VOLUME III: CHAPTER 24

CONDUCTING SURVEYS FOR AREA SOURCE INVENTORIES

Final Report

December 2000

Prepared by:
Eastern Research Group

Prepared for:
Area Sources Committee
Emission Inventory Improvement Program

DISCLAIMER

As the Environmental Protection Agency has indicated in Emission Inventory Improvement Program (EIIP) documents, the choice of methods to be used to estimate emissions depends on how the estimates will be used and the degree of accuracy required. Methods using site-specific data are preferred over other methods. These documents are non-binding guidance and not rules. EPA, the States, and others retain the discretion to employ or to require other approaches that meet the requirements of the applicable statutory or regulatory requirements in individual circumstances.

ACKNOWLEDGEMENT

This document was prepared by Eastern Research Group, Inc., for the Area Sources Committee of the Emission Inventory Improvement Program and for Charles O. Mann of the Air Pollution Prevention and Control Division, U.S. Environmental Protection Agency (EPA). Members of the Area Sources Committee contributing to the preparation of this document are:

Kristin Abraham, West Virginia Department of Environmental Protection
Kwame Agyei, Puget Sound Air Pollution Control Agency
Dan Brisko, New York State Department of Environmental Conservation
Orlando Cabrera-Rivera, Wisconsin Department of Natural Resources
Andy Delao, California Air Resources Board
Laurel Driver, Emission Factor and Inventory Group, U.S. Environmental Protection Agency
Mark Eastburn, Delaware Department of Natural Resources
Charles Mann, Air Pollution Prevention and Control Division, U.S. Environmental Protection Agency
Sally Otterson, Washington Department of Ecology
Kenneth Santlal, Massachusetts Department of Environmental Protection
Walter Simms, Maryland Department of the Environment
Jack Sipple, Delaware Department of Natural Resources and Environmental Control
Karla Smith-Hardison, Texas Natural Resources Conservation Commission
Angel Thompson, South Carolina Department of Health and Environmental Control
Lee Tooly, Emission Factor and Inventory Group, U.S. Environmental Protection Agency

CONTENTS

Section **Page**

CONTENTS (CONTINUED)

Section **Page**

CONTENTS (CONTINUED)

Section **Page**

FIGURES AND TABLES

Figure **Page**

Table **Page**

1

INTRODUCTION

1.1 WHAT IS A SURVEY?

A survey is a method of gathering information from a number of individuals (a "sample") in order to learn something about the larger population from which the sample has been drawn. Surveys can be conducted using different tools and may have variety of purposes, but all surveys have two characteristics in common:

- Information is collected from only a sample of the population; and
- Information is collected by means of standardized questions so that every individual surveyed responds to exactly the same question(s).

The steps involved in the survey process are presented in Figure 24.1-1. Inventory preparers often use survey questionnaires to gather point source emissions inventory data. Emission inventories for area sources are usually not compiled using the same methods as emission inventories for point sources. The level of effort required to collect data and estimate emissions from the large number of individual facilities or activities would be very high, especially with respect to the relatively low levels of pollutants emitted by each. To estimate emissions from area sources, the individual facilities or activities are grouped with like facilities or activities into broad source categories so that emissions can be collectively estimated using one methodology. A survey approach can be used to gather information needed to calculate area source emission estimates or used to develop region-specific emission factors for the development of emission estimates.

1.2 HOW CAN THIS CHAPTER HELP ME?

This chapter is intended to help state and local air pollution control agency personnel determine if a survey is needed as part of their area source emissions inventory development effort, and if so, will assist them in planning and implementing each of the steps in the survey process.

1.3 HOW IS THIS CHAPTER ORGANIZED?

This chapter consists of 10 sections and presents information you can use to:

- Determine if you need to conduct a survey (see Section 2);

Steps in the Surveying Process

1. Define the purpose. Be specific!
2. Determine if a survey is needed.
3. Define the population.
4. Develop the survey plan.
5. Develop survey questions.
6. Develop introduction and instructions.
7. Pretest instrument.
8. Edit and revise questionnaire.
9. Obtain approvals as required.
10. Distribute the survey.
11. Follow-up as required.
12. Quality control/data reduction.
13. Analyze the data.
14. Compile the results.

Figure 24.1-1. Steps in the Surveying Process

- Prepare a written plan for your survey effort (see Section 3);
- Choose the appropriate type of survey for your project (see Section 4);
- Select an appropriate sample for your survey (see Section 5);
- Identify the appropriate contacts (see Section 5);
- Design an effective survey document (see Section 6 and 7);
- Distribute and track the survey (see Section 8);
- Compile the data; and
- Conduct quality control/quality assurance (QA\QC) throughout the survey process.

1.4 WHERE CAN I GO FOR ADDITIONAL INFORMATION?

While this document attempts to compile the information necessary to plan and conduct a survey, you may wish to refer to more in-depth references for additional information on some aspect the of the survey process. Section 10 of this chapter presents complete citations for all of the references used to prepare this document. In addition, you can find valuable information in:

Babbie, E. (1990). *Survey Research Methods*. Belmont, CA: Wadsworth, Inc.

Braverman, M.T. & Slater, J.K. (Eds) (1996). *Advances In Survey Research. New Directions for Evaluation, 70*. San Francisco, CA: Jossey-Bass.

Dillman, D. (2000). *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Fink, A. (1998). *How to Design Surveys*. The Survey Kit No. 5. Thousand Oak, CA: Sage.

Fink, A. & Kosecoff, J. (1998). *How to Conduct Surveys: A Step-by-Step Guide*. Thousand Oaks, CA: Sage.

Fowler, F.J. (1993). *Survey Research Methods. Sage Applied Social Research Methods Series Volume 1*. Thousand Oaks, CA: Sage.

Kalton, G. (1983). *Introduction to Survey Sampling. Sage Series on Quantitative Applications in the Social Sciences*, Volume 35. Thousand Oaks, CA: Sage.

Litwin, M.S. (1995). *How to Measure Survey Reliability and Validity*. The Survey Kit No. 7. Thousand Oaks, CA: Sage.

Oppenheim, A.N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*. New York: Pinter Publishers.

Patton, M.L. (1998). *Questionnaire Research*. Los Angeles, CA: Pryczak Publishing.

Rea, L.M. & Parker, R.A. (1997). *Designing and Conducting Survey Research: A Comprehensive Guide*. San Francisco, CA: Jossey-Bass.

Smith, M.L. & Glass, G.V. (1987). *Research and Evaluation in Education and the Social Sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Sudman, S. & Bradburn, N.M. (1982). *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass.

2

HOW DO I DETERMINE IF A SURVEY IS NECESSARY?

In general, you should remember that any survey requires an extensive level of effort. The decision to undertake a survey should be made only after you have considered all other possibilities.

Surveys can be useful tools for collecting information from a large population—they can provide data to make calculations or information on which to base decisions. Although surveys can be expensive, difficult, or time-consuming, there are times when a survey is the most appropriate tool to use to gather data. To determine if a survey is required as part of the data collection effort for an area source emissions inventory for a specific category, you should:

- Assess your data needs;
- Evaluate the relative importance of the category to total emissions inventory;
- Evaluate your data quality objectives;
- Assess data availability;
- Evaluate resource availability;
- Consider the administrative clearances needed under federal or state rules to conduct a survey of the private sector; and
- Consider if it is possible to coordinate the information collection activity with other inventory efforts.

Each of these criteria is discussed in the following sections.

2.1 ASSESS YOUR DATA NEEDS

In order to determine if you must conduct a survey to collect the data needed for your area source emission inventory effort, you must first develop a specific definition of what data are required. The more specific you can make your definitions, the easier it will be to collect the

appropriate data. The objectives of a survey can usually be phrased in the form of a question, such as:

- Which control devices do autobody refinishers use?
- How much solvent do dry cleaning facilities purchase on an annual basis?
- How much wood do households burn per week in the winter?

Surveys may be used to collect qualitative information for the source category, but should be designed to produce quantitative results—that is, results that can be expressed numerically and can be used in rigorous data analysis.

For area source inventory purposes, keep in mind that the data collected from the sample will be scaled up for the entire inventory region. Even if not specifically stated in the inventory preparation plan, you will need to identify and collect reasonable surrogate data.

2.2 EVALUATE THE RELATIVE IMPORTANCE OF THE CATEGORY TO YOUR TOTAL EMISSIONS INVENTORY PROGRAM

You should use existing inventory data and your knowledge of federal, state, and local regulations to determine the importance of an individual emissions inventory project within the scheme of the overall responsibilities of your agency.

Questions to ask include:

- Does the source category emit a large percentage of the pollutant(s) of interest? Is the category a significant source of volatile organic compounds, hazardous air pollutants (HAPs) nitrogen oxides, sulfur dioxide, carbon monoxide, particulate matter, or ammonia?
- Is an accurate inventory of the pollutants of interest of particular importance to your agency (e.g., PM-2.5, ozone precursors, specific HAPs)?
- What is the end use of the inventory? An inventory with significant regulatory implications such as a residual risk study for HAP sources may require a survey component for data collection, while an inventory for source characterization may not.

Inventory efforts identified as high priority should be afforded sufficient resources to ensure that complete and accurate data are compiled. For many area sources, this data collection effort may require a survey.

2.3 EVALUATE THE DATA QUALITY OBJECTIVES FOR THE INVENTORY

The first step in planning any inventory is to define the purpose and intended end use of the inventory. This information is used to determine the data quality objectives (DQOs) and the quality control/quality assurance (QA/QC) requirements for the inventory. DQOs are qualitative and quantitative statements of the uncertainty that a decision maker is willing to accept in the estimates and/or decisions made with inventory data. For a more complete discussion of DQOs, refer to EIIP Volume VI, Chapter 2 (Documentation).

Preparation of a written DQO statement should be part of the initial planning stages of the inventory process. The DQO statement should address:

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

For inventory efforts with strict DQOs, it may be necessary to conduct a survey in order to ensure that appropriate data are collected.

You can use the Data Attribute Rating System (DARS) to evaluate the merits of one emission estimation method relative to another. DARS defines certain classifying attributes that are believed to influence the accuracy, appropriateness, and reliability of an emission factor or activity, and assigns a numerical score to each of these components that are combined to arrive at an overall confidence rating—an uncertainty estimate—for the inventory. You can develop DARS scores for several potential estimation methods and use this information when planning your inventory. For example, the DARS scores for two alternative methods to estimate emissions from the architectural surface coatings area source category are shown in Table 24.2-2. One method is based on a survey of paint distributors, the second uses a national per capita factor. The more resource-intensive survey method results in a much higher overall DARS score. You can use this information when considering questions such as:

- How much better can an inventory get if a survey is used compared to other methods?

TABLE 24.2-1

DARS SCORES (UNITLESS) 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 Specifity	1.0	1.0	1.00
Spatial	1.0	1.0	1.00
Temporal	1.0	1.0	1.00
Composite	0.925	0.975	0.908
National Per Capita Factor			
Measurement/Method	0.3	0.4	0.12
Source Specifity	1.0	0.3	0.30
Spatial	0.3	0.3	0.09
Temporal	0.7	1.0	0.70
Composite	0.575	0.500	0.30

- Does this amount of improvement justify the additional cost?

You will need to balance the inventory quality objectives and the available resources in order to formulate a workable strategy for your inventory. Refer to EIIP Volume VI (Chapter 4 and Appendix F) for detailed discussions of DARS.

2.4 ASSESS DATA AVAILABILITY

Before taking on any data collection effort, survey or non-survey, you should always check to see if the data have already been compiled. Each of the chapters in this volume suggest references of information for the respective area source categories. In general, you should also check for relevant data in:

- Permit files or compliance files;
- Case studies and site visit reports;
- U.S. Department of Commerce publications including *County Business Patterns, Census of Population, Census of Manufacturers, Census of Agriculture, County and City Data Book, Current Industrial Reports, Annual Housing Survey,* and *Census of Retail Trade;*
- U.S. Department of Energy publications such as *State Energy Data Reports*, *Natural Gas Annual*, and *Petroleum Marketing Annual;*
- State Departments of Transportation and State Energy Offices (for information on gasoline consumption and paving activities);
- State Departments of Labor (for employment data by SIC [Standard Industrial Classification] code);
- Local industrial directories (these are often organized by SIC code and provide employment data);
- Trade and professional association publications;
- Regional planning commission publications;
- Agency-sponsored surveys;
- National and state directories of manufacturers; and
- Data compiled by private research and development companies such as the *Directory of Chemical Producers* compiled by SRI International.

2.5 EVALUATE RESOURCE AVAILABILITY

As with any task, the resources required to conduct a survey will be determined by the scope of the project. You will need to evaluate your agency's ability to commit the appropriate resources—both personnel and money—to ensure that the survey is designed and conducted properly. Taking shortcuts can invalidate the results. The cost of a survey is a function of the completeness and specificity of the questionnaire, the size of the target audience, and the thoroughness of the QA/QC follow-up activities (Radian, 1996).

The survey process consists of a series of steps including planning, sample design, sample selection, contact identification, questionnaire preparation, pre-testing, mail-out, response tracking, follow-up activities, data collection, data reduction, data processing, data extrapolation, and QA/QC at each step. Each of these steps may require experience not readily available in an emissions inventory agency. You may need to factor in the cost of additional personnel or "learning curve" inefficiencies. Examples of resources required to conduct a survey include (Ferber, et al., 1994):

- Managerial staff time for planning the study and supervision through the various stages;
- Labor and material costs for design and pretest of the questionnaires;
- Computer hardware and software for data management;
- Telephone charges;
- Postage (surveys are often sent via registered mail);
- Reproduction and printing costs;
- Labor and materials cost for mail-out including compilation of up-to-date mailing lists, production of labels and cover letters, stuffing envelopes;
- Labor cost for programming e-mail or Internet surveys; and
- Labor costs for tracking responses, including data logging and follow-up with non-respondents.

You should recognize that allowing ample resources for quality checks at each step of the survey process is critical to a well-designed and well-conducted survey project. Be sure to include the cost of a rigorous QA/QC program when developing a project budget.

You must also consider the amount of time available to prepare the inventory. It may take several months to conduct a survey and analyze the data. If an emissions estimate must be prepared in 3 months and the survey process requires 6 months, you will need to use a non-survey method, even if a survey would result in a more accurate estimate of emissions.

2.6 CONSIDER THE ADMINISTRATIVE CLEARANCES

Before deciding to conduct a survey, you must identify any regulations that relate to the conduct a survey of the private sector. Specifically:

- Federal - Under the Paperwork Reduction Act, any Federal Government entity must obtain an approval from the Office of Management and Budget (OMB) to collect substantially similar information from ten or more respondents in any 12 month period. If EPA decides to collect information, it must prepare an Information Collection Request and submit it for OMB approval.

- Typically, the Director of the air quality regulatory agency has the authority to request information. However, you should identify the procedures that apply to your agency prior to planning and mailing the survey.

2.7 CONSIDER IF IT IS POSSIBLE TO COORDINATE THE INFORMATION COLLECTION ACTIVITY WITH OTHER INVENTORY EFFORTS

Because survey efforts are time-consuming and expensive, you should consider the possibility of coordinating your data collection efforts with other activities being conducted by your agency. Surveys are frequently conducted by permitting groups. If another survey is being planned that will include an appropriate population to collect the information that you require for your emissions inventory estimate, it may be most efficient to work cooperatively to collect the required information.

Two points must be carefully considered when making the decision to combine data collection efforts:

- You must make certain that the surveyed population represents a sample that is appropriate for all of the data collection efforts. Both the sample size and sample representativeness must be evaluated for each the data collection efforts.

- You must design the survey with extreme care. Unless response to the survey is mandatory, you will need to carefully balance the effort to collect information for more than one project with the need to keep the survey short and simple. If the questionnaire becomes too long or too confusing, recipients may not be willing to complete and return the form.

This page is intentionally left blank.

3

WHAT IS THE MOST IMPORTANT ASPECT OF CONDUCTING A SURVEY?

Simply put, ***the key to a successful survey effort is good planning.***

Careful and thorough planning of survey procedures will greatly facilitate the process and can prevent the need for costly revisions to the survey while in progress or emissions estimates that are generated from the survey data. Before the inventory process begins, your agency should prepare an inventory preparation plan to identify the required staffing levels and resource allocations. The inventory preparation plan will also specify the methods and procedures to be used by each member of the inventory team to collect, handle, review, and report emissions data. Refer to the *Handbook for Criteria Pollutant Inventory Development: A Beginner's Guide for Point and Area Sources* (EPA, 1999b) for additional information on inventory preparation plans.

While careful planning and survey design take time and may add front-end costs to the inventory effort, you should keep in mind that good planning actually reduces overall costs by preventing:

- Time wasted by the repeated need to make short-term decisions on what to do next;
- Duplication of data collection efforts;
- Time wasted collecting and analyzing irrelevant data; and
- Unplanned data analysis in the hope of finding relevant information in an incomplete data set (GAO, 1991).

As part of the inventory preparation plan, as a stand-alone document, or as a series of documents, you should **prepare a written plan that covers every phase of the survey process**. The written plan should clearly identify the goals, methods, and resources required for each step in the survey process. The written plan should:

- Define the data quality objectives;

- Define the data to be collected;
- Determine appropriate sample size;
- Identify the sample selection technique;
- Identify the survey technique;
- Outline techniques for design, pilot test, and revision of the survey;
- Identify the mail out and tracking techniques;
- Identify the data entry procedures;
- Identify the statistical methods to be used in the data analysis; and
- Identify the QA/QC procedures to be conducted during all phases of the survey process.

These survey procedures are discussed in the following sections.

4

WHAT SURVEY TECHNIQUES ARE AVAILABLE?

There are a variety of methods that can be used to collect survey data:

- Mail surveys;
- Electronic (E-mail or website) questionnaires;
- In-person interviews; and
- Telephone interviews.

The advantages and disadvantages of each method are described in the following sections. The information below was summarized from several documents (Fink and Kosecoff, 1998; Creative Research Systems, 2000; and Parker, 1999).

4.1 MAIL SURVEYS

Mail surveys, also referred to as "paper questionnaires", are documents mailed to the sample population that include a cover letter, instructions, and a form for the recipient to complete and return.

Advantages of the mail survey include:

- Mail surveys are one of the least expensive survey techniques. Unlike interview techniques, you will not need a team of trained interviewers;
- Mail survey forms can contain graphics. Use of diagrams, photographs, and tables is not possible with telephone interviews and may be limited with e-mail survey forms;
- Unlike interview techniques, mail surveys allow the survey recipients to respond when it is convenient for them. Mail surveys are considered less intrusive than telephone or personal interviews;

- This method gives survey recipients time to consider and research responses increasing the accuracy and completeness of the responses; and

- There is no possibility for interviewer bias to influence the respondents.

Disadvantages of the mail survey include:

- Mail surveys require more time than other survey techniques. You may need to wait for several weeks after mailing the questionnaires before you can assess your response rate. It may take several more weeks to follow-up with non-respondents;

- It may be difficult to obtain up-to-date and accurate mailing lists;

- Respondents may misinterpret or omit questions; and

- Mail surveys require motivated or interested respondents. Without legal requirements or incentives, respondents may choose to ignore a survey form.

4.2 E-MAIL SURVEYS

E-mail surveys are documents electronically distributed to the sample population. These messages include an introduction (similar to the cover letter of a paper survey), instructions, and a form for the recipient to complete and return via e-mail. More people have e-mail than have full Internet access, making e-mail a better choice than web page surveys for some populations.

Advantages of the e-mail survey include:

- E-mail surveys are one of the least expensive survey techniques. Because there are not fees for mailing or costs for interviewers time, it does not cost more to collect large samples;

- E-mail surveys can be fast. You can send the survey and receive responses in a period of days;

- The novelty element and convenience of an e-mail survey might stimulate higher levels of response than an ordinary paper survey;

- With many systems, you can attach picture and sound files; and

- Like paper surveys, e-mail surveys allow recipients to respond at their convenience and to carefully consider their responses.

Disadvantages of the e-mail survey include:

- It may be difficult or expensive to obtain up-to-date and accurate e-mail address lists;
- The sample population is limited to those with e-mail access. You will need to either acknowledge this when you identify the sample or be willing to distribute paper surveys to those who do not have e-mail access;
- It can be difficult to track e-mail responses. It is not unusual for recipients to forward the survey to multiple associates—or to respond more than once. It is important that the size and the representativeness of the sample that the responses be carefully monitored;
- Many people dislike unsolicited e-mail even more than unsolicited regular mail; and
- Respondents may misinterpret or omit questions.

4.3 WEB PAGE SURVEYS

Surveys can be posted on Internet web pages. The web pages include an introduction (similar to the cover letter of a paper survey), instructions, and a form for the recipient to complete and submit.

Advantages of web page surveys include:

- Web page surveys can be extremely fast and reach a large population;
- Like e-mail surveys, it does not cost more to collect large samples;
- These surveys can be designed to be interactive. The programs can be designed to provide respondents with explanations of terms or complex questions. Web page questionnaires can use complex question skipping logic and randomizations;
- The forms can be designed so that only "legal" answers are accepted, reducing the resources required for QA/QC and follow-up;

- Data in the submitted responses can be automatically entered into programs for analysis; and

- Web page questionnaires can use colors, fonts, and formatting options not available with an e-mail surveys.

Disadvantages of web page surveys include:

- The sample population is limited to those with Internet access. You will need to either acknowledge this when you identify the sample or be willing to distribute paper surveys to those who do not have Internet access;

- Depending on your software, you may have no control over who replies. This issue can be addressed by posting the survey on a page that can only be addressed directly (there are no links to it) or by restricting access by requiring a password; and

- Constructing an Internet web page may require a software engineer or programmer.

4.4 IN-PERSON AND TELEPHONE INTERVIEWS

In-person and telephone interviews are conducted by trained individuals who collect information from individuals using a written script and prepared data forms.

Advantages of in-person and telephone interviews include:

- The interviewer can explore answers given by respondents or provide additional information to ensure that the respondent understands each of the questions;

- The interviewer can prompt the respondent to prevent incomplete or inappropriate responses;

Disadvantages of in-person and telephone interviews include:

- Conducting interviews is resource intensive. You will need to find up-to-date phone numbers, schedule the interviews, ensure that all of the interviewers are well trained, and supervise dispersed personnel performing a complex task; and

- This type of survey has the greatest possibility that the interviewer can influence (bias) the results.

This page is intentionally left blank.

5

HOW DO I DETERMINE HOW MANY SURVEYS TO SEND - AND TO WHOM?

5.1 WHAT ARE THE ADVANTAGES OF SAMPLING FROM THE ENTIRE POPULATION?

A complete enumeration (or census) of the entire population may not be practical and is almost never economical. Collecting data from a properly identified sample of the population requires fewer resources than a census. You can use inferential statistics to determine a population's characteristics by directly observing only a portion (or sample) of the population.

Additionally, a sample may be more accurate than a census. A poorly designed and/or poorly conducted census can provide less reliable information than a carefully obtained sample.

Keep in mind that two things are absolutely necessary, however, to ensure a high level of confidence that the sample represents the population:

- A sufficiently large sample; and
- An unbiased sample.

5.2 HOW DO I IDENTIFY THE POPULATION?

As a starting point, you need to be familiar with the following terms (Fridah, 1998):

- Population - a group of individual persons, objects, or items from which samples are taken for measurement. For example, a "population" from an area source inventory might be all dry cleaners.
- Sample - a finite part of a statistical population whose properties are studied to gain information about the whole. For example, a "sample" for an area source inventory might be 10 percent of the dry cleaners in the non-attainment counties within a given state.

- Sampling - the act, process, or technique of selecting a suitable sample, or a representative part of a population for the purpose of determining parameters or characteristics of the whole population

The first step in the survey process is to define the relevant population (Ferber et al., 1994). This is particularly important for area source inventory efforts. Unlike surveys conducted for point source inventories, the best recipient may not be at the point of emissions. For example, an area source survey on solvent use could be sent to manufacturers, distributors, retailers, users—or a combination of these groups. You will need to identify the appropriate facilities/subcategories or point in production/use within the surveyed source category to serve as the population.

To identify the appropriate population, you should carefully evaluate:

- The data requirements for the survey effort; and

- The practical considerations of surveys of the different potential populations. In the solvent use example, it might be most practical to survey the manufacturers or the distributors.

In order to identify the relevant population, you will need to compile a list of the names, address, and general process category of each facility in the inventory area that could be included in the survey. This list could be based on information resources including:

- **Existing Inventories.** A recent or recently updated, well-documented, existing air emissions inventory is a good starting point. However, many existing inventories may focus on pollutants other than those needed in the inventory being prepared. Thus, certain sources that emit only one type of pollutant may not be well represented.

- **Other Inventories.** In addition to emissions inventories, other environmental inventories may be useful in identifying plants in various Standard Industrial Classification (SIC) Codes. Information in the Toxic Release Inventory System (TRIS), gathered annually under the "Community Right-to-Know" Law of the Superfund Amendments and Reauthorization Act (SARA Title III), and facility inventories developed under Title V may be useful. The TRIS database gives plant locations and SIC Codes, as well as quantitative information on emissions of specific toxic chemicals including many solvents. Title V permits may provide specific information about area source processes taking place at a facility.

In addition, listings of water pollution sources and hazardous waste generators are maintained by state water pollution and hazardous waste agencies. These may be used to identify potential sources in various SIC codes.

- **Air Pollution Control Agency Files.** Compliance, enforcement, permit application, or other air pollution control agency files may provide valuable information on the location and types of sources in the area of concern. These files can also be used later to cross-check certain information supplied on questionnaires.

- **Other Government Agency Files.** Files maintained by labor departments and tax departments frequently aid in the preparation of the mailing list. Such files will include various state industrial directories in which companies are listed alphabetically by SIC code and county. The information available in these files will vary from state to state. Thus, it is advisable to contact the appropriate personnel with these agencies to become familiar with which listings are available.

- **Other Local Information Sources.** The following local information sources can be consulted, where available:

  - Local industrial directories—A local industrial development authority may provide a current list of the sources that operate in the inventory area. These are often organized by SIC code and provide employment data.

  - *Yellow Pages*—The local telephone directory will have names, addresses, and telephone numbers of many industrial/commercial facilities that may be emissions sources. However, telephone directory areas often do not correspond to county or community boundaries.

  - Manufacturers and suppliers—Firms that make or supply equipment and materials such as solvents, storage tanks, gasoline pumps, incinerators, or emissions control equipment maybe used to identify industries emitting VOCs, HAPs, CO, and nitrogen oxides.

- **National Publications.** The national publications listed below can be used when available. However, the information in them may be older and less accurate than local primary references.

  - Dun & Bradstreet, *Million Dollar Directory and Middle Market Directory* compile lists of companies by SIC code and county (refer to http://www.dnb.com);

- Dun & Bradstreet, *Industrial Directory*;
- National Business Lists—Companies are listed by SIC Code and county with information on financial strength and number of employees; and
- Trade and professional society publications—Names and addresses of members are listed along with their type of business.

5.3 HOW DO I DETERMINE THE SAMPLE SIZE?

The sample size is dependent on:

- The data quality objectives established for the accuracy of the results of the survey project;
- Available resources;
- The nature of the analyses to be performed; and
- The degree of heterogeneity of the population.

Ideally, the sample size chosen for a survey should be based on how reliable the final estimates must be. In practice, a trade-off is usually made between the ideal sample size and the expected cost of the survey. You can get a general idea of the sample size needed for your survey project from the checklist in Table 24.5-1.

Risk, as it relates to sample size determination, is specified by two interrelated factors:

- The confidence level; and
- The precision (or reliability) range.

To minimize risk, you should have a high confidence (say 95 percent) that the true value you seek (the actual value in the population) lies somewhere within a small interval (say + or - 5 percent) around your sample value (your precision). The desired degree of precision and confidence level are established as part of the data quality objectives.

TABLE 24.5-1

FACTORS IMPACTING SAMPLE SIZE

	Large sample	Small Sample
The calculations or decisions to be made based on the survey data have important or costly consequences	✓	
The data quality objectives for level of confidence are high	✓	
The data quality objectives for level of confidence are low - only a rough estimate is required		✓
The population to be sampled is relatively heterogeneous (high level of variance)	✓	
The population to be sampled is relatively homogeneous (little variance)		✓
Project costs increase dramatically with sample size		✓
Project costs and time required vary only slightly with increases in sample size	✓	
Financial and staffing resources are limited		✓
Time allowed for project completion is limited		✓

Basic statistics books present the formulas that you can use to determine sample size.

On-line calculators are available to determine sample size, for example:

- http://www.au.af.mil/au/hq/selc/samplsiz.htm
- http://www.surveysystem.com/sscalc.htm

Also, you should adjust the computed sample size (n) by dividing by the expected response rate. For instance, if you determined that the appropriate sample size for your survey effort was 125 and you expect a 75 percent response rate, you should make your sample size equal 125/0.75 or 166.

If more sources are identified on the mailing list than can be realistically handled with available resources, your agency should review the mailing list to reduce the number of facilities to be sent questionnaires.

One way to reduce the size of the mailout is to develop an initial estimate of emissions by facility. If the number of employees in a company is known, then an estimate of the emissions potential can be made using available per-employee emission factors. This will provide a rough estimate of the emissions potential of each facility, which can then be used to select a sample of facilities that represent a range of emissions to receive the questionnaire. Any bias that this selection process introduces to the returned surveys should be considered when scaling up the survey results. Another way to reduce the size of the mailout is to contact the intended recipients of the survey by telephone before mailing the survey. These brief contacts with plant managers or other appropriate employees will indicate whether the pollutant-emitting process takes place at the facility (or if the facility is even operating), and reduce the number of surveys that are sent out unnecessarily.

5.4 HOW DO I SELECT THE SAMPLE?

Just as important as the size of the sample is the determination of the appropriate sampling method. Random sampling always produces the smallest possible sampling error—the size of the sampling error in a random sample is affected only by random chance. Because a random sample contains the least amount of sampling error, it is referred to as an unbiased sample. Note that this does not mean that the sample contains no error, but rather the minimum possible amount of error.

Sampling techniques range from simple random selection of the population units to highly complex samples involving multiple stages or levels of selection with stratification and/or clustering of the units into various groupings. Whether simple or complex, a properly designed sample always has two distinguishing characteristics:

- All the units in the target population have a known, nonzero chance of being included in the sample; and

- The sample design is described in sufficient detail to permit reasonably accurate calculation of sampling errors.

These features make it scientifically valid to draw inferences from the sample results about the entire population which the sample represents (Ferber et al., 1994).

The sampling method will be determined by the objectives of the survey effort. Table 24.5-2 lists some of the common sample selection techniques. The two most useful random sampling techniques are simple random and stratified random sampling methods.

Stratified random sampling is often conducted for area source survey projects (Radian, 1996). For example, most urban areas with a diversified economy contain numerous, small manufacturing facilities that may use solvents in coating, degreasing, or wipe cleaning operations. These facilities could include wood products manufacture and coating, plastics coating, miscellaneous metal parts manufacture and coating. The large number of operations and the differences in raw materials and production characteristics require that you develop a survey approach to accurately collect information that can be statistically extrapolated to the entire population of non-point source facilities. You could use a stratified random survey to solve this problem:

- The first stratum might be to use source category codes such as the Area and Mobile Source (AMS) Codes and the Source Classification Codes (SCC) to group facilities manufacturing similar materials.

- A second stratum might be to use number of employees to distinguish between larger and smaller facilities to account for different rates of material usage.

Stratified random sampling requires a detailed knowledge of the distribution of attributes or characteristics of interest within the population to determine the homogeneous groups that compose the population. A stratified random sample is superior to a simple random sample because the population is divided into smaller homogeneous groups before sampling, resulting in less variation the samples. This enables you to reach the desired degree of accuracy with a smaller sample size. But, if you cannot accurately identify the homogeneous groups, you are better off using the simple random sample because improper stratification can lead to serious error.

TABLE 24.5-2

TYPES OF SURVEY SAMPLING METHODS

Method	Description	Conditions When the Sampling Design is Useful
Haphazard sampling	"Any sampling location will do" Take samples at convenient locations or times. Can lead to biased estimates.	A very homogeneous population over time and space is essential if unbiased estimates of population parameters are needed. This method of selection is *not* recommended due to difficulty in verifying this assumption.
Judgement sampling	Subjective selection by an individual. Select samples that appear to be "representative" of average conditions. Can lead to biased estimates. Accuracy is difficult to measure.	The target population should be clearly defined, homogeneous, and completely assessable so that sample selection bias is not a problem. Conversely, specific samples are selected for their unique value and interest rather than for making interferences to a wider population.
Simple random sampling	Each population unit has an equal chance of being selected for measurement. Selection of one unit does not influence selection of other units. The best way to choose a simple random sample is to use a random number table or a computer-generated series of random numbers. As the first step, you assign each member of the population a unique number. The members of the population chosen for the sample will those whose numbers are identical to the ones on the random number list in succession until the desired sample size is reached.	The simplest random sampling design. Other designs below will frequently give more accurate estimates if the population contains trends or patterns of emission rates.
Stratified random sampling	Divide target population into non-overlapping parts. If one group is proportionally larger than another, its sample size should also be proportionally larger. As appropriate, different sampling techniques can be used in each of the different groups.	Useful when a heterogeneous population can be broken down into parts that are internally homogeneous. For example, solvent usage might be stratified according to the end product produced.

TABLE 24.5-2

(CONTINUED)

Method	Description	Conditions When the Sampling Design is Useful
Multi-stage sampling	Divide target population into primary units. Select a set of primary units using simple random sampling. Randomly subsample each of the selected primary units. Example: collect soil samples (primary units) at random, then select one or more aliquots at random from each subsample.	Needed when measurements are made on subsamples of the field sample. This technique has limited applicability to emissions inventory development.
Cluster sampling	Clusters of individual units chosen at random. All units in chosen clusters are measured.	Useful when population units cluster together (schools of fish, clumps of plants, etc.) and every unit in each randomly selected cluster can be measured. This technique has limited applicability to emission inventory development.
Systematic sampling	Samples are selected at intervals, locations, or times according to a predetermined spatial or temporal pattern. For example, assign each member of the population a unique number, choose a random number as a starting point and then survey every n^{th} member. This is a non-random sampling method! Regardless of how much you mix the population before selecting a starting point, the fact remains that once that point is chosen, further selection of members for the sample is non-random (no independence).	Usually the method of choice when estimating trends or patterns of emissions over space. Also useful for estimating the mean when trends and patterns are not present or they are known a priori or when strictly random methods are impractical.
Double sampling	If data using one measurement technique has a strong linear relationship to data obtained with less expense or effort using another measurement technique, more samples can be taken using the less expensive method. The linear relationship between the two techniques is then applied to estimate the mean for the more expensive method.	Useful when there is a strong linear relationship between the variable of interest and a less expensive or more easily measured variable.

TABLE 24.5-2

(CONTINUED)

Method	Description	Conditions When the Sampling Design is Useful
Search sampling	Used to geographically locate pollution sources or to find "hot spots" of elevated contamination.	Useful when historical information, site knowledge, or prior samples indicate where the object of the search may be found. This technique has limited applicability to emissions inventory development. This approach, however, could be used to develop information that would describe the spatial characteristics of emissions in relationship to a specific parameter (for example, lawn and garden equipment use versus household income).

Source: Gilbert, 1987

6

WHAT SHOULD I CONSIDER WHEN PREPARING THE QUESTIONNAIRE FORM?

You will need to develop a questionnaire specifically tailored for each source category you survey. While this may require significant resources, it will allow you to use industry-specific terminology and ask only relevant questions thus reducing confusion and increasing accuracy.

6.1 WHAT ARE THE MOST IMPORTANT ASPECTS OF PREPARING THE QUESTIONNAIRE FORM?

Many factors are critical to designing a survey form that will result in a high response rate and usable data. The most important points to keep in mind are:

- **Make certain that you ask the right questions;**
- **Design the questionnaire for the person who will be asked to fill it out; and**
- **Be as brief as possible.**

The goal is to design a survey that will provide you with accurate information that meets your data needs. Chapters 2, 3, 4, 13, 14, and 17 of this volume all contain example survey forms.

6.2 WHERE DO I START?

As the first step in designing a survey, you must set the survey boundaries so that you can write the correct questions. You need to have a specific definition of all of the information that the survey is being developed to collect. This should be included in the inventory preparation plan.

Construct a list of potential questions and use the following criteria to determine which questions should be included in the final survey:

- Does the question pertain to a stated survey goal? If a question is not necessary, do not include it.

- Place your questions into three groups: need to know, useful to know, and nice to know. Discard the last group.

For area source inventory purposes, keep in mind that the data collected from the sample will be scaled up for the entire inventory region. Even if not specifically stated in the inventory preparation plan, you will need to identify and collect reasonable surrogate data.

6.3 WHAT SHOULD I CONSIDER FOR THE SURVEY FORMAT?

6.3.1 KEEP THE SURVEY BRIEF

Keep the survey as **short and simple** as possible—this is very important to both the response rate and the accuracy of the responses. You should invest the resources to make the survey as understandable, simple, and quick as possible for the recipient. Remember, they are doing you a favor and you want to maintain a good agency-industry working relationship.

Carefully consider the physical size and format of the document. Again, keep the form as short as possible. If you need only a few specific data items, consider a pre-printed, postage-paid postcard.

While brevity is important, you should not design a survey that looks crowded or confusing. Techniques for wise use of space on a page include:

- Using columns in the page layout;
- Using different fonts and bold text for emphasis;
- Defining sections and emphasizing key items using lines, boxes, or shading; and
- Printing double-sided pages.

6.3.2 DESIGN THE QUESTIONNAIRE TO FIT THE MEDIUM

You need to design the questionnaire to fit the medium. Each survey type (paper, e-mail, web site, or interview) has advantages and limitations. For example, only some survey formats allow you to incorporate graphics. Review the descriptions of each survey type presented in Section 4 of this chapter. Use the survey medium to its fullest potential.

6.3.3 CONSIDER ALL OF THE SURVEY "USERS"

You want to make the survey attractive and easy for the recipient to complete. However, there are other people who will be working with the survey, and if you keep them in mind while

designing the forms, you can save time and money. Once the survey is completed and returned, it will go through several processes:

- Log-in;
- QA/QC to check for completeness and reasonableness of responses; and
- Data entry.

Design the survey so that each of these processes can be conducted quickly and accurately (Creative Research Systems, 2000):

- For questions requiring text answers, you should allow sufficient space for handwritten answers. Lines should be about one half inch apart.
- Try to keep the answers in a straight line, either horizontal or vertical. Studies show that the best place for answers is the right hand edge of the page.

6.3.4 MISCELLANEOUS TIPS

Other techniques that can improve response rate or the accuracy of the responses for your survey include:

- Leave a space at the end of the questionnaire entitled "Other Comments". A respondent might include a remark about an issue you had not even considered.
- Include the return address on each page of the survey. It is not unusual for return envelopes and forms to get separated.

6.4 WHAT IS MOST IMPORTANT FOR THE QUESTIONS?

You need to make sure that all of the questions are simple and well-worded. Keep in mind that the way you phrase a question can change the answers you get.

6.4.1 OPEN-END AND CLOSED-END QUESTIONS

An open-end question is one that the recipients answer in their own words. In many area source inventory surveys, you will be requesting that the recipient provide "fill in the blank" information. When you write open-end questions:

- Be specific;

- Clearly indicate the relevant time period for the requested data; and
- Clearly indicate the units of measure for the response.

A closed-end question lists the possible answers from which the respondents choose the response appropriate for their facilities. When you write closed-end questions:

- Provide references or lists so that recipients use appropriate chemical trade names, synonyms, Chemical Abstract Service (CAS) numbers, etc.;
- Make certain that you include all of the possible responses;
- Include a blank space for "Other" responses (just in case!); and
- Clearly indicate the relevant time period for the requested data.

6.4.2 "DON'T KNOW" OR "NOT APPLICABLE"

Include instructions for items that recipients do not want to, or cannot, answer. Allow a "Don't Know" or "Not Applicable" response to all questions, except to those in which you are certain that all respondents will have a clear answer.

While this may seem like an invitation for incomplete survey forms, it actually makes the QA/QC review of the returned surveys much simpler. Without "Don't Know" or "Not Applicable" as response choices, the respondent may simply skip the question. This requires either follow-up by the inventory staff or a decision by the QA/QC and data analysis groups on how to enter and analyze the data.

6.4.3 MISCELLANEOUS TIPS

Other techniques that can improve response rate or the accuracy of the responses for your survey include:

- Use industry-appropriate terminology.
- Define all acronyms the first time you use them.
- Each question should be self-explanatory or accompanied by clear directions.
- Make certain that each question addresses only one issue. Ask separate questions rather than try to collect multiple data points with one question.
- Clearly identify the time period (months, year) you need data for.

- Consider asking for the raw data. That way you'll know how values were calculated and will be able to perform QA/QC checks.

6.5 INSTRUCTIONS

Prepare a set of procedures and instructions to accompany the survey. The instructions should clearly and completely (yet briefly):

- Explain, in general terms, how to fill out the questionnaire;
- Provide specific directions for how to complete each type of question;
- Tell the recipient the date the survey should be returned to the agency; and
- Provide names and contact information for agency personnel who can provide technical assistance with survey issues.

6.6 PILOT TESTING

The purpose of pilot testing is to see how well your cover letter motivates your respondents and how clear your instructions, questions, and answers are. You should always have the survey reviewed by peer reviewers within your agency (experts) **and** a focus group from the appropriate industry. Pilot testing takes time and may be expensive, but it is best to identify areas for improvement before the survey is distributed. Once the survey has been sent out, changing the directions, questions, or potential responses is no longer an option.

After explaining the purpose of the pilot test, let the peer reviewers read and answer the questions without interruption. When they are through, ask them to critique the cover letter, instructions, and each of the questions and answers. Rigorous pilot testing:

- Reveals whether people understand the directions;
- Reveals whether people can answer each of the questions;
- Helps you confirm that you are using the appropriate industry-specific terminology;
- Tells you how long it takes to complete the survey;

Use the reviewers' recommendations and comments to finalize your survey document.

This page is intentionally left blank.

7

WHAT SHOULD I CONSIDER WHEN PREPARING THE COVER LETTER?

The cover letter is key to a successful survey effort. If the cover letter does not command attention, the attached questionnaire might be not be considered a priority and it may be filed away or discarded.

Start with an introduction or welcome message. In the case of mail questionnaires, this message can be in a cover letter or on the questionnaire form itself. If you are sending e-mails that ask people to take a web page survey, put your main introduction or welcome message in the e-mail. State who you are and why you want the information in the survey. A good introduction or welcome message will encourage people to complete your questionnaire.

7.1 WHO SHOULD SEND THE LETTER?

Whenever possible, survey forms should be sent by the state or local air pollution control agency. Most of the surveys conducted to collect information required to compile an area source inventory are voluntary—recipients are not legally required to respond. Recipients are more likely to respond to a survey from a state agency than to a request from a contractor.

If you send the survey cover letters on state or local air pollution control agency letterhead and have them signed by a government official, you may have a positive influence on the response rate.

7.2 WHO SHOULD RECEIVE THE LETTER?

There are two possible strategies for the distribution of survey forms: approaching the facilities directly, or dealing with trade associations and requesting that they collect the information from their members.

If you choose to send surveys directly to facilities, your effort will be more successful if you have specific contact names. If it is not feasible or economical to compile contact information, determine an appropriate job title and clearly mark the outer envelope (e.g., “Attention: Plant Manager”) to direct the survey to the proper supervisory personnel.

A trade association that supports the inventory effort can be a valuable ally in a data collection effort. These groups are used to working with government agencies and may have access to contact information that is not publicly available. Facilities may be more comfortable sharing information with a trade group than sending it to a government agency. However, keep in mind that if you adopt this strategy, you may lose control over the techniques used to choose the sample from the population. If the size and composition of the sample are critical to the survey study design, this strategy may not be appropriate.

7.3 WHAT SHOULD BE INCLUDED IN THE LETTER?

The cover letter should be as short and direct as possible. A strong statement about any existing and applicable regulations which require the recipient to respond is your most powerful tool for maximizing return rate. If response to your survey is not a legal requirement, you should include a statement explaining the potential benefit of the survey effort to individual facilities and the industry as a whole. Explain why the survey is important to your agency and your state.

Another important item to include in the cover letter is the response due date. To improve the return rate, you should present the final due date in the cover letter so that it will not be overlooked by the recipients who do not read the instructions. Provide a reasonable amount of time for recipients to complete the survey—but not so much that they set it aside to do it later and then forget it.

You should clearly state that all Confidential Business Information (CBI) will be handled appropriately. In addition, you should provide names and contact information for agency personnel who can provide technical assistance with survey issues.

8

WHAT IS REALLY INVOLVED IN THE MAIL OUT AND TRACKING STEPS?

The process of mailing and tracking the survey includes:

- Preparing the mailing list;
- Pre-screening;
- Assembling survey packets;
- Mailing the survey packets;
- Tracking responses; and
- Following-up with non-respondents and incomplete responses.

8.1 PREPARATION OF THE MAILING LIST

The first step in the survey distribution process is to compile a mailing list that tabulates the name and address of each facility to be surveyed. Assign each facility a unique identification code. Whenever possible, identify a specific individual at each facility. You should have collected most of this information while conducting research to identify the relevant population. Refer to Section 5.2 of this chapter for a list of potential data sources.

Invest the time and effort to ensure that the appropriate facility and contact information has been identified. Check and double check addresses—it will save you lots of time in the long run if you don't have to deal with lots of returned survey packages.

You can greatly increase the efficiency and accuracy of the survey process if you use a database or spreadsheet program to construct your mailing list. A well-designed database can be used to generate mailing forms, create identification labels for each form, track survey returns, and format data for analysis and reporting. This type of tool will enable you to enter facility-specific information only once, rather than repeating data entry at several steps of the survey process.

8.2 PRESCREENING

You can improve the response rate by contacting the survey recipients before you send out the survey forms. Prescreening can be used to:

- Collect information (does the process of interest occur at the facility?) to help limit the number of surveys that are sent to inappropriate facilities;
- Confirm mailing addresses and ensure that appropriate contacts have been identified; and
- Inform the recipients about the upcoming survey project and foster support for the effort.

8.3 PREPARATION AND MAILING

The survey package will include the envelope; cover letter; survey form; and a pre-addressed, postage-paid return envelope.

To expedite tracking and data processing procedures, print duplicate mailing labels and place one on the outer envelope and the second one on the survey form. This will ensure that the proper identification code is on the returned survey. Mailing labels should contain the following information:

- SIC code (if applicable);
- Unique identification number;
- Contact name (or appropriate job title if a specific contact has not been identified);
- Contact title;
- Facility name;
- Street address;
- City;
- County;
- State; and

- Zip code.

If you were not able to identify specific contact information for a facility, determine an appropriate job title and clearly mark the outer envelope (e.g., "Attention: Plant Manager") to direct the survey to the proper supervisory personnel.

Your agency will need to carefully consider whether to send the surveys via first class or registered mail. With registered mail, the sending agency is informed when a questionnaire is received. Registered mail is more expensive and labor-intensive, but it does positively impact the response rate—simply because the recipient knows that the agency knows that the survey was delivered. Whether to spend money on registered mail depends on the importance of high response rate of success of the survey project. As an alternative, you might consider using registered mail for the largest or "most important" sources.

8.4 TRACKING AND FOLLOW-UP

Responses can begin arriving within a few days after mailing. The majority of the early returns will be from companies that are not sources of the emissions being studied and questionnaires returned by the postal service as undeliverable.

The following records must be kept for every survey form that you send:

- Facility information—all of the information included on the mailing label;
- Date mailed;
- Date returned;
- Whether follow-up is required; and
- Status of follow-up effort.

Follow-up will be required if the survey is returned by the postal service; the facility receives the survey but does not respond; or if the response is inadequate.

For all surveys "Returned to Sender", you will need to conduct research to find up-to-date information and re-send as appropriate. Log in corrected addresses, contact information, date mailed, and date returned for all surveys re-sent.

Approximately 2 weeks before the survey response deadline, you should begin to contact the facilities that have not responded. If the number of non-respondents is small, you can collect the information through telephone contacts or plant visits. If a facility refuses to complete the

survey, your agency might take legal action (if appropriate) to force a response; or you will need to follow the data handling procedures for non-responses outlined in the QA/QC plan.

Returned surveys will be checked by the QA/QC staff to ensure that the information on each form is complete and reasonable. If the data are determined to be inadequate, your agency will need to re-contact the surveyed facility. Telephone calls or plant visits are the most efficient means to collect the complete or revised information. Direct contact will provide your agency with the opportunity to clarify any misunderstood questions and assist the facilities to complete the survey form.

9

I GOT RESPONSES - NOW, WHAT DO I DO WITH ALL OF THAT DATA?

As survey responses are received, they are logged in and put through an initial data check to determine if responses are complete and the data are reasonable. The QA/QC procedures and criteria for this check should be presented in the inventory preparation plan. As described in Section 4 of this chapter, you will need to follow-up with facilities that do not respond or that submitted inadequate data.

Following the initial QA/QC check and any required follow-up, your data entry team can input the information into the appropriate data files. Consider designing look-up tables or pull-down menus in the data entry programs to define and fill in permissible entries—this is particularly helpful when the data entry involves lots of chemical names.

Your QA/QC staff will then review the entered data as required by the QA/QC procedures and criteria to verify the accuracy of the data entry. You can also use automated computerized checks to:

- Ensure against entering inappropriate data in a field—for example characters where numbers are expected;
- Conduct range checks to confirm that values are within a specified minimum and maximum for a specific variable; and
- Highlight outliers or suspect data.

Refer to Volume VI of the EIIP series, *Quality Assurance Procedures*, for additional information on QA/QC of data entry and analysis.

You will need to address any outliers identified in the data set by using apply programming solutions, statistical techniques, or your knowledge of the sources and processes. These procedures should be defined in the inventory preparation plan.

A properly designed area source survey will collect the necessary data and include a mechanism for "scaling up" the survey results. By the nature of area sources, it may not be possible to survey the entire population of a source category. The method extrapolating from a sample to

the population will depend on the type of source that is covered by the survey effort. Refer to the source-specific chapters in this volume for additional information.

A suggested method on how to scale up your survey results can best be illustrated using an industrial source category example. You conduct a survey of facilities reporting under SIC code 2711, Newspapers. In the county of interest, a publicly available business database shows 150 facilities in this SCC code. You randomly select 25 facilities and mail surveys:

- 5 do not respond; and
- 20 respond.
  - 2 are listed under wrong SIC codes;
  - 14 are listed under correct SIC codes, and complete the survey; and
  - 4 are listed under correct SIC code, but do not complete the survey adequately.

So, out of the 150 facilities in this SCC code:

- You know nothing about 20% (the non-respondent facilities);
- You know something about the respondent 80%:
  - 8% report under the wrong SIC code;
  - 56% you can develop emission estimates for; and
  - 16% you can NOT develop emission estimates for.

Thus, your sample data can be extrapolated to 72% (108 facilities) of the total population reporting under that SIC code. Emission estimates can also be extrapolated to the unknown 20% by assuming that 8% of these facilities report under the wrong SIC code, and 92% (27 additional facilities) report under the correct SIC code.

10

REFERENCES

Creative Research Systems. 2000. *The Survey System Tutorial: Survey Design*. http://www.surveysystem.com/sdesign.htm

EIIP. 1997. Evaluating the Uncertainty of Emission Estimates. In: *EIIP Volume VI, Quality Assurance Procedures*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-454/R-97-004f. Research Triangle Park, North Carolina.

EPA. 1999a. *Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Haze Regulations*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-454/R-99-006. Research Triangle Park, North Carolina.

EPA. 1999b. *Handbook for Criteria Pollutant Inventory Development: A Beginner's Guide for Point and Area Sources*. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, EPA-454/R-99-037. Research Triangle Park, North Carolina.

Ferber, R., P. Sheatsley, A. Turner, and J. Waksberg. 1994. *What Is A Survey?* Subcommittee of the Section on Survey Research Methods. American Statistical Association Washington, D.C. http://asio.jde.aca.mmu.ac.uk/resdesgn/survey3.htm

Fink, A. and J. Kosecoff. 1998. *How To Conduct Surveys: A Step-By-Step Guide*. Sage Publications, Thousand Oaks, CA.

Fridah, M.W. 1998. *Sampling In Research*. Cornell University. http://trochim.human.cornell.edu/tutorial/Mugo/TUTORIAL.HTM

GAO, 1991. *Designing Evaluations*. U.S. General Accounting Office, Program Evaluation and Methodology Division, GAO/PEMD-10.1.4.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold Company, New York.

Parker, R. 1999. *Survey Research: Sampling and Design* (Part of a course in Planning Analysis). http://darkwing.uoregon.edu/~rgp/PPPM613/class11.htm

Radian International, LLC. 1996. Mexico Emissions Inventory Program Manuals. Volume III: *Basic Emission Estimating Techniques*. Prepared for the Western Governors' Association and the Binational Advisory Committee. Radian International, LLC., Sacremento, CA.

U.S. Air Force. 1996. *Air University Sampling and Surveying Handbook*. United States Air Force, Air University. Maxwell AFB, AL 36112-6335. http://www.au.af.mil/au/hq/selc/smplntro.htm