RCRA Waste Management

Planning, Implementation, and Assessment of Sampling Activities

William M. Cosgrove, Michael P. Neill, and Katharine H. Hastie, editors

ASTM

RCRA Waste Management:

Planning, Implementation, and Assessment of Sampling Activities

Prepared by
Committee D-34 on Waste Management

William M. Cosgrove, Michael P. Neill, and Katharine H. Hastie, Editors

ASTM Stock Number: MNL42

ASTM
100 Barr Harbor Drive
West Conshohocken, PA 19428–2959

Printed in the U.S.A

Library of Congress Cataloging-in-Publication Data

RCRA waste management : planning, implementation, and assessment of sampling activities / prepared by Committee D-34 on Waste Management ; William M. Cosgrove, Michael P. Neill, Katharine H. Hastie, editors.
p. cm.—(ASTM manual ; 42)
"ASTM stock number: MNL42."
Includes bibliographical references and index.
ISBN 0-8031-2085-0
1. Hazardous wastes—Analysis—Handbooks, manuals, etc. 2. Hazardous wastes—United States—Management—Handbooks, manuals, etc. I. Cosgrove, William M., 1956- II. Neill, Michael P., 1962- III. Hastie, Katharine H., 1973- IV. ASTM Committee D-34 on Waste Management. V. ASTM manual series ; MNL 42.

TD1032 .R37 2000
628.4'2—dc21 00-028895

NOTE: This manual does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this manual to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.

Printed in Baltimore, MD
May 2000

Foreword

THIS PUBLICATION, *RCRA Waste Management: Planning, Implementation, and Assessment of Sampling Activities*, was sponsored by Committee D-34 on Waste Management. The editors were William M. Cosgrove, Michael P. Neill, and Katharine H. Hastie. This is Manual 42 in ASTM's manual series.

Preface

THIS MANUAL, *RCRA Waste Management: Planning, Implementation, and Assessment of Sampling Activities*, was prepared by William M. Cosgrove, Michael P. Neill, and Katharine H. Hastie under the direction of ASTM's Committee D-34 on Waste Management. The purpose of the manual is to make available to practitioners a basic reference regarding the development of a sampling strategy to meet the objectives of projects associated with common RCRA waste management activities. It is intended to be a companion document to EPA's SW-846, the guidance manual for planning and conducting sampling activities under RCRA. The planning (data quality objectives), implementation (sampling and analysis), and assessment (data quality assessment) phases are discussed in this manual for a variety of waste management scenarios. This manual provides a summary of the step-by-step process for completing a sampling investigation associated with a data collection activity for waste identification purposes under RCRA. As a basis, many of the ASTM standards and guides developed by Committee D-34 are referenced as well as others from committees such as D-18 on Soil and Rock and D-19 on Water. Guidance documents from sources outside ASTM such as the U.S. Environmental Protection Agency (EPA) are also included where appropriate, as well as helpful textbooks and technical manuals. This manual uses a practical "waste pile" example to illustrate the planning, implementation, and assessment process. The authors encourage the readers to consult the references listed at the end of each chapter and appropriate experts in the areas of sample collection and handling, sample analysis, and statistical methods for data assessment.

Contents

Chapter 1—Introduction **1**
References 1

Chapter 2—Sampling for Waste Management Activities: Planning Phase **2**
Introduction 2
Data Quality Objectives (DQOs) 2
DQO Steps 3
Step 1—Stating the Problem 3
Step 2—Identifying Possible Decisions 3
Step 3—Identifying Inputs to Decisions 4
Step 4—Defining Boundaries 5
Step 5—Developing Decision Rules 5
Step 6—Specifying Limits on Decision Errors 6
Step 7—Optimizing Data Collection and Design 6
Sampling Designs 9
Authoritative Sampling Designs 9
Probabilistic (Statistical) Sampling Designs 14
Summary 15
References 15

Chapter 3—Sampling for Waste Management Activities: Implementation Phase **16**
Introduction 16
Data Collection 16
Project Preparations 16
Selection of Sampling Equipment 18
Field Activities 19
Sampling Waste Units 23
Post Sampling Activities 25
Field Documentation 28
Technical Assessments 32
References 32

Chapter 4—Sampling for Waste Management Activities: Data Assessment Phase **33**
Introduction 33
Overview of Data Quality Assessment 33
DQA and the Data Life Cycle 33
Overview of the Five Steps of the DQA Process 33

Step 1—Review the DQOs and the Sampling Design 33
Step 2—Prepare Data for Statistical Analysis 34
Step 3—Conduct Preliminary Analysis of the Data and Check Statistical Assumptions 34
Step 4—Select and Perform Statistical Tests 35
Step 5—Draw Conclusions and Report Results 36
Summary 36
References 37

Appendix A: Confidence Intervals and Hypothesis Tests 41

Appendix B: Q/A/G-4 Chapter 6: Specify Tolerable Limits on Decision Errors 45

Appendix C: Waste Pile Example 53
Introduction 55
Planning Phase 55
Implementation Phase 60
Assessment Phase 60
For Case 1—Authoritative Sampling Design 61
For Case 2A (Normal Data Distribution) 61
For Case 2B (Non-Normal Data Distribution) 64
For Case 3—Systematic Grid Without Compositing Sampling Design 65
For Case 4—Systematic Grid with Compositing Sampling Design 66
For Case 5—Stratified Random Sampling Design 67
References 67

Appendix D: Statistical Tables 69

Index 75

1

Introduction

EACH YEAR the EPA and the regulated community expend a significant amount of resources collecting waste management data for research, regulatory decision making, and regulatory compliance. While these investigations are required for accurate decision making and effective environmental protection, it is the goal of EPA and the regulated community to optimize these studies by eliminating unneeded, duplicative, or overly precise data [*1,2*]. At the same time, however, the data collected must be of sufficient quantity and quality to meet the objectives of the study.

There are numerous difficulties that can complicate efforts to meet this goal including: lack of definition of the data users objectives, inadequate identification of the decisions and alternate actions that may be taken based on the findings, lack of information on the sources of contamination, appropriate action levels or sampling/analytical approaches, undefined boundaries (spatial and temporal) including the types of media to be sampled, undefined scale of decision making, practical constraints to sample collection including equipment limitations, access to all areas of the target population, and extreme variability or heterogeneity associated with the media being sampled, undefined decision errors that are acceptable to the data users, inadequate optimization of the study design including resource limitations, lack of consideration of the study objectives, and insufficient incorporation of quality assurance into the sampling and analysis plan [*1–3*].

Specific difficulties associated with sampling a population can be classified into five general categories:

- population access problems making it difficult to sample all or portions of the population,
- sample collection difficulties due to physical properties of the population (for example, unwieldy large items or high viscosity),
- planning difficulties caused by insufficient knowledge regarding population size,
- heterogeneity of the contaminant of interest, or item size, or a combination thereof, and
- budget considerations that prevent implementation of a workable, but too costly, sampling design.

The most efficient way to accomplish the goal of optimizing waste management studies is to determine the type, quality, and quantity of data required to address the problem before the sampling study is initiated. In order to meet these requirements, EPA developed and refined the Data Quality Objectives (DQO) process, a systematic planning tool for determining the type, quantity, and quality of data that will be sufficient and appropriate for the data's intended use [*1*]. ASTM has also developed a standard guide for the DQO process [*2*]. Data generation efforts involve three phases: *planning* with DQO development and sampling design optimization [*2,3*], the *implementation* of sampling and analysis strategies, and the *assessment* of data quality [*4,5*]. This manual uses a RCRA waste identification case history to illustrate the development of a sampling design and subsequent data assessment. This manual does not provide comprehensive sampling procedures, but references are given for locating guidance and standards where sampling procedures are discussed in more detail. It is the responsibility of the user to ensure appropriate procedures are used.

REFERENCES

[*1*] U.S. EPA, "Guidance for the Data Quality Objectives Process," QA/G-4, EPA/600/R-96/055, Office of Research and Development, Washington, DC, September 1994.

[*2*] ASTM, "Standard Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives," D 5792-95, 1995.

[*3*] U.S. EPA, "Guidance on Implementation of the Data Quality Objectives Process for Superfund," OSWER Directive 9355.9-01, EPA 540/R-93/071, Washington, DC, August 1993.

[*4*] U.S. EPA, "Guidance for Data Quality Assessment—Practical Methods for Data Analysis," QA/G-9, EPA/600/R-96/084, Office or Research and Development, Washington, DC, 1998.

[*5*] ASTM, "Standard Guide for Data Assessment for Environmental Waste Management Activities," D 6233-98, 1998.

Sampling for Waste Management Activities: Planning Phase

2

INTRODUCTION

PERHAPS THE MOST IMPORTANT of the three phases to completing a study is the planning phase. Without careful consideration during the planning phase, the implementation and assessment phases may result in data that are not of sufficient quantity and quality to meet study objectives. To facilitate the planning phase, EPA developed the Data Quality Objectives (DQO) process [*1*]. ASTM has further refined the process and included additional examples of DQO applications related to waste management activities [*2*].

DATA QUALITY OBJECTIVES (DQOs)

The development of DQOs is the first of three phases of data generation activities (Fig. 2.1). The others are implementation of the sampling and analysis strategies and data quality assessment [*2*].

By using the DQO process to plan waste management data collection efforts, study planners can improve the effectiveness, efficiency, and defensibility of decisions in a resource effective manner [*1*]. DQOs are qualitative and quantitative statements that:

- clarify the study objective,
- define the most appropriate type of data to collect,
- determine the most appropriate conditions from which to collect the data, and
- specify tolerable limits on decision errors.

To determine the level of assurance necessary to support a decision, this iterative process must be used by decision makers, data collectors, and data users. Objectives may need to be re-evaluated and modified as information concerning the data collection activity is gained. This means that DQOs are the product of the DQO process and are subject to change as data are gathered and assessed (Fig. 2.2).

DQOs are actually statements generated as outputs from each step of the process, although all of the DQOs are considered together during the data collection design step. The impacts of a successful DQO process on the project are as follows: (1) consensus on the nature of the problem and the desired decision shared by all the decisionmakers, (2) data quality consistent with its intended use, (3) a resource efficient sampling and analysis design, (4) a planned approach to data collection and evaluation, (5) quantitative criteria for knowing when to stop sampling, and (6) known measure of risk of making an incorrect decision based on the data collected [*2*].

The DQO process is a logical sequence of seven steps that leads to decisions with a known level of uncertainty. It is a planning tool used to determine the type, quantity, and adequacy of data needed to reach a decision. It allows the users to collect proper, sufficient, and appropriate information for the intended decision. The output from each step of the process is stated in clear and simple terms and agreed upon by all affected parties. The overall output consists of clear and concise presentation of the DQO process and complete documentation of the logic involved in the development of decision rules and associated limits on decision errors. As a useful tool, the DQO process can be integrated into a typical decision tree or logic flow diagram that clearly indicates actions to be taken as the result of implementation of the decision rules. The seven steps of the DQO process are as follows:

(1) stating the problem,
(2) identifying decisions,
(3) identifying inputs to decisions,
(4) defining boundaries,
(5) developing decision rules,
(6) specifying limits on decision errors, and
(7) optimizing data collection design.

All outputs from steps one through six are assembled into an integrated package that describes the project objectives (the problem and desired decision rules). These objectives summarize the outputs from the first five steps and end with a statement of a decision rule with a specified level(s) of the decision error (Step 6). In the last step of the process, various approaches to a sampling and analysis plan for the project are developed that allow the decisionmakers to select a plan that balances resource allocation considerations (personnel, time, and capital) with the project's technical objectives. Taken together, the outputs from these seven steps comprise the DQO process. The relationship of the DQO process to the overall process was illustrated in Fig. 2.1. At any stage of the project or during the field implementation phase, it may be appropriate to revisit the DQO process, beginning with the first step based on new information.

As noted in QA/G-4, the DQO process:

- has both qualitative and quantitative aspects,
- is flexible and iterative,
- can be applied more or less intensively as needed and is useful for "small studies,"
- helps develop the "conceptual site model,"
- does not always result in a statistical design,
- helps the transition from authoritative designs to more complicated statistical designs, and
- promotes good planning.

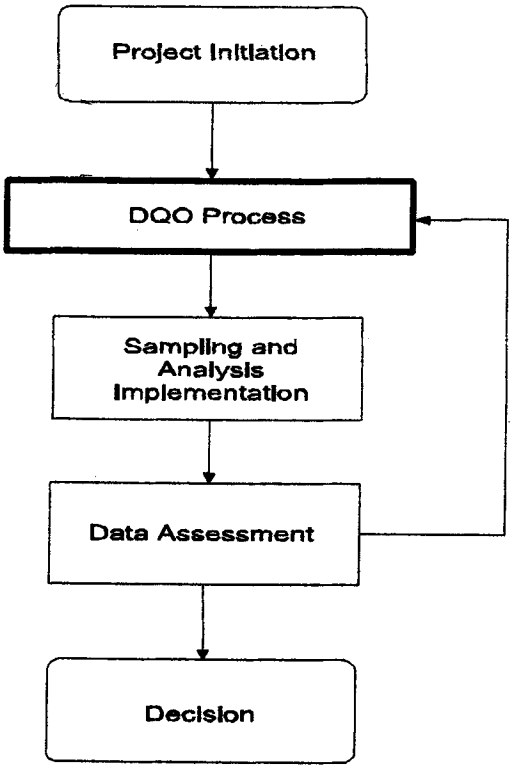


FIG. 2.1—DQO's process and overall decision process.

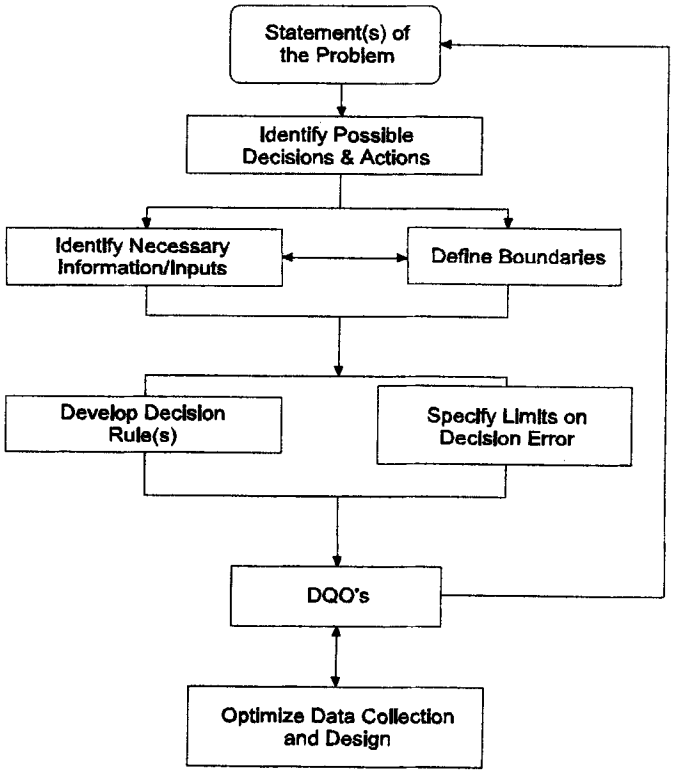


FIG. 2.2—DQO process.

DQO STEPS

The purpose of each of the seven DQO steps is discussed in the following section:

Step 1—Stating the Problem

The purpose of this step is to state the problem clearly and concisely. The first indication that a problem (or issue) exists is often articulated poorly from a technical perspective. A single event or observation is usually cited to substantiate that a problem exists. The identity and role of key decision-maker(s) and technical qualifications of the problem-solving team may not be provided with the first notice. Only after the appropriate information and problem-solving team are assembled can a clear statement of the problem be made [*2*].

The following elements of the problem description should be considered [*1*]:

- nature of the problem,
- study objectives/regulatory context,
- persons or organizations involved in the study,
- persons or organizations that have an interest in the study,
- political issues surrounding the study,
- sources and amounts of funding,
- previous study results, and
- existing sampling design constraints.

A brief description of the contamination problem that presents a threat or potential threat to human health and the environment may also be helpful during this step [*3*]. Included in this description would be the regulatory and program context of the problem, such as the regulatory basis for the field investigation, appropriate action levels for evaluating and responding to releases or exposures, and appropriate response actions. The development of a "conceptual site model" using existing data and information is needed to define affected media, contaminants, and receptors [*3*]. The conceptual site model is a non-mathematical model that provides an initial assessment of the contaminant sources, types, and concentrations of contaminants, migration/exposure pathways, and potential receptors. An initial review of resource issues, particularly those involving the budget and time constraints, should be completed during this step.

Step 2—Identifying Possible Decisions

The purpose of this step is to identify the decisions that will address the problem once it has been clearly stated. This step will help focus the efforts of the planning team towards a common objective. Multiple decisions are required when the problem is complex, and these may be arranged in the sequence in which they will be resolved with each decision being addressed separately from Step 2 through Step 7. Information required to make decisions and to define the domain or boundaries of the decision will be determined in later steps. Each potential decision is evaluated to ensure that it is worth pursuing further in the process. A series of one or more decisions will result in actions that resolve the problem. Figure 2.3 illustrates the activities that lead to identification of the decision [*2*].

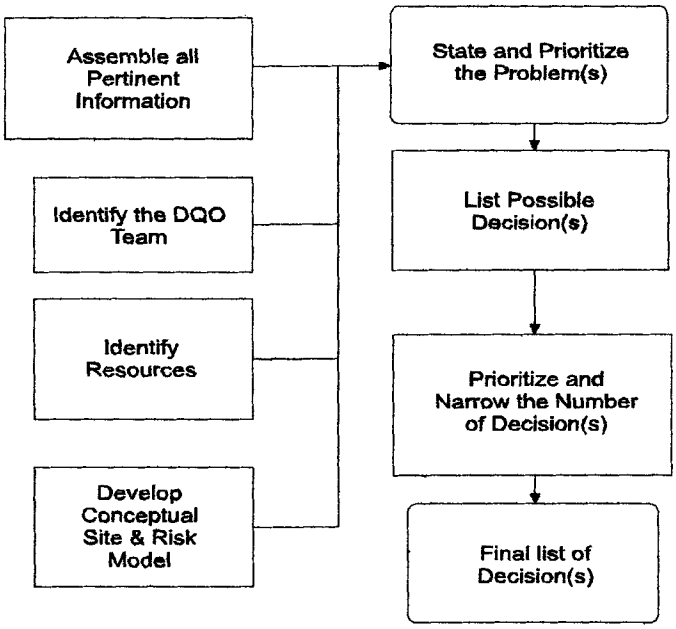


FIG. 2.3—Stating the problem and identifying the decisions.

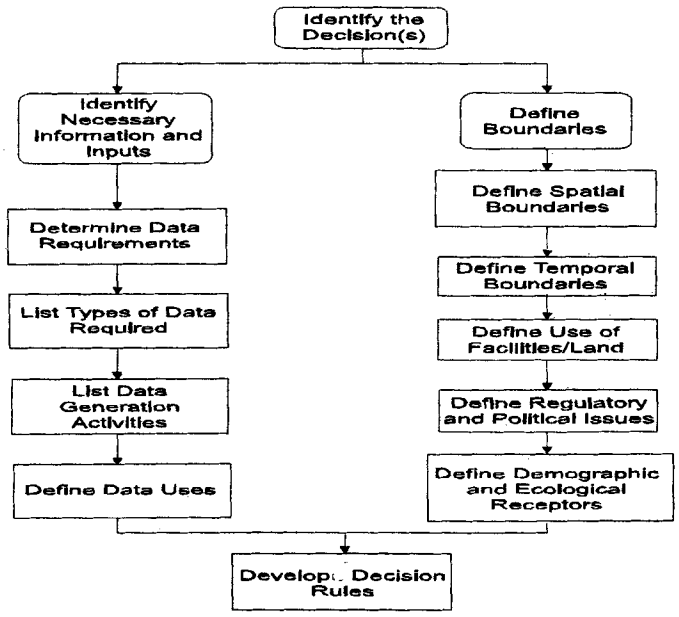


FIG. 2.4—Determination of information inputs and study boundaries.

A number of alternative decisions could be considered during this step including the determination of the following decisions [*3*]:

- Is a material hazardous by a characteristic?
- Does a material exceed a specific regulatory threshold?
- Has a release of contamination occurred from a process unit or waste management unit?
- Does a material exceed a risk-based number or remediation goal?
- What is the volume of contaminated material?
- Has a clean-up level been achieved?

At the conclusion of this step the planning team should be able to develop for each decision a clear decision statement that includes the principal study question and the alternative actions. An example would be: "Determine if the waste pile contains lead at a level (using the TCLP test) which will require management under the provisions of Subtitle C of RCRA."

Step 3—Identifying Inputs to Decisions

The answers to each of the questions identified by the previous step in the DQO process may be resolved through the collection of data via a sampling investigation [*2*]. The output of this step will be (1) a list of informational inputs needed to resolve the decision statement, and (2) a list of environmental variables or characteristics that will be measured [*1*]. Figure 2.4 shows the key activities that lead to development of the data requirements, as well as the study boundaries (Step 4). This sequence of activities must be performed for each question. Note that the limits of the study (or boundary conditions) are determined in a parallel step identified as "defining boundaries."

Activities during the input identification step are as follows (1):

1. Identify the informational inputs needed to resolve the decision. The information gathered during this phase would include:
   - historical waste generation and disposal practices,
   - hazardous substances associated with the site or process/waste management unit,
   - physical attributes of the waste management unit (size, accessability, shape),
   - known or anticipated variability in the distribution or nature of the contaminants, and
   - critical sampling locations that can be identified prior to sampling design consideration.
2. Identify sources for each information input and list those inputs that are obtained through previous data collection, historical records, regulatory guidance, professional judgment, scientific literature, or new data collection. Qualitatively determine if existing data are appropriate for the study (quantitative evaluation will occur in DQO Step 7: Optimizing Data Collection and Design).
3. Identify the information that is needed to establish the action level. The action level is the threshold value which provides the criterion for choosing between alternative actions. Action levels may be based on regulatory thresholds or standards, or they may be derived from problem-specific considerations such as risk analysis. In this step determine the criteria that will be used to set the numerical value.
4. Confirm that appropriate measurement methods exist to provide the necessary data, including the detection limit and limit of quantitation for each constituent of concern.
5. Identify potential sampling approaches and begin a preliminary evaluation of whether a non-probabilistic (authoritative) or probabilistic (statistical) sampling design is appropriate.

Step 4—Defining Boundaries

This step of the DQO process determines the boundaries to which the decisions will apply [2]. Boundaries establish limits on the data collection activities identified in Step 3. These boundaries include, but are not limited to, spatial boundaries (physical and geographical), temporal boundaries (time periods), demographic, regulatory, political, and budget boundaries.

Activities associated with this step include [1]:

1. Specify the characteristics that define the population of interest. It is important to clearly define the attributes that make up the population by stating them in a way that makes the focus of the study unambiguous. For instance, the population may be the sludge in a surface impoundment with the TCLP results for lead being the attribute. Note that typically RCRA waste identification decisions are made on samples collected at the point of generation rather than once the solid waste is located to a waste pile. However, in this case the material could have been identified as a Solid Waste Management Unit (SWMU) during the RCRA Facility Assessment process. Consequently the facility could be attempting to determine if the material exhibits a characteristic in addition to containing hazardous constituents.
2. Define the spatial boundary of the decision statement. This step has two components:

- Define the geographic area to which the decision statement applies. The geographic area is a region distinctively marked by some physical features (i.e., volume, length, width, boundary). This could be an exposure unit on a site, the limits of a waste pile, or soil to a depth of three inches.
- When appropriate, divide the population into strata that have relatively homogeneous characteristics. Using existing information, stratify or segregate the elements of the population into subsets or categories that exhibit relatively homogeneous properties or characteristics that may have an influence on the outcome of the study, such as contaminant concentrations or distributions. Dividing the population into strata will have a significant affect on the sampling design and is desirable for studying sub-populations, reducing variability within subsets of data, or reducing the complexity of the problem by breaking it into more manageable pieces.

3. Define the temporal boundary of the problem. This also has two components for consideration:

- Determine the time frame to which the decision applies. The planning team should decide when and over what period the data should reflect.
- Determine under what conditions the data should be collected. Conditions may vary over the course of the study, which may affect the success of data collection and the interpretation of results. Determine when conditions will be most favorable for collecting data and select the most appropriate time period to collect data that reflect those conditions.

4. Define the scale of decision making: which is the smallest area, volume, or time frame of the media in which the planning team will make a decision? The size of the scale of decision is usually based on either (1) risk (exposure unit), (2) technological considerations (area or volume that can be physically removed, treated, or disposed), or (3) other considerations such as the presence of hot spots of unknown size and location. Under RCRA the scale of decision making could be defined operationally, for example the decision could be made on each individual roll-off, drum, or other container prior to manifesting the waste.
5. Identify any practical constraints on data collection such as inability to gain physical access to the population under consideration, equipment limitations, matrix interferences (large particle sizes, extremely heterogeneous material, difficult to handle material), or seasonal/meteorological conditions.

Step 5—Developing Decision Rules

The purpose of this step is to integrate outputs from previous steps into a set of statements that describe the logical basis for choosing among alternative outcomes/results/actions. These statements are decision rules that define the following: (1) how the sample data will be compared to a regulatory threshold or action level, (2) which decisions will be made as a result of that comparison, and (3) what subsequent action(s) will be taken based on the decisions. The format for these rules is either an "if (criterion) . . ., then (action)" statement, or a decision tree.

The decision rule will include four main elements:

- The parameter of interest, which is a descriptive measure (such as a mean, median, or proportion) that specifies the characteristic or attribute that the decisionmaker would like to know about the population. The purpose of the data collection design is to produce environmental data that can be used to develop a reasonable estimate of the population parameter.
- The scale of decisionmaking that was defined in Step 4: Defining Boundaries.
- The action level, a measurement threshold value of the parameter of interest that provides the criterion for choosing among alternative actions. The action level can be based on regulatory standards, an exposure assessment, technology based limits, or reference-based standards.
- The alternative actions that the decisionmaker would take, depending on the true value of the parameter of interest (these were identified in Step 2: Identifying Possible Decisions).

Specific activities for this step include [1]:

1. Specify the statistical parameter of interest such as mean, median, or percentile. For instance, the decisionmaker may want to determine if the contamination level in a waste pile exceeds the regulatory threshold (i.e., the TC Rule regulatory level for lead of 5.0 mg/L) by using the mean of the data set, or by using an upper percentile. The statistical parameter may be dictated by a regulation and therefore not subject to change by the decisionmakers. Information about the positive and negative attributes of the alternate statistical parameters is available in EPA guidance manuals [1,3].
2. Specify the action level for the study that will direct the decisionmakers to choose between alternative actions. For instance, the decisionmakers may choose one alternative action if the TCLP result for material in the waste pile exceeds 5.0 mg/L for lead (i.e., managed under Subtitle C),

whereas a result under 5.0 mg/L may lead to a different action (i.e., managed under Subtitle D).

3. Formulate the decision rule. The output of this step in the DQO process is a decision rule using an "if . . . . . ., then . . . ." format that incorporates the parameter of interest, scale of decision making, action level, and the action(s) that would result from the decision. For example, "If the mean TCLP result for lead from the waste pile exceeds 5.0 mg/L, then the material is hazardous and must be managed under Subtitle C of RCRA; otherwise the material will be managed under Subtitle D."

Note that a "two-step" decision rule may be applied in certain situations, for example, to determine whether soil in an area exceeds an action level for a contaminant of concern, but where the decisionmaker also wants to prevent a hot spot from being left on the site without being removed. Let's say the site is four acres in size and the sampling design has a composite sample being collected in each quadrant of each acre (total of 16 samples). In this case the first step of the decision rule could be "If the 90% (one-tailed) upper confidence level for the mean concentration of lead is equal to or exceeds 400 mg/kg, then the soil will be removed and disposed." The scale of decision making in this case is the entire four-acre site. However, a second step to the decision rule could be added by saying, "If any one composite sample exceeds two times the action level (i.e., 800 mg/kg), then the soil in that quadrant will be removed and disposed." This approach allows for an overall decision to be made on the entire four acres, while allowing for the removal of a "hot" quadrant on any of the four acres.

Step 6—Specifying Limits on Decision Errors

An essential part of the DQO process is to establish the degree of uncertainty (decision error) that decisionmakers are prepared to accept in making a decision concerning the problem. The purpose of this step is to define the acceptable decision error rates (probabilities) based on a consideration of the consequences of making the incorrect decision. It is possible that the regulatory framework under which the data collection activity is being conducted will determine the decision error rate (i.e., the toxicity characteristic (TC) rule—40 CFR 261.24). In this case a relatively simple "confidence interval" method for decisionmaking may be used rather than a more complicated hypothesis testing method. This manual and the accompanying example discuss a "confidence interval" method for decisionmaking rather than formal hypothesis testing [*3*]. However, the reader is encouraged to consider the advantages of each method as they are addressed in Appendix A. A complete discussion of the use of formal hypothesis testing for Step 6 is included in Appendix B (an excerpt from QA/G-4).

The goal of the planning team is to develop a data collection design that reduces the chance of making a decision error to a tolerable level. There are two reasons why the decisionmaker cannot know the true value of a population parameter:

- *Sampling error*—due to the natural variability associated with a population over space and time. This error occurs because it is usually impossible to measure all portions of the population of interest.
- *Measurement error*—due to a combination of random and systematic errors that arise during the sampling and analysis (implementation) step. Examples include sample collection, sample handling, sample preparation, sample analysis, data reduction, and data handling. These potential error sources may be minimized through the use of a comprehensive Quality Assurance Project Plan (QAPP).

In order to evaluate the decision error associated with the data collection activity, an initial assumption or "null hypothesis" must be selected. For the TC Rule example in Appendix C, the null hypothesis is that the material in the waste pile is hazardous. For this null hypothesis the data collection activity may lead the decisionmaker to under-estimate the concentration of lead in the waste pile, thereby concluding that the material is not hazardous when it actually should be managed under Subtitle C of RCRA. This is a Type I or "false positive" error because it makes the "alternate hypothesis" (the material in the waste pile is not hazardous) true when in fact it is not. In making a hazardous waste determination under the TC Rule you set the Type I error rate (denoted by α) equal to 0.10. In doing so, you have specified a 10% chance of making a Type I error (note that 0.10 is a Type I error rate historically used for TC Rule applications). As a general rule, the lower you set the probability of making an error, a greater number of samples is required.

On the other hand, the decisionmaker may over-estimate the concentration of lead when the material is actually under the regulatory level and therefore should not be considered hazardous. This is called a Type II or "false negative" error. It is important to note that the confidence interval method for decision making included in this manual sets the Type II error rate (denoted by β) at a default of 50% or 0.50. The confidence interval method does not fully consider the implications of a Type II error on the data collection activity when compared to the formal hypothesis testing method.

Although a full treatment of the advantages and disadvantages of each statistical method is beyond the scope of this manual, we have included in Appendix B an excerpt from EPA's QA/G-4 DQO guidance manual that provides a complete discussion of the hypothesis testing method. The Appendix includes a discussion of a graphical approach (Decision Performance Goal Diagram) developed by EPA to evaluate the decision errors associated with a data collection activity. EPA has also developed a computer program (DEFT) for developing the diagrams that is based on the hypothesis testing approach [*4*]. DEFT assumes that the estimated mean is normally distributed and that a one sample t-test is the selected statistical test for comparing the result with a fixed standard.

Step 7—Optimizing Data Collection and Design

Prior to beginning this step of the process, the output from the first six steps must be assembled and provided to DQO team members who will optimize the sampling design for data collection. Care must be taken to separate the factual material from the DQO team's assumptions or estimates of factors important to development of the output from each step. The data collection effort must gather sufficient data to confirm (if possible/feasible) the accuracy of these assumptions.

The objective of this step is to generate the most resource-effective sampling design that will provide adequate data for decisions to be made. In this step, sampling designs are developed based on the outputs of the first six steps of the process, assumptions made during those steps, and applicable statistical techniques. The reader is encouraged to consult several excellent references that explain the advantages of alternate sampling designs [*1,7,8*]. A discussion of alternate sampling designs is included in a subsequent section of Chapter 2 on Sampling Designs.

An understanding of the sources of variability and levels of uncertainty is essential in developing the sampling design alternatives. The focus of the DQO process is the balancing of the limits of decision errors against the resources available to complete the project. Many of the sampling design alternatives will address different strategies for balancing the acceptable level of decision error with the resources available (time, money, and personnel) to resolve the problem. If a resource-effective sampling design to provide adequate data for the decision rule cannot be found among the sampling design alternatives, it may be necessary to alter the decision or revise the inputs into the DQO process. The steps for optimizing the sampling design is presented in Fig. 2.5. Activities associated with this step include [*1*].

1. Review DQO inputs and existing environmental data to determine the number of samples to be collected, the location of the samples, and the time of sample collection (if appropriate). A list of logistical concerns (equipment, access, personnel, resource constraints, etc.) should be assembled at this step.
2. Develop general sampling and analysis design alternatives. Although a complete discussion of the merits of alternate sampling designs, both probabilistic and authoritative, is beyond the scope of this manual, a brief overview is included later in this chapter. Examples of general data collection design alternatives include: authoritative (non-probabilistic) and several probabilistic designs: simple random, stratified random, sequential random, and systematic sampling. Several excellent references on the optimization of a sampling design are available from ASTM [*5*], EPA [*6*], and the private sector [*7*].
3. Define the sampling and analysis methods, including which SOPs may be used.
4. Select the optimal sample size that satisfies the DQOs for each alternative design. The planning team should evaluate each alternative design to determine how it performs when the assumptions are changed (i.e., increased variability over what was anticipated). To calculate the appropriate number samples, it is necessary to assemble existing data identified in DQO Step 3 ("Identify Inputs to the Decision") and Step 6 ("Specify Limits on Decision Errors"). If the population parameter of interest is the mean and a normal distribution is assumed, you can calculate the number of samples required using equations presented in the following sections and the example. Alternative equations can be found in the statistical literature and EPA guidance [*1,7,8,9*].
5. For each design alternative, verify that the DQOs are satisfied, including limits on decision errors, budget, schedule, and practical constraints (experience level of personnel, equipment limitations, site access, health and safety concerns, scheduling). If none of the designs satisfy the DQOs, the planning team may need to increase the acceptable decision error rates, relax other project constraints, such as time requirements or personnel limits, increase funding for sampling and analyses, or change the boundaries (spatial, temporal scale of decisionmaking).
6. Select the most resource effective design that satisfies all the DQOs.
7. Document the operational details of the selected design in the Quality Assurance Project Plan (QAPP). This will insure that the study is conducted as efficiently and effectively as possible [*6*]. Following completion of the planning step, the DQOs and sampling design are used to develop the Quality Assurance Project Plan [*6*]. The QAPP should clearly provide a link between the project objectives and how they will be met through the execution of the data collection activity. The QAPP will discuss the project objectives, project management (who is responsible for developing project documents, coordinating the field and laboratory support, and reviewing/assessing the final data), sampling requirements (locations, equipment, sampling procedures, preservation, shipping), analytical requirements (procedures, analyte lists, detection limits, regulatory requirements, and required precision and bias), quality assurance and quality control requirements (field and laboratory), and project documentation.

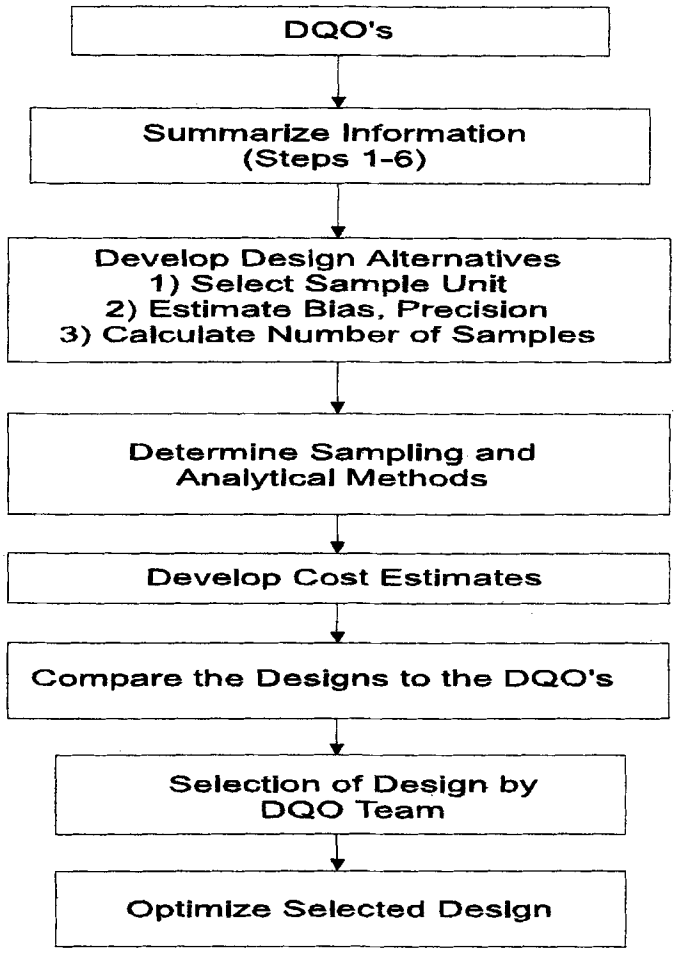


FIG. 2.5—Development of sampling design alternatives.

Design elements that must be documented include:

- sample types (composite versus grab),
- general collection techniques (equipment used),
- amount of sample to be collected,
- size of the aliquot from the sample that will be measured,
- sample locations and how they were selected (i.e., the sampling design),
- timing issues for sample collection, handling and analyses,
- analytical methods, and
- quality assurance and quality control needs.

Estimating the Required Sample Size

The sample size equations presented here should yield the approximate minimum number of samples required to achieve the DQOs for the assumptions mentioned earlier (mean is of interest, normal distribution, default Type II error rate of 0.5, etc.). However, it is prudent to collect a somewhat greater number of samples than indicated by the equations to protect against poor preliminary estimates of the mean and standard deviation that could result in an underestimate of the appropriate number of samples. It is important to note that the sample size equations do not account for the number or type of control samples (or quality assessment samples) required to support the QC program.

A key assumption for use of the sample size equations is that you have some prior estimates of parameters, such as the sample mean ($\bar{x}$) and sample standard deviation (s). To resolve this question, you may conduct a pilot study, use "real time" field analytical techniques (XRF, immunoassay kits, etc.) to evaluate variability, apply process knowledge and conduct a materials balance study, or use data from a study of a similar site or waste stream. If none of the above options can provide a suitable estimate of the standard deviation (s), a crude approximation of s still can be obtained. The approximation is based on the judgment of a person knowledgeable of the waste and their estimate of the range within which constituent concentrations are likely to fall. Given a range of constituent concentrations in a waste, but lacking the individual data points, an approximate value for s may be computed by dividing the range (the estimated maximum concentration minus the minimum concentration) by 6. Note that this estimate assumes that the data are normally distributed.

Post-Study Assessment of the Number of Samples Collected

Upon completion of the sampling effort, the data obtained is reviewed (see Chapter 4 on Data Quality Assessment). It can then be determined if an adequate number of samples were collected with respect to the relative error and confidence interval selected during the planning process. This determination is completed by calculating the appropriate sample size using the actual standard deviation obtained during the study. If this second value for "n" is less than or equal to the number of samples collected during the study, then the site has been characterized with the desired confidence level and margin of error. If the second value for "n" is significantly greater, then additional sampling is necessary, or an adjustment to the margin of error or confidence level may be considered. If the collection of additional samples is deemed necessary by the investigation team, the data that have been generated may be used to plan for a more efficient and cost-effective re-sampling of the site. Areas of the site where higher than anticipated variabilities were obtained may be segregated from areas of lower variability (stratified design).

Information pertaining to the estimate of sample numbers for alternative designs is included in the following sections:

Simple Random Sampling Designs

In order to estimate the number of samples required for a simple random sampling design, one approach requires that you determine the absolute margin of error (Δ) and an acceptable probability for the occurrence of decision error (α). Using this information, along with an estimate of the standard deviation, you may calculate the appropriate number of samples (n) for simple random sampling using the following equation [*4,8*]:

$$n = \frac{(t_{1-\alpha} + t_{1-\beta})^2 s^2}{\Delta^2}$$

where:

$t_{1-\alpha}$ = percentile value for the Student's t distribution for $n - 1$ degrees of freedom, where α is the probability of making a Type I error (the significance level of the test set in DQO Step 6).

$t_{1-\beta}$ = percentile value for the Student's t distribution for $n - 1$ degrees of freedom; where β is the probability of making a Type II error. Note that in the Appendix C example the Type II error rate is set at 0.50, the associated t value becomes zero, and the term drops from the equation.

s = an estimate of the standard deviation, and

Δ = the absolute "margin of error" defined as: $\Delta = RT - \bar{x}$.

An example application of the sample size equation is presented in the waste pile example (Appendix C). Note that an iterative procedure is required to obtain a final value of n.

Systematic Sampling Designs

One approach to calculating the appropriate number of samples (n) for systematic sampling designs is to use the same equation used for the simple random example, with the understanding that the sample locations will be arranged systematically with a "random" starting point. Such an approach should provide reasonable results as long as there are no strong cyclical patterns, periodicities, or significant spatial correlations between pairs of sample locations. If such features are present or suspected to be present, consultation with a professional statistician is recommended. As with all the sampling designs described in this section, you should have a preliminary estimate of the sample variance before using the sample size equation.

Stratified Sampling Designs

In general, there are two approaches for determining the number of samples to take when stratified random sampling is used to estimate the true mean for all strata combined: *optimal allocation* and *proportional allocation.* In optimal allocation, the number of samples assigned to a stratum (n_h) is proportional to the relative variability within each stratum and the relative cost of obtaining samples from each stratum. The number of samples can be determined to minimize the

variance of the estimated mean for a fixed cost, or to minimize the cost for a prespecified variance of the estimated mean. Optimal allocation requires considerable advance knowledge about the relative variability within each stratum, the relative size of the strata, and the costs associated with obtaining samples from each stratum. For this reason proportional allocation is recommended. In proportional allocation, the number of samples assigned to a stratum (n_h) is proportional to the stratum size.

Composite Sampling

Composite sampling is a tool that can be used with any of the authoritative or probabilistic sampling designs to increase the efficiency of the design when an estimate of average conditions is needed. The appropriate number of composite samples to be collected can be estimated by the equation used for simple random sampling. The sample variance with compositing is equal to the variance without compositing divided by the number of aliquots (k), aliquots being defined as the number of grab samples used to form each sample [*1,7*]. This assumes that the analytical variability is small relative to the sampling uncertainty. In comparison to non-composite sampling, composite sampling may have the effect of reducing between-sample variation, thereby reducing somewhat the total number of samples that must be submitted for analysis. Any preliminary or pilot study conducted to estimate the appropriate number of composite samples should be generated using the same compositing scheme planned for the confirmatory study. See Appendix C for an example of composite sampling.

Table 2.1 is designed to illustrate the general relationship between the margin of error and standard deviation versus the required sample size using the formula for a simple random design. The number of samples required at a 90% confidence interval (one tailed) with varying margin of error (Δ), and standard deviation (s) has been calculated assuming a normal distribution. Note that as the standard deviation increases at a set margin of error, the number of samples required increases. A similar relationship is observed for the margin error, with the number of samples increasing as the margin of error decreases for any selected standard deviation.

The *important point* to note is that to achieve a smaller margin of error, more samples are required for a fixed value of the standard deviation. This table applies *only* to the simple random sampling design example and is not intended as a substitute for calculating the appropriate number of samples.

If the stakeholders change the confidence interval, then the numbers in the table provided would change accordingly. If the confidence level is decreased below 90%, then the required number of samples reflected in this table would be lower for each margin of error and standard deviation combination.

SAMPLING DESIGNS

Information on the various types of sampling designs included in this section has been summarized from a number of Chapter 2 references [*1,7,8,9*]. This section discusses some basic concepts involved in selecting a sampling design that meets the study objectives (DQOs). Table 2.2 summarizes the advantages and limitations of several sampling design alternatives. Figure 2.6 illustrates the general pattern of sampling locations for each design. It's important to recognize that the U.S. Department of Energy (DOE) and the EPA are developing web-based software tools to assist investigators in identifying and selecting appropriate sampling designs including Visual Sampling Plan (VSP) and SampTool [*10*]. These programs are in their formative stages at the time of publication for this manual, but should be available for use in the near future.

Authoritative Sampling Designs

Non-probabilistic or "authoritative" sampling designs are based on the expertise of the investigator(s) and the knowledge that they have concerning the waste stream or site that is being studied. In practice, authoritative designs are frequently used because they meet the objectives of the primary decision maker while minimizing the complexity of the study. Authoritative designs are primarily developed based on site history, process knowledge, regulatory/programmatic

TABLE 2.1—Number of Samples Matrix.

Confidence Level	**Standard Deviation (s)**				
0.90 ($t_{0.90}$ = 1.282)	**0.5**	**1.0**	**2.1**	**3.0**	**4.0**
Margin of Error (Δ)	**Number of Samples (n)**				
0.1	**42**	**164**	**725**	**1479**	**2630**
0.25	**8**	**28**	**117**	**237**	**421**
0.5	**2**	**8**	**30**	**60**	**105**
1.0	**-**	**2**	**9**	**16**	**28**
2.0	**-**	**-**	**3**	**5**	**8**

TABLE 2.2—Guidance for Selection of Sampling Designs.

Sampling Design	*Appropriate Conditions for Use*	*Advantages*	*Limitations*
Probability Sampling			
Simple Random Sampling	Useful when the population of interest is relatively homogeneous (i.e., there are no major patterns or "hot spots" expected).	• Provides statistically unbiased estimates of the mean, proportions, and the variability. • Easy to understand and implement.	• Least preferred if patterns or trends are known to exist and are identifiable. • Localized clustering of sample points can occur.
Stratified Random Sampling	Most useful for estimating a parameter (e.g., the mean) of wastes exhibiting high heterogeneity (e.g., there are distinct portions or components of the waste with high and low constituent concentrations or characteristics).	• Ensures more uniform coverage of the entire target population. • Potential for achieving greater precision in estimates of the mean and variance. • May reduce costs over simple random and systematic sampling designs because fewer samples may be required. • Enables computation of reliable estimates for population subgroups of special interest.	• Requires some prior knowledge of the waste or media to define strata and to obtain a more precise estimate of the mean. • Statistical procedures for calculating the number of samples, the mean, and the variance are more complicated than for simple random sampling.
Systematic Sampling	Useful for estimating spatial patterns or trends over time.	• Preferred over simple random when sample locations are random within each systematic block or interval. • Practical and easy method for designating sample locations. • Ensures uniform coverage of site, unit, or process. • May be lower cost than simple random sampling because it is easier to implement.	• May be misleading if the sampling interval is aligned with the pattern of contamination, which could happen inadvertently if there is inadequate prior knowledge of the pattern of contamination. • Not truly random, but can be modified through use of the "random within blocks" design.

TABLE 2.2—(continued).

Sampling Design	*Appropriate Conditions for Use*	*Advantages*	*Limitations*
Probability Sampling (continued)			
Ranked Set Sampling	• Useful for reducing the number of samples required. • Useful when the cost of analysis is much greater than the cost of collecting samples. • Inexpensive auxiliary variable (based on expert knowledge or measurement) is needed and can be used to rank randomly selected population units with respect to the variable of interest. • Useful if the ranking method has a strong relationship with accurate measurements.	• Can reduce analytical costs.	• Requires expert knowledge of waste or process or use of auxiliary quantitative measurements to rank population units.
Sequential Sampling	• Applicable when sampling and/or analysis are quite expensive, when information concerning sampling and/or measurement variability is lacking, when the waste and site characteristics of interest are stable over the time frame of the sampling effort, or when the objective of the sampling effort is to test a specific hypothesis. • May not be especially useful if multiple waste characteristics are of interest or if rapid decision making is necessary.	• Can reduce the number of samples required to make a decision. • Allows a decision to be made with less sampling if there is a large difference between the two populations or between the population mean and the standard.	• If the concentration of the constituent of concern is only marginally different from the action level, sequential procedures will require an increasing number of samples approaching that required for other designs such as simple random or systematic sampling.

TABLE 2.2—(continued).

Sampling Design	*Appropriate Conditions for Use*	*Advantages*	*Limitations*
Authoritative Sampling			
Judgmental	• Useful for generating rough estimates of the average concentration or typical property. • To obtain preliminary information about a waste stream or site to facilitate planning or to gain familiarity with the waste matrix for analytical purposes. • To assess the usefulness of samples drawn from a small portion of the waste or site. • To screen samples in the field to identify "hot" samples for subsequent analysis in a laboratory.	• Can be very efficient with sufficient knowledge of the site or waste generation process. • Easy to do and explain.	• The utility of the sampling design is highly dependent on expert knowledge of waste. • Nonprobability-based so inference to the general population is difficult. • Cannot determine reliable estimates of variability.
Biased	• Useful to estimate "worst-case" or "best-case" conditions (e.g., to identify the composition of a leak, spill, or waste of unknown composition).		

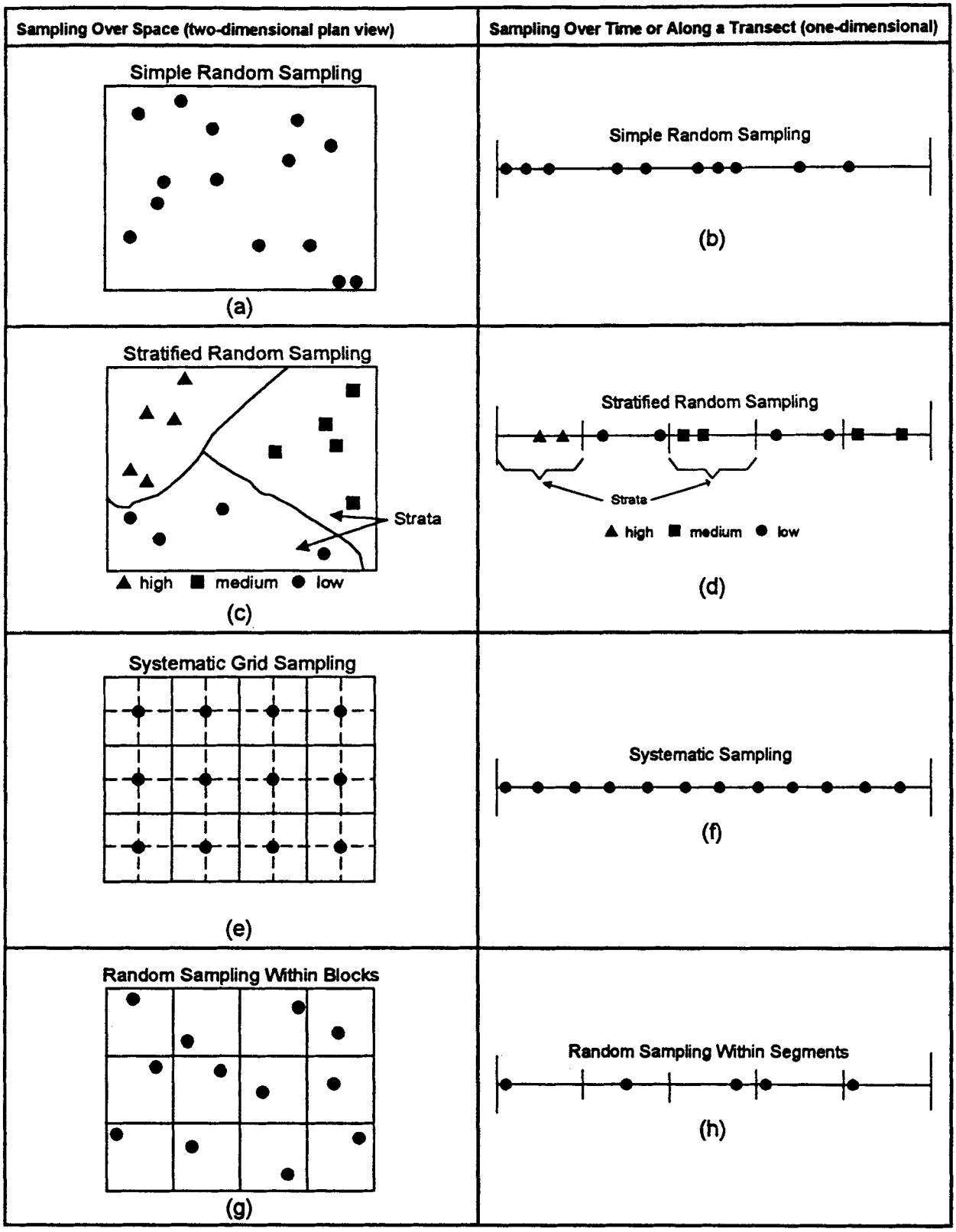


FIG. 2.6—Probability sampling designs over space or along an interval (modified after Cochran (1977) and Gilbert (1987)).

issues, and additional information identified in the conceptual site model. Information generated from subsequent DQO process steps (inputs, boundaries) is also used to optimize an authoritative design. This type of design is generally appropriate until the investigator's expert knowledge of the site or waste stream is exhausted.

Authoritative sampling designs are typically divided into two types: biased and judgmental (Table 2.2). Biased sampling is characterized by the selection of sampling locations in order to estimate "best case" (i.e., site background samples) or "worst case" conditions (i.e., at a known or suspected location of spill or point of release from a waste management unit). Biased sampling is commonly conducted in the early stages of a site assessment when little preliminary data exist and the site is being screened to determine if a further assessment or response action is warranted. Judgmental samples are typically collected to generate a rough estimate of the average concentration of a contaminant in a waste stream or on a site. However, judgmental designs may not be appropriate when the expected average contaminant concentration of a population is near the action level (see Appendix C, Case 1). Also, it is important to note that statistical measures of uncertainty cannot be developed with authoritative sampling designs.

Probabilistic (Statistical) Sampling Designs

Probabilistic sampling designs allow the results from a set of samples to be generalized to the entire decision unit. They have an element of randomization which allows probability statements to be made about the quality of estimates derived from the data, and every potential sampling point within the sampling unit has a probability of being sampled. Therefore, probabilistic samples are useful for testing hypotheses about whether a waste stream or site is contaminated, the level of contamination, and other questions common to RCRA sites. There are many different probabilistic sampling designs, each with advantages and disadvantages (see Table 2.2). A few of the most basic designs include simple random sampling, systematic sampling, and stratified sampling.

Simple Random Sampling

The simplest probabilistic sample is the simple random sample (Table 2.2). With a random sample, every possible sampling point has an equal probability of being selected, and each sample point is selected independently from all other sample points. Random sample locations are usually generated using a random number table or through computer generation of random numbers. Simple random sampling is appropriate when little or no information is available for a waste stream or a site, the population does not contain any trends, and it is acceptable to leave some portions of the population of interest less intensively sampled than other portions. If some information is available, simple random sampling may not be the most cost-effective sampling design available.

Systematic Sampling

Systematic sampling achieves a more uniform spread of sampling points than simple random sample by selecting sample locations using a spatial grid. It is useful for estimating spatial patterns or trends over time. To determine sample locations, a random starting point is chosen, the grid is laid out using this starting point as a guide, then all points on the grid (grid nodes) are sampled. Since sampling locations are located at equally spaced points, they may be easier to locate in the field than with simple random samples or other probability samples. However, a systematic sampling design should not be used if the contamination exhibits any cyclical patterns.

Stratified Sampling

Stratification of the study area may be used to improve the precision of a sampling design when areas of distinct variability exist. To create a stratified sample, divide the study area into two or more non-overlapping subsets (strata) that cover the entire site. Strata should be defined so that measurements within a stratum are more similar to each other than to measurements from other strata. Sampling depth, concentration level, previous sampling events, or contaminants present can be used as the basis for creating strata. Once the strata have been defined, each stratum is then sampled separately using either a random or systematic approach. A stratified sample can control the variability due to media, terrain characteristics, etc., if the strata are internally homogenous. Therefore a stratified random sample may provide more precise estimates of the mean contaminant level for the combined strata than those estimates obtained from a simple random sample. Even with imperfect information, a stratified sample can be more cost effective. In addition, stratification can be used to ensure that important areas of the site are represented in the sample. However, analysis of the data may be more complicated than other sampling designs. The boundaries for the decision must be determined prior to the development of the sampling design [7]. The purpose of defining strata for a stratified random sample is different from the purpose of defining strata for a scale of decisionmaking. The strata in a stratified random sample are sampled separately; then the data may be combined to create estimates for the entire site or scale of decisionmaking. Stratum estimates are also available; however, decisions made using individual stratum estimates will not have the same decision error rate unless the number of samples for each stratum was determined with that goal in mind.

Composite Sampling

If analysis costs are high compared to sampling cost and the parameter of interest is the mean, then the use of composite samples should be considered [7]. Composite sampling involves physically mixing two or more grab samples to create one sample for analysis. This method must be used in conjunction with a previously selected sampling design in order to determine sample locations (for instance, random composite sampling). Compositing samples can be a cost-effective way to incorporate a large number of sampling units (grabs) in one sample, and it provides better coverage of the site without analyzing each unit when the DQOs specify estimating average site conditions.

SUMMARY

The planning process must begin during the earliest stages of sampling plan development for a data collection activity. The DQO process may be followed in a strict, step-by-step fashion, or by a more informal approach that incorporates the seven elements of the DQO process in a less structured fashion [*4*]. Thoughtful planning will not only facilitate the implementation step, but also prepare for a successful data assessment step.

REFERENCES

[*1*] U.S. EPA, "Guidance for the Data Quality Objectives Process," QA/G-4, EPA 600/R-96/055, Office of Research and Development, Washington, DC, September 1994.

[*2*] ASTM, "Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives," D 5792-95, 1995.

[*3*] U.S. EPA, "Guidance on Implementation the Data Quality Objectives Process for Superfund," OSWER Directive 9355.9-01, EPA 540/R-93/071, Office of Solid Waste and Emergency Response, Washington, DC, August 1993.

[*4*] U.S. EPA, "Data Quality Objectives Decision Error Feasibility Trials (DQO/DEFT)—Users Guide—Version 4.0," EPA QA/G-4D, Office of Research and Development, Washington, DC, 1994.

[*5*] ASTM, "Guide for Generation of Environmental Data Related to Waste Management Activities: Selection and Optimization of Sampling Design," D 6311-98, 1998.

[*6*] U.S. EPA, "EPA Guidance for Quality Assurance Project Plans," QA/G-5, EPA/600/R-98/018, Office of Research and Development, Washington, DC, 1998.

[*7*] Gilbert, R. O., *Statistical Methods for Environmental Pollution Monitoring*, John Wiley and Sons, New York, NY, 1987.

[*8*] U.S. EPA, "RCRA Waste Sampling Draft Technical Guidance SW-846 Chapter Nine—Planning, Implementation, and Assessment," EPA/530/R-99/015, Solid Waste and Emergency Response, Washington, DC, 1999.

[*9*] U.S. EPA, "Data Quality Objectives for Hazardous Waste Site Investigations," Final Peer Review Draft, QA/G-4HW, Office of Research and Development, Washington, DC, 1999.

[*10*] US-DOE, Data Quality Objectives Home Page, *http://etd.pnl.gov:2080/DQO/*, Pacific Northwest National Laboratory (PNNL), Richland, WA.

Sampling for Waste Management Activities: Implementation Phase

3

INTRODUCTION

THE IMPLEMENTATION PHASE follows the planning stage of a sampling project and is comprised of data collection activities and technical assessment. While the analytical requirements are a part of the implementation stage of a project and are crucial for the success of the investigation, it is beyond the scope of this manual to provide the recommended analytical procedures for waste investigations. The implementation phase follows the planning step in the data generation process (Fig. 3-1).

The objective of the implementation phase is to collect and analyze the physical samples that will produce the data which will satisfy the DQO's developed in the planning stage. Field samplers should be able to minimize sampling bias (systematic error) and generate data that are of known quality by the proper selection and use of correct field sampling equipment, sample handling techniques, and unbiased subsampling methods. Data collection consists of project coordination, selection of sampling equipment, field activities, sampling waste units, post-sampling procedures, and field documentation.

Technical assessments are quality assurance (QA) tools and are conducted to ensure that the data collection activities meet the requirements as well as the intent of the QAPP developed in the planning stage. Some aspects of technical assessments may originate in the planning phase and extend into the data assessment portions of a project. However, it is important that there is verification that the data collection activities used were conducted appropriately. Technical assessment tools such as technical system audits, surveillance, and performance evaluations may be used to evaluate the effectiveness of the implementation phase of a project.

This manual does not purport to address all of the safety concerns, if any, associated with it use. It is the responsibility of the user of this manual to establish appropriate safety and health practices and determine the applicability of regulatory limits prior to use.

DATA COLLECTION

Project Preparations

Laboratory Coordination

Most field investigators have protocols to procuring a laboratory(s) that will satisfy the analytical requirements of an investigation. In fact, many samplers may have a contact or support staff that will fill this role. Additionally, laboratory analytical methods as well as other analytical needs associated with a sampling investigation should always be specified in the QAPP. However, there are many issues that a field project leader still needs to be aware of in order to effectively coordinate the sampling investigation. Some of these concerns for a sampler to address prior to sampling are: funding for the analytical services, deliverables, data quality objectives, minimum quantitation limits, turn-around times, scheduling, laboratory contact and phone number, laboratory capacity (if additional samples are collected), sample containers, preservatives, quality control blanks/spikes, laboratory's proximity to the site, and the laboratory's reputation and certification. As an investigation progresses, field investigators need to keep the laboratory contacts appraised of developments.

Site Entry and Site Reconnaissance

All sampling activities must be done in accordance with the appropriate statutory and regulatory authority. Site investigators and field samplers do *not* have the right to enter on private property without permission from the owner/operator/occupant of a facility/site, or a search warrant. All field investigators should explain the nature of the investigation prior to or at the time of the visit. If an investigation could lead to regulatory enforcement activities, investigators should show the owner/operator of the site identification. If the visit is not enforcement in nature, the facility should be contacted prior to any site reconnaissances or sampling event so that arrangements may be made to access all portions of the site.

Site reconnaissance of large-scale investigations are typically required and are recommended for smaller studies. If time or conditions do not permit a site reconnaissance, a walk through of the site should be conducted prior to any sampling. At least one member (usually the field project leader) of the potential field sampling crew should take part in the site reconnaissance. During a site reconnaissance, the following information may be obtained;

- verification of preliminary data
- site logistics (site sketches, maps, and photographs)
- site topography/drainage
- site conditions
- conditions and uses of adjoining property
- waste generation, storage or unit processes
- interviews with owners/operators/occupants
- available technical literature

- collect samples for variability
- collect samples for analytical screening levels
- target potential sample locations
- screen waste media for selection of sampling equipment
- air-monitoring readings

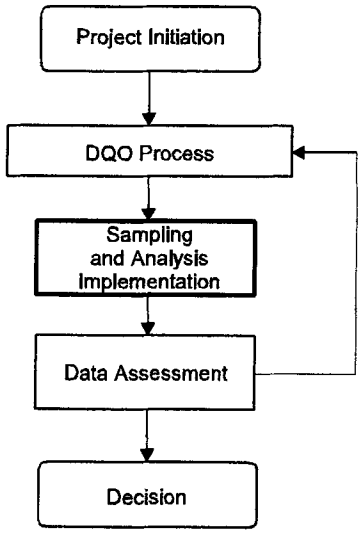


FIG. 3.1—Planning, implementation, and assessment steps.

- determine levels of personnel protection required to conduct investigation
- available utilities to conduct investigation (water, electricity, phones, etc.)
- conduct non-intrusive surveys (i.e., geophysical surveys)

Mobilization

Mobilization is considered the resources (time/money) it takes to get a sampling crew and their associated equipment to a facility/site and the time it takes to establish the essential components of the site so that the process of collecting samples may begin.

Oftentimes, the QAPP may include all members of a field sampling crew and may list each's responsibilities. However, due to the lengthy process of obtaining approval of work plans, personnel changes occur frequently. Each field investigation team will usually have their own policies for traveling and travel reimbursements, but the field project leader should make sure that all members of the team are aware of times, places, and modes of transportation prior to initiating a sampling investigation. In addition, it is important that the field project leader host a meeting with the complete sampling team to clarify the study's objective(s) and to define each member's responsibilities prior to traveling to the site so that everyone will arrive prepared.

Prior to departure, it is necessary to estimate the amount and type of equipment that will be required to conduct a sampling investigation. In addition to the equipment and containers that will actually be used to collect the samples, other ancillary equipment that may be required also needs to be included in the equipment estimate. Examples of the ancillary equipment may include-mixing pans and utensils, air-monitoring instruments/calibration gases, protective clothing, respiratory protection, field-screening instruments, container-opening tools, grounding wires, extension chords, generators, batteries, flash light, shipping supplies, decontamination supplies, garbage bags, oil wipes/towels, investigative derived waste containers, water coolers, first aid kit, vehicles, etc. . . . . If heavy equipment (drill rig, back-hoe, etc.) is required for an investigation and the services will be subcontracted, the field project leader needs to communicate clearly the responsibilities and expectations of the contractor in the statement of work (SOW). Even if field decontamination is going to be required as part of a study, it is desirable to have all sampling precleaned before arriving at a site because it is usually more effective and efficient to clean equipment in a control setting.

As long as the sampling is not being conducted as part of an on-going chemical spill or release, a walk-through shall be conducted prior to collecting samples so that all portions of the site under consideration are examined to determine if they are accessible. After the walk-through has been conducted to address health and safety and site security issues, mobilization can be completed by establishing the components of a site.

Components of a site may vary considerably depending on the site/facility, the potential hazards, the study's objectives, and size of the investigation. However, essential components of a site are a support zone (comprised of a command post, equipment storage, sample records processing), contamination reduction zone (also known as the decontamination area), and the exclusion zone (where sampling of waste media occurs) (Figure 3.2). These zones should always be delineated so that contaminated equipment can be segregated from clean areas. Even for an investigation requiring one sample, the essential components of the site should be established. For example, consider a small study which requires one sample from a waste unit. The support zone may be a vehicle with the front seat serving as the command post, the back seat as the equipment storage area, and the dash board may be a record-processing area. The contamination reduction zone may be located near the sample collection point and consist of bagging up disposable sampling equipment or washing reusable equipment in a bucket. For large-scale investigations, trailer(s), buildings, or structures may be constructed to be used as designated as areas for specific site functions.

The support zone may consist of many areas depending on the scope of the investigation. The primary function of a command post would be to serve as a place for internal communication and coordination for the sampling team. Some of these communications may include task assignments, daily progress and safety meetings, and changes in the scope of the investigation. Other activities that may occur at the command post are communications with the laboratory as well as other stake holders and concerned individuals, ordering of expendable sampling supplies, and data entry/management. Also, an area in the support zone should be designated as a clean area which should be used to store clean equipment and instruments, and perform calibrations and maintenance on field instruments. Later in this chapter, procedures for contamination reduction zone activities and decontamination of personnel and sampling equipment are discussed.

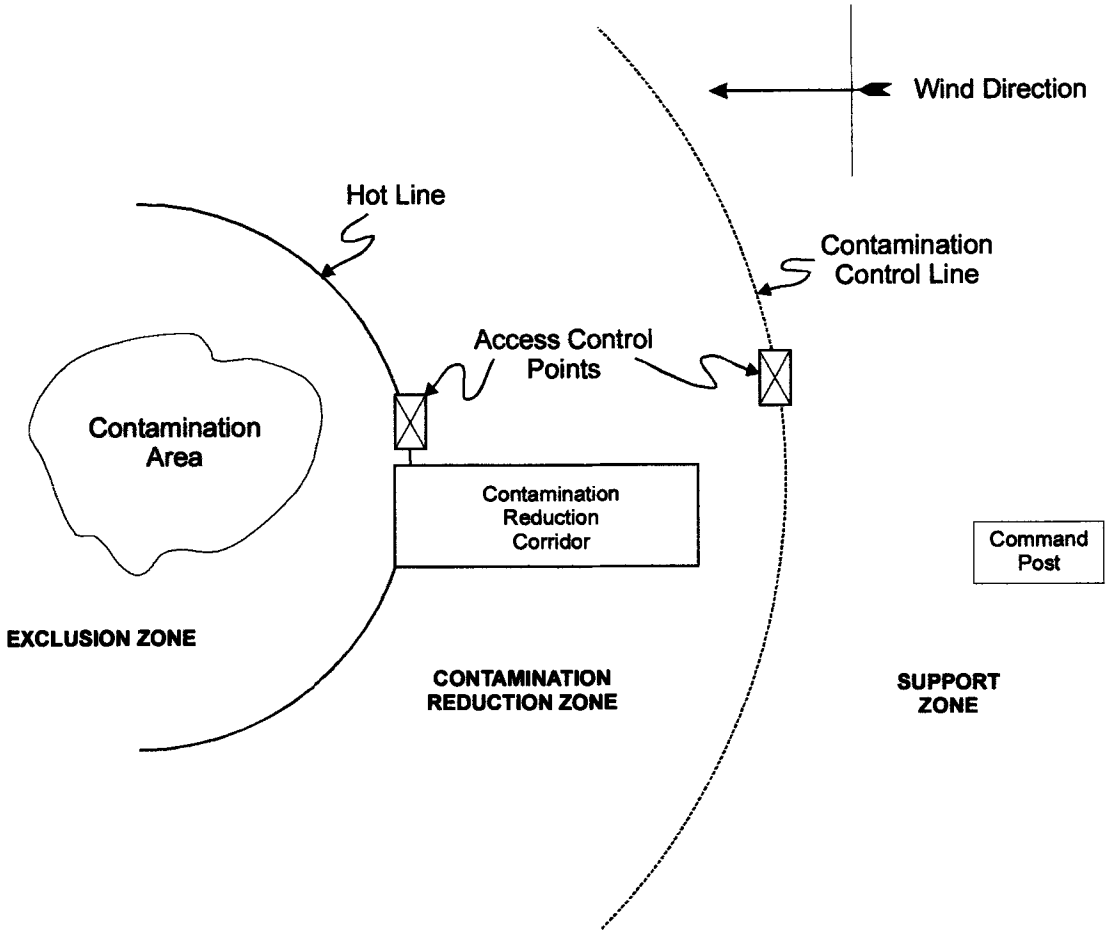


FIG. 3.2—Diagram of site work zones.

Selection of Sampling Equipment

Selecting appropriate sampling equipment for waste investigations can be a challenging task. Sampling equipment should be selected to accommodate all of the known physical characteristics of concern or chosen such that the effect of any sampling bias is understood [*1*]. Often because of a lack of preliminary information, varying field conditions, or waste heterogeneity, a piece of equipment selected during the sampling design may be unsuccessful for collecting a particular waste sample and another piece of equipment will be required as a substitute. All substitutions should be based on the study's DQOs, and any sampling bias or deficiencies resulting from the use of substituted equipment should be documented and reviewed with the data.

An extremely important factor in collecting samples of waste and contaminated media will be determined by the physical characteristics of the waste material. By selecting sampling equipment that will not discriminate against certain physical characteristics (e.g., phase, particle size, etc.), sampling bias can be minimized during waste sampling. Because wastes often stratify due to different densities of phases, settling of solids, or varying wastes constituents generated at different times, it may also be important to obtain a vertical cross section of the entire unit. Other considerations in the selection of sampling equipment are:

- the ability to access and extract from relevant location in the target population,
- the ability to collect a sufficient mass of sample such that the distribution of particle sizes in the population are represented,
- the compatibility (the ability to collect a sample without the addition or loss of constituents of interest),
- the ease of operation,
- the cost of the equipment, and
- the ability to properly decontaminate the sampling apparatus.

In addition to these considerations, analytical requirements such as sample handling and preparation to correctly analyze physical samples need to be considered. For consolidated/solidified wastes, samples will often be required to undergo particle size reduction (PSR) prior to chemical analyses. Any influences that these types of sample preparation/handling procedures or ancillary equipment may have on the data should be evaluated and reported as necessary. PSR will be discussed in a later section in this chapter.

There are many types and manufactures of sampling equipment that may be used to collect samples of wastes and contaminated media. ASTM D 6232, Standard Guide for Selection of Sampling Equipment for Waste and Contaminated Media Data Collection Activities, provides criteria for selecting sampling equipment [*1*]. The guide also provides lists of common, readily available sampling devices and their advantages/disadvantages, line drawings, and narratives describing their op-

eration. Tables 3.1, 3.2, and 3.3 are from this ASTM standard. A limited list of sampling equipment is presented in the tables. The list attempts to include a variety of different types of equipment. However, the list is not all inclusive. Table 3.1 lists matrices (surface and ground water, sediment, soil, and mixed phased wastes) and indicates which sampling devices are appropriate for use of these matrices. Table 3.2 indicates ASTM method references: physical requirements (such as batteries, electrical power, and weight); physical and chemical compatibility; effect on matrix; ease of operation; decontamination; and reusability. Table 3.3 provides a sampler-type selection process based upon the sample type and matrix to be sampled.

After careful evaluation of the waste unit and the study's objective, the experienced field sampler will usually be able to narrow the preferred choice to one or two pieces of sampling equipment. However, occasionally site-specific conditions may dictate that only one approach will work, even though that sampling equipment might not have been the preferred choice.

Field Activities

Selection of Sample Locations

Sample locations are usually specified in the QAPP. Oftentimes the locations might be depicted on a figure. However, when the sampler arrives at a site/facility, it may be difficult to transpose a point on a figure to one in the field, especially when many figures and sample location symbols are not to scale.

When the unit under consideration is containerized (i.e., drum, tank, etc.), there may be limited access points into the unit. This will restrict the initial sample location to the available access points. If there are multiple containers present, field screening may be required to help determine which ones would be suited to meet the study's objective.

Uncontainerized units may require some type of spatial measurements or the establishment of a grid to determine the appropriate sampling locations. Having the number of samples to collect specified in the QAPP, the project leader should then determine how to disperse the samples within the site if the information has not been specified. Commonly, a grid system is used for both probabilistic and non-probabilistic sampling designs. Sometimes the method of laying out the grid or the accuracy required to lay out a grid are not specified in the QAPP, or sometimes the grid pattern and logistics specified in the QAPP do not match up with the physical features at a site/facility. With the study's DQOs in mind, the field project leader must make the appropriate modifications to the proposed sample locations and then document it accordingly.

TABLE 3.1—Equipment Selection—Matrix Guide.

Equipment[A]	Water and Waste Water			Sediment	Soil	Liquid	Waste			
	Surface Water	Ground Water	Point Discharges				Multi-Layer Liquid	Mixed Phase Solid/Liquid	Consolidated Solid	Unconsolidated Solid
Pumps and Siphons										
Automatic Sampler - Non Volatiles	X[B]	N[C]	X[B]	...[D]	...[D]	N[C]	N[C]	N[C]	...[D]	...[D]
Automatic Composite - Sampler Volatiles	X[B]	N[C]	X[B]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Air/Gas Displacement Pump	X[B]	Guide D 4448	X[B]	...[D]	...[D]	X[B]	X[B]	...[D]	...[D]	...[D]
Piston Displacement Pump	X[B]	Guide D 4448	X[B]	...[D]	...[D]	X[B]	N[C]	...[D]	...[D]	...[D]
Bladder Pumps	X[B]	Guide D 4448	X[B]	...[D]	...[D]	N[C]	N[C]	...[D]	...[D]	...[D]
Peristaltic Pump	X[B]	Guide D 4448	X[B]	...[D]	...[D]	X[B]	X[B]	N[C]		
Centrifugal Submersible Pump	X[B]	X[B]	X[B]	...[D]	...[D]	X[B]	X[B]	...[D]	...[D]	...[D]
Dredges										
Eckman Dredge	...[D]	...[D]	...[D]	Guide D 4387	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Petersen Dredge	...[D]	...[D]	...[D]	Guide D 4387	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Ponar Dredge	...[D]	...[D]	...[D]	Guide D 4387	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Discrete Depth Samplers										
Bacon Bomb	X[B]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Kemmerer Sampler	X[B]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]
Syringe Sampler	...[D]	...[D]	N[C]	...[D]	...[D]	X[B]	X[B]	X[B]	...[D]	...[D]
Lidded Sludge/Water Sampler	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]	X[B]	X[B]	...[D]	N[C]
Discrete Level Sampler	X[B]	X[B]	X[B]	...[D]	...[D]	X[B]	X[B]	...[D]	...[D]	...[D]
Push Coring Devices										
Temporary G.W. Sampler	...[D]	X[B]	...[D]	...[D]	...[D]	N[C]	...[D]	...[D]	...[D]	...[D]
Penetrating Probe Sampler	...[D]	...[D]	...[D]	N[C]	X[B]	...[D]	...[D]	N[C]	...[D]	X[B]
Split Barrel Sampler	...[D]	...[D]	...[D]	X[B]	Test Method D 1586/ Guide D 4700	...[D]	...[D]	N[C]	...[D]	N[C]
Concentric Tube Thief	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]

TABLE 3.1—(continued).

Equipment[A]	Water and Waste Water			Sediment	Soil	Liquid	Waste			
	Surface Water	Ground Water	Point Discharges				Multi-Layer Liquid	Mixed Phase Solid/Liquid	Consolidated Solid	Unconsolidated Solid
Trier	...[D]	...[D]	...[D]	...[D]	X[B]	...[D]	...[D]	N[C]	...[D]	Practice D 4541
Thin Walled Tube	...[D]	...[D]	...[D]	Guide D 4823	Test Method D 1587/ Guide D 4700	...[D]	...[D]	...[D]	...[D]	X[B]
Coring Type w/Valve	...[D]	...[D]	...[D]	N[C]	X[B]	...[D]	...[D]	X[B]	...[D]	X[B]
Rotating Coring Devices										
Bucket Auger	[D]	...[D]	...[D]	N[C]	Practice D 1452/ Guide D 4700	...[D]	...[D]	...[D]	...[D]	X[B]
Screw Auger	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]	...[D]
Rotating Coring Device	...[D]	...[D]	...[D]	Guide D 4823	Guide D 4700	...[D]	...[D]	...[D]	X[B]	...[D]
Liquid Profile Devices										
COLIWASA	...[D]	...[D]	...[D]	...[D]	...[D]	Practice D 5495	Practice D 5495	Practice D 5495	...[D]	...[D]
Reuseable Point Sampler	N[C]	...[D]	N[C]	...[D]	...[D]	X[B]	X[B]	X[B]	...[D]	...[D]
Drum Thief	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]	X[B]	X[B]	...[D]	...[D]
Valved Drum Sampler	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]	X[B]	X[B]	...[D]	...[D]
Surface Sampling Devices										
Bailer	N[C]	Guide D 4448	...[D]	...[D]	...[D]	N[C]	N[C]	...[D]	...[D]	...[D]
Dipper	Practice D 5358	...[D]	Practice D 5013	...[D]	...[D]	Practice D 5358	...[D]	Practice D 5358	...[D]	...[D]
Impact Devices	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	...[D]	X[B]	...[D]
Spoon	N[C]	...[D]	N[C]	...[D]	Guide D 4700	N[C]	N[C]	...[D]	...[D]	N[C]
Scoops and Trowel	...[D]	...[D]	...[D]	N[C]	Guide D 4700	N[C]	...[D]	N[C]	...[D]	X[B]
Shovels	...[D]	...[D]	...[D]	N[C]	Guide D 4700	...[D]	...[D]	N[C]	...[D]	X[B]

[A] May be used for discrete sample collection.

[B] Equipment may be used with this matrix.

[C] Not equipment of choice but use is possible.

[D] Not recommended.

Field Screening

Field screening has been used successfully on many waste and contaminated media sampling investigations. Special statistical designs, such as double sampling and rank set sampling, utilize screening (auxiliary) data to increase the statistical power over simple random designs. The field-screening methods can and will vary considerably depending on the waste material and the DQOs of a particular project. Some of these successfully demonstrated field screening and analytical techniques include:

- colorimetric test strips,
- gas chromatography,
- Fourier-transformed infrared (FTIR),
- X-ray fluorescence,
- mercury vapor analyzer, and
- immunoassay.

Field screening can be very effective in waste characterization and extremely valuable in selecting appropriate sampling locations and chemical analyses when little preliminary data exist. Field investigators routinely use observations of the physical characteristics of waste units, air monitoring equipment, pH meters/paper, and for field flash point analyzers to confirm preliminary data or to decide on sampling locations during waste investigations. Figure 3.3 (RCRA Waste Characterization) is a flow diagram that depicts the process that field investigators may use to decide which waste containers to sample and what analyses to perform on particular samples. Such field screening techniques can be incorporated into the DQOs for a particular investigation. Results from the field screening would then be the basis for decisions made during implementation about sample locations and analyses.

Composite Sampling

When composite samples are going to be collected during a sampling investigation, they should be specified in the QAPP. Compositing is a physical averaging process that tends to produce samples containing constituents that are more normally distributed than grab samples. There are several advantages to collecting composite samples, such as:

- reduction in the variance of an estimated average,
- increasing the efficiency locating/identifying hot spots, and,
- reduction of sampling and analytical costs.

The sample mixing and subsampling procedures described in this manual are inappropriate for samples to be analyzed

TABLE 3.2—Sampling Equipment Selection Guide.

Equipment	Chemical[A,B]	Physical	Effect on Matrix	Range of Volume[C]	Physical Requirements[D]	Ease of Operation	Decon	Disposal and Reuse[E]
Pumps and Siphon								
Automatic Sampler - Non Volatiles	•	•	√	U	B/P	•	•	R
Automatic Composite Sampler - Volatiles	•	•	√	U	B/P	•	•	R
Air/Gas Displacement Pump	•	•	•	U	P/S/W	•	•	R
Piston Displacement Pump	•	•	•	U	P/S/W	•	•	R
Bladder Pumps	√	•	√	U	P/S/W	•	•	R
Peristaltic Pump	•	•	√	U	B/P	•	√	R
Centrifugal Submersible Pump	•	•	•	U	P/S/W	√	•	R
Dredges								
Eckman Dredge	√	√	•	0.5-3.0	N	•	•	R
Petersen Dredge	√	√	•	0.5-3.0	W	•	•	R
Ponar Dredge	√	√	•	0.5-3.0	W	•	•	R
Discrete Depth Samplers								
Bacon Bomb	•	√	√	0.1-0.5	N	√	•	R
Kemmerer Sampler	•	√	√	1.0-2.0	N	√	•	R
Syringe Sampler	√	√	√	0.2-0.5	N	√	√	R
Lidded Sludge/Water Sampler	√	•	•	1.0	S/W	•	•	R
Discrete Level Sampler	√	•	√	0.2-0.5	N	√	•	R
Push Corning Devices								
Temporary G.W. Sampler	√	√	√	0.1-0.3	P/S/W	•	•	R
Penetrating Probe Sampler	√	√	√	0.2-2.0	S/W	•	√	R
Split Barrel Sampler	√	√	•	0.5-30.0	S/W	√	√	R
Concentric Tube Thief	√	√	√	0.5-1.0	N	√	√	R
Trier	√	√	√	0.1-0.5	N	√	√	R
Thin Walled Tube	√	√	•	0.5-5.0	S/W	√	√	R
Coring Type w/Valve	√	√	√	0.2-1.5	N	√	√	R
Rotating Coring Devices								
Bucket Auger	√	√	•	0.2-1.0	N	•	√	R
Screw Auger	√	√	•	0.1-0.3	N	•	√	R
Rotating Coring Device	√	√	•	0.5-1.0	B/P	√	√	R
Liquid Profile Devices								
COLIWASA	√	•	√	0.5-3.0	N	√	•	D/R
Reuseable Point Sampler	√	√	√	0.2-0.6	N	√	√	R
Drum Thief	√	•	√	0.1-0.5	N	√	•	D/R
Valved Drum Sampler	√	√	√	0.3-1.6	N	√	√	D/R
Surface Sampling Devices								
Bailer	•	√	•	0.5-2.0	N	√	√	D/R
Dipper	√	√	√	0.5-1.0	N	√	√	R
Impact Devices	•	•	•	N/A	B/P	√	√	R
Spoon	√	√	•	N/A	N	√	√	R
Scoops and Trowel	√	√	•	0.1-0.6	N	√	√	R
Shovels	√	√	•	1.0-5.0	N	√	√	R

[A] • Significant operational consideration.
[B] √ Not a significant operational consideration.
[C] Range of Volume (litres)—U -Unlimited, and N/A -Not Applicable.
[D] Physical Requirements—B -Battery, S -Size, W -Weight, N -No limitations, and P -Power.
[E] Disposal and Reuse—R -Reusable, and D -Single use.

TABLE 3.3—Cross Index of Sampling Equipment.

Media Type	Sampler Type	Section	Sample Type
Consolidated	Rotating Corer	7.6.7	Surface or Depth, Undisturbed
Solid	Screw Auger	7.6.4	Surface, Disturbed
	Lidded Sludge	7.4.8	Discrete, Composite
	Penetrating Probe	7.5.4	Discrete, undisturbed
	Split Barrel	7.5.7	Discrete, Undisturbed
	Concentric Tube Thief	7.5.10	Surface, Disturbed, Selective
	Trier	7.5.10	Surface, Relatively Undisturbed, Selective
Unconsolidated	Thin Walled Tube	7.5.13	Surface or Depth, Undisturbed
Solid	Coring Type w/Valve	7.5.16	Surface or Depth, Disturbed
	Bucket Auger	7.6.1	Surface or Depth, Disturbed
	Spoon	7.8.10	Surface, Disturbed, Selective
	Scoops/Trowel	7.8.13	Surface, Disturbed, Selective
	Shovel	7.8.16	Surface, Disturbed
	Penetrating Probe	7.5.4	Discrete, Undisturbed
	Split Barrel	7.5.7	Discrete, Undisturbed
	Trier	7.5.10	Surface, Relatively Undisturbed, Selective
	Thin Walled Tube	7.5.13	Surface or Depth, Undisturbed
Soil	Coring Type w/Valve	7.5.16	Surface or Depth, Disturbed
	Bucket Auger	7.6.1	Surface or Depth, Disturbed

TABLE 3.3—(continued).

Media Type	Sampler Type	Section	Sample Type
Soil (continued)	Rotating Corer	7.6.7	Surface or Depth, Undisturbed
	Spoon	7.8.1	Surface, Disturbed, Selective
	Scoops/Trowel	7.8.13	Surface, Disturbed, Selective
	Shovel	7.8.16	Surface, Disturbed
Mixed Solid/Liquid	AutoSampler, Non V.	7.2.1	Shallow, Composite-Suspended Solids only
	Peristaltic Pump	7.2.10	Shallow, Discrete or Composite-Suspended Solids Only
	Syringe Sampler	7.4.7	Shallow, Discrete, Disturbed
	Lidded Sludge/Water	7.4.8	Discrete
	Penetrating Probe	7.5.4	Depth, Discrete, Undisturbed
	Split Barrel	7.5.7	Depth, Discrete, Undisturbed
	Trier	7.5.10	Surface, Semi-solid only, Selective
	Coring Type w/Valve	7.5.16	Depth, Disturbed
	COLIWASA	7.7.1	Shallow, Composite, Semi-liquid only
	Reuseable Point	7.7.1	Shallow, Discrete
	Drum Thief	7.7.4	Shallow, Composite
	Valved Drum	7.7.7	Shallow, Composite
	Dipper	7.8.4	Shallow, Composite
	Scoops/Trowel	7.8.13	Shallow, Composite, Semi-solid only
	Shovel	7.8.16	Shallow, Composite, Semi-solid only
Sediments	Eckman Dredge	7.3.1	Bottom Surface, Soft only, Disturbed
	Petersen Dredge	7.3.2	Bottom Surface, Rocky or Soft, Disturbed
	Ponar	7.3.3	Bottom Surface, Rocky or Soft, Disturbed
	Penetrating Probe	7.5.4	Bottom Surface or Depth, Undisturbed
	Split Barrel	7.5.7	Bottom Surface or Depth, Undisturbed
	Thin Walled Tube	7.5.13	Bottom Surface or Depth, Undisturbed
	Coring Type w/Valve	7.5.16	Bottom Surface or Depth, Disturbed
	Bucket Auger	7.1.6	Bottom Surface, Disturbed
	Rotating Corer	7.6.7	Bottom Surface, Undisturbed if solid
	Scoops, Trowel	7.8.13	Exposed Surface only, Disturbed, Selective
	Shovel	7.8.16	Exposed Surface only, Disturbed
Surface Water	AutoSplr. -Non Vols.	7.2.1	Shallow (25 in.), Discrete or Composite
	Auto Splr. - Vols.	7.2.1	Shallow (25 in.), Discrete
	Air/Gas Displacement	7.2.4	Depth, Discrete
	Piston Displacement	7.2.4	Depth, Discrete
	Bladder Pump	7.2.7	Depth, Discrete
	Peristaltic Pump	7.2.10	Shallow(25 in.), Discrete
	Centrifugal Sub. Pump	7.2.13	Depth, Discrete
	Bacon Bomb	7.4.1	Depth, Discrete
	Kemmerer	7.7.4	Depth, Discrete
	Discrete Level	7.4.11	Depth, Discrete
	Reuseable Point	7.7.1	Shallow (8 in.), Discrete
	Bailer	7.8.1	Depth, Discrete
	Dipper	7.8.4	Shallow (10 in.), Composite
	Spoon	7.8.1	Shallow (1 in.), Composite
Ground Water	AutoSplr. -Non Vols.	7.2.1	Shallow (25 in.), Discrete or Composite
	Auto Splr. - Vols.	7.2.1	Shallow (25 in.), Discrete
	Air/Gas	7.2.4	Depth, Discrete
	Piston Displacement	7.2.4	Depth, Discrete
	Bladder Pump	7.2.7	Depth, Discrete
	Peristaltic Pump	7.2.10	Shallow(25 in.), Discrete
	Centrifugal Sub.	7.2.13	Depth, Discrete
	Discrete Level	7.4.11	Depth, Discrete
	Temp. Ground Water	7.5.1.1	Depth, Discrete
	Bailer	7.8.1	Depth, Discrete
Point Discharges	AutoSplr. -Non Vols.	7.2.1	Shallow (25 in.), Discrete or Composite
	Auto Splr. - Vols.	7.2.1	Shallow (25 in.), Discrete
	Air/Gas	7.2.4	Depth, Discrete
	Piston Displacement	7.2.4	Depth, Discrete
	Bladder Pump	7.2.7	Depth, Discrete
	Peristaltic Pump	7.2.10	Shallow (25 in.), Discrete
	Centrifugal Sub.	7.2.13	Depth, Discrete
	Syringe Sampler	7.4.7	Shallow (8 in.), Discrete
	Discrete Level	7.4.11	Depth, Discrete
	Reuseable Point	7.7.1	Shallow (8 in.), Discrete
	Dipper	7.8.4	Shallow (10 in.), Composite
	Spoon	7.8.1	Shallow (1 in.), Composite
Liquid	AutoSplr. -Non Vols.	7.2.1	Shallow (25 in.), Discrete or Composite
	Air/Gas	7.2.4	Depth, Discrete
	Piston Displacement	7.2.4	Depth, Discrete
	Bladder Pump	7.2.7	Depth, Discrete
	Peristaltic Pump	7.2.10	Shallow (25 in.), Discrete

TABLE 3.3—(continued).

Media Type	Sampler Type	Section	Sample Type
	Centrifugal Sub.	7.2.13	Depth, Discrete
	Syringe Sampler	7.4.7	Shallow (8 in.), Discrete
	Lidded Sludge/Water	7.4.8	Shallow (8 in.), Discrete
Liquid	Discrete Level	7.4.11	Depth, Discrete
(continued)	Temp. Ground Water	7.5.11	Depth, Discrete
	COLIWASA	7.7.1	Shallow (4 in.), Composite
	Reuseable Point	7.7.1	Shallow (8 in.), Discrete
	Drum Thief	7.7.4	Shallow (3 in.), Composite
	Valved Drum Sampler	7.6.7	Shallow (8 in.), Composite
	Bailer	7.8.1	Depth, Discrete
	Dipper	7.8.4	Shallow (10 in.), Composite
	Spoon	7.8.1	Shallow (1 in.), Composite
	Scoops & Trowel	7.8.10	Shallow, (1 in.), Composite
	AutoSplr. -Non Vols.	7.2.1	Shallow (25 in.), Discrete or Composite
	Air/Gas	7.2.4	Depth, Discrete
	Piston Displacement	7.2.4	Depth, Discrete
	Bladder Pump	7.2.7	Depth, Discrete
	Peristaltic Pump	7.2.10	Shallow(25 in.), Discrete
	Centrifugal	7.2.13	Depth, Discrete
	Syringe Sampler	7.4.7	Shallow (8 in.), Discrete
Multi Layer	Lidded Sludge/Water	7.4.8	Shallow (8 in.), Discrete
Liquid	Discrete Level	7.4.11	Depth, Discrete
	Temp. Ground Water	7.5.1.1	Depth, Discrete
	COLIWASA	7.7.1	Shallow (4 in.), Composite
	Reuseable Point	7.7.1	Shallow (8 in.), Discrete
	Drum Thief	7.7.4	Shallow (3 in.), Composite
	Valved Drum	7.6.7	Shallow (8 in.), Composite
	Bailer	7.8.1	Depth, Discrete
	Spoon	7.8.1	Shallow (1 in.), Composite

for volatile organic compounds. Volatile organics are typically lost through volatilization during the sample collection and handling procedures. Other limitations to composite sampling include the loss of discrete information contained in a single sample and the potential for dilution of contaminants in a sample with uncontaminated material.

ASTM D 6051, Standard Guide for Composite Sampling and Field Subsampling for Environmental Waste Management Activities, discusses the advantages and appropriate use of composite sampling, and field procedures and techniques to mix the composite and procedures to collect an unbiased and precise subsample(s) from a larger sample [*2*].

Field mixing of composite sampling is considered essential. The following are some common methods for mixing solid and semi-solid samples: pan mixing/quartering, mixing square/kneading, sieving, and mixing. Field sub-sampling procedures include: rectangular scoop, alternate scoop, and slab cake.

Heterogeneous Waste

Sampling of any population may be difficult. However, with all other variables being the same, non-random heterogeneous populations are usually more difficult. The increased difficulty in sampling heterogeneous populations is due to the existence of unidentified or numerous strata, or both. If the existence of strata are not considered when sampling a non-random heterogeneous population, the resulting data will average the measured characteristics of the individual strata over the entire population. ASTM D 5956, Standard Guide for Sampling Strategies for Heterogeneous Waste, serves as a guide to develop sampling strategies for heterogeneous waste [*3*]. Sometimes there is little preliminary data available to the field investigator when collecting waste samples or contaminated media. If a heterogeneous waste population is encountered, the sampler must consider its impact on the investigation. The objectives of the investigation may have to be modified. When collecting waste samples, the field investigator must be aware of some of the physical signs that might reveal that material is a heterogeneous waste. Waste can be heterogenous in particle size or composition, or both, allowing for the existence of the following:

- strata of different size items of similar composition,
- strata of similar sized items of different composition, and,
- strata of different size items of different composition.

Sampling Waste Units

Waste management units can be generally categorized into two types: uncontainerized and containerized. In practice, uncontainerized units are larger than containerized units. Uncontainerized units include waste piles and surface impoundments, whereas containerized units include containers and tanks as well as ancillary tank equipment. Besides containers and tanks, sumps may also be considered containerized units because they are designed to collect the spillage of liquid wastes and are sometimes configured as a confined space.

Although both may pose hazards, units that are uncontainerized to the environment are generally less hazardous than containerized units. Sampling of containerized units is considered a higher hazard risk because of the potential of exposure to toxic gases and flammable/explosive atmospheres. Because containerized units prevent the dilution of the wastes by environmental influences, they are more likely to contain materials that have concentrated levels of hazardous constituents. While opening containerized units for

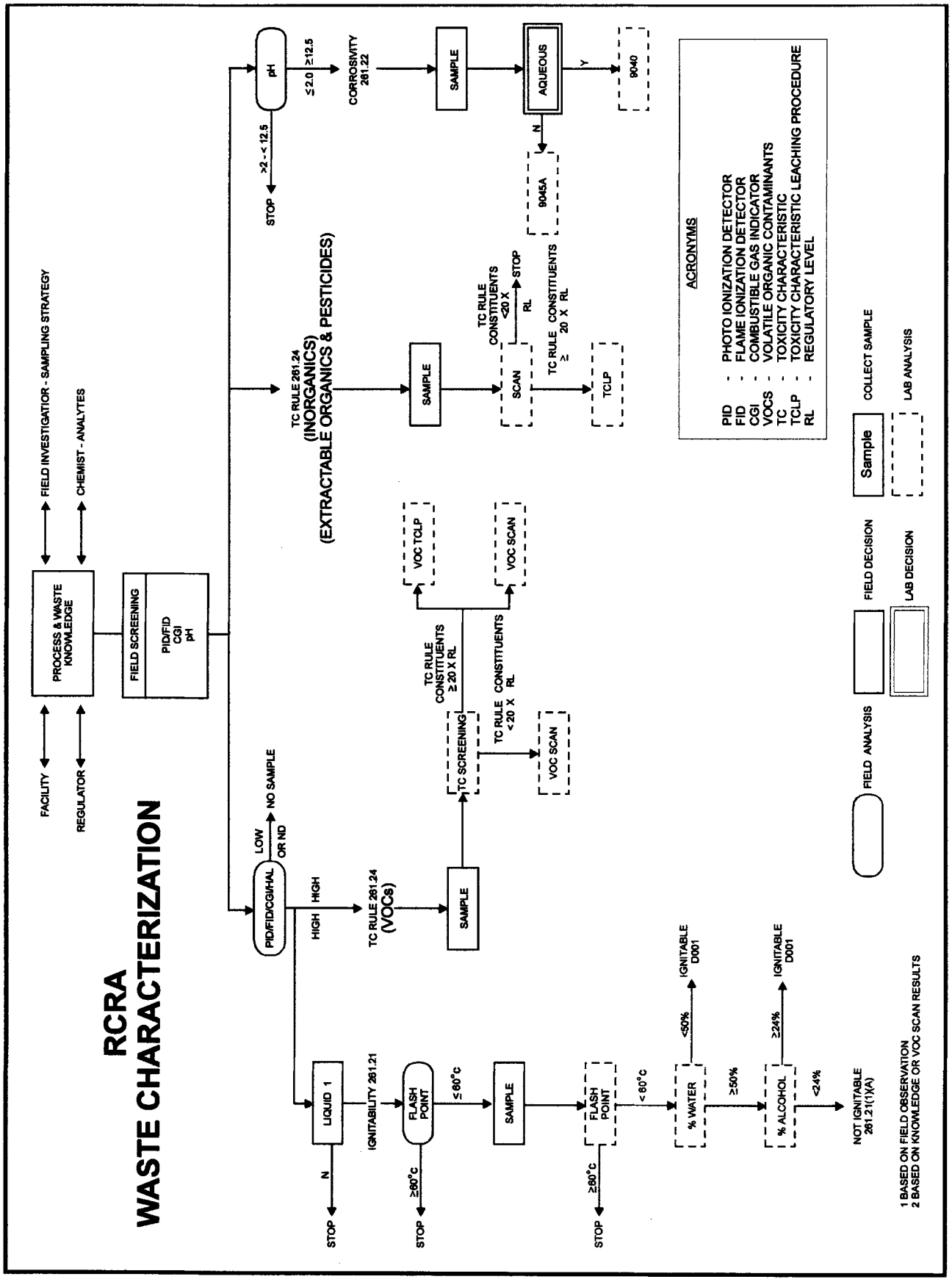


FIG. 3.3—RCRA waste characterization.

sampling purposes, investigators should use Level B personnel protective equipment, air-monitoring instruments to ensure that the working environment does not contain hazardous levels of flammable/explosive gasses or toxic vapors, and follow the appropriate safety requirements stipulated in the site specific safety plan.

Uncontainerized Waste Units

While uncontainerized units may contain many types of wastes and come in a variety of shapes and sizes, they can be generally regarded as either waste piles or surface impoundments. Definitions of these two types of uncontainerized units from 40 CFR Part 260.10 are:

- *Waste pile*—any non-containerized accumulation of solid, non-flowing hazardous waste that is used for treatment or storage and that is not a containment building.
- *Surface impoundment*—". . . a facility or part of a facility which is a natural topographic depression, man-made excavation, or diked area formed primarily of earthen materials (although it may be lined with man-made materials), which is designed to hold the accumulation of liquid wastes or wastes containing free liquids, and which is not an injection well. Examples of surface impoundments are storage, settling and aeration pits, ponds, and lagoons."

One of the distinguishing features between waste piles and surface impoundments is the state of the waste. Waste piles typically contain solid or non-flowing materials, whereas liquid wastes are usually contained in surface impoundments. The nature of the waste will also determine the mode of delivering the waste to the unit. Wastes are commonly pumped or gravity fed into impoundments, while heavy equipment or trucks may be used to dump wastes in piles. Once the waste has been placed in an uncontainerized unit, the state of the waste may be altered by environmental factors (e.g., temperature, precipitation, etc.). Surface impoundments may contain several phases such as floating solids, liquid phase(s), and sludges. Waste piles are usually restricted to solids and semi-solids.

Containerized Units

There are a variety of designs, shapes, sizes, and functions of containerized units. In addition to the challenges of the various designs and the safety requirements for sampling them, containerized units are difficult to sample because they may contain liquid, solid, semi-solid/sludge, or any combination of phases. Based on the study's design, it may be necessary to obtain a cross-sectional profile of the containerized unit in an attempt to characterize the unit. The following are definitions of types of containerized waste units described in 40 CFR Part 260.10:

- *Container*—any portable device in which waste is stored, transported, treated, disposed, or otherwise handled. Examples of containers are drums, overpacks, pails, totes, and roll-offs. Portable tanks, tank trucks, and tank cars vary in size and may range from simple to extremely complex designs. Depending on the unit's design, it may be convenient to consider some of these storage units as tanks for sampling purposes even though they meet the definition of a container.
- *Tank*—a stationary device, designed to contain an accumulation of waste which is constructed primarily of non-earthen materials which provide structural support.
- *Ancillary tank equipment*—any device including, but not limited to, such devices as piping, fittings, flanges, valves, and pumps that are used to distribute, meter, or control the flow of waste from its point of generation to a storage or treatment tank(s), between waste storage and treatment tanks to a point of disposal on-site, or to a point of disposal off-site.
- *Sump*—any pit or reservoir that meets the definition of a tank and those troughs/trenches connected to it that serve to collect liquid wastes. (Note: some outdoor sumps may be considered uncontainerized units/surface impoundments.)

Although any of the containerized units may not be completely sealed and may be partially uncontainerized to the environment, the unit needs to be treated as a containerized unit for sampling purposes until a determination can be made. Once a containerized unit is opened, a review of the proposed sampling procedures and level of protection can be performed to determine if the personal protection equipment is suitable for the site conditions. Samples collected from different waste units should not be composited into one sample container without additional analytical and/or field screening data to determine if the materials in the units are compatible and will not cause an inadvertent chemical reaction.

Post Sampling Activities

Particle Size Reduction

Particle size reduction (PSR) of waste samples is periodically required in order to complete an analytical scan or the Toxicity Characteristic Leaching Procedure (TCLP) test. Samples that may require PSR include slags, bricks, glass/mirror cullet, wire, etc. PSR is performed on a sample to decrease the maximum item size of the field sample so that the field sample then can be split or subsampled. The difficulties in applying particle size reduction to waste samples are the following:

- Not all materials are easily amenable to PSR (i.e., stainless steel artifacts).
- Adequate PSR capabilities and capacities do not exist in all laboratories.
- PSR can change the properties of material (i.e., leachability).
- PSR can be a source of cross-contamination.
- PSR often is not applicable to volatile compounds.
- Large mass/volumes may have to be shipped, handled, and disposed.

SW-846 Method 1311 (TCLP) states "Particle size reduction is required, unless the solid has a surface area per gram of material equal to or greater than 3.1 cm^2, or is smaller than 1 cm in its narrowest dimension (i.e., capable of passing through a 9.5-mm (0.375-in.) standard sieve). If the surface area is smaller or the particle size larger than described above, prepare the solid portion of the waste for extraction by crushing, cutting, or grinding the waste to a surface area or particle size as described above" (55 *Federal Register* 26990). The method also states that the surface criteria are meant for filamentous (paper, cloth, etc.) waste materials, and that "Actual measurement of the surface area is not required, nor is it recommended." Also, the loss of volatile organic compounds could be significant during particle size reduction.

Waste samples that require particle size reduction are often too large for standard sample containers. If this is the case, the sample should be secured in a clean plastic bag and

processed using normal sample identification and chain-of-custody procedures.

Because of the difficulties in conducting particle size reduction, it may be completed in the field or at the laboratory where the conditions can be controlled. There are several commercial grinding devices available for sample preparation prior to laboratory analysis. However, these devices may be expensive, particularly if the sampling of the consolidated waste matrix is not a routine operation.

When trace levels of contaminants are of concern, special procedures and equipment may have to be developed for the PSR to meet the objectives of the investigation. If trace levels of contaminants are not a concern, the following procedure may be used for crushing and/or grinding a solid sample:

1. Remove the entire sample, including any fines that are contained in the plastic bag, and place them on the standard cleaned stainless steel pan.
2. Using a clean hammer, carefully crush or grind the solid material (safety glasses are required), attempting to minimize the loss of any material from the pan. Some materials may require vigorous striking by the hammer, followed by crushing or grinding. The material may be subject to crushing/grinding rather than striking.
3. Continue crushing/grinding the solid material until the sample size approximates 0.375 in. (9.5 mm). Attempt to minimize the creation of fines that are significantly smaller than 0.375 in. (9.5 cm) in diameter.
4. Pass the material through a clean 0.375-in. (9.5-cm) sieve into a glass pan.
5. Continue this process until a sufficient sample is obtained. Thoroughly mix the sample. Transfer the contents of the glass pan into the appropriate containers.
6. Attach the previously prepared tags and submit for analyses.

Personnel and Sampling Equipment Decontamination

For most investigations involving hazardous waste and concentrated, contaminated waste media, personnel and equipment decontamination will be required by all personnel/equipment leaving the exclusion zone. Sampling equipment should also be cleaned prior to the sampling event, and, possibly, field decontaminated if a device will have to be reused to collect more than one sample. Properly designed and executed decontamination procedures offer:

- reducing the potential for worker exposure,
- minimizing the spread of contamination, and
- improved data quality and reliability.

The following reagents may be used during decontamination procedures:

- *detergent*—non-phosphate detergent solution
- *acid rinse*—10% nitric or hydrochloric acid solution
- *solvent rinse*—isopropanol, acetone, or methanol; pesticide grade
- *control rinse water*—preferably from a water system of known chemical composition
- *deionized water*—organic-free reagent grade

Personnel Decontamination—Prior to exiting the exclusion zone at a hazardous waste site, all personnel and equipment (as needed) must undergo a thorough decontamination. Decontamination should be conducted in an organized, stepwise manner. If certain pieces of the protective equipment are removed prior to the elimination of potential problems by decontamination, the worker may suffer damage due to inhalation or skin contact with contaminants. It is therefore important that persons doing the decontamination work know the proper procedures and the order in which to perform them to insure that such potential personal injuries do not occur.

Personnel decontamination procedures will differ from site to site depending on the level of protection and if the protective clothing is disposable or not. Generally, reusable protective clothing/equipment should be washed with a detergent solution and rinsed with control water.

Sampling Equipment Decontamination—Prior to initiating a field sampling investigation, equipment that will contact the sample population should be washed with a detergent solution followed by a series of control water, desorbing agents, and deionized water rinses. Non-sample contacting equipment should be washed with a detergent solution and rinsed with control water. Although such techniques may be difficult to perform in the field, they may be necessary to most accurately evaluate low concentrations of the chemical constituent(s) of interest.

The following procedures are recommended for sampling equipment [*4*];

1. Wash with detergent solution using an inert brush to remove particles or film (for equipment like tubing, the solution may be circulated through the equipment).
2. Rinse thoroughly with control water.
3. Rinse with an inorganic desorbing agent (may be deleted for field decontamination due to safety considerations; and may also be deleted if samples will not undergo inorganic chemical analysis).
4. Rinse with control water.
5. Rinse with an organic desorbing agent (may be deleted if samples will not undergo organic chemical analysis, or if equipment is made of plastic material).
6. Rinse with deionized water.
7. Allow equipment to air dry prior to next use.
8. Wrap equipment for transport with inert material (aluminum foil or plastic wrap) until ready for use.

For non-contact sampling equipment, Steps 1, 2, 7, and 8 above should be employed. If the heavy equipment is the non-contact equipment, a portable power washer or steam-cleaning machine may be used.

It is also recommended that QA/QC samples be collected and analyzed to document the effectiveness of the decontamination procedures. Collection of rinse or wipe samples after decontamination will vary depending on the scope of the project.

Investigation Derived Waste (IDW)

Materials which may become IDW are:

- *Personnel protective equipment* (PPE)—This includes disposable coveralls, gloves, booties, respirator canisters, splash suits, etc.
- *Disposable equipment*—This includes plastic ground and equipment covers, aluminum foil, conduit pipe, composite liquid waste samplers (COLIWASAs), Teflon® tubing, broken or unused sample containers, sample container boxes, tape, etc.
- Soil cuttings from drilling or hand auguring.

TABLE 3.4—Management of IDW.

TYPE	HAZARDOUS	NON-HAZARDOUS
PPE-Disposable	Containerize in plastic 5-gallon bucket with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately.	Double bag waste. Place in dumpster with permission of site operator, otherwise make arrangements for appropriate disposal.
PPE-Reusable	Decontaminate. If the equipment cannot be decontaminated, containerize in plastic 5-gallon bucket with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately.	Decontaminate.
Spent Solvents	Containerize in original containers. Clearly identify contents. Leave on-site with permission of site operator, otherwise characterize and dispose of appropriately.	N/A
Soil Cuttings	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately..	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise arrange with site manager for testing and disposal.
Groundwater	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately..	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise arrange with site manager for testing and disposal.
Decontamination Water	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately.	Containerize in 55-gallon drum with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise arrange with site manager for testing and disposal.
Disposable Equipment	Containerize in 55-gallon drum or 5-gallon plastic bucket with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise characterize and dispose of appropriately.	Containerize in 55-gallon drum or 5-gallon plastic bucket with tight-fitting lid. Identify and leave on-site with permission of site operator, otherwise arrange with site manager for testing and disposal.
Trash	N/A	Double bag waste. Place in dumpster with permission of site operator, otherwise make arrangements for appropriate disposal.

- Drilling mud or water used for water rotary drilling.
- Groundwater obtained through well development or well purging.
- Cleaning fluids such as spent solvents and washwater.
- Packing and shipping materials.

Table 3.4 lists the types of IDW commonly generated during waste investigations, and current management practices. Disposal of non-hazardous IDW from hazardous waste sites should be addressed in the QAPP. To reduce the volume, it may be necessary to compact the waste into a reusable container, such as a 55-gal drum.

If the waste is from an active facility, permission should be sought from the operator of the facility to place the non-hazardous PPE, disposable equipment, and/or paper/cardboard wastes into the facility's dumpsters. These materials may be

placed into municipal dumpsters with the permission of the owner, or these materials may also be taken to a nearby municipal landfill. On larger studies, waste hauling services may be obtained and a dumpster located at the study site.

Disposal of non-hazardous IDW such as drill cuttings, purge or development water, decontamination washwater, drilling muds, etc. should be placed into a unit with an environmental permit such as a landfill or sanitary sewer. These materials must not be placed into dumpsters. If the facility at which the study is being conducted is active, permission should be sought to place these types of IDW into the facilities, treatment system. It may be feasible to spread drill cuttings around the borehole, or, if the well is temporary, to place the cuttings back into the borehole. Nonhazardous cuttings, purge water, or development water may also be placed in a pit in or near the source area. Nonhazardous monitoring well purge or development water may also be poured onto the ground downgradient of the monitoring well. Purge water from private potable wells which are in service may be discharged directly onto the ground surface.

Disposal of hazardous or suspected hazardous IDW must be specified in the approved QAPP. Hazardous IDW must be disposed as specified in US-EPA regulations. If appropriate, these wastes may be placed back in an active facility waste treatment system.

If on-site disposal is not feasible, and if the wastes are suspected to be hazardous, appropriate tests/analyses must be conducted to make that determination. If they are determined to be hazardous wastes, they must be properly contained and labeled. They may be stored on the site for a maximum of 90 days before they must be manifested and shipped to a permitted treatment or disposal facility. Generation of hazardous IDW must be anticipated, if possible, to permit arrangements for proper containerization, labeling, transportation, and disposal/treatment in accordance with US-EPA regulations.

The generation of hazardous IDW should be minimized to conserve resources. Care should be taken to keep non-hazardous materials segregated from hazardous waste-contaminated materials. The volume of spent solvents produced during equipment decontamination should be controlled by applying only the minimum amount of solvent necessary and capturing it separately from the washwater.

At a minimum the requirements of the management of hazardous IDW are as follows:

- Spent solvents must be properly disposed or recycled.
- All hazardous IDW must be containerized. Proper handling and disposal should be arranged prior to commencement of field activities.

Shipping Samples

Samples collected during field investigations or in response to a hazardous materials incident must be classified prior to shipment as either environmental or hazardous materials samples. In general, environmental samples include drinking water, most groundwater and ambient surface water, soil, sediment, treated municipal and industrial wastewater effluent, biological specimens, or any samples *not* expected to be contaminated with high levels of hazardous materials.

Samples collected from process wastewater streams, drums, bulk storage tanks, soil, sediment, or water samples from areas suspected of being highly contaminated may require shipment as dangerous goods. Regulations for packing, marking, labeling, and shipping of dangerous goods by air transport are promulgated by the International Air Transport Authority (IATA), which is equivalent to the United Nations International Civil Aviation Organization (UN/ICAO) [5]. The project leader is responsible for determining if samples collected during a specific field investigation meet the definitions for dangerous goods.

Field Documentation

Field Records and Sample Identification

Detailed and accurate field records are integral elements of the field investigation process and are too often overlooked, both in the implementation and data assessment phases. Good field records will allow the pending data to be adequately evaluated, and, if need be, reconstruct the sampling effort.

The details of an investigation should be recorded in a site-dedicated, bound logbook. The project leader's name, the sample team leader's name (if appropriate), and the project name and location should be entered on the inside of the front cover of the logbook. It is recommended that each page in the logbook be numbered and dated. The entries should be legible and contain accurate and inclusive documentation of an individual's site activities. At the end of all entries for each day, or at the end of a particular event if appropriate, the investigator should draw a diagonal line and initial indicating the conclusion of the entry. Since field records are the basis for later written reports, language should be objective, factual, and free of personal feelings or other terminology which might prove inappropriate. All aspects of sample collection and handling, as well as visual observations, shall be documented in the field logbooks.

Information included in the logbook should include the following:

- address/location of sampling,
- name and address of field contact,
- generator of waste and address,
- waste generation process (if known),
- sample collection equipment (where appropriate),
- field analytical equipment, and equipment utilized to make physical measurements shall be identified,
- calculations, results, and calibration data for field sampling, field analytical, and field physical measurement equipment,
- serial numbers of any sampling equipment/monitoring used, if available,
- sampling station identification,
- date and time of sample collection,
- description of the sample location,
- description of the sample,
- who collected the sample,
- how the sample was collected,
- diagrams of processes,
- maps/sketches of sampling locations, and
- weather conditions that may affect the sample.

The method of sample identification used depends on the type of sample collected. Samples collected for specific field analyses or measurement data are recorded directly in bound field logbooks with identifying information. Examples include pH, temperature, and conductivity. Samples collected for laboratory analyses are identified by using standard sample tags/labels which are attached to the sample containers.

The following information shall be included on the sample tag/label using waterproof, non-erasable ink:

- field identification or sample station number,
- date and time of sample collection,
- designation of the sample as a grab or composite,
- type of sample (water, wastewater, leachate, soil, sediment, etc.) and a very brief description of the sampling location,
- the signature of either the sampler(s) or the designated sampling team leader and the field sample custodian (if appropriate),
- whether the sample is preserved or unpreserved,
- the general types of analyses to be performed (checked on front of tag), and
- any relevant comments (such as readily detectable or identifiable odor, color, or known toxic properties).

When samples are collected from vessels or containers which can be moved (drums for example), mark the vessel or container with the field identification or sample station number for future identification, when necessary. The vessel or container may be labeled with an indelible marker (e.g., paint stick or spray paint). The vessel or container need not be marked if it already has a unique marking or serial number; however, these numbers shall be recorded in the bound field logbooks. In addition, it is suggested that photographs of any physical evidence (markings, etc.) be taken and the necessary information recorded in the field logbook.

All field sample identification and field records should be recorded with waterproof, non-erasable ink. If errors are made in any of these documents, corrections should be made by crossing a single line through the error and entering the correct information. All corrections should be initialed and dated. If possible, all corrections should be made by the individual making the error.

Electronic data recorders, portable computers, and computer software have become widely available, which has greatly enhanced the amount of data acquisition that can be obtained during field investigations. As a result, the time it takes to adequately document and produce corresponding paperwork has been reduced. When using unfamiliar equipment to store crucial field records, it is prudent to confirm that the records will meet the study's objectives and that the data can be backed up.

Chain-of-Custody Procedures for Samples

Chain of custody procedures are used to maintain and document the possession of samples from the time of collection until sample disposal [*4*]. The procedures are intended to document sample possession during each stage of a sample's life cycle (i.e., collection, shipment, storage, and the process of analysis). Chain-of-custody procedures are comprised of the following elements: (1) maintaining sample custody, and (2) documentation of samples for evidence. To document chain-of-custody, an accurate record must be maintained to trace the possession of each sample from the moment of collection to its disposal.

A sample is in custody if:

- it is in the actual possession of an investigator,
- it is in the view of an investigator, after being in their physical possession,
- it was in the physical possession of an investigator and then they secured it to prevent tampering, and/or
- it is placed in a designated secure area.

Custody seals should be used to document that the sample container has not been tampered with prior to analyses. Samples should be sealed as soon as possible following collection utilizing an appropriate custody seal. The use of custody seals may be waived if field investigators keep the samples in their custody from the time of collection until the samples are delivered to the laboratory analyzing the samples.

The field Chain-of-Custody Record is used to record the custody of all samples or other physical evidence collected and maintained by investigators. All sample sets shall be accompanied by a Chain-of-Custody Record. This Chain-of-Custody Record documents transfer of custody of samples from the sample custodian to another person, to the laboratory, or other organizational elements. To simplify the Chain-of-Custody Record and eliminate potential litigation problems, as few people as possible should have custody of the samples during the investigation. A separate Chain-of-Custody Record should be used for each final destination or laboratory utilized during the investigation.

A typical field Chain-of-Custody Record would be Fig. 3.4. The following information should be supplied in the indicated spaces to complete the field Chain-of-Custody Record.

- The project number.
- The project name.
- All samplers and sampling team leaders (if applicable) should sign in the designated signature block.
- The sampling station number, date, and time of sample collection, grab or composite sample designation, and a brief description of the type of sample and/or the sampling location must be included on each line. One sample should be entered on each line and a sample should not be split among multiple lines.
- The Remarks section may be used to record air bill numbers, registered or certified mail serial numbers, or other pertinent information. The total number of sample containers must be listed in the "Total Containers" column for each sample. The number of individual containers for each analysis must also be listed. There should not be more than one sample type per sample. Required analyses should be circled or entered in the appropriate location as indicated on the Chain-of-Custody Record.
- The tag numbers for each sample and any needed remarks are to be supplied in the "Tag No./Remarks" column.
- The sample custodian and subsequent transferee(s) should document the transfer of the samples listed on the Chain-of-Custody Record. The person who originally relinquishes custody should be the sample custodian. Both the person relinquishing the samples and the person receiving them must sign the form. The date and time that this occurred should be documented in the proper space on the Chain-of-Custody Record.
- Usually, the last person receiving the samples or evidence should be the laboratory sample custodian or their designee(s).
- Any errors made on the field Chain-of-Custody Record should be corrected by crossing a single line through the error and entering the correct information. All corrections should be initialed and dated.

If custody of samples will be transferred with shipment, the samples shall be properly packaged for shipment in accordance with the appropriate US DOT and IATA procedures and regulations. All samples shall be accompanied by the Chain-of-Custody Record. The original and one copy of the

PROJ. NO.	PROJECT NAME
SAMPLERS (Signature)	

STA NO.	DATE	TIME	COMP	GRAB	STATION LOCATION	NO OF CONT-AINERS								REMARKS

Relinquished by (Signature)	Date/Time	Received by (Signature)	Relinquished by (Signature)	Date/Time	Received by (Signature)
Relinquished by (Signature)	Date/Time	Received by (Signature)	Relinquished by (Signature)	Date/Time	Received by (Signature)
Relinquished by (Signature)	Date/Time	Received by Laboratory by	Date/Time	Remarks	
Distribution Original Accompanies Shipment Copy to Coordinator Field Files					

FIG. 3.4—Chain-of-custody form.

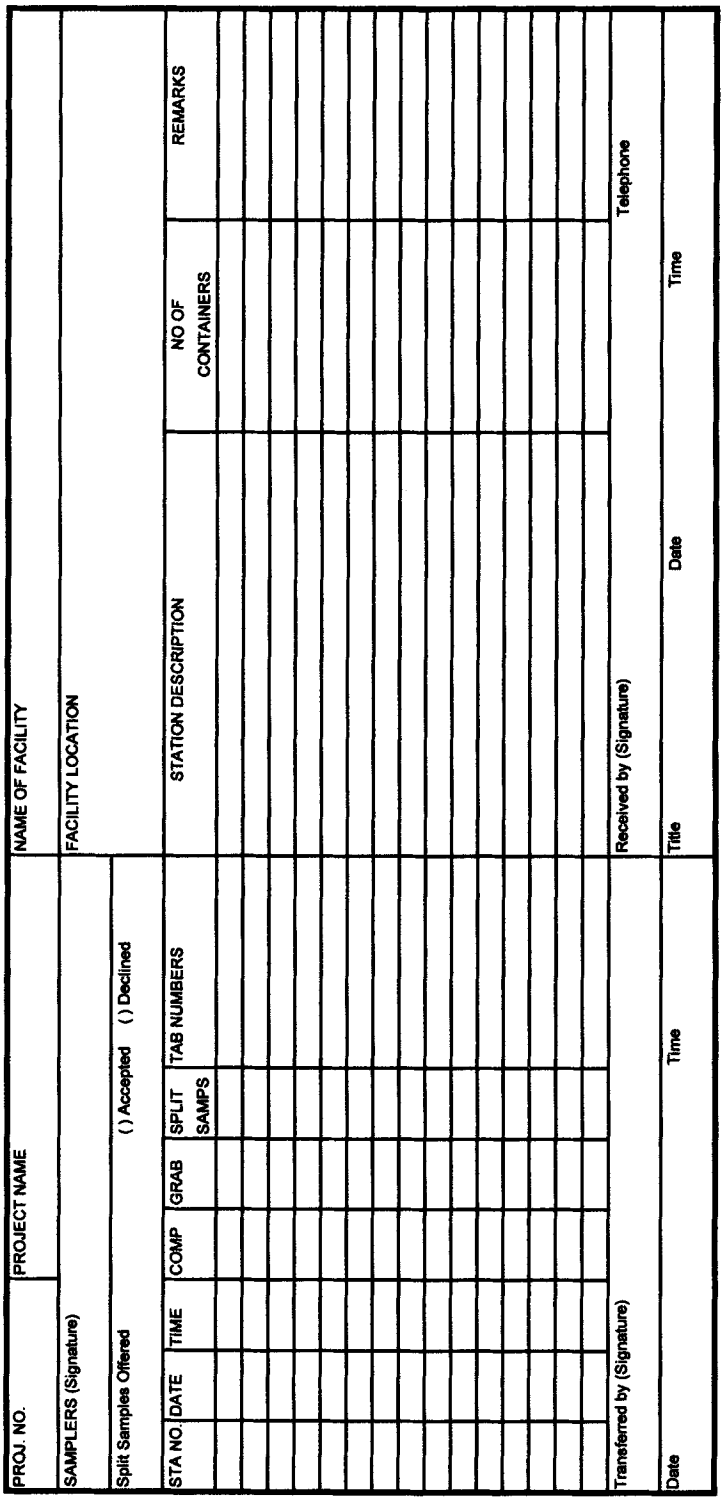


FIG. 3.5—Receipt for samples form.

Record will be placed in a plastic bag inside the secured shipping container if samples are shipped. When shipping samples via common carrier, the "Relinquished By" box should be filled in; however, the "Received By" box should be left blank. The laboratory sample custodian is responsible for receiving custody of the samples and will fill in the "Received By" section of the Chain-of-Custody Record. One copy of the Record will be retained by the project leader. The original Chain-of-Custody Record will be transmitted to the project leader after the samples are accepted by the laboratory. This copy will become a part of the project file.

If sent by mail, the package shall be registered with return receipt requested. If sent by common carrier, a Government Bill of Lading (GBL) or Air Bill should be used. Receipts from post offices, copies of GBLs, and Air Bills shall be retained as part of the documentation of the chain-of-custody. The Air Bill number, GBL number, or registered mail serial number shall be recorded in the remarks section of the Chain-of-Custody Record or in another designated area.

Section 3007 of the Resource Conservation and Recovery Act (RCRA) of 1976 and Section 104 of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA or Superfund) of 1980 require that a "receipt" for all facility samples collected during inspections and investigations be given to the owner/operator of each facility before the field investigator departs the premises. The Toxic Substances Control Act (TSCA) contains similar provisions. Figure 3.5 depicts a typical Receipt for Samples Form.

TECHNICAL ASSESSMENTS

Technical assessments are used to check that a data collection activity is conducted as planned, that is, producing data and information of the type and quality specified in the QAPP [*6*]. While they rely on the data quality criteria developed during the planning process, technical assessments occur during the data collection phase. Technical assessments play an important role in documenting the implementation of the QAPP and are used to check whether DQOs and other quality goals are being met. If the quality of the data generated by the project is found to be inadequate, corrective action is taken.

Technical assessments are usually performed by personnel external to the organization conducting the data collection activities. However, internal audits may be appropriate depending on an organization's structure. Regardless if a technical assessment is being performed, the field project leader or the QA officer designee are responsible for continually monitoring individual compliance with the QA and QC programs and planning documents. A QA officer shall review the field records, results, and findings for compliance with the QA and QC programs and planning documents.

EPA uses several types of technical assessments to evaluate the effectiveness of QA implementation: technical system audits (TSAs), surveillance, and performance evaluations (PEs) as well as audits of data quality (ADQs).

Technical system audits (TSAs) are a thorough, systematic, on-site, qualitative audit of the total measurement system used to collect data. Auditors examine facilities/sites, equipment, personnel, training, field procedures (sampling/sample handling/decontamination), and record keeping. TSAs can also be extended to assess data validation, data management, and data assessment, but are most effective when they are initiated during the data collection phase. TSAs are ideally performed near the beginning of large-scale projects so that any deficiencies may be addressed quickly. Usually one or more assessors with the appropriate technical expertise conduct the audit. TSAs can reveal shortcomings in a project's management structure, policy, practices, or procedures.

Surveillance occurs in field and is an oversight activity to determine that field procedures are being implemented according to the QAPP. The frequency, duration, and who (internal/external) performs the surveillance activities may be prescribed in the QAPP or may be unannounced to the field personnel. Surveillance may be performed by a QA or technical expert(s) familiar with field procedures being implemented. First, the assessor(s) reviews the QAPP and any related project documents, which is then followed by observing the field personnel conducting the data collection activities. Surveillance should identify if field procedures are being implemented correctly and consistently.

Unlike TSAs and surveillance, PEs are used to attempt to quantify the effectiveness of a sampling investigation by using samples of known composition. EPA considers PE audits capable of determining if a measurement system is operating within the specified standards for precision and bias as prescribed in the QAPP.

A PE sample mimics production samples in all possible aspects. Blind PE samples refer to PE samples that are not distinguishable in any way to the individuals operating the measurement system.

Therefore, blind PE samples are treated routinely and are not subjected to any special treatment that an undisguised PE sample may receive. As a result, blind PEs may provide a more objective performance assessment than nonblind PE samples.

Other QC samples which may be collected as part of an audit to help determine the effectiveness of a sampling investigation are:

- duplicates,
- splits,
- equipment (rinse) blanks, and
- field blanks.

Results of a technical assessment should be reported to management with recommended requirements to correct any observed deficiencies.

REFERENCES

[*1*] ASTM, "Standard Guide for: Selection of Sampling Equipment for Waste and Contaminated Media Data Collection Activities," D 6232-98, 1998.

[*2*] ASTM, "Standard Guide for Composite Sampling and Field Subsampling for Environmental Waste Management Activities," D 6051-96, 1996.

[*3*] ASTM, "Standard Guide for Sampling Strategies for Heterogenous Wastes," D 5956-96, 1996.

[*4*] EPA, "Environmental Investigations Standard Operating Procedures and Quality Assurance Manual (EISOPQAM)," Science and Ecosystem Support Division, Athens, GA, May 1996.

[*5*] International Air Transport Association (IATA), "Dangerous Goods Regulations," International Air Transport Authority (IATA), 31st ed., effective 1 Jan. 1995.

[*6*] US-EPA, "Guidance on Technical Assessments for Environmental Data Operations," External Review Draft, EPA QA/G-7, Washington, DC, 1998.

Sampling for Waste Management Activities: Data Assessment Phase

4

INTRODUCTION

DATA QUALITY ASSESSMENT (DQA) is a scientific and statistical process that determines whether environmental data are of the right type, quality, and quantity to support a decision. The DQA process consists of five steps that parallel the activities of a statistician analyzing a data set for the first time. However, the DQA process makes use of statistical and graphical tools that non-statisticians can apply to data sets [*1–4*].

DQA is built on a fundamental premise: data quality, as a concept, is meaningful only when it relates to the intended use of the data. Data quality does not exist without some frame of reference; one must know the context in which the data will be used to establish a yardstick for judging whether or not the data set is adequate.

By using the DQA process, environmental scientists and managers can answer two fundamental questions: (1) Can the decision (or estimate) be made with the desired confidence, given the quality of the data set, and (2) How well can the sampling design used to collect the data set be expected to perform in other data collection events under different conditions? The first question addresses the data user's immediate needs. For example, if the data provide evidence strongly in favor of one course of action over another, then the decisionmaker can proceed knowing that the decision will be supported by unambiguous data. If the data do not show sufficient evidence to favor one alternative, the data analysis alerts the decisionmaker to this uncertainty.

The second question addresses the data user's future needs. Often, investigators decide to use a certain sampling design at a location different from that for which it was first designed. In these cases, they should determine how well the design is expected to perform given that the outcomes and environmental conditions will differ from those of the original event. By estimating the outcomes before the sampling design is implemented, investigators can make any necessary modifications and thus prevent costly additional follow-up rounds of sampling to supplement inadequate data.

DQA completes the project's life cycle begun in the planning phase of the project level of EPA's Quality System. Investigators use DQA to examine the data collected during the implementation stage of the project as well as any quality control (QA/QC) information associated with the data. Investigators also examine the outputs and assumptions of the Data Quality Objectives (DQO) Process to determine if the data meet the user's performance criteria.

OVERVIEW OF DATA QUALITY ASSESSMENT

DQA and the Data Life Cycle

In the project level of EPA's Quality System, the Data Life Cycle comprises three phases: planning, implementation, and assessment. During the planning phase, the DQO Process (Chapter 2) is used to define quantitative and qualitative criteria for determining when, where, and how many samples (measurements) to collect and a desired level of confidence. This information along with the sampling design, sampling methods, analytical procedures, and appropriate QA/QC procedures are documented in the Quality Assurance Project Plan. During the implementation phase, data (which include primary data and QC data) are collected following the QAPP specifications (Chapter 3). During the assessment phase, the data are first verified and validated to ensure that the sampling and analysis protocols specified in the QAPP were followed. The DQA Process is then conducted on the validated data set to determine if the data are sufficient and adequate for their intended use.

Overview of the Five Steps of the DQA Process

The DQA Process consists of five steps that roughly parallel the actions that an environmental statistician takes when analyzing a set of data. Although the DQA Process is presented in a linear fashion, it is important to note that it is iterative by nature.

Using the validated data, the DQA process includes five steps: (1) review the DQOs and study design; (2) prepare the data for statistical analysis; (3) conduct a preliminary review of the data and check statistical assumptions; (4) select and perform statistical analyses; and (5) draw conclusions from the data.

This section provides guidance on performing the five steps of the DQA Process. Supplemental guidance on statistical analysis of data can be found in the Appendix C example to this manual and in EPA's publication QA/G-9, *Guidance for Data Quality Assessment* [*2*]. Software tools, such as EPA's DataQUEST, are available to help with DQA implementation [*5*]. DataQUEST is free software, available for download on EPA's website.

Step 1—Review the DQOs and the Sampling Design

Review the DQO outputs to assure that they are still applicable. Refer back to Chapter 2 of this manual for more infor-

mation on the DQO Process, or see EPA QA/G-4 and QA/G-9 guidance documents. Pose and answer questions about the data. Do the data obtained meet the requirements specified during the DQO process? Were the data collected of the correct type, quality, and quantity as specified during the DQO process? A clear understanding of the original project objectives, as determined during the DQO Process, is critical in selecting the appropriate statistical test and interpreting the results relative to the RCRA regulatory requirements.

Step 2—Prepare Data for Statistical Analysis

After the data validation and verification, and before the data are available in a form for further analysis, several intermediate steps usually are required. For most situations, EPA recommends you prepare the data in computer-readable format. Steps in preparing data for statistical analysis are outlined below (modified from Ott [*6*]):

Receive the verified and validated source from the QA reports. Data are supplied to the user in a variety of formats and readiness for use, depending on the size and complexity of the study and the type of analyses requested. Most laboratories supply a QA evaluation package including the validation/verification review; a narrative; tabulated summary forms (including the results of analyses of field samples, laboratory standards, and QC samples); copies of logbook pages; and copies of chain-of-custody records. From this information, you can create a database for analysis.

Create a database from the verified and validated data source. For most studies in which statistical analyses are scheduled, a computer-readable database is the most efficient method for managing the data. The steps required to create the database and the format used will depend upon the software systems used to perform the analysis. For example, the database may be as simple as a string of concentration values for a single constituent input into a spreadsheet or word processor (such as required for use of EPA's DataQUEST software) or it may be more complex, requiring multiple and related data inputs such as sample number, location coordinates, depth, data and time of collection, constituent name and concentration, units of measurements, test method, detection limit achieved, QC information, etc.

If the database is created via manual data entry, the verified and validated data should be checked for legibility. Any questions pertaining to illegible information should be resolved before the data are entered. Any special coding considerations should be specified in a coding guide or in the QAPP. For very large projects, it may be appropriate to prepare a separate detailed data management plan in advance.

Check and edit the data base. After creation of the data set, the database should be checked against the data source to verify accurate data entry and to correct any errors discovered. Even if the database is received from the laboratory in electronic format, it should be checked for obvious errors, such as unit errors, decimal errors, missing values, and detection limits.

Create data files from the data base. From the original data files, work files are created for use within the statistical software package. When creating the final data files for use in the statistical software, be sure to use a file naming and storage convention that facilitates easy retrieval for future use, reference, or reporting.

Step 3—Conduct Preliminary Analysis of the Data and Check Statistical Assumptions

Many statistical tests and procedures require that certain assumptions be met for their use. Failure to satisfy these assumptions can result in biased estimates of the parameters of interest. Therefore, it is important to conduct preliminary analyses of the data to learn about the characteristics of the data set. We recommend you compute statistical quantities, determine the proportion of the data reported as "non-detect" for each constituent of concern, check whether the data exhibit a normal distribution, and determine if there are any "outliers" that deserve a closer look. The outputs of these activities are used to select and perform the appropriate statistical tests.

Statistical Quantities

To help "visualize" and summarize the data, calculate basic statistical quantities such as the:

- mean,
- maximum,
- percentiles,
- variance,
- standard deviation,
- coefficient of variation (CV), and
- standard error of the mean.

Calculate appropriate parameters for each constituent of concern. Example calculations of the mean, variance, coefficient of variation, and standard deviation are given in the waste pile example in Appendix C of this manual. Detailed guidance on the calculation of statistical quantities is provided in Chapter 2 of EPA's QA/G-9 guidance document, and the useful quantities can easily be computed using EPA's DataQUEST software or other statistical package [*2,5*]. When calculating statistical quantities, determine which data points were reported as below a limit of detection, or "non-detect" (ND). See EPA's QA/G-9 for guidance on handling data sets which contain non-detects [*2*].

Checking Data for Normality

Check the data sets for normality by using graphical methods such as histograms, box and whisker plots, and normal probability plots, or use statistical methods such as the Shapiro-Wilk test for normality. Table 4-1 provides a summary of recommended methods. An example employing several of these techniques can be found in Appendix C of this manual (waste pile example). Detailed guidance on the use of graphical and statistical methods can be found in EPA's QA/G-9 guidance document [*2*]. Graphical methods allow you to visualize the central tendency of the data, the variability in the data, the location of extreme data values, and any obvious trends in the data.

The Shapiro-Wilk test is recommended as a superior method for testing normality of the data. The specific method for implementing the Shapiro-Wilk Test is provided in Gilbert's *Statistical Methods for Environmental Pollution Monitoring* [*3*], and it can also be performed with DataQUEST software or other commercially available statistical software.

TABLE 4.1—Recommended Graphical and Statistical Methods for Checking Distributional Assumptions.

Test	Use	Reference
	GRAPHICAL METHODS	
Histograms and frequency plots	Provides visual display of probability or frequency distribution.	See EPA QA/G-9 (USEPA 1998). Construct via EPA's DataQUEST software (USEPA 1997), or use a commercial software package.
Normal probability plot	Provides visual display of deviation from expected normality.	See EPA QA/G-9. Construct via EPA's DataQUEST software, or use a commerical software package.
Box and Whisker Plot	Provides visual display of potential "outliers" or extreme values.	See EPA QA/G-9. Construct via EPA's DataQUEST software, or use a commerical software package.
	STATISTICAL TESTS FOR NORMALITY	
Shapiro-Wilk W Test	Use for sample sizes of ≤50	See procedure in Appendix C. This test can be performed using EPA's DataQUEST software.
Filliben's Statistic	Use for sample sizes of >50	See EPA QA/G-9 guidance. This test can be performed using EPA's DataQUEST software.

How to Assess "Outliers"

A measurement that is very different from other values in the data set is sometimes referred to as an "outlier" [*4*]. EPA and ASTM caution that the term "outlier" should be used advisedly, since a common reaction to the presence of "outlying" values has been to "cleanse the data," thereby removing any "outliers" prior to further analysis [*4*]. In fact, such discrepant values can occur for many reasons, including: a catastrophic event such as a spill or process upset that impacts measurements at the sampling point, inconsistent sampling or analytical chemistry methodology that may result in laboratory contamination or other anomalies, errors in the transcription of data values or decimal points, and *true* but extreme hazardous constituent measurements.

While any one of these events can cause an apparent "outlier," it should be clear that the appropriate response to an outlier will be very different depending on the origin. Because high values due to contaminated media or waste are precisely what one may be trying to identify, one would clearly not want to eliminate such data in the guise of "screening for outliers." Furthermore, depending on the form of the underlying population, unusually high concentrations may be real but infrequent, as in lognormally distributed data. Again, one would not want to remove such data without adequate justification.

A *statistical outlier* is defined as a value originating from a different underlying population than the rest of the data set. If the value is not consistent with the distributional behavior of the remaining data and is "too far out in one of the tails" of the assumed underlying population, it may test out as a statistical outlier. However, defined as it is strictly in statistical terms, an outlier test may identify values as discrepant when no physical reason can be given for the aberrant behavior. For this reason, one should be especially cautious about indiscriminate testing for statistical outliers.

If an outlier is suspected, an initial and helpful step is to construct a quantile-quantile probability plot (*Q-Q* plot) of the data set. A *Q-Q* plot plot is designed to judge whether the sample data are consistent with an underlying normal population model. If the rest of the data follow normality, but the outlier comes from a distinctly different population with higher concentrations, this behavior will tend to show up on a probability plot as a lone value "out of line" with the remaining observations. If the data are lognormal instead, but the outlier is again from a distinct population, a probability plot on the logged observations should be constructed. Neither of these plots is a formal test; still they provide invaluable visual evidence as to whether the suspected outlier should really be considered as such. Methods for conducting outlier tests are described in Chapter 4 of QA/G-9, and statistical tests are available in the DataQUEST software (for example, Rosner's Test, and Walsh's Test).

Step 4—Select and Perform the Statistical Tests

This section provides guidance on how you can select the appropriate statistical test to make a decision about the waste under study. The decisions and conclusions derived from incorrectly used statistics can be expensive [*7*]. For example, incorrect use of a statistical test may result in the conclusion that a waste is nonhazardous, when in fact it is hazardous (see Table 4-1). See Chapter 9 of EPA's SW-846 RCRA sampling and analysis guidance manual for additional information [*8*].

Prior to selecting the statistical test, consider the following factors:

- the objectives of the study (identified in DQO Step 2),
- whether assumptions of the test are fulfilled,
- the nature of the underlying distribution,
- the decision rule and null hypothesis (identified in DQO Step 5),
- the relative performance of the candidate tests (for example, parametric tests generally are more efficient than their nonparametric counterparts), and
- the proportion of the data that are reported as non-detects (ND).

See EPA's QA/G-9 and SW-846 guidance documents to obtain information about the importance of these factors [*2,8*].

Data Transformations in Statistical Tests

Users of this guidance may encounter data sets that show significant evidence of non-normality. Due to the assumption of underlying normality in most parametric tests, a common statistical strategy when encountering this predicament is to search for a mathematical transformation that will lead to normally distributed data on the transformed scale. Unfortunately, because of the complexities associated with interpreting statistical results from data that have been transformed

to another scale and the common occurrence of lognormal patterns in environmental data, EPA generally recommends that the choice of scale be limited to either the original measurements (for normal data) or a log-transformed scale (for lognormal data). If neither of these scales results in approximate normality, it is typically easiest and wisest to switch to a non-parametric (or "distribution-free") version of the same test [7].

Treatment of Non-Detects in Confidence Intervals

According to ASTM if no more than 15% of the samples are non-detect (i.e., reported as below a detection limit), the results of parametric statistical tests will not be substantially affected if non-detects are replaced by half their detection limits [4]. This procedure is known as a substitution method. However, when more than 15% of the samples are non-detect, the treatment of non-detects is more crucial to the outcome of statistical procedures. Indeed, simple substitution methods (such as replacing the detection limit with one-half the detection limit) tend to perform poorly in statistical tests when the non-detect percentage is substantial [9]. If the percentage of non-detects is between 15 and 50%, we recommend use of Cohen's Adjustment (see QA/G-9). If the percentage of non-detects is greater than 50%, a nonparametric test should be used.

Table 4-2 provides guidance on selecting an approach for handling non-detects in statistical intervals. See EPA's QA/G-9 guidance document [2] for descriptions of appropriate nonparametric tests.

Step 5—Draw Conclusions and Report Results

The final step in the DQA Process is to draw conclusions from the data, determine if further sampling is required, and report the results. This step brings the planning, implementation, and assessment process "full circle" in that you attempt to resolve the problem and make the decision identified in DQO Steps 1 and 2. Additional guidance on this step can be found in Chapter 5 of EPA's QA/G-9 guidance [2].

In the DQO Process, you establish a "null hypothesis" and attempt to gather evidence via sampling that will allow you to reject that hypothesis (see Chapter 2). Otherwise, the null hypothesis must be accepted (Fig. 4-1). In most RCRA waste testing programs, the null hypothesis is that the "constituent concentration in the waste exceeds the standard." For us to reject that hypothesis (in other words, conclude that the constituent concentration is below the standard), the entire confidence interval must fall below the standard, and we must show that a sufficient number of samples were taken.

TABLE 4.2—Guidance for Handling Non-Detects in Statistical Intervals.

Percentage of Data Reported as "Nondetect"	Recommended Treatment of Dataset
<15%	Replace Non-Detects with DL/2
15 to 50%	Use Cohen's adjustment
>50%	Use a Nonparametric Test

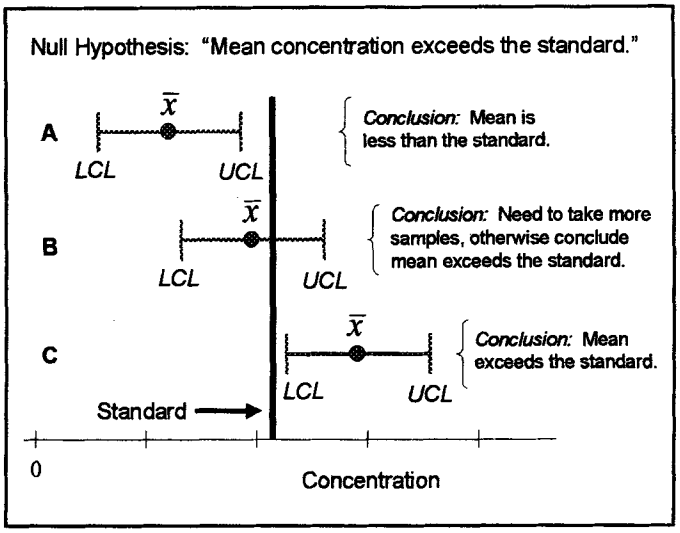


FIG. 4.1—Using confidence limits to compare waste concentrations to a fixed standard.

Figure 4-1 demonstrates how we can interpret the results relative to the null hypothesis: in the situation depicted at "A," the data have provided the evidence needed to reject the null hypothesis because the UCL is less than the standard. The decision can be made that the waste concentration is below the standard with sufficient confidence and without further analysis. In situation "B," we cannot reject the null hypothesis. However, because the interval "straddles" the standard, it is possible the true mean lies below the standard, and there is a possibility of a false negative error (i.e., to conclude the concentration is above the standard, when in fact it is not). One possible remedy to this situation is to obtain more data by collecting and analyzing more samples. In situation "C," the Type II decision error rate is satisfied and we must conclude that the mean concentration exceeds the standard.

One simple method for checking the performance of the statistical test is to use the variance obtained from the samples to retrospectively estimate the number of samples required. The variance can be input into the sample size equation used (DQO Step 7). If this theoretical sample size is less than or equal to the number of samples actually taken, then the test is sufficiently powerful. If the required number of samples is greater than the number actually collected, then additional samples would be required to satisfy the data user's performance criteria for the statistical test. See EPA's QA/G-9 guidance [2] for additional guidance on this topic.

SUMMARY

The assessment phase "closes the loop" on a data collection activity by insuring that the type, quality of the data collected meets the data quality objectives generated by the primary decision maker and the planning team during the Planning Phase. The graphical and statistical techniques that were re-

viewed in this chapter provide the tools needed by the investigators to accurately assess the data and successfully complete the data collection activity.

REFERENCES

[1] U.S. EPA, "The EPA Quality System," QA/G-0, Office of Research and Development, Office of Research and Development, Washington, DC, 1997.

[2] U.S. EPA, "Guidance for Data Quality Assessment—Practical Methods for Data Analysis," QA/G-9, EPA/600/R-96/084, Office of Research and Development, Washington, DC, 1998.

[3] Gilbert, R. O., *Statistical Methods for Environmental Pollution Monitoring*, John Wiley and Sons, New York, NY, 1987.

[4] ASTM, "Standard Guide for Data Assessment for Environmental Waste Management Activities," D 6233-98, 1998.

[5] U.S. EPA, Data Quality Assessment Statistical Toolbox (DataQUEST), User's Guide and Software, <http://es.epa.gov/ncerqa/qa/qa_docs.html#G-9d> EPA QA/G-9D, EPA/600/R-96/085, December, 1996.

[6] Ott, L., *An Introduction to Statistical Methods and Data Analysis*, 3rd ed., PWS-Kent Publishing Co., Boston, 1988.

[7] Singh, A. K., Singh, A., and Engelhardt, M., "The Lognormal Distribution in Environmental Applications," EPA/600/R-97/006, *Technology Support Center Issue*, U.S. EPA Office of Research and Development, 1997.

[8] U.S. EPA, "RCRA Waste Sampling Draft Technical Guidance SW-846 Chapter Nine—Planning, Implementation, and Assessment," EPA/530/R-99/015, *Solid Waste and Emergency Response*, Washington, DC, 1999.

[9] Gilliom, R. J. and Helsel, D. R., "Estimation of Distributional Parameters for Censored Trace Level Water Quality Data: Part I—Estimation Techniques," *Water Resources Research*, Vol. 22, No. 2, 1986, pp. 135–146.

Appendices

Appendix A: Confidence Intervals and Hypothesis Tests

A

Confidence Intervals and Hypothesis Tests

CONFIDENCE INTERVALS and formal hypothesis tests are two statistical methods that can be used for decisionmaking. A hypothesis test controls both the false positive decision error rate (α) and false negative decision error rate (β). A confidence interval controls only the probability of making a false positive decision error (α) (for example, concluding that a site is clean when it is truly dirty). However, the probability of making a false negative decision error (β) is fixed at 50% for confidence intervals (i.e., $\beta = 0.5$).

A confidence interval and a hypothesis test can be very similar. Consider the problem of determining whether the mean concentration (μ) of a site exceeds a cleanup standard (CS), where the contaminant is normally distributed. A confidence interval could be constructed for the mean, or a t-test could be used to test the statistical hypothesis:

$$H_0: \mu > CS \text{ vs. } H_a: \mu < CS$$

If the site manager's false negative decision error rate is 0.5 (i.e., $\beta = 0.5$), these methods are the same. Additionally, with a fixed α, the sample size of a confidence interval influences only the width of the interval (since $\beta = 0.5$). Similarly, the sample size of a t-test influences β and δ (where δ = upper value of the gray region minus the lower value of the gray region). However, by solving for the sample size using a t-test, one can substitute back into the sample size equation for a confidence interval and compute a width corresponding to this sample size. Then the results of the two methods will be identical.

Although the results of the hypothesis test and the confidence interval may be identical, the hypothesis test has the added advantage of a power curve. The power curve is defined as the probability of rejecting the null hypothesis. An ideal power curve is 1 for those values corresponding to the alternative hypothesis (all $\mu < CS$ in the example above) and 0 for those values corresponding to the null hypothesis (all $\mu > CS$ in the example above). The power curve is thus a way to tell how well a given test performs and can be used to compare two or more tests. Additionally, if the null hypothesis is not rejected, the power curve gives the decisionmaker some idea of whether or not the design could actually reject the null hypothesis for a given level (μ).

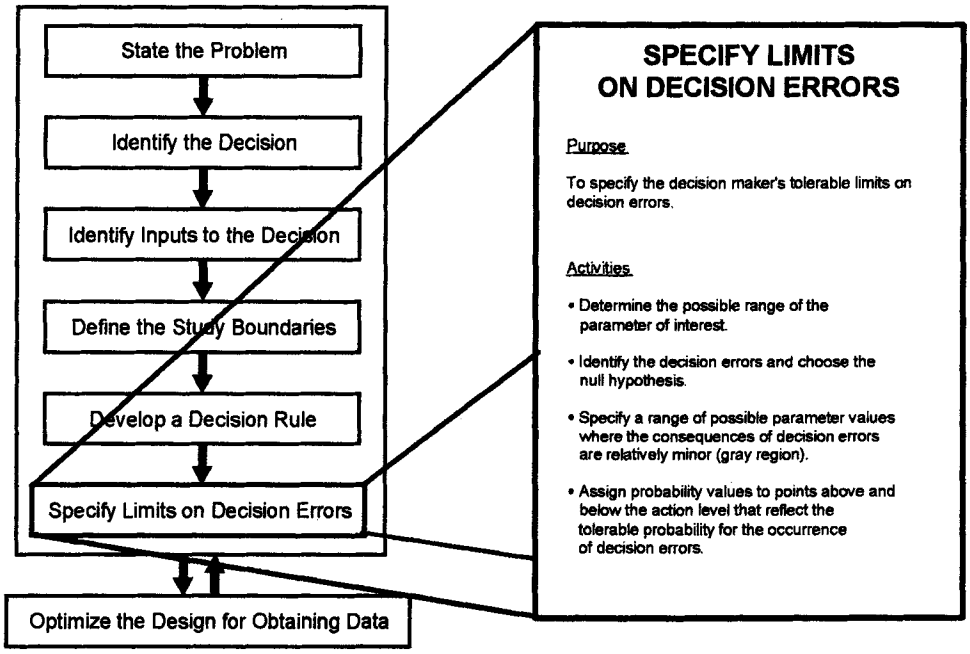
There is no corresponding idea of a power curve in terms of confidence intervals. To derive a power curve, one would need to translate the confidence interval into the corresponding test (i.e., a t-test) and then compute the power curve. Additionally, whereas a statistical test accounts directly for the false negative decision error, a confidence interval does not ($\beta = 0.5$). Finally, a confidence interval and a statistical test almost always are based on distributional assumptions, independence assumptions, etc. If these assumptions are violated, it may be easier to select an alternative test (for example, a non-parametric test) than it is to derive an alternative confidence interval. For these reasons, this document concentrates its discussion on hypothesis testing.

Appendix B: Q/A/G-4 Chapter 6: Specify Tolerable Limits on Decision Errors

Step 6: Specify Tolerable Limits on Decision Errors

B

THE DATA QUALITY OBJECTIVES PROCESS



PURPOSE

The purpose of this step is to specify the decisionmaker's tolerable limits on decision errors, which are used to establish performance goals for the data collection design.

EXPECTED OUTPUTS

- The decisionmaker's tolerable decision error rates based on a consideration of the consequences of making an incorrect decision.

BACKGROUND

Decisionmakers are interested in knowing the true state of some feature of the environment. Since data can only *estimate* this state, decisions that are based on measurement data could be in error (decision error). Most of the time the correct decision will be made; however, this chapter will focus on controlling the less likely possibility of making a decision error. The goal of the planning team is to develop a data collection design that reduces the chance of making a decision error to a tolerable level. This step of the DQO process will provide a mechanism for allowing the decisionmaker to define tolerable limits on the probability of making a decision error.

There are two reasons why the decisionmaker cannot know the true value of a population parameter (i.e., the true state of some feature of the environment):

(1) The population of interest almost always varies over time and space. Limited sampling will miss some features of this natural variation because it is usually impossible or impractical to measure every point of a population. *Sampling design error* occurs when the sampling design is unable to capture the complete extent of natural variability that exists in the true state of the environment.

(2) Analytical methods and instruments are never absolutely perfect, hence a measurement can only estimate the true

* Pages 32–36 from EPA's QA/5-4 (Ch. 2, Ref *1*).

value of an environmental sample. *Measurement error* refers to a combination of random and systematic errors that inevitably arise during the various steps of the measurement process (for example, sample collection, sample handling, sample preparation, sample analysis, data reduction, and data handling).

The combination of sampling design error and measurement error is called *total study error,* which may lead to a decision error. Since it is impossible to eliminate error in measurement data, basing decisions on measurement data will lead to the possibility of making a decision error.

The probability of decision errors can be controlled by adopting a scientific approach. In this approach, the data are used to select between one condition of the environment (the *null hypothesis,* H_0) and an alternative condition (the *alternative hypothesis,* H_a). The null hypothesis is treated like a baseline condition that is presumed to be true in the absence of strong evidence to the contrary. This feature provides a way to guard against making the decision error that the decisionmaker considers to have the more undesirable consequences.

A decision error occurs when the decisionmaker rejects the null hypothesis when it is true, or fails to reject the null hypothesis when it is false. These two types of decision errors are classified as *false positive* and *false negative* decision errors, respectively. They are described below.

False Positive Decision Error—A false positive decision error occurs when the null hypothesis (H_0) is rejected when it is true. Consider an example where the decisionmaker presumes that a certain waste is hazardous (i.e., the null hypothesis or baseline condition is "the waste is hazardous"). If the decisionmaker concludes that there is insufficient evidence to classify the waste as hazardous when it truly is hazardous, then the decisionmaker would make a false positive decision error. A statistician usually refers to the false positive error as a "Type I" error. The measure of the size of this error is called alpha (α), the level of significance, or the size of the critical region.

False Negative Decision Error—A false negative decision error occurs when the null hypothesis is *not* rejected when it is false. In the above waste example, the false negative decision error occurs when the decisionmaker concludes that the waste is hazardous when it truly is *not* hazardous. A statistician usually refers to a false negative error as a "Type II" error. The measure of the size of this error is called beta (β), and is also known as the complement of the *power* of a hypothesis test.

The definition of false positive and false negative decision errors depends on the viewpoint of the decision maker.[1] Consider the viewpoint where a person has been presumed to be "innocent until proven guilty" (i.e., H_0 is "innocent"; H_a is "guilty"). A false positive error would be convicting an innocent person; a false negative error would be not convicting the guilty person. From the viewpoint where a person is presumed to be "guilty until proven innocent" (i.e., H_0 is "guilty"; H_a is "innocent"), the errors are reversed. Here, the false positive error would be not convicting the guilty person, and the false negative error would be convicting the innocent person.

While the possibility of a decision error can never be totally eliminated, it can be controlled. To control the possibility of making decision errors, the planning team must control total study error. There are many ways to accomplish this, including collecting a large number of samples (to control sampling design error), analyzing individual samples several times, or using more precise laboratory methods (to control measurement error). Better sampling designs can also be developed to collect data that more accurately and efficiently represent the population of interest. Every study will use a slightly different method of controlling decision errors, depending on where the largest components of total study error exist in the data set and the ease of reducing those error components. Reducing the probability of making decision errors generally increases costs. In many cases controlling decision error within very small limits is unnecessary for making a decision that satisfies the decisionmaker's needs. For instance, if the consequences of decision errors are minor, a reasonable decision could be made based on relatively crude data (data with high total study error). On the other hand, if the consequences of decision errors are severe, the decisionmaker will want to control sampling design and measurement errors within very small limits.

To minimize unnecessary effort controlling decision errors, the planning team must determine whether reducing sampling design and measurement errors is necessary to meet the decisionmaker's needs. These needs are made explicit when the decision maker specifies probabilities of decision errors that are tolerable. Once these tolerable limits on decision errors are defined, then the effort necessary to analyze and reduce sampling design and measurement errors to satisfy these limits can be determined in Step 7: Optimize the Design for Obtaining Data. It may be necessary to iterate between these two steps before finding tolerable probabilities of decision errors that are feasible given resource constraints.

ACTIVITIES

Determine the possible range of the parameter of interest. Establish the possible range of the parameter of interest by estimating its likely upper and lower bounds. This will help focus the remaining activities of this step on only the relevant values of the parameter. For example, the range of the parameter shown in Figs. 6-1 and 6-2 at the end of this chapter is between 50 and 200 ppm. Historical and documented analytical data are of great help in establishing the potential parameter range.

Identify the decision errors and choose the null hypothesis. Define where each decision error occurs relative to the action level and establish which decision error should be defined as the null hypothesis (baseline condition). This process has four steps:

(1) *Define both types of decision errors and establish the true state of nature for each decision error.* Define both types of decision errors and determine which one occurs above and which one occurs below the action level. A decision error occurs when the data mislead the decisionmaker

[1] Note that these definitions are not the same as false positive or false negative instrument readings, where similar terms are commonly used by laboratory or field personnel to describe a fault in a single result; false positive and false negative *decision* errors are defined in the context of hypothesis testing, where the terms are defined with respect to the null hypothesis.

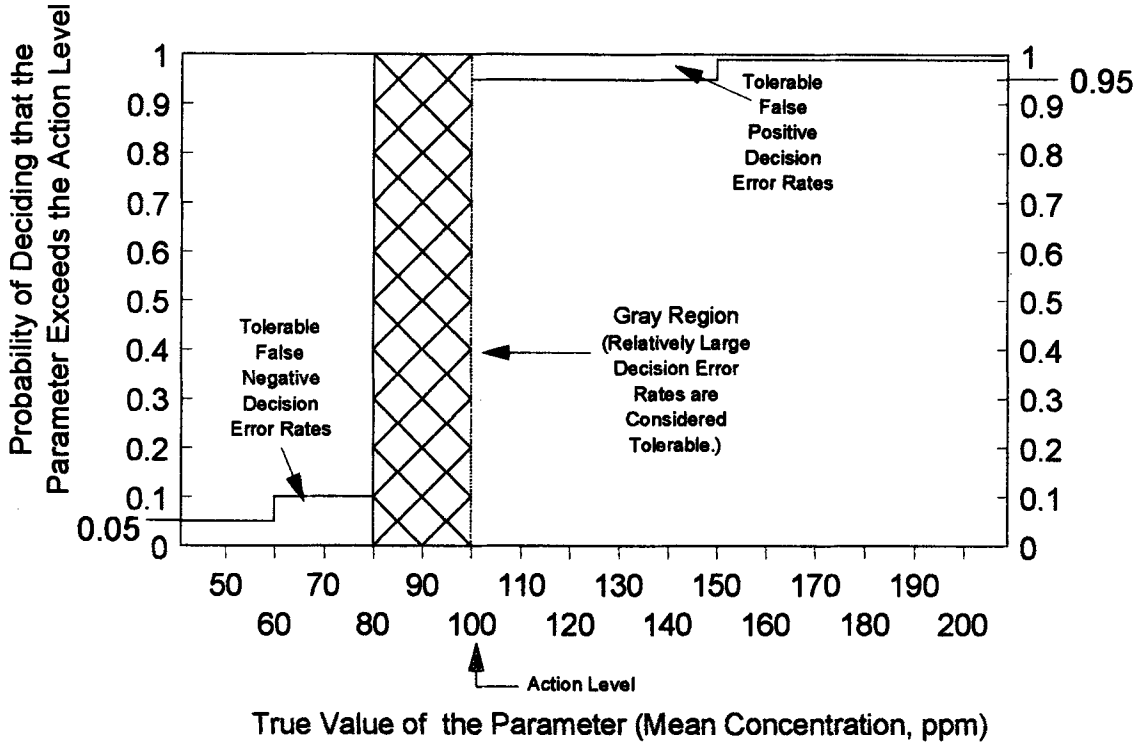


FIG. 6.1—An example of a decision performance goal diagram baseline condition: Parameter exceeds action level.

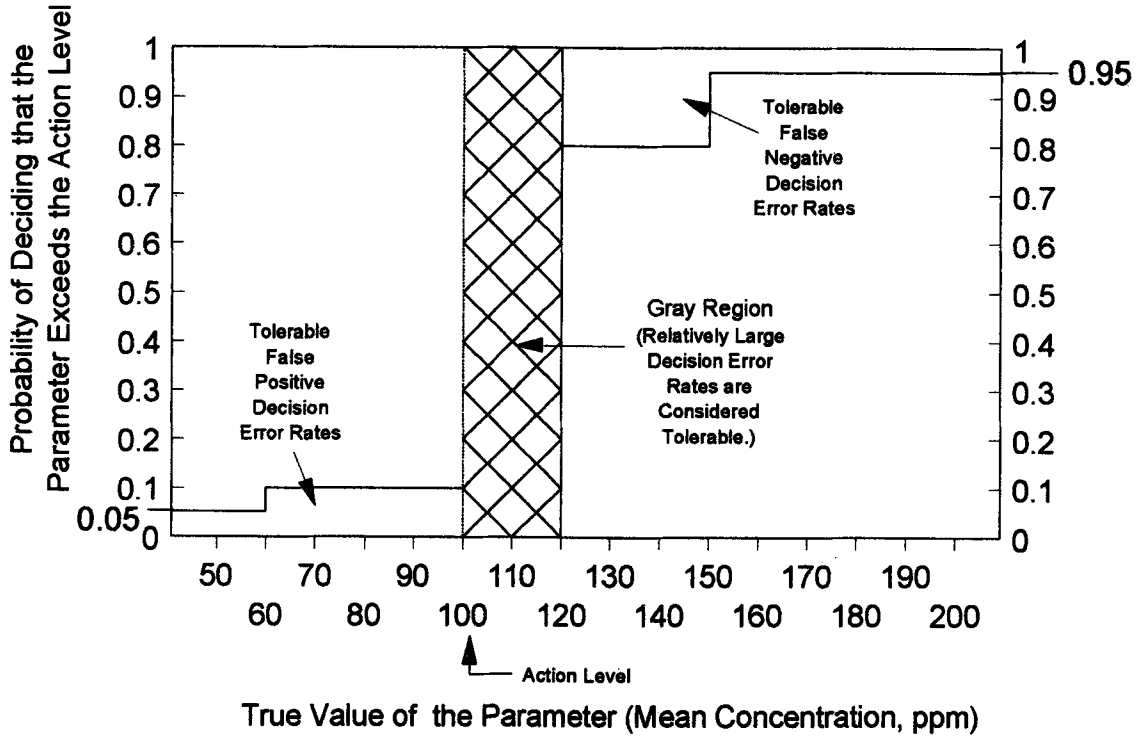


FIG. 6.2—An example of a decision performance goal diagram baseline condition: Parameter is less than action level.

into concluding that the parameter of interest is on one side of the action level when the true value of the parameter is on the other side of the action level. For example, consider a situation in which a study is being conducted to determine if mercury contamination is creating a health hazard and EPA wants to take action if more than 5% of a population of fish have mercury levels above a risk-based action level. In this case, a decision error would occur if the data lead the decisionmaker to conclude that 95% of the mercury levels found in the fish

population were below the action level (i.e., the parameter is the "95th percentile" of mercury levels in the fish population) when the true 95th percentile of mercury levels in the fish population was above the action level (which means that more than 5% of the fish population contain mercury levels greater than the action level). The other decision error for this example would be that the data lead the decisionmaker to conclude that the 95th percentile of mercury levels in the fish population is greater than the action level when the true 95th percentile is less than the action level. The "true state of nature" is the actual condition or feature of the environment that exists, but is unknown to the decisionmaker. Each decision error consists of two parts, the true state of nature and the conclusion that the decisionmaker draws. Using the example above, the true state of nature for the first decision error is that the 95th percentile of mercury levels in the fish population is above the action level.

(2) *Specify and evaluate the potential consequences of each decision error.* Specify the likely consequences of making each decision error and evaluate their potential severity in terms of economic and social costs, human health and ecological effects, political and legal ramifications, and so on. Consider the alternative actions that would be taken under each decision error scenario, as well as secondary effects of those actions. For example, in determining whether or not 95% of a fish population contain mercury levels above a risk-based action level, there may be a variety of potential consequences of committing a decision error. In the first decision error described above, where the decisionmaker concludes that the 95th percentile is below when the true 95th percentile was above the action level, the decisionmaker may decide to continue to allow fishing in the waters and not undertake any cleanup activity. The resulting consequences might include human health and ecological effects from consumption of contaminated fish by humans and other animals, economic and social costs of health care and family disruption, and damaged credibility of EPA when (and if) the decision error is detected. If the other type of decision error is committed, where the decisionmaker decides that the 95th percentile exceeds the action level when the true 95th percentile is below the action level, the decisionmaker might ban all fishing in the local waters and initiate cleanup activities. The consequences might include economic and social costs of lost revenues and job displacement in the fishing industry, damaged credibility for EPA when the cleanup activities expose the nature of the decision error, and the threat of lawsuits by fishing interests.

Evaluate the severity of potential consequences of decision errors at different points within the domains of each type of decision error, since the severity of consequences may change as the parameter moves further away from the action level. Consider whether or not the consequences change abruptly at some value, such as a threshold health effect level; the decisionmaker may want to change the tolerable limit on the decision error at such a point.

(3) *Establish which decision error has more severe consequences near the action level.* Based on the evaluation of potential consequences of decision errors, the decisionmaker should determine which decision error causes greater concern when the true parameter value is near the action level. It is important to focus on the region near the action level because this is where the true parameter value is most likely to be when a decision error is made (in other words, when the true parameter is far above or far below the action level, the data are much more likely to indicate the correct decision). This determination typically involves value judgments about the relative severity of different types of consequences within the context of the problem. In the fish contamination problem above, the decisionmaker would weigh the potential health consequences from allowing people to consume contaminated fish versus the economic and social disruption from banning all fishing in the community. In this case, the decisionmaker might carefully consider how uncertain or conservative the risk-based action level is.

(4) *Define the null hypothesis (baseline condition) and the alternative hypothesis and assign the terms "false positive" and "false negative" to the appropriate decision error.* In problems that concern regulatory compliance, human health, or ecological risk, the decision error that has the most adverse potential consequences should be defined as the null hypothesis (baseline condition).[2] In statistical hypothesis testing, the data must conclusively demonstrate that the null hypothesis is false. That is, the data must provide enough information to authoritatively reject the null hypothesis (disprove the baseline condition) in favor of the alternative. Therefore, by setting the null hypothesis equal to the true state of nature that exists when the more severe decision error occurs, the decisionmaker guards against making the more severe decision error by placing the burden of proof on demonstrating that the most adverse consequences will *not* be likely to occur.

It should be noted that the null and alternative hypotheses have been predetermined in many regulations. If not, the planning team should define the null hypothesis (baseline condition) to correspond to the true state of nature for the more severe decision error and define the alternative hypothesis to correspond to the true state of nature for the less severe decision error.

Using the definitions of null and alternative hypotheses, assign the term "false positive" to the decision error in which the decisionmaker rejects the null hypothesis when it is true, which corresponds to the decision error with the more severe consequences identified in task (3). Assign the term "false negative" to the decision error in which the decisionmaker fails to reject the null hypothesis when it is false, which corresponds to the decision error with the less severe consequences identified in task (3).

[2] Note that this differs somewhat from the conventional use of hypothesis testing in the context of planned experiments. There, the alternative hypothesis usually corresponds to what the experimenter hopes to prove, and the null hypothesis usually corresponds to some baseline condition that represents an "opposite" assumption. For instance, the experimenter may wish to prove that a new water treatment method works better than an existing accepted method. The experimenter might formulate the null hypothesis to correspond to "the new method performs no better than the accepted method," and the alternative hypothesis as "the new method performs better than the accepted method." The burden of proof would then be on the experimental data to show that the new method performs better than the accepted method, and that this result is not due to chance

Specify a range of possible parameter values where the consequences of decision errors are relatively minor (gray region). The gray region is a range of possible parameter values where the consequences of a false negative decision error are relatively minor. The gray region is bounded on one side by the action level and on the other side by that parameter value where the consequences of making a false negative decision error begin to be significant. Establish this boundary by evaluating the consequences of not rejecting the null hypothesis when it is false. The edge of the gray region should be placed where these consequences are severe enough to set a limit on the magnitude of this false negative decision error. Thus, the gray region is the area between this parameter value and the action level.

It is necessary to specify a gray region because variability in the population and unavoidable imprecision in the measurement system combine to produce variability in the data such that a decision may be "too close to call" when the true parameter value is very near the action level. Thus, the gray region (or "area of uncertainty") establishes the minimum distance from the action level where the decisionmaker would like to begin to control false negative decision errors. In statistics, the width of this interval is called the "minimum detectable difference" and is often expressed as the Greek letter delta (Δ). The width of the gray region is an essential part of the calculations for determining the number of samples needed to satisfy the DQOs, and represents one important aspect of the decision maker's concern for decision errors. A more narrow gray region implies a desire to detect conclusively the condition when the true parameter value is close to the action level ("close" relative to the variability in the data). When the true value of the parameter falls within the gray region, the decisionmaker may face a high probability of making a false negative decision error, since the data may not provide conclusive evidence for rejecting the null hypothesis, even though it is actually false (i.e., the data may be too variable to allow the decisionmaker to recognize that the presumed baseline condition is, in fact, *not* true).

From a practical standpoint, the gray region is an area where it will not be feasible or reasonable to control the false negative decision error rate to low levels because of high costs. Given the resources that would be required to reliably detect small differences between the action level and the true parameter value, the decisionmaker must balance the resources spent on data collection with the expected consequences of making that decision error. For example, when testing whether a parameter (such as the mean concentration) exceeds the action level, if the *true* parameter is near the action level (relative to the expected variability of the data), then the imperfect data will tend to be clustered around the action level, with some values above the action level and some below. In this situation, the likelihood of committing a false negative decision error will be large. To determine with confidence whether the true value of the parameter is above or below the action level, the decisionmaker would need to collect a large amount of data, increase the precision of the measurements, or both. If taken to an extreme, the cost of collecting data can exceed the cost of making a decision error, especially where the consequences of the decision error may be relatively minor. Therefore, the decisionmaker should establish the gray region, or the region where it is not critical to control the false negative decision error, by balancing the resources needed to "make a close call" versus the consequences of making that decision error.

Assign probability limits to points above and below the gray region that reflect the tolerable probability for the occurrence of decision errors. Assign probability values to points above and below the gray region that reflect the decisionmaker's tolerable limits for making an incorrect decision. Select a possible value of the parameter; then choose a probability limit based on an evaluation of the seriousness of the potential consequences of making the decision error if the true parameter value is located at that point. At a minimum, the decisionmaker should specify a false positive decision error limit at the action level, and a false negative decision error limit at the other end of the gray region. For many situations, the decision maker may wish to specify additional probability limits at other possible parameter values. For example, consider a hypothetical toxic substance that has a regulatory action level of 10 ppm, and which produces threshold effects in humans exposed to mean concentrations above 100 ppm. In this situation, the decisionmaker may wish to specify more stringent probability limits at that threshold concentration of 100 ppm than those specified at 10 ppm. The tolerable decision error limits should decrease further away from the action level as the consequences of decision error become more severe.

Given the potentially high cost of controlling sampling design error and measurement error for environmental data, Agency decision making is rarely supported by decision error limits more stringent than 0.01 (1%) for both the false positive and false negative decision errors. This guidance recommends using 0.01 as the starting point for setting decision error rates. The most frequent reasons for setting limits greater (i.e., less stringent) than 0.01 are that the consequences of the decision errors may not be severe enough to warrant setting decision error rates that are this extreme. The value of 0.01 should *not* be considered a prescriptive value for setting decision error rates, nor should it be considered as the policy of EPA to encourage the use of any particular decision error

TABLE 6.1—Decision Error Limits Table Corresponding to Figure 6-1. (Action Level = 100 ppm).

True Concentration	Correct Decision	Type of Error	Tolerable Probability of Incorrect Decision
<60 ppm	Not exceed	F(−)	5%
60 to 80	Not exceed	F(−)	10%
80 to 100	Not exceed	F(−)	gray region
100 to 150	Does exceed	F(+)	5%
>150	Does exceed	F(+)	1%

TABLE 6.2—Decision Error Limits Table Corresponding to Figure 6-2. (Action Level = 100 ppm).

True Concentration	Correct Decision	Type of Error	Tolerable Probability of Incorrect Decision
<60 ppm	Not exceed	F(+)	5%
60 to 100	Not exceed	F(+)	10%
100 to 120	Does exceed	F(−)	gray region
120 to 150	Does exceed	F(−)	20%
>150	Does exceed	F(−)	5%

rate. Rather, it should be viewed as a starting point from which to develop limits on decision errors that are applicable for each study. If the decisionmaker chooses to relax the decision error rates from 0.01 for false positive or false negative decision errors, the planning team should document the reasoning behind setting the less stringent decision error rate and the potential impacts on cost, resource expenditure, human health, and ecological conditions.

The combined information from the activities section of this chapter can be graphed onto a "Decision Performance Goal Diagram" or charted in a "Decision Error Limits Table" (see Figs. 6-1 and 6-2 and Tables 6-1 and 6-2). Both are useful tools for visualizing and evaluating all of the outputs from this step. Figure 6-1 and Table 6-1 illustrate the case where the null hypothesis (baseline condition) is that the parameter of interest exceeds the action level (e.g., the waste is hazardous). Figure 6-2 and Table 6-2 illustrate the case where the null hypothesis (baseline condition) is that the parameter is less than the action level (e.g., the waste is not hazardous).

Appendix C: Waste Pile Example

C

Waste Pile Example

INTRODUCTION

IN THIS EXAMPLE five case studies with varying waste pile characteristics and alternate sampling designs are presented through the planning (DQO process), implementation, and assessment phases. For purposes of these case studies, the stakeholders have different prior knowledge for each case. However, for consistency and to clearly present the development of the alternate sampling designs, each waste pile has the same characteristics, as described in the following paragraph.

The waste pile in these examples consists of material that has been generated from a metals recovery process. The dimensions of the waste pile are approximately 100 by 100 ft (38.48 m) with a maximum height of 10 ft (3.048 m); however, more material was deposited in the front corner of the pile (see Fig. 1—Topographic Base Map). The material in the pile was generated from the same source and contaminated with lead. It is also known that no containerized waste has been disposed of in the waste pile. The waste pile is now a Solid Waste Management Unit (SWMU) under investigation as part of a RCRA Facility Investigation (RFI). Specific guidance is provided in ASTM's Standard Guide for Sampling Waste Piles, D 6009. Note that the sampling design for each case is denoted in the text of the example for clarification purposes; the appropriate sampling design is actually selected at Step Seven in the DQO process.

For *Case 1 (authoritative),* the stakeholders expect the lead concentration to be extremely elevated due to process knowledge (perhaps several times the Toxicity Characteristic (TC) Rule regulatory level of 5.0 mg/L), and it is likely that the TCLP results will designate the material as hazardous. If the lead concentration in the TCLP greatly exceeds the TC Rule regulatory level, then a statistical evaluation of the data would not be necessary. Thus, a complex sampling design would probably not be warranted in this case. In this case, the stakeholders have set a limit of $2,000 for the analytical costs of the study.

For *Case 2 (simple random),* preliminary data indicate that the mean lead concentration is near the regulatory limit. The stakeholders expect the pile to be relatively homogeneous; therefore, information on the distribution of lead is not important. (The entire waste pile will be considered the "remediation unit" in this case. (See Identifying Inputs to Decision section).) Although the degree of stratification is not known (either over space or by component), it is not expected to be significant because the recovery process that generated the waste was reportedly constant over the time period that the pile was generated and the particle sizes of the material in the pile could be considered homogeneous for the purposes of this investigation (also known as practically homogeneous). The stakeholders have decided that a limit of $8,000 for the analytical costs of the study will be set in this case.

For *Case 3 (systematic grid),* a minimal amount of data exists on the material in the waste pile so that no assumptions concerning probable contaminant concentrations can be made initially. Information regarding contaminant distribution across the waste pile is a primary objective of the study. The stakeholders have decided that a limit of $5,000 for the analytical costs of the study will be set in this case.

For *Case 4 (systematic grid with compositing),* a minimal amount of data exists on the material in the waste pile so that no assumptions concerning probable contaminant concentrations can be made initially. Specific information regarding distribution of contamination across the waste pile is not an objective of the study. The degree of stratification is not known, but it is not expected to be significant. The stakeholders have set a limit of $2,000 for the analytical costs of the study in this case.

For *Case 5 (stratified with systematic grid),* it is discovered that a recent process change was incorporated in the metals recovery process which significantly increased the lead concentration in the waste. Information exists suggesting that approximately the front 20% of the pile (note slightly greater elevation) was generated by the new process, while the material generated by the previous process is located in the remainder of the pile. Although two areas of different concentrations, or strata, exist within the waste pile, the two individual strata are internally homogeneous. One decision will be made on the entire waste pile. The stakeholders have decided on an analytical cost limit of $5,000.

PLANNING PHASE

The DQO process and sampling design optimization process are outlined in the Planning Step section of this manual. The following information pertains to all five cases described in the introduction unless otherwise stated. Figures illustrating the location of the samples for each case are included at the end of the example.

Data Quality Objectives (DQO) Process

Step One: Stating the Problem

The waste pile contains material that may be considered hazardous due to elevated lead content. Therefore, in each case the

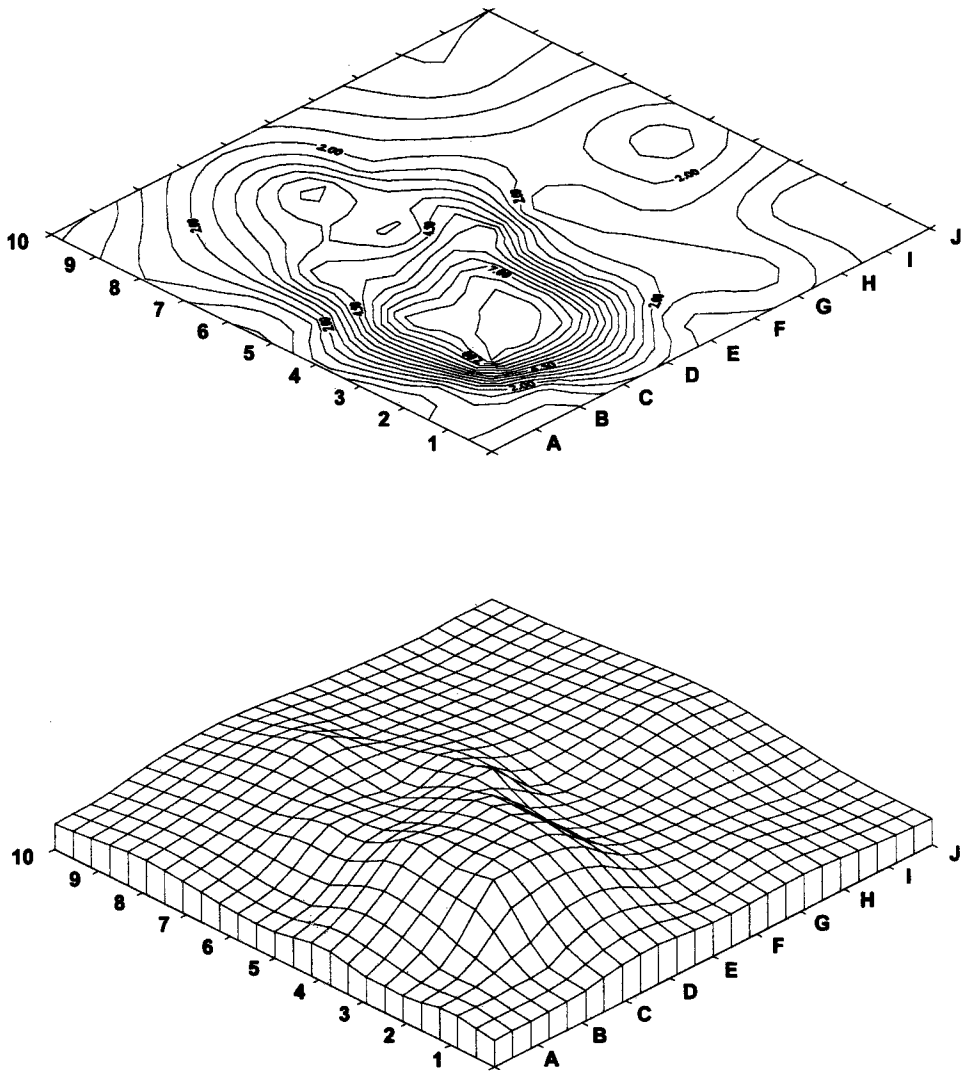


FIG. 1—Topographic base map.

company needs to determine if the material should be disposed of in a hazardous waste landfill under Subtitle C of RCRA (@ $500 per ton) versus a Subtitle D landfill (@ $50 per ton). The stakeholders in this study are the company that generated the waste (and will be conducting the sampling and analysis), the appropriate regulatory agencies, and in some cases representatives from local communities. The company will be required to develop a sampling design that meets the objectives of the study and satisfies all pertinent regulatory requirements.

Step Two: Identifying Possible Decisions

The principal study question is: Is the material in the waste pile a RCRA hazardous waste (per 40 CFR 261.24)? The potential alternate actions are: (a) the material must be managed under Subtitle C of RCRA as hazardous waste or (b) the material may be disposed of in a permitted Subtitle D Municipal Solid Waste Landfill (MSWLF).

Step Three: Identifying Inputs to the Decision

- The decision on whether the material is hazardous or not will depend on the results of the Toxicity Characteristic Leaching Procedure (TCLP) test on the samples collected. The regulatory level for lead under the TC Rule is 5.0 mg/L. If the sample results exceed this value, the material will be considered hazardous. Totals results may be used to determine if the lead concentration is elevated enough—at least 20 times the regulatory level—to warrant completion of the TCLP test. (See EPA Method 1311, Section 1.1.) Note that the totals results may also be necessary to provide information for a subsequent risk assessment to determine the need to characterize soil and/or groundwater in areas adjacent to the waste pile if it is determined to be non-hazardous, and, in the case when the material is determined to be hazardous, for characterization required for off-site disposal by a permitted Treatment, Storage, Disposal Facility (TSDF). For purposes of this example, only Cases 1 and 2 will include totals results; however, they may be included during the planning step based on the objectives of the study.
- In each case, the decision will be based on the entire waste pile; in other words, there will not be smaller "remediation units" within the pile where a Subtitle C versus D decision will be made. Either the entire pile is hazardous, or the entire pile is not. In certain situations, however, it may prove advantageous to employ different scales of decisionmak-

ing, such as with a two-part decision rule. An example of a two-part decision rule that could be used in this situation would be to (1) compare the mean of the pile to a regulatory level and (2) make a decision on smaller remediation units of the pile if they contained lead greater than three standard deviations above the regulatory level.

- For Cases 1–4 the material in the waste pile was generated by the same process, while two different processes were used in Case 5.
- Lead is the contaminant of concern, although the exact distribution across the pile is unknown.
- Access to the pile is not limited, and traditional sampling equipment is expected to be adequate.
- The analytical methods for lead (SW-846 Method 6010B for total lead and SW-846 Method 1311 for the TCLP) should be able to meet the required detection limits as the sample matrix is not expected to be difficult from a sample preparation or analysis standpoint. The totals results, if being used for a subsequent risk assessment, must meet the quantitation limits required for the assessment. Also, an acceptable approach for addressing non-detects must be decided upon prior to the investigation (see Data Quality Assessment section in the Manual).
- The particle size of the material in the waste pile (approximately 0.05 cm) could be considered homogeneous for purposes of this investigation.
- "Real-time" field analytical techniques and innovative approaches (such as XRF, field atomic adsorption or gas chromatography, immunoassay-based test kits, direct push technologies, etc.) could be used to improve decision-making in the field. These techniques would be incorporated into the DQO process to provide flexibility in the field based on the information being generated on-site. They would also assist the investigators in determining the presence and nature of contaminant heterogeneity.

Step Four: Defining Boundaries

The waste pile will be sampled using an appropriate design and analyzed for lead (totals and TCLP). The spatial boundary of the waste pile has been defined by the obvious elevation above the surrounding terrain, the discoloration associated with the material, and the practically homogeneous particle size of the material. The samples will be collected from the surface to a 1-ft (0.30 m) depth, although in every case locations should be sampled to the base of the waste pile to obtain information about potential vertical stratification (Case 1 illustrates this approach). Samples will be collected within a reasonable time frame; however, a temporal boundary for an inorganic contaminant such as lead is generally not a concern.

Step Five: Developing Decision Rules

The decision rule will differ depending on the case under consideration.

With an authoritative design (Case 1), the decision rule will be:

If the average lead concentration for the data set, based on a judgmental approach, greatly exceeds the regulatory level of 5.0 mg/L using the TCLP, then the material in the waste pile will be considered hazardous, and it will be managed under Subtitle C of RCRA. If the average concentration is near or below the regulatory level, a more complex sampling design will be developed. Since an authoritative design is being considered for this investigation, a statistical test would not be applicable and, in fact, unnecessary if the results significantly exceed the regulatory level.

With a probabilistic design (Cases 2–5), the decision rule will be:

If the 90% (one-tailed) upper confidence level (UCL) of the mean concentration is equal to or exceeds the regulatory level of 5.0 mg/L using the TCLP, then the material in the waste pile will be considered hazardous, and it will be managed under Subtitle C of RCRA. If the 90% UCL is below the regulatory level, the material will not be considered hazardous and will be managed under Subtitle D for Municipal Solid Waste Landfills. The use of the term "mean" assumes a normal distribution of the data, an assumption that must be checked. A lognormal distribution could also be evaluated, but the UCL would be computed differently. (See Data Quality Assessment section of this example.)

Step Six: Specifying Limits on Decision Errors

The sampling design error and measurement error will be minimized by using a well-prepared Project Plan (QAPP). The acceptable decision error is decidedly smaller for a Type I error (the material is actually hazardous when the study indicates it is not); therefore, the stakeholders have decided that any outcome where the lead concentration is near or below the regulatory level will result in the need for further investigation using a more complex sampling design. However, because the risk associated with a Type II error (the material is determined to be hazardous when it is not) from an environmental or human health standpoint is less, a result that is significantly above the regulatory level will result in a decision that is protective. Note that the decision error is more important when the mean of the data set is near the regulatory level of 5.0 mg/L of lead.

For a study implementing a probabilistic design, limits on decision errors will be set as follows:

In the case of making a hazardous waste determination, we are comparing the 90% UCL of the mean concentration of the TCLP results for the sample to the Toxicity Characteristic (TC) Rule regulatory level of 5 mg/L. SW-846 suggests that the decision be based on a 90% one-tailed test [*1*]. The Type I error rate is set at 0.10 (10%). That is the probability of rejecting the null hypothesis when it is actually true. See Appendix B for additional information on hypothesis testing.

Step Seven: Optimizing Data Collection Design

Initial Design Selection

The initial design selection for the Case 1 study is:

Since available information strongly suggests that the lead concentration in the waste pile is elevated, an authoritative design is chosen initially for this case. However, if the sample results reveal values close to the regulatory limits, the sample design will need to be reconsidered in light of the new data. Two types of authoritative designs are to be considered: *biased,* where the investigation targets worst case conditions,

or *judgmental,* where the investigator uses professional judgment and site information/observations to collect samples that reflect average conditions on the site. The determination of average conditions would be appropriate in this case because the facility has conceded that the lead concentrations are elevated. Note that worst case conditions would be difficult to determine in a waste unit such as this but would be appropriate when process or site knowledge can be used to identify areas of highest contamination. Therefore, the specific sampling locations and the number of samples will be determined by the investigators in the field. As a general rule, at least four to six samples should be collected. This number allows for one sample to be taken in each of four quadrants and provides a minimum degree of coverage for the pile.

The initial design selection for the Case 2 study is:

The stakeholders expect the lead concentration to be near the regulatory limit; thus, a probabilistic approach will be chosen to validate data results. Simple random, stratified, and systematic (grid-based) designs provide information on the mean concentration of lead. Since the existence of strata is not expected (although could be discovered during the investigation), the stratified design is at this time eliminated from consideration. Information on spatial distribution of lead in the pile is not a primary objective of this study, although it would confirm the investigators, assumptions concerning a non-stratified contaminant distribution. A simple random design is the simplest of the probabilistic sampling methods, but it is not ideally suited for providing information on spatial distribution. The systematic design, both without compositing or with compositing, provides some spatial distribution information and is typically easy to implement. Compositing may increase precision and reduce decision errors by reducing the variability of the estimated mean. The design team will further consider all three alternatives in the Practical Evaluation step of the optimization process.

The initial design selection for the Case 3 and Case 4 study is:

The stakeholders do not have enough information to predict the lead concentration; thus, a probabilistic approach will be chosen to validate data results. Simple random, stratified random, and systematic (grid-based) designs will provide information on the mean concentration of lead. Since the existence of distinct strata is not expected, the stratified design is at this time eliminated from consideration. The design team will further consider the remaining alternatives in the Practical Evaluation step.

The initial design selection for the Case 5 study is:

Due to the existence of a process change that affected the characteristics of the waste, and the expected stratification of the waste pile, a stratified sample design is chosen.

Practical Evaluation

The practical considerations that should be reviewed for each alternative include site access and conditions, equipment selection/use, experience needed, special analytical needs, and scheduling. The remaining alternatives do not have significant practical considerations that would limit their potential use for this study. However, the systematic design may result in sampling locations that are easier to survey and locate in the field, and it would provide better spatial coverage, if needed. Problems with access to all sampling locations, difficult matrices (resistant to penetration by an auger, for example, or containing large pieces of debris or material), and sampling into native material below the pile should all be considered during the development of the Quality Assurance Sampling Plan. A standard operating procedures (SOP) manual for conducting the field sampling will influence the collection of a representative sample.

Estimating the Number of Samples Required for the Study

The designs are evaluated for the number of samples that will be required:

Step One: Determination of the Number of Samples

Based on the use of an *authoritative* approach (Case 1):

Samples will be collected within each quadrant of the waste pile and at the center of the pile. The boring at the center will be advanced to the base of the pile at two-foot intervals to provide information on the vertical concentration profile. The TCLP will be conducted on the top one-foot interval of the boring.

Based on the use of a *probabilistic* approach (Cases 2 to 5):

Simple random design (Case 2):

An acceptable margin of error (Δ) and acceptable probability of exceeding that error (α) must be set. Then the appropriate number of samples to collect may be calculated by [*1*]:

$$n = \frac{(t_{1-\alpha} + t_{1-\beta})^2 s^2}{\Delta^2}$$

where:

n = number of samples to collect,
$t_{1-\alpha}$ = percentile value for the Student t distribution for $n - 1$ degrees of, freedom where α is the probability of making a Type I error,
$t_{1-\beta}$ = percentile value for the Student t distribution for $n - 1$ degrees of, freedom where β is the probability of making a Type II error,
s^2 = estimate of the variance (for individual samples), and
$\Delta = RT - \bar{x}$ (RT is the regulatory threshold, $\bar{x}$ is the estimated mean).

Note that values of the Student t distribution may be obtained from Table 3 in Appendix D. Because the Type II error rate (the chance of deciding the waste is hazardous when it is not) is set at 50% (i.e., $\beta = 0.50$), the associated t value becomes zero and the $t_{(1-\beta)}$ term drops from the equation. The discussion in Appendix B addresses the advantages obtained by setting the Type II error rate at a value less than 0.50. The resulting equation is used to calculate the number of samples:

$$n = \frac{t_{1-\alpha}^2 \cdot s^2}{\Delta^2}$$

In a preliminary pilot study, five samples were collected at random. Results for TCLP were 5.8, 10.5, 4.9, 2.1, and 5.4 mg/L. The mean and standard deviation were estimated to be 5.74 and 3.03, respectively. Note that the regulatory level for

lead is 5.0 mg/L, and α was set at 0.10. Thus, the acceptable margin of error is calculated as $\Delta = RT - \bar{x} = -0.74$. Using this sample size equation and the t value with $n - 1 = 4$ degrees of freedom,

$$n = \frac{1.533^2 \cdot 3.03^2}{(5 - 5.74)^2} = 40$$

An iteration of the equation is then performed to stabilize the result using $n = 40$ and a t value for $n - 1 = 39$ degrees of freedom. The final sample size is calculated as:

$$n = \frac{1.303^2 \cdot 3.03^2}{(5 - 5.74)^2} = 29$$

Systematic grid design (Case 3):

The minimum number of samples for a systematic grid sampling design may be estimated using the same approach described above for the Simple Random design. Such an approach should provide acceptable results if no strong cyclical patterns, periodicities, or significant spatial correlations exist between sample locations [*1*].

In Case 3, a preliminary pilot study was utilized to calculate the number of samples using the method described above for Case 2. With five samples, the estimated mean and standard deviation were 4.42 and 1.37, respectively. The "*n*" necessary to achieve a 10% probability of exceeding the absolute margin of error was calculated (after several iterations to stabilize the result) to be 11 samples.

Systematic grid design with compositing (Case 4):

Compositing samples, when appropriate, reduces decision errors and increases the precision of the estimated sample mean by reducing variability associated with that mean. With the assumption that the analytical variation is negligible compared to the spatial variation, the sample variance with compositing is equal to the variance without compositing divided by the number of aliquots (k). The necessary number of samples to achieve a desired α is inversely proportional to the number of aliquots. The number of aliquots (k) refers to the number of individual grab samples used to form each composite. For a simple random design, the number of samples may be calculated by:

$$n = \frac{t_{1-\alpha}^2 \cdot (s^2/k)}{\Delta^2}$$

Using the same pilot study data for this case as used for Case 3 and choosing k to be 5, the number of samples necessary with compositing would be reduced to 4. In summary, four composite samples will be collected and each will be comprised of five aliquots that are distributed in four quadrants around a center point, with the last aliquot for each sample coming from the center point.

Stratified systematic design (Case 5):

It is known that the waste pile consists of two different types of internally homogeneous material, so the total waste pile is divided into $L = 2$ nonoverlapping strata. The number of population units in each of the two strata is denoted by N_1 and N_2, and the number of necessary samples in h^{th} stratum may be calculated by $N_h = N \cdot W_h$, where W_h represents the weight or volume of material in the h^{th} stratum. Since it is known that approximately 20% of the waste pile was generated by a new process, W_1 will be set equal to 0.2 and W_2 will be 0.8. Preliminary data was collected from the pile. Three samples were collected from Strata 1, and five samples were collected from Strata 2. The mean and standard deviation for Strata 1 was calculated to be 9.9 and 0.7, respectively. For Strata 2, the mean and standard deviation were 3.5 and 0.7, respectively. The optimum number of samples may be determined using proportional allocation by [*1*]:

$$n = \frac{(t_{1-\alpha,df} + t_{1-\beta,df})^2}{\Delta^2} \cdot \sum_{h=1}^{L} W_h \cdot s_h^2$$

where

$t_{1-\alpha}$ = percentile value for the Student t distribution for $n - 1$ degrees of freedom where α is the probability of making a Type I error,

$t_{1-\beta}$ = percentile value for the Student t distribution for $n - 1$ degrees of freedom where β is the probability of making a Type II error,

$\Delta = RT - \bar{x}$ (RT is the regulatory threshold, $\bar{x}$ is the estimated mean),

s^2 = estimate of the variance (for individual samples),

W_h = weight or volume of material in the h^{th} stratum,

df = the degrees of freedom connected with each t-quantile.

The value of df may be calculated by:

$$df = \left(\sum_{h=1}^{L} W_h \cdot s_h^2 \right)^2 \Big/ \left(\sum_{h=1}^{L} \frac{W_h^2 \cdot s_h^4}{(n \cdot W_h) - 1} \right)$$

Using the preliminary pilot data results and the weighting values for the two strata, df is calculated to be 2, and the corresponding number of samples is 30. The equations must be solved iteratively, so the same calculations are repeated using $n = 30$. After several iterations, the total number of samples is set at 17. Using proportional allocation with $n = 17$ samples, $0.2 \cdot 17 = 3$ samples should be taken from Stratum 1, while $0.8 \cdot 17 = 14$ samples should be collected from Stratum 2. The pilot study data may be used as a portion of the final data set. Thus, no additional samples need to be collected from Stratum 1, and nine additional samples are needed from Stratum 2.

The mean of the data set will be evaluated using the approach in SW-846, Chapter Nine, where the upper bound of the 90% (one-tailed) UCL of the mean is compared to the regulatory level (in this case 5.0 mg/L for lead using the TCLP). The 90% one-tailed approach has been determined by the EPA to provide an adequate margin of safety against making a wrong decision.

Cost Evaluation

This section evaluates the cost associated with the alternate sampling designs.

For Case 1 (authoritative sampling design):

A judgmental authoritative design meets the requirements for the study; that is, it estimates the average lead concentration (via the TCLP) for the material in the waste pile. "Average" is used here rather than "mean," which is associated with a probabilistic design. Seven samples will be collected at

an analytical cost of $250 per sample plus an additional 10% for various quality assurance samples. The total analytical cost for each remaining sampling design will be approximately $1,925, which is under the analytical budget target of $2,000. Because a judgmental authoritative design provides information on the average concentration of lead in the waste pile (without the establishment of a confidence interval), it is selected as the preferred sampling design. Note that if this simple design did not meet the study objectives, then a modification in either the design or the study objectives would be required.

For Case 2 (simple random sampling design):

The simple random design as well as both approaches to the systematic design (with and without compositing) meet the statistical requirements for the study in determining the estimated mean lead concentration (via the TCLP) for the material in the waste pile. If a simple random design or a systematic grid design without compositing is chosen, 30 samples will be collected. The analytical cost per sample is $250 including the totals and TCLP, and various quality assurance samples would increase the cost by approximately 10%. Both the simple random design and the systematic grid design without compositing would generate a total analytical cost of about $8,250 (30 samples at $250 for the totals and TCLP plus 10% for quality assurance). The stakeholders decide on the simple random design because they expect the waste pile to be relatively homogeneous; therefore, information on the distribution of lead is not important.

For Cases 3–4 (systematic grid sampling designs):

Again the simple random design and both approaches to the systematic design (with and without compositing) meet the statistical requirements for the study in determining the estimated mean lead concentration (via the TCLP) for the material in the waste pile. If a simple random design or a systematic grid design without compositing is chosen, 15 samples will be collected, to exceed the estimated number of necessary samples. The analytical cost per sample is $250 for the TCLP, and various quality assurance samples would increase the cost by approximately 10%. Both simple random design and the systematic grid design without compositing would generate a total analytical cost of about $4,125 (15 samples at $250 each for the TCLP plus 10% for quality assurance). A systematic grid design with compositing may improve precision over the systematic design without compositing. For Case 3, the analytical costs of each of the alternate sample designs are within the budget of $5,000. The stakeholders decide to use the systematic grid design because spatial information is desired. For Case 4, the systematic grid with compositing is chosen to improve precision and study efficiency (fewer samples collected). Four composite samples will be collected. The cost for that design is approximately $1,100).

For Case 5 (stratified random sampling design):

A stratified random approach is chosen due to the expected stratification of the waste pile. This approach should improve the efficiency of the final determination on the entire waste pile. The analytical costs are estimated at $4,675 (17 samples at $250 each for the TCLP plus 10% for quality assurance) and are within the proposed analytical budget of $5,000.

(What if the Alternate Designs Do Not Meet the DQOs?)

Note that if the sampling designs do not meet the study objectives for each case, then a modification in either the design (more samples, use of sampling tools such as compositing or double sampling) or study objectives (change in the confidence interval, study boundaries, allowable decision error, or budget constraints) will then be required.

IMPLEMENTATION PHASE

For All Cases

Implementation of the authoritative design, simple random design, systematic grid design, and the stratified random design should not present any significant problems. The samples will be collected using decontaminated hand augers, and glass pans will be used for sample mixing. The samples will be collected to a depth of 1 ft (0.61 m) at each location. Note that for Case 1 information will be collected to evaluate the potential presence of vertical stratification in the waste pile. In that Case, samples for vertical profiling will be collected at one location by a boring advanced to the base of the waste pile. Individual samples will be collected at 2-ft (0.61 m) intervals. The simple and stratified random samples may require careful surveying to determine the location of the specific sampling locations. See Figs. 5–9 at the end of this chapter for the sample locations.

ASSESSMENT PHASE

This section illustrates some of the graphical and statistical techniques available for completing the data quality assessment (DQA) step of a data collection activity. The U.S. EPA publication on Data Quality Assessment (QA/G-9) and the accompanying software (DataQUEST) may be utilized as a tool by the investigator in this step [*2,3*]. Other references provided in Chapter 4 of the manual should also be consulted. More detail is presented for Case 2 in order to illustrate a range of graphical and statistical assessment options.

Review of the DQOs and the Sampling Design

In each case, the data collected during the study have met the DQOs. Sampling error was minimized through the selection and use of correctly designed sampling devises, careful implementation of the field sampling and handling procedures, and use of minimally biased subsampling procedures within the laboratory (e.g., using guidance in ASTM D 6051) as specified in the QAPP and SOPs. The material that was sampled does not appear to have presented any special problems concerning access to sampling locations, equipment usage, particle size distribution, or matrix interferences. The analytical package has been validated and the data generated are acceptable for their intended purpose.

FOR CASE 1—AUTHORITATIVE SAMPLING DESIGN:

Preliminary Data Review

Results for the data collection effort are listed in Table 1-1.

Statistical Quantities:

Table 1-2 lists the totals and TCLP mean and range of values for lead. As expected, the TCLP concentration for lead greatly exceeds the TC Rule regulatory level of 5.0 mg/L. Totals and TCLP results of the vertical boring indicate that there is not a discernable difference in the lead concentration at the 1 to 3 and 3 to 5 ft intervals versus the surface interval (0 to 1 ft). This confirms the original assumptions concerning vertical stratification that was based on knowledge of the waste generated and the management practices of the facility.

Graphical Representation for Case 1 data:

Because of the limited amount of data collected and the authoritative nature of the study design, no graphical depictions were prepared.

Conclusion

Based on the established decision rule, the material in the waste pile was determined to be hazardous for lead for Case 1. The totals results could be used for profiling the waste to ensure compliance with the Subtitle C permit (see Identifying Inputs to the Decision).

TABLE 1-1—Total and TCLP Results for Case 1.

Location	C3	C7	E5	G3	G7
Totals result (mg/kg)	1400	975	1420	1800	1500
TCLP result (mg/L)	26	20	30	42	32

Vertical Boring	Total Results, mg/kg	TCLP Results, mg/L
E5 (1–3 feet)	1600	28
E5 (3–5 feet)	1350	32

NOTE: 1 ft = 0.3048 m.

TABLE 1-2—Totals and TCLP Statistical Results—Case 1.

Totals Results, mg/kg		TCLP Results, mg/L	
Average	Range	Average	Range
1419	975–1800	30	20–42

FOR CASE 2—SIMPLE RANDOM SAMPLING DESIGN

FOR CASE 2 CONSIDER TWO DIFFERENT DATA SETS, TERMED 2A (NORMAL DISTRIBUTION) AND 2B (NON-NORMAL DISTRIBUTION)

FOR CASE 2A (NORMAL DISTRIBUTION):

Preliminary Data Review

The results for the data collection effort are listed in Table 2a-1. Thirty samples were collected to exceed twenty nine (the number of samples calculated to achieve the specified margin of error). Note that the samples collected from the two vertical cores (Locations H8 and C4) indicate that no significant vertical stratification is present.

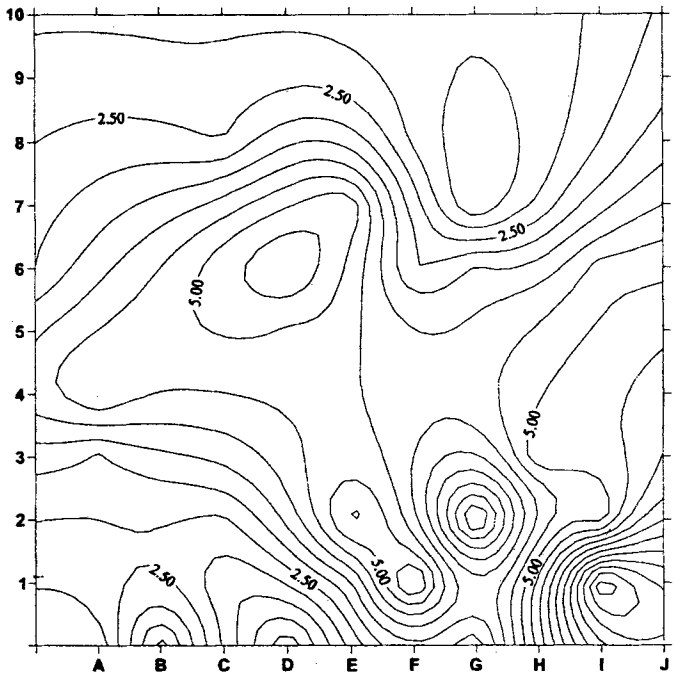


FIG. 2a-1—Lead concentration distribution—Case 2a.

TABLE 2a-1—Totals and TCLP Analytical Results for Case 2a.

Location	Totals Result, mg/kg	TCLP Result, mg/L	Location	Totals Result, mg/kg	TCLP Result, mg/L
A5	1574	4.34	F3	1478	5.73
A7	1047	2.95	F8	1678	5.36
B1	405	1.58	G2	1415	6.34
B4	328	2.86	G7	452	3.05
B5	1234	5.03	G9	24	1.92
B9	661	2.65	H1	219	2.57
C1	1359	4.31	H3	189	0.74
D2	327	1.61	H7	358	3.57
D3	129	2.40	H8	89	1.00
D7	924	5.29	I4	1592	5.36
D9	1012	2.54	I8	2015	10.50
E1	24	0.11	J2	861	6.30
E6	1310	4.89	J3	654	4.61
E7	605	6.04	J7	1014	4.70
F2	1319	3.42	J9	689	2.55

Graphical Representation:

Figure 2a-1 shows the lead concentration isopleth based on the data generated. Although the graphical depiction has inherent limitations, the distribution of lead across the waste pile can be readily observed. No spatial trends or distinct strata are apparent.

Statistical Evaluation of the Data

TCLP versus Totals Results

Figure 2a-2 is provided to evaluate the general relationship between the TCLP and Totals results. The data presented is provided for illustrative purposes, and conclusions should not be drawn about any relationship between the totals and the TCLP data for other data sets. However, the information concerning this relationship could be useful in the future to estimate in very general terms at what totals concentration is this waste likely to exceed the TCLP regulatory level (approximately ≥1,600 mg/kg). Remember, use the results of this comparison with caution, even with a similar waste stream. Note also that in most cases the investigators would not have completed the TCLP on samples collected at the following locations since the Total results were below 100 mg/kg—E1, G9, and H8.

Histogram

Figure 2a-3 is a histogram of the totals data, which provides a picture of the shape of the data and aids in identifying the symmetry and variability of the data set. Using a histogram, one may visually estimate the underlying distribution using binned data plotted against relative frequency of occurrence. If the data are symmetric, then the structure of the histogram will be symmetric around a central point, such as the mean, if the data set is sufficiently large ($n > 25$). Thus, using a histogram, a normal distribution or a skewed distribution may be visually identified. The histogram provides a tool for preliminary data assessment but is inadequate for verification of distributional assumptions. TCLP data is used to test distributional assumptions since the final decision will be made using this data set. EPA's QA/G-9 (Guidance for Data Quality Assessment) provides guidance in creating a histogram. In this case, the histogram appears to display symmetric data [*2*].

Coefficient of Variation

The coefficient of variation (CV) may be used to quickly check if the data may be modeled by the normal curve by comparing the sample CV to 1. If the CV is greater than 1, then the data should not be modeled by a normal curve. However, this method should not be used to conclude the opposite. (If CV < 1, the test is inconclusive). The CV is computed by dividing the standard deviation by the mean of the data set. In this case, the CV of the TCLP data is computed to be 0.6, so the test is inconclusive.

Box and Whiskers Plot

An additional visual method of evaluating the shape of the data is a box and whiskers plot; it is useful in determining the

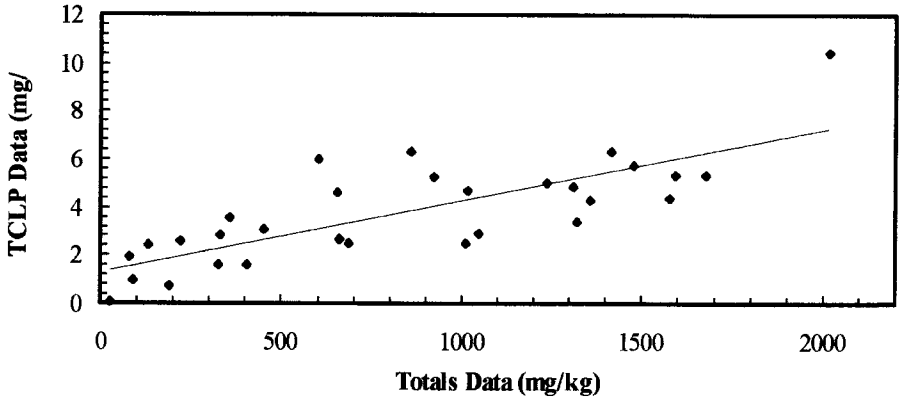


FIG. 2a-2—TCLP vs. total data—Case 2a.

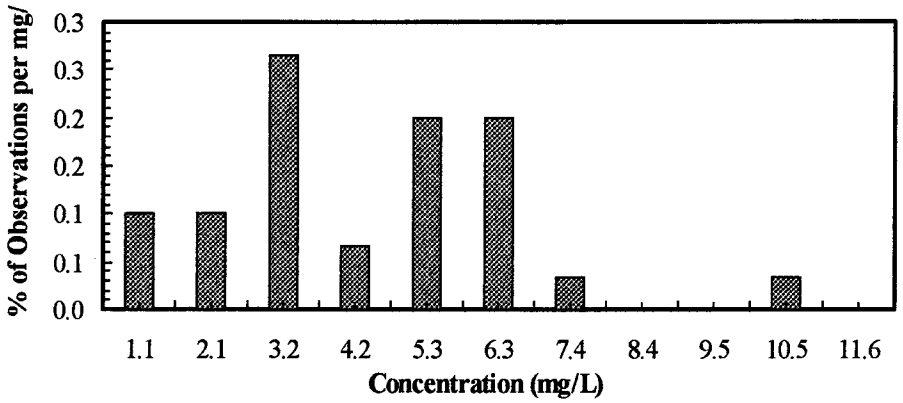


FIG. 2a-3—Histogram—Case 2a.

symmetry of the data. See QA/G-9 for guidance on constructing Box and Whiskers plots. The TCLP data was used to generate the box and whiskers plot for Case 2a seen in Fig. 2a-4.

The box and whiskers plot consists of a central box, whose length denotes the spread of the bulk of the data (the central 50%) and whiskers, whose length indicates the spreading of the distribution tails. The width of the box is arbitrary. The plus sign marks the sample mean, and the sample median is displayed as a line through the box. Any outlying data points are marked by a "*" on the plot. In Case 2 the identified "outlier" is the TCLP result at Location J2 (10.5 mg/L). Techniques and approaches for determining when to keep or discard an identified outlier are discussed in Chapter 4 of the manual. Just because this technique identifies the data point as an outlier does not mean that the data point should be discarded. It could be an actual hot-spot within the pile rather than an error introduced through cross contamination of the sample or laboratory problems. If a valid reason for the "outlier" cannot be identified, then further investigation at this location in the waste pile may be warranted.

If the distribution is symmetrical, the box is divided into two equal halves; the whiskers are about the same length, and any extreme data points are equally distributed. According to the box and whiskers plot shown here, the data set appears to be symmetrical with one identified outlier.

Normal Probability Plot (Quantile-Quantile Plot)

A normal probability plot, or Q-Q plot (Fig. 2a-5), may be used to visually check if a sample data set fits a specified probability model. The n TCLP data values, x_i, are plotted against the expected data value, y_i, from the parent model probability distribution. A normal probability plot, which may be used to test the assumption of normality, is the graph of the quantiles of a data set against the quantiles of the normal distribution. If the data follow an approximate linear trend on the plot, the validity of the normality assumption is probable. Refer to EPA QA/G-9 for guidance on generating a normal probability plot. The data set appears to be normally distributed from the Q-Q plot in Fig. 2a-5. However, the plot is a visual quantifier of the data and may not be used to finalize distributional assumptions.

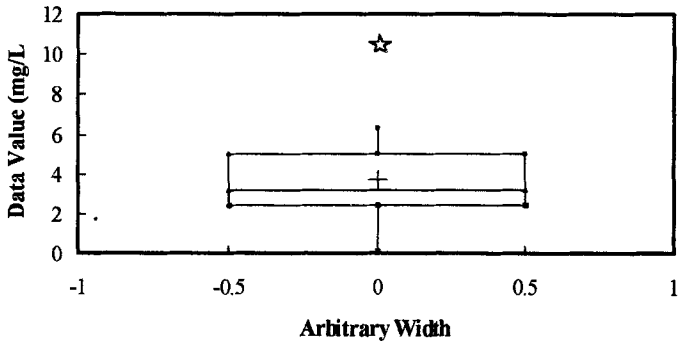


FIG. 2a-4—Box and whiskers plot—Case 2a.

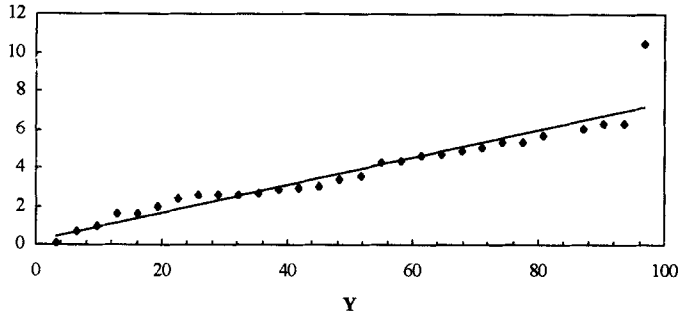


FIG. 2a-5—Normal probability plot—Case 2a.

Shapiro-Wilk Test for Normality

A more precise test for distributional assumptions is the Shapiro-Wilk test, which is conducted on the TCLP data to check for normality as follows:

Compute d, the denominator of the test statistic, using the n data.

$$d = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = 132$$

Compute k, where

$k = n/2$ If n is even.
$k = (n - 1)/2$ If n is odd.

In this case, $n = 30$ and $k = 15$. From Table 1 in Appendix D (Table A-6 in Gilbert's Statistical Methods for Environmental Pollution Monitoring (1989)), the coefficients for the test may be obtained as $a_1, a_2, \ldots, a_k$. [4]. Then compute the W value.

$$W = \frac{1}{d}\left[\sum_{i=1}^{k} a_i \left(x_{[n-i+1]} - x_{[i]}\right)\right]^2 = 0.948$$

If the computed W value is greater the tabled quantile at the given alpha significance level, then the assumption of normality cannot be rejected. In this case, alpha is taken to be 0.01. Because the W value for this example is higher than the 0.01 quantile of 0.900, the assumption of normality cannot be rejected. W values may be obtained from Table 2 in Appendix D of this manual (also found in Gilbert, Table A-7 "Shapiro-Wilk Tables").

Characterization of the Distribution

The statistical analysis of the TCLP data upheld the distributional assumption of normality. Statistical quantities may now be calculated based on the assumption of normality. The results are displayed in Table 2a-2.

To calculate the 90% UCL when the true standard deviation is not known, use the t distribution from Table 3 in Appendix D. Calculate the 90% UCL by

$$\begin{aligned} 90\%\ \text{UCL} &= \bar{x} + t_{1-a}\left(\frac{s}{\sqrt{n}}\right) \\ &= \bar{x} + t_{0.90}\left(\frac{s}{\sqrt{n}}\right) \\ &= 3.8 + 1.311\left(\frac{2.1}{\sqrt{30}}\right) \\ &= 4.3\ \text{mg/L} \end{aligned}$$

The tabulated "t value" (1.311) is based on a 90% one-tailed confidence interval with a probability of 0.10, $t_{0.90}$ (see Table 1 in Appendix D).

TABLE 2a-2—Totals and TCLP Results—Case 2a.

	Mean	Range	Standard Deviation	Variance	Coefficient of Variation	90% UCL (one-tailed)
Totals Result, mg/kg	833	24–2015				
TCLP Result, mg/L	3.8	0.1–10.5	2.1	4.6	0.6	4.3

TABLE 2b-1—Totals and TCLP Analytical Results—Case 2b.

Location	Totals Result, mg/kg	TCLP Result, mg/L	Location	Totals Result, mg/kg	TCLP Result, mg/L
A5	308	1.7	F3	1283	3.4
A7	474	1.7	F8	320	1.7
B1	570	2.3	G2	869	3.2
B4	709	1.9	G7	331	3.0
B5	415	2.7	G9	540	1.6
B9	363	1.1	H1	502	1.7
C1	516	3.0	H3	1118	4.3
D2	72	1.2	H7	268	2.4
D3	654	2.4	H8	348	1.5
D7	643	2.0	I4	498	5.2
D9	336	1.2	I8	461	4.6
E1	777	2.2	J2	2259	7.1
E6	234	1.0	J3	453	1.4
E7	334	1.5	J7	2587	6.9
F2	474	4.5	J9	283	1.9

Conclusion

The 90% UCL for the mean of the TCLP data is calculated to be 4.3 mg/L, which is less than the regulatory level of 5.0 mg/L. Thus, in Case 2a the material in the waste pile is determined not to be hazardous for lead based on the established decision rule. Note that the TCLP result for the pilot study (5.7 mg/L) indicated that the waste pile was hazardous; however, the more comprehensive evaluation using a simple random approach shows that the waste pile is actually non-hazardous. This illustrates the potential advantage of an expanded characterization effort based on a probabilistic sampling design.

A quick check may be performed to determine if an adequate number of samples was collected to satisfy specified error limits. Refer to Chapter 2 of the Manual to review the sample size equation. The standard deviation and sample mean are entered into the sample size equation with $n - 1 = 29$ degrees of freedom and $\alpha = 0.10$.

$$n = \frac{t^2_{1-\alpha} \cdot s^2}{\Delta^2} = \frac{1.311^2 \cdot 2.1^2}{(5 - 3.8)^2} = 6$$

Five is less than thirty; therefore, the test was sufficiently powerful and achieves the Type I error rate specified in the DQOs.

FOR CASE 2B (NON-NORMAL DATA DISTRIBUTION):

Preliminary Data Review

The results for the data collection effort are listed in Table 2b-1.

Graphical Representation:

See Fig. 2a-1 for an example of concentration isopleths based on the data generated.

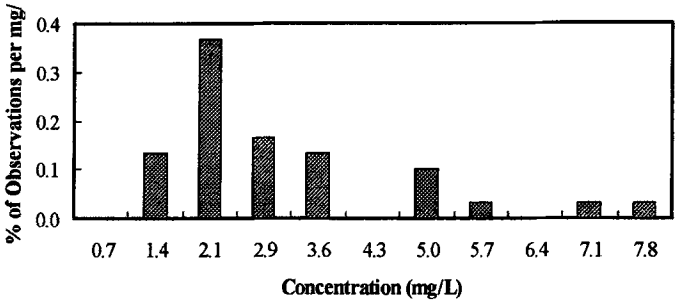


FIG. 2b-1—Histogram—Case 2b.

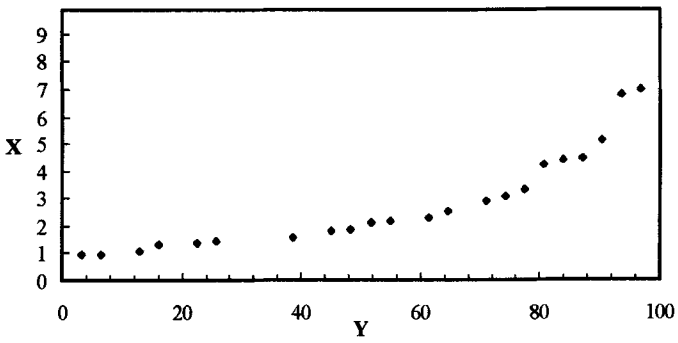


FIG. 2b-2—Normal probability plot—Case 2b.

Statistical Evaluation of the Data

The CV test yields a value of 0.6 for the TCLP data. The CV value is less than 1. Thus, this method is inconclusive, and additional statistical evaluation is needed. Figure 2b-1 is a histogram of the totals data.

The histogram does not appear to display normally distributed data. A normal probability plot is constructed to further test the distribution (Fig. 2b-2).

The data set does not follow a linear trend; thus, the distribution may not be normal. The Shapiro-Wilk test is performed to further verify the deviation from normality at a 0.01 significance level. The test estimated a W value of 0.827, which is less than the 0.01 quantile, 0.900 (found in Appendix D). Thus, the Shapiro-Wilk test confirms the non-normality of the data. To check for lognormality, a lognormal probability plot may be created (Fig. 2b-3) in which the natural logarithms of the data are plotted against the calculated Y. If the data lie linearly on the lognormal plot, the assumption of a lognormal distribution is strengthened.

The natural logarithms of the data follow an approximately linear trend on a logrithmic scale. Thus, the plot agrees with the assumption of log-normality. The Shapiro-Wilk test is a more accurate way to access lognormality by conducting the test on the natural logrithms of the data. This method produces a W value of 0.946. Because the W value for this example is higher than the 0.10 quantile of 0.939 (found in Appendix D), the assumption of log-normality may be accepted as valid.

Characterization of the Distribution

The statistical analysis of the data indicates a log-normal data distribution. Statistical quantities are calculated for the TCLP data assuming a log-normal data distribution. The resulting values are displayed in Table 2b-2. The 90% upper confidence limit for the mean is then compared to the regulatory limit of 5.0 mg/L. Several methods exist for estimating the mean of a log-normal distribution [4]. A simple method for estimating the mean and variance of lognormally distributed data is illustrated below.

Compute the log-transformed data set $y_i = \ln x_i$ where x_i is the original data set. Then compute the mean and variance of the log-transformed data.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 0.8$$

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2 = 0.3$$

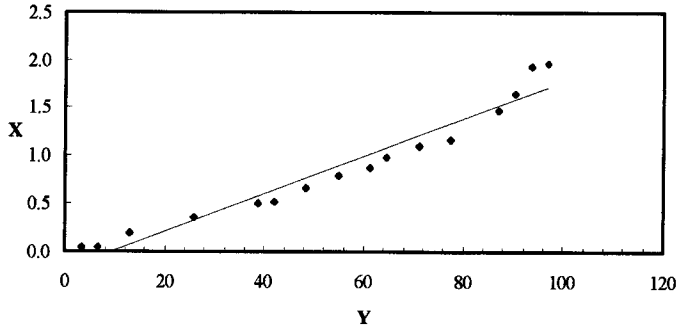


FIG. 2b-3—Lognormal probability plot—Case 2b.

The upper one-sided $100(1 - \alpha)\%$ confidence limit for the mean of log-normally distributed data is calculated by:

$$\mathrm{UCL}_{1-\alpha} = \exp\left(\bar{y} + 0.5s_y^2 + \frac{s_y H_{1-\alpha}}{\sqrt{n-1}}\right)$$

where $\bar{y}$ and s_y^2 are the mean and the variance, respectively, of the log-transformed data set, n is the number of samples, and $H_{1-\alpha}$ is an empirical constant that is provided in tables by Land and Gilbert [4]. For $\alpha = 0.1$, $H_{1-\alpha} = 1.505$, and the UCL_{90} is calculated to be 3.1 mg/L. Note that this formula for estimating the UCL on the mean of a lognormal distribution can give unreliable results if n is small even when the data are truly lognormally distributed. Refer to Singh for further information on the lognormal distribution [5].

Conclusion

The 90% UCL for the mean of a log-normal distribution was calculated to be 3.1 mg/L, which is less than the regulatory level of 5.0 mg/L. Thus, in Case 2b the material in the waste pile was determined not to be hazardous for lead based on the established decision rule.

FOR CASE 3—SYSTEMATIC GRID WITHOUT COMPOSITING SAMPLING DESIGN:

Preliminary Data Review

Fifteen samples were collected to exceed eleven (the calculated number of samples to achieve the desired margin of error). The results for the data collection effort are listed in Table 3-1.

Graphical Representation:

A graphical depiction of the data could be completed. (See Case 2a for an example.)

Statistical Evaluation of the Data

A histogram is not constructed because the number of samples is too small to accurately use this quantifier ($n < 25$). A normal probability plot is constructed to test the assumption

TABLE 3-1—Totals and TCLP Results—Case 3.

Location	TCLP Result, mg/L	Location	TCLP Result, mg/L
B2	0.7	F2	3.6
B4	4.5	F4	5.2
B6	7.9	F6	6.1
B8	6.0	F8	7.4
D2	4.1	H2	1.1
D4	2.3	H4	9.6
D6	5.2	H6	5.6
D8	9.2		

TABLE 2b-2—Totals and TCLP Statistical Result—Case 2b.

	Mean	Range	Standard Deviation	Variance	Coefficient of Variation	90% UCL (one-tailed)
Totals Results, mg/kg	633	72–2587				
TCLP Results, mg/L	2.7	1.0–7.1	1.6	2.6	0.6	3.1

of normality (Fig. 3-1). Again, the TCLP data is used to test for normality.

The data set appears to be normally distributed from the Q-Q plot. The Shapiro-Wilk test is conducted on the TCLP data to further validate the distributional assumption of normality. The W value is 0.939, which is higher than the 0.01 quantile of 0.855 (found in Table 2 of Appendix D), so the assumption of normality cannot be rejected.

Characterization of the Distribution

The statistical analysis of the data upheld the distributional assumption of normality. Statistical quantities may now be calculated based on the assumption of normality. The results are displayed in Table 3-2.

To calculate the 90% UCL, use the t-distribution:

$$\text{90\% UCL for TCLP data} = \bar{x} + t_{1-\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 6.3 + t_{0.90,14}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 6.3 + 1.345\left(\frac{2.6}{\sqrt{15}}\right)$$

$$= 7.2 \text{ mg/L}$$

The tabulated "t value" (1.345) is based on a 90% one-tailed confidence interval with a probability of 0.10 and 14 degrees of freedom, $t_{0.90,14}$ (Table 3 in Appendix C).

Conclusion

The 90% UCL for the mean of the TCLP data is 7.2 mg/L, which is greater than the regulatory level of 5.0 mg/L. Thus, in Case 3 the material in the waste pile is determined to be hazardous for lead based on the established decision rule.

A quick check is performed to determine if a sufficient number of samples were collected to satisfy specified decision error limits on the test for whether the waste pile is hazardous. The standard deviation and sample mean are entered into the sample size equation with $n - 1 = 14$ degrees of freedom and $\alpha = 0.10$. The calculated number is six samples, which is less than fifteen, therefore a sufficient number of samples was collected.

FOR CASE 4—SYSTEMATIC GRID WITH COMPOSITING SAMPLING DESIGN:

Preliminary Data Review

Four samples were collected as specified by the sample size equation. The results for the data collection effort are listed in Table 4-1.

Statistical Evaluation of the Data

A histogram is not constructed because the number of samples is too small to accurately use this quantifier. A normal probability plot is constructed on the TCLP data to test the assumption of normality (Fig. 4-1).

The data set appears to be normally distributed from the normal probability plot. The Shapiro-Wilk test is conducted to further validate the distributional assumption. The W value (Table 2 in Appendix D) is 0.903, which is higher than the 0.01 quantile for the sample size of 0.707, so the assumption of normality cannot be rejected. However, it should be noted that both the Q-Q plot and the Shapiro-Wilk test have low power to detect small deviations from normality when n is so small.

Characterization of the Distribution

The statistical analysis of the totals data upheld the distributional assumption of normality. Statistical quantities may

TABLE 4-1—Totals and TCLP Results for Case 4.

Location	TCLP Result, mg/L
C2	4.8
C8	3.4
H2	4.1
H8	4.9

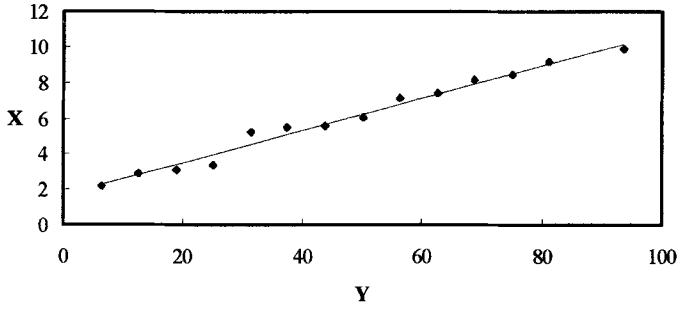


FIG. 3-1—Normal probability plot.

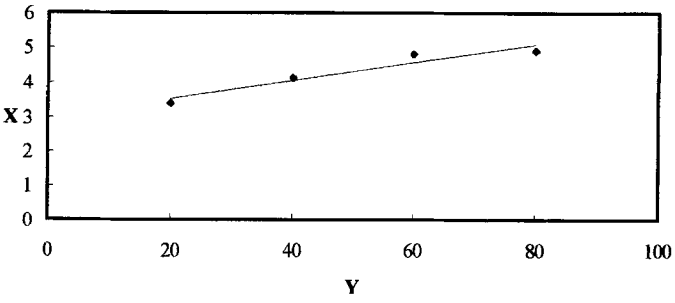


FIG. 4-1—Normal probability plot for Case 4.

TABLE 3-2—Totals and TCLP Statistical Result—Case 3.

	Mean	Range	Standard Deviation	Variance	Coefficient of Variation	90% UCL (one-tailed)
TCLP Results, mg/L	6.3	2.2–9.9	2.6	6.6	0.4	7.2

TABLE 4-2—Totals and TCLP Statistical Results—Case 4.

	Mean	Range	Standard Deviation	Variance	Coefficient of Variation	90% UCL (one-tailed)
TCLP Results, mg/L	4.3	3.4–4.9	0.3	0.1	0.1	4.6

now be calculated based on the assumption of normality. The results are displayed in Table 4-2.

Conclusion

The 90% UCL for the mean of the TCLP data is 4.6 mg/L, which is less than the regulatory level of 5.0 mg/L. Thus, in Case 4 the material in the waste pile is determined to be non-hazardous for lead based on the established decision rule.

A quick check is performed to determine if a sufficient number of samples were collected to satisfy specified decision error limits on the test for whether the waste pile is hazardous. The standard deviation and sample mean are entered into the sample size equation with $n - 1 = 3$ degrees of freedom and $\alpha = 0.10$. The calculated number is one sample, which is less than four, therefore a sufficient number of samples was collected.

FOR CASE 5—STRATIFIED RANDOM SAMPLING DESIGN:

Preliminary Data Review

Three samples are collected for stratum one, and fourteen samples are collected from Stratum 2 as calculated in the sample size equation for proportional allocation. The results for the data collection effort are listed in Table 5-1.

Characterization of the Distribution

Statistical quantities may now be calculated. The results are displayed in Table 5-2.

For a stratified design which considers multiple strata, the overall mean concentration for the waste pile, $\bar{x}_{total}$, may be calculated using the following formula [6]:

$$\bar{x}_{total} = \sum_{h=1}^{L} W_h \cdot \bar{x}_h = 0.8 \cdot 3.7 + 0.2 \cdot 9.9 = 4.9$$

where $\bar{x}_h$ is equal to the mean of the individual stratum (computed as shown above for Case 2a—Simple Random), W_h is equal to the weight of the individual stratum, h is the individual stratum, and L is the total number of strata.

The standard deviation of the overall waste pile may be calculated by:

$$s_{total} = \sqrt{\sum_{h=1}^{L} W_h^2 \cdot \frac{s_h^2}{n_h}} = 0.2$$

where $_{Nh}$ is the number of samples collected in the h^{th} stratum. To calculate the upper confidence limit (UCL) on the mean, the degrees of freedom (df) must first be calculated using the formula

$$df = \frac{s_{total}^4}{\sum_{h=1}^{L} \frac{(W_h \cdot s_h)^4}{n_h^2 (n_h - 1)}} = 469$$

The upper confidence limit on the mean can then be calculated using the specified alpha error rate and the degrees of freedom calculated using the above equation.

$$UCL_\alpha = \bar{x}_{total} + t_{1-\alpha,df} \cdot s_{total} = 4.9 + 1.284 \cdot 0.2 = 5.1 \text{ mg/L}$$

Conclusion

The 90% UCL for the mean of the TCLP data is 5.1 mg/L, which is greater than the regulatory level of 5.0 mg/L. Thus, material in the waste pile is determined to be hazardous for lead based on the established decision rule.

TABLE 5-1—Totals and TCLP Results—Case 5.

Location	TCLP Result, mg/L	Location	TCLP Result, mg/L
Stratum 1 (A1):	9.2	Stratum 2 (F4):	4.8
Stratum 1 (B3):	10.5	Stratum 2 (F7):	3.0
Stratum 1 (C2):	9.9	Stratum 2 (G8):	4.4
Stratum 2 (A8):	3.5	Stratum 2 (H1):	3.7
Stratum 2 (B7):	4.2	Stratum 2 (H6):	3.1
Stratum 2 (C5):	3.8	Stratum 2 (I9):	5.0
Stratum 2 (D7):	3.6	Stratum 2 (J3):	2.8
Stratum 2 (E9):	2.3	Stratum 2 (J6):	3.4
Stratum 2 (F2):	4.0		

REFERENCES

[1] U.S. EPA, "RCRA Waste Sampling Draft Technical Guidance SW-846 Chapter Nine—Planning, Implementation, and Assessment," EPA/530/R-99/015, Solid Waste and Emergency Response, Washington, DC, 1999.

[2] U.S. EPA, "Guidance for Data Quality Assessment—Practical Methods for Data Analysis," QA/G-9, EPA/600/R-96/084, Office of Research and Development, Washington, DC, 1998.

[3] U.S. EPA, Data Quality Assessment Statistical Toolbox (DataQUEST), User's Guide and Software, <http://es.epa.gov/

TABLE 5-2—Totals and TCLP Statistical Results for Case 5.

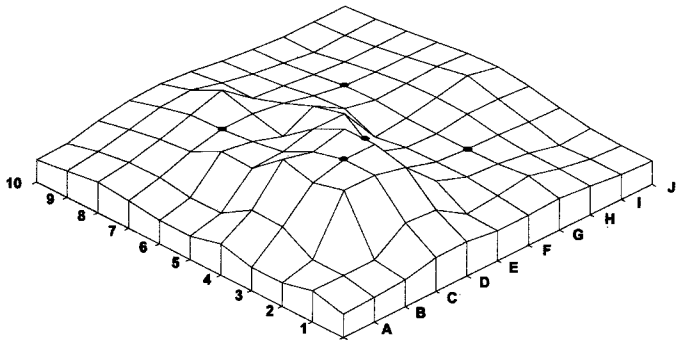
	Standard	Coefficient of	90% UCL

ncerqa/qa/qa_docs.html#G-9d> EPA QA/G-9D, EPA/600/R-96/085, December 1996.

[4] Gilbert, R. O., *Statistical Methods for Environmental Pollution Monitoring*, John Wiley and Sons, New York, NY, 1987.

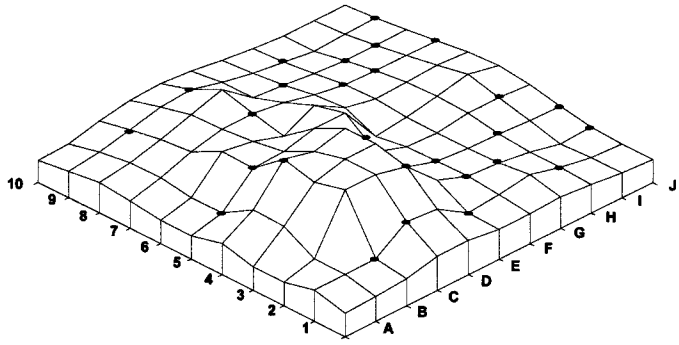
[5] Singh, A. K., Singh, A., and Engelhardt, M., "The Lognormal Distribution in Environmental Applications," (EPA/600/R-97/006), Technology Support Center Issue, USEPA Office of Research and Development, 1997.

[6] U.S. EPA, "Methods For Evaluating the Attainment of Cleanup Standards—Volume 1: Soils and Solid Media," EPA 230/02-89-042, Office of Policy Planning and Evaluation, Washington, DC, 1989.



• Sample Location

FIG. 5—Sample location map, Case 1: Authoritative Sampling Design.



• Sample Location

FIG. 6—Sample location map, Case 2a and 2b: Simple Random Design.

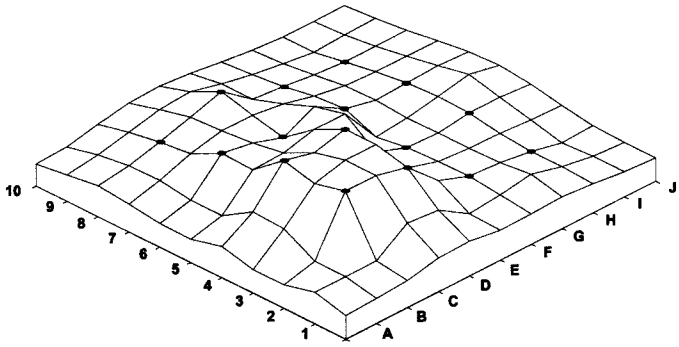
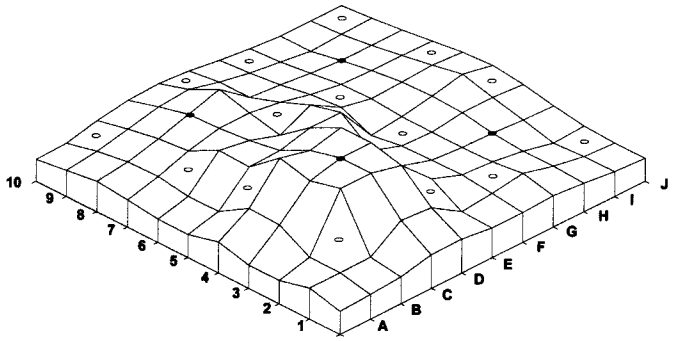


FIG. 7—Sample location map, Case 3: Systematic Grid Sampling Design (without compositing).



• Sample Location (center point)
○ Alliquot Locations

FIG. 8—Sample location map, Case 4: Systematic Grid Sampling Design (with compositing).

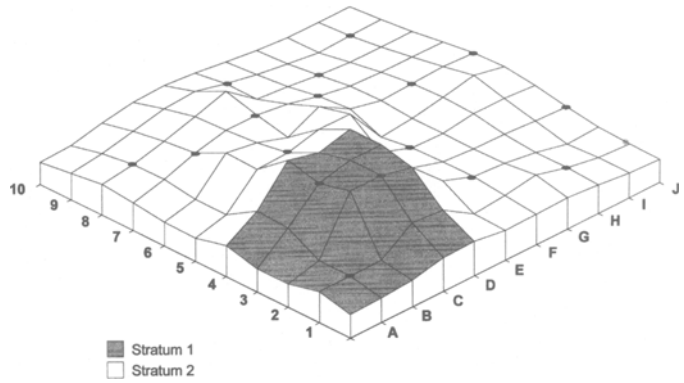


FIG. 9—Sample location map, Case 5: Stratified Random Sampling Design.

Appendix D: Statistical Tables

TABLE 1—Coefficients of a_i for the Shapiro-Wilk Test for Normality.

i \ n	2	3	4	5	6	7	8	9	10	
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739	
2	—	0.0000	0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291	
3	—	—	—	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141	
4	—	—	—	—	—	0.0000	0.0561	0.0947	0.1224	
5	—	—	—	—	—	—	—	0.0000	0.0399	
i \ n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	—	—	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	—	—	—	—	0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9	—	—	—	—	—	—	0.0000	0.0163	0.0303	0.0422
10	—	—	—	—	—	—	—	—	0.0000	0.0140
i \ n	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836	0.1848	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	—	—	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	—	—	—	—	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	—	—	—	—	—	—	0.0000	0.0084	0.0159	0.0227
15	—	—	—	—	—	—	—	—	0.0000	0.0076
i \ n	31	32	33	34	35	36	37	38	39	40
1	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
2	0.2921	0.2898	0.2876	0.2854	0.2834	0.2813	0.2794	0.2774	0.2755	0.2737
3	0.2475	0.2462	0.2451	0.2439	0.2427	0.2415	0.2403	0.2391	0.2380	0.2368
4	0.2145	0.2141	0.2137	0.2132	0.2127	0.2121	0.2116	0.2110	0.2104	0.2098
5	0.1874	0.1878	0.1880	0.1882	0.1883	0.1883	0.1883	0.1881	0.1880	0.1878
6	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
7	0.1433	0.1449	0.1463	0.1475	0.1487	0.1496	0.1505	0.1513	0.1520	0.1526
8	0.1243	0.1265	0.1284	0.1301	0.1317	0.1331	0.1344	0.1356	0.1366	0.1376
9	0.1066	0.1093	0.1118	0.1140	0.1160	0.1179	0.1196	0.1211	0.1225	0.1237
10	0.0899	0.0931	0.0961	0.0988	0.1013	0.1036	0.1056	0.1075	0.1092	0.1108
11	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0986
12	0.0585	0.0629	0.0669	0.0706	0.0739	0.0770	0.0798	0.0824	0.0848	0.0870
13	0.0435	0.0485	0.0530	0.0572	0.0610	0.0645	0.0677	0.0706	0.0733	0.0759
14	0.0289	0.0344	0.0395	0.0441	0.0484	0.0523	0.0559	0.0592	0.0622	0.0651
15	0.0144	0.0206	0.0262	0.0314	0.0361	0.0404	0.0444	0.0481	0.0515	0.0546
16	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
17	—	—	0.0000	0.0062	0.0119	0.0172	0.0220	0.0264	0.0305	0.0343
18	—	—	—	—	0.0000	0.0057	0.0110	0.0158	0.0203	0.0244
19	—	—	—	—	—	—	0.0000	0.0053	0.0101	0.0146
20	—	—	—	—	—	—	—	—	0.0000	0.0049

TABLE 1—(continued).

i^{n}	41	42	43	44	45	46	47	48	49	50
1	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
2	0.2719	0.2701	0.2684	0.2667	0.2651	0.2635	0.2620	0.2604	0.2589	0.2574
3	0.2357	0.2345	0.2334	0.2323	0.2313	0.2302	0.2291	0.2281	0.2271	0.2260
4	0.2091	0.2085	0.2078	0.2072	0.2065	0.2058	0.2052	0.2045	0.2038	0.2032
5	0.1876	0.1874	0.1871	0.1868	0.1865	0.1862	0.1859	0.1855	0.1851	0.1847
6	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
7	0.1531	0.1535	0.1539	0.1542	0.1545	0.1548	0.1550	0.1551	0.1553	0.1554
8	0.1384	0.1392	0.1398	0.1405	0.1410	0.1415	0.1420	0.1423	0.1427	0.1430
9	0.1249	0.1259	0.1269	0.1278	0.1286	0.1293	0.1300	0.1306	0.1312	0.1317
10	0.1123	0.1136	0.1149	0.1160	0.1170	0.1180	0.1189	0.1197	0.1205	0.1212
11	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
12	0.0891	0.0909	0.0927	0.0943	0.0959	0.0972	0.0986	0.0998	0.1010	0.1020
13	0.0782	0.0804	0.0824	0.0842	0.0860	0.0876	0.0892	0.0906	0.0919	0.0932
14	0.0677	0.0701	0.0724	0.0745	0.0765	0.0783	0.0801	0.0817	0.0832	0.0846
15	0.0575	0.0602	0.0628	0.0651	0.0673	0.0694	0.0713	0.0731	0.0748	0.0764
16	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
17	0.0379	0.0411	0.0442	0.0471	0.0497	0.0522	0.0546	0.0568	0.0588	0.0608
18	0.0283	0.0318	0.0352	0.0383	0.0412	0.0439	0.0465	0.0489	0.0511	0.0532
19	0.0188	0.0227	0.0263	0.0296	0.0328	0.0357	0.0385	0.0411	0.0436	0.0459
20	0.0094	0.0136	0.0175	0.0211	0.0245	0.0277	0.0307	0.0335	0.0361	0.0386
21	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
22	—	—	0.0000	0.0042	0.0081	0.0118	0.0153	0.0185	0.0215	0.0244
23	—	—	—	—	0.0000	0.0039	0.0076	0.0111	0.0143	0.0174
24	—	—	—	—	—	—	0.0000	0.0037	0.0071	0.0104
25	—	—	—	—	—	—	—	—	0.0000	0.0035

Source: From Shapiro and Wilk, 1965. Used by permission.
This table is used in Section 12.3.1

TABLE 2—Quantiles of the Shapiro-Wilk W Test for Normality (values of W such that $100p\%$ of the distribution of W is less than W_p).

n	$w_{0.01}$	$w_{0.02}$	$w_{0.05}$	$w_{0.10}$	$w_{0.50}$
3	0.753	0.756	0.767	0.789	0.959
4	0.687	0.707	0.748	0.792	0.935
5	0.686	0.715	0.762	0.806	0.927
6	0.713	0.743	0.788	0.826	0.927
7	0.730	0.760	0.803	0.838	0.928
8	0.749	0.778	0.818	0.851	0.932
9	0.764	0.791	0.829	0.859	0.935
10	0.781	0.806	0.842	0.869	0.938
11	0.792	0.817	0.850	0.876	0.940
12	0.805	0.828	0.859	0.883	0.943
13	0.814	0.837	0.866	0.889	0.945
14	0.825	0.846	0.874	0.895	0.947
15	0.835	0.855	0.881	0.901	0.950
16	0.844	0.863	0.887	0.906	0.952
17	0.851	0.869	0.892	0.910	0.954
18	0.858	0.874	0.897	0.914	0.956
19	0.863	0.879	0.901	0.917	0.957
20	0.868	0.884	0.905	0.920	0.959
21	0.873	0.888	0.908	0.923	0.960
22	0.878	0.892	0.911	0.926	0.961
23	0.881	0.895	0.914	0.928	0.962
24	0.884	0.898	0.916	0.930	0.963
25	0.886	0.901	0.918	0.931	0.964
26	0.891	0.904	0.920	0.933	0.965
27	0.894	0.906	0.923	0.935	0.965
28	0.896	0.908	0.924	0.936	0.966
29	0.898	0.910	0.926	0.937	0.966
30	0.900	0.912	0.927	0.939	0.967
31	0.902	0.914	0.929	0.940	0.967
32	0.904	0.915	0.930	0.941	0.968
33	0.906	0.917	0.931	0.942	0.968
34	0.908	0.919	0.933	0.943	0.969
35	0.910	0.920	0.934	0.944	0.969
36	0.912	0.922	0.935	0.945	0.970
37	0.914	0.924	0.936	0.946	0.970
38	0.916	0.925	0.938	0.947	0.971
39	0.917	0.927	0.939	0.948	0.971
40	0.919	0.928	0.940	0.949	0.972
41	0.920	0.929	0.941	0.950	0.972
42	0.922	0.930	0.942	0.951	0.972
43	0.923	0.932	0.943	0.951	0.973
44	0.924	0.933	0.944	0.952	0.973
45	0.926	0.934	0.945	0.953	0.973
46	0.927	0.935	0.945	0.953	0.974
47	0.928	0.936	0.946	0.954	0.974
48	0.929	0.937	0.947	0.954	0.974
49	0.929	0.937	0.947	0.955	0.974
50	0.930	0.938	0.947	0.955	0.974

Source: After Shapiro and Wilk, 1965.

The null hypothesis of a normal distribution is rejected at the α significance level if the calculated W is less than W_α.

This table is used in Section 12.3.1

TABLE 3—Quantiles of the t Distribution (values of t such that $100p\%$ of the distribution is less than t_p).

Degrees of Freedom	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.990}$	$t_{0.995}$
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.841
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.086	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576

Source: From Fisher and Yates, 1974. Used by permission.
This table is first used in Section 4.4.2

Subject Index

A
Alternative hypothesis, 48
  defining, 50
Ancillary tank equipment, definition, 25
Assessment phase, waste pile example, 60
ASTM D 1452, 20
ASTM D 1586, 19
ASTM D 1587, 20
ASTM D 4387, 19
ASTM D 4448, 19–20
ASTM D 4700, 19–20
ASTM D 4823, 20
ASTM D 5013, 20
ASTM D 5358, 20
ASTM D 5495, 20
ASTM D 5956, 24
ASTM D 6009, 55
ASTM D 6051, 23, 60
ASTM D 6232, 18

B
Boundaries, 5
  defining, 57
Box and whiskers plot, 62–63
40 CFR 260.10, 25
40 CFR 261.24, 6, 56

C
Chain of custody form, 29–30
Chain-of-custody procedures, samples, 29–31
Coefficient of variation, 62, 64
Comprehensive Liability Act of 1980, 29
Confidence intervals, 36, 43
Container, definition, 25
Contamination problem, description, 3

D
Data
  checking for normality, 34
  preparing for statistical analysis, 34
  transformations in statistical tests, 35–36
Data assessment phase, 33–37
Data collection, 16–29
  chain-of-custody procedures for samples, 29–31
  composite sampling, 20, 23
  field records and sample identification, 28
  field screening, 20
  heterogeneous waste, 23
  investigation derived waste, 26–28
  laboratory coordination, 16
  mobilization, 17–18
  particle size reduction, 25
  personnel decontamination, 26
  sample location selection, 19
  sample shipment, 28
  sampling equipment
    decontamination, 26
    selection, 18–19
  sampling waste units, 25
  site entry and reconnaissance, 16–17
  waste pile example, 57–60
Data life cycle, Data Quality Assessment, 33
Data Quality Assessment, 33
  conclusions and reports, 36
  data life cycle, 33
  data preparation for statistical analysis, 34
  overview, 33
  preliminary data analysis, 34–35
  review DQOs and sampling design, 33–34
  statistical assumption check, 34–35
  statistical testing, 35–36
Data Quality Objectives, 1–4
  boundaries, defining, 5
  composite sampling, 9
  data collection and design optimization, 6–9
  decision errors, specifying limits on, 6
  decision rule development, 5–6
  defining, 2
  identifying decisions, 3–4
  identifying inputs to decisions, 4
  impact of process on project, 2
  process and overall decision process, 3
  review, 33–34
    waste pile example, 60
  sample size
    estimating requirements, 8
    post-study assessment, 8
  sampling designs, see Sampling designs
  stating the problem, 3–4
  waste pile example, 55–57
DataQUEST, 33–35
Decision errors
  false positive and negative, 48, 50
  identifying, 48–50
  parameter values where consequences are minor, 51
  potential consequences, 50
  specifying tolerable limits, 6, 47–52
    background, 47–48
    determining possible range of parameter of interest, 48
    expected outputs, 47
    waste pile example, 57
  tolerable probability, 51
Decision rules, 5–6
  developing, 57
Disposable equipment, 26

E
EPA QA/G-4, 2, 6, 34
  see also Decision errors
EPA QA/G-9, 33–36, 60, 62–63
EPA SW-846, 25, 35–36, 56–57

F
False positive and negative, 48, 50
55 Federal Register 26990, 25
Field documentation, 28–31
Field records, 28–29
Field screening, 20

H
Histogram, 62
Hypothesis tests, 43

I
Implementation phase, see Data collection
Investigation derived waste, 26–28

M
Minimum detectable difference, 51
Mobilization, data collection, 17–18

N
Null hypothesis, 36, 48
  choosing, 48–50
  defining, 50

O
Optimal allocation, 8–9
Outliers, assessing, 35

P
Particle size reduction, 18, 25
Performance evaluation, 32
Personnel, decontamination, 26
Personnel protective equipment, 26
Planning phase, 2–15
  see also Data Quality Objectives
Population sampling, difficulties, 1
Proportional allocation, 8–9

Q
Quality Assurance Project Plan, 7
  sample locations, 19
Quantile-quantile probability plot, 35, 63

R
Reagents, used during decontamination, 25–26
Receipt for samples form, 29, 31
Resource Conservation and Recovery Act of 1976, 29
  waste characterization, 23

S

Samples
  chain-of-custody procedures, 29–31
  estimating number required, 8
    waste pile example, 58
  identification, 28–29
  post-study assessment of number, 8
  shipping, 28
Sampling designs, 9–14
  authoritative, 9, 14
    waste pile example, 57–64
  composite, 9, 14
    data collection, 20, 23–24
  not meeting Data Quality Objectives, 60
  probabilistic, 14
  probability, 13
  review, 33–34
  selection, 10–12
  simple random, 8, 14
    waste pile example, 58–59, 60
  stratified, 8–9, 14
    waste pile example, 67–68
  stratified systematic, waste pile example, 59–60
  systematic, 8, 14
  systematic grid
    with compositing, waste pile example, 66–67
    waste pile example, 59–60
    without compositing, waste pile example, 65–66
  waste pile example, 57–58
Sampling equipment, decontamination, 25–26
Shapiro-Wilk test for normality, 34, 63, 65
  coefficients of a^i, 70–71
  quantiles of t distribution, 73
  quantiles of W test, 72
Site
  components, 18
  entry, 16
  reconnaissance, 16–17
  work zones, 17
Statistical analysis, preparing data for, 34
Statistical outlier, 35
Statistical quantities, 34
Statistical tests
  data transformations, 35–36
  selecting and preforming, 35–36
Sump, definition, 25
Superfund, 29
Support zone, 18
Surveillance, 32

T

Tank, definition, 25
Technical assessments, 29, 32
Technical system audits, 32
Toxicity Characteristic Leaching Procedure, 25, 55–56
  versus total results, waste pile example, 62–63
Toxicity Characteristic Rule, 55–56

U

Upper confidence limit, 63–67

W

Walk-through, 18
Waste
  heterogeneous, data collection, 24
  investigation derived, 26–28
Waste pile, 24
  example, 55–68
    authoritative sampling design, 57–64
    cost of sampling, 59–60
    data collection, 57–60
    Data Quality Objectives, 55–57
    implementation phase, 60
    inputs to decision, 56–57
    non-normal data distribution, 64–65
    topographic base map, 55–56
Waste units
  containerized, 25
  sampling, 23
  uncontainerized, 25

About the editors

William M. Cosgrove is a science advisor to EPA's Director, Science and Ecosystem Support Division (SESD) in Athens, GA. SESD provides field, analytical, and quality assurance support to the Region 4 enforcement and permitting programs. Mr. Cosgrove has served as a technical authority for the RCRA program while conducting numerous field investigations for the civil and criminal enforcement programs and he has assisted in the development and instruction of RCRA training courses.

Mr. Cosgrove has a B.S. in Environmental Health Science from the University of Georgia, Athens, GA and an M.S. in Environmental Systems Engineering from Clemson University, Clemson, SC. He serves on the adjunct faculty at the University of Georgia, and is active with the Water Environment Federation and ASTM.

Michael P. Neill is an Environmental Scientist at the U.S. EPA in Athens, GA. He serves as senior project leader in the hazardous waste section's team of field investigation specialists, responsible for enforcement actions and other investigations involving RCRA.

Prior to joining EPA in Athens, Mr. Neill was an On-Scene Coordinator for the U.S. EPA in Edison, NJ directing CERCLA removal actions at metal plating facilities, asbestos dumps, solvent recovery sites and chemical manufacturing plants.

Mr. Neill received his B.S. in Geology from Virginia Tech, Blacksburg, VA and his M.S. in Environmental Sciences from Rutgers University, New Brunswick, NJ.

Katharine Hastie is currently employed as a consultant at Booz Allen & Hamilton, Inc. One of her primary responsibilities at Booz Allen involves supporting the U.S. EPA, Region 9 with program management for RCRA permitting and enforcement. Previously, she worked in the EPA, Region 4 RCRA Enforcement Program conducting field investigations in support of criminal and civil enforcement activities. Additionally, she served as an instructor in the Region 4 hazardous waste training program.

Ms. Hastie received her B.S. in Physics from University of the South, Sewanee, TN and her M.S. in Environmental Engineering from Clemson University, Clemson, SC.

ISBN 0-8031-2085-0