand uncertainties in the LMOS emission inventory. Their initial analysis of VOC/NO_x ratios indicated differences of more than a factor of two with the estimated ratio higher than the observed ratio. However, after a prioritized review and reassessment of the sources of VOC emissions contributing most significantly to the estimated ratio, the difference between the observed and estimated ratio was reduced to under 30 percent. This study is a good example of the role that observed monitoring data can play in prioritizing enhancements to emission inventories.

Peer et al. (1992) used field sampling data to develop a regression model and to validate and calibrate a physical model of emissions of methane from landfills. In this study, data from 21 landfills with gas recovery systems were collected and used to develop a regression model of methane emissions. Information included in the model were methane recovery rate, landfill size, refuse mass, site age, and climatic measures (average precipitation, temperature, and dewpoint temperature).

The data for each landfill were also used to validate the EPA's Landfill Air Emissions Estimation Model. This model is used to estimate methane (CH_4), carbon dioxide

TABLE 3.9-1

COMPARISON OF AMBIENT AND INVENTORY RATIOS OF CO/NO_x AND $NMOG/NO_x$ FOR LOS ANGELES FOR 1987

Pollutant Ratio	Season	Time (PST)	Amb/EI Ratio[a]	Required MV Adj Factor[b]
CO/NO_x	Summer	06 - 08	1.4	1.5
		20 - 08	1.6	1.9
	Fall	06 - 08	1.1	1.2
		20 - 08	1.4	1.7
$NMOG/NO_x$	Summer	06 - 08	2.5	4.4
		20 - 08	3.0	5.5
	Fall	06 - 08	1.7	2.8
		20 - 08	2.3	4.0

Source: Fujita et al., 1992.

[a] Ratio of the pollutant ratio estimated from ambient monitoring data to that estimated from the motor vehicle emissions inventory.
[b] Adjustment factor required to bring the hot motor vehicle emissions inventory ratio to equality with the ambient ratio.

(CO_2), and VOC emissions for individual landfills; it is sensitive to the input values for k (decay rate) and L_0 (methane potential of the waste). Table 3.9-2 shows a comparison of the performance of the deterministic model and the regression model. The numerical values shown are the ratios of actual methane flows (at each of the 21 landfills identified by "site number") to predicted methane emissions for the models. The closer the ratio is to 1, the better the predicted value.

The Landfill Air Emissions Estimation Model was run using three different sets of input values for k and L_0; Run 3 used the recommended defaults. Based on this study, these defaults were shown to overpredict emissions (mean ratio is 1.97). Subsequent versions of the model recommended new default values similar to the input values used for Run 2.

Field studies provide the information needed for direct estimates of both emissions and the uncertainty associated with those estimates. However, there are two significant

TABLE 3.9-2

VALIDATION OF LANDFILL METHANE EMISSIONS MODELS[a]

Site Number	Landfill Air Emissions Estimation Model			Regression Model
	Run 1 Pred./Actual	Run 2 Pred./Actual	Run 3 Pred./Actual	Pred./Actual
1	0.16	0.40	0.73	0.52
2	0.48	1.21	2.23	1.55
3	0.28	0.71	1.31	0.83
4	0.22	0.55	1.01	0.62
5	0.58	1.44	2.66	1.95
6	0.24	0.60	1.10	0.73
7	0.46	1.16	2.14	1.50
8	0.37	0.93	1.71	1.15
9	0.36	0.90	1.67	1.15
10	0.25	0.64	1.17	0.83
11	0.23	0.57	1.05	0.85
12	0.54	1.34	2.47	1.72
13	0.16	0.39	0.72	0.57
16	0.33	0.82	1.52	1.02
17	0.49	1.23	2.26	1.73
20	0.41	1.02	1.88	1.24
21	0.15	0.36	0.67	0.47
22	0.19	0.47	0.87	0.57
23	1.74	4.35	8.00	6.32
24	0.54	1.34	2.46	1.60
25	0.82	2.06	3.79	2.34
Mean	0.43	1.07	1.97	1.39
Standard Deviation	0.34	0.85	1.56	1.24

[a]From Peer et al., 1992.

limitations of field studies that must be addressed in any field program. First, the emissions data are only valid for the conditions and sources that were present during the field measurements. Because it is not possible to make measurements under all conditions for all sources, there are inherent biases that will be present in all emission models developed based on the field data if these models are applied outside the development bounds. For example, if the traffic in the tunnel during the Van Nuys tunnel study were atypical of traffic in the Los Angeles area and that represented internally in the emission factor model, the results of the analysis of the data from the study would have limited applicability.

A second limitation is that field studies must limit their effort to generally well-defined source classes or source types. It is not possible to design a comprehensive field program to measure all, or even most, of the emission sources in an urban area. Consequently, there will be inherent limitations in the applicability of any emission model or uncertainty estimates developed using the data.

9.3 RECEPTOR MODELING

Receptor modeling, also called source apportionment, is a "top-down" method of estimating the relative contribution of individual source categories to an emission inventory. The Chemical Mass Balance Model (CMB) (Watson et al., 1984) is a widely used model for performing receptor modeling. It is called top-down because it uses information about the entire inventory and modeling domain to estimate the relative contribution of emissions from each source category rather than building up contribution estimates on a source-by-source basis.

In receptor modeling, least squares estimation is used to obtain the best fit of emissions from each source modeled that reproduces the chemical composition of observed monitoring data at a given monitoring site. The two key input requirements for receptor modeling are the chemical composition of ambient monitoring data and the chemical composition ("fingerprints") of emissions from each source with the corresponding same level of detail as the monitoring data.

The model is able to resolve contributions from sources that have unique chemical compositions. Unless a given source (facility) has a unique emissions chemical composition, it is not possible with receptor modeling to resolve the impact of an individual source. In addition, only relative impacts are produced by the model indicating the relative contribution from each source class. Typical source classes modeled include motor vehicle exhaust, fuel combustion, electric utilities, construction activities, marine aerosol, and fugitive dust of geological origin.

Receptor modeling in the past has been used primarily for examining sources of relatively stable pollutants such as particulate matter. For example, Chow et al. (1992) used the CMB to examine PM_{10} source apportionment in the San Joaquin Valley in California. They concluded that fugitive dust of geological origin (e.g., fugitive dust from agricultural tilling, roadways, and construction activity) exceeded 50 percent of the observed PM_{10} at Bakersfield in the summer and fall. Onroad motor vehicles contributed only approximately 10% of the observed PM_{10} while 4 percent of the PM_{10} mass was unexplained.

Much recent work has been performed to apply receptor modeling to reactive pollutants. For example, Scheff et al. (1995) applied the CMB to evaluation of emissions of non-methane organic compounds for the Southeast Michigan Ozone Study (SEMOS). Sheff selected chemical species with relatively low reactivity for use in the modeling, and selected his modeling scenarios to reflect relatively small travel times. Using this methodology, Sheff et al. were able to overcome the issue of reactivity and loss of emitted mass times and thus limited the potential influence of atmospheric photochemistry on the modeling results. They concluded that the relative proportion of the observed NMOC concentrations was consistent with current estimates of emissions for architectural coating and coke oven sources. However, there were significant differences between the CMB results and the current inventory on the relative proportion of emissions from refineries and graphic arts.

A limitation of the methodology is the resolution available from the CMB model is limited by the quality of the available source emission composition profiles. Because of the large variability between sources and potentially from test to test for a given source, these profiles are highly variable. Consequently, significant preparatory work must be performed to ensure that the most representative source profiles are used as input to the CMB model. In addition, the monitoring data used in CMB modeling must be representative of ambient concentrations in the area in which the emissions are occurring. These monitoring data cannot be primarily influenced by local sources for which source profiles are not available. Nor can chemical reactions or losses such as deposition or decay deplete the emitted substances prior to their measurement as ambient concentrations.

9.4 Inverse Air Quality Modeling

The inverse air quality modeling method relies on the use of a validated air quality dispersion model to back calculate what the emissions should be to produce the observed air pollutant concentrations. These general inverse methods are based upon an iterative application of a Kalman filter (Hartley and Prinn, 1993). They use a recursive, least squares approach to minimize the difference between the observed concentrations field and the modeled concentration field produced using adjusted

emissions. These methods are very powerful and can yield direct estimates of temporal and spatial emissions and their uncertainty, both overall and stratified by source type or location.

Chang et al. (1993; 1995) used the Kalman filter approach to examine the ability of this methodology to yield estimates of emission uncertainty. Work in progress is continuing the testing of this methodology and applying it to estimate uncertainty in isoprene emissions in the Atlanta area (Chang et al., 1995).

Mulholland and Seinfeld (1995) used the Kalman filter inverse modeling approach to examine potential errors in the CO inventory for the Los Angeles area. In this study, estimates of both temporal and spatial errors in the CO emissions inventory for Los Angeles were developed. Figure 3.9-1 presents the estimated overall temporal adjustment factor for air basin-wide CO emissions that resulted from the best fit between the modeled and observed CO concentrations. The weekday adjustment varies from approximately a factor of 3 during the day to considerably less than one at night. The weekend adjustment factor approaches a factor of ten. In addition to these temporal biases, they also identified potential spatial biases in the Los Angeles CO emissions inventory, including a postulated significant underestimate of recreational traffic along the coast on the weekend day modeled (Saturday).

There are inherent limitations in the use of air quality models for estimating emissions uncertainty. Air quality simulation models are imperfect representations of the real world in that many simplifying assumptions must be made during model development. These models, in their attempt to simulate very complex atmospheric behavior, introduce their own uncertainty to the emission estimates. These uncertainties can be due to inherent assumptions in the model, the numerical algorithms used to implement the assumptions, inadvertent biases introduced in the model formulation, and imprecise or erroneous input data. As a result, there are limitations to the accuracy of air quality models in use today.

Any use of the inverse air quality modeling methodology to estimate emissions uncertainty must account for the uncertainty introduced by the air quality model itself. For example, the EPA performance goal for the Urban Airshed Model is for the gross error of all observed-predicted pairs to be less than 30 to 35 percent and normalized bias to be within ±15 percent (U.S. EPA, 1991b). However, when the model is used to examine differences between modeling scenarios or incremental differences in model estimates due to small perturbations in model input (such as occurs in these general inverse methods), the model errors and biases can tend to cancel or drop out. Consequently, the uncertainty introduced into the estimates of emission uncertainty can be made small compared to the other sources of uncertainty.

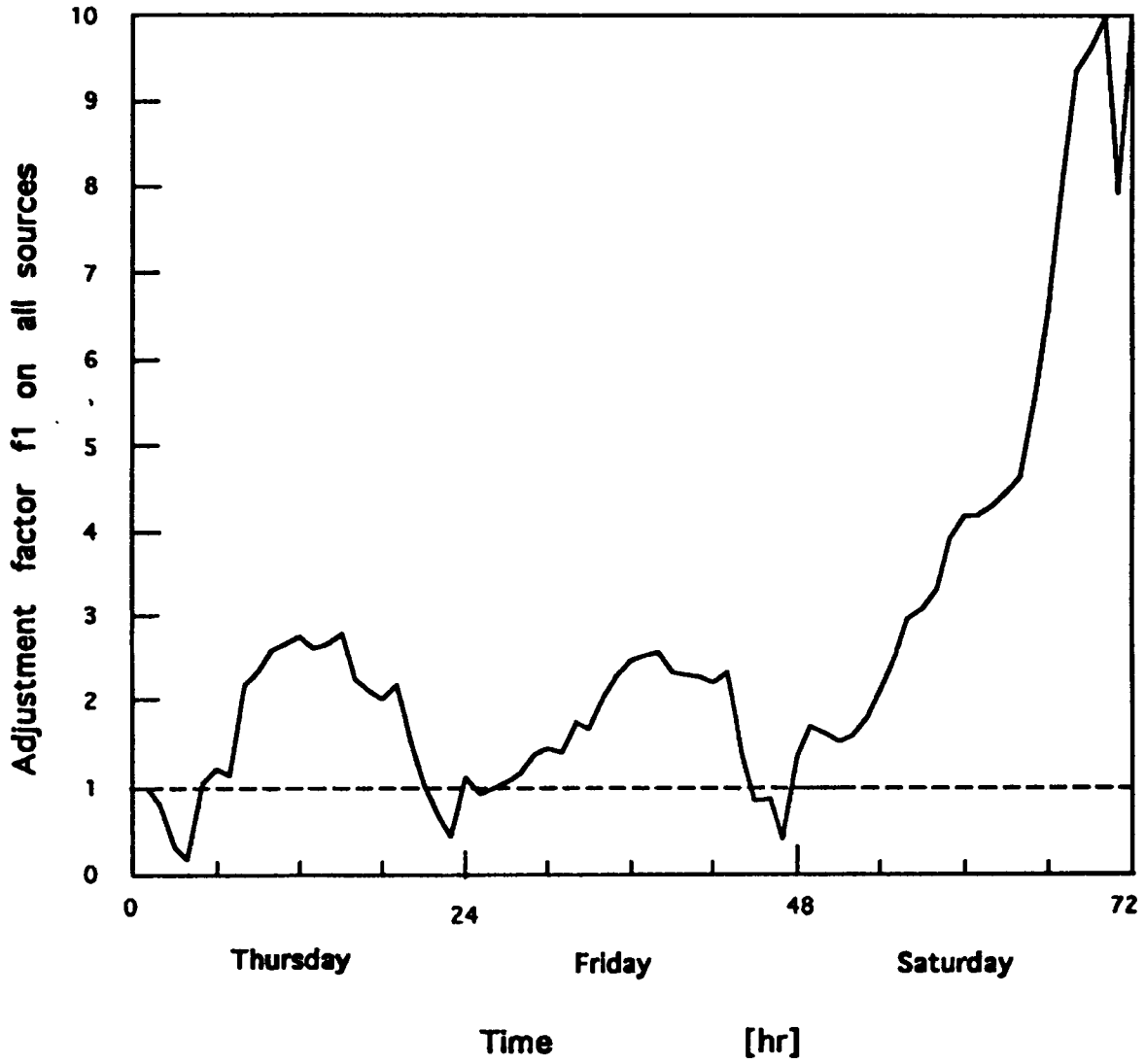

Fig. 5. Single adjustment factor for all CO sources: time-variation to achieve best CIT model fit to CO concentration observations.

FIGURE 3.9-1. SINGLE ADJUSTMENT FACTOR FOR ALL CO SOURCES

Reprinted from *Atmospheric Environment*, Vol. 29, M. Mulholland and J.H. Seinfeld, Inverse Air Pollution Modelling of Urban-Scale Carbon Monoxide Emissions, pp. 497-516, 1995, with kind permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington OX5 1GB, UK.

10

REFERENCES

Birth, T.L. 1995. *User's Guide to the Personal Computer Version of the Biogenic Emissions Inventory System (PC-BEIS2).* Prepared by Computer Sciences Corporation for U.S. Environmental Protection Agency, Office of Research and Development, EPA Interagency Agreement DW47936118-01. Washington, D.C.

Cardelino, C. 1995. "The Application of Land Use Patterns to Study the Impact of Temperature, Biogenic Emissions and Land Use Pattern Changes on Ozone Precursor Emissions and Ozone Concentrations." In: *Proceedings: Regional Photochemical Measurement and Modeling Studies. Volume 3*, VIP-48, Air and Waste Management Association, Pittsburgh, Pennsylvania.

Chameides, W.R., R. Lindsay, J. Richardson, and C. Kiang. 1988. The Role of Biogenic Hydrocarbons in Urban Photochemical Smog: Atlanta as a Case Study. *Science* 241:1473-1475.

Chang, M., C. Cardelino, W. Chameides, and W. Change. 1993. "An Iterative Procedure to Estimate Emission Inventory Uncertainties." Presented at the Regional Photochemical Measurement and Modeling Studies Meeting of the Air and Waste Management Association at San Diego, California. November.

Chang, M., C. Cardelino, D. Hartley, and W. Change. 1995. "Inverse Techniques for Developing Historical and Future Year Emission Inventories." Presented at the Emission Inventory: Programs and Progress Specialty Conference of the Air & Waste Management Association, Raleigh, North Carolina. October 11-13.

Chow, J.C., J.G. Watson, D.H. Lowenthal, P.A. Soloman, K.L. Magliano, S.D. Ziman, and L.W. 1992. PM_{10} Source Appointment in California's San Joaquin Valley, *Atmospheric Environment* 26A:3335-3354.

Claiborn, C., A. Mitra, G. Adams, L. Bamesberger, G. Allwine, R. Kantamaneni, B. Lamb and H. Westberg. 1995. Evaluation of PM_{10} Emission Rates from Paved and Unpaved Roads Using Tracer Techniques. *Atmospheric Environment* 29:1075-1089.

Fujita, E.M., B.E. Croes, C.L. Bennett, D.R. Lawson, F.W. Lurmann and H.H. Main. 1992. Comparison of Emission Inventory and Ambient Concentration Ratios of CO, NMOG, and NO_x in California's South Coast Air Basin. *J. Air Waste Manage. Assoc.* 42(3):264-276.

Gaudioso, D., M.C. Cirillo, C. Trozzi, and R. Vaccaro. 1994. "Uncertainty of NMVOC emission estimates from vegetation." In: *The Emission Inventory: Perception and Reality*. Proceedings of an International Specialty Conference, VIP-38. Air & Waste Management Association, Pittsburgh, Pennsylvania.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Von Nostrand Reinhold, New York.

Grant, E.L., and R.S. Leavenworth. 1988. *Statistical Quality Control*. McGraw-Hill Book Company, New York.

Harley, R.A., A.G. Russell, G.J. McRae, G.R. Cass, and J.H. Seinfeld. 1993. Photochemical Modeling of the Southern California Air Quality Study. *Environ. Sci. Technol.* 27:378-388.

Hartley, D., and R.J. Prinn. 1993. Feasibility of Determining Surface Emissions of Trace Gases Using an Inverse Method in a Three-Dimensional Chemical Transport Model. *J. Geophys. Res.* 98:5183-5197.

Janssen, M., and M. Koerber. 1995. "Developing a Future Year Regional Attainment Demonstration Inventory for Ozone." In: *The Emission Inventory: Programs and Progress*, Proceedings of a specialty conference, VIP-56. Air & Waste Management Association, Pittsburgh, Pennsylvania.

Lamb, B., D. Gay, H. Westberg, and T. Pierce. 1993. A Biogenic Hydrocarbon Emission Inventory for the U.S.A. Using a Simple Forest Canopy Model. *Atmospheric Environment* 27A(11):1673-1690.

Lawrence, R.J. 1988. "Applications in Economics and Business." In: *Lognormal Distributions: Theory and Applications*. E.L. Crow and K. Shmizu, editors. Marcel Dekker, Inc. New York and Basel. pp. 229-266.

Misenheimer, D.C. 1996. Personal communication from D.C. Misenheimer, U.S. EPA, to L.Y. Cooper, Eastern Research Group, Inc., regarding transmission of example auditing tool (Voyager) figure. June 6, 1996.

Mulholland, M., and J.H. Seinfeld. 1995. Inverse Air Pollution Modelling of Urban-Scale Carbon Monoxide Emissions. *Atmospheric Environment* 29:497-516.

National Acid Precipitation Assessment Program. 1991. *Acidic Deposition: State of Science and Technology; Volume I - Emissions, Atmospheric Processes, and Deposition.* P.M. Irving, editor, Government Printing Office, ISBN 0-16--0361144-3. Washington, D.C.

Neece, J.D., and J.H. Smith. 1994. "Automated Quality Assurance Of Point Source Emission Inventories For Urban Airshed Modeling." In: *The Emission Inventory: Perception and Reality.* Proceedings of an International Specialty Conference, VIP-38. Air & Waste Management Association, Pittsburgh, Pennsylvania.

Peer, R.L., D.L. Epperson, D.L. Campbell., and P. von Brook. 1992. *Development of an Empirical Model of Methane Emissions from Landfills.* U.S. Environmental Protection Agency, Air and Energy Engineering Laboratory, Research Triangle Park, North Carolina.

Pierce, T.E., B.K. Lamb, and A.R. Van Meter. 1990. "Development of a Biogenic Emissions Inventory System for Regional Scale Air Pollution Models." Paper 90-94-3, presented at the 83rd Annual Meeting of the Air & Waste Management Association, Pittsburgh, Pennsylvania, June 24-29.

Pierson, W.R., A.W. Gertler, and R.L. Bradow. 1990. Comparison of the SCAQS Tunnel Study with Other On-road Vehicle Emission Data. *J. Air Waste Manage. Assoc.* 40:1495-1504.

Scheff, P.A., R.A. Wadden, D.M. Kenski, and J. Chang. 1995. "Receptor Model Evaluation of the SEMOS Ambient NMOC Measurements" Paper 84-72.3. Presented at the 77th Annual Meeting of the Air Pollution Control Association, San Francisco, California, June.

Shapiro, S.S., and M.B. Wilk. 1965. An Analysis of Variance Test for Normality (complete samples). *Biometrika* 52:591-611.

Smith, A.E., J.T. Schanz, and P. Fuja. 1995. Air Quality Utility Information System (AQUIS) Emission Algorithms. Prepared by Argonne National Laboratory for USAF Headquarters Air Force Material Command, Wright-Patterson Air Force Base, Environmental Assessment Division. Dayton, Ohio.

Sokal, R.R., and F.J. Rohlf. 1969. *Biometry.* W.H. Freeman and Company, San Francisco, California.

South Coast Air Quality Management District. 1994. 1994 Air Quality Management Plan, Final Technical Report V-B, Ozone Modeling - Performance Evaluation. Diamond Bar, California.

Spellicy, F.L., J.A. Draves, W.L. Crow, W.F Herget, and W.F. Buchholtz. 1992. "A Demonstration of Optical Remote Sensing in a Petrochemical Environment." Presented at the Air & Waste Management Association Specialty Conference on Remote Sensing, Houston, Texas. 6-9 April.

U.S. Air Force. 1994. Air Quality Utility Information System (AQUIS) User's Guide and System Information. Prepared by Argonne National Laboratory for USAF Headquarters Air Force Material Command, Wright-Patterson Air Force Base, Environmental Assessment Division. Dayton, Ohio.

U.S. EPA. 1988. *Guidance for the Preparation of Quality Assurance Plans for O_3/CO SIP Emission Inventories*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-88-023. Research Triangle Park, North Carolina.

U.S. EPA. 1989. *Quality Assurance Program for Post-1987 Ozone and Carbon Monoxide State Implementation Plan Emission Inventories*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-89-004. Research Triangle Park, North Carolina.

U.S. EPA. 1991a. *Quality Review Guidelines for 1990 Base Year Emission Inventories*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-91-022. Research Triangle Park, North Carolina.

U.S. EPA. 1991b. *Guideline for Regulatory Application of the Urban Airshed Model*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-450/4-91-013. Research Triangle Park, North Carolina.

Watson, J.G., J.A. Cooper, and J.J. Huntzicker. 1984. The Effective Variance Weighing for Least Squares Calculations Applied to the Mass Balance Receptor Model. *Atmospheric Environment* 18:1347-1355.

Woodall, W.H., and B.M. Adams. 1990. "Statistical Process Control." In: *Handbook of Statistical Methods for Engineers and Scientists*. H.M. Wadsworth, editor. McGraw-Hill Publishing Company, New York. pp. 7.1 to 7.28.

APPENDIX A
EXAMPLE AUDIT REPORT

This page is intentionally left blank.

EXAMPLE AUDIT REPORT

The audit report includes a summary of the findings from assessing the data, procedures, and results from the interviews held with the inventory project team members. Recommendations for corrective actions are discussed immediately after the audit; however, a formal report which summarizes the auditor's activities and findings should be distributed within two weeks. The example shows the format and information that could be included in a technical systems audit report which summarizes the findings from a site visit and QA training session.

This page is intentionally left blank.

AUDIT REPORT

TO: Project Team

FROM: Auditor's name

DATE:

SUBJECT: **Technical Systems Audit and Quality Assurance Training of State Air Agency's Ozone and Carbon Monoxide Point Sources Emissions Inventory Development Program**

1.0 INTRODUCTION/BACKGROUND

Under the Clean Air Act Amendments (CAAA) of 1990, state and local air pollution agencies are required to inventory emissions contributing to National Ambient Air Quality Standards nonattainment, including ozone and carbon monoxide (CO), using 1990 as the base year. Because the data will serve as the basis for State Implementation Plans (SIPs), the inventories must be accurate and complete.

On *date*, a technical systems audit and quality assurance (QA) training session were conducted by the Project Quality Assurance Coordinator (QAC) at the state air agency office. The audit and training are part of the QA program designed to help produce an accurate and complete point sources inventory.

The state air agency works with contract personnel on the development of the point sources inventory. The state collects data from permitted sources, assesses its accuracy, and determines seasonal adjustments of total emissions. Contractors perform the same tasks for nonpermitted sources. The data collected by the state air agency are then entered into the State Implementation Plan Air Pollution Inventory Management System (SAMS) database for range checking and emissions calculations. The emissions

data from both the contractor and the agency are combined to yield a final point sources emissions inventory report.

A draft of the QA Plan for the state was discussed with the State Emissions Inventory Development Manager and approved a week prior to the audit. Because the QA Plan was not approved and distributed to the inventory development team prior to commencing the inventory activities, the auditor did not expect to find complete compliance with the approved quality control (QC) procedures. However, the QC and documentation procedures in use at the time of the audit were assessed and compared to the QA requirements established by the United States Environmental Protection Agency (EPA) for emissions inventory development work. The ultimate goals of the QA/QC program developed for emissions inventory development are data accuracy, procedural consistency, and good documentation of the data and all inventory development activities. When the potential for problems or deficiencies in the QC program were found, recommendations were made by the auditor for improvements.

2.0 AUDIT PROCEDURE

The technical systems audit was the first of two audits to be conducted in the state. The objective of this audit was to review and assess the effectiveness of the QC procedures established and implemented in the following areas:

- Data collection;
- Data analysis;
- Data documentation;
- Data management;
- Personnel training; and
- Senior technical supervision.

All survey forms received from permitted facilities are forwarded to the State Emissions Inventory Development Manager and maintained in a master data file. The survey forms are later forwarded to the inventory staff members. Each person prioritizes the data and focuses more attention on data from those facilities known to be high emitters. The information recorded on the survey forms are assessed for completeness and reasonableness. Calls are made to the submitter, if needed, to request or clarify data prior to making seasonal adjustments and totaling emissions results.

During the audit, the auditor met with each person involved in permitted sources inventory development and asked them to describe the procedures followed after data are forwarded from the State Emissions Inventory Development Manager. Some personnel were asked to review, analyze, and enter data into the SAMS database. While this was being done, the auditor assessed each person's experience using the database and ease in assessing the information recorded on the forms. Data documentation procedures, data management procedures, and use of senior technical resources were also evaluated. The findings from these individual assessments and the recommendations to improve the QC procedures are presented in the next section of this report.

3.0 AUDIT FINDINGS

The auditor confirmed the use of sufficient adequately trained personnel and the presence of sufficient senior technical supervision to develop an accurate point sources inventory. However, additional trained personnel would help the staff meet the tight deadline imposed by EPA and provide a more objective validation of the data.

Data documentation procedures could be improved to facilitate referencing data obtained via telephone or added/corrected as a result of engineering judgement.

The improvement in data documentation would also facilitate reconstruction of inventory development activities and thus provide a means to more thoroughly assess data quality and the accuracy of the inventory.

The data produced by each inventory staff member are not peer reviewed. According to the State Emissions Inventory Development Manager, limited time and personnel resources will not allow this objective validation of the data; therefore, more effort is being placed on technical over-sight during inventory development.

Although the audit findings do not suggest major deficiencies in the QC program, recommendations for improvement were made to further verify the accuracy of the results and integrity of the data.

4.0 RECOMMENDATIONS TO IMPROVE THE QUALITY CONTROL PROGRAM

As a result of the audit findings, recommendations are being made to improve the QC program. These recommendations will be discussed with the State Emissions Inventory Development Manager and the EPA Project Officer prior to implementing any revised procedures. Upon agreement, the procedural changes needed to provide the quality of data and level of documentation expected by EPA will be implemented by the staff and verified by the auditor during the next audit.

The auditor's recommendations are as follows:

4.1 Develop and standardize a procedure for rounding numbers in order to provide consistency in calculating and reporting results

During the audit, the auditor evaluated the calculation of seasonal data and other derivations by inventory development team members and found inconsistencies in the rounding of numbers. Although these inconsistencies will not grossly impact the accuracy of the final emission results, a consistent approach to rounding numbers is recommended. Clarity and consistency are data quality goals that are encouraged by EPA.

4.2 Develop and implement standard procedures for documenting and correcting data on the survey forms and in notebooks

The inventory staff members are documenting data very well; however, a set of guidelines for this documentation would improve data retrievals and help reconstruct inventory development activities. The following guidelines should be considered:

- **Document data in black ink because it photocopies well.** Several colors of ink and pencil are being used to record data in notebooks and on the survey forms. Data could be inadvertently lost if the raw data are copied because some ink colors and pencil entries do not reproduce well.

- **Distinguish between data recorded by the submitter and data recorded by the inventory staff members.** Data recorded on survey forms are not always initialed and dated when recorded by inventory staff members on the survey forms. It will be difficult to reconstruct inventory development activities if one cannot determine the source of the data entries. The validity of the data may also be questioned if the source of the information is unclear.

- **Do not obscure original entries when correcting data so that all inventory development activities can be reconstructed.** Some data

corrections are made by using correction fluid or erasing pencil entries. This practice may compromise the integrity of the data. Being able to determine the original entry, the reason for the correction, and the identity of the person revising the data enhances the validity of the correction. Corrections should be made by drawing a single line through the data and entering the correction next to, above, or below the original entry. If the data on a form includes entries by more than one person, corrections should be initialed and dated.

- **Assign book and page numbers to the notebooks used to record information received by telephone and information regarding other critical inventory development activities.** Data are currently documented in spiral stenographer notebooks. Because the books are not paginated and assigned unique identification numbers, it is difficult to reference the source of information when data are transcribed to the survey forms. Book and page numbers can be used as references on the survey forms when there is a need to verify data sources or provide clarifying comments.

- **For consistency, develop a standard procedure that describes the data to be recorded in the notebook and on the survey forms.** The type of information recorded in the notebooks and on the survey forms varies between the inventory staff members. If the type of data to be recorded in the notebooks and on the survey forms are clarified and standardized for all inventory staff members, it would be easier to validate the accuracy of the data that are finally entered into the database to calculate the emissions.

- **When information is received via telephone, document the name of the contact person and date the data are received.** Information received by telephone is often recorded on the survey form or in notebooks without providing the name of the person contacted or the date the information was received. To allow verification of the data, the source must always be provided. The date allows the reviewer to reconstruct the activities associated with data collection and analysis. The ability to reconstruct these activities will provide a means of further assessing data quality and accuracy.

4.3 Require the establishment of a file folder for each survey form as it is forwarded to the staff member and the maintenance of the survey form in the folder as the data are analyzed to help avoid data loss

Inventory staff members have devised different procedures for maintaining survey forms received from the State Emissions Inventory Development Manager. The organization of the data and handling practices ultimately determines how complete the master file will be after the inventory is developed. Forms kept loosely on a desk or in desk drawers could be easily misplaced. The maintenance of the forms in folders with the identification of the facility clearly marked will help improve data management procedures and help prevent data loss.

4.4 Document all inventory development training

Although the documentation of training may not be a state requirement, training records further qualify the data by verifying that the personnel developing the inventory are adequately trained to perform the duties assigned. According to the State Emissions Inventory Development Manager, training was provided by having less experienced persons work with a more experienced person until their performance was at a level that did not require direct supervision. Because this training further qualifies a staff member to make technical judgements, it should be verified in writing. The documentation should include the dated signatures of the trainer and trainee, a description of the training, training dates, and the date the trainee was considered to be performing at a level that allowed him/her to independently assume inventory development responsibilities.

5.0 DISCUSSION

A discussion of the findings related to data collection, data analysis, data documentation, data management, personnel training, and senior technical supervision are presented in this section of the audit report. As mentioned earlier, the audit findings and recommendations to improve the QC program were discussed with the State Emissions Inventory Development Manager immediately after the audit.

5.1 Data Collection

The survey forms are forwarded to the State Emissions Inventory Development Manager and the data receipt information is logged into a database. After the receipt of the data is documented, the data are assigned to personnel responsible for the geographical area within which the facility forwarding the survey form is located.

The database provides adequate tracking of the completeness of the responses from the permitted sources. The database can also be used to assess the completeness of the master data file. No recommendations for improving the existing data collection procedures were made.

5.2 Data Analysis

After receiving the data, each staff member assesses the completeness of the form and reviews the information provided for reasonableness. Seasonal adjustments are made and total emissions are manually calculated. The units are checked for consistency and compatibility with those used in the SAMS database. Calculated results are added to the survey form and the information required to calculate emissions is entered into the SAMS database. If the results are out of range, they are not accepted

by the system. Data appearing to be unreasonable are double-checked by calling the facility contact person. No recommendations were made to improve the data analysis procedures.

5.3 Data Documentation

As data are manually calculated or received via conversations with the submitter, inventory staff members record the information directly on the survey forms or in a spiral stenographer notebook. The notebooks are not uniquely numbered or paginated. The information recorded in the notebook or on the survey forms varied between the inventory staff members. Some inventory staff members recorded the information received by telephone in the notebook and then transferred it to the survey forms. Others recorded the information directly on the survey forms. Others only recorded information from telephone conversations held with permitted source personnel in the notebook.

Additional data documentation concerns that led to recommendations for improvement included:

- Not always including the source, date, and initials of the person adding data to the survey forms;
- Using correction fluid; and
- Not using black indelible ink to record data.

Recommendations were made to standardize the data documentation procedures so that data are easy to review and calculated results are easily verified. Recommendations were made to help validate the accuracy and integrity of the data used to calculate emissions results and facilitate the reconstruction of emissions

development activities. The auditor also recommended the use of bound rather than spiral notebooks for future work to help avoid data loss.

5.4 Data Management

Data are managed by each staff member in his/her designated office area until they are filed in the master data file. Most of the survey forms are kept loosely in a desk drawer or mail tray. In order to help eliminate the possibility of the data being misplaced or lost and improve the completeness of the master data file, the auditor made recommendations to improve data handling and management procedures.

5.5 Personnel Training

All personnel appeared to be adequately trained to perform the tasks reviewed during the audit. The technical accuracy of engineering judgement and data analysis were not assessed; however, familiarity with the database, priority of the data, and use of EPA guidance material were verified to be acceptable. All personnel appeared to be readily able to use the database and move between screens when explaining how data entries were made.

The State Emissions Inventory Development Manager commented that the training conducted on his newest employee was not documented; therefore, a recommendation to document this training in the personnel files was made. Each staff member was very confident about his/her abilities to perform the tasks assigned.

5.6 Senior Technical Supervision/Peer Review

All staff members spoke very highly of the accessibility of senior technical advice from the State Emissions Inventory Development Manager. When there are questions about the acceptability of the data, the manager is asked to make the final decisions. No recommendations to improve the system were made because the staff members and the auditor agreed that adequate supervision is provided.

6.0 QUALITY ASSURANCE TRAINING

The QA training provided by the QAC included an overview of basic QC principles related to the inventory development process and a discussion of the sections of the project QA Plan that are applicable to point sources inventory development. The bulk of the presentation was taken from material developed by the auditor along with EPA's Emissions Inventory Branch, Technical Division. An EPA representative presented this material at an inventory development training session held last year. An outline of the presentation is included as Attachment A. The outline and over-heads were distributed during the training session.

The training session was attended by the State Emissions Inventory Development Manager and other inventory staff members. Based on the comments made by the staff members during the audit, the information provided was helpful, although it would have been more useful had it been presented during the planning phase of the work. Interest was also expressed in more EPA training on emissions inventory and QA program development.

This page is intentionally left blank.

ATTACHMENT A

This page is intentionally left blank.

OUTLINE: STATE AIR AGENCY SIP EMISSIONS INVENTORY DEVELOPMENT QUALITY ASSURANCE TRAINING SESSION

Presenter: QAC

I. Introduction
- A. Objective of training session and systems audit
- B. Purpose of QA Plan and encouragement to implement it

II. Definition of key terms
- A. Quality assurance and quality control
- B. QA Plan
- C. QA Coordinator
- D. Systems audit

III. Purpose of Emissions Inventory QA Program
- A. Reduce Errors
- B. Maximize consistency in inventory preparation
- C. Improve data documentation, accuracy, and completeness
- D. Foster confidence in data
- E. Facilitate inventory review process
- F. Decrease inventory development cost

IV. Technical Considerations
- A. Planning
  1. Resources for QA
  2. Personnel training and project staffing
  3. Standardizing inventory development procedures
  4. Prioritizing data
  5. Development of data validation procedures
- B. Data Collection and Analysis
  1. Reputable data source
  2. Documentation of data reduction and calculations
  3. Senior Technical Review
  4. Completeness assessments
- C. Data Handling
  1. Logging upon receipt
  2. Maintenance of a master data file
  3. Coding and tracing data to monitor completeness and facilitate retrievals
  4. Data corrections

D. Reporting
E. Internal and External (EPA) Systems Audits

V. Previous Problems with Inventories
A. Double counting of point sources
B. Poor data organization
C. Incomplete inventory

VI. Successful Inventory Development Procedures
A. Use of standard forms
B. Documentation of validation procedures
C. Use of the EPA guidance to develop and implement a QA/QC program
D. Independence of the QA Coordinator from inventory development activities

APPENDIX B
TECHNICAL SYSTEMS AUDIT CHECKLIST EXAMPLE

This page is intentionally left blank.

TECHNICAL SYSTEMS AUDIT CHECKLIST EXAMPLE

The Area and Non-mobile Inventory Quality Assurance Checklist provides guidelines for a thorough evaluation of the inventory development program implemented by a contractor and state agency. It includes assessments of the following critical phases of the work:

- Planning/Management;
- Inventory Development;
- Documentation/Data Entry; and
- Reporting.

This page is intentionally left blank.

NYSDEC AREA AND NON-MOBILE EMISSIONS INVENTORY QUALITY ASSURANCE CHECKLIST

Auditor _____ Date _____

Personnel Interviewed _____

_____

This checklist is to be used to document the findings from the audit of activities and data associated with area and non-road mobile emissions inventory development. Use the applicable parts of the checklist to identify the quality concerns associated with each task. Document the results and use them to generate the audit report.

	Yes	No
Planning/Management 1. Work Plan a. Was the work plan approved prior to commencing work?		
b. Was the work plan revised?		
c. If yes, were the revisions documented and approved?		
2. Are the resources required for the work adequate?		
3. Were the staff adequately trained?		
4. Was an adequate number of project personnel available to perform the tasks assigned?		
5. Are project meetings held two times a week or at an appropriate interval to keep project personnel abreast of the status of the work?		
6. Are NYSDEC adn EPA routinely kept abreast of inventory development milestones and problems?		
7. Senior Technical Review a. Were the results from senior technical peer reviews documented?		
b. Were problems identified?		

	Yes	No
c. If yes, were the problems documented and resolved to the satisfaction of the reviewer and project personnel involved?		
d. If no, attach a description of the problems and how they will be or were resolved?		
8. Was conformance to EPA guidelines checked by completing applicable sections of the EPA *Quality Review Guidelines* checklist?		
Area and Non-Road Mobile Source Inventory 1. Were 1990 estimates made for the following: a. VOC sources emitting <10 tons per year (TPY) using *Procedures for the Preparation of Emission Inventories for Carbon Monoxide and Precursors of Ozone, Volume I: General Guidance for Stationary Sources* (EPA-450/4-91-016)?		
b. Stationary sources of nitrogen oxides (NO_x) that emit <25 TPY using the above document?		
c. Stationary sources of CO that emit <100 TPY using the above document?		
d. Other sources of VOCs, NO_x, and CO by county or equivalent city emission areas and non-attainment designations using demographic, 1990 census, and industrial and commercial data with appropriate AP-42 factors or other valid sources?		
e. Non-road sources of VOCs, NO_x, and CO using *Volume IV of the Procedures for Emission Inventory Preparation*?		
2. Were area source estimates adjusted to account for point sources to avoid double-counting?		
3. Were estimates provided for VOCs, NO_x, and CO for non-attainment areas?		
4. Was the approach used to determine the completeness of the determination of all non-attainment areas described?		
5. Were EPA guidance documents used and referenced for area and non-road estimates?		

	Yes	No
6. Were the references used to determine non-road mobile and area sources given?		
7. Are 1990 emissions for other non-road engine and vehicle categories extrapolated to other NY non-attainment areas using population as recommended by EPA?		
8. Are the data for all of the counties applicable to this work included? (See work plan and revisions for list of counties.)		
9. Were all estimates reported as "0" explained?		
10. Were estimates made from biogenic and surface impoundment data prepared by NYSDEC reviewed prior to use?		
11. Were pre-processed (LOTUS spreadsheets) data reviewed prior to entry into AIRS-AMS?		
12. Database a. Are electronic data and hard copies compared to determine if they agree?		
b. Are all defaults clearly explained?		
13. Was the Inventory Preparation Plan developed by NYSDEC followed?		
a. If yes, is an approved copy available in the project file?		
14. Reporting a. Was a comprehensive draft report describing the development of the area sources inventory written and submitted to EPA by 9/30/92?		
b. Was a comprehensive final report be submitted after completion of the point source inventory (10/15/92)?		
c. Is EPA kept abreast of the need to revise these deadlines?		

	Yes	No
On-Road Mobile Sources 1. Were procedural reviews of the methodology used by NYSDEC to determine vehicle miles traveled data (VMT) and MOBILE4.1 inputs conducted by Radian?		
a. Were the results from the review documented?		
b. If problems were found, were they resolved to the satisfaction of NYSDEC and Radian?		
2. Were MOBILE4.1 input files checked for accuracy and reasonableness prior to running models?		
3. Were representative numbers of the VMT and MOBILE4.1 emission factors checked?		
4. Was a representative number of the data entered into AIRS-AMS checked for accuracy and completeness?		
Documentation/Data Entry 1. Are the data documented in a manner that will allow reconstruction of project activities? (notebooks, contact reports, memos, notes to the file)		
2. Was the receipt of data from NYSDEC and others documented and the data controlled to provide a complete master file?		
3. Are data sources and guidance documents identified?		
4. Are telephone calls documented in project dedicated notebooks or on telephone contact report forms?		
5. Were data entry worksheets used and reviewed prior to input into AIRS-AMS?		
a. Was this review adequate to assess the accuracy and completeness of the data?		
b. Were the results from the reviews documented?		
6. Were spreadsheet estimates compared to AIRS-AMS estimates to assess the reasonableness and accuracy of the database calculations?		

	Yes	No
Reporting 1. Was each report formatted as required in the EPA guidance document for emissions inventory reports?		
2. Was a report describing the development and implementation of the QA program written?		
3. Were the procedures for determining the completeness and accuracy of the database discussed in the inventory development report?		
4. Were the final electronic files compared to the hardcopies prior to delivery to NYSDEC and EPA to ensure that they both agree?		
5. Did the reports meet the 9/30/92 deadline?		

Comments:

This page is intentionally left blank.

APPENDIX C
DATA QUALITY AUDIT CHECKLIST
(EXAMPLE 1)

This page is intentionally left blank.

DATA QUALITY AUDIT CHECKLIST (EXAMPLE 1)

The Chicago Inventory QA Checklist was used to assess the quality of the planning documents, adequacy of the data gathering procedures, and thoroughness of the technical review process during the development of a toxics inventory from an existing inventory. This was a Level 4 inventory (see Chapter 2, Section 1 for explanation of inventory categories), and the objective was to develop a preliminary estimate of HAP emissions based on existing inventories. Determining the validity of the existing data would be critical to the accuracy and completeness of this inventory.

This page is intentionally left blank.

CHICAGO INVENTORY QA CHECKLIST

1.0 GENERAL PROCEDURES

1.1 Written Instructions, Plans

1.1.1 Were project instructions prepared and distributed to all team members? _____

Date: _____

1.1.2 Was an Inventory Preparation Plan prepared? _____

Reviewed, and if necessary, revised? _____

Distributed to team members? _____

1.2 Data Gathering

1.2.1 Was project file set up? _____

Date: _____

1.2.2 Are separate files maintained for each of the following source categories?

Storage, Transportation, and Marketing of VOL	_____	Consumer/Commercial Solvent Use	_____
VOL Transfer	_____	Agricultural Pesticide Application	_____
Barge and Tanker Cleaning	_____	Architectural Surface Coatings	_____
Service Station Unloading	_____	Automobile Refinishing	_____
Vehicle Refueling	_____	Traffic/Maintenance Coatings	_____
Gasoline Tanker Truck Leaks	_____	Landfills	_____
Underground Tank Losses	_____	TSDFs	_____
Degreasing	_____	POTWs	_____
Dry Cleaning	_____	Incinerators	_____
Graphic Arts	_____	Structure Fires	_____
Asphalt Paving	_____		

1.2.3 Are sufficient document/access controls in place? _____
If so, what are they? _____

_____

1.2.4 Are copies of completed calculation sheets placed in the appropriate files? _____

_____

2.0 TECHNICAL REVIEW

2.1 General

2.1.1 Were "Source Category Review" forms completed (QC) for all categories? _____

If not, list those not completed. _____

_____

_____

2.1.2 Are project files complete? _____

_____

_____

2.2 Individual Source Category Reviews

2.2.1 Storage, Transportation, and Marketing of Volatile Organic Liquids (VOL)

- Were emissions for the following categories completed?

_____ Barge and Tanker Cleaning

_____ Service Station Loading (Stage I)

_____ Vehicle Refueling

_____ Gasoline Tank Truck Leaks

_____ UST Breathing

- Were VOC emission factors adjusted for temperature and RVP before applying speciation profiles? _____

  _____

- Does QC review for this category include assessment of reasonableness of estimates? _____

  _____

2.2.2 Other Solvent Use

- Are assumptions used for HAP cleaning machine solvent use clearly stated and referenced? _____

  _____

- Were graphic arts estimates adjusted by subtracting emissions from point sources < 100 tons? _____

  _____

- Were asphalt paving emissions estimated using IL DOT data and assuming 95% diluent evaporation from rapid cure cutbacks, 70% from medium cure cutbacks? _____

  _____

- Was the consumer products emission estimate adjusted to include nonreactive VOC HAPs (like methylene chloride)?__

  _____

  Was it done correctly? _____

- Did agricultural emissions of captan, carbaryl, lindane, and trifluoralin correctly address seasonality (i.e., confined to growing season) and application rate (i.e., accounted for frequency and amount applied)? _____

  _____

- Verify the source of VOC emissions and speciation profiles for each of the following:

	Proposed VOC Emission Source	✓ if used, or show source	Proposed Speciation Profile	✓ if used, or show source
Architectural Surface Coatings	IL SIP		CARB (1991)	
Automobile Refinishing	IL SIP		CARB (1991)	
Traffic Markings	IL SIP		CARB (1992)	
Bridge Painting	IL SIP		CARB (1992)	

Explain any deviations from plan:

_____

_____

_____

- Was information from the Methane Recovery Yearbook (1991) used to adjust uncontrolled landfill VOC emissions? _____ If not, how was the adjustment made? _____

_____

- For TSDF emissions, what was the source of growth factor used to adjust emissions? _____

  Is it reasonable? _____

  Was sufficient justification for its use given? _____

- The IPP methodology for POTW emissions estimation proposed use of emission factors from California. Was any information obtained and documented that shows the use of these factors is reasonable for Chicago? _____

_____

- Were the assumptions, methods, and data sources used to develop structure fires emission factors thoroughly described and referenced? _____

  Are they reasonable? _____

This page is intentionally left blank.

APPENDIX D
DATA QUALITY AUDIT CHECKLIST
(EXAMPLE 2)

This page is intentionally left blank.

DATA QUALITY AUDIT CHECKLIST (EXAMPLE 2)

The Continuous Emissions Monitoring Systems Audit Checklist provides an example of the type of questions that could be asked when evaluating the collection of measurement data that are used to develop a facility's point source inventory. The critical phases of inventory development of concern during this audit are management/planning, availability of resources, instrument testing, data acquisition, and the implementation of QC measures to ensure the accuracy of the measurement data.

This page is intentionally left blank.

CONTINUOUS EMISSION MONITORING SYSTEMS AUDIT CHECKLIST

Contract: _____ Date: _____

Site: _____ Auditor: _____

Operation	Yes	No	Comments
General Check Points 1. Qualified personnel?			
2. QAPP or work plan on site? Revision # ___			
3. Spare parts and support equipment available?			
4. Instrumentation and apparatus maintained in good condition?			
5. Adequate facilities?			
6. Instrument certification and/or calibration documentation available?			
7. Sample line(s) properly located to obtain representative sample?			
8. Instrument logbook and maintenance record properly maintained (up-to-date, entries dated and initialed)?			
Calibration Procedures 1. Calibration frequency appropriate?			
2. Multipoint calibration/linearity check performed regularly?			
3. Zero point included in calibration?			

Operation	Yes	No	Comments
4. Calibration gases of acceptable quality, analyzed within last 12 months?			
5. Adequate supply of calibration gases?			
6. Calibration gas mixture appropriate for atmosphere being sampled?			
7. Appropriate record keeping procedures used for calibration documentation?			
8. Instrument range (span) used appropriate for measured concentrations?			
9. Calibration gas introduced so as to check entire sampling interface (sample lines, conditioning system, etc.)?			
Data Acquisition and Recording System 1. Strip chart recorder used?			
2. Data logging system used?			
3. Appropriate sensitivity?			
4. Instrument output signal checked regularly at zero and span values?			
5. Appropriate data averaging procedure?			
6. Appropriate zero offset (10%)?			
7. All pertinent data recorded (e.g., date, time, etc.)?			
8. Hard copy data clearly identified and properly stored?			

Operation	Yes	No	Comments
Quality Control Procedures 1. Calibration frequency specified?			
2. Control standard used to document day-to-day variability (precision)?			
3. Specific acceptance criteria specified for calibrations and control sample analyses?			
4. Control sample introduced so as to check entire sampling interface?			
5. Drift checks performed?			
6. Control charts and/or data summary for control sample analyses?			

Instrumentation

Parameter	Manufacturer	Model	Serial Number	Full Scale Range
Data System				

Comments: _____

_____

_____

_____

_____

APPENDIX E
PERFORMANCE EVALUATION AUDIT CHECKLIST EXAMPLE

This page is intentionally left blank.

PERFORMANCE EVALUATION AUDIT CHECKLIST EXAMPLE

The Computer Audit Checklist provides guidelines for evaluating the adequacy of a computer that could be used to calculate emissions, store inventory data, and generate tables of results for the final inventory report. The testing, calibrating, and evaluation of operating procedures of instruments are all important for assuring the quality of data; a computer is also an instrument and should be treated as such. Because of the dependence of the inventory staff on computers to complete the inventory, an assessment of the testing, operation, and use of the computer would be one of the most important audits conducted during the planning and development phases of the inventory.

This page is intentionally left blank.

AUDIT CHECKLIST
COMPUTER DATA ACCURACY

Procedure: Select a date with results passing and not passing quality acceptance criteria, if possible. Determine if reported results accurately reflect the raw data. Use an alternate method to verify the accuracy of computer manipulations to yield calculated results.

	Yes	No	NA
1. Identify data maintained only on the hard drive _____ _____ _____ _____			
2. Locate and identify data maintained on hard copies _____ _____ _____ _____			
3. Are quality control data referenced to sample results? (standards, blanks, calibrations, replicates, duplicates, spikes, instrument conditions, surrogates, internal standards, etc.) - LAB DATA _____ _____ _____			
a. Are the references to quality control data protected or can they be easily changed?			

	Yes	No	NA
b. Are the references sufficient to associate quality data with individual sample results?			
4. Are data outside of acceptance criteria flagged?			
a. If yes, are actions taken described?			
5. Are the detection limits for target analytes clearly defined or referenced in the raw data? (LAB)			
a. Are they accurately represented in the report?			
b. Were changes made to reflect dilutions?			
6. Are the data reduction steps definined in a written procedure?			
a. Is the procedure sufficient to recalculate results?			
b. Are the data sufficient to verify results?			
c. Were errors found in the data verified?			
If Yes, explain and identify results: _____ _____ _____ _____ (use additional comment sheet if necessary)			
7. Are the data accurately reported?			
a. Are the units accurate?			
b. Are the data accurately flagged to note bias or confidence in the results reported?			
8. Can results be traced back to the original data associated with a specific test article or data set?			
If not, describe problem: _____ _____ _____ _____			

AUDIT CHECKLIST
COMPUTER PROGRAM DOCUMENTATION AND VALIDATION

	Yes	No	NA
1. Are the following documents present a. Software management plan			
b. Software development plan and results			
c. Software test and acceptance plan			
d. Software User's operations document			
e. Software maintenance document			
f. Assessment of hardware			
2. Are the following documented: a. Program			
b. Table of definitions			
c. System size and timing requirements			
d. Definitions of subsystems			
e. Requirements for: -- Hardware			
-- Electricity			
-- Security			
f. Backup and disaster recovery procedures			
g. Quality requirements for: -- Reliability			
-- Maintainability			
-- Flexibility for expansion			
h. Testing procedures			
3. Does software management include the following: a. Independent validation			

	Yes	No	NA
b. Definitions/identifications of interfaces			
c. Definition of software tools -- Identification of program language			
-- Identification of network software requirements			
d. Configuration control (Control, release, and storage of master copies)			
4. Is there a flow chart or text showing functional flow?			
5. Are input and output fields identified?			
6. Are there written procedures for software revisions?			
a. Are revisions tested to determine how the entire program is affected? _____			
b. How are revisions implemented? _____			
c. Are revisions documented? _____			
7. Does the User's Guide or software description include: a. Description of system (hardware)			
b. Identification of person to contact when problems occur			
c. How to access the system			
d. How to input data			
e. How to generate reports			
f. How to update data			
g. Description of error codes			
h. Procedures to follow if the system goes down			

	Yes	No	NA
8. Do testing procedures include the following: a. Description of test procedures to perform			
b. Expected outcome			
c. Documentation of results			
d. Recommendation for handling problems			
9. Has security be addressed (Statements or passwords to safeguard accuracy of computer program operation)?			
10. Can someone knowledgeable of the program language understand from reading the program what it does to the data to yield the expected or desired results (if not, explain on separate page)			

AUDIT CHECKLIST
COMPUTER PROGRAM OPERATION

	Yes	No	NA
1. Is a password required to access the system?			
2. Operator Training a. Are operators adequately trained?			
b. Is training documented?			
3. When testing the system, are there system delays that hamper the job?			
4. Are error messages given if an entry error is made (ex. data out-of-range)?			
5. Does the system prevent entry of data whic are out-of-range?			
6. Are there prompts if fields are missed that are required for data manipulations?			
7. How are default parameters assigned by the system? _____ _____ _____			
8. Does the system carry over data from one screen to the next in order to minimize entry errors?			
9. Are data entered into the central database via a computer readable media? If yes, does the data include information concerning:			
a. The source of the data			
b. When the data were collected			
c. Conditions under which collected			
d. Pointers to link data to quality control data			

	Yes	No	NA
e. Quality control (QC) flags indicating the level of data acceptability			
10. If the data are entered by prompting the system to access a previously existing data file, are the data validated by a. Comparison of #/size of files transferred			
b. A log which documents files transferred			
c. Documenting or creating a record of the data, date, and name of the person transferring the data			
d. Periodic audits of data transfer			
-- Are audits documented?			
11. Are data entered into the central data system via a link from an instrument? If yes, is the following information given			
a. Data generated			
b. Date and time generated			
c. Identification of the instrument			
d. QC flags indicating the level of acceptability of the data			
e. Instrument conditions (operating) (Lab)			
12. If data are entered by direct link from an instrument is the data validated by... a. Voltage or calibration checks			
b. Comparison of results			
c. Comparison of hardcopy output with database contents			
d. Are data reasonableness checks either built into the data capture system or instruments?			
13. Data changes a. How are the data corrections made? _____			
b. Are corrections verified?			

	Yes	No	NA
c. Are corrections documented on a written log (who, when, authorization)			
d. Is there a computer generated record of the changed and unchanged data?			
e. If changes were made to data transferred from another source, was the original source corrected?			
f. If changes are made in flags from a central data base -- Who determines the need to make the change? _____			
-- Is authorization for revision documented?			
-- Is the change adequately documented?			
14. Data Reduction, Analysis and Assessment a. If data quality flags are used, are they defined			
b. Are qualifying flags correct?			
c. Can new flags be created? If so, how are they created? _____ _____ _____			
d. Are the mathematical expressions used by the system available in a written format?			
e. Were the mathematical expressions reviewed for accuracy?			
f. Was the validation of mathematical expression documented?			
g. Did revisions affect the over-all performance of data manipulations?			
h. If mathematical expressions are modified, -- Is the reason for the modification documented?			
-- Are old data recalculated with new formulas?			

	Yes	No	NA
i. Are printouts of report modifications routinely checked for accuracy? -- By whom _____ -- Documented _____ -- Percent checked ____%			
15. Data output a. Are there written procedures for generating data output (graphs, charts, reports)			
b. Are the data used to generate the output adequately identified			
c. After final output are generated, is the database "locked" so that no further changes can be made without managerial consent			
d. Are output generated in a timely fashion (unusual delays be database)			
e. Are final copies of output properly archived			
-- Secure			
-- Precaution against fire			
-- Limited access			
f. Are reports electronically generated?			
g. If reports are electronically generated, how is it done? -- Direct computer link			
-- Magnetic media			
-- Verified upon transmission			
16. Backups a. Are system backups performed?			

	Yes	No	NA
b. If backups are performed, indicate frequency _____			
c. Are backups partial or total?			
d. Is there an individual responsible for backups? If so, who? _____			
e. On what media are backups stored -- Magnetic tape			
-- Disks			
-- Diskettes			
-- Other _____			
f. Are media storing backups properly labelled?			
g. Are backups documented? -- Written log			
-- Project identified			
-- Computer system identified			
-- Other _____			
h. How are data from backups stored short term _____			
i. How are backup data stored long term _____			
j. Are the backup data arranged for expedient retrieval from the storage area(s)?			
k. Are long-term backup data stored off site or in a location other than the original data?			
l. Is the data storage area adequate?			
-- Security assured			

	Yes	No	NA
-- Access limited			
-- Fire precautions			
-- Environmentally controlled			
17. Hardware maintenance a. Are there procedures for conducting and documenting preventative maintenance?			
b. Is there a regularly scheduled preventative maintenance program? Frequency _____			
c. Is preventative maintenance documented (who, what, when)? _____			
d. Is non-routine maintenance performed by in-house staff? If yes, how is it documented? _____ _____ _____			
e. How is non-routine maintenance documented? _____ _____ _____			
f. What provisions, if any, are made for system downtime? ____ _____ _____			

	Yes	No	NA
g. Has downtime adversely afffected the project? If yes, explain _____ _____ _____			
h. If the system fails because of electrical glitches or power outage what happens to the system? _____ _____ _____			
-- Backup power source available _____			
-- Backup starts automatically or manually _____			
-- How are power failures indicated if the system is running? _____			
-- Does the system lose the data being processed? _____			
-- Does the system start where it left off? _____ _____			

	Yes	No	NA
-- If data are lost, does the system indicate the loss? _____			
-- While the system is running, is there a backup procedure to minimize the data loss if the system goes down? _____			
-- If the system goes down is the time down and time restored easy to determine? _____			

This page is intentionally left blank.

VOLUME VI: CHAPTER 4

EVALUATING THE UNCERTAINTY OF EMISSION ESTIMATES

July 1996

Prepared by:
Radian Corporation

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

DISCLAIMER

This document was furnished to the U.S. Environmental Protection Agency by Radian Corporation, Research Triangle Park, North Carolina. This report is intended to be a final document and has been reviewed and approved for publication. The opinions, findings, and conclusions expressed are those of the authors and not necessarily those of the U.S. Environmental Protection Agency. Mention of company or product name is not to be considered as an endorsement by the U.S. Environmental Protection Agency.

CONTENTS

		Page
1	Introduction	4.1-1
	1.1 Background	4.1-2
	1.2 Uncertainty Analysis	4.1-3
	1.3 EIIP Recommendations	4.1-6
2	Sources of Uncertainty in Emission Inventories	4.2-1
	2.1 Variability	4.2-2
	2.2 Parameter Uncertainty	4.2-5
	2.3 Model Uncertainty	4.2-8
3	Qualitative Uncertainty Analysis	4.3-1
4	Semiquantitative Data Quality Rankings	4.4-1
	4.1 DARS	4.4-1
	4.2 *AP-42* Emission Factor Rating System	4.4-4
	4.3 Other Grading Systems	4.4-5
	4.4 GEIA Reliability Index	4.4-7
5	Quantitative Uncertainty Analysis	4.5-1
	5.1 Expert Estimation Method	4.5-1
	5.2 Propagation of Error Method	4.5-4
	5.3 Direct Simulation Method	4.5-5
	5.4 Other Methods	4.5-12
6	References	4.6-1

FIGURES

Page

4.5-1. Recommended IPCC Procedures for Quantifying Uncertainty 4.5-6

TABLES

Page

4.1-1 Overview of Methods Used to Estimate Emissions Uncertainty 4.1-4

4.1-2 Preferred and Alternative Methods for Qualifying Emission Inventory Data 4.1-7

4.1-3 Comparison of Uncertainty and Data Quality for Three Estimation Methods for Industrial Surface Coatings 4.1-8

4.2-1 Examples of Variability in Emission Source Estimates 4.2-3

4.2-2 Examples of Parameter Uncertainty in Emission Source Estimates 4.2-6

4.2-3 Examples of Model Uncertainty in Source Emission Estimates 4.2-9

4.3-1 Summary of Uncertainties Associated with the OCS Production-Related Emissions Inventory 4.3-2

4.4-1 Preferred and Alternative Methods for Ranking Systems 4.4-2

4.4-2 DARS Scores for Architectural Surface Coating Emissions Estimated by Two Different Methods 4.4-3

4.4-3 *AP-42* Rating System for Emissions Test Data 4.4-5

4.4-4 *AP-42* Rating System for Emission Factors 4.4-6

4.4-5 Data Quality Codes Recommended by the IPCC 4.4-8

4.5-1 Preferred and Alternative Methods for Quantifying Uncertainty 4.5-1

4.5-2 Comparison of VOC Emission Uncertainty Estimates Derived Using Three Alternative Uncertainty Estimation Methods 4.5-3

4.5-3 Estimated Coefficient of Variation for Parameters Used in Estimating SO_2 Emissions from Industrial and Commercial Sources 4.5-9

TABLES (CONTINUED)

Page

4.5-4 Estimates of Uncertainty of CO_2 Emissions in Canada Preliminary 1990 CO_2 Estimates in Kilotonnes 4.5-10

1

INTRODUCTION

The quality assurance and quality control (QA/QC) procedures described in other chapters of this volume are designed to ensure that the appropriate methods and data are used, that errors in calculations or data transcriptions are minimized, and that documentation is adequate to reconstruct the estimates. It is important to recognize that the resulting quality of the emission estimates is only partly determined by adherence to a good QA program. The quality of the emission estimates is also determined by the uncertainty inherent in the estimates.

This chapter deals with the determination and evaluation of the uncertainty in emission estimates and the methodology available to do this. The goal is always to reduce uncertainty. To do so, the inventory preparer must first know the sources of bias and imprecision in the estimates; Section 2 discusses the sources of uncertainty and gives specific examples. The next step in certifying the emissions inventory is to conduct a qualitative assessment of the sources of uncertainty in the inventory. Section 3 provides an example of how this can be done. The third step is to develop subjective quality indicators for the source categories; Section 4 describes alternative approaches that produce subjective quality indicators. Finally, Section 5 describes several approaches for quantitative uncertainty analysis, arranged in order of increasing complexity.

The intended uses of the emissions data should be considered before spending significant resources on quantifying the uncertainty and reducing it. For example, if the inventory for a point source is being used to show compliance with an emissions limit, the relative accuracy is usually part of the reporting requirements. The uncertainty associated with the estimate is low.

On the other hand, a national inventory to identify and rank the relative importance of sources of a specific hazardous air pollutant (HAP) may not be as concerned with the uncertainty of specific estimates. This is especially true of smaller emissions sources. If an estimate is highly uncertain, but at worst represents only 1 percent of all the emissions, accurately quantifying the uncertainty is probably not a high priority. However, a source that is insignificant at a national level can be very important at a local level. When viewed from the local community's perspective, high uncertainty in the estimated emissions may be unacceptable.

1.1 BACKGROUND

As discussed in Chapter 2 of this volume, the desired quality of an inventory is a function of the intended end use. If a Level IV inventory is being prepared, the users must be willing to accept that the estimates are not necessarily of the best possible quality, whereas a Level I inventory implies the highest possible data quality.

It is not always possible to achieve the desired level of quality. In some instances, the state-of-the-science may not be sufficient to provide the level of detail desired. In other situations, unforeseen problems (e.g., equipment failure, survey responses not as high as expected, activity data not available) may be encountered in the process of preparing the estimates. In any event, an important step in preparing an emissions inventory is to "qualify the data." This term means to provide an assessment of how closely the desired level of quality, or data quality objective (DQO), is met by the inventory preparer. Ideally, the target data quality can be evaluated using a quantitative data quality indicator (DQI).

When discussing the quality of an estimate, the term "uncertainty" is often used as an indicator of quality, rather than "accuracy" because there is no reasonable or practical way to determine to emission values for comparison. Confidence in an estimate is generally determined by our perception of the reliability of the underlying data and model used to generate the emissions estimate. For example, an annual boiler nitrogen oxides (NO_x) emission estimate generated using continuous emission monitor (CEM) data is generally held to be more reliable (less uncertain) than an estimate based on fuel consumption and an accepted emission factor. However, this logic implicitly assumes that the CEM is maintained properly, that QA and calibration procedures are rigorously followed, and that the data capture is near 100 percent. So, assuming that appropriate QA procedures are followed in both cases, the CEM estimate is assumed to be of higher quality (i.e., more reliable and less uncertain) than the estimate based on an emission factor.

Calculating the range, confidence interval, or other error bounds for an emission estimate is a very important tool for assessing the uncertainty of the estimate. However, these statistics are not complete measures of quality because there may be systematic errors (biases) associated with the emission estimate that are not bounded by the range or confidence interval estimates. In addition, uncertainty is due to many causes, one of which is the inherent variability in the process or processes that cause the emissions. Even if all other sources of uncertainty were removed, the variability remains. Because some processes are more variable than others, some will always have larger error bounds than others. That does not mean that the estimates are of lower quality. It

does mean that we do not have as much confidence in our ability to predict the emissions at a particular point in time, but that we can confidently predict a range.

Emission inventory development and uncertainty analysis should be an iterative process. Once estimates of uncertainty are developed, the inventory preparer should review the inventory and target the significant sources with the largest uncertainty for more research. The objective of this iterative process is a minimization of overall uncertainty in the inventory. Several factors make this process difficult to implement:

- Data are not available (and not readily measurable) to quantify the uncertainty;
- The available data are insufficient to meet the data input needs of the statistical or numerical methods to be used to estimate uncertainty; and
- Reducing the uncertainty requires more resources (i.e., money and time) than are available.

The solutions to the second and third problems require the expenditure of resources that may not be available. However, the Emission Inventory Improvement Program (EIIP) has developed recommendations for methods to be used to develop better uncertainty estimates if the necessary resources are available. The EIIP recommendations for implementing uncertainty analyses are presented in the next section.

1.2 UNCERTAINTY ANALYSIS

The first step towards reducing the uncertainty associated with emission estimates is to understand and quantify the various sources of variability and inaccuracies in the data used to estimate the emissions. This analysis should include an assessment of both bias and imprecision in the estimates. When identified, bias should be eliminated while imprecision should be minimized. The remaining sources of uncertainty in the inventory should be identified and quantified if possible.

The initial task in any emissions uncertainty analysis is the definition of the analysis methodology to be used to estimate emissions uncertainty. Table 4.1-1 presents a list of eight general types of analyses that have been used or are currently being used to evaluate emissions inventory uncertainty. A brief overview of each general method, with references, is given in Table 4.1-1. Additional discussion and examples of each method are described in Sections 3 through 5 of this chapter. The inventory specialist must be aware that each of the methods in Table 4.1-1 may provide different estimates of

TABLE 4.1-1

OVERVIEW OF METHODS USED TO ESTIMATE EMISSIONS UNCERTAINTY

Methodology	References	Description	Approximate Level of Effort[a]
Qualitative Discussion	Steiner et al., 1994	Sources of uncertainty are listed and discussed. General direction of bias, and relative magnitude of imprecision are given if known.	<100 hrs
Subjective Data Quality Ratings	U.S. EPA, 1995 Saeger, 1994	Subjective rankings based on professional judgement are assigned to each emission factor or parameter.	<100 hrs
Data Attribute Rating System (DARS)	Beck et al., 1994	Numerical values representing relative uncertainty are assigned through objective methods.	<500 hrs
Expert Estimation Method	Linstene and Turoff, 1975 SCAQMD, 1982 Horie, 1988 Horie and Shorpe, 1989	Emission distribution parameters (i.e., mean, standard deviation, and distribution type) are estimated by experts. Simple analytical and graphical techniques can then be used to estimate confidence limits from the assumed distributional data. In the Delphi method, expert judgement is used to estimate uncertainty directly.	<500 hrs
Propagation of Errors Method	Mangat et al., 1984 Benkovitz, 1985 Benkovitz and Oden, 1989 Balentine et al., 1994 Environment Canada, 1994	Emission parameter means and standard deviations are estimated using expert judgement, measurements, or other methods. Standard statistical techniques of error propagation typically based upon Taylor's series expansions are then used to estimate the composite uncertainty.	<500 hrs
Direct Simulation Method	Freeman et al., 1986 Iman and Helton, 1988 Oden and Benkovitz, 1990 Efron and Tibshirani, 1991 Environment Canada, 1994 Gatz and Smith, 1995a Gatz and Smith, 1995b	Monte Carlo, Latin hypercube, bootstrap (resampling), and other numerical methods are used to estimate directly the central value and confidence intervals of individual emission estimates. In the Monte Carlo method, expert judgement is used to estimate the values of the distribution parameters prior to performance of the Monte Carlo simulation. Other methods require no such assumptions.	<1,000 hrs

TABLE 4.1-1

(CONTINUED)

Methodology	References	Description	Approximate Level of Effort[a]
Direct or Indirect Measurement (Validation) Method[b]	Pierson et al., 1990 Spellicy et al., 1992 Fujita et al., 1992 Peer et al., 1992 Mitchell et al., 1995 Claiborn et al., 1995	Direct or indirect field measurement of emissions are used to compute emissions and emissions uncertainty directly. Methods include direct measurement such as stack sampling and indirect measurement such as tracer studies. These methods also provide data for validating emission estimates and emission models.	>1,000 hrs
Receptor Modeling (Source Apportionment) Method[b]	Watson et al., 1984 Lowenthal et al., 1992 Chow et al., 1992 Scheff et al., 1995	Receptor modeling is an independent means to estimate the relative contribution of specific source types to observed air quality measurements. The method works best for nonreactive pollutants for which unique emission composition "fingerprints" exist for all significant source categories. The method provides a measure of the relative contribution of each source type but not absolute emission estimates.	>1,000 hrs
Inverse Air Quality Modeling Method[b]	Hartley and Prinn, 1993 Chang et al., 1993 Chang et al., 1995 Mulholland and Seinfeld, 1995	Air quality simulation models are used in an inverse, iterative approach to estimate the emissions that would be required to produce the observed concentrations fields.	>1,000 hrs

[a] The levels shown are a relative level of effort, including data collection. The actual effort will depend upon the scope of work implemented.

[b] These methods are described in Chapter 3, Section 9, "Emission Estimation Validation" of this volume. They can be used to develop estimates of uncertainty as well.

uncertainty when applied to the same data set. These differences range from slight to significant. A method should be chosen and applied consistently to the inventory categories. If different methods are used to develop different source groups, comparisons between the uncertainty results may not be meaningful. The overall goal of any emissions uncertainty analysis is likely to be the development of confidence limits about the mean of emission estimates from each source type analyzed. The significance level assumed for the confidence limits, generally 90 or 95 percent, is a function of the quality of the input data available and the use to which the uncertainty estimates will be put. It is up to the analyst for each study to determine the appropriate significance level for his or her study.

1.3 EIIP RECOMMENDATIONS

The preferred and alternative methods of qualifying emissions inventory data are summarized in Table 4.1-2. Note that there are two aspects to these recommendations. The first is that all three elements--qualitative assessment, ranking, and quantitative uncertainty--are included; the second is that different methods are preferred for completing these three elements. Inventory preparers are not constrained to the combinations of elements shown in this table; rather, they should develop a plan for qualifying the data that is most suitable for the specific situation. The methods shown are the minimum recommended for the level shown.

As discussed above, the lack of necessary data is a significant limitation in the development of emission inventory uncertainty estimates. For these instances, the EIIP recommends use of ranking methods. The EIIP preferred ranking method is the Data Attribute Rating System (DARS). Because of its potential to provide significant information on emissions inventory uncertainty, the EIIP has focused on the DARS method for further development.

The DARS method addresses four major sources of uncertainty or error in both the emission factor and the activity data. Numerical scores are assigned based on predefined criteria, and a composite score is computed to provide an overall indicator of the relative quality of the estimate. While DARS does address uncertainty in a subjective way (i.e., the higher the DARS score, the more confidence or certainty we have about the estimate), it does not quantify the imprecision of the estimate. For this reason, the EIIP strongly encourages the additional use of quantitative methods to calculate confidence intervals (or other measures of the dispersion of the estimate).

For many area sources, the preferred method is to conduct a survey of facilities in the inventory region to gather more accurate activity and emissions data. The value of using this more resource-intensive approach is shown in Table 4.1-3. This table shows

TABLE 4.1-2

PREFERRED AND ALTERNATIVE METHODS FOR QUALIFYING EMISSION INVENTORY DATA

	Qualitative	Ranking[a]	Quantitative Uncertainty
Preferred (Level 1)	Provide a qualitative assessment of uncertainty, addressing bias and imprecision of key data elements; indicate direction of bias and relative magnitude of imprecision where possible. Provide any statistical measures of data dispersion that are available.	For each source contributing to the top 90% of emissions, provide a subjective relative ranking of the quality of the estimate.	Quantify the range of the estimates as a 90% confidence level for all sources.
Alternative (Level 1 or 2)	Provide a qualitative assessment of uncertainty, addressing bias and imprecision of key data elements; indicate direction of bias and relative magnitude of imprecision where possible. Provide any statistical measures of data dispersion that are available.	For each source contributing to the top 90% of emissions, provide a subjective relative ranking of the quality of the estimate.	Quantify the range of estimates at the 90% confidence level for the top 10 sources in the point, area, on-road mobile, non-road mobile, and biogenic categories.
Other Methods (Level 3)	Provide a qualitative assessment of uncertainty, addressing bias and imprecision of key data elements; indicate direction of bias and relative magnitude of imprecision where possible. Provide any statistical measures of data dispersion that are available.	Rank sources from largest to smallest; provide subjective relative ranking for as many as possible (starting with largest).	None.
Other Methods (Level 4)	Provide a qualitative assessment of uncertainty, addressing bias and imprecision of key data elements; indicate direction of bias and relative magnitude of imprecision where possible. Provide any statistical measures of data dispersion that are available.	None.	None.

[a] The EIIP preferred ranking method is DARS.

TABLE 4.1-3

COMPARISON OF UNCERTAINTY AND DATA QUALITY FOR THREE ESTIMATION METHODS FOR INDUSTRIAL SURFACE COATINGS

Method	Emissions of VOCs (tpy)	Assessment of Imprecision	DARS Score	Level of Effort (hours)
Per Capita	7423	Very high[a]	0.15	1
Per Employee	589	High[a]	0.43	200
Survey	198	±40%[b]	0.86	300

[a] Qualitative assessment.
[b] 90% confidence interval based on survey data.

the estimated emissions for industrial surface coatings in the Houston area. The first two estimates were calculated using volatile organic compound (VOC) per capita and per employee factors, respectively, and then using standard speciation profiles to allocate the emissions to individual chemical species (speciation). The third estimate was based on a survey of Standard Industrial Classification (SIC) codes included in this category. A telephone survey of 198 facilities was first used to determine what fraction of the surveyed facilities actually were sources of hazardous air pollutant (HAP) emissions. (Note: this survey was designed primarily as a survey of organic HAP emissions, but data on total VOCs were also collected. Only the VOC results are used here as an example.) The total number of facilities (416) in the SIC group was then multiplied by the survey fraction of emitting sources to give an estimate of the total number of emitting sources. A subset of 32 sites were then visited and solvent use data were collected. This information included material safety data sheets (MSDSs) documenting solvent composition and the total annual volume of solvent. From this data set, emission factors for each pollutant were developed on a per facility basis, and used with the estimated number of facilities to estimate VOC emissions in the Houston area.

The estimated VOC emissions as calculated in tons per year (tpy) by each method are given in Table 4.1-3, along with an analysis of the uncertainty, the DARS scores for each, and an estimate of the number of labor hours required for each method. As is clearly shown, the per capita estimate requires very little effort, but produces an estimate of very high uncertainty. This low-cost estimate is also shown to overestimate emissions by an order of magnitude for this case.

The second approach requires an intermediate expenditure of time but produces estimates closer to the best estimate. (For this specific example, however, the estimated emissions would have been twice as high if the number of employees had not been adjusted using the phone survey results.) The third approach targeted uncertainty in both the factor and in the activity data; while considerably more resources were required to generate this estimate, the results are dramatic both in the decrease in estimated emissions and in the increase in quality.

In this example, the higher the quality of the estimate, the lower the emissions. This will not always be the case; because the per capita and per employee factors are based on national averages, these factors will over- and underestimate emissions for specific regions (assuming that the national estimates are not biased in some way). The table does not include an assessment of possible bias. All methods potentially overestimate emissions (i.e., positive bias) because they do not account for non-air losses. However, the latter two possibly have a negative bias in that potential respondents are more likely to decline to answer when they *are* a source than when they are not.

This page is intentionally left blank.

2

SOURCES OF UNCERTAINTY IN EMISSION INVENTORIES

Estimates in emission inventories are nearly always the result of modeling of one form or another. The simplest emissions modeling method is the use of an emission factor multiplied by an activity level to approximate emissions. Statistical models (such as regression models) are a more sophisticated way to achieve the same objective. Or, more complex models such as the Biogenic Emissions Inventory System (BEIS) or the Mobile Source Emissions Model MOBILE5a use detailed input data to generate emission estimates or factors. Temporal and spatial allocation of emissions may require further modeling through the use of statistical analysis or surrogate variables to distribute the emissions data or underlying activity to a grid at a specified temporal resolution. In all cases, uncertainty is associated with the development and adjustment of emission estimates.

Uncertainty in emission estimates is due to a variety of causes. First, there is inherent variability in the processes producing the emissions. For example, sulfur dioxide (SO_2) emissions from combustion sources fluctuate with the sulfur content of the fuel and the process load. For other sources, uncertainty results from variation in the environmental factors that produce the emissions (e.g., biogenic emissions vary with temperature, sunlight intensity, exposed leaf surface area, and other environmental factors). Other sources of uncertainty in emission estimates stem from the methods, models, and assumptions used to fill in our incomplete knowledge about the emission process and allow simplistic estimation of emissions from highly complex processes. Still other uncertainty comes from the measurement methods and instruments themselves. Finally, random errors--usually stemming from human errors or naturally occurring but unforeseeable events--introduce uncertainty.

The term *uncertainty* comprises two types of error in estimation: bias and imprecision. A *bias* is a consistent difference between a measurement and its true value that is not due to random chance. In an emissions inventory, bias can result from an emissions estimation process in which a systematic error occurs because some aspect of the emissions inventory process is misrepresented or is not taken into account. For example, if the emission factor for a given source category was developed from a nonrepresentative sample of source types, the emission factor will produce a biased estimate of emissions. Bias may also result due to one's inability to obtain a

comprehensive set of measurements for all conditions producing emissions (i.e., one cannot perform source sampling for all conditions under which the source may operate). A common example of bias is the use of solvent consumption as a surrogate for emissions; if the disposal of waste solvent or other nonair releases are ignored, this approach consistently overestimates emissions (i.e., positive bias).

In contrast to bias, *imprecision* in a parameter is the difference due to random error or fluctuations between a measurement and its true value. Multiple measurements of the parameter will differ, but--if the measurements are nonbiased--the measurements will cluster about the true value of the parameter. This imprecision is caused by sampling error and human error, as well as by the natural fluctuations in the process being measured. Emissions data variability results from a number of causes including temporal or spatial fluctuations in data used to estimate emissions (e.g., the temporal variation in the fuel sulfur content, heating value, and load for an industrial boiler). In addition, there are inherent differences in individual emission sources in that no two sources or operations can be exactly identical.

The factors producing uncertainty in emissions data can be separated into three general classes: variability, parameter uncertainty, and model uncertainty. This system of classifying uncertainty is based on a discussion by Finkel (1990) of uncertainty in risk assessment. In the following section, these concepts are applied to the understanding and estimation of uncertainty in emission estimates.

2.1 VARIABILITY

Variability is inherent in the process that produces emissions. If all other sources of uncertainty were removed, the inherent variability would still make it impossible to precisely specify emissions at a certain point in time and space. Some processes have very little natural variability, others have a lot. There are two major components of the variability that occur in emissions estimates and the data used to create the emissions estimates. The first component is the uncertainty introduced by variation from source to source (spatial uncertainty) and the second component is within source variation (temporal uncertainty). Table 4.2-1 presents examples of these two sources of variability in emission sources.

Source-to-source differences, such as the vehicle fleet composition between urban areas, differences in the process operation of two refineries, and differences in the physical attributes of similar boilers, introduce imprecision into estimates of emissions, emission factors, and activity data. However, even if all source-to-source uncertainty were eliminated, there would still be uncertainty in emission estimates resulting from within-source (temporal) variability. Factors such as the change in equipment operating

TABLE 4.2-1

EXAMPLES OF VARIABILITY IN EMISSION SOURCE ESTIMATES

Source of Variability	Examples of Causes	Ways to Minimize Effect
Inherent variability between sources	Environmental factors vary spatially (e.g., application rate, soil moisture, ambient temperature, isolation when determining VOC emissions from a pesticide application). Hourly/daily/weekly/seasonal variations in activity (e.g., seasonal agricultural activities, business day versus weekend activities, morning/evening commute, batch processing operations). Annual variability in activity (e.g., heating and cooling demand, economic growth). Processes or activities included in the category are not uniform (e.g., product formulations vary between manufacturers, product can be produced using several processes).	Identify environmental factors responsible for variation in source emissions or activity. Make sure the averaging time of the emission factor and activity data are appropriate for temporal scale of emission estimates desired. If possible, subdivide category to create more uniform subcategories.
Inherent variability within a source	Source emissions and effectiveness of emission control systems on a source can be a function of age and maintenance history of the source. Load or production variability of a source (e.g., dry cleaning emissions depend upon demand that can vary day to day). Variation in fuel characteristics and raw materials input to an industrial process (both within-specification and outside-specification). Inherent differences in two similar pieces of equipment (i.e., no two boilers can be exactly the same).	Detail age and maintenance history of all sources. Document variability in load and production for the time scale of interest. Document fuel, raw material, and processing variability for a given source, particularly in batch processing operations. Quantify physical differences between individual pieces of equipment (e.g., type of emission control system, modifications to original equipment).

characteristics with age, variation in fuels, load fluctuations, and maintenance history all contribute to the variability of emission estimates for a single source.

An incomplete understanding of the variability in a process can lead to systematic errors in estimation. For example, emissions due to the application of pesticides are highly variable. Rather than being entirely random, however, the emissions are a complex function of the volatility of the solvents in the pesticide, existing meteorological conditions, the amount and type of vegetation sprayed, the method of application, and the effect of biological organisms that can metabolize the pesticide. However, because the form and magnitude of these complex relationships are unknown, the inventory preparer tends to "be conservative" and assume that all the solvent applied is emitted. Even when an adjustment is made (e.g., assume 90 percent is emitted), that adjustment is often an expert judgement that may still produce biased results. Because adjustments that are not supported by data introduce an *unknown* bias, the tendency is to estimate high so that the direction (if not the magnitude) is known.

Most sources show some sort of temporal variation because of variability in activity patterns. For example, residential fuel consumption is higher in the winter than in the summer. Commercial or industrial activity tends to be greater on weekdays than on weekends. Other sources have variable emissions due to variability in load, operation, or fuel composition. For example, municipal solid waste combustors are characterized by spikes in SO_2 emissions that are associated with the random feed into the combustor of individual waste elements with large, and highly variable, sulfur contents. For these variable sources, activity data (i.e., fuel composition and feed rate) must be known to at least the temporal resolution required for the emission estimates in order to minimize imprecision in the emission estimates.

For many sources, the main recognized source of variability in emissions is temporal fluctuations in activity, which are usually greatest on a daily or weekly basis (e.g., weekday activity rates tend to be higher than weekend rates). Some sources vary significantly between years, particularly if emissions are driven by extreme events (e.g., chemical spills and extreme meteorological conditions).

The uncertainty due to source variability should be quantified and minimized whenever possible. Many times it is possible to attribute a portion of the emission uncertainty to a given source of variation. However, it is never possible to eliminate all imprecision in emission estimates. For example, it would not be feasible to obtain hourly use rates of dry cleaning fluids at all dry cleaning establishments in an urban area. If an estimate of the confidence interval (or other measure of dispersion) is available for a given parameter, that portion of uncertainty that is attributable to that parameter can potentially be quantified. However, other sources of variability may not be quantifiable;

for example, source production data may be available on an hourly or daily basis, but detailed fuel sulfur content is known only as an annual average.

Good inventories will minimize the uncertainty due to temporal variability by ensuring that input emission factors and activity data match the scale of the inventory. If factors or activity have to be scaled up or down, adjustments must be made that account for temporal variability. Similarly, any other adjustments to the calculation to account for variability should be performed.

2.2 PARAMETER UNCERTAINTY

Parameter uncertainty is caused by three types of errors: measurement errors, sampling errors, and systematic errors (also called nonrandom error or bias). Examples of these types of parameter errors are given in Table 4.2-2.

Measurement errors occur because of the imprecision of the instrument or method used to measure the parameters of interest. Where emissions are measured directly, the measurement error of a particular method is usually known; the U.S. Environmental Protection Agency (EPA) typically uses the concept of relative accuracy to describe the performance of a measurement method (or device) with respect to an EPA Reference Method.

A more common measurement error for area sources occurs due to misclassification. For example, area source categories are frequently identified by SIC group, and the number of employees or facilities in a particular SIC group are used as the activity data. However, some SIC groups encompass a wide variety of industrial processes and activities, not all of which are really sources of emissions. This issue can still be a problem even when a survey is used to gather activity data within a SIC group if the sample design does not account for subpopulations adequately. For example, different manufacturing processes may be used to produce the same product; the ratio of emissions to employees may be different for these processes. In addition, facilities are sometimes listed under an incorrect SIC or may have more than one SIC. Any of these errors results in misclassification of data and adds to our uncertainty about the emissions estimates.

Sampling error is an important factor when one or more of the parameters (i.e., activity, factors, or emissions) are to be estimated from a sample of the population. While most people recognize the importance of an adequate sample size, obtaining an adequate sample size is often difficult. Furthermore, sample data are usually used to estimate the arithmetic mean value from which the population mean is extrapolated. This approach assumes that the underlying data are normally distributed--an assumption that is often

TABLE 4.2-2

EXAMPLES OF PARAMETER UNCERTAINTY IN EMISSION SOURCE ESTIMATES

Source of Parameter Uncertainty	Examples of Causes	Ways to Minimize Effect
Measurement errors in activity data and emission factors	Inherent random error in measurement equipment (e.g., selected anemometer is only accurate to the nearest 0.1 m/sec, air flow meter is accurate to only 10% of measured flow, CEM has relative accuracy of 12%). Monitoring equipment error tolerance too high. Misclassification of activity data (e.g., wrong area source SIC category used).	Use monitoring equipment adequate to gather data required (i.e., do not use an instrument accurate to only 1.0 ppm if 0.1 ppm levels are required). Establish and follow a data collection protocol including performance QA measurements (i.e., field blanks, duplicate samples or data entry). Verify appropriateness of all activity and emission factor data.
Sampling (random) error in activity data and emission factors	Inadequate sample size. Errors in performance of the sampling (e.g., improper probe placement in stack, misread dials, failure to follow the sampling protocol). Sampling equipment or source not in a stable or steady-state mode (i.e., monitoring equipment not at a stable temperature, source subject to load fluctuations, data collection during atypical production periods). Sampling protocol or sampling equipment inadequate to produce required resolution in collected data.	Establish sample size required to meet analytical needs as part of sampling protocol. Establish and follow a monitoring (or sampling) protocol. Audit all monitoring results to ensure compliance with proper procedures and sampling protocols. Perform a defined number of measurements as QA measurements (i.e., field blanks, duplicate samples or data entry).

TABLE 4.2-2

CONTINUED

Source of Parameter Uncertainty	Examples of Causes	Ways to Minimize Effect
Systematic errors (bias)	Inherent bias in a survey (e.g., only largest facilities are surveyed and they do not reflect activities at smaller facilities).	Develop and follow a sampling or inventory development protocol.
	Misclassification of data (e.g., SIC group used does not accurately define activities of facility).	Obtain external review of methods by a qualified expert.
	Incorrect assumption (e.g., assuming 100% rule compliance and ignoring rule effectiveness).	Make sure that characteristics of the source population are understood and accounted for in the sampling or emission estimation methods.
	Improper calibration of monitoring equipment.	Validate all assumptions.
	Sampling methodology improper for sources sampled.	Compare emission estimation or sampling results to similar data from other studies.
	Use of nonrepresentative meteorological data in estimation procedures (e.g., temperature or wind speed data used are not valid for the situation).	Perform mass balance or other simple, common sense checks to ensure reasonableness of data.

violated (see Chapter 3, Section 7, of this volume). If the underlying data are extremely skewed, a small sample size can lead to very large errors in estimating means. Again, sampling error can be minimized if proper statistical approaches are used, QA procedures are followed, and sample sizes are adequate and properly obtained.

Systematic errors (bias) are the most problematic sources of parameter uncertainty because they are the most difficult to detect and reduce. They occur primarily because of an inherent flaw in the data-gathering process or in the assumptions used to estimate emissions. A common way that this happens is if the population to be sampled is not well defined, and a sample (thought to be random) is actually nonrandom. This is a fairly common problem for certain types of industries. For example, consider a local survey of solvent use by auto body refinishing shops. One approach would be to develop a list of facilities from business registration or other state/local business listings. However, this industry has a very large number of "backyard" operations that are not identified in these official lists. Therefore, any sample that did not recognize this fact would have systematic sampling errors.

As part of the emissions inventory development process, the goal is to reduce all known sources of bias, both across and within sources. If a bias is known to exist, then effort should be initiated to quantify and remove the bias. However, in practice this may be difficult to accomplish because of a lack of resources, data, or other factors. For example, the source testing used to develop the emission factors for a given class of sources could potentially exclude a key source type. Because this key source type would not be represented in the emission factor, the emission factor would potentially contain a known (or suspected) bias. However, resources may not be available to perform the needed source testing to develop a revised emission factor incorporating this key source type. Consequently, a known bias would exist in the emission inventory but would not be readily susceptible to elimination.

2.3 MODEL UNCERTAINTY

Model uncertainty applies to most emission estimates. In this context, a model is a simplified representation of the processes leading to the emissions. Model uncertainty stems from the inability to simulate the emission process completely due to the use of surrogate variables, exclusion of variables from the computation process, and over-simplification of emission process by the model. Table 4.2-3 presents examples of model uncertainty in emission estimates.

TABLE 4.2-3

EXAMPLES OF MODEL UNCERTAINTY IN SOURCE EMISSION ESTIMATES

Source of Model Uncertainty	Examples of Causes	Ways to Minimize Effect
Use of surrogate variables	Surrogate variable is an incomplete representation of the activity or variable desired (e.g., factors in addition to heating degree days [HDD] contribute to the demand for space heating). Use of surrogates can mask underlying relationships between activity data and emissions.	Enhance emission models to account for more fundamental parameters. Develop emission factors based on statistically correlated surrogates. Obtain site-specific data through statistically valid surveys and with site units so that surrogate data use can be minimized.
Model simplification/ over-simplification	Data parameterized (separated into classes) rather than used as discrete values (e.g., speeds in motor vehicle emission factor models input as discrete classes, traffic network input as links and nodes). Reduction of a complex dependency to a single factor (e.g., emissions of biogenic isoprene are a complicated function of the temperature and the wavelength distribution of incoming sunlight at the leaf surface but are typically modeled as a function of ambient temperature and sunlight intensity at a single wavelength). Emission model contains an invalid representation of the process producing emissions (e.g., older versions of motor vehicle emission factor models significantly underestimated evaporative VOC emissions).	Verify the theoretical basis for all models. Validate emissions models against independent data. Use continuous variable representations where feasible and appropriate.

Most emission estimates are the product of a series of quasi-independent parameters (i.e., emission factor, activity data, control factor, temporal adjustment) of the form

$$ER_t = p_1 \times p_2 \times \ldots \times p_n \qquad (1)$$

where:

ER_t	=	Emission rate for time t;
p_i	=	Parameter used to estimates emissions, where i = 1, 2, . . ., n;
n	=	Number of parameters.

This same general equation, or linear model, applies to the simple case of an emission factor (e.g., grams per kilogram combusted) times an activity datum (e.g., number of kilograms combusted per day) as well as to the complex case where a model such as MOBILE5a or BEIS are used (although nonlinear terms may be introduced as well).

There are a number of real-world problems and complexities associated with estimating emissions uncertainty when the linear model is used to develop the emission inventory. These problems include the inherent (and generally erroneous) assumption of independence of the individual parameters, the complications inherent in obtaining temporal, spatial, and speciated estimates of emissions from average values of emissions (e.g., obtaining gridded, speciated, hourly emission estimates from annual county-wide emission estimates), the limited amount of data that may be available for validation of estimates, and the difficulty posed by temporal and spatial data dependencies in validating those estimates even when data are available.

The use of surrogate variables is common in area source methods where population or the number of employees are used as surrogates for emission activities. The uncertainty in using these surrogates is especially high when emissions for a small region (i.e., county or smaller area) are estimated using a national average factor. Local variations in activity are not necessarily accounted for by population or employment. A common example is found in large cities that have the corporate headquarters for an industry. The number of employees may be high, but all of the manufacturing may be occurring in other areas.

Per capita emission factors are often an oversimplification of emission processes. For example, the consumer/commercial solvent use factors are based on a national survey of the solvent constituents of various consumer products. From this data set, the national average consumption per person was calculated for various product groups (see Volume III, Chapter 5 of the EIIP series). These factors may not account for solvent that is not emitted because it is either disposed of (in containers) or washed down the drain; furthermore, publicly owned treatment works (POTW) or landfill emissions may

include these solvents, and emissions may therefore be double-counted. The factors also do not reflect regional variation in product usage, so when used to calculate emissions on a county basis, they are likely to over- or underestimate.

Point source emissions are also often based on the use of surrogates, although usually the surrogate is very closely related to the emissive activity. Fuel consumption, for example, is a surrogate for fuel combustion in a specific type of boiler. When an emission factor is used to estimate point source emissions, the assumption is that the design, processes, and test conditions of the original boiler (from which test data were derived) are good approximations of the boiler to which the factor is being applied. The further this assumption is from reality, the more uncertainty there is regarding the accuracy of the emission estimate.

This discussion of uncertainty in emissions inventories is by no means exhaustive. More details are provided in the specific volumes and chapters for point, area, mobile, and biogenic source categories. The EIIP has sought to encourage the reduction in uncertainty in their selection of "preferred" methods wherever possible. Emission factors are usually not the best choice if reducing uncertainty is the criterion; direct or indirect measurements, surveys, and other methods targeting the specific source are preferred. Unfortunately, this is not always practical. It is important that inventory preparers recognize the sources of uncertainty, quantify it, and reduce it as much as is practical.

This page is intentionally left blank.

3

QUALITATIVE UNCERTAINTY ANALYSIS

The simplest approach for estimating uncertainty is to discuss all known and suspected sources of bias and imprecision in the inventory. If possible, the direction (over- or underestimates) of any biases and relative magnitude (e.g., factor of two, order of magnitude) of the specific source of uncertainty should be stated. Sometimes standard deviations, confidence limits, or other statistics are available for some of the variables used to develop the inventory; if so, those statistics should be documented and their contribution to overall accuracy of the estimates should be discussed.

The qualitative uncertainty assessment can be presented in narrative form. However, tables provide a more systematic and concise method of summarizing the uncertainty. An example of a qualitative uncertainty assessment is shown in Table 4.3-1. This table is part of a report describing the results of the QA/QC procedures used during development of an inventory of emissions from offshore oil production facilities (Steiner et al., 1994). Many of the key sources of uncertainty shown are generally applicable to any inventory (e.g., survey respondent expertise and applicability/usage components). It is more important to list and discuss issues that are particularly relevant. For example, the authors of this study do a good job of describing uncertainties in their survey data.

A table such as this one is a good method for presenting the results of a qualitative assessment. One additional column that describes the direction (positive or negative) of any biases or the relative magnitude of any imprecision (if these are known) would provide additional valuable information to the assessment.

TABLE 4.3-1

SUMMARY OF UNCERTAINTIES ASSOCIATED WITH THE OCS PRODUCTION-RELATED EMISSIONS INVENTORY[a]

Inventory Component	Basis of Uncertainty	Description
Survey	Survey Respondent Expertise	Different levels of expertise of survey recipients could lead to incorrect or incomplete survey answers because of lack of understanding or incorrect interpretation.
	Unknown Answers	Some of the equipment on the platforms is very old and equipment ratings cannot be read or the equipment has been modified and manufacturers' ratings no longer are applicable.
	Incorrect Responses	Most likely some respondents did not read the directions, which could lead to aberrant or incomplete answers. Many of the problems corrected in the database were a result of incorrect units. Some of the flow rates in the survey were metered, others were not metered, and survey respondent had to guess activity levels.
	Data Entry	Even though we used a double data entry system to enter the data to minimize typographical and data omission errors, some may have occurred. In addition, some respondents had their survey responses typed onto forms by support staff, which could lead to data entry errors.
	Omitted Sources	15 percent of the companies operating platforms in the GOM contacted did not return the survey. Some of those companies may have multiple platforms. All of the major corporations operating in the Gulf returned their surveys. Some emissions sources (e.g., equipment) on the platforms may have been omitted because the survey respondent neglected to include information necessary. 6 percent of the helicopter companies contacted did not return the survey. Only the smaller helicopter companies did not return their survey. 26 percent of the vessel companies contacted did not return the survey. The companies that did not return the surveys are the smaller operations.
Emissions Methodology	Emission Factors	Emission factors represent an average population. Gulf population may not be representative of the emission factor mix.
	Fugitive Emissions	An empirical formula derived from Pacific OCS facilities was used. Gulf OCS platforms were not exactly configured as those in Pacific and product mix of oil and gas was different.
	Applicability/Usage	Even though the methodologies were reviewed for applicability, there is the possibility that a more applicable emissions methodology exists or that the methodology was applied incorrectly because of an incorrect assumption.

[a] Source: Steiner et al., 1994. OCS = outer continental shelf. GOM = Gulf of Mexico.

4

SEMIQUANTITATIVE DATA QUALITY RANKINGS

Semiquantitative ranking methodologies are relatively easy to implement and can be used where detailed data on emissions are unavailable. A drawback of their use is that it can be difficult to prevent logical inconsistencies (i.e., A > B, B > C, and C > A) because subjective criteria are applied by different people at different times. Some older methods such as those used for *AP-42* emission factors (U.S. EPA, 1995) rely on a ranking for each emission factor from A (best) to E (worst). No numerical uncertainty values are associated with each rating. Newer methods such as DARS (Beck et al., 1994) assign a numerical value to the quality of the various components of the emissions inventory and allow numerical manipulation of the uncertainty estimates of the system.

The DARS and other ranking methods are discussed below. Table 4.4-1 summarizes the preferred and alternative methods for ranking systems.

TABLE 4.4-1

PREFERRED AND ALTERNATIVE METHODS FOR RANKING SYSTEMS

Preferred	Provide DARS scores for all sources.
Alternative 1	Provide DARS scores for the largest sources (specify criteria used to identify "largest").
Alternative 2	Use a letter grade or numerical scheme to rank data quality; provide rules and rationale used to develop scores and make sure system is used consistently throughout the inventory.

4.1 DARS

The Data Attribute Rating System or DARS is currently under evaluation by the EIIP's Quality Assurance Committee (QAC). EPA originally developed DARS to assist in evaluating country-specific inventories of greenhouse gases. The system disaggregates

emission estimates into emission factors and activity data, then assigns a numerical score to each of these two components. Each score is based on what is known about the factor and activity parameters, such as the specificity to the source category, spatial (geographical) congruity, measurement of estimation techniques employed, and temporal congruity. The resulting emission factor and activity data scores are combined to arrive at an overall confidence rating for the inventory.

DARS defines certain classifying attributes that are believed to influence the accuracy, appropriateness, and reliability of an emission factor or activity and derived emission estimates. This approach is quantitative in that it uses numeric scores; however, scoring is based on qualitative and often subjective assessments. DARS also disaggregates specific attributes of the data and methods utilized in development of the inventory, thus providing perspective on the reason for the overall rating.

The DARS approach, when applied systematically by inventory analysts, can be used to provide a measure of the merits of one emission estimate relative to another. The proposed inventory data rating system cannot guarantee that an emission inventory with a higher overall rating is of better quality, or more accurate, or closer to the true value. The inventory with the higher overall rating is *likely* to be a better estimate, given the techniques and methodologies employed in its development.

An example of DARS scores for the architectural surface coatings area source category is shown in Table 4.4-2. Two alternative methods were used to estimate emissions from an urban area; one was based on a survey of paint distributors (conducted several years prior to the inventory) in the area, the other used a national per capita factor based on data from within one year of the inventory year. The more labor-intensive method gives a much higher overall DARS score. More information on considerations in using DARS scores for paints and coatings emission sources is presented in Appendix F.

EIIP members have recognized the potential utility of DARS for inventories at all levels. Among the proposed uses of DARS are:

- To identify the weakest areas of an inventory for further research and improvement;
- To use as one of several methods to quickly compare different inventories;
- To rank alternative emission estimation methods (the EIIP Area and Point Source Committees have used DARS as one of several tools to select the best method);
- To set DQO targets during the inventory planning stage; and

TABLE 4.4-2

DARS SCORES FOR ARCHITECTURAL SURFACE COATING EMISSIONS ESTIMATED BY TWO DIFFERENT METHODS

Attribute	Factor	Activity	Emissions
Local Survey			
Measurement/Method	0.7	0.9	0.63
Source Specificity	1.0	1.0	1.00
Spatial	1.0	1.0	1.00
Temporal	0.7	1.0	0.70
Composite	0.85	0.975	0.83
Per Capita Factor			
Measurement/Method	0.3	0.4	0.12
Source Specificity	1.0	0.3	0.30
Spatial	0.3	0.3	0.09
Temporal	0.7	1.0	0.70
Composite	0.575	0.5	0.30

- To provide a means of ranking inventories.

A more thorough discussion of the recommended EIIP approach for DARS is provided in Appendix F of this volume.

4.2 *AP-42* EMISSION FACTOR RATING SYSTEM

The U.S. EPA's *Compilation of Air Pollutant Emission Factors, AP-42,* is the primary reference for emission factors in the United States (U.S. EPA, 1995). Each *AP-42* emission factor is given a rating of A through E, with A being the best. A factor's rating is a general indication of the reliability, or robustness, of that factor. This rating is assigned using expert judgement. That judgement is based on the estimated reliability of the methods used to develop the factor, and on both the amount and the representative characteristics of the data.

In general, emission factors based on many observations, or on more widely accepted test procedures, are assigned higher rankings. Conversely, a factor based on a single observation of questionable quality, or one extrapolated from another factor for a similar process, is usually rated much lower. Because emission factors can be based on source tests, modeling, mass balance, or other information, factor ratings can vary greatly. In addition, there is a wide variation in the amount of QA to which each factor has been subjected.

Because the ratings do not consider the inherent scatter among the data used to calculate factors, the ratings do not imply statistical error bounds or confidence intervals about each emission factor. At most, a rating should be considered an *indicator* of the accuracy and precision of a given factor. This indicator is largely a reflection of the professional judgement of *AP-42* authors and reviewers concerning the reliability of any estimates derived with these factors.

Two steps are involved in factor rating determination. The first step is an appraisal of data quality or the reliability of the basic emission data that will be used to develop the factor. The second step is an appraisal of the ability of the factor to stand as a national annual average emission factor for that source activity. The *AP-42* rating system for the quality of the test data consists of four categories and is presented in Table 4.4-3.

The quality rating of *AP-42* data helps identify satisfactory data, even when it is not possible to extract a factor representative of a typical source in the category from those data. For example, the data from a given test may be good enough for a data quality rating of "A," but the test may be for a unique feed material, or the production specifications may be either more or less stringent than at the typical facility.

TABLE 4.4-3

***AP-42* RATING SYSTEM FOR EMISSIONS TEST DATA**

Rating	Description
A	Tests are performed by a sound methodology and are reported in enough detail for adequate validation.
B	Tests are performed by a generally sound methodology, but lacking enough detail for adequate validation.
C	Tests are based on an unproven or new methodology, or are lacking a significant amount of background information.
D	Tests are based on a generally unacceptable method, but the method may provide an order-of-magnitude value for the source.

The *AP-42* emission factor rating is an overall assessment of how good a factor is, based on both the quality of the test(s) or information that is the source of the factor and on how well the factor represents the emission source. Higher ratings are for factors based on many unbiased observations, or on widely accepted test procedures. For example, a factor based on 10 or more source tests on different randomly selected plants would likely be assigned an "A" rating if all tests are conducted using a single valid reference measurement method. Likewise, a single observation based on questionable methods of testing would be assigned an "E", and a factor extrapolated from higher-rated factors for similar processes would be assigned a "D" or an "E." A description of the *AP-42* emission factor quality ratings is given in Table 4.4-4.

The *AP-42* emission factor scores are of some value as indicators of the quality of emissions estimates. At best, they rate the quality of the original data as applied to estimates for that original point source. However, when applied to other sources or to groups of sources (i.e., area sources) the *AP-42* factor score is less meaningful because it does not consider how similar the original source and the modeled source(s) are, and it does not address the quality of the activity data at all.

4.3 OTHER GRADING SYSTEMS

A review of inventory quality rating systems was recently completed for the EPA (Saeger, 1994). Several systems similar to the *AP-42* system are described.

TABLE 4.4-4

AP-42 RATING SYSTEM FOR EMISSION FACTORS[a]

Ranking	Quality Rating	Discussion
A	Excellent	Factor is developed from A- and B-rated source test data taken from many randomly chosen facilities in the industry population. The source category population is sufficiently specific to minimize variability.
B	Above Average	Factor is developed from A- or B-rated test data from a "reasonable number" of facilities. Although no specific bias is evident, it is not clear if the facilities tested represent a random sample of the industry. As with an A rating, the source category population is sufficiently specific to minimize variability.
C	Average	Factor is developed from A-, B-, and/or C-rated test data from a reasonable number of facilities. Although no specific bias is evident, it is not clear if the facilities tested represent a random sample of the industry. As with the A rating, the source category population is sufficiently specific to minimize variability.
D	Below Average	Factor is developed from A-, B-, and/or C-rated test data from a small number of facilities, and there may be reason to suspect that these facilities do not represent a random sample of the industry. There also may be evidence of variability within the source population.
E	Poor	Factor is developed from C- and D-rated test data, and there may be reason to suspect that the facilities tested do not represent a random sample of the industry. There also may be evidence of variability within the source category population.

[a] Source: U.S. EPA, 1995.

A method used in Great Britain is based on letter ratings assigned to both emission factors and the activity data. The combined ratings are then reduced to a single overall score following an established protocol. The emission factor criteria for the letter scores are similar to those applied in the U.S. EPA's approach and scores for the activity data are based largely on the origin of the data. Published data either by a government agency or through an industry trade association are assigned C ratings and extrapolated data based on a surrogate would receive an E rating.

The Intergovernmental Panel on Climate Change (IPCC) uses a rating scheme in its guidelines for reporting of greenhouse gas emissions. The IPCC system incorporates an assessment of completeness and of overall data quality in a code. Table 4.4-5 shows the codes used for each of four characteristics. These codes are entered in an inventory review table (such as the one shown in Figure 2.4-1, Chapter 2 of this volume).

4.4 GEIA Reliability Index

The Global Emissions Inventory Activity (GEIA) group is a consortium of research institutions that is attempting to develop common data sets for use in developing global emissions inventories. Data are supplied to this group from many different sources. The person supplying the data is asked to categorize it into one of three reliability categories of <50 percent, 50-100 percent, or >100 percent that represent the estimated error in the data.

This categorization relies entirely in the subjective judgements of the data originator. If a system like this is used, the inventory developer should clearly define each category and provide a rationale for the assignment of each category. This type of approach may be most useful as a relative indicator for categories within a given inventory (particularly if used in combination with a qualitative assessment as described above). Without some standardization of the reliability category definitions, this method is not suitable for comparisons between inventories.

TABLE 4.4-5

DATA QUALITY CODES RECOMMENDED BY THE IPCC[a]

Estimates		Quality		Documentation		Disaggregation	
Code	Meaning	Code	Meaning	Code	Meaning	Code	Meaning
Part	Partly estimated	H	High confidence in estimation	H	High (all background information included)	1	Total emissions estimated
All	Full estimate of all possible sources	M	Medium confidence in estimation	M	Medium (some background information included)	2	Sectoral split
NE	Not estimated	L	Low confidence in estimation	L	Low (only emission estimates included)	3	Subsectoral split
IE	Estimated but included elsewhere						
NO	Not occurring						
NA	Not applicable						

[a] Source: Intergovernmental Panel on Climate Change (IPCC), 1995.

5

QUANTITATIVE UNCERTAINTY ANALYSIS

This section describes several methods for generating statistically based uncertainty estimates. They differ from previous methods in that they give quantitative, or numerical, estimates of the error associated with emission estimates. Table 4.5-1 summarizes the preferred and alternative methods for conducting quantitative uncertainty analysis. As discussed in the introduction to this chapter, the intended uses of the emissions data should be considered before spending significant resources on quantifying the uncertainty associated with the estimates.

TABLE 4.5-1

PREFERRED AND ALTERNATIVE METHODS FOR QUANTIFYING UNCERTAINTY

Preferred	Use expert judgment (based on as much data as are available) to estimate standard deviation (or coefficient of variation) and distribution for key variables for each source type or category. Conduct probabilistic modeling (e.g., Monte Carlo), accounting for dependencies between variables.
Alternative 1	Develop standard deviations (as above), assume independence, and use error propagation to estimate uncertainty limits.
Alternative 2	Use Delphi Method or other survey of experts to generate upper and lower bounds in estimates.

5.1 EXPERT ESTIMATION METHOD

In general, information on the distributional nature of emissions data is required for a quantitative analysis. These data include the type of distribution that best fits the data, and values of the key distribution parameters (i.e., mean or median and variance) are generally unavailable. Typically, no information is available to define the distribution of activity data as being normal, lognormal, or some other distribution, and there are no

estimates of mean or standard deviation of the parameter of concern. The most readily available source of data for use in emission uncertainty analysis is "expert judgement." Consequently, experts are asked to estimate key parameters associated with an emission inventory such as the qualitative lower and upper bounds of an emission estimate or the shape of a particular parameter distribution.

One approach is the highly formalized Delphi method (Linstene and Turoff, 1975) in which the opinion of a panel of experts working separately but with regular feedback converges to a single answer. The Delphi approach does not require an assumption of either independence or distribution of the underlying emissions data and is a very powerful technique when used properly and is focused on the "right" question. However, its capability is limited by the quality of the "experts" selected and the care with which the analysis protocol is followed. The work at the South Coast Air Quality Management District (SCAQMD, 1982) is an example of application of a simple Delphi technique to assess uncertainty in a large-scale inventory.

Expert judgement outside a formal Delphi framework is also used to estimate emissions uncertainty. In these methods, which can be relatively simplistic to highly structured, one or more experts make judgements as to the values of specific distributional parameters for a number of sources. For example, Horie (1988) used graphical techniques to estimate confidence limits once estimates of upper and lower limits of emissions were developed through expert judgement. Dickson and Hobbs (1989) applied three separate methods, including Horie's, to estimate the confidence limits for a number of source categories after developing estimates of the uncertainty parameters based upon questionnaires filled out by a panel of emission inventory experts.

Table 4.5-2 presents a portion of the results of Dickson and Hobbs. This table presents alternative estimates of uncertainty in VOC emissions in the San Joaquin Valley for 1985 using the lognormal method of Mangat et al. (1984), the probability method of Horie (1988), and the error propagation method as implemented by Benkovitz (1985), which is discussed in the next section. The estimates of uncertainty for emission factors and activity data for each individual source type were obtained through a polling of experts. Once these data were compiled and processed, a simple Monte Carlo simulation was used to estimate uncertainty in the entire inventory for the Mangat and Horie approaches. For the error propagation method, estimates of overall inventory uncertainty were obtained directly through error propagation. For most categories (only four are presented in the table), the lognormal and error propagation methods yield similar results with slightly larger differences produced using the probability method.

All three of these methods require the assumptions of independence of the activity data and emission factors, assumptions that are not often met. In addition, each method makes the explicit assumption of normality (or lognormality) of the emissions data. A

TABLE 4.5-2

COMPARISON OF VOC EMISSION UNCERTAINTY ESTIMATES DERIVED USING THREE ALTERNATIVE UNCERTAINTY ESTIMATION METHODS[a]

VOC Area Source Category	Median Emission Estimate (Med)			90% Upper Confidence Limit (UCL_{90})			Relative Percentage Difference[b] [(UCL_{90}-Med)/Med*100]		
	LN	P	EP	LN	P	EP	LN	P	EP
On-Road Motor Vehicles	30	29	29	35	41	36	17	41	22
Surface Coating	9.2	9.6	9.1	11	16	11	18	63	18
Pesticide Use	21	22	21	25	27	26	19	25	22
Oil Production	150	150	140	200	190	200	33	27	37
All Sources	210	210	200	260	260	260	23	21	27

[a] Source: Dickson and Hobbs, 1989. Emissions are for 1985 for the San Joaquin Valley and units are in thousands of tons per year.
[b] Computed prior to rounding median and upper confidence limits to two significant figures.

LN Lognormal Method of Mangat et al., 1984.
P Probability Method of Horie, 1988.
EP Error Propagation Method, Benkovitz, 1985.

consequence of a violation of any of these basic assumptions is that the uncertainty estimates that result are typically biased low. The fact that emissions data often violate these assumptions is a major weakness in most simple emission uncertainty estimation methodologies such as the three listed in Table 4.5-2. A strength of these methods, however, is their relatively low implementation cost when compared to the next two methods discussed in this section. In many circumstances, reasonable estimates of uncertainty for a multicounty or regional inventory can be developed for less than 2,000 hours of effort.

Note that these methods are different from the GEIA reliability index (and other ranking systems) discussed in the previous section. While all rely on expert judgement, the methods described in this section rely on sampling expert opinions and using that data to develop statistical indicators.

5.2 PROPAGATION OF ERROR METHOD

Error propagation methods follow traditional statistical methodology to estimate the composite error introduced by the joint action of a number of individual factors each with their own uncertainty. These error propagation methods are based upon the twin assumptions that:

- Emission estimates are equal to the product of a series of parameters; and
- Each of the parameters is independent (i.e., no temporal or spatial correlations among the parameters).

A good example of an error propagation analysis used to estimate emissions uncertainty in a large-scale emissions inventory is the National Acid Precipitation Assessment Program (NAPAP, 1991). For NAPAP, Benkovitz (1985) used a Taylor's series expansion of the equation describing the variance of a series of products to develop an analytic closed-form to an otherwise intractable problem. In particular, the assumption of independence allows the variance of multiplicative products to be expressed in terms of the individual variances. There is general agreement that the uncertainty in the NAPAP inventory is underestimated, in part because of the incorrect assumption of independence of the emission parameters used in the NAPAP error propagation analysis (EPA, 1986).

The IPCC proposes that this approach be used only when the ranges in the emission factor and uncertainty do not exceed more than 60 percent above or below the mean emission estimate. The uncertainty in each component (i.e., the factor and activity) is first established using classical statistical analysis (Chapter 3, Section 7 of this volume),

probabilistic modeling (described in next section), or the formal expert assessment methods (described in the previous section). Figure 4.5-1 presents an excerpt from the IPCC guidelines (IPCC, 1995). [Note that the nomenclature used in the IPCC example is not always consistent with EIIP terms. In particular, "point estimate" refers to the statistical concept of a single number that may be an average or an engineering judgement; it does not refer to "point sources" or single facilities.]

An example of the input data used in an error propagation analysis for the Grand Canyon Visibility Transport Commission emissions inventory is given in Table 4.5-3 (Balentine et al., 1994). In this study, Balentine et al. used expert judgement to estimate emissions uncertainty for all significant source classes and then developed refined estimates of uncertainty for nine source categories using error propagation methodology. Emissions from each source category were assumed to be a multiplicative function of the underlying emission parameters. For each of the nine categories for which refined uncertainty estimates were made, estimates of the coefficient of variation of each emission parameter contributing to the uncertainties were developed based upon the analysis of surrogate parameters and expert judgement. Including data acquisition (but not inventory development), this study required less than 300 staff hours to complete.

For the example source category given in Table 4.5-3 (industrial and commercial fuel combustion), the emission estimate, and hence uncertainty, was assumed to be a function of the number of sources, the distillate oil demand, the average sulfur content, and the variability in the *AP-42* emission factor. Application of the error propagation method with the data in Table 4.5-3 yields an overall composite coefficient of variation of approximately 40 percent, estimated as the square root of the sum of the square of the individual coefficients of variation. Again, this method requires the generally poorly met assumptions of independence and normal (lognormal) distribution of the individual emission parameters.

5.3 Direct Simulation Method

Direct simulation methods are statistical methods in which the uncertainty and confidence limits in emission estimates are directly computed using statistical procedures such as Monte Carlo (Freeman et al., 1986), bootstrap and related resampling techniques (Efron and Tibshirani, 1991), and Latin hypercube approaches (Iman and Helton, 1988). A major benefit of these statistical procedures is that the lack of independence in emission parameters is not a limitation. If set up and performed properly, the analysis methodology explicitly accounts for any dependencies as part of the statistical formulation of the problem.

TABLE A1-1
UNCERTAINTIES DUE TO EMISSION FACTORS AND ACTIVITY DATA

1	2	3	4	5
Gas	**Source category**	**Emission factor** U_E	**Activity data** U_A	**Overall uncertainty** U_T
CO_2	Energy	7%	7%	10%
CO_2	Industrial Processes	7%	7%	10%
CO_2	Land Use Change and Forestry	33%	50%	60%
CH_4	Biomass Burning	50%	50%	100%
CH_4	Oil and Nat. Gas Activities	55%	20%	60%
CH_4	Coal Mining and Handling Activities	55%	20%	60%
CH_4	Rice Cultivation	3/4	1/4	1
CH_4	Waste	2/3	1/3	1
CH_4	Animals	25%	10%	25%
CH_4	Animal waste	20%	10%	20%
N_2O	Industrial Processes	35%	35%	50%
N_2O	Agricultural Soils			2 orders of magnitude
N_2O	Biomass Burning			100%

Note: Individual uncertainties that appear to be greater than ± 60% are not shown. Instead judgement as to the relative importance of emission factor and activity data uncertainties are shown as fractions which sum to one.

A1.2 Procedures for Quantifying Uncertainty

Estimating Uncertainty of Components

To estimate uncertainty by source category and gas for a national inventory, it is necessary to develop information like that shown in Table A1-1, but specific to the individual country, methodology and data sources used. In scientific and process control literature the 95 per cent (± 2) confidence limit is often regarded as appropriate for range definition. Where there is sufficient information to define the underlying probability distribution for conventional statistical analysis, a 95 per cent confidence interval should be calculated as a definition of the range. Uncertainty ranges can be estimated using classical analysis (see Robinson) or the Monte Carlo technique (in Eggleston, 1993). Otherwise the range will have to be assessed by national experts.

FIGURE 4.5-1. RECOMMENDED IPCC PROCEDURES FOR QUANTIFYING UNCERTAINTY (SOURCE: IPCC, 1995)

If possible ranges should be developed separately for

- emission factors (and other assumptions in the estimation method) (column 3 of Table A1-1).
- socio-economic activity data (column 4 of Table A1-1)

Combining Uncertainties

It is necessary to derive the overall uncertainty arising from the combination of emission factor and activity data uncertainty. IPCC/OECD suggest that emission factor and activity data ranges are regarded as estimates of the 95 per cent confidence interval, expressed as a percentage of the point estimate, around each of two independent components (either from statistically based calculations or informal *ex ante* judgements).

On this interpretation (for quoted ranges extending not more than 60 per cent above or below the point estimate) the appropriate measure of overall *percentage* uncertainty U_T for the emissions estimate would be given by the square root of the sum of the squares of the *percentage* uncertainties associated with the emission factor (U_E) and the activity data (U_A). That is, for each source category:

$$U_T = \pm\sqrt{(U_E^2 + U_A^2)}; \text{ so long as } |U_E|, |U_A| < 60\% \text{ [1]}$$

For individual uncertainties greater than 60 per cent the sum of squares procedure is not valid. All that can be done is to combine limiting values to define an overall range, though this leads to upper and lower limiting values which are asymmetrical about the central estimate[2].

Estimated total emission for each gas is of course the summation ΣC_i where C_i is the central estimate of the emission of the gas in the source category. The appropriate measure of *uncertainty* in total emissions in emissions units (not percentages) is then:

$$E = \pm(1/100)\cdot\sqrt{(\Sigma U_{T,i}^2 \cdot C_i^2)}$$

where $U_{T,i}$ is the overall percentage uncertainty for the source category of the gas from Table A1-1. Source categories for which symmetrical limiting values cannot be defined (because $|U_E|$ or $|U_A|$ exceeds 60 per cent) cannot sensibly be treated in this way. The uncertainty might be handled by reporting that total emissions from gas X are estimated to be Y Mt, of which Y_1 Mt had an estimated uncertainty of ± E_1 Mt and Y_2 Mt had a range of uncertainty between - L Mt and + U Mt.

---

[1] The 60% limit is imposed because the rule suggested for U_T requires σ to be less than about 30% of the central estimate, and we are interpreting the quoted range as $\pm 2\sigma$

[2] If uncertainties due to the emission factor and the activity data are ± E% and ± A% respectively, and the upper and the lower limits of overall uncertainty are U% and L% respectively, then U% = (E+A+E·A/100) and L% = (E+A-E·A/100).

FIGURE 4.5-1. CONTINUED

A1.3 Implications

If the assumptions in Table A1.1 are correct then typical uncertainties in national emissions estimates range between:

- ± 10% for CO_2 from fossil fuels although this may be lower for some countries with good data and where source categories are well defined (IPCC, 1993; von Hippel et al., 1993)
- ± 20% and ± 100% for individual methane sources (though the overall error might be ± 30%)
- perhaps two orders of magnitude for estimates of nitrous oxide from agricultural soils

These uncertainties will affect the level of quantitative understanding of atmospheric cycles of greenhouse gases that can be derived using the summation of inventories.

The situation is less critical for monitoring emissions mitigation options, because the profile of the emissions time series will be relatively insensitive to revisions to the emissions estimation methodology. However very different levels of uncertainty for different gases will be inevitable for some time to come, and this will need to be recognised in any move towards a comprehensive approach to greenhouse gas mitigation.

A1.4 References

(IPCC) Intergovernmental Panel on Climate Change (1992), *Climate Change 1992: The Supplement to the IPCC Scientific Assessment.*

The method for combining errors in a multiplicative chain are given in many statistical textbooks, but note Jennifer Robinson's discussion (On uncertainty in the computation of global emissions from biomass burning, *Climatic Change*, 14, 243-262) about the difficulties which arise at high coefficients of variation.

H S Eggleston (1993), "Uncertainties in the estimates of emissions of VOCs from Motor Cars." Paper presented at the *TNO/EURASAP Workshop on the Reliability of VOC Emission Databases*, June 1993, Delft, The Netherlands.

IPCC (1993), "Preliminary IPCC national GHG inventories: in depth review." Report presented at the *IPCC/OECD Workshop on National GHG Inventories*, October 1993, Bracknell, UK.

von Hippel et al. (1993), "Estimating greenhouse gas emissions from fossil fuel combustion", *Energy Policy*, 691-702, June 1993.

FIGURE 4.5-1. CONTINUED

TABLE 4.5-3

ESTIMATED COEFFICIENT OF VARIATION FOR PARAMETERS USED IN ESTIMATING SO_2 EMISSIONS FROM INDUSTRIAL AND COMMERCIAL SOURCES[a]

Parameter	Source	Discussion	Coefficient of Variation (%)
Number of industrial and commercial sources	Dickson et al., 1992	Variation in day-specific emissions from annual day emissions for 33 facilities in Wisconsin.	15
Distillate oil demand	Oil and Gas Journal, 1992, 1993, and 1994	1992-1994 quarterly variability in nontransportation distillate fuel demand.	5
Distillate oil average sulfur content	El-Wakil, 1984	Average sulfur content for No. 2 and No. 6 fuel oils.	25
Emission factor variability	Assumption based on *AP-42*, Table 1.3-1; EPA, 1985	*AP-42* uncertainty in emission factor given as an "A" rating.	20

[a] Source: Balentine et al., 1994.

The common Monte Carlo technique is a powerful direct simulation method. Freeman et al. (1986) applied this technique to evaluate uncertainty in the input parameters, including emissions, in an air quality model. Environment Canada (1994) applied the methodology to estimate uncertainty in greenhouse gas emissions for Canada. Table 4.5-4 presents the Environment Canada Monte Carlo results for carbon dioxide (CO_2) emissions in Canada. The estimated uncertainty in emissions in individual source categories varied from 5 percent for diesel fuel combustion to 40 percent for coal combustion. Because the Environment Canada study made the assumption of independence between parameters, the resultant uncertainty estimates should be considered lower limits.

One limitation of the Monte Carlo approach is that a distribution type and distribution values for each emission parameter must be specified. Typically, expert judgement is required to make some or all the estimates of distribution type and parameters. A second limitation (but also a strength because it gets around the assumption of independence) is that all underlying dependencies between the various parameters must be accounted for when formulating the model. These dependencies can be taken into account during randomization of each parameter because the same random number can

TABLE 4.5-4

ESTIMATES OF UNCERTAINTY OF CO_2 EMISSIONS IN CANADA PRELIMINARY 1990 CO_2 ESTIMATES IN KILOTONNES[a]

Source Group	At 95% Confidence Level		
	Range of Emissions (Rounded) From/To	Range Width	Overall Uncertainty (± %)
Industrial Processes			
• Cement Process Only	4,700/6,000	1,300	(12)
• Lime & Other Inorganics	1,900/2,600	700	(15)
• Stripped Natural Gas	5,100/7,300	1,800	(18)
• Non Energy Use	9,900/18,400	8,500	(30)
Subtotal	23,350/32,254	8,903	(16)
Fuel Combustion			
• Power Generation	89,000/99,000	10,000	(5)
• Residential	38,000/43,000	5,000	(6)
• Commercial	22,500/25,500	3,000	(6)
• Industrial/Steam	72,500/85,500	13,000	(8)
• Agriculture	2,280/2,680	400	(8)
• Public Administration	1,900/2,200	300	(8)
• Refinery Use	12,500/18,700	6,200	(20)
• Oil & Gas Production	22,700/34,000	11,300	(20)
• Pipeline	6,000/7,400	1,400	(10)
• Coal	240/560	320	(40)
• Miscellaneous	200/500	300	--
Subtotal	282,409/304,074	21,666	(4)
Transportation			
• Gasoline	73,900/81,600	7,700	(5)
• Jet Fuels	11,900/14,500	2,600	(10)
• Diesel	44,900/49,700	4,800	(5)
• Natural Gas & Propane	1,500/1,900	400	(10)
Subtotal	135,024/144,540	9,515	(3.5)
Overall Total	448,185/473,467	25,283	(3)

[a] Source: Table 4.2.1-1 in Environment Canada, 1994.

be used to estimate multiple parameters that are correlated (e.g., if population is a factor in multiple emission parameters, the same random number can be used to estimate each factor that is dependent on population rather than allowing use of independent random numbers).

The Latin hypercube methodology (Iman and Helton, 1988) has evolved from the Latin square methodology (Cox, 1958) commonly used for planning and analyzing field measurements. A typical application of the Latin square approach would be an agricultural experiment examining the role of application of *n* alternative pesticides to *m* plant varieties in an effort to determine the combination that maximizes crop yield.

In the Latin hypercube approach, the methodology is expanded beyond the simple two-dimensional relationship to higher dimensions with multiple parameters and potential interactions. In an example involving emission estimates, the data set may consist of stack test measurements of emissions from industrial boilers and three of the variables included are the boiler type, load, and control device. Using the Latin hypercube approach, the internal relationship between load, boiler type, and control device would be approximated by the various random samples selected from the data set. This approach allows one to estimate directly the uncertainty of the parameter of interest (i.e., emission rate) as a function of the causative factors examined.

The numerical method developed by Oden and Benkovitz (1990) allows one to estimate uncertainty in the typical situation in which autocorrelations and covariances that occur in the parameters responsible for producing emissions. Using their methodology, it is possible to estimate the uncertainty accounting for the lack of independence between the parameters and to determine what this lack of independence contributes to the overall uncertainty.

Resampling methodologies such as the bootstrap method (Efron and Tibshirani, 1991) involve performing random sampling (with replacement) from a data set in order to estimate some statistical parameter such as the standard error or variance. For a small data set, a direct computation of the parameter of interest can be highly uncertain given the small sample size or may not even be possible because there is no simple formula with which to compute the value. However, in bootstrap and other resampling methods, resampling with replacement both increases the effective size of the data set and allows direct estimation of parameters of interest. While there are difficulties in applying resampling techniques to emissions data because they are temporally correlated, recent work by Carlstein (1992) has allowed bootstrap techniques to be applied in situations with temporally correlated data.

The major drawback of the direct simulation methodologies is the computationally intensive nature of the techniques. However, as computing costs decrease with the

advent of increasingly more powerful desktop computers, this limitation is becoming less important as a selection criteria for an uncertainty estimation methodology. Because of the complexity of the statistical analyses required, staff members with advanced degrees in statistics are typically involved in studies using direction simulation methods to estimate uncertainty. Also, the level of effort required can approach (or exceed) 1,000 staff hours depending upon the complexity of the analysis and the data collection required.

5.4 OTHER METHODS

In addition to the above methods, direct and indirect field measurement, receptor modeling, and inverse air quality modeling can be used to produce estimates of uncertainty (or relative uncertainty) in emission inventories. However, such methods can provide significantly more information than estimates of uncertainty. Each can produce emission estimates completely independent of standard emission computation methods. Typically, these other methods are very labor and data intensive, and can easily require 1,000 staff hours or more to collect the required data and perform the analysis.

Because these emission estimates are independent, they can be used as an independent verification of the emission estimates. This potential role in validating emission estimates is perhaps the most important use of information resulting from application of these methods. A detailed discussion of each method is given in Chapter 3, Section 9 of this volume.

6

REFERENCES

Balentine, H.W., R.J. Dickson and W.R. Oliver. 1994. *Development of Uncertainty Estimates for the Grand Canyon Visibility Transport Commission Emissions Inventory.* Radian Corporation Technical Memorandum, Sacramento, California. December.

Beck, L., R. Peer, L. Bravo, and Y. Yan. 1994. *A Data Attribute Rating System.* Presented at the Air & Waste Management Association's Emission Inventory Conference, Raleigh, North Carolina.

Benkovitz, C.M. 1985. *Framework for Uncertainty Analysis of the NAPAP Emissions Inventory.* U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory. EPA/600/7-85/036. Research Triangle Park, North Carolina.

Benkovitz, C.M., and N.L. Oden. 1989. Individual Versus Averaged Estimates of Parameters used in Large Scale Emissions Inventories. *Atmospheric Environment* 23:903-909.

Carlstein E. 1992. Resampling techniques for stationary time series: Some recent developments. In: *The IMA Volumes in Mathematics and its Applications: New Directions in Time Series Analysis I.* Springer-Verlag, New York.

Chang, M., C. Cardelino, W. Chameides, and W. Chang. 1993. *An iterative procedure to estimate emission inventory uncertainties.* Presented at The Regional Photochemical Measurement and Modeling Studies Meeting of the Air & Waste Management Association at San Diego, California. November.

Chang, M., C. Cardelino, D. Hartley, and W. Chang. 1995. *Inverse Techniques for Developing Historical and Future Year Emission Inventories.* Presented at The Emission Inventory: Programs and Progress Specialty Conference of the Air & Waste Management Association, Research Triangle Park, North Carolina. October 11-13.

Chow, J.C., J.G. Watson, D.H. Lowenthal, P.A. Soloman, K.L. Magliano, S.D. Ziman, and L.W. Richards. 1992. PM_{10} Source Apportionment in California's San Joaquin Valley. Atmospheric Environment 26A:3335-3354.

Claiborn, C., A. Mitra, G. Adams, L. Bamesberger, G. Allwine, R. Kantamaneni, B. Lamb, and H. Westberg. 1995. Evaluation of PM_{10} Emission Rates from Paved and Unpaved Roads using Tracer Techniques. *Atmospheric Environment* 29:1075-1089.

Cox, D.R. 1958. *Planning of Experiments.* John Wiley and Sons, Inc., New York, pp. 35-44.

Dickson, R.J., and A.D. Hobbs. 1989. *Evaluation of Emission Inventory Uncertainty Estimation Procedures.* Paper 89-24.08, presented at the 82nd Annual Meeting of the Air & Waste Management Association, Anaheim, California.

Dickson, R.J., K.K. Mayenkar, and J. Laas. 1992. *Increasing the Accuracy of Hourly Emission Estimates Using Day-Specific Data in the Lake Michigan Ozone Study.* Paper 92-87.04, presented at the 85th Annual Meeting of the Air & Waste Management Association, Kansas City, Missouri.

El-Wakil, M.M. 1984. *Power Plant Technology.* McGraw-Hill, Inc., New York, New York. p. 149.

Efron, B., and R. Tibshirani. 1991. Statistical Data Analysis in the Computer Age. *Science* 253:390-395.

Finkel, A.M. 1990. *Confronting Uncertainty in Risk Management - A Guide for Decision-Makers.* Center for Risk Management, Resources for the Future, Washington, D.C.

Freeman, D.L., R.T. Egami, N.F. Robinson, and J.G Watson. 1986. A Method for Propagating Measurement Uncertainties through Dispersion Models. *Journal of the Air Pollution Control Association* 36:246-253.

Fujita, E.M., B.E. Croes, C.L. Bennett, D.R. Lawson, F.W. Lurmann, and H.H. Main. 1992. Comparison of Emission Inventory and Ambient Concentration Ratios of CO, NMOG, and NO_x in California's South Coast Air Basin. *Journal of the Air & Waste Management Association 42(3):264-276.*

Gatz, D.F., and L. Smith. 1995a. The Standard Error of a Weighted Mean Concentration - I. Bootstrapping vs Other Methods. *Atmospheric Environment* 29:1185-1193.

Gatz, D.F., and L. Smith. 1995b. The Standard Error of a Weighted Mean Concentration - I. Estimating Confidence Intervals. *Atmospheric Environment* 29:1195-1200.

Hartley, D., and R.J. Prinn. 1993. Feasibility of Determining Surface Emissions of Trace Gases Using an Inverse Method in a Three-Dimensional Chemical Transport Model. *Journal of Geophysical Research.* 98:5183-5197.

Horie, Y. 1988. *Handbook on Procedures for Establishing the Uncertainties of Emission Estimates.* Prepared for the California Air Resources Board Research Division, by Valley Research Corporation, ARB Contract A5-184-32. Sacramento, California. February.

Horie, Y., and A.L. Shorpe. 1989. *Development of Procedures for Establishing the Uncertainties of Emission Estimates.* Paper 89-24.7. Presented at the 82nd Annual Air & Waste Management Association Meeting, Anaheim, California, June.

Iman, R.L., and J.C. Helton. 1988. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Analysis* 8:71-90.

Intergovernmental Panel on Climate Change (IPCC). 1995. *IPCC Guidelines for National Greenhouse Gas Reporting Instructions.* IPCC WGI Technical Inventories, Volume 1: Support Unit, Bracknell, United Kingdom.

Linstene, H.A., and M. Turoff. 1975. *The Delphi Method, Techniques and Applications.* Edited by Linstene and Turoff. Addison - Wesley Publishing Company.

Lowenthal, D.H., J.C. Chow, J.G. Watson, G.R. Neuroth, R.B. Robbins, B.P. Shafritz, and R.J. Countess. 1992. The Effects of Collinearity in the Ability to Determine Aerosol Contributions from Diesel- and Gasoline-Powered Vehicles using the Chemical Mass Balance Model. *Atmospheric Environment* 26A:2341-2351.

Mangat, T.S., L.H. Robinson, and P. Switzer. 1984. *Medians and Percentiles for Emission Inventory Totals.* Presented at the 77th Annual Meeting of the Air Pollution Control Association, San Francisco, California. June.

Mitchell, W.J., J.C. Suggs, and E.W. Streib. 1995. A Statistical Analysis of Stationary Source Compliance Test Audit Data. *Journal of the Air & Waste Management Association* 45:83-88.

Mulholland, M., and J.H. Seinfeld. 1995. Inverse Air Pollution Modelling of Urban-Scale Carbon Monoxide Emissions. *Atmospheric Environment* 29:497-516.

National Acid Precipitation Assessment Program. 1991. *Acidic Deposition: State of Science and Technology; Volume I - Emissions, Atmospheric Processes, and Deposition.* ISBN 0-16-036144-3. P.M. Irving, ed. Washington, D.C.

Oden, N.L., and C.M. Benkovitz. 1990. Statistical Implications of the Dependence Between the Parameters Used for Calculations of Large Scale Emissions Inventories. *Atmospheric Environment* 24A:449-456.

Oil and Gas Journal. Annual Forecast and Review, July 27, 1992, p. 63; July 26, 1993, p. 51, July 25, 1994, p. 55.

Peer, R.L., and D.L. Epperson, D.L. Campbell., and P. von Brook. 1992. *Development of an Empirical Model of Methane Emissions from Landfills.* United States Environmental Protection Agency, Air and Energy Engineering Laboratory, EPA-600/R-92-037, Research Triangle Park, North Carolina.

Pierson, W.R., A.W. Gertler and R.L. Bradow. 1990. Comparison of the SCAQS Tunnel Study with Other On-Road Vehicle Emission Data. *Journal of the Air & Waste Management Association.* 40(11):1495-1504.

Saeger, M. 1994. *Procedures for Verification of Emissions Inventories, Final Report.* Prepared for Emissions Inventory Branch, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina (EPA Contract No. 68-D3-0030).

Scheff, P.A., R.A. Wadden, D.M. Kenski, and J. Chang. 1995. *Receptor Model Evaluation of the SEMOS Ambient NMOC Measurements.* Paper 95-113C.03. Presented at the 88th Annual Meeting of the Air Pollution Control Association, San Antonio, Texas. June.

South Coast Air Quality Management District (SCAQMD). 1982. *Uncertainty of 1979 Emissions Data, Chapter IV: 1983 Air Quality Management Plan Revision, Appendix IV-A; 1979 Emissions Inventory for the South Coast Air Basin, Revised October 1982.* Planning Division, El Monte, California. October.

Spellicy, F.L., J.A. Draves, W.L. Crow, W.F Herget, and W.F. Buchholtz. 1992. *A Demonstration of Optical Remote Sensing in a Petrochemical Environment.* Presented at the Air & Waste Management Association Specialty Conference on Remote Sensing, Houston, Texas. April 6-9.

Steiner, C.K.R., L. Gardner, M.C. Causley, M.A. Yocke, and W.L. Steorts. 1994. Inventory Quality Issues Associated with the Development of an Emissions Inventory for the Minerals Management Service Gulf of Mexico Air Quality Study. In: *The Emission Inventory: Perception and Reality.* Proceedings of an International Specialty Conference, VIP-38. Air & Waste Management Association, Pittsburgh, Pennsylvania.

Environment Canada. *Uncertainties in Canada's 1990 Greenhouse Gas Emission Estimates, A Quantitative Assessment.* Prepared by T.J. McCann & Associates. Unpublished report. Ottawa, Ontario, March, 1994.

U.S. EPA. 1986. *Estimation of Uncertainty for the 1980 NAPAP Emission Inventory.* U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory, EPA/600/S7-86/055. Research Triangle Park, North Carolina.

U.S. EPA. 1985-1993. *Compilation of Air Pollutant Emission Factors, Volume I: Stationary Point and Area Sources, Fourth Edition and Supplements, AP-42.* U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. Research Triangle Park, North Carolina.

U.S. EPA. 1995. *Compilation of Air Pollutant Emission Factors, Volume I: Stationary Point and Area Sources, Fifth Edition, AP-42.* U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. Research Triangle Park, North Carolina.

Watson, J.G., J.A. Cooper, and J.J. Huntzicker. 1984. The Effective Variance Weighing for Least Squares Calculations Applied to the Mass Balance Receptor Model. *Atmospheric Environment* 18:1347-1355.

This page is intentionally left blank.

APPENDIX F
EIIP RECOMMENDED APPROACH TO USING THE DATA ATTRIBUTE RATING SYSTEM (DARS)

This page is intentionally left blank.

DARS BASICS

The Data Attribute Rating System (DARS) was originally developed as a research tool for rating national and global greenhouse gas inventories. The theoretical basis of DARS is described in Beck et al., 1994. EIIP has made some changes in the original system based in part on the results of several pilot studies. State agency personnel were trained in the DARS method, and then used DARS to rate their base year State Implementation Plan (SIP) ozone precursor inventories. In addition, particulate matter (PM-10) inventories (state and national levels) were evaluated by inventory developers trained in the use of DARS. The experiences and recommendations of field testers were incorporated in the version of DARS presented here. Key changes from the original are:

1. Rating criteria have been expanded to include point and mobile source emission estimation methods. The original DARS was developed for area source-type methods.

2. The definitions of the attributes have been made more specific. In particular, the full range of emission estimation methods and source types found in a state or regional inventory have been taken into account.

3. The assignment of scores within an attribute have been made less flexible. It is important that the scoring system not be too rigid because the inherent uncertainty in emissions varies among source types. Therefore, a method that is considered poor in most cases may actually produce very good estimates in certain other cases. An example is the use of mass balance. If the emissive process is the result of complex chemical reactions, mass balance produces a rough approximation. If the process is a simple physical one (e.g., evaporation), mass balance is a much more acceptable method.

4. The original DARS had five attributes, the EIIP version has four. Two attributes--measurement and pollutant specificity--were combined. This change actually improves the discriminating power of DARS because the pollutant specificity attribute was nearly always the same value in SIP-type inventories.

The DARS score is based on the perceived quality of the emission factor and activity data. Scores are assigned to four data attributes: measurement/method, source specificity, spatial congruity, and temporal congruity. A key feature of DARS is that these attributes are orthogonal; that is, they are independent of each other, and therefore the score for each attribute is independent of the other scores. However, the emission factor and activity scores for a given attribute are not necessarily independent. This is because the choice of one is

usually limited by the selection of the other. For example, if a per capita factor is being used to estimate architectural surface coating emissions, then the activity must be population.

Table F-1 shows a DARS scoring box. The procedures for filling in the scores for emission factors and activity are described below. The emissions scores for each attribute (i.e., the right-hand column of the box) are computed by first dividing each score by 10, and then multiplying the factor score times the activity score. The composite scores for factor, activity, and emissions (i.e., the bottom row of the box) are computed by averaging the scores in a column. Scoring of each attribute is discussed below with specific examples. In general, the following guidelines should be used:

1. The specific scores and descriptions shown in the attribute scoring flow charts (Figures F-1 through F-8) are to be used as set-points. Users can interpolate between the values shown.

2. The scores are shown on a 1 to 10 basis, although the final scores are always less than 1 because the scores are divided by the maximum possible score of 10. In general, it is easier to think and talk in terms of 1 to 10, so that convention is used in the following descriptions and examples. However, the composite scores shown are always presented as fractions.

3. For the beginner, a good approach to selecting a score is to start at the beginning of the flow chart and work down to find the lowest number that most nearly fits the situation. Then adjust up to factor in other considerations (examples are given in later sections).

4. In the absence of sufficient information on the derivation of factors, activity, or emissions, choose the highest score that can be confidently made with the information provided. If the source or derivation of the data is totally undocumented, the highest possible score is 1. (One objective of DARS is to encourage good documentation of inventory data.)

DARS Scores Using Statistical Correlations

Many of the DARS attributes are scored based on presumed correlations between the target category and a surrogate. Unfortunately, very few of these correlations have been demonstrated statistically. If a statistical correlation is available, the correlation coefficient (usually expressed as r or sometimes R) can be used to help determine the DARS score. However, statistical correlations should be used very carefully. The data should apply directly to the region and source category being scored. Also, the data should be adequate and a representative sample should be chosen.

TABLE F-1

DARS SCORING BOX

Attribute	Factor	Activity	Emissions
Measurement/Method	e_1	a_1	$e_1 * a_1$
Source Specificity	e_2	a_2	$e_2 * a_2$
Spatial Congruity	e_3	a_3	$e_3 * a_3$
Temporal Congruity	e_4	a_4	$e_4 * a_4$
Composite	$\frac{\sum_{i=1}^{4} e_i}{4}$	$\frac{\sum_{i=1}^{4} a_i}{4}$	$\frac{\sum_{i=1}^{4} (e_i * a_i)}{4}$

The spatial and temporal attributes deal with scaling issues (in part). For example, many area source emission factors are based on annual national consumption that is then apportioned using population or employment. If the inventory uses daily emissions in a county, uncertainty is introduced by scaling down. If the activity and emissions are very uniform, then the uncertainty is low (and the DARS score relatively high). But many emissive activities vary in nature and importance geographically; in this case, using a national factor (or a mean value) will result in over- and underestimates of emissions at a small scale (i.e., at the county level).

Note that the same spatial concerns apply when scaling up. If the emissions from a small number of facilities are used to estimate emissions for the entire region, the representativeness of those sources in the entire population is important.

No formal relationship between DARS attribute scores and statistical variability or correlation measures has been developed. Unfortunately, it has not been the practice to publish statistical measures of emission factor variability in the past, although this is changing.

ASSIGNING ATTRIBUTE SCORES

Measurement/Method Attribute

The key to correctly scoring this attribute is to remember that it deals explicitly with measurement. The score is based on the quality of the factor itself--not on how it has been used (that is covered in the next section under source specificity). The presumption is that the best results are usually obtained by direct measurement of either emissions (either by source testing or continuous emission monitors [CEMs]) or by measurement of surrogate parameters that have a strong, statistically documented correlation with the pollutant of interest. The term "factor" is appropriate even when source testing was used because emission measurement data are usually expressed per unit of time. If a concentration is measured, the emissions per unit of time must be calculated for use in an inventory, or the original data may be expressed based on fuel consumed (or other variable). Figures F-1 and F-2 show the flow chart decision process used to score this attribute.

Very often, *AP-42* or other emission factors are used to estimate emissions. If possible, the appropriateness of the test data used to develop the factor should be studied to determine the DARS score. Alternatively, the default DARS scores for *AP-42* factors shown in Table F-2 can be used for point source estimates.

Area source emission factors are treated the same as point source factors when scoring this attribute. It is very unlikely that an area source emission factor will receive a score of 10 for this attribute. A 9 is possible if a large number of samples covering a representative portion of the source were used to develop the factor.

Some additional comments are warranted for emission factors based on mass balance. As seen in Figure F-1, this method can get a score varying from 3 to 5 depending on the source types and thoroughness. However, the score may be pushed even higher for some types of sources and if endpoints (other than air) have been fully quantified. For example, evaporative losses from solvent use can be reliably estimated using this method, provided that accounting for all the nonemissive losses is done. It seems reasonable to assume that volatile compounds will evaporate. The problem is that for surface coatings or graphic arts, some solvent may remain in the substrate. Some solvent may also be released to publicly owned treatment works (POTWs) or, if released inside a building, it may be absorbed by living tissue (e.g., plants or lungs of animals). For all of those reasons, the scorer is allowed to exercise some judgment. If there is some empirical basis for the mass balance factor (and especially if some of these other sinks for the solvents have been included in some way), the score can be raised to a 5 for area sources; higher scores may be given for point sources.

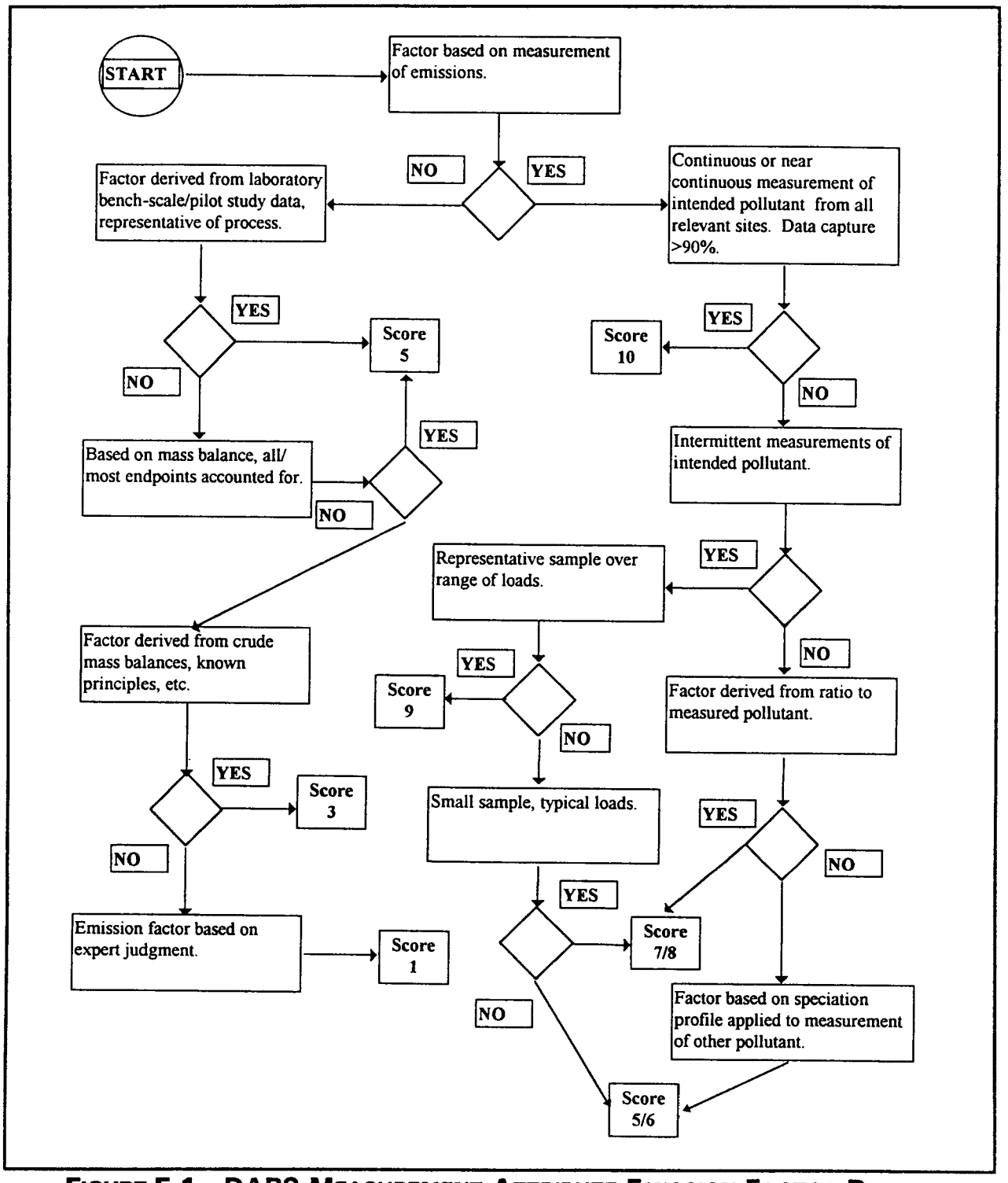

FIGURE F-1. DARS MEASUREMENT ATTRIBUTE EMISSION FACTOR RATING FLOW CHART

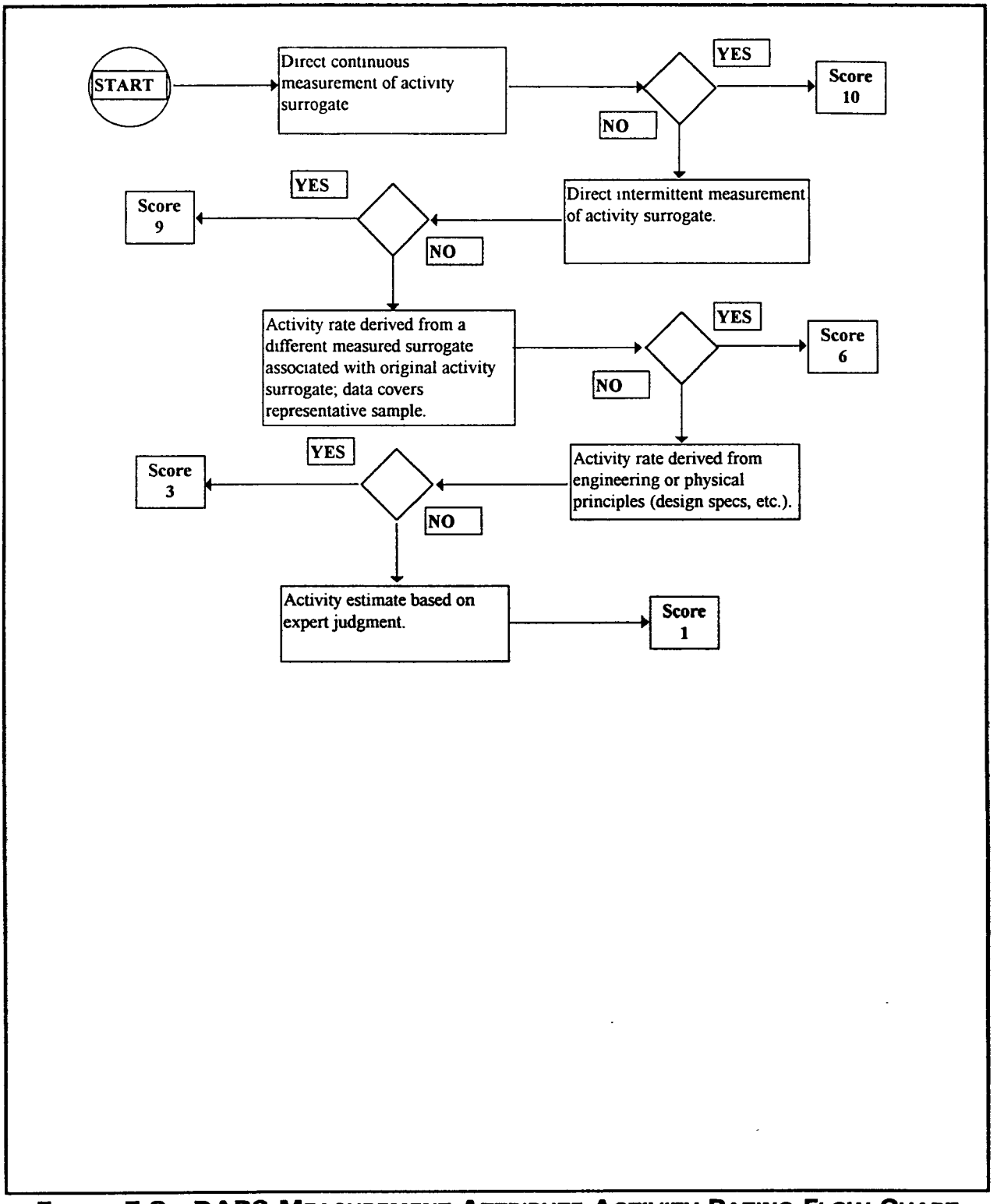

FIGURE F-2. DARS MEASUREMENT ATTRIBUTE ACTIVITY RATING FLOW CHART

TABLE F-2

AP-42 LETTER CODES AND CORRESPONDING DARS FACTOR MEASUREMENT ATTRIBUTE SCORES

AP-42 Factor Rating	Pollutant Factor			
	NO_x	CO	VOC	PM-10
A	6	6	5	5
B	6	6	5	5
C	5	5	4	4
D	5	5	4	4
E	4	4	3	3

A 10 will rarely be given for the emission factor measurement score; however, they will be fairly common for the activity measurement score in point source inventories. For example, fuel use by a boiler is usually known for an industrial site and, assuming no uncertainty or gaps in the data, will receive a 10. Total county-wide fuel use by small boilers may not be directly measured, or may be difficult to obtain. A common source of state-level data is the *State Energy Data Report Consumption Estimates* (published annually by the U.S. Department of Energy or DOE); the methods used to compile these data are discussed in the technical appendices in that volume, and some known sources of errors are acknowledged. Generally, the values are based on either sales data (which is a surrogate, so this gets a score of 6 if the correlation is good), or shipments (by weight or volume) that might be construed as a direct intermittent measurement (and assigned a score of 8 or 9). The correct score will generally fall between 5 and 9 for this example.

If the oil is being consumed by industry other than in boilers, it is probably for heaters or other combustion devices. If no adjustment is made for these other uses, the DARS score is a 7. If the other uses have been subtracted from the total (or are known not to be important), the potential score can be raised to 8. If the DOE *State Energy Data Report Consumption Estimates* is the source of fuel oil consumption, then the highest possible score is an 8 given the uncertainties in the DOE method.

The two other fuels commonly included in the combustion source categories are natural gas and coal. The distinction between the industrial/commercial sectors used by the gas industry is not consistent with the definitions used by EPA (the gas industry definition is not based on Standard Industrial Classification [SIC] codes). Unless adjustments have been made to the data (based on state information, for example), the DARS score is a 5 for industrial or commercial natural gas combustion. Coal use by industry falls somewhere in between the natural gas and fuel oil DARS sources. The DOE reports that these are the most uncertain numbers because it is difficult to track at a state level. So, there is the potential for error in allocating national coal consumption to states (this comes into play when scoring the spatial congruity attribute). The allocation to industrial uses (versus commercial or residential) is pertinent to the measurement attribute; it is also difficult to track, but because very little coal is used by any sector other than utilities and industry, it is generally safe to assume that nonutility users are primarily industrial. The best possible score for coal is a 7.

Source Specificity

The source specificity attribute concerns how specific the original factor or activity surrogate is to the source being estimated. This attribute is easily confused with the previous one. The key point to remember is NOT to be concerned with whether or not the emission factor or activity is measured; the question to ask is "was this emission factor (or activity parameter) specifically developed for this source category?" To answer this question will require a clear definition of the source category and a good understanding of the source of the emission factor and activity parameter. Figures F-3 and F-4 provide the details needed to score this attribute.

It is common practice to borrow emission factors from similar processes if none are available for the intended source category. For example, no emission factors are available for small industrial reciprocating engines (SIREs) less than 250 hp, so it is common practice to use the factors intended for SIREs in the range of 250-600 hp. Using the rating flow chart shown in Figure F-3, the factor score in this example falls between 5 and 8. If nothing is known about the relative variation in engine emissions (particularly as related to size), the only option would be to choose the low score of 5. However, if standard references (such as *AP-42*) are being used, it is usually possible to find additional information. For example, if we compare the nitrogen oxides (NO_x) emission factor for a SIRE to that for a large bore engine (LBE), the ratio is 4.41 to 3.1 or roughly 1.4. If the same relationship can be assumed to apply to the smaller SIREs when compared to larger SIREs, then the variability is not likely to be high (where high is an order of magnitude or more). The score could be raised to 6. In most cases, the expected variability values and ranges shown on Figure F-3 will be subjective rather than actually quantifiable.

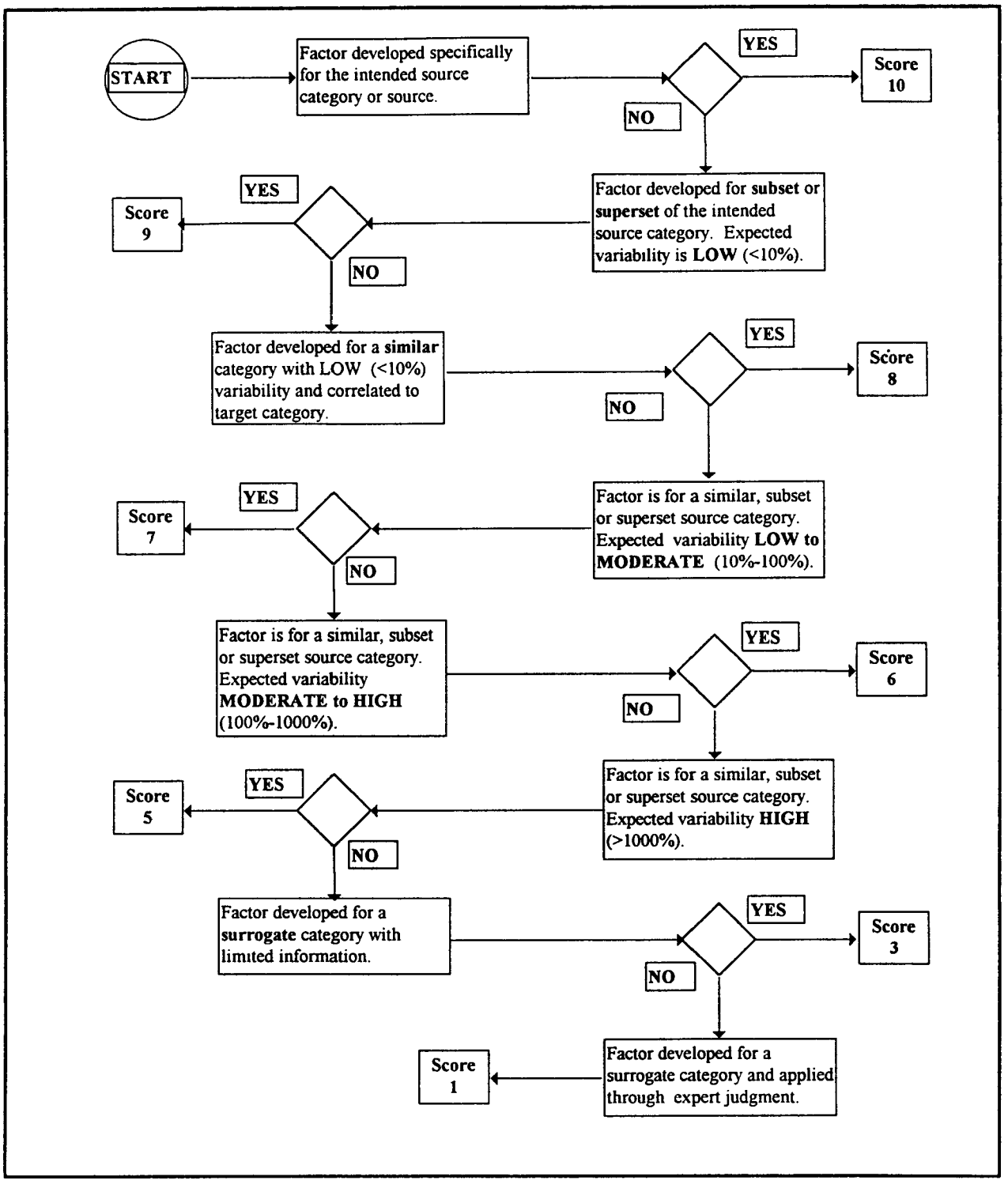

FIGURE F-3. DARS SOURCE CATEGORY SPECIFICITY ATTRIBUTE EMISSION FACTOR RATING FLOW CHART

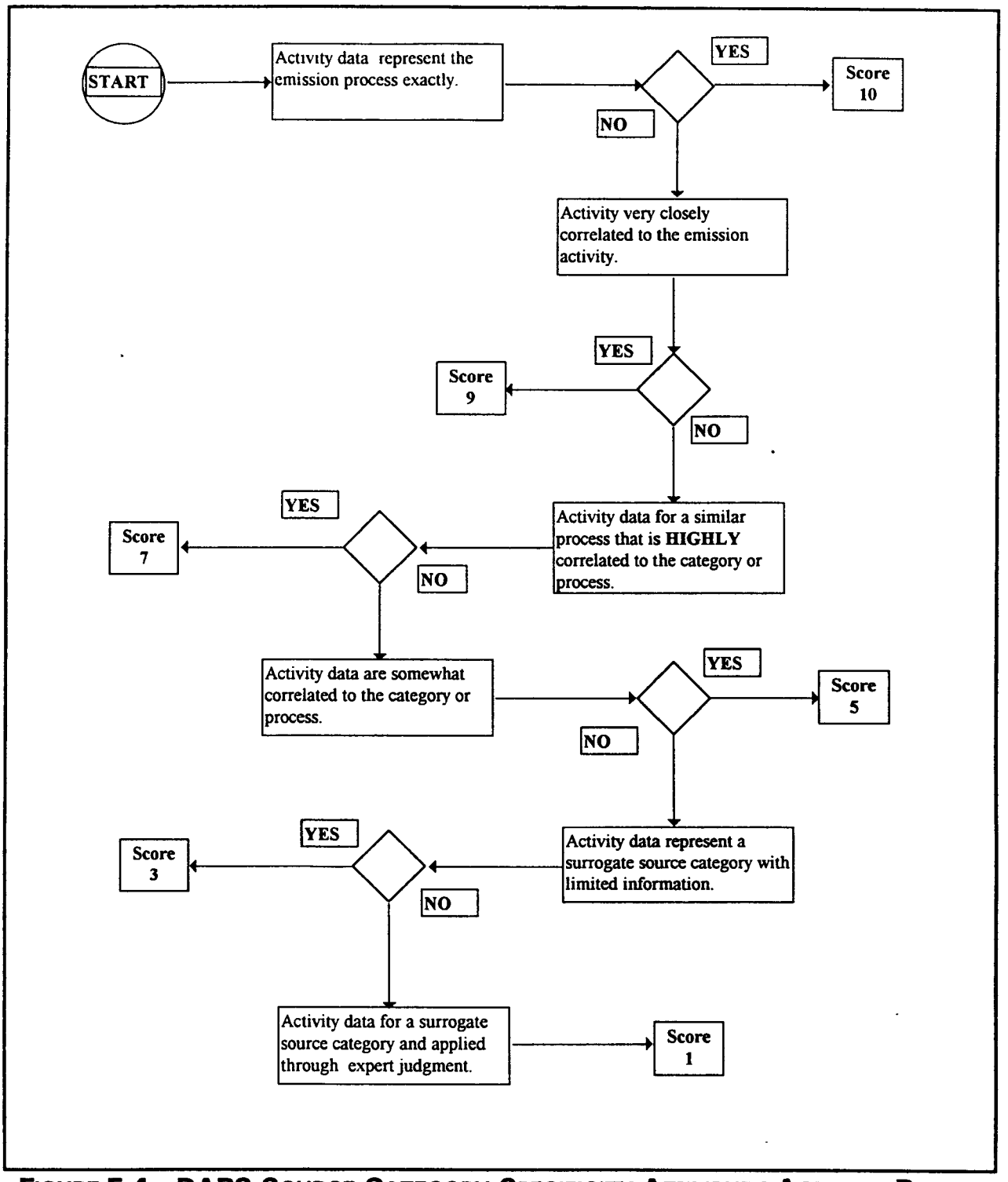

FIGURE F-4. DARS SOURCE CATEGORY SPECIFICITY ATTRIBUTE ACTIVITY RATING FLOW CHART

The activity score for this attribute is determined by the denominator in the emission factor. Scoring should be based on how *specifically* that activity variable applies to the emissive process. The use of annual industrial fuel oil consumption to estimate combustion area source emissions from industry provides a good example. Oil consumption is a surrogate for oil combustion, but it is a very good surrogate (activity source specificity attribute score is 9).

Very often, area source methods use an easily obtained surrogate variable as the activity variable. Commonly, population or employment data are used. The use of population as a surrogate will usually be scored as 1. However, if the inventory preparer (or provider of the emission factor) can demonstrate statistically a correlation between population and a specific activity, the score could be raised. One exception is consumer/commercial solvent use where population could reasonably be expected to correlate with product usage and therefore with emissions. This would still only get a 3 because other demographic factors (e.g., age, gender, ethnic background) are likely to affect types and quantities of products used.

Many adjustments to the factor should be included in scoring this attribute. These include rule effectiveness and rule penetration as well as others that are determined by the definition or characteristics of the source. Accounting for adjustments to estimates is discussed in the section entitled "Adjustments to Estimates." The main point is that any adjustments that improve the match between source category and factor or activity adds to the DARS score.

Spatial Congruity Attribute

This attribute deals with the spatial scaling of factors and activity data that is common to inventories. Figures F-5 and F-6 show the criteria used to score this attribute. For example, in the previous section, the use of state-level DOE fuel consumption data was discussed. With the exception of utilities' fuel consumption, sector fuel use by state is not measured directly. Various databases and assumptions must be used to allocate national fuel use to state level.

Furthermore, to use the state data at a county level, some method of apportionment must be used. Typically, the ratio of county industrial employment to state industrial employment is used. Unless there are studies demonstrating a correlation between employment and emissions, an activity score of 3 is indicated. The activity in this case is representative of a larger scale, and the scaling factors are not correlated well with activity. If information or data can be used to verify or adjust the scaled data, the score can be increased. A lot of judgment is required for scoring this attribute.

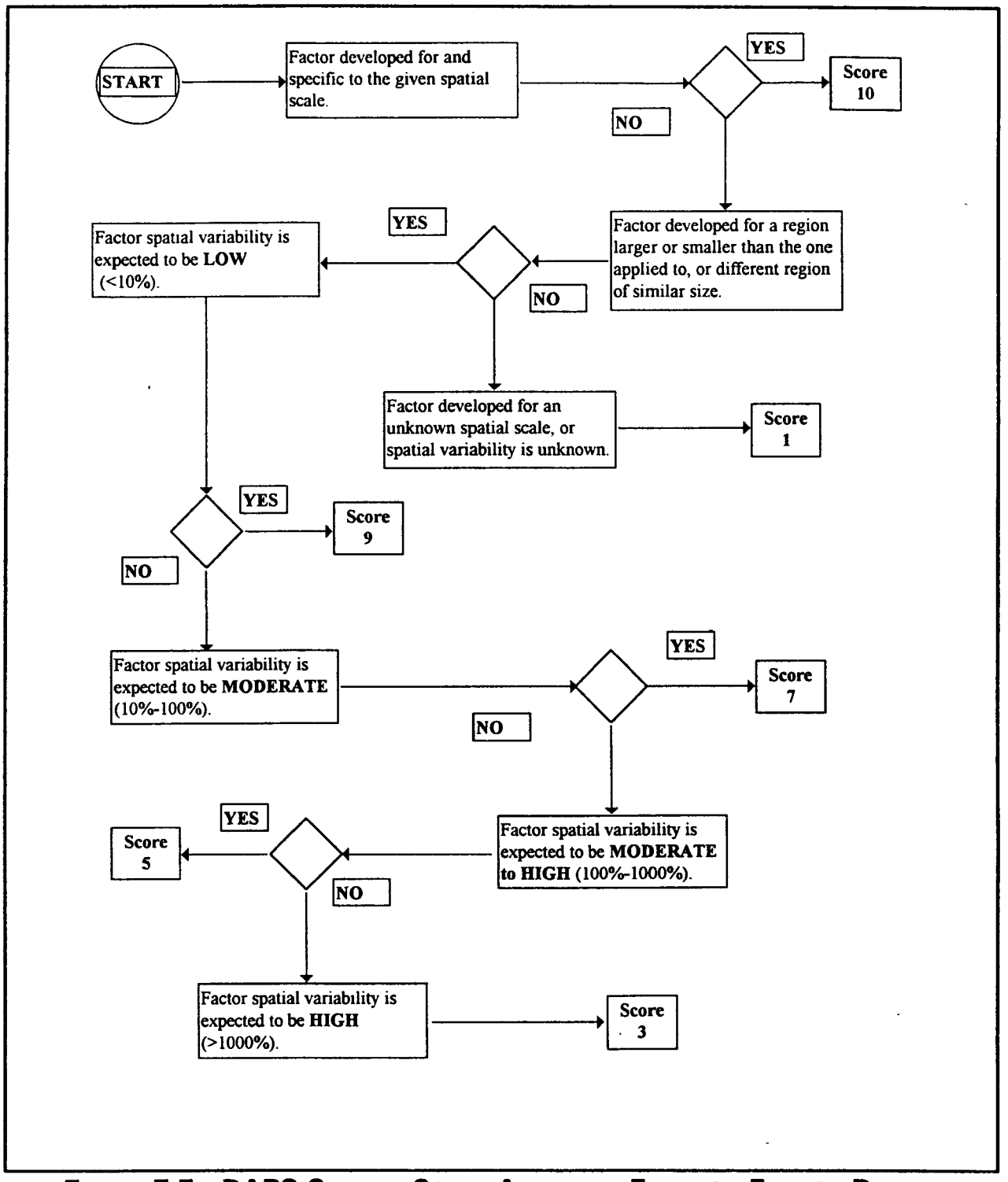

FIGURE F-5. DARS SPATIAL SCALE ATTRIBUTE EMISSION FACTOR RATING FLOW CHART

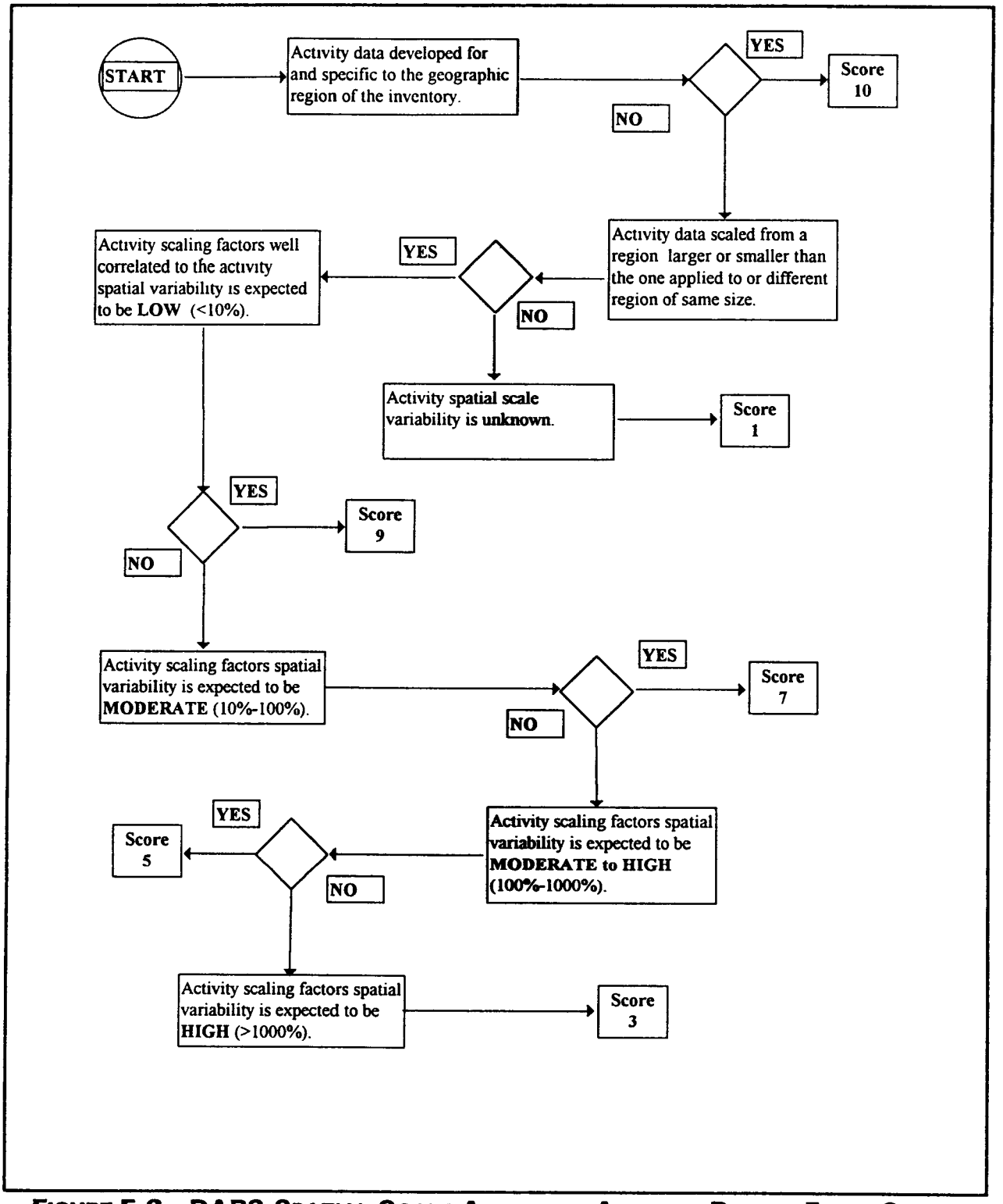

FIGURE F-6. DARS SPATIAL SCALE ATTRIBUTE ACTIVITY RATING FLOW CHART

Spatial scale considerations should include instances where emissions or activity from the same scale are adapted for use in another region. An example is the use of non-road mobile emission studies for specific metropolitan areas applied to other metropolitan areas. Activity scores will typically fall between 3 and 7 in this example, depending on how well matched the two areas are. Clearly, if some additional work has been done to make the data match the intended source region better, that should be reflected in the DARS scores.

The variability in the emission factor that is caused by spatial scaling problems is easy to confuse with Source Specificity issues. For this attribute, regional or local variability in emissions that are attributable to climate, terrain, or other physical (environmental) factors is included. For example, evaporative losses of volatile organic compounds (VOCs) are affected by temperature. The emission factor equations for evaporative losses from petroleum product storage and distribution allow for adjustments based on local meteorological data. If these adjustments have not been made, and if the potential error is high, the DARS score is 3. However, if local conditions are not very different from the values used to calculate default emission factors, the DARS score could go as high as 8.

Temporal Congruity

This attribute describes the match between emission factor, activity, and temporal scale of the inventory. The scoring criteria are shown in Figures F-7 and F-8. The potential mismatches between an inventory estimate and the data used to calculate it that are included in this attribute are:

1. Emission factor or activity based on annual totals used to estimate hourly emissions.

2. Emission factors or activity based on short-term measurements are extrapolated to longer time frames.

3. Emissions projected into the future based on estimates of future growth.

The guidance for DARS scoring here is probably the most subjective of any of the attributes. The approach to use is:

1. Determine if there are any temporal incongruities for the source categories (such as those described above).

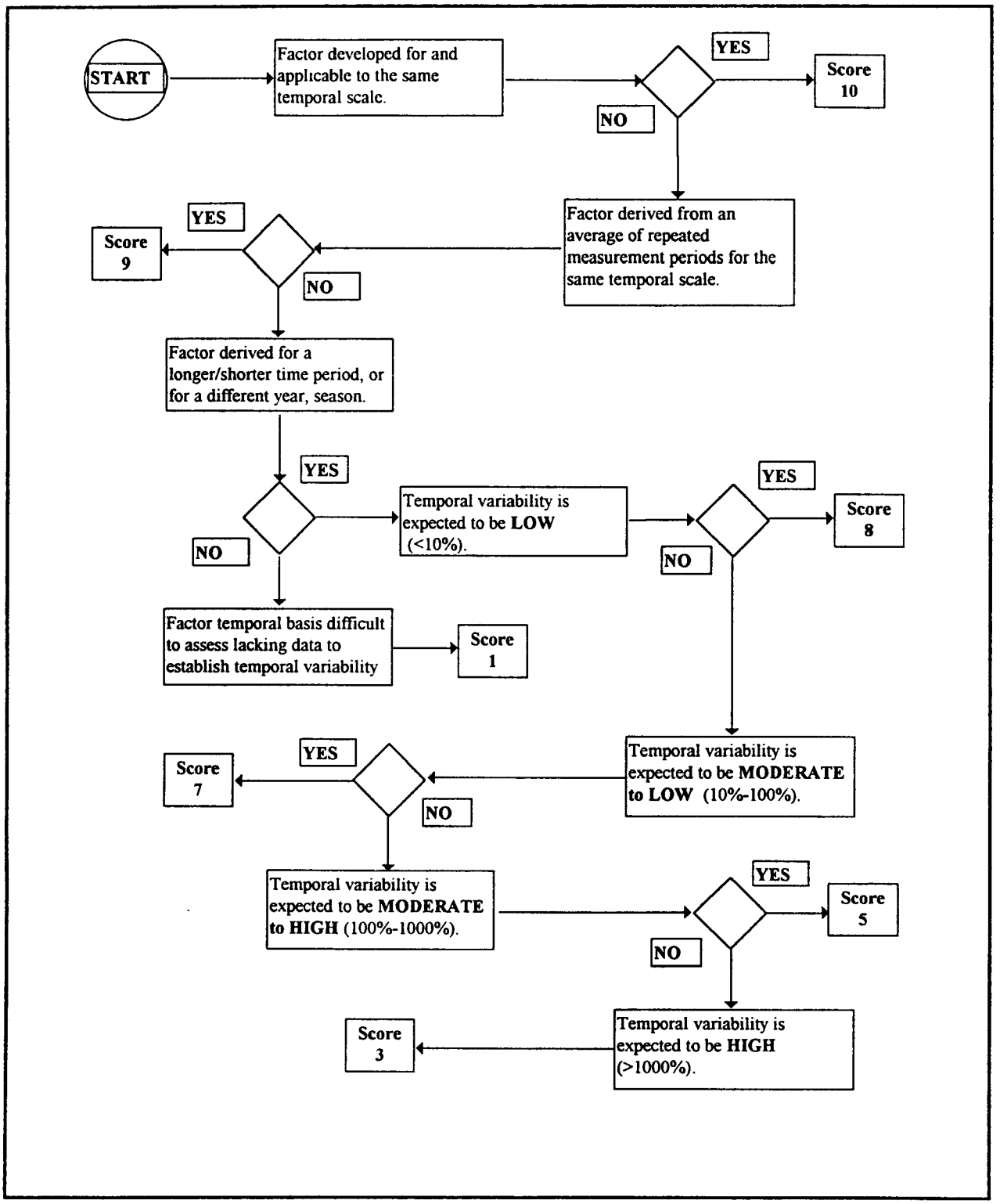

FIGURE F-7. DARS TEMPORAL ATTRIBUTE EMISSION FACTOR RATING FLOW CHART

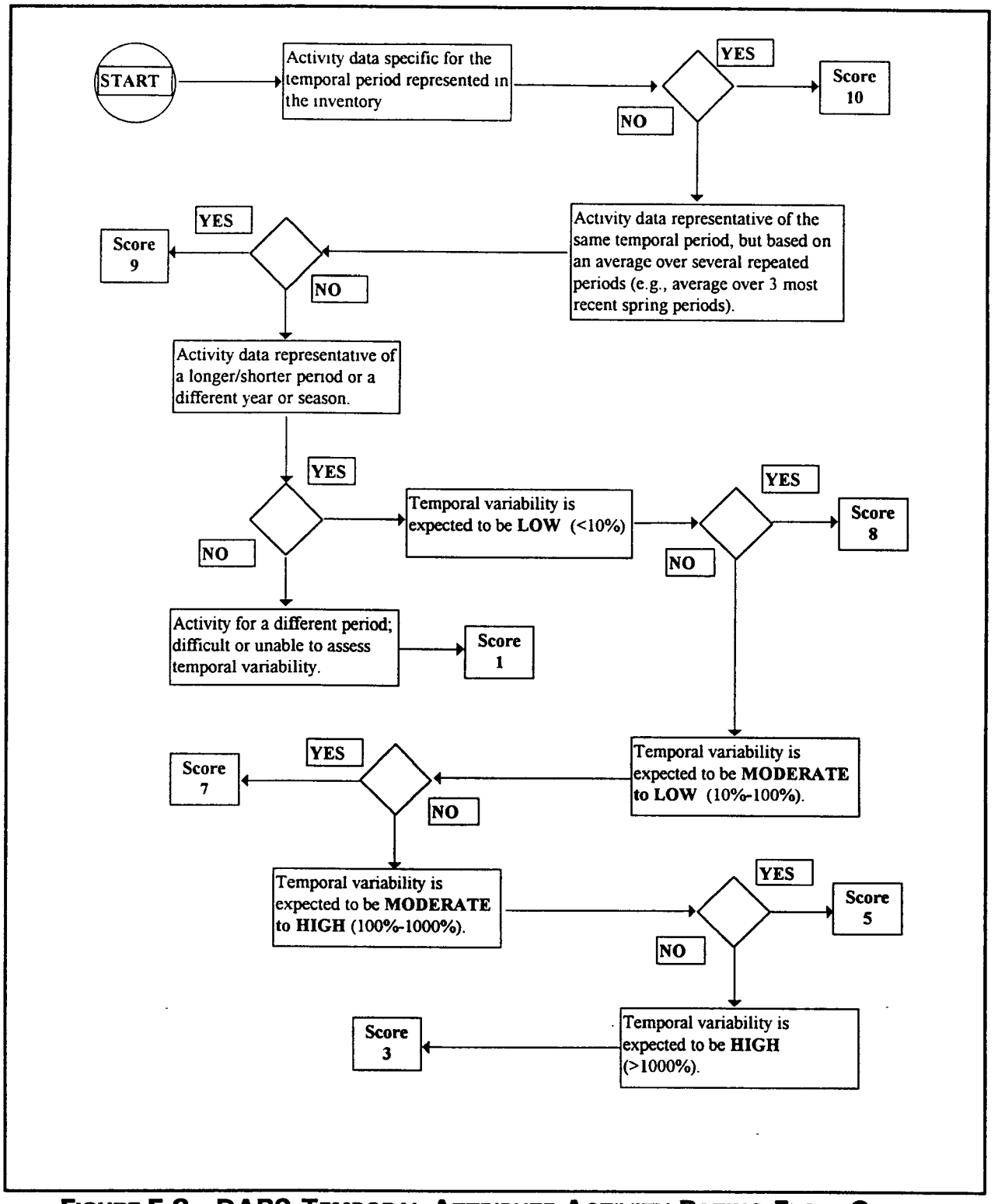

FIGURE F-8. DARS TEMPORAL ATTRIBUTE ACTIVITY RATING FLOW CHART

2. Evaluate the likelihood that these incongruities have the potential to affect the emissions.

There is no simple answer here. Some processes are fairly constant from year to year, and fairly uniform throughout the year. Others change dramatically from year to year, and may fluctuate widely throughout the year.

Activity in many industries fluctuates with demand for their products. These facilities do not necessarily reduce employment; instead, the plant may shut down for a few weeks, or they may go to shortened work weeks. Emissions estimated using per employee factors will not necessarily reflect this reduction in activity.

If the DARS scorer has reason to believe any of these issues (and others) apply to a source category, the DARS scores should be kept in the 3 to 5 range. As with spatial congruity, considerable judgment is required unless actual data are available.

Comparison of Measurement and Source Specificity Attributes

Some additional guidance is needed to clarify how the measurement and source specificity attributes are scored for both the emission factor and the activity:

- The source specificity attribute for the activity applies to the original choice of activity variable used when the factor was developed; for example, using "lb of coating" as the activity for architectural surface coating emissions would receive a higher source specificity score than using "population."

- The measurement attribute for the activity applies to the actual data used to estimate emissions in the inventory being scored; if population is required by the emission method chosen (i.e., a per capita factor is being used), and if population is measured directly, a score of 10 is possible. If coating consumption is being used, but had to be estimated, a score lower than 10 will be given.

For the emission factor scores, it is important to keep in mind that it is the *numerator* in the factor that you should consider. This may require some research to determine how the emission factor was developed. The general approach when developing a factor is to first quantify emissions from the source, and then to express the emissions in terms of some commonly available variable that is directly related to the emissive activity itself. The original emissions (i.e., the numerator in the factor) were probably expressed in terms of time and space (e.g., VOCs/day/spray booth, NO_x/year/globally). Two very different examples of emission factor derivations are given as examples:

- Company XYZ's coal-fired boiler emissions might be estimated using cumulative NO_x emissions from the boiler's CEM data over 1 year's time; emissions are expressed as "tons NO_x emitted from XYZ's boiler annually."

- Total VOCs from architectural surface coating use might be estimated by collecting national paint and coatings consumption data in 1 year, determining the average VOC content of those coatings, and assuming that all VOCs evaporate. The emissions might be expressed as "tons VOCs emitted from the use of architectural surface coatings in the United States annually."

In both cases, the temporal interval is 1 year, but the spatial scales are quite different. Either could be used to develop an emission factor.

The best factors are expressed in terms of a variable that is directly or indirectly related to the emissive activity itself. However, sometimes convenience is weighted more heavily, and a less-well-correlated surrogate (such as population) is used. Using the previous examples:

- Company XYZ's total coal use for the year is divided into the total NO_x emissions to develop an emission factor. The units of the factor are "tons of NO_x per ton of coal consumed."

- The VOCs from architectural surface coating are divided by the total national population. The units of this factor are "tons of VOCs per capita per year."

(Note that neither factor conveys any information about their original spatial scales. The NO_x factor also does not give any indication about the original temporal scale of the data.)

When assigning a DARS score for the emission factor measurement attribute, only the original emissions data should be rated. The denominator (i.e., the activity) is rated using the activity source specificity attribute. Figure F-9 illustrates this; the boxes that are applicable to the original emission factor data are shaded. This means that *no matter how poor the activity surrogate is, the emission factor measurement score will be a 10 if the factor is based on valid, near-continuous data.* This is illustrated using the NO_x example.

Attribute	Factor	Activity
Measurement		
Source Specificity		
Spatial		
Temporal		

FIGURE F-9. ATTRIBUTE SCORES BASED ON ORIGINAL FACTOR SHOWN AS SHADED BOXES

Consider the following emission factors, all developed using the continuous CEM data described above:

- lb NO_x/ton coal burned;
- lb NO_x/hours of operation;
- lb NO_x/rated capacity of boiler; and
- lb NO_x/boiler.

For each of these, the DARS emission factor measurement attribute is a 10. The activity source specificity score for the first is a 9 (at least), for the second is a 6, for the third is a 3, and for the last one is a 1. This approach suggests that a partial DARS score could be used to rank emission factors (irrespective of how they are later applied). If DARS becomes a standard tool used by the states and EPA, the partial DARS scores could be supplied with the factor. EIIP guidance has already started to use this approach.

ADJUSTMENTS TO ESTIMATES

Some inventories require certain adjustments to the emission estimates or to the data used in the estimates. These adjustments may be prescribed for certain types of inventories (e.g., rule efficiency in SIP inventories). Or, they may be applied after the inventory was created to make it suitable for a new use (e.g., allocation of emissions to a grid for modelling purposes). Table F-3 lists some of these adjustments and shows which attribute score is affected. A brief description of each type of adjustment and its potential effect on DARS scores is given below. The reason for mapping an adjustment to a particular attribute may not always be apparent; in fact, some of the pairings are

TABLE F-3

DARS ATTRIBUTES AFFECTED BY INVENTORY ADJUSTMENTS[a]

Adjustments	Measurement/ Method		Source Specificity		Spatial Congruity		Temporal Congruity	
	e	a	e	a	e	a	e	a
Control efficiency (CE)			X					
Rule effectiveness (RE)			X					
Rule penetration (RP)				X				
Speciation profiles	X							
Seasonal activity factors								X
Allocation to grid cell (modeling inventory)						X		
Subtraction of point sources from area		X						
Projections (or backcasting) of emissions	---[b]	X	---[b]	---[b]				

[a] e = emission factor score; a = activity score.
[b] One or more of these sources may be impacted (see text).

debatable. However, it is more important that the effect of an adjustment be accounted for only once, and that it be done consistently.

Control Efficiency (CE) and Rule Effectiveness (RE)

RE is an adjustment to CE that is used to account for deterioration, improper maintenance, or other factors that lower the effectiveness of control equipment. The use of RE affects the DARS emission factor source specificity score in the following ways:

1. If RE is *not* used at all, the base score is lowered unless justification is provided to show that RE is not applicable;

2. If RE has been calculated specifically for the source in the inventory region, the base score is either unchanged or may be raised (default is no change); and

3. If the EPA's default RE has been used, the base score is either unchanged or may be lowered (default is no change).

The decision to raise or lower a score is situation-dependent. For example, if the base DARS score source specificity factor score is already high (9 or 10) or very low (1 to 3), developing a source-specific RE value may have little effect on user confidence in the estimate. On the other hand, ignoring RE completely should produce some doubt about the estimate.

The reason that source specificity and not measurement/method is affected is that the measurement/method score is always based on the quality of the original factor. The way in which the factor has been applied is addressed in the source specificity attribute.

Rule Penetration (RP)

RP represents the fraction of the source population that is affected by a rule. The activity parameter is adjusted using RP, so it is accounted for in the DARS score for the activity source specificity attribute. This attribute score is affected as follows:

1. If a control requirement exists but RP is not addressed, lower the score at least one point.

2. If RP is included (where appropriate), the base score is unaffected or possibly raised if inclusion of RP has a significant impact on the estimate.

The decision to raise or not raise the score is situation-dependent.

Speciation Profiles

The use of speciation profiles to estimate a specific pollutant type (e.g., specific hazardous air pollutants, PM-2.5) from a more general pollutant category (e.g., VOCs, PM) is accounted for in the emission factor measurement/method score. Generally, the method used to estimate the general pollutant is scored first; the application of a speciation profile usually results in lowering the score. However, if the speciation profile can be shown to be very accurate, the effect on the score may be minimal.

Seasonal Activity Factors (SAFs)

SAFs are often used to adjust an annual value to a daily or seasonal estimate. Because it is usually more accurate to estimate a short-term value using data appropriate to that time period, using an SAF will negatively impact the activity temporal congruity score.

Allocation to Grid Cell

Modelling inventories for area or mobile sources are often prepared by allocating emissions to grid cells; the resulting inventory is more finely resolved (spatially) than the original inventory. Any impact on emissions certainty is accounted for in the activity spatial congruity attribute because allocation is usually achieved by adjusting the activity variable.

Point Sources Adjustment to Area Sources Emissions

Certain types of processes may be represented by both point and area sources. Typically, the activity factor is adjusted by subtracting any point source activity from the total activity before calculating the area source emissions. This should be considered when scoring the activity measurement/method attribute.

Projections (or Backcasting) of Emissions

If emissions at some future date are needed, they must be estimated by projecting into the future. Determining the DARS score for projection inventories will generally be a two-step process. First, determine the scores for the inventory used as the basis for the projections. Second, modify the attributes affected by the use of growth factors.

Usually, the activity measurement/method attribute is adversely affected by projections because the activity is "grown" by some multiplier. However, in some circumstances,

more than one attribute will be impacted. In an example using utility boilers, forecasts may predict an increase in energy demand in the inventory region. This percentage of increase in demand may be used to "grow" the utility emissions from the base year. This approach assumes that the base year proportions of processes (e.g., coal, oil, hydroelectric, solar) will not change and, furthermore, that the emission factors for those processes will not change. If these assumptions cannot be confidently made, the DARS score for emission factor measurement/method and for activity source specificity may be lowered (from the base year's) to reflect the decreased confidence in the estimate.

A current example is any industry using paints or coatings that is subject to one of the new National Emission Standards for Hazardous Air Pollutants (NESHAPs) developed since the 1990 Clean Air Act Amendments were passed. These proposed Maximum Achievable Control Technology Standards (MACT) and Control Techniques Guidelines have resulted in significant changes in coatings formulations. These reformulated products are already in use and will increase in use over the next few years. Changes in the formulations will result in changes in emissions that will not be predictable. However, once all affected facilities are in compliance, the emissions per unit activity should be more consistent from facility to facility. Using a 1990 emission factor to estimate future emissions will not produce reliable results.

DARS: Point Sources

Although DARS was originally conceived as an application for global inventories developed using area source inventory methods, not as an application for point sources, the method can be applied to inventories at any scale. However, the amount of time required to assign DARS scores to an entire point source inventory for just one county can be quite high. Even if simplified methods (as described below) are used, scoring an entire point source inventory will nearly always require more effort than required for the area and mobile source inventories.

Point source inventories are generally expected to be "better" (which usually is assumed to imply "more accurate") than area source inventories. Although the activity data are certainly likely to be more accurate, the same cannot be said for the emission factors. In fact, emission factors are usually based on a small sample of individual units in the source category. Using these factors to estimate emissions from *another* individual emission unit has a higher probability of error than using it to estimate the sum of emissions of units (i.e., an area source approach). This occurs because an average factor will either under- or overestimate the emissions about half of the time. However, summing up these individual estimates tend to cancel out the errors. (Note that according to this argument, the sum of the point source emissions is "better" than the individual estimates.)

Another problem with emission factors, however, is that they may be based on a biased data set. (The example cited earlier for SIREs applies here.) If the emission factor is based on a subset of the technologies covered by the source category, the factor may in fact be misapplied (or at best a tenuous match) to units not represented in the original data set.

For these reasons and others, it is generally not productive to evaluate every emission unit within a facility, or even every facility, individually, if the overall objective is to rate the point source inventory. An easier (and just as accurate) approach is to sort the point source inventory emissions into categories based on methods used, e.g., source testing, *AP-42* factors, state factors, mass balance, and engineering judgment. Then, sort each of these categories by the method used for the activity data (fuel/materials consumption, production rate, hours usage).

DARS scores are then applied to these groups of sources rather than to individual sources. The scorer may want to spend more time on sources where source testing or nonstandard methods are used. Also, keep in mind that mass balance used at an individual facility may be more accurate than when used to develop a national estimate. If other losses (i.e., non-air releases) are accurately accounted for, the DARS activity measurement score could reach a 9 (and possibly 10 if well supported).

Keeping the above comments in mind, the general guidance given for area sources will apply to point sources. For a given source type (e.g., industrial fuel consumption), if *AP-42* emission factors are used, the measurement attribute score will be the same as for area sources. The scores for the other attributes will usually be different.

As stated above, several different approaches can be used to apply DARS to a point source inventory. If a comparison of individual facilities is needed, each point source must be evaluated separately, applying DARS scores to each emission unit (or collection of similar units). The individual scores are then weighted by the percentage contribution to total facility emissions. A more common approach is to assess the point source inventory overall, rather than dealing with individual point sources. Sources are grouped in some logical way (such as by Source Classification Code), and DARS scores are assigned to those groupings. An example using this approach is shown in the next section.

APPLICATION OF DARS TO RICHMOND, VA, POINT SOURCE INVENTORY

In this inventory, five different estimation methods were used to estimate emissions for point sources (Ballou, 1995). These methods and the rationale for the DARS scores are

discussed in the following subsections in order from highest to lowest composite emissions score.

Continuous Emissions Monitoring

This method is generally considered the best method for emission estimates if the associated data quality and coverage goals are met. However, no estimate is 100% accurate and a margin of equipment downtime or data loss is allowed when accepting this type of data. As a result, a score of 9 was applied to the measurement attribute to reflect this. This could be raised or lowered if actual coverage and completeness data were available. If monitored correctly, all other scores could achieve scores of 1.0. In the case of this inventory, there were no estimates based on this method.

Attribute	Factor	Activity	Emissions
Measurement	0.9	0.9	0.81
Specificity	1.0	1.0	1.0
Spatial	1.0	1.0	1.0
Temporal	1.0	1.0	1.0
Composite	0.975	0.975	0.95

Source Stack Testing

This is the next highest-rated method for developing emission estimates. However, because most stack test data used for inventories come from compliance tests based on small samples running "typical" loads, lower scores were assigned to the measurement attribute. Lower scores were also assigned to the temporal attribute since it is likely that the factor and activity data are based on different time periods or seasons/years. A small number of emission estimates in the inventory were based on stack test results.

Attribute	Factor	Activity	Emissions
Measurement	0.8	0.9	0.72
Specificity	1.0	1.0	1.0
Spatial	1.0	1.0	1.0
Temporal	0.8	0.8	0.64
Composite	0.9	0.925	0.84

Mass Balance Calculation

In a deviation from the DARS flow charts, mass balance calculations were assigned a higher score for evaporative and/or process emissions at point sources. This was done assuming that all losses are accounted for as part of these calculations. A substantial number of VOC estimates in the inventory were based on mass balance calculations as well as a lesser amount of process NO_x emissions.

Attribute	Factor	Activity	Emissions
Measurement	0.7	0.9	0.63
Specificity	1.0	1.0	1.0
Spatial	1.0	1.0	1.0
Temporal	0.8	0.8	0.64
Composite	0.875	0.775	0.8175

Source-specific Emission Factors

For this evaluation, source-specific emission factors were rated a step above *AP-42* factors but below the methods already discussed. The reason for this is that these factors are derived from source-specific information, based on other more accurate estimation methods (stack tests and mass balance) or tests made on similar equipment elsewhere. Therefore, source specificity and the overall score is slightly lower for this reason. Relatively few estimates were based on these types of calculations.

Attribute	Factor	Activity	Emissions
Measurement	0.7	0.9	0.63
Specificity	0.8	1.0	0.8
Spatial	1.0	1.0	1.0
Temporal	0.8	0.8	0.64
Composite	0.825	0.925	0.7675

AP-42 Emission Factors

The official *AP-42* emission factors, and the computerized factors based on *AP-42* are rated at this point in the hierarchy of methods. The major dilemma in assigning a single DARS score to *AP-42* factors is the variation in the quality of the factors. To try to address this quality variability, different scores were assigned to the measurement attribute factor rating based on the factor rating in *AP-42* (if available). This was an easier task in the case of NO_x because the *AP-42* factors are generally combustion related, well documented, and of better quality. This was much more difficult for VOC factors that had little or no documentation of quality, and had no ratings in many cases. When no rating could be found, the lowest score was assigned. The scores assigned were based on the information shown in Table F-2.

Attribute	Factor	Activity	Emissions
Measurement	0.4 to 0.6	0.9	0.36 to 0.54
Specificity	0.7	1.0	0.7
Spatial	1.0	1.0	1.0
Temporal	0.8	0.8	0.64
Composite	0.725 to 0.775	0.925	0.675 to 0.72

Expert Judgment/Guess

For obvious reasons, this method of emission calculation is ranked lowest because of the lack of any information on the data and calculations used to make such estimates. A small number of emission estimates were developed using this method.

Attribute	Factor	Activity	Emissions
Measurement	0.1	0.8	0.08
Specificity	0.6	0.8	0.48
Spatial	1.0	1.0	1.0
Temporal	0.7	0.7	0.49
Composite	0.6	0.825	0.51

These scores were then applied to the source emissions at the process level. The DARS scores for each process were multiplied by the relative contribution of that process to *total* point source inventory emissions to produce a weighted DARS score. An example for two sources is shown in Table F-4. The sum of these weighted emissions is the overall DARS score for the inventory. In this example, the composite DARS score for the NO_x inventory was 0.76; for the VOC inventory, 0.72.

APPLICATION OF DARS TO COMPLEX MODELS

Emission inventories are not usually thought of as models, but in fact they are. A model is a representation of reality. In an emission inventory, the model may be as simple as

emissions = emission factor x activity

Or, a complex, computer-based model may be used to estimate the emission factor, the activity, or both. The same guidance given previously in this appendix for simple models (i.e., emission factor x activity) can be applied to the use of more complex models such as those used to estimate mobile source emissions.

One factor to be considered when rating complex emission factor models is whether the model is based solely on theory (i.e., first principles). Theoretical models generally rate a 3 for the measurement attribute factor score (see Figure F-1). If the model has been validated or calibrated using real-world measurements, the model results are considered better. Empirical models (e.g., statistical regressions) are not necessarily better or worse than theoretical or deterministic models. In both cases, the degree to which key explanatory variables have been included and the amount of variability reduced affect the quality.

A key point to keep in mind is that *increased model complexity does not necessarily imply better quality emission factors*. If default input values are used to estimate an emission

TABLE F-4

EXAMPLE OF 1990 POINT SOURCE DARS EVALUATION FOR NO_x (RICHMOND OZONE NONATTAINMENT AREA)

Plant Name	Emissions	% of Total	DARS Score	Weighted Score
Facility 1	95.15	0.000329	0.72	0.00023739
	2685.27	0.009304	0.72	0.00669958
	0.8	0.000002	0.72	0.00000199
	2270.52	0.007867	0.72	0.00566480
	86.35	0.000299	0.72	0.00021543
	0.6	0.000002	0.72	0.00000149
	2420.04	0.008385	0.72	0.00603785
	91.85	0.000318	0.72	0.00022916
	0.8	0.000002	0.72	0.00000199
	1846.72	0.006399	0.72	0.00460750
	70.4	0.000243	0.72	0.00017564
	0.6	0.000002	0.72	0.00000149
	2404.08	0.008330	0.72	0.00599803
	91.3	0.000316	0.72	0.00022778
	0.8	0.000002	0.72	0.00000199
	55	0.000190	0.72	0.00013722
	0.2	0.000000	0.72	0.00000049
	56.1	0.000194	0.72	0.00013996
	0.2	0.000000	0.72	0.00000049
	57.2	0.000198	0.72	0.00014271
	0.06	0.000000	0.72	0.00000014
	486.66	0.001686	0.675	0.0011383
	217.25	0.000752	0.675	0.00050814
	159.47	0.000552	0.675	0.00037300
Facility 2	161.2717	0.000558	0.6975	0.00038978
	6175.324	0.021398	0.6975	0.01492558
	6815.7	0.023617	0.72	0.01700475
	16.32	0.000056	0.675	0.00003817
	16298.25	0.056476	0.72	0.04066312

Plant Name	Emissions	% of Total	DARS Score	Weighted Score
	16.8	0.000058	0.675	0.00003929
	36635.85	0.126950	0.72	0.09140418
	34.56	0.000119	0.675	0.00008083
	62202.75	0.215544	0.72	0.15519201
	68.64	0.000237	0.675	0.00016054
Facility 3	0.66	0.000002	0.72	0.00000164
	105.374	0.000365	0.675	0.00024647
	6	0.000020	0.5125	0.00001065
Facility 4	101.2	0.000350	0.72	0.00025248
	484	0.001677	0.72	0.00120755
	14.72	0.000051	0.675	0.00003443
	422.4	0.001463	0.72	0.00105386
	2051.5	0.007108	0.72	0.00511836
	37.95	0.000131	0.675	0.00008876
	4	0.000013	0.72	0.00000997
	4	0.000013	0.72	0.00000997
	583	0.002020	0.8175	0.00165151
	11	0.000038	0.84	0.00003201
	126	0.000436	0.84	0.00036675
	29	0.000100	0.84	0.00008441
	23	0.000079	0.84	0.00006694
	6434	0.022295	0.84	0.01872783
	291	0.001008	0.84	0.00084703
	823	0.002851	0.84	0.00239555
	12229	0.042375	0.84	0.03559569
	12817	0.044413	0.84	0.03730722
	15806	0.054770	0.84	0.04600749
	10360	0.035899	0.84	0.03015548
	8903	0.030850	0.84	0.02591450
	4640	0.016078	0.84	0.01350593

factor, the resulting computer model-generated factor may be no better than a national average factor based on measurements or mass balance; if, however, site-specific input data are used, the emission factor produced by the model is of better quality than the default.

For the modeled emission factor, most of the effort will focus on the measurement attribute score. However, for some models such as MOBILE5a that provide national default inputs, the spatial congruity attribute should be carefully evaluated as well. If specific local inputs were used, the spatial congruity score for the factor may rate as high as 10. If national defaults were used, the score will be lower.

When scoring the measurement attribute for an emission factor or activity developed using a model, ask the following questions:

Step 1. Is the model entirely theoretical? This means it has never been validated or calibrated using real-world data, and no empirical measurement data were used in developing the model equations. If yes, then the factor measurement attribute score is 3; if no, go to Step 2.

Step 2. Were some parts of the model based on empirical data, either limited field data, or laboratory or bench-scale data? If yes, do these "parts" have a significant impact on the model results? If no, the score may be raised slightly, but not greater than 4.

If yes, go to Step 3.

Step 3. Has significant real-world validation/calibration of the model been done and has this demonstrably improved the model's capability to produce accurate emission factors? If yes, have these capabilities been used by the inventory developer to their maximum capability? If no, do not increase the score any higher. If yes, the score may be raised as high as 9 (but only if the model has been demonstrated to be highly accurate and its full potential has been used).

If a model is used to calculate activity, the same considerations apply. For example, travel demand modules may be used to calculate vehicle miles travelled (VMT) for a mobile source inventory. The first step is to determine a score for source specificity: how good a surrogate is VMT for combustion of fuel in an internal combustion engine? This score does *not* consider how VMT was calculated. Developing the measurement score requires consideration of the quality of the model.

By now, it should be clear that the scores for complex models are assigned using essentially the same criteria as for any area source. Furthermore, no matter how complex the model, the relevance of its results with respect to real-world data must be demonstrated in order for higher scores to be achieved.

REFERENCES

Ballou, T., R.L. Peer, and V. Chandler. 1995. *Evaluation of an Inventory Rating System: An EIIP QAC Project.* Presented at the Air & Waste Management Association's Emission Inventory Conference, Research Triangle Park, North Carolina.

Beck, L., R. Peer, L. Bravo, and Y. Yan. 1994. *A Data Attribute Rating System.* Presented at the Air & Waste Management Association's Emission Inventory Conference, Raleigh, North Carolina.

VOLUME VI: CHAPTER 5

MODEL QA PLAN

November 1996

Prepared by:
Eastern Research Group

Prepared for:
Quality Assurance Committee
Emission Inventory Improvement Program

DISCLAIMER

This document was furnished to the Emission Inventory Improvement Program and the U.S. Environmental Protection Agency by Eastern Research Group, Inc., Morrisville, North Carolina. This report is intended to be a final document and has been reviewed and approved for publication. The opinions, findings, and conclusions expressed represent a consensus of the members of the Emission Inventory Improvement Program. Mention of company or product names is not to be considered an endorsement by the U.S. Environmental Protection Agency.

Eastern Research Group
Post Office Box 2010
Morrisville, NC 27560

1

INTRODUCTION

This is the final chapter of the *Quality Assurance Procedures* volume. The model quality assurance (QA) plan presented here illustrates the principles described in the other chapters of this volume. The plan uses fictitious Ozoneville for a QA plan for a State Implementation Plan (SIP) emissions inventory. Because it is intended to be a model QA plan, the remainder of this chapter does not adhere to Emission Inventory Improvement Program (EIIP) format and style used elsewhere in this volume.

This model QA plan was prepared by the EIIP Quality Assurance Committee. The EIIP Quality Assurance Committee was formed to develop (1) a plan for the EIIP's QA program, (2) a comprehensive QA procedures volume, and (3) an emission inventory quality rating system.

Throughout the EIIP QA procedures volume, the concept was developed of delineating inventories into one of four inventory levels. Ultimately, it is the end use of an inventory that delineates the inventory level, which in turn determines the minimum requirements needed in the QA program, for staffing assignments, and for documentation and reporting.

A SIP emissions inventory is considered by EIIP to be a Level II inventory because it will provide supportive data for strategic decisions making. Although the requirements for a Level II inventory are less stringent than those for a Level I inventory (which will be used to support enforcement, compliance, or litigation activities), the minimum elements required for QA and technical work plans are identical. These elements are reflected in this model QA plan.

As discussed in other chapters of the QA procedures volume, there is flexibility in selecting preferred or alternative staffing options, developing separate or combined QA and technical work plans, and delineating activities that will be part of the QA/Quality Control (QC) program. Because of this flexibility, the model QA plan presented in this chapter should not be viewed as a template that must be strictly followed by a state agency preparing an ozone nonattainment area emissions inventory. The level of detail presented in this QA plan and level of effort described in the QA/QC program may vary in a state's actual QA plan, and the plan would still be acceptable to the U.S. Environmental Protection Agency.

For example, in preparing the point source emissions inventory for Ozoneville, the model QA plan states that the Inventory Development Team will send the permit branch's emissions data to each permitted facility for verification. An option would be to send the emissions data only to facilities that have been flagged by the enforcement branch as having potential problems.

While this model QA plan may appear to provide an unattainable QA/QC goal for a SIP emissions inventory, it can serve as the basis from which a hierarchy of procedures and methods can be developed. An agency can identify their most critical inventory needs, and focus their resources in these areas. QA/QC activities listed in this model QA plan can then be selected on the basis of these priorities.

1996 OZONE NONATTAINMENT AREA
STATE IMPLEMENTATION PLAN EMISSION INVENTORY

QUALITY ASSURANCE PLAN

Prepared By:

Ozoneville Department of Environmental Quality

November 1996

TABLE OF CONTENTS

Section — Page

1.0 INTRODUCTION 1-1

1.1 Purpose of Inventory 1-1
1.2 Data Quality Objectives and Indicators 1-1
1.3 Summary of Quality Assurance Plan Organization 1-3

2.0 PROGRAM SUMMARY 2-1

2.1 Major Program Components 2-1
2.2 Inventory Constraints 2-5

3.0 PURPOSE AND POLICY STATEMENT FOR THE 1996 OZONEVILLE STATE IMPLEMENTATION PLAN OZONE NONATTAINMENT EMISSIONS INVENTORY 3-1

4.0 EMISSIONS INVENTORY PREPARATION PLAN 4-1

4.1 Managerial Responsibilities 4-1
4.2 Inventory Development Task Leaders 4-3
4.3 Inventory Development Team Members 4-3
4.4 Technical Reviewers 4-3
4.5 Quality Assurance Coordinator 4-4

5.0 GENERAL QA/QC PROCEDURES 5-1

5.1 QC Activities 5-1

5.1.1 Data Gathering 5-1
5.1.2 Data Documentation 5-2
5.1.3 Calculating Emissions 5-4
5.1.4 Data Checking and DARS Scoring 5-4
5.1.5 Reporting 5-6
5.1.6 Maintenance of Master File 5-9

5.2 QA Activities 5-9

5.2.1 Training/Education 5-11
5.2.2 Audits 5-12
5.2.3 Reporting Audit Findings 5-13

6.0 CORRECTIVE ACTION MECHANISMS 6-1

TABLE OF CONTENTS (continued)

Section | Page

7.0 POINT SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES 7-1

8.0 AREA SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES . . 8-1

9.0 ON- AND NONROAD MOBILE SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES 9-1

9.1 Onroad Mobile Sources 9-1
9.2 Nonroad Mobile Sources 9-5

10.0 BIOGENIC SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES 10-1

11.0 DATA REPORTING 11-1

12.0 REFERENCES 12-1

LIST OF FIGURES

Figure		Page
1-1	Area Map for the Ozoneville MSA Emissions Inventory - 1996	1-2
2-1	Inventory Development Process and Identification of QA Checkpoints	2-3
2-2	Ozoneville MSA 1990 Emissions Inventory Audit Schedule	2-4
2-3	Recommendation for Corrective Action Form	2-6
4-1	Ozoneville MSA 1996 Emission Inventory QA Organization	4-2
5-1	Contact Report	5-5
5-2	Area Source Category QA/QC Checklist	5-7
5-3	Master File Sign-Out Sheet	5-10
7-1	Point Source Survey Form for Stationary Combustion Sources	7-4
8-1	Area Source Data Log Sheet	8-5

LIST OF TABLES

Table		Page
1-1	Data Quality Objectives	1-4
1-2	Data Quality Indicators	1-4
2-1	Possible Effect of Constraints on Ozoneville MSA 1996 Emissions Inventory	2-7
5-1	Data Collection Guidance Documents	5-2
8-1	Area Sources to Include in 1996 Inventory and Proposed Emission Estimation Method	8-2

1.0 INTRODUCTION

1.1 Purpose of Inventory

The 1996 Ozone Nonattainment Area State Implementation Plan (SIP) emissions inventory is being developed in response to requirements specified in the Clean Air Act Amendments of 1990. The inventory addresses volatile organic compounds (VOCs), carbon monoxide (CO), and oxides of nitrogen (NO_x) from point, area, and mobile emission sources. Emissions of VOCs are also addressed for biogenic sources.

The area covered by the inventory includes the Ozoneville Metropolitan Statistical Area (MSA), which was designated by the U.S. Environmental Protection Agency (U.S. EPA) as a serious nonattainment area for ozone. The geographical area delineated by the MSA is shown on the map in Figure 1-1. The map shows the area boundary that was established to avoid unnecessary judgment calls pertaining to the precise location of particular facilities in relation to the MSA borders.

In addition to the regulatory requirements specified by the U.S. EPA, this periodic inventory is being developed to meet the following objectives:

- Determine trends in emission levels, both historically and prospectively;
- Track the 3 percent annual emission reduction requirement for nonattainment pollutants;
- Develop and evaluate air quality-related indicators for measuring progress in attaining ambient standards; and
- Determine the effect of transportation and other control measures on the region's emissions.

1.2 Data Quality Objectives and Indicators

The 1996 Ozone Nonattainment Area SIP emission inventory for Ozoneville is considered a Level II inventory, based on guidance provided by the Emission Inventory

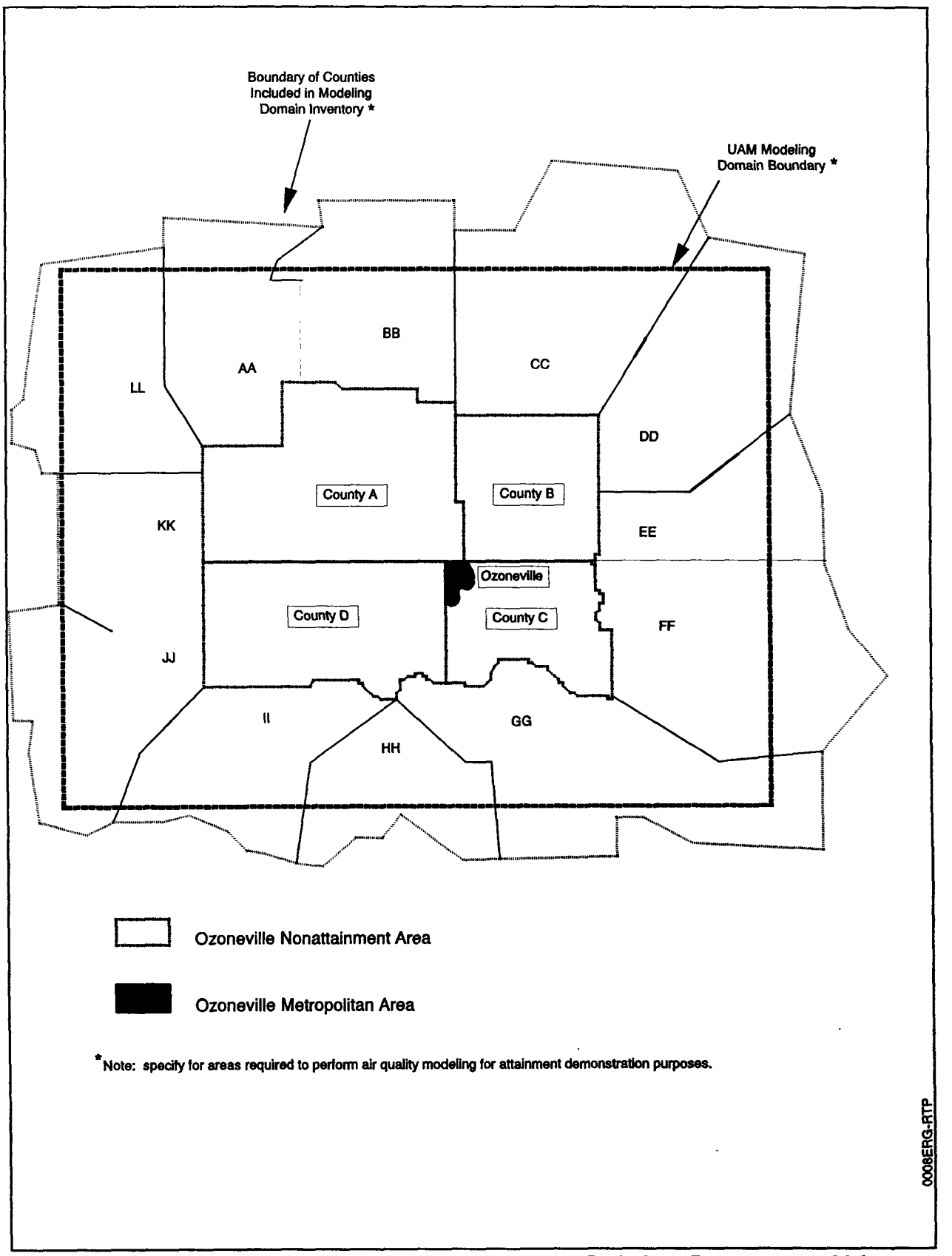

Figure 1-1. Area Map for the Ozoneville MSA Emissions Inventory - 1996

Improvement Program (EIIP) (EIIP, 1996). It is a Level II inventory because it will provide supportive data for strategic decision making. The end use of this inventory, therefore, drives the minimum QA and work plan requirements.

As shown in Table 1-1, data quality objectives (DQOS) were established to help ensure the accuracy, completeness, representativeness, and comparability of the inventory, in keeping with the EIIP's guidance for Level II inventories.

Table 1-2 presents the data quality indicators (DQIs) that will be used to measure progress towards each DQO. The Data Attribute Rating System (DARS) will be used to verify the desired inventory accuracy DQO.

1.3 Summary of Quality Assurance Plan Organization

The remainder of this quality assurance plan (QAP) is organized as follows: Section 2.0 contains the program summary that describes the major components of the inventory development and QA/Quality Control (QC) program; Section 3.0 presents the purpose and policy statement. Section 4.0 contains the emissions inventory preparation plan, which contains details on the organizational structure, roles, and training of inventory development and QA/QC team members. Section 5.0 discusses QA/QC procedures that will be implemented throughout the project, and Section 6.0 describes the corrective action mechanisms that will be implemented as needed. Sections 7.0 through 10.0 discuss the methods that will be used to prepare the point, area, onroad mobile, nonroad mobile, and biogenic source inventories, as well as planned QA/QC activities for each source category. Section 11.0 presents the data reporting procedures that will be followed, and Section 12.0 presents reference citations for all data sources discussed in this QAP.

TABLE 1-1. DATA QUALITY OBJECTIVES

Data Quality Objective	Procedure for Achieving Objective
Accuracy	For point and onroad mobile sources, 100% of the calculations will be checked by the data generator, and 20% of the calculations will be checked by another equally qualified inventory development team (IDT) member. For area, nonroad mobile, and biogenic sources, 100% of the calculations will be checked by the data generator, and 10% of the calculations will be checked by another equally qualified IDT member. In all cases, the data validator will develop a written summary of his or her activities, and will conduct follow-up activities to ensure that data are corrected as needed. If more than 5% of the calculations checked by an equally qualified IDT member need to be revised, then 100% of the calculations will be checked.
Completeness	Extensive planning will be conducted prior to data collection to identify all applicable emission sources. After identifying these sources, the goal will be to determine 100% of the emissions from the largest emitting sources from each source category and as many of the minor sources as possible within the time frame allotted for the work. Those sources identified but not included in the inventory will be identified in the data file and final report.
Representativeness	Senior technical personnel will review all of the primary source data and compare it to previous emissions results and similar results from comparable regions to determine the reasonableness of the emissions estimates and representativeness of the data.
Comparability	To ensure that the data are comparable, standard procedures will be followed and results will be presented in the same units that were used in the 1993 inventory. If a new or improved emission estimation method is used, the 1993 estimate will be recalculated or adjusted to ensure comparability.

TABLE 1-2. DATA QUALITY INDICATORS

DQO	Inventory DQI Target Values
Accuracy	• Achieve DARS score ≥0.7 for all area sources contributing >10% of total emissions of VOCs or NO_x. • Achieve DARS score ≥0.8 for all point sources ≥100 tons per year (tpy). • Achieve DARS score ≥0.7 for onroad mobile source inventory. • Achieve DARS score of ≥0.5 for nonroad mobile source inventory.
Completeness	• 100% of all point sources ≥100 tpy. • 90% of all other point sources. • Top 15 area sources listed in 1990 base year SIP inventory.
Representativeness	• Counties A, B, C, and D. • 1996 daily ozone season.
Comparability	• Results to be compared to 1993 inventory.

2.0 PROGRAM SUMMARY

This QAP provides written instructions for the technical and quality aspects associated with development of the 1996 Ozoneville emissions inventory. It is designed so that QA/QC procedures are implemented throughout the whole inventory development process. This will ensure that the inventory is as complete as possible, accurate, comparable, and representative of the MSA. Personnel involved with the inventory and their responsibilities are discussed in Section 4.0.

2.1 Major Program Components

Inventory tasks and QC procedures will include data checking by the inventory development team (IDT) throughout the development of the inventory and final emission report. These procedures include, but are not limited to, the following:

- The development and implementation of written procedures for data gathering, data assessment, data handling, calculation of emissions, and reporting;
- Adequate management and supervision of the work;
- Review of all calculations for technical soundness and accuracy, including verification that the appropriate emission factors were used and the impacts of controls were correctly addressed;
- Correct assignment of source category codes;
- Assignment of DARS scores;
- Use of technically sound approaches when developing results based on engineering judgment;
- Documentation of the data in a manner that will allow reconstruction of all inventory development activities; and
- Maintenance of an orderly master file of all the data gathered and a copy-ready version of the final inventory submitted to the U.S. EPA.

QA activities are distinguished from QC activities in that they provide a more objective assessment of data quality because QA personnel are not directly involved in the development of the inventory. QA activities are usually more comprehensive because they include assessments of the effectiveness and appropriateness of the systems established by management to control data quality.

The QA program is equal in importance to inventory development and QC procedures, and includes a series of well-planned audits and training sessions that will be conducted by a trained QA staff. QA staff will not be involved in the development of the emission inventory in order to provide an objective assessment of the quality of the work and data.

The value of the QA program is that it highlights the effectiveness of the inventory development program. It can identify points in the process that require improvements in quality in a timely manner.

Audits will be conducted during the collection of the data, calculation of emissions, and development of the final report to determine whether QC requirements specified in this QAP are met. The audit schedule includes quality evaluations at critical points in the inventory development process where data quality could be compromised. The QA Coordinator is also authorized to conduct additional audits, if needed, to ensure that corrective actions are implemented as planned and the work is progressing according to the proposed schedule. The critical phases of the inventory development process and points at which data quality and technical systems audits will be conducted are identified in Figure 2-1. The audit schedule, the identities of the QA auditors, and estimated level of effort for each audit are provided in Figure 2-2.

QA training will be conducted to ensure that the IDT is aware of the quality issues of concern and expectations of the auditors. This training will include an overview of QC requirements and items on the audit checklist. Training will also be conducted to improve compliance with QA/QC requirements, as needed.

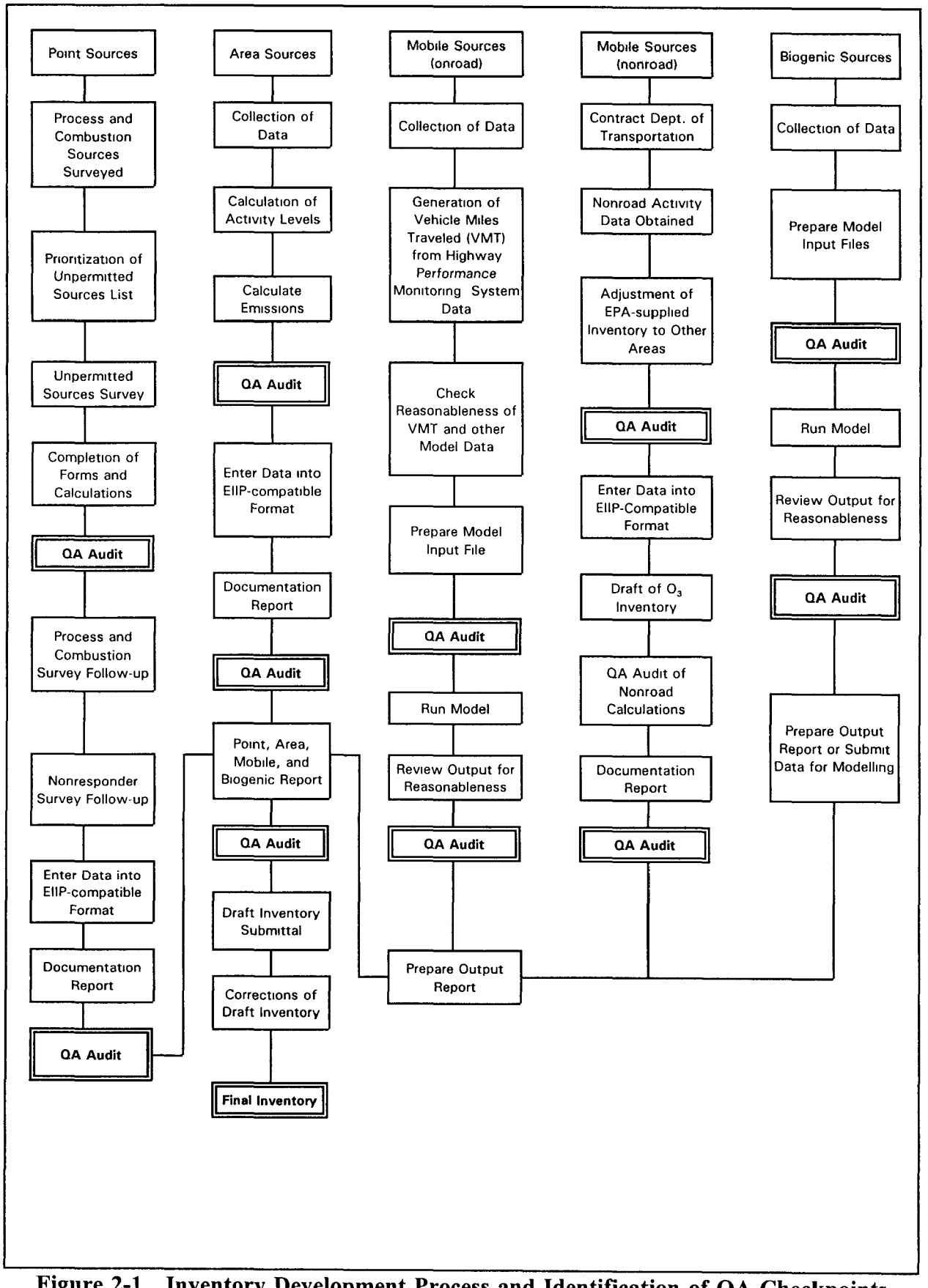

Figure 2-1. Inventory Development Process and Identification of QA Checkpoints

Date*	Audit Type	Objective of Audit	Auditor	Estimated Level of Effort (Hours)
March	Technical Systems Audit	Evaluate data documentation and data file maintenance procedures; supervise activities; train personnel.	S. Carr	36
May	Data Quality Audit	Determine accuracy of sets of area, point, mobile, and biogenic data entry into spreadsheets and emission calculations.	G. Marts	≤8
July	Technical Systems Audit	Evaluate adequacy of computer system: accessibility, memory capacity, ease of use, response time, etc.	S. Carr	24
August	Technical Systems Audit	Evaluate effectiveness of data review and corrective action process; check documentation of reviews, corrective actions, and follow-up activities.	G. Marts	12
October	Data Quality Audit	Determine accuracy of sets of area, point, mobile and biogenic data entry into spreadsheets and emission calculations.	S. Carr	≤8
November	Data/Report Audit	Verify accuracy of emission results included in draft report and compliance to reporting requirements.	G. Marts	24
December	Data/Report Audit	Check accuracy of at least 20% of the emission results included on each page of the report; check completeness of data file.	S. Carr	12

*Note: The schedule will be revised, as needed, to include critical phases of inventory development for each source type (point, area, mobile, and biogenic).

Figure 2-2. Ozoneville MSA 1996 Emissions Inventory Audit Schedule

2.2 Inventory Constraints

The QA Coordinator and Inventory Development Manager have been made aware of several constraints that will impact the inventory development process. The primary constraints are the availability of adequate state funds, inadequacies in the state's computer system, and limited access to the state's computer system. In response, the QA Coordinator has made the Director of the Ozoneville Department of Environmental Quality (ODEQ) and the Inventory Development Manager aware of the following impacts that these constraints could have on the DQOs and deadline for completing the inventory:

(1) Limited funds could mean that the resources required are not available to complete the inventory by the deadline and meet the data quality goals. Additional personnel recommended to complete the inventory on time and within budget cannot be hired.

(2) The computer system deficiencies identified during a previous audit cannot be improved at this time. (See previous audit finding on recommendation for corrective action form included as Figure 2-3.) The computer system currently used by the state does not have sufficient memory to handle some of the emissions programs designed to perform calculations and manage the data. This could potentially affect the accuracy and completeness of the inventory because the computer will not be used exclusively to manage the data, perform routine data searches, and calculate emissions.

(3) Limited access to the computer system also may mean that many calculations, data evaluations, and data searches will have to be conducted manually. Consequently, there is a higher risk of human error and it will take longer to complete the inventory.

Table 2-1 summarizes the possible impacts that these constraints may have on the inventory development process as well as the deadline for submitting the inventory to the U.S. EPA.

The U.S. EPA has been made aware of these constraints and the need for additional funds. The QA Coordinator has designed a QA program to minimize the negative impacts that these constraints could have on the quality of the data. The program includes modification of the QA/QC approach and procedures to include the following:

RECOMMENDATION FOR CORRECTIVE ACTION

A. Initial Information

RCA NO 0001	DATE: 9/2/93	URGENCY LEVEL 1,2
ORIGINATOR Steve Carr	APPROVED BY: NHD	1 Potential for major revisions needed. 2. Potential for failure to achieve data quality objective. 3 Suggested improvement.
ORGANIZATION/INDIVIDUAL RESPONSIBLE FOR ACTION· ODEQ, Dave Jones		

B. Problem Identification

SITE/LAB· Agency Office (DEQ)	SYSTEM. Computer Support	DATE PROBLEM IDENTIFIED. 9/1/93
DESCRIPTION OF PROBLEM: The computer system does not have enough memory to store emissions data and calculate emissions results. The response time is also very slow. Therefore, time is wasted waiting for the system to respond to commands.		

C. Recommended Corrective Action

DESCRIPTION·	IMPLEMENT BY:
Provide sufficient memory and improve the response time of the computer system to allow the users to use the system to calculate emissions and manage the data in an effective manner.	

D. Problem Resolution

PLANNED CORRECTIVE ACTION:	PROPOSED BY Jim Hall, Monica White, D. Jones	DATE PROPOSED: 9/15/93	SCHEDULE IMPLEMENTATION: 2/15/94
The computer system will be upgraded to provide sufficient memory to manage the emission inventory data and calculate emissions results. This upgrade will also improve the response time.			
IMPLEMENTED CORRECTIVE ACTIONS:			DATE IMPLEMENTED:

E. QA Verification

VERIFIED BY.	DATE:	COMMENTS:

Figure 2-3. Recommendation for Corrective Action Form

TABLE 2-1. POSSIBLE EFFECT OF CONSTRAINTS ON OZONEVILLE MSA 1996 EMISSIONS INVENTORY

Identification of Constraint	Impact on Inventory			
	Not Representative	Incomplete	Less Accurate	Deadline Not Met
Insufficient Funds	✓	✓	✓	✓
Inadequate Computer System		✓	✓	✓
Limited Access to Computer System			✓	✓

- Prioritization of categories so that resources will be allocated preferentially to critical data and sources;
- Additional meetings to discuss the status of the work, results from data reviews/audits, and corrective actions;
- More comprehensive QA audits; and
- Cross training of the IDT.

3.0 PURPOSE AND POLICY STATEMENT FOR THE 1996 OZONEVILLE STATE IMPLEMENTATION PLAN OZONE NONATTAINMENT EMISSIONS INVENTORY

This point, area, mobile, and biogenic emissions inventory is being developed in response to the requirements specified in the Clean Air Act Amendments of 1990 for VOCs, CO, and NO_x emission estimates. Because of the potential usefulness of the information gathered, every effort will be made to generate data that are accurate, representative, and comparable.

In order to provide data of sufficient quality, ODEQ has developed this QAP. It includes all of the critical elements recommended in the U.S. EPA document, *Guidance for the Preparation of Quality Assurance Plans for Ozone/Carbon Monoxide State Implementation Plan Emission Inventories* (U.S. EPA, 1988), as well as guidance provided through the EIIP (EIIP, 1996).

Implementation of the QA and QC procedures described in this QAP is fully supported by the IDT, the Inventory Development Manager, and the QA Coordinator. This support is evidenced by their commitment to implement the procedures as described in this QAP and generate data of known quality.

QC procedures described in this document were developed with the consensus of the QA Coordinator. The procedures were developed to provide a comprehensive program that includes QC measures that are implemented by the IDT, as well as QA measures that are implemented independently by the QA Coordinator and other QA personnel.

It is the responsibility of the IDT to report deviations from the procedures described in the QAP immediately to the Inventory Development Manager and QA Coordinator. The impact of the deviations on the inventories will be evaluated and the appropriate corrective actions will be taken to ensure that the technical and DQOs are met.

The Inventory Development Manager and QA Coordinator have worked with the Air Pollution Control Director and the Director of ODEQ to assure that adequate resources and a sufficient number of trained personnel are provided to meet the objectives of the work and to meet the deadline for submitting the inventory to the U.S. EPA.

Air Pollution Control Director	_____	Date _____
Director of ODEQ	_____	Date _____
Inventory Development Manager	_____	Date _____
QA Coordinator	_____	Date _____

4.0 EMISSIONS INVENTORY PREPARATION PLAN

The organization of the IDT and the QA team results in supervision of all aspects of the inventory development activities and independent QA audits, and establishes a direct line of communication to the State Director of Air Pollution on unresolved QA issues. The organization chart is presented as Figure 4-1. The responsibilities of the key personnel identified on the organization chart are discussed below.

4.1 Managerial Responsibilities

The Air Pollution Control Director, Dr. Mary Clean, will work with the U.S. EPA to obtain additional resources to complete the inventory by the agreed upon deadline. Meetings will be held routinely with the Environmental Quality Department Director, Dave Jones, and the QA Coordinator, Steve Carr, to keep Dr. Clean informed of the status of work. Dr. Clean will also handle any unresolved quality issues that may arise as a result of a QA audit.

Mr. Dave Jones, Director of ODEQ, is ultimately responsible for compliance with the requirements specified in the Clean Air Act Amendments. He has designated the person in the agency who will be responsible for the management of the inventory development activities (the Inventory Development Manager) and will work with the Inventory Development Manager, Air Pollution Control Director, and EPA to provide the resources needed to complete the inventory.

Ms. Monica White, Inventory Development Manager, will plan and manage all inventory development activities. She is responsible for the development of the QAP and final emissions report. Ms. White has assigned the work to four Task Leaders, will inform the ODEQ Director of the need for additional resources, and will routinely hold status meetings with the Task Leaders to keep them informed of the progress of the work and quality concerns. Recommendations for corrective actions will be forwarded to Ms. White for resolution, and she will conduct internal follow-up activities until QA concerns are

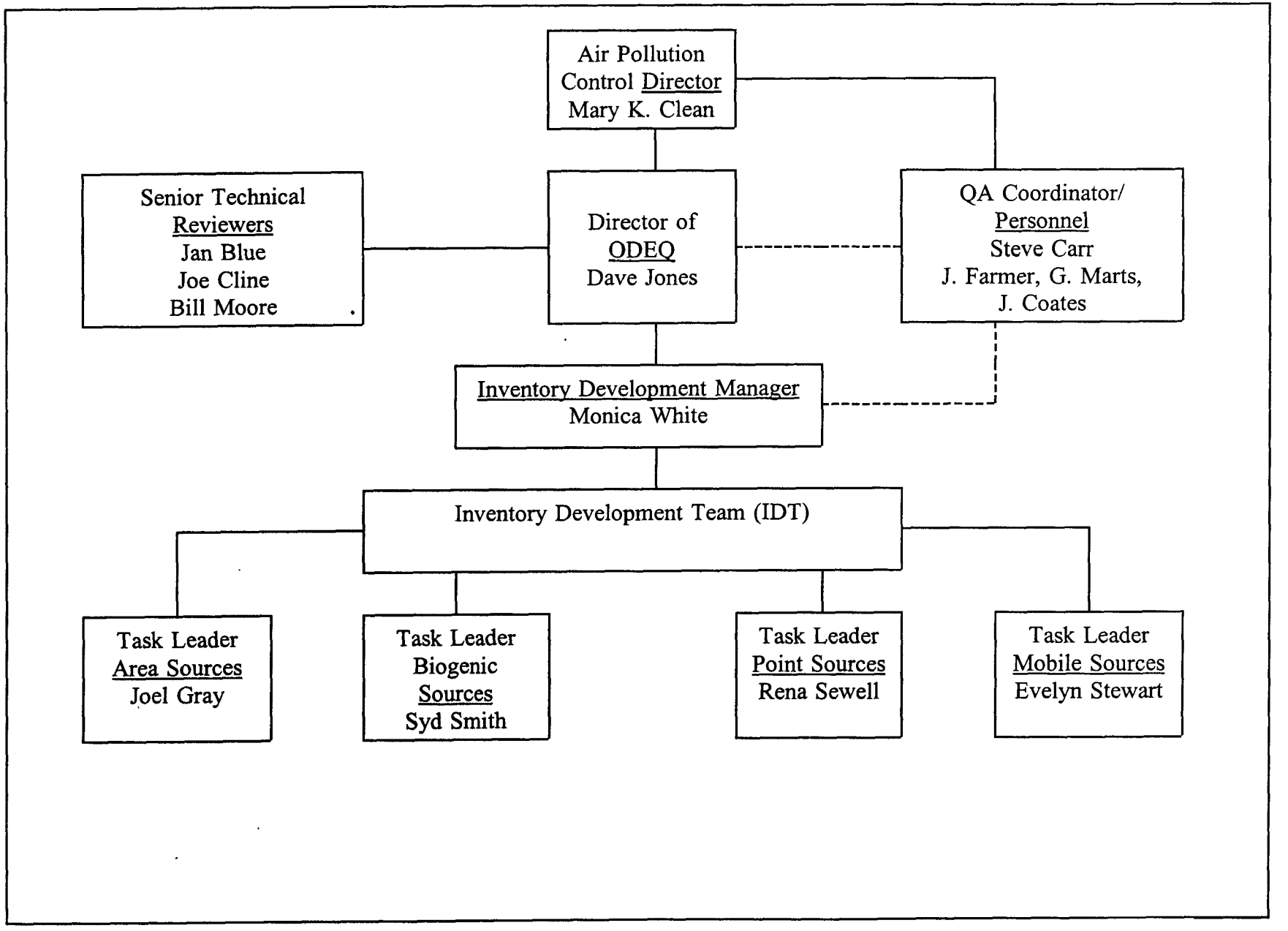

Figure 4-1. Ozoneville MSA 1996 Emission Inventory QA Organization

eliminated. She will provide technical training and meet regularly with the QA Coordinator to discuss any trends that may be obvious from reviewing the audit findings. QC procedures will be revised accordingly to maintain continuous improvement of data quality.

4.2 Inventory Development Task Leaders

The inventory development Task Leaders are Joel Gray, Syd Smith, Rena Sewell, and Evelyn Stewart. They will help plan and conduct the technical aspects of the development of the inventory, supervise daily IDT activities, and help develop the emission inventory report. They will work closely with the IDT to help identify ineffective QC procedures, and, if necessary, make recommendations to the Inventory Development Manager on revisions needed in QC procedures. Task Leaders will also monitor the status of the work, review data calculations, and make arrangements to have the data reviewed by senior technical reviewers.

4.3 Inventory Development Team Members

IDT members report to their respective Task Leaders. They will collect data, calculate emissions, and participate in QC reviews of work completed by other IDT members. Each team member is responsible for maintaining a complete data file and documenting all of their activities in a manner that will allow verification of the emissions reported and source of the supporting data. Team members have been or will be trained by the Task Leaders and the Inventory Development Manager.

4.4 Technical Reviewers

Senior technical reviewers for the inventory are Jan Blue, Joe Cline, and Bill Moore. These reviewers were selected because of their experience collecting emission inventory data and calculating results. Their experience will be used to assess the technical soundness, accuracy, reasonableness, representativeness, completeness, and comparability of

the data. They will summarize the results from their reviews in written reports and inform the Inventory Development Manager of findings that could compromise data quality.

4.5 Quality Assurance Coordinator

The QA Coordinator, Steve Carr, helped develop this QAP. His input ensures that adequate QA/QC procedures are incorporated into the inventory development process. He will conduct QA training and revise the audit schedule as needed so that all critical phases of the inventory development process are audited prior to generation of the emissions report. Steve will routinely attend status meetings held by the Inventory Development Manager and use the information from these meetings to revise the audit schedule, when appropriate, to ensure that the audit objectives are met.

The QA Coordinator will categorize and use the audit findings to evaluate the effectiveness of QC measures and QA audits. The QA/QC program will be revised to address trends that suggest that the technical and data quality objectives are not being achieved.

A summary of the QA Coordinator's responsibilities and activities follows and will be to:

- Help develop the QAP;
- Develop the audit checklists and audit schedule;
- Provide QA training to inventory development and QA personnel;
- Attend inventory development status meetings;
- Schedule audits, conduct audits, and report findings;
- Evaluate audit findings to determine if trends exist, and keep management informed of the results;
- Follow up on recommendations for corrective actions;

- Keep the Inventory Development Manager, Environmental Quality Department Director, and Air Pollution Control Director informed of the audit results;

- Work with the Air Pollution Control Director to resolve any quality concerns that cannot be resolved at the inventory management level; and

- Maintain a file of audit findings and corresponding corrective actions.

The QA Coordinator reports directly to the Air Pollution Control Director and indirectly to the managers overseeing development of the inventory. These reporting lines will help provide an objective approach to implementation of the QA program and reporting of quality issues.

QA personnel assisting the QA Coordinator may be employees of ODEQ who are also not actively involved in the development of the inventory, and have the technical experience to evaluate the technical soundness of the data and inventory development systems. However, the QA Coordinator is ultimately responsible for implementing the audit program, reporting audit findings, and conducting follow-up activities.

5.0 GENERAL QA/QC PROCEDURES

QA/QC procedures described in this QAP were developed to help ensure data accuracy, completeness, representativeness, and comparability. These procedures have been incorporated in the technical procedures, where applicable, and will be implemented by the IDT throughout the planning, data collection, emission estimation, and reporting phases of the inventory development program.

5.1 QC Activities

QC procedures will be implemented by the IDT during inventory development to meet the technical and DQOs. These activities will be conducted at critical steps in the inventory development process where the successful outcome of inventory development could be compromised. These critical steps are presented below and discussed in the following subsections of this QAP:

- Data gathering;
- Data documentation;
- Calculating emissions;
- Data checking and DARS scoring;
- Reporting; and
- Maintenance of the master file.

5.1.1 Data Gathering

Data gathering will be conducted according to U.S. EPA-approved procedures. The approach and supporting documents or references will be thoroughly documented and included in the emissions report.

The documents identified in Table 5-1 will be used to determine the best data-gathering approach for each emissions source type. Some data sources identified in these documents are also listed in Table 5-1. All data sources will be thoroughly documented in bound notebooks by the IDT. The IDT will also document when required data needed for specific source categories cannot be obtained or do not apply. The reason for not including a source or source category in the inventory will be clearly explained in the documentation.

TABLE 5-1. DATA COLLECTION GUIDANCE DOCUMENTS

Source Type	Guidance Document	Suggested Data Sources
Point	EPA-450/4-91-016 AP-42 EIIP Volume II	Existing inventories, state permit files, facility surveys, county business directories, telephone directory
Area	EPA-450/4-91-016 EIIP Volume III	Existing inventories, example cases, and data sources
Nonroad Mobile	EPA-450/4-81-026d	Existing inventories, example cases, and data source
Onroad Mobile	EPA-450/4-91-010 EPA-450/4-91-011 EIIP Volume IV	Transportation or planning agency data
Biogenic	EPA-450/4-91-010 EIIP Volume V	Crop acreage and land use; meteorological data by county

5.1.2 Data Documentation

Previous audit findings and comments from data reviewers have helped to emphasize the need for good data documentation procedures when developing an emissions inventory. Therefore, data documentation requirements have been developed for the IDT to facilitate the validation of the final emissions results. The documentation requirements will help ensure that all data needed for the emissions final report are gathered, maintained, and retained in the master file.

All activities conducted by the IDT on a daily basis will be documented in bound notebooks with indices to facilitate the retrieval of recorded information. A notebook will be assigned to each team member and it will only be used to record information relative to the development of the inventory. This daily log of activities will help another IDT member reproduce the emission results and allow an evaluation of data accuracy and completeness.

The following procedures are to be followed when documenting data in the notebooks:

- Data will be recorded legibly and in black ink;
- Entries will be corrected by drawing a single line through the data and writing the correct data above or below the correction (with initials, date, and explanation of corrections to allow reconstruction of the work);
- Complete descriptions of all data sources will be included (references to be included in final inventory report);
- Units of measurement will be provided with each data value;
- An explanation will be provided for emission sources that are omitted from the final inventory (justification required in report);
- The procedures used to calculate emissions will be described and example calculations will be provided;
- The approach used to determine completeness for each source type will be described;
- Documents from which emission factors are taken will be identified and referenced; and
- The source, agency, group, or company providing information by telephone will be identified (include telephone number and date information was provided).

Worksheets and contact reports may also be used to maintain records of data sources or calculations; however, the same guidelines must be followed when recording

information on them. A file will be developed specifically for these forms to ensure that they are retained and are easily located when the data are needed to calculate emissions. The contact report form that will be used is shown in Figure 5-1.

All worksheets, electronic spreadsheets, and notebooks will be reviewed periodically by the inventory development task leaders to determine whether the procedures described above are being followed. This review should be evidenced by a dated signature on the notebook pages or worksheets reviewed (i.e., reviewed by _____ on _____.)

Examples of acceptable documentation practices for each source type are included in the next section of this QAP.

5.1.3 Calculating Emissions

Information on how point, area, mobile, and biogenic emissions will be calculated is provided in Sections 7.0 through 10.0.

5.1.4 Data Checking and DARS Scoring

Data checking by the IDT is used to ensure data accuracy. Data will be checked at logical steps in the development of the inventory where transcription or calculation errors are likely to be found. Data checking will also be used to assess the technical soundness of the data. QA checkpoints are depicted in Figure 2-1.

Although different types of data will be reviewed at each checkpoint, the type of review may also vary. For example, when a document containing information is first received and logged in, it will first be checked to see if it was generated in the correct year and is for the correct location. Later, as data are used in calculating emissions, checking will include evaluations of data accuracy, reasonableness, and completeness.

CONTACT REPORT

Date _____ Originator _____

CONTACT BY: TELEPHONE ___ MEETING ___ OTHER _____

NAME, TITLE, & ORGANIZATION
ADDRESS & TELEPHONE NUMBER
PURPOSE OR SUBJECT (Give project number if appropriate)

SUMMARY:

ACTION:

Figure 5-1. Contact Report

The most logical checkpoints for each review are after data entry and calculations are performed. Data can be checked by another IDT member, by the Task Leader, or a senior technical reviewer (see Figure 4-1). If errors are found during these reviews, the person generating the data and reviewer must agree on the corrective action to be taken and see to it that the error is eliminated. They must also determine the impact, if any, that the error will have on other relevant data, and revise the affected data accordingly.

The results from data checking will be documented to further qualify the emissions estimates. In addition to the DARS scores assigned, the number of data points checked assists reviewers in evaluating the accuracy of the completed emissions report. Documentation of DARS scoring and data checking should include descriptions of the rationale for scoring, the data checked, and the dated signature of the reviewer. Figure 5-2 presents an example area source form that will be used to document the data checked and the findings.

5.1.5 Reporting

The emissions inventory report will be formatted according to the instructions provided by the U.S. EPA. Prior to finalizing the report, all of the actions taken in response to the recommendations for corrective actions will be evaluated to determine whether the report accurately reflects the corrections made. The report will be reviewed for technical soundness, completeness, accuracy, comparability, and representativeness by senior technical reviewers, editors, and QA Personnel.

It is the responsibility of the Inventory Development Manager to ensure that the report accurately reflects the data and that the master file provides sufficient data to verify the results reported. A copy-ready master of the report will be retained in the master file and made available to all project personnel.

CATEGORY DESCRIPTION:

SOURCE CODE:

INVENTORY REGIONS COVERED:

NAME (person responsible for calculations):

SIGNATURE (QC Review 1) _____ DATE _____

IS SECOND QC REVIEW NEEDED (YES/NO)? _____

SIGNATURE (QC Review 2) _____ DATE _____

SIGNATURE (QA Review) _____ DATE _____

Checklist	QC1	QC2	QA
1. Were all emission factor sources clearly referenced?			
2. Were correct emission factors used?			
3. Were sources of activity data clearly referenced?			
4. Were correct activity data used?			
5. Were the correct seasonal adjustment factors (SAFs) and activity days per week used?			
6. Were all calculations documented?			
7. Are calculations correct?			
8. Were applicable regulations cited?			
If so, were rule effectiveness (RE) and rule penetration (RP) correctly applied?			
9. Were nonreactive VOCs excluded?			

Figure 5-2. Area Source Category QA/QC Checklist

The following problems need to be corrected (indicate whether QC1, QC2, or QA and sign):

Figure 5-2. (Continued)

5.1.6 Maintenance of the Master File

The master file is a compilation of all data gathered and produced during development of the inventory. It should include sufficient supporting data to verify the accuracy of the emissions results reported. Indexing procedures must facilitate data retrieval.

Maintenance of the master file will begin with retention of this QAP. All correspondence and data received concerning development of the inventory will be filed by source and county. References will be maintained, along with applicable data contained within each reference.

The master file will be maintained by one of the IDT members or an administrative assistant. Access to the file will be limited to the IDT and controlled so that it is maintained in an orderly manner and is complete. A sign-out sheet will be used and is shown in Figure 5-3. A log will also be used to document data receipt and retrieval. File identification numbers will be assigned to the data and used to facilitate retrieval.

Copies of pertinent data will be made to provide working copies for the IDT; however, the original documents will remain in the master file.

5.2 QA Activities

QA activities are distinguished from QC activities in that they provide a more objective assessment of data quality because QA personnel are not directly involved in the development of the inventory. QA activities are usually more comprehensive because they include assessments of the effectiveness and appropriateness of the systems established by management to control data quality. This includes evaluations of the management and supervision of the work.

File Number	Check Out		Check In	
	Date	Initials	Date	Initials

Figure 5-3. Master File Sign-Out Sheet

QA activities will include QA training and the conduct of a series of independent audits to assess the effectiveness of the entire QC system and management of inventory development activities. These activities will be conducted by the QA Coordinator and other adequately trained personnel who are not involved in the inventory development process.

5.2.1 Training/Education

Initial training sessions will be conducted to discuss the items on the audit checklist and QA/QC requirements specified in this document. The QA Coordinator will conduct additional training when audits reveal the need for more QC measures or revision of existing procedures.

Occurrences that may lead to additional QA training include the following:

- An audit reveals a lack of understanding of QA/QC requirements or the need to develop additional QC measures.
- An audit reveals the need to provide guidance on acceptable data handling procedures because data are not maintained in a manner that allows easy verification of the accuracy of emission results and source of the supporting data.
- An audit reveals unacceptable data documentation practices that lead to entry errors and an inability to recreate emission results.
- An audit reveals that the internal data reviews do not adequately control data entry and calculation errors.
- The Inventory Development Manager requests QA training for new IDT members.

The QA Coordinator will conduct the training and maintain records of each training session.

5.2.2 Audits

Audits will be conducted according to the schedule presented in Figure 2-2. They will include assessments as the inventory development process is being planned, during data collection, as emissions are being calculated, and when the results are reported. The primary goal of the audit program is to prevent quality concerns. Opportunities to incorporate preventive measures will be included during each audit.

Prior to announced audits, the auditor will inform the persons to be interviewed of the date and time of the audit and data/systems to be reviewed. All of the personnel involved in inventory activities being audited will be asked to be available to respond to questions about their duties. The responses will then be compared to the requirements specified in this document and other referenced documents to determine compliance with approved procedures.

Questions that will be asked by the auditor and data to be evaluated during the audits vary by source type. The QA Coordinator will use the results of previous audits and work with Inventory Development Manager to develop source-specific audit checklists. Example checklists for technical systems and data audits are provided in Appendices A, B, and C. After developing the checklists, QA training will be held to inform the IDT of the steps in the inventory development process that are considered to be of concern. These meetings will be used to increase the team's awareness of the points in the inventory development process that may require more QC and managerial oversight.

Data audits will be conducted after major data transcriptions and calculations. The auditor will evaluate consistency in data entry and manipulation between IDT members. The results from the audits will be immediately shared with the data generator, Task Leader(s), and Inventory Development Manager. IDT members involved in the audit will be asked to respond immediately to the findings that will be informally discussed after the audit.

In addition to data quality audits, systems audits will be conducted to determine whether the procedures used are effective to collect data, document inventory development activities, and maintain the data. Systems audits will also include assessments of the supervision of the work and review of the data by the senior technical reviewers. These systems audits are used to evaluate the need to revise or develop additional procedures.

As the audits are conducted, the auditors will also assess how well the IDT is meeting the DQOs. Audit findings that reveal any compromises in data integrity or identify something that interferes with the achievement of these objectives will be brought immediately to the attention of the Task Leader and Inventory Development Manager. Solutions will be found for the problems identified, and the actions taken will be monitored until the quality issues are resolved.

5.2.3 Reporting Audit Findings

Audit findings will be documented on the audit checklist. The audit checklist and notes will be used to summarize the preliminary findings for the IDT, Task Leaders, and Inventory Development Manager after the audit. The checklist and notes will also be used to develop the audit report, which will be included in the QA documentation section in the inventory report.

The audit report will describe each deviation from approved procedures or finding that could compromise the successful outcome of the inventory. Documentation of each finding will include a description of the action or data reviewed that led to the quality concern and recommendation for corrective action.

The audit report will at least include the following information:

- Name of auditor, Inventory Development Manager, and IDT members audited;
- Audit date;

- Audit type;
- Audit objectives;
- Audit findings; and
- Recommendations for corrective actions.

Audit reports will be distributed within two weeks of the conduct of each audit to the persons interviewed, Task Leaders, and Inventory Development Manager. A summary of the types of quality concerns found will be periodically forwarded to the ODEQ Director to keep him informed of the quality issues found and actions being taken to resolve them. Audit reports will be retained in a QA file and used to conduct subsequent audits and plan follow-up activities.

6.0 CORRECTIVE ACTION MECHANISMS

Recommendations for corrective actions are made as quality concerns are identified. Recommendations for corrective actions will be formally presented in the audit reports. The corrective action form included as Figure 2-3 will be used to document the findings and actions implemented in response to each recommendation for corrective action. Original forms will be retained in the audit file, and copies will be distributed to the Task Leaders and the Inventory Development Manager. The information on the corrective action forms will be used by the auditor to monitor the types of problems found and the phases of the inventory development process that may need additional QC to eliminate recurring quality concerns.

The urgency of the responses is determined by the category of the finding. The categories are:

Priority 1: Potential for major revisions needed;

Priority 2: Potential for failure to achieve DQOs; and

Priority 3: Suggested improvements.

Priority 1 and 2 findings will be immediately brought to the attention of upper management and the planned implementation dates for the corrective actions will be as soon as possible after the audit. The Task Leader and Inventory Development Manager will meet routinely with the QA Coordinator to discuss the actions taken in response to the recommendations for corrective actions for Priority 1 and 2 findings until the quality concerns are resolved. The planned implementation date for Priority 3 findings may be later than the dates proposed for implementing actions taken in response to Priority 1 and 2 findings.

Follow-up activities will be conducted as frequently as required by the Task Leader and auditors to determine whether the recommended actions are taken. Each effort to

assess the implementation of the corrective action will be documented and maintained by the QA Coordinator in an audit file established for these records.

Follow-up activities could include the conduct of additional audits or informal assessments of the data or system of concern. The type of reevaluation will be determined by considering the impact that the quality concern could have on the technical and DQOs.

7.0 POINT SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES

For the purposes of this emissions inventory, point sources are defined as stationary, commercial, government, and industrial operations that emit more than 10 tpy of VOCs or 100 tpy or more of NO_x or CO. Point source data will be collected from two basic sources:

- The permitted sources database, and
- The small point sources database.

The latter was developed initially in 1992 by ODEQ for preparing the 1990 Base Year Inventory. It is updated every three years. The last update was completed in 1995 (the 1993 Periodic Inventory). Previous inventories included all sources not included in the permitted sources database.

The permitted sources database is maintained by ODEQ's Permits Branch. All facilities that have been issued permits are required to submit a comprehensive annual inventory that includes emissions from permitted as well as insignificant sources. However, exempt sources (e.g., wastewater treatment, housekeeping activities, on-site mobile sources)[1] are not included; exempt emission sources will be covered in the 1996 inventory as area, nonroad, or onroad mobile sources. The Permits Branch conducts a review of each facility inventory received, including checking for completeness, verifying emission factors, performing sample calculations of emissions from larger sources, and performing reasonableness checks of emissions and throughput values. Each facility is asked to verify questionable data to correct any errors. Inspectors within the Permits Branch continually check for new businesses and industries, as well as ones that have been closed recently. Their database, therefore, accurately reflects the 1996 operations of each permitted facility.

---

[1] The Permits Branch is currently evaluating the feasibility of requiring that facilities include these sources in their annual inventories.

ODEQ has found that this database is generally adequately reviewed for completeness, accuracy, and reasonableness. However, additional reviews will be conducted before it is used in the ozone nonattainment point source inventory:

- Records from ODEQ's Enforcement Branch will be reviewed to determine if facility inspections have identified any emission estimation inconsistencies; and
- The database will be reviewed to determine if any emission source data are missing or out-of-date.

The data for each facility will be printed (Facility Report 1) and sent to the contact listed in the database. The facility contact will be given 60 days to review the data, make any needed updates, and return the Facility Report to ODEQ. Follow-up calls will be made to resolve any issues found in the additional reviews described above.

After 90 days, the list of facilities that have not responded to the request for information will be prioritized by size. Calls will be made to as many facilities as possible to determine why the forms were not returned. Each call will be documented using the Contact Report form (Figure 5-1).

A similar procedure will be followed for the small point sources database. To identify any new sources that need to be added, the following will be checked:

- Construction permit records for 1993 to 1996 (for commercial or industrial properties);
- State employment data for 1993 to 1996; and
- Dun & Bradstreet and other listings of businesses in the Ozoneville MSA.

The lists generated from follow-up on the two databases will then be cross-checked with the existing list of unpermitted facilities. Any new source with a

commercial or industrial Standard Industrial Classification (SIC) Code will be sent an initial survey (Figure 7-1).[2] The existing sources will be sent a copy of their current data (Facility Report 1) and asked to update the report with new information. The same procedures described above for follow-up will be used.

Data for new sources or updates to existing sources will be entered into the Ozoneville Emissions Modeling System (OEMS). OEMS is an EIIP-compatible data system. For new sources, OEMS data entry forms will be generated by the IDT. Data will be transferred from the reports or forms sent by the facilities to OEMS data entry forms. Each of these forms will be reviewed by the IDT member who generated them and 100 percent of the data elements will be checked. Then, 20 percent of the data elements will be checked by another IDT member. If errors are found, they will be corrected and a second review will be made. This process will be repeated until the data forms are verified to be 100 percent correct. Forms with completed and verified data will then be sent to ODEQ's Data Management Branch for entry into OEMS.

ODEQ aspires to attain an error-free database; however, it will not be possible to perform checks on 100 percent of all the data elements. Furthermore, data updates occur every year in the Permits Branch database, introducing the possibility of new errors.

---

[2] Figure 7-1 presents an example survey form for stationary combustion sources only.

FORM EI-1 REPORTING YEAR 1996

ODEQ EMISSION INVENTORY QUESTIONNAIRE
STATIONARY COMBUSTION SOURCES

Please complete this form with data representing operations for the calendar year listed above.

BUSINESS NAME _____
BUSINESS ADDRESS _____
REPORTED BY _____
ODEQ FACILITY NUMBER _____ DATE _____
TELEPHONE _____ FAX _____

Company ID Number				
ODEQ Source No. (i.e., B001)				
Type of Fuel Used				
Coal, tons/year				
Ash, % (as fired)				
Oil (indicate type), gal/yr				
Fuel Sulfur Content, % (as fired)				
Gas, Million cubic ft/yr (please report MMCF)*				
Other Fuel (indicate), units/yr				
Fuel Heat Value, Btu/unit (as fired)				
Last Test Date				
% Annual Throughput Each Quarter: January - March				
April - June				
July - September				
October - December				
Operating Schedule: Hours/Day				
Days/Week				
Weeks/Year (or Hours/Year)				
Hours Operated Without Control Equipment				

* **Natural gas usage should be reported in million cubic feet (MMCF). The unit MCF is thousand CF. If billed in units other than MMCF, usage may be reported in those units provided the new units are clearly noted.**

COMMENTS:

Figure 7-1. Point Source Survey Form For Stationary Combustion Sources

8.0 AREA SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES

Based on a review of the 1993 periodic inventory, a checklist of area sources has been prepared, evaluated, and prioritized for inclusion in the 1996 ozone nonattainment area SIP inventory. Source categories have been ranked by VOC followed by NO_x emissions to identify the most significant area source categories for inclusion in the inventory. The area sources have been generally grouped from largest to smallest, with priority given to the larger sources. U.S. EPA and EIIP guidance, as well as inventories from two nearby states, were also reviewed to identify sources excluded from the 1993 periodic inventory. Emission estimates will also be developed for 1993 for these two categories for comparability. Two categories were added as a result of this analysis: slash burning and industrial wastewater treatment. Emission estimates will also be developed for 1993 for these two categories for comparability. Two source categories were also eliminated from the 1996 inventory: open burning and on-site incineration. The reasons for excluding these sources will be documented in the inventory report. The list of area sources to be included in the 1996 periodic inventory is summarized in Table 8-1.

Data will be collected by the IDT in the following formats:

(1) Mail and phone surveys. For a limited number of area source categories, data will be collected by means of survey forms mailed to facilities, or used as phone questionnaires by ODEQ personnel.

(2) Reports. Examples are population data from the U.S. Census Bureau, and energy use data from the U.S. Department of Energy and the Ozoneville Energy Office.

(3) Letters and memoranda. In many cases, information will be sent in letters or memoranda, or as database hard-copy reports with a transmittal letter.

(4) By phone or facsimile. Some information will be received over the phone and recorded on telephone contact log sheets; other information will consist of facsimiles showing hand calculations or data inputs. The information received in this manner is likely to come from agencies in ODEQ, other state and local agencies, and federal agencies.

TABLE 8-1. AREA SOURCES TO INCLUDE IN 1996 INVENTORY AND PROPOSED EMISSION ESTIMATION METHOD

Category	Estimation Method
Architectural surface coating	Use results of survey conducted in 1993. Update by resurveying five largest distributors.
Asphalt paving	Use U.S. EPA-recommended emission factor/gallon paving material. Obtain paving data from the state Department of Transportation.
Commercial bakeries	Use U.S. EPA-recommended production-based emission factor. Obtain production data from U.S. Census Bureau.
Consumer/commercial solvent use	Use preferred method in EIIP Volume III, Chapter 5 (EIIP, 1996). Obtain population data from U.S. Census Bureau.
Dry cleaning	Use preferred method in EIIP Volume III, Chapter 4 (EIIP, 1996).
Gasoline distribution	Use preferred method in EIIP Volume III, Chapter 11 (EIIP, 1996). Use MOBILE model to calculate vehicle refueling emission factors (U.S. EPA, 1996a). Information on gasoline dispensing outlets is available from the ODEQ.
Graphic arts	Use alternative method 1 in EIIP Volume III, Chapter 7 (EIIP, 1996). Obtain population data from U.S. Census Bureau.
Industrial surface coating	Conduct survey of facilities with SIC Codes identified by U.S. EPA in the Procedures document as industrial surface coating facilities (U.S. EPA, 1991a). Survey using a two-stage approach: (1) determine proportion of facilities with surface coating activity, and (2) obtain solvent consumption/disposal data from representative sample of applicable sources.
Industrial wastewater treatment	Use Surface Impoundment Modeling System (SIMS) emission estimation model (U.S. EPA, 1990).
Landfills	Use U.S. EPA's Landfill Air Emission Estimation Model (U.S. EPA, 1996b). Obtain activity data from ODEQ's Department of Solid Waste.
Pesticide application	Use alternative method 1 in EIIP Volume III, Chapter 9 (EIIP, 1996).
Slash burning	Use AP-42 emission factors (U.S. EPA, 1995). Obtain activity data from the state Department of Forestry.
Small stationary source fossil fuel combustion (residential, commercial, and industrial)	Use AP-42 emission factors (U.S. EPA, 1995). Obtain activity data from state Energy Office and Energy Information Agency. Apportion fuel consumption data to the county level based on residential population and employment (from U.S. Census Bureau data).
Solvent cleaning	Use preferred method in EIIP Volume III, Chapter 6 (EIIP, 1996).
Structure fires	Use U.S. EPA-recommended emission factors from Procedures document (U.S. EPA, 1991a). Obtain activity data from Ozoneville Fire Marshal.

All data will be retrieved from the appropriate agencies or libraries. Copies of all data, including paper and electronic formats, will be logged in, assigned a file identification number, and filed in the project file. Data sources will be clearly documented in the database and on worksheets.

The quality and completeness of data will be ensured by using the worksheets and completing the QA/QC checklist shown as Figure 5-2. Alternative sources of data will be evaluated; the basis for choosing one data source over another will be clearly documented. QA/QC procedures (if known) used by the agencies or individuals supplying the data will also be documented.

Emissions calculations will require the acquisition of data from a variety of different sources. Preferred data sources are shown in Table 8-1. Any deviations from the methods or data sources shown will be documented and explained. In addition, because improvements will be made to the emission estimates based on EIIP guidance, 1993 estimates will be prepared for any source categories for which revised emission estimation methods are used. The inventory report will also include a discussion of other area source categories for which improved estimation methods are needed.

Calculations will be performed on spreadsheets as much as possible. If handwritten calculations are necessary, they will be recorded and maintained in the project file. Emissions factors for most sources are available from EPA guidance documents; the EIIP area source volume will be the primary reference. Values for seasonal adjustment factors (SAFs) and activity days per week were developed for the 1993 periodic inventory. These values will be verified for applicability before use in this inventory. Emissions will be adjusted (if needed) to exclude nonreactive VOCs. Federal and state regulations will be reviewed to determine which area source categories are affected, particularly those for which rules were promulgated between 1993 and 1996. Rule penetration (RP) and rule effectiveness (RE) adjustments to the emissions will be made where appropriate.

Care will taken to avoid double-counting of emissions. The area source emissions will be adjusted after completion of the point source inventory to account for point source emissions.

All area source calculations will be reviewed as part of the QC program by the inventory developer; 10 percent will be reviewed by another IDT member. The person doing the QC will check data accuracy and the reasonableness of assumptions, and review the calculations. In general, the following considerations will be used to assess the accuracy of all calculations:

- Are the equations used for each method or procedure consistent?
- Are assumptions and engineering judgments documented and reviewed?

Spreadsheet audit functions will be used to verify that spreadsheet functions were used correctly.

If revisions are needed, the person who did the calculations is responsible for making them. If a second QC review is needed, it will be performed. Revisions to the computer files will be indicated on the "DATA LOG SHEET" and current copies of all computer files will be kept in the project file (Figure 8-1).

After emissions are calculated, they will be reviewed by the Task Leader. Outliers will be checked to ensure that the data, data processing, and calculations used are correct and acceptable. It should be noted that simply because the estimate is an outlier, i.e., it does not fall within the expected range, this does not necessarily denote an error. Such outliers will be examined, but additional time will be taken to assess outliers that occur in categories that have large emissions and those that differ greatly from the expected range. If an outlier is identified and the emission methodology and calculations are reasonable, the data will be flagged and appropriate EPA personnel will be contacted for guidance.

DATA LOG SHEET

SOURCE:

SOURCE CODE:

NAME (Person responsible for calculations):

Item	Date Completed	Name of Person Responsible
Calculations		
QC Review (1)		
Revisions		
QC Review (2)		
Data coded for OEMS		
Data entered in OEMS		
OEMS data QA		

Page 1 of 2

Figure 8-1. Area Source Data Log Sheet

DATA LOG SHEET

AREA SOURCE CATEGORY:

SOURCE CODE:

NAME (Person responsible for calculations):

Spreadsheet file names (copy of current versions should be kept in project file):

File Name	Date of Current Version	Contents

Page 2 of 2

Figure 8-1. Continued

When all QC reviews are complete and all necessary revisions have been made, the data will be entered into OEMS. If the data must be hand entered, assistance will be provided by the person who calculated the estimates. This person will also help QC the data after OEMS entry. The data entry personnel are responsible for completing the OEMS-related item in the table on the "DATA LOG SHEET."

9.0 ON- AND NONROAD MOBILE SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES

9.1 Onroad Mobile Sources

Onroad mobile emissions will be estimated using emission factors generated from the U.S EPA's MOBILE program and vehicle miles traveled (VMT) data generated primarily from the Highway Performance Monitoring System (HPMS). The Ozoneville nonattainment area encompasses the Federal Aid Urbanized Area (FAUA) and extends beyond it. Therefore, in addition to using Ozoneville FAUA HPMS data, HPMS data representing the rural portion of the state will be factored in to account for the portion of the nonattainment area outside of the Ozoneville FAUA. Local-road VMT within the nonattainment area will be calculated from countywide local road VMT estimates and added to HPMS-derived VMT.

Most of the resources expended in this portion of the inventory will be spent developing and reviewing (QC level) the data inputs for the MOBILE model, checking the reasonableness of the VMT estimates, and reviewing the final emission estimates. Some specific input data needed include vehicle registration and vehicle type distributions, fuel characteristics, vehicle operating conditions, and ambient temperatures. Data on these and other input data needed will be obtained primarily from the U.S. EPA's Office of Mobile Sources, the Ozoneville Department of Transportation, and the Ozoneville Department of Motor Vehicles.

QC procedures that will be implemented to ensure that the onroad mobile source inventory is of high quality will include data verification checks, comparison of model input data with original data, and comparison of overall emissions estimates with estimates developed for similar urban areas.

Developing an onroad mobile sources emission inventory requires data from a number of diverse organizations. These groups do not normally interact with one another, so

to help ensure the production of a high-quality onroad motor vehicle inventory, a working group of representatives from each organization will be formed. This group will identify the type and source of the needed data, and coordinate the data acquisition process. The group members and the primary data each will contribute will be:

- Mr. Dale Earnhart, Ozoneville Association of Governments. The Association of Governments will have primary responsibility for calculation of the onroad vehicle emissions, and will also contribute data developed from the regional transportation model. Mr. Earnhart will chair the working group.

- Ms. Shirley Muldowney, Ozoneville Department of Transportation. The Department of Transportation will contribute HPMS data and countywide VMT estimates from local roads.

- Mr. Al Unser, Ozoneville Department of Motor Vehicles. The Department of Motor Vehicles will contribute vehicle registration distribution data.

- Mr. A.J. Foyt, ODEQ. ODEQ will contribute input data for the MOBILE model, including vehicle inspection and maintenance parameters and fuel vapor pressure data.

- Mr. Michael Andretti, ODEQ. Mr. Andretti will provide independent review of the data flow and calculations.

The VMT data needed to cover the nonattainment area are:

- HPMS data for the Ozoneville FAUA, which includes Ozoneville as well as Counties A, B, C, and D.

- HPMS data for the portion of the Ozoneville nonattainment area in Counties A and B that is outside of the Ozoneville FAUA; and

- Local-road VMT for urban and rural roads for the nonattainment area.

HPMS data for the Ozoneville FAUA will be used in its entirety, because all of the FAUA is contained in the nonattainment area. For the portions of Counties A and B that are outside of the Ozoneville FAUA, HPMS data representing samples from rural roads in the state will be factored in based on the ratio of road mileage (by functional class) within the affected area versus the rural statewide total.

Local-road VMT will be estimated from existing countywide VMT estimates generated by the Ozoneville Department of Transportation. For those areas of Counties A and B that are only partially contained with the Ozoneville nonattainment area, local-road VMT will be estimated by factoring the countywide total by the ratio of local-road mileage within the nonattainment area to the total local-road mileage in the county. Note that these local-road VMT represent annual averages and must be factored in to correctly represent seasonal daily VMT. To achieve consistency with the treatment of the HPMS data, SAFs used in the Ozoneville FAUA to factor collector road VMT will be used to adjust the local-road VMT.

Specific QC measures to be taken to check the accuracy and reasonableness of the VMT estimates will include, at a minimum:

- Independent review of all calculations to verify application of the ratio techniques, use of correct expansion factors, and correct transferral of VMT input and output data;
- Comparison of the distribution of nonattainment area VMT by functional class to the statewide totals published in *Highway Statistics* (U.S. DOT) for 1996;
- Calculation of ratios of VMT on a per capita basis, per gallon of fuel used (by vehicle type after applying VMT ratios), and per road mile (by functional class). These ratios will be compared for reasonableness to similar ratios calculated for the state and other similar states from *Highway Statistics*; and
- Verification of the inclusion of all input HPMS and Ozoneville Department of Transportation local-road VMT data.

If an additional check is deemed necessary, a real-time simulation using one of the available vehicular flow models will be performed. This additional check will be performed if it appears that HPMS data have significantly underestimated truck VMT (and overestimated automobile VMT).

Emission factors will be generated using the MOBILE model from EPA. Inputs to the model will include a number of region-specific parameters as well as some national defaults. Region-specific parameters will include:

- Model year vehicle registrations distributions by vehicle class;
- Diesel sales fractions by model year;
- Statewide VMT distributions by vehicle class;
- Trip length distributions;
- Vehicle inspection and maintenance and antitampering program parameters;
- Vehicle speeds by functional class;
- Vehicle operating modes;
- Daily minimum and maximum temperatures; and
- Fuel characteristics including vapor pressure and oxygenate content.

MOBILE defaults of tampering rates and mileage accumulation rates will be used.

One of the critical areas of QC at this stage is to ensure the correct entry of data into the MOBILE model. The required input files are strictly formatted, but not amenable to easy interpretation. This raises the possibility of errors in the inputs going undetected. Therefore, two primary methods will be used to QC the use of the appropriate data as MOBILE inputs:

- Independent review of MOBILE inputs to ensure correct entry into the model, utilizing the MOBILE Input Data Analysis System (MIDAS) to provide a structure for checking the model inputs. MIDAS formats the model inputs for easier review, and performs additional QA to determine if the model inputs conform to specifications for the inventory type.
- Independent review of MOBILE model output files to determine if the model echoes the correct input parameters.

Once emissions have been calculated, a number of QC measures will be implemented to check the accuracy of the calculations and the reasonableness of the results. The planned QC measures will include, at a minimum:

- Independent review of the calculations, including review of data transfer from the MOBILE results to the calculation database, checking a sample of the calculations by hand for errors in the process, and review of the data transfer from the calculation database to the summary formats.

- Comparison of the results versus other independent variables to check for abnormalities in the calculations. These will include checks to:

  - Summarize emissions by pollutant on a per capita basis for each county;

  - Summarize emissions by pollutant versus VMT for each county. This provides a back-calculation of an overall emission factor that can be compared to separate MOBILE runs using average speeds and daily minimum and maximum temperatures, to check for unusual results.

  - Summarize emissions by pollutant and road functional class;

  - Compare the per capita, per VMT, and road functional class results to other inventories for the same year; and

  - Compare the percentage contribution by pollutant of onroad mobile emissions to the overall inventory to other inventories for the same year.

Variation of the results from the other inventories do not automatically indicate errors, but provide reasons to perform additional QC of the results.

9.2 Nonroad Mobile Sources

The U.S. EPA hired Energy and Environmental Analysis, Inc., to update the nonroad equipment category inventories for 33 ozone nonattainment areas in the United States. Although the Ozoneville nonattainment area is not one of the 33, the U.S. EPA concluded that the Washington, DC-Maryland-Virginia Metropolitan Statistical Area (MSA)

is similar in climate and economic conditions to Ozoneville. Therefore, a ratio of the populations of the two areas will be used in combination with the Washington, DC-Maryland-Virginia MSA inventory to develop the nonroad emissions estimates in the Ozoneville inventory.

There are three distinct U.S. EPA inventories that can be used to extrapolate nonroad emissions to the Ozoneville nonattainment area:

- Inventory A
- Inventory B
- (A+B)/2

The average of Inventories A and B will be used for nonroad mobile sources in Ozoneville based on guidance provided by the U.S. EPA.

QC procedures for the nonroad mobile source emissions estimates will consist of:

- Verification by the Task Leader and senior peer reviewers that the Washington, DC-Maryland-Virginia MSA is appropriate as a surrogate for Ozoneville;
- 100 percent QC of the spreadsheets provided by the U.S. EPA; and
- QC level reviews of data input (to identify transcription errors), calculations, and data conversions.

In assessing the applicability of the Washington, DC-Maryland-Virginia MSA inventory for Ozoneville, an overall comparison of the two regions and a detailed examination of key sources will be conducted. In the overall assessment, it will be determined if each of the sources present in the Washington, DC-Maryland-Virginia MSA inventory is also present in Ozoneville, or if potentially significant activities in Ozoneville are not present in the Washington, DC-Maryland-Virginia MSA inventory. For those sources that should be added to the Ozoneville inventory, a rough estimate of emissions will be made

based on data in the U.S. EPA inventories for more applicable regions. This rough estimate will be used to guide the level of effort employed to develop a more refined estimate.

In the detailed assessment, the key source types in the Washington, DC-Maryland-Virginia MSA inventory will be identified with the largest contribution of each pollutant. Additional attention will then focus on developing independent assessments of these sources for comparison purposes. For example, the key sources that contribute more than 10 percent of the total offroad emissions for either VOC or NO_x are:

- Construction equipment;
- Airport service equipment; and
- Lawn and garden equipment.

Developing emission estimates for these sources based on population ratios may or may not result in reasonable estimates. For example, for construction equipment, data will be collected on the number of permitted construction projects, and the Washington, DC-Maryland-Virginia MSA and Ozoneville data will be compared. Average emissions by equipment category will be compared between the two regions. If there is a difference between the two, the emissions estimates developed based on the permit data will be used provided it is complete.

10.0 BIOGENIC SOURCE INVENTORY PREPARATION AND QA/QC ACTIVITIES

There are three computer models that can be used to estimate biogenic emissions:

- Biogenic Emission Inventory System-2 (BEIS-2);
- The personal computer (PC) version of BEIS, PC-BEIS2.2; and
- Biogenic Model for Emissions (BIOME).

The BEIS models, BEIS-2 and PC-BEIS2.2, have default land use files for all counties in the U.S. (except those in Alaska and Hawaii). An alternative approach to using these models is to collect local information to substitute for model defaults. In the initial inventory planning phase, each of these methods was reviewed. On the basis of these reviews, it was determined that the preferred method for estimating biogenic emissions is with the U.S. EPA's BEIS-2 computer model. BEIS-2 is the most scientifically advanced model for estimating biogenic ozone precursor emissions.

BEIS-2 estimates VOC emissions for forested areas by multiplying the foliar density for each forest type by the appropriate emission factors. The emission rates are then adjusted for specific environmental conditions using user-supplied temperature and solar energy change data and output files of the Urban Airshed Model (UAM) temperature preprocessor. Similar calculations are used to estimate emissions from nonforested areas.

Meteorological data will be obtained from the National Weather Service and the U.S. EPA. U.S. EPA-supplied land use and biomass data sets provided with the model will be used.

In reviewing the estimates developed with BEIS-2, priority will be given to verifying the modeling days selected. In addition, the selection of the meteorological station will be reviewed. Typically the nearest station is recommended and used, but elevation and topography differences must also be considered in order to determine if there is a more

suitable station other than the closest. The default land use data in BEIS-2 will be evaluated to determine if they are representative of base year data, reality checks will be performed on genus/land use proportions, and the Federal Information Procedures System (FIPS) code will be checked to verify that the correct county data were used.

11.0 DATA REPORTING

Reporting will be accomplished by submittal to the U.S. EPA of written documentation and emissions summaries. The procedures, assumptions, sample calculations, and summary tables of emissions will be thoroughly documented in the ozone inventory report.

The report will include summary tables, raw listings of equipment, activity levels and emissions from individual sources, and a QA documentation section. A detailed inventory report allows comparison of baseline inventories from one area to another, the evaluation of the impact of control strategies, and facilitates updates to the inventory and development of projection inventories.

In addition to EIIP guidance, the 1992 U.S. EPA report *Example Documentation Report for 1990 Base Year Ozone and Carbon Monoxide State Implementation Plan Emission Inventories* will be followed (U.S. EPA, 1992a). These documents provide guidance for presenting and documenting SIP emissions inventories, and contain examples of how to present and verify inventory development efforts. The QA documentation section of the emissions inventory report will provide enough detail so that the inventory development described in the report can be compared to the information provided in this QAP. Any discrepancies will be identified and explained.

The QA documentation section of the inventory report will also include the audit report. As discussed previously, the audit report will describe each deviation from approved procedures or findings that could compromise the successful outcome of the inventory. Documentation of each finding will include a description of the action or data reviewed that led to the quality concern, along with a recommendation for corrective action. The QA documentation section of the inventory report will then discuss how the recommended corrective actions were implemented.

12.0 REFERENCES

Birth, T.L. 1995 User's Guide to the Personal Computer Version of the Biogenic Emissions Inventory System (PCBEIS2.2). Prepared for the U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, North Carolina.

Emission Inventory Improvement Program (EIIP). 1996. Volumes II-VI. Prepared by the State and Territorial Air Pollution Program Administrators and the Association of Local Air Pollution Control Officials (STAPPA/ALAPCO).

Federal Highway Administration. Highway Performance Monitoring System (HPMS). U.S. Department of Transportation, Washington, D.C.

Highway Statistics, U.S. Department of Transportation, Federal Highway Administration, Washington, D.C. Annual Publication.

U.S. EPA, 1996a. Mobile Source Emission Factor Model (MOBILE). Office of Mobile Source Air Pollution Control, Ann Arbor, Michigan.

U.S. EPA, 1996b. Landfill Air Emissions Estimation Model (LAEEM). Office of Research and Development, Research Triangle Park, North Carolina.

U.S. EPA, 1995. Compilation of Air Pollutant Emission Factors, 5th Edition and Supplements, AP-42, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1993. User's Guide for the Urban Airshed Model, Volume IV: User's Manual for the Emissions Preprocessor System 2.0, Part A: Core FORTRAN System, EPA-450/4-90-007D(R). Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1992a. Example Documentation Report for 1990 Base Year Ozone and Carbon Monoxide State Implementation Plan Emission Inventories. EPA-450/4-92-007. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1992b. Procedures for Emission Inventory Preparation, Volume IV: Mobile Sources, EPA-450/4-81-026d. (Revised). Office of Mobile Sources, Ann Arbor, Michigan.

U.S. EPA, 1991a. Procedures for the Preparation of Emission Inventories for Carbon Monoxide and Precursors of Ozone, Volume I: General Guidance for Stationary Sources, EPA-450/4-91-016. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1991b. Emission Inventory Requirements for Ozone State Implementation Plans, EPA-450/4-91-010. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1991c. Emission Inventory Requirements for Carbon Monoxide State Implementation Plans, EPA-450/4-91-011. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1991d. Nonroad Engine and Vehicle Emission Study Report. EPA-21A-2001. Office of Mobile Sources, Ann Arbor, Michigan.

U.S. EPA, 1991e. Procedures for the Preparation of Emission Inventories for Carbon Monoxide and Precursors of Ozone, Volume II: Emission Inventory Requirements for Photochemical Air Quality Simulation Models, EPA-450/4-91-014. Research Triangle Park, North Carolina.

U.S. EPA, 1990. Surface Impoundment Modeling System (SIMS) User's Manual. EPA-450/4-90-019a. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

U.S. EPA, 1988. Guidance for the Preparation of Quality Assurance Plans for O_3/CO SIP Emission Inventories, EPA-450/4-88-023. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

APPENDIX A

QUALITY ASSURANCE INVENTORY CHECKLIST

Auditor _____
Date _____

Personnel Interviewed _____

_____

This audit checklist is to be used to document the findings from the audit of activities and data associated with the Ozoneville emissions inventory. Use applicable parts of the checklist to identify the quality concerns associated with each task. Document the results and use them to generate the audit report.

I. MANAGEMENT OF THE WORK

A.	Is the QAP available to the personnel audited?	Y/N
B.	Are the procedures applicable to their work understood and followed?	Y/N
C.	Are the procedures adequate for the desired outcome of the work performed?	Y/N
D.	Are meetings held routinely to discuss the progress of the work and any quality problems that were found?	Y/N
E.	Are the personnel adequately trained to perform the duties assigned?	Y/N
F.	Are the resources required to perform the duties assigned available and adequate to achieve the objective of the work?	Y/N
G.	Is the work on schedule?	Y/N

II. DATA MAINTENANCE AND COLLECTION

A. Are the data used for the inventory coded to facilitate tracking? Y/N

B. Are the data organized to facilitate retrievals? Y/N

C. Does the data file include of all of the data required to estimate the emissions from a given source? (Check about 4-5 sources) Y/N

D. Are the data in a place where access is controlled and limited? Y/N

E. Are the data copied when requests are made for retrievals? Y/N

F. If originals are released to inventory development personnel, is the location of the original data documented in the data tracking database? Y/N

G. Is the data tracking database operational and used to track the receipt and distribution of the data? Y/N

H. Are the state permit applications and supporting data completed in a manner that will not lead to misinterpretation of the data? (Check for obscuring of data when making corrections, insufficient data to discern the identity and level of emissions of a given pollutant, unclear labels on attachments, etc.) Y/N

I. Are the data documented in black ink so that reproductions will include all of the data recorded on the data forms? Y/N

J. Are the data request forms complete? If not, what is done to acquire the missing data? Y/N

_____

_____

III. DATA EVALUATION

A. What steps were taken to ensure that the data collected are complete?

_____

_____

_____

B. What steps were taken to evaluate the accuracy, completeness, comparability, and representativeness of the data?

_____

_____

_____

C. What procedures were followed to eliminate double counting of sources or points within a source?

_____

_____

_____

D. How were sources below the cutoff point handled?

_____

_____

_____

E. Were task activities prioritized to provide emissions data about the highest emitters first? Y/N

F. Were discrepancies found in the data? If yes, what were they and how were they eliminated? Y/N

_____

_____

_____

G. Were calculations reviewed by another IDT member for technical soundness and accuracy? Y/N

Were results documented? Y/N

_____

_____

_____

H. Were evaluated data reviewed by a senior technical reviewer prior to entering it into the emissions database? Y/N

Were results from the data reviews documented and corrective actions implemented as requested? Y/N

If corrections were made, will the corrections affect other emissions data? Y/N

How was the impact of the erroneous data evaluated?

_____

_____

_____

I. Were the data validation procedures and activities adequately documented in the bound project notebook assigned to the persons evaluating the data? Y/N

If no, describe the problems found.

_____

_____

_____

IV. EMISSIONS DATABASE DEVELOPMENT

A. Were the data validated prior to being entered into the database? Y/N

B. Were the data presented to the entry personnel recorded in a manner that facilitated entry into the database? Y/N

C. Was all of the information required to be entered in the database included on the data form? Y/N

D. If data are missing from data request forms, how are data gaps handled?

_____

_____

E. Were results in the units to be reported? If not, were calculations performed manually or electronically? Y/N

_____

F. Were the database activities documented in the bound project notebooks? Y/N

Did the data recorded allow reconstruction of the activities? Y/N

Were pages in the notebook reviewed and signed by the senior technical reviewer? Y/N

G. Were data entries reviewed for transcription errors by someone other than the person entering the data into the database? Y/N

If problems were found, were the resolution of them documented and the revision of the data indicated in the electronic file? Y/N

H. Was the database developed so that revised versions of the database are identified? Y/N

I. Were the software and hardware evaluated to determine whether they are adequate to achieve the objectives of the computer database activities prior to using them? Y/N

What tests were performed and were the results from the tests documented? (response time, available memory, available power, accessibility for use)

_____

_____

_____

J. How often are files backed up?

_____

Is the schedule appropriate to minimize data loss? _____

K. Was a log maintained of database revisions? Y/N

L. Are the computer manuals available for use by the operators? Y/N

Does the manual include all of the data needed to log into the system and perform the duties required to develop the emissions database? Y/N

V. REPORTING

A. Was the report formatted as required by U.S. EPA? Y/N

B. Was the report clearly written and inclusive of the applicable emission source identified during the planning phase of the work? Y/N

If a source was missing, can the reason for the omission be verified to be acceptable? Y/N

C. Did the report accurately reflect the data included in the database? (Compare the results in the report to the information included in the database for 5-10 sources). Y/N

D. Was there evidence in the data file of editorial and technical review of the document? Y/N

E. Was a copy-ready version of the report included in the master data file? Y/N

VI. QUALITY CONTROL

A. Were the QC measures taken adequate to ensure data quality? Y/N

B. Were the project and quality goals met? Y/N

C. Were actions taken in response to all previous recommendations for corrective actions? Y/N

Did the actions taken adequately address the quality concerns found? Y/N

_____

_____

_____

VII. RECOMMENDATIONS FOR CORRECTIVE ACTIONS

VIII. COMMENTS

APPENDIX B

QUALITY CONTROL CHECKLIST

Auditor: _____ Date _____

Data/Procedure Reviewed: _____

_____

Inventory Development Personnel Involved in Work: _____

_____

Select a facility or source category with high emissions and evaluate the quality of the data and adequacy of the data handling procedures (access, organization, completeness, etc.). Record the findings and recommendations for corrective actions, if any, on the checklist and comment sheet provided.

If recommendations for corrective actions are made, discuss them with the Task Leader immediately following the audit. Conduct follow-up activities to determine if the actions taken in response to the recommendations appropriately resolved the quality issues identified.

I. DATA

A. Identify the source evaluated.

_____

B. Describe the data included in the master file for the facility or source category.

_____

_____

_____

C. Are the data documented in a manner that will not have the potential to be misinterpreted? Y/N

Were the instructions for documenting the data followed? Y/N

D. Are there missing data fields? Y/N

What procedures are taken by the Data Manager and Task Leaders to ascertain missing data?

_____

_____

_____

At what point in the inventory process are requests for missing data made?

_____

_____

_____

How is the receipt of the missing data handled? (Are original data sheets placed in the master file?)

_____

_____

_____

Is the procedure followed to ascertain missing data efficient and adequate? Y/N

E. Are emissions types given (e.g., actual, allowable, maximum design capacity)? Y/N

F. Are the procedures used to calculate emissions described in the data provided? Y/N

G. Are the emissions determined in a technically sound manner? Y/N

H. Are sufficient data provided to recalculate the emission results? Y/N

Verify the accuracy of the calculations of the emissions for some of the pollutants. (Attach calculation sheets to the checklist.)

If any of the values are incorrect, explain how the emissions data were corrected.

_____

I. How are unavailable data identified? Are they mentioned in the report?

_____

II. EMISSIONS DATABASE

A. Do the values reported on the data sheets reviewed agree with the entries in the database? Y/N

B. Who provided the data to the data entry personnel?

_____

C. Was there evidence that the data were reviewed for accuracy and completeness prior to submittal to the data entry personnel? Y/N

D. Were the data sheets complete when they were received? Y/N

E. Were copies or original data sheets submitted to the data entry personnel? Y/N

If original data sheets were used, do the data tracking records show the release of the original data to the data entry personnel? Y/N

F. Were the QAP and a user's manual accessible to the data entry personnel? Y/N

G. Were the personnel adequately trained to perform the duties assigned? Y/N

H. Were the procedures followed in agreement with those specified in the QAP? Y/N

I. Is the database routinely backed up at the end of each updating event? Y/N

J. Does the computer allow double entries for the same source? Y/N

K. Are default values understood and properly documented? Y/N

L. Are key data fields flagged when data are not entered or are not available? Y/N

M. Ask the data entry personnel to explain the QC procedures followed to ensure data quality.

Do they agree with the procedures described in the QAP? Y/N

N. Does the computer system appear to be adequate for its intended use? (Ask the data entry personnel about the problems they have experienced with the system.) Y/N

O. Is the data entry progressing as expected and are the procedures followed adequate to ensure data quality? Y/N

_____

_____

_____

III. RECOMMENDATIONS FOR CORRECTIVE ACTIONS

IV. COMMENTS

APPENDIX C

EMISSIONS INVENTORY DATA ELEMENT CHECKLIST

Name _____

Data Reviewed _____

Date _____

A permit will not be considered complete for use in the emissions inventory if the following data elements are not in the permit application. Look in the permit application for each data element and check the space next to the data element to determine if the information requested was provided.

(1) Facility name _____
(2) Facility address, including city _____
(3) Zip code _____
(4) County or city _____
(5) Business description or SIC Code _____

(6) Design capacity emissions (if applicable)
Explanation if not applicable:

_____

(7) Projected emissions (hourly and annually, if applicable) _____
Explanation if not applicable:

_____

(8) Allowable emissions (hourly and annually, if applicable) _____
Explanation if not applicable:

_____

In some cases, an emission type, such as design capacity emissions, may not be applicable for that process or source. If that is the case and no emissions are reported for that emission type, provide a brief explanation below.

_____

_____

_____

APPENDIX C (CONTINUED)

EMISSIONS INVENTORY DATABASE CHECKLIST

Name _____

Data Reviewed _____

Date _____

Data entered into the Air Pollutant database should be checked for missing data elements. Use this list to check representative facilities in the database. Look at the facility records for the following data elements, and check the space next to the data element name if the information was provided. In some cases, an emission type, such as design capacity emissions may not be applicable for that process or source. If that is the case and no emissions are reported for that emission type, provide a brief explanation on this form.

(1) Facility name _____
(2) Facility address, including city _____
(3) Zip code _____
(4) County or city _____
(5) County code _____
(6) Facility code _____
(7) Business description or SIC Code _____
(8) Design capacity emissions (if applicable) _____
Explanation if not applicable:

_____

_____

(10) Projected emissions (hourly and annually, if applicable) _____
Explanation if not applicable:

_____

_____

(11) Allowable emissions (hourly and annually, if applicable)
Explanation if not applicable: _____

_____

_____

☆U.S. GOVERNMENT PRINTING OFFICE: 1997 -529-018

TECHNICAL REPORT DATA

(PLEASE READ INSTRUCTIONS ON THE REVERSE BEFORE COMPLETING)

1. REPORT NO. EPA-454/R-97-004f	**2.**	**3. RECIPIENT'S ACCESSION NO.**
4. TITLE AND SUBTITLE Emission Inventory Improvement Program Quality Assurance Procedures Preferred And Alternative Methods		**5. REPORT DATE** **7/25/97** **6. PERFORMING ORGANIZATION CODE**
7. AUTHOR(S) Emission Inventory Improvement Program Quality Assurance Committee		**8. PERFORMING ORGANIZATION REPORT NO.**
9. PERFORMING ORGANIZATION NAME AND ADDRESS U S Environmental Protection Agency Office Of Air Quality Planning And Standards (MD-14) Research Triangle Park, NC 27711		**10. PROGRAM ELEMENT NO.** **11. CONTRACT/GRANT NO.** 68-D2-0160
12. SPONSORING AGENCY NAME AND ADDRESS Office Of Air Quality Planning And Standards, Office Of Air And Radiation, U S. Environmental Protection Agency Research Triangle Park, NC 27111		**13. TYPE OF REPORT AND PERIOD COVERED** Technical **14. SPONSORING AGENCY CODE** **EPA/200/04**

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The Emission Inventory Improvement Program (EIIP) was established in 1993 to promote the development and use of standard procedures for collecting, calculating, storing, reporting, and sharing air emissions data. The EIIP is designed to promote the development of emission inventories that have targeted quality objectives, are cost-effective, and contain reliable and accessible data for end users. To this end, the EIIP is developing inventory guidance and materials which will be available to states and local agencies, the regulated community, the public and the EPA.

Volume VI presents preferred and alternative methods for applying quality control and quality assurance procedures to estimating and compiling emissions data.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI FIELD/GROUP
Air Emisisons Air Pollution Emission Inventory Inventory Guidance	Air Pollution Control Emission Inventory Guidance	

18. DISTRIBUTION STATEMENT		**21. NO. OF PAGES** 449
		22. PRICE