Designation: D 5792 – ~~95~~02

Standard Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives[1]

This standard is issued under the fixed designation D 5792; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice covers the process of development of data quality objectives (DQOs) for the acquisition of environmental data. Optimization of sampling and analysis design is a part of the DQO process. This practice describes the DQO process in detail. The various strategies for design optimization are too numerous to include in this practice. Many other documents outline alternatives for optimizing sampling and analysis design. Therefore, only an overview of design optimization is included. Some design aspects are included in the practice's examples for illustration purposes.

1.2 DQO development is the first of three parts of data generation activities. The other two aspects are (*1*) implementation of the sampling and analysis strategies, see Guide D 6311 and (*2*) data quality assessment, see Guide D 6233.

1.3 This guide should be used in concert with Practices D 5283, ~~which~~ D 6250, and Guide D 6044. Practice D 5283 outlines the quality assurance (QA) processes specified during planning and used during implementation. Guide D 6044 outlines a process by which a representative sample may be obtained from a population, identifies sources that can affect representativeness and describes the attributes of a representative sample. Practice D 6250 describes how a decision point can be calculated.

1.~~3~~4 Environmental data related to waste management activities include, but are not limited to, the results from the sampling and analyses of air, soil, water, biota, process or general waste samples, or any combinations thereof.

1.~~4~~5 The DQO process is a planning process and should be ~~developed and initiated~~ completed prior to ~~the application of planning, implementation, and assessment of~~ sampling and analysis activities.

1.~~5~~6 This practice presents extensive requirements of management, designed to ensure high-quality environmental data. The words "must" and "shall" (requirements), "should" (recommendation), and "may" (optional), have been selected carefully to reflect the importance placed on many of the statements in this practice. The extent to which all requirements will be met remains a matter of technical judgement.

1.~~6~~7 The values stated in SI units are to be regarded as the standard. The values given in parentheses are for information only.

1.~~7~~8 *This standard 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 standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*

C ~~970 Guide~~ 1215 Guide for ~~Sampling Special Nuclear Materials~~ Preparing and Interpreting Precision and Bias Statements in

---

[1] This practice is under the jurisdiction of ASTM Committee D ~~34~~ on Waste Management and is the direct responsibility of Subcommittee D34.0~~2~~1 on Physical and Chemical Characterization.

Current edition approved ~~Nov.~~ Feb. 10, ~~1995.~~ 2002. Published ~~January 1996.~~ May 2002. Originally published as D 5792–95. Last previous edition D 5792–95.

Multi-Container Lots Test Method Standards Used in the Nuclear Industry[2]
C 1215 Guide for Preparing and Interpreting Precision and Bias Statements in Test Method Standards Used in the Nuclear Industry[2]
D 5283 Practice for Generation of Environmental Data Related to Waste Management Activities: Quality Assurance and Quality Control Planning and Implementation[3]
D 6044 Guide for Representative Sampling for Management of Waste and Contaminated Media[3]
D 6233 Guide for Data Assessment for Environmental Waste Management Activities[3]
D 6250 Practice for Derivation of Decision Point and Confidence Limit for Statistical Testing of Mean Concentration in Waste Management Decisions[3]
D 6311 Guide for Generation of Environmental Data Related to Waste Management Activities: Selection of Optimization of Sampling Design[3]

3. Terminology

3.1 *Definitions*—Where applicable, the originating reference is associated with the definition and follows the definition in boldface type. Definitions:

3.1.1 *accuracy (see bias)*—(*1*) bias; (*2*)bias, *n*—the difference between the closeness of a measured sample value to the true value; (*3*) the closeness of a measured value to the test results and an accepted reference or standard value.

3.1.1.1 *Discussion*—For many investigators, *accuracy* is attained only if a procedure is both precise and unbiased (see *bias*). Because this blending of *precision* and *accuracy* can lead to confusion, ASTM requires a statement on *bias* instead of *accuracy*.

D 5283

3.1.2 *action level*—the numerical value that causes the decision maker to choose one of the alternative actions (for example, compliance or noncompliance). It may be a regulatory threshold standard, such as maximum contaminant level for drinking water, a risk-based concentration level, a technological limitation, or reference-based standard.

EPA QA/G-4 (1)[4]

3.1.3 *bias (see accuracy)*—the difference between the population mean of the test results and an accepted reference value.

[2] *Annual Book of ASTM Standards*, Vol 12.01.
[3] *Annual Book of ASTM Standards*, Vol 11.04.

3.1.3.1 *Discussion*—*Bias* represents a constant error as opposed to a *random error*. A method *bias* can be estimated by the difference (or relative difference) between a measured average and an accepted standard or reference value. The data from which the estimate is obtained should be statistically analyzed to establish *bias* in the presence of *random error* . A thorough *bias* investigation of a measurement procedure requires a statistically designed experiment to repeatedly measure, under essentially the same conditions, a set of standards or reference materials of known value that cover the range of application. *Bias* often varies with the range of application and should be reported accordingly. **C 1215, D 5283**

3.1.4 **C 1215**

3.1.2 *confidence interval, n*—an interval used to bound the value of a population parameter with a specified degree of confidence (this is an interval that has different values for different samples).

3.1.42.1 *Discussion*—When providing a *confidence interval*, analysts should give the number of observations on which the interval is based. The—The specified degree of confidence is usually 90, 95, or 99 %. The form of a *confidence interval* depends on underlying assumptions and intentions. *Confidence intervals* are usually taken to be symmetric, but that is may or may not necessarily so, as in be symmetric about the case of mean, depending on the underlying statistical distribution. For example, *confidence intervals* for the variances are not symmetric. **C 1215**

3.1.53 *confidence level, n*—the probability, usually expressed as a percent, that a *confidence interval* willis expected to contain the parameter of interest (see discussion of *confidence interval*).

3.1.64 *data quality objectives (DQOs), n*— qualitative and quantitative statements derived from the DQO process describing the decision rules and the uncertainties of the decision(s) within the context of the problem(s).

3.1.64.1 *Discussion*—DQOs clarify the study objectives, define the most appropriate type of data to collect, determine the most appropriate conditions from which to collect the data, and establish acceptable levels of decision errors that will be used as the basis for establishing the quantity and quality of data needed to support the decision. The DQOs are used to develop a sampling and analysis design.

3.1.75 *data quality objectives process, n*— a quality management tool based on the scientific method and developed by the U.S. Environmental Protection Agency (EPA) to facilitate the planning of environmental data collection activities. The DQO process enables planners to focus their planning efforts by specifying the use of the data (the decision), decision criteria (action level), (decision point), and decision maker's acceptable decision error rates. The products of the DQO process are the DQOs.

3.1.75.1 *Discussion*—DQOs result from an iterative process between the decision makers and the technical team to develop qualitative and quantitative statements that describe the problem and the certainty and uncertainty that decision makers are willing to accept in the results derived from the environmental data. This acceptable level of uncertainty should then be used as the basis for the design specifications for project data collection and data assessment. All of the information from the first six steps of the DQO process are used in designing the study and assessing the data adequacy. **EPA QA/G-4**

3.1.86 *decision error*

3.1.86.1 *false negative error, n*—this occurs when environmental data mislead decision maker(s) into not taking action specified by a decision rule when action should be taken.

3.1.86.2 *false positive error, n*—this occurs when environmental data mislead decision maker(s) into taking action specified by a decision rule when action should not be taken.

3.1.97 *decision point, n*—the numerical value that causes the decision-maker to choose one of the alternative actions point (for example, compliance or noncompliance). **D 6250**

3.1.7.1 *Discussion*—In the context of this practice, the numerical value is calculated in the planning stage and prior to the collection of the sample data, using a specified hypothesis, decision error, an estimated standard deviation, and number of samples. In environmental decisions, a concentration limit such as a regulatory limit usually serves as a standard for judging attainment of cleanup, remediation, or compliance objectives. Because of uncertainty in the sample data and other factors, actual cleanup or remediation, may have to go to a level lower or higher than this standard. This new level of concentration serves as a point for decision-making and is, therefore, termed the *decision point*.

3.1.8 *decision rule, n*—a set of directions in the form of a conditional statement that specify the following: (*1*) how the sample data will be compared to the action level, decision point, (*2*) which decision will be made as a result of that comparison, and (*3*) what subsequent action will be taken based on the decisions.

3.1.109 *precision, n*—a generic concept used to describe the dispersion of a set of measured values.

3.1.109.1 *Discussion*—It is important that some quantitative measure be used to specify *precision*. A statement such as "the *precision* is 1.54 g" is useless. Measures—Measures frequently used to express *precision* are standard deviation, relative standard deviation, variance, repeatability, reproducibility, confidence interval, and range. In addition to specifying the measure and the *precision*, it is important that the number of repeated measurements upon which the estimated *precision* is based also be given. **D 5283**

3.1.11

3.1.10 *quality assurance (QA), n*—an integrated system of management activities involving planning, quality control, quality assessment, reporting, and quality improvement to ensure that a process or service (for example, environmental data) meets defined standards of quality with a stated level of confidence. **EPA QA/G-4**

3.1.121 *quality control (QC), n*—the overall system of technical activities whose purpose is to measure and control the quality

of a product or service so that it meets the needs of users. The aim is to provide quality that is satisfactory, adequate, dependable, and economical. **EPA QA/G-4**

3.1.12 *population, n*—the totality of items or units of materials under consideration.

3.1.13 *random error, n*—(*1*) the chance variation encountered in all measurement work, characterized by the random occurrence of deviations from the mean value; (*2*) an error that affects each member of a set of data (measurements) in a different manner. **D 5283**

3.1.14 *risk, n*—the probability or ~~expectation that~~ an expected loss associated with an adverse ~~effect will occur.~~ effect.

3.1.14.1 *Discussion—Risk* is frequently used to describe the adverse effect on health or on economics. Health-based *risk* is the probability of induced diseases in persons exposed to physical, chemical, biological, or radiological insults over time. This *risk* probability depends on the concentration or level of the insult, which is expressed by a mathematical model describing the dose and *risk* relationship. *Risk* is also associated with economics when decision makers have to select one action from a set of available actions. Each action has a corresponding cost. The *risk* or expected loss is the cost multiplied by the probability of the outcome of a particular action. Decision makers should adopt a strategy to select actions that minimize the expected loss.

3.1.15 *sample standard deviation, n*—the square root of the sum of the squares of the individual deviations from the sample average divided by one less than the number of results involved.

$$S = \sqrt{\frac{\sum_{j=1}^{n} (X_j - \bar{X})^2}{n-1}}$$

where:

S = sample standard deviation,
n = number of results obtained,
X_j = jth individual result, and
$\bar{X}$ = sample average.

~~3.1.15.1 *Discussion*—The use of the *standard deviation* to describe *precision* implies that the uncertainty is independent of the measurement value. The practice of associating the ± symbol with *standard deviation* (or RSD) is not recommended. The ± symbol denotes an interval. The *standard deviation* is not an interval, and it should not be treated as such.~~ **D 5283**

4. Summary of Practice

4.1 This practice describes the process of developing and documenting the DQO process and the resulting DQOs. This practice also outlines the overall environmental study process as shown in Fig. 1. It must be emphasized that any specific study scheme must be conducted in conformity with applicable agency and company guidance and procedures.

4.2 For example, the investigation of a Superfund site would include feasibility studies and community relations plans, which are not a part of this practice.

5. Significance and Use

5.1 Environmental data are often required for making regulatory and programmatic decisions. Decision makers must determine whether the levels of assurance associated with the data are sufficient in quality for their intended use.

5.2 Data generation efforts involve three parts: development of DQOs and subsequent project plan(s) to meet the DQOs, implementation and oversight of the project plan(s), and assessment of the data quality to determine whether the DQOs were met.

5.3 To determine the level of assurance necessary to support the decision, an iterative process must be used by decision makers, data collectors, and users. This practice emphasizes the iterative nature of the process of DQO development. Objectives may need to be reevaluated and modified as information related to the level of data quality is gained. This means that DQOs are the product of the DQO process and are subject to change as data are gathered and assessed.

5.4 This practice defines the process of developing DQOs. Each step of the planning process is described.

5.5 This practice emphasizes the importance of communication among those involved in developing DQOs, those planning and implementing the sampling and analysis aspects of environmental data generation activities, and those assessing data quality.

5.6 The impacts of a successful DQO process on the project are as follows: (*1*) a consensus on the nature of the problem and the desired decision shared by all the decision makers, (*2*) data quality consistent with its intended use, (*3*) a more 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 for making an incorrect decision.

6. Data Quality Objective Process

6.1 The DQO process is a logical sequence of seven steps that leads to decisions with a known level of uncertainty (Fig. 1). It is a planning tool used to determine the type, quantity, and adequacy of data needed to support 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 seven steps are as follows:

(1) Stating the problem,

(2) Identifying possible decisions,

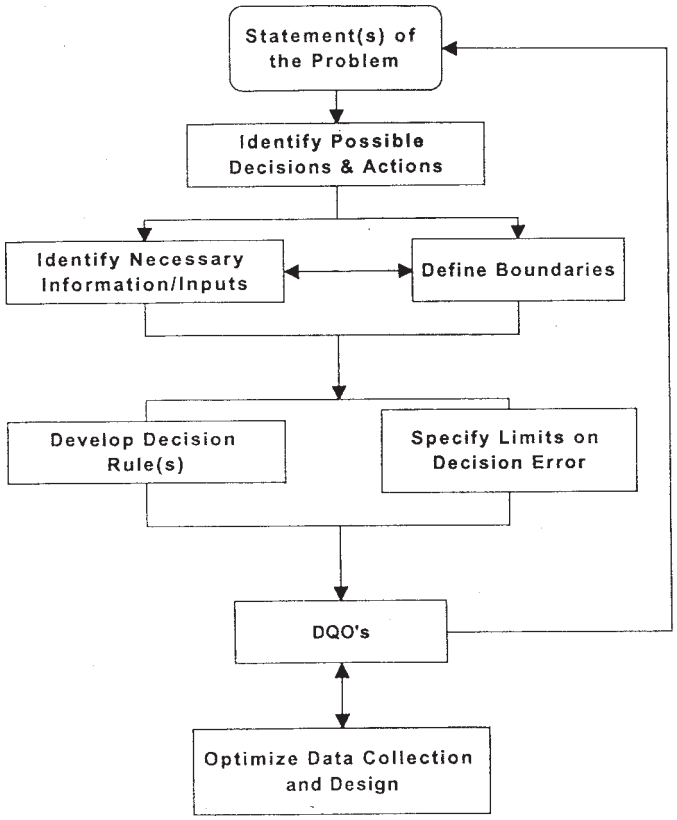

FIG. 1 DQO Process

(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 specified levels of the decision errors (from the sixth step). In the last step of the process, various approaches to a sampling and analysis plan for the project are developed that allow the decision makers 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 project process is shown in Fig. 2. At any stage of the project or during the field implementation phase, it may be appropriate to reiterate the DQO process, beginning with the first step based on new information. See Refs **(2, 3)** for examples of the DQO process.

6.2 *Step 1—Stating the Problem* :

6.2.1 *Purpose*—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 roles of key decision makers and technical qualifications of the problem-solving team may not be provided with the first notice. Only after the appropriate information and problemsolving team are assembled can a clear statement of the problem be made.

6.2.2 *Activities*:

6.2.2.1 *Assembling of all Pertinent Information*—The necessary first action to describe a problem is to verify the conditions that indicate a problem exists. The pertinent information should be assembled during this phase of problem definition. A key source is any historical record of events at the site where the problem is believed to exist. This enables the decision makers to understand the context of the problem. A series of questions need to be developed concerning the problem.

(1) What happened (or could happen) that suggests a problem?
(2) When did it (could it) happen?
(3) How did it (could it) happen?
(4) Where did it (could it) happen?
(5) Why did it (could it) happen?
(6) How bad is (might be) the result or situation?

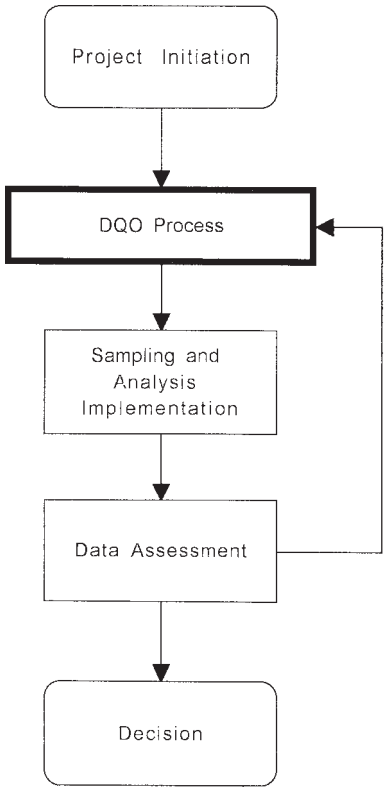

FIG. 2 DQOs Process and Overall Decision Process

(7) How fast is (might be) the situation changing?
(8) What is (could be) the impact on human health and the environment?
(9) Who was (could be) involved?
(10) Who knows (should know) about the situation?
(11) Has anything been (might anything be) done to mitigate the problem?
(12) What contaminants are (could be) involved?
(13) How reliable is the information?
(14) What regulations could or should apply?
(15) Is there any information that suggests there is not a problem?

This list of potential information is not exhaustive, and there may be other data applicable to the definition of the problem.

6.2.2.2 *Identification of the DQO Team*— Even as information is being gathered, it is necessary to begin assembling a team of decision makers and technical support personnel to organize and evaluate the information. These individuals become the core of the DQO team and may be augmented by others as information and events dictate. The ~~identities and roles of the DQO team members are usually determined by the~~ decision makers who have either jurisdiction over the site and personnel or financial resources that will be used in resolving the problem usually determine the identities and roles of the DQO team members. The DQO team is usually made up of the following key individuals:

(1) Site Owners or Potentially Responsible Parties—These individuals have authority to commit personnel and financial resources to resolve the problem and have a vital interest in the definition of the problem and possible decisions.

(2) Representatives of Regulatory Agencies—These individuals are usually responsible for enforcing the standards that have been exceeded, leading to classifying the observations or events as a problem. Additionally, they have an active role in characterizing the extent of the problem, approving any proposed remedial action, and concurring that the action mitigated the problem.

(3) Project Manager—This individual generally has the responsibility for overseeing resolution of the problem. This person may represent either the regulatory agency or the potentially responsible parties.

(4) Technical Specialists—These individuals have the expertise to assess the information and data to determine the nature and extent of the potential problem and may become key players in the design and implementation of proposed decisions.

It is important that these individuals be assembled early in the process and remain actively involved to foster good communications and to achieve consensus among the DQO team on important decision-related issues.

6.2.3 *Outputs*:

6.2.3.1 *Statement of Problem and Context*— Once the initial information and data have been collected, organized, and evaluated, the conclusions of the DQO team should be documented. If it is determined that no problem exists, the conclusion must

be supported by a summary of the existing conditions and the standards or regulatory conditions that apply to the problem.

(1) If a problem is found to exist, the reasons must be stated clearly and concisely. Any standards or regulatory conditions that apply to the situation must be cited. If the initial investigation concludes that the existing conditions are the result of a series of problems, the DQO team should attempt to define as many discrete problems (or issues) as possible.

(2) The following are examples of problem statements:

(*a*) A former pesticide formulation facility is for sale, but it is unknown whether it meets local environmental standards for property transfer.

(*b*) An industrial site is known to be contaminated with low levels of lead, but it is unknown whether levels are below risk-based standards.

(*c*) Most of a vacant lot is believed to be uncontaminated with PCBs (<2 ppm), but it is unknown whether abandoned, leaky transformers in the vacant lot make it necessary to remove any of the top layer of soil.

(*d*) The former industrial site has contaminated soil areas that may be contaminating ground water, and it is necessary to decide which type of monitoring program will satisfy local health requirements.

(*e*) The city would like to use local ground water on an athletic field near a Superfund site, but must know how this water will impact the health of the athletes and spectators.

(3) Complex problems should be broken down into manageable smaller problems that are linked together to form the final decision. As an example, the sale of a piece of property may involve solving the following problems:

(*a*) Is the site contaminated? If yes, then,

(*b*) Is off-site disposal required? If no, then

(*c*) Which of two allowable on-site treatment options should be used?

6.2.3.2 *Identification of Resources*—As the nature and magnitude of the problem is being documented, the decision makers should be conferring to determine the type and amount of resources that can be committed. Preliminary budget, personnel assignments, and schedule should be established. Preliminary milestones, timelines, and approvals should be documented and concurred upon by affected decision makers. The DQO team leader and technical specialists should be included in these discussions where possible. At a minimum, they should be kept informed of these issues so their impact can be anticipated in the definition of the problem.

(1) Fig. 3 shows the primary components of the problem statement step. After this step is completed, the DQO team moves on to the next step, where the process to resolve the problem continues.

(2) It is important to remember that the DQO process is an iterative one. New information is collected as projects proceed. The DQO team members associated with the problem-statement step should remain involved with the DQO process. If new data, unavailable to the DQO team during the development of the problem statement, demonstrates that the statement is incomplete or otherwise inadequate, the problem statement should be reconsidered.

6.3 *Step 2—Identifying Possible Decisions*:

6.3.1 *Purpose*—The purpose of this step is to identify the possible decision(s) that will address the ~~problem once it has been stated clearly.~~ problem. Multiple decisions are required when the problem is complex. Information required to make decisions and

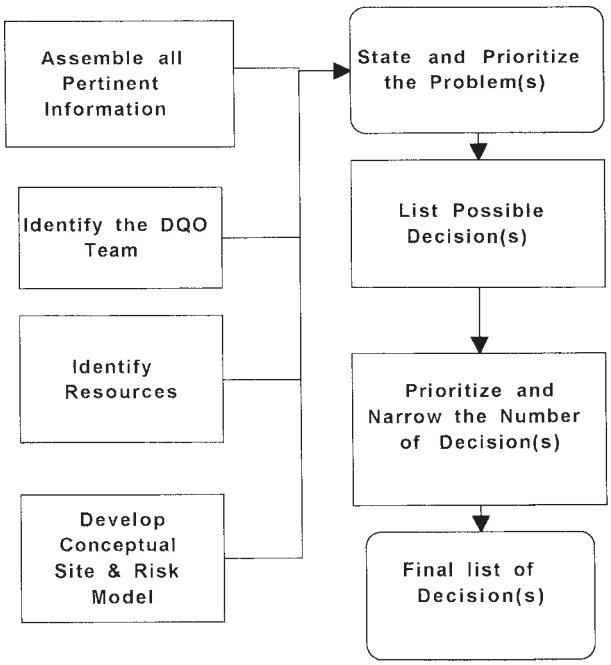

FIG. 3 Stating the Problem and Identifying the Decisions

to define the domain or boundaries of the decision will be determined in later steps (6.4 and 6.5, respectively). Each potential decision is tested 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. The activities that lead to identifying the decision(s) are shown in Fig. 3 and discussed in 6.3.2.

6.3.2 *Activities*:

6.3.2.1 *Listing of Possible Questions Leading to Decisions* —All possible decisions concerning the problem should be listed. Choices should not be eliminated at this time. Possible decision statements are presented in the form of a series of questions that, when answered, result in actions that will resolve the problem. Examples of questions related to problems given in 6.2.3 (Step 1) are as follows:

(1) Are possible contaminants on the site below regulatory thresholds?

(2) Must all of the surface soil be remediated to less than 5 ppm lead?

(3) Can only locations with PCB levels above 2 ppm be remediated?

(4) Will a ground water monitoring program at the site capable of detecting contaminants at the 5-ppm level satisfy regulatory requirements?

(5) Will a single monitoring point on or near the athletic field be sufficient?

6.3.3 *Output*—After all possible decisions that might be made have been documented, those determined to be most appropriate to resolve the problem should be prioritized by the DQO team in decreasing order of level of effort (available resources and technical challenge). Justification for the rankings should be provided. The recommended sequence in which the decisions are made should also be listed. In cases in which a complex decision statement has been broken down into a series of simpler decisions, the DQO team should identify whether the individual decisions should be addressed sequentially or in parallel. After the possible decisions have been identified, the DQO team focuses on gathering the information necessary to formulate the decision statements in Step 3 (6.4).

6.4 *Step 3—Identifying Inputs to Decisions*:

6.4.1 *Purpose*—The answers to each of the questions identified by the previous step in the DQO process must be resolved with data. Fig. 4 shows the key activities that lead to development of the data requirements. 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 “define boundaries” in Fig. 1. This is another type of data requirement and is discussed in 6.4.

6.4.2 *Activities*:

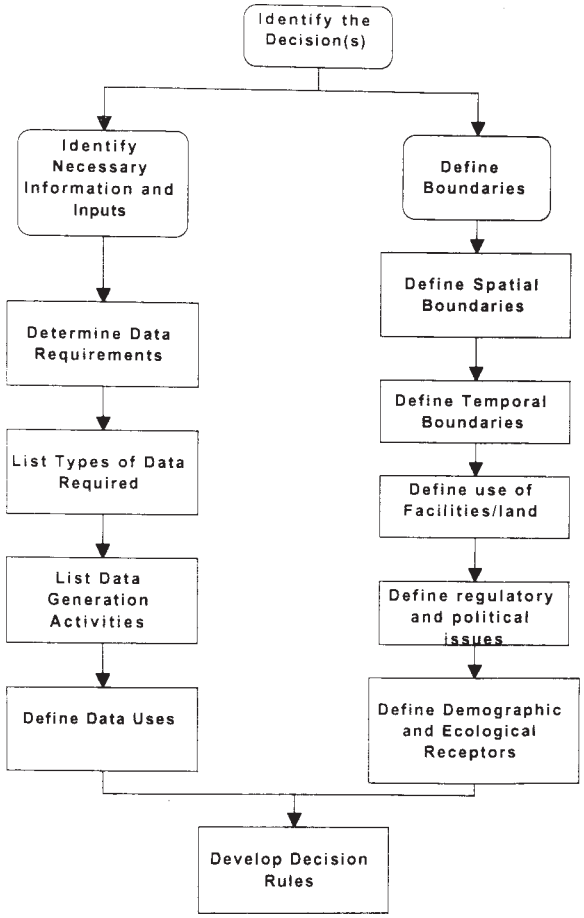

FIG. 4 Determination of Information Inputs and Study Boundaries

6.4.2.1 *Determination of Data Requirements*— At this stage of the process, it is important to carefully examine the complete set of data requirements needed to support each of the decisions. Each possible decision to be made should be considered independently of others to ensure that no omissions have occurred. After all possible questions concerning the decisions have been considered, group the data requirements together to determine overall data needs for the project. It may be possible to plan efficiencies in collecting and processing data to meet multiple needs and thereby lower overall project costs or reduce the time necessary to meet important milestones, or both.

(*1*) When considering whether specific information is needed for making a decision, test the data to ensure that it is appropriate for the decision statement. If no use of the data can be identified, it may be extraneous to the needs.

(*2*) The following list is indicative of some of the information needs that may be considered for each decision. It is not inclusive of all important data, but it provides examples common to many environmental problems.

(a) ~~(*a*) What~~What regulatory limits may be associated with the problem or regulatory issue?
(b) ~~(*b*) Does~~Does contamination exceed regulatory limits?
(c) ~~(*c*) What~~What tests must be performed for the type of waste in question?
(d) ~~(*d*) What~~What are the hydrogeological considerations?
(e) ~~(*e*) What~~What populations are at risk?
(f) ~~(*f*) What~~What are the ecological considerations?
(g) ~~(*g*) What~~What process knowledge is available?
(h) ~~(*h*) What~~What historical/background data (past uses or spills) are available?
(i) ~~(*i*) What~~What are the budget constraints?
(j) ~~(*j*) What~~What is the time schedule?
(k) ~~(*k*) What~~What potential health, political, and social factors must be considered?
(l) ~~(*l*) What~~What is the potential for legal action?
(m) ~~(*m*) Who~~Who is the end-user of the data?
(n) ~~(*n*) What~~What data validation criteria will be used?
(o) ~~(*o*) What,~~What, if any, limitations exist on the data collection process (detection limits, matrix interferences, or no known measurement technology)?

6.4.3 *Outputs*:

6.4.3.1 The DQO team must specify data needs for each problem/decision that has been identified in the first two steps.

6.4.3.2 List the types of data required. Some example data types include, but are not limited to, the following:

(1) Chemical,
(2) Physical (including site hydrogeology and meteorology),
(3) Biological,
(4) Toxicological,
(5) Historical,
(6) Economic (time, budget, and manpower),
(7) Demographic,
(8) Toxicity characteristics, and
(9) Fate and transport model output.

6.4.3.3 *Listing of Data Generation Activities*—Determine which data can be acquired from historical records and which new data must be obtained in the field or laboratory, or both. If the DQO team determines that no new data are necessary to make a decision, they should document their reasoning. If new information is necessary, activities that will be required to generate inputs (data) affecting the decision should be listed. Examples of these include, but are not limited to, the following:

(1) Assembly of historical data,
(2) Sampling and chemical analysis,
(3) Physical testing, and
(4) Modeling.

6.4.3.4 *Definition of Data Use(s)*—Each set of data will be used for some purpose. This purpose must be defined. For example, will ~~action levels~~ regulatory thresholds for contaminants be determined by a risk-based calculation, by reference dose, or by pre-defined threshold values established by regulators? If so, ensure that data requirements are consistent with the criteria against which they will be compared. Data collected at the parts per million level may not be useful if they are to be compared to criteria at the parts per billion level.

6.5 *Step 4—Defining Boundaries* :

6.5.1 *Purpose*—This step of the DQO process determines the boundaries to which the decisions will apply. Boundaries establish limits on the data collection activities identified in Step 3 (6.4). These boundaries include, but are not limited to, spatial boundaries (physical and geographical), temporal boundaries (time periods), demographic, regulatory, political, and budget. The activities for this step of the DQO process are shown in Fig. 4.

6.5.2 *Activities*:

6.5.2.1 *Definition of Spatial Boundaries*— Define the boundaries of the total area and smallest increment of concern. Examples

of items affecting the boundary definition are as follows:

(1) Horizontal or lateral areas,

(2) Vertical boundaries (depth/height),

(3) Discrete locations (hot spots),

(4) Media/matrix (air, soil, water, biota, and waste),

(5) Number of containers of waste, and

(6) Volume.

6.5.2.2 *Definition of Temporal Boundaries (Time Period)* —This activity determines the time interval over which environmental data will be collected for use in the decision-making process. If current or future real-time data are used to represent or model previous conditions, the basis of these assumptions or models must be documented and agreed upon between the decision makers and the technical team. The same constraint is also placed on the extrapolation of historical or real-time data, or both, to future time periods.

(*1*) The duration of new data collection activities must be established. In addition, the following factors should be considered:

(*a*) Availability and reliability of existing historical data,

(*b*) Access to the site or impacted area,

(*c*) Exposure potential, and

(*d*) Budgetary constraints.

6.5.2.3 *Definition of the Demographic Receptors*—The DQO team must frequently define the receptor population that may be effected. All affected populations and the mode of their anticipated exposure should be identified. These populations include the following:

(1) Known/Anticipated Population(s)—Human (children, adults, age, gender, and so forth), plant/animal (wetlands, endangered species, and so forth), and global;

(2) Population activity patterns; and

(3) Exposure pathway for each population.

6.5.2.4 *Definition of Nontechnical Boundaries*—Decision makers also have to consider nontechnical boundaries that can impact the resolution of the problem seriously. These nontechnical boundaries include the following:

(1) Regulatory considerations, and

(2) Political or legal action(s).

6.5.3 *Outputs*—The results from each of the activities in this step must be documented. Care must be taken to identify which boundary conditions apply to each decision being made. It may be that similar information is needed for several decisions but different boundary conditions may apply. It is important that decision makers understand and concur on the boundaries; otherwise, the ability to make decisions may be compromised.

6.6 *Step 5—Developing Decision Rules*:

6.6.1 *Purpose*:

6.6.1.1 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 the ~~action level,~~ regulatory threshold or to the decision point,

(2) Which decision(s) will be made as a result of that comparison, and

(3) What subsequent action(s) will be taken based on the decisions.

Greater details on how a decision rule is formulated can be found in Practice D 6250.

6.6.1.2 The formats for these rules are either "if (criterion) ..., then (action)" statements or a decision tree, as shown in Fig. 5. The decision criteria should be stated as clearly and concisely as possible. The rule(s) must contain both a decision point (or ~~action level~~) decision point) and an action. The decision rule is generated through a cooperative effort among the DQO team. If an acceptable decision rule cannot be formulated, the process returns to the appropriate previous step of the DQO process.

6.6.1.3 Decision rules usually contain the following elements: measurement of interest, sample statistic, ~~action level,~~ decision point, and a resultant action. "Measurement of interest" is the variable or attribute to be measured. It can be concentration of a contaminant, volume/mass of a waste, or physical property, such as flash point of a waste. "Sample statistic" is the quantity computed from the sample data. It can be average value, median, present/absent, or some other expression of quantity. If that data are not normally distributed, statistical methods based on other distributions or non-parametric methods can be used.

6.6.1.4 The ~~"action level"~~ "decision point" is the limit against which the sample statistic will be compared (see X1.2.7.5 for example). Depending on whether the ~~action level~~ decision point is exceeded or not, the specified action will result. If the ~~action level~~ decision point equals the regulatory threshold, the probability of a false positive error equals the probability of a false negative error. For unequal probabilities of the decision errors, the ~~action level can be~~ decision point can be either less or greater than the regulatory threshold. The degree to which the ~~action level~~ decision point is different from the regulatory threshold depends on the acceptable level of uncertainty for the decision errors that the decision makers are willing to accept. The ~~action level is determined by the~~ levels of false positive error, false negative error, ~~measurement~~ variability, and number of samples determine the decision point. Derivation of ~~an action level~~ a decision point for a given level of false positive and false negative error is included as part of Appendix X1.

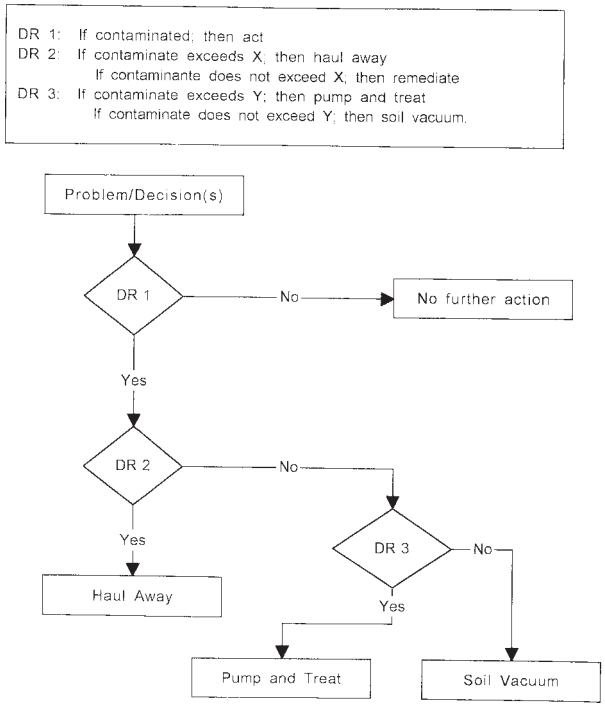

FIG. 5 Decision Tree for Three Sequential Decision Rules (DRs)

6.6.1.5 The decision rule is completed by stating the "resultant action" to be taken based on comparison of the sample statistic with the ~~action level.~~

~~6.6.1.6 Two illustrations~~ decision point.

6.6.1.6 An illustration of general decision rule formats are as follows:

(1) "If the average concentration of a contaminant in waste is greater than the ~~action level~~ decision point for that contaminant, then the waste will be classified as a 'hazardous' waste and will be disposed of according to the governing regulations."

(2) "If the average concentration of a contaminant in a waste is lower than the ~~action level~~ decision point for that contaminant, then the waste is classified as 'nonhazardous' and there are no special limitations placed on the disposal options.

6.6.1.7 In this illustration, the measurement of interest is "concentration of a contaminant." The sample statistic is the "average concentration." The ~~action level~~ decision point is some value to be specified. The resultant action is "disposal according to governing regulations." There may be separate decision rules for each medium, each domain (site), or other designated collections of data.

6.6.1.8 The ~~action level~~ decision point may be an observation or occurrence in some cases. An example of this type of decision rule is as follows:

(*1*) If soil exhibits a visible dark spot as compared to the surrounding soil, use the portable organic monitor to screen for organics in the dark spot.

6.6.2 *Activities*—The activities that must be completed to establish a decision rule are: specification of a regulatory threshold, agreement on acceptable false positive and false negative error rates, estimation of a sample standard deviation, calculation of the sample statistic and the decision point, and specification of alternative actions as a result of the decision. After these activities have been completed, a decision performance curve can be graphed as in Fig. 6. Decision performance curve is discussed in 6.7.2.5 and X1.2.8.1.

6.6.2.1 *Determination of Measurement of Interest*—A clear expression of the measurement (parameter) upon which the decision is based must be provided.

6.6.2.2 *Specification of ~~Action Level~~ Decision Point*— The ~~sample statistic of the measurement or observation of interest that initiates the agreed-upon action must be specified. The~~ determination of the ~~action level~~ decision point for any decision is a combination of the total variability in the data acquisition process and the level of decision errors that decision makers will accept in the final decision. The role of decision makers and decision errors is discussed in 6.7 (Step 6), and the derivation of ~~an action level~~ a decision point is illustrated in Appendix X1.

6.6.2.3 *Specification of Sample Statistic (if Applicable)* —Prior to the statement of a decision rule, it is necessary to determine how the sample statistic will be calculated and expressed (units of measure). The statistical approach chosen can be the ~~average,~~ mean, median, high, low, range, present/absent, and so forth. The unit of measurement must correspond to those of the decision criteria, and the limit of detection (measurement) must be lower than the ~~action level. A statistic may or may not be applicable in stating observations.~~ decision point.

6.6.2.4 *Specification of Mode of Comparison*—After the sample statistic is derived from historical or ~~real-time~~ new sample data

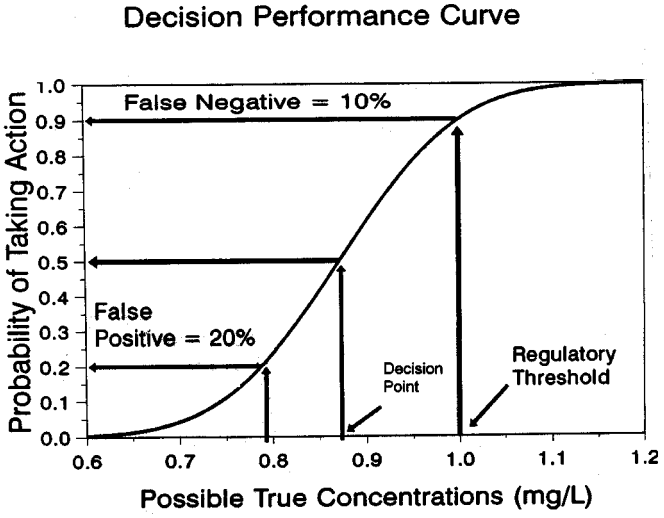

FIG. 6 Decision Rule Development

and ~~an action level~~ a decision point has been identified, they must be compared. This comparison is usually stated as greater than ..., less than ..., ~~equal to ...,~~ or ~~present/absent.~~ equal to. Depending on the results of the comparison, a specific action is indicated by the decision rule.

6.6.2.5 *Specification of Action*—When the result of the comparison of the sample statistic with the ~~action level~~ decision point is known, an action ~~must be specified.~~ will result. It should be sufficient to resolve the problem. In complex situations, the action may direct decision makers to another problem (addressed by an additional set of DQOs) that must also be resolved. This type of logical pathway is described frequently as a decision tree. These situations should have been identified in Step 2 (6.3). Fig. 5 shows the decision tree derived from the application of a set of three sequential decision rules.

6.6.3 *Outputs*—An example showing the application of a decision rule is presented in Appendix X1. Some additional examples of decision rules that might apply to waste problems and possible actions discussed in 6.2 and 6.3, respectively, are given as follows:

6.6.3.1 If the historical record of site monitoring activities shows the absence of any regulated constituent above 1 ppm, then the site can be left as is.

NOTE 1—A value of 1 ppm selected for this example only.

6.6.3.2 If site characterization indicates that 20 % of the soil (top 30 cm) is contaminated above 5 ppm lead, then the entire soil layer (1 m) must be remediated.

6.6.3.3 If site characterization data show that 95 % of the total surface area (10 cm deep) of the site contains less than 2 ppm PCB, then only those areas exceeding that value need to be remediated.

6.6.3.4 If the levels of contaminants found in the monthly ground water monitoring program total less than 1000 ppm in each well, then no additional corrective action needs to be instituted.

6.6.3.5 If no contaminate above 1 ppm is observed in a ground water monitoring well located downgradient and within 100 m of the site boundary during monthly monitoring events, then additional monitoring wells will not be required.

6.7 *Step 6—Specifying Limits on Decision Errors*:

6.7.1 *Purpose*—An essential part of the DQO process is to establish the degree of uncertainty (decision errors) that decision makers are prepared to accept in making a decision concerning the problem (Refs ~~(Refs 4-6)~~4-6). The purpose of this step is to define the acceptable decision errors based on a consideration of the consequences of making an incorrect decision. The perspective of the decision makers or baseline assumption must be stated clearly, that is, the site is considered contaminated or the site is not ~~contaminated. There are two kinds of decision errors: false positive error and false negative error.~~ contaminated (see Practice D 6250).

6.7.2 *Activities*:

6.7.2.1 *Specifications of Decision Errors*— It should be understood that, when a decision is made based on empirical data, there is no way to reduce either type of decision error to zero. Furthermore, there is usually a tradeoff between the two decision errors, meaning that a lower false negative error would lead to a higher false positive error, and vice-versa (for a given amount of data or number of samples). Decision makers should understand the consequences of decision errors and the tradeoffs between a false positive error and a false negative error. Error rates (false positive and false negative errors) must be specified relative to an agreed-upon concentration regulatory threshold or health-risk level.

6.7.2.2 *Consequences of an Incorrect Decision*—The random variability for empirical data is often composed of (but not limited to) sample variability and measurement variability. ~~Sampling variability is composed of both environmental variability (for~~

~~example, spatial, temporal, matrix, and so forth) and sample collection variability. Measurement variability is a function of extraction efficiencies, matrices effects, and analyte interferences.~~ Taken together, ~~sample variance and measurement variance components~~ they comprise the total variability in the data that contributes to errors in the decision under consideration. Decision makers must make an a priori judgement regarding how often they are willing to be wrong because of data variability. This uncertainty is the "acceptable error" in the decision. In the context of a decision designed to be protective of human health and the environment, they can be wrong by taking a prescribed action when none was necessary (false positive error), or they can fail to take action when it was necessary (false negative error).

6.7.2.3 *False Positive Error*—If the true concentration is lower than the regulatory threshold, but the decision makers conclude that the waste is hazardous because the sample average concentration is equal to or higher than the ~~action level,~~ decision point, then a false positive error has been made. The consequence of this error is that the nonhazardous waste will be remediated or disposed of according to stricter requirements than ~~required by regulations.~~ what is truly needed. A false positive error is undesirable because it will incur unnecessary costs and result in inefficiency.

6.7.2.4 *False Negative Error*—If the true concentration is equal to or greater than the regulatory threshold, but the decision makers conclude that the waste is nonhazardous because the sample average concentration is below the ~~action level,~~ decision point, then a false negative error has been made. The consequence of this error is that the waste will be disposed of by a less stringent method. This error is undesirable because ~~the allowed waste management method~~ this error may ~~allow~~ lead to consequences harmful to health or the environment.

6.7.2.5 The relationship between the probability of taking action on a decision rule and the possible true value of the measurement of interest is illustrated graphically by a decision performance curve in Fig. 7 ~~based on the~~ (see example ~~described~~ in Appendix X1). The decision performance curve depends on the decision makers' willingness to accept false positive and false negative errors, the total variability of the measurement process, the number of samples, and a regulatory threshold. The interval between the ~~action level~~ decision point and the regulatory threshold represents the range of possible true measurement values over which decision makers are willing to take more than a 50 % chance of sending a nonhazardous waste to a regulated landfill to ensure a specified false negative error (if the true value happens to be at the regulatory threshold). The curve is derived from the following:

(1) Acceptable errors (either a false positive error or a false negative error) agreed upon between the decision makers,

(2) Total variability of the system,

(3) Number of samples analyzed, and

(4) Statistical distribution of sample data (normal, lognormal and so forth).

6.7.2.6 In some cases, the ~~action level~~ decision point may equal a regulatory ~~level, or risk~~ level. In these cases, all of the decision makers ~~must~~ should understand that the value of a false positive error and false negative error associated with making a decision would be equal if the true value ~~q~~ happens to be at the regulatory threshold.

6.7.2.7 Specification of false positive error and false negative error is typically made on the basis of the relative importance of the consequence of an incorrect decision of either type. If the costs of environmental disposal or remediation are substantial and the potential environmental impact is relatively minor, then the emphasis may be on the control or reduction of false positive error (cost control). If the reverse is the case, then the emphasis may be on the control or reduction of the false negative error (control of environmental risk and liability). This important issue must be negotiated and resolved on a case-by-case basis for each problem identified in Step 1 by all decision makers.

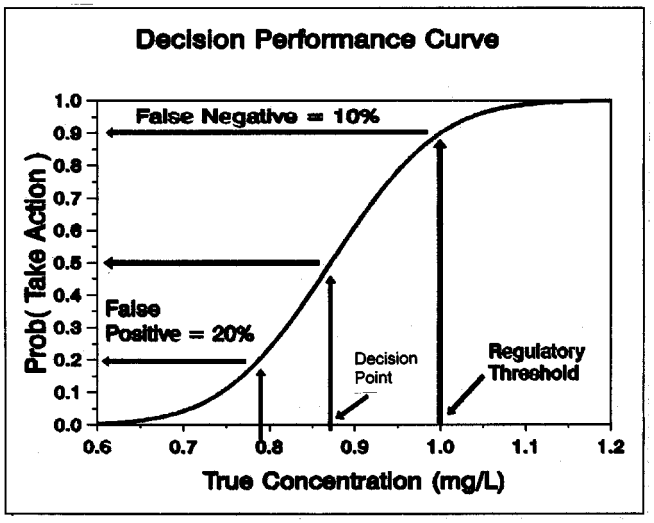

FIG. 7 Decision Performance Curve for Appendix X1 Example

~~6.7.2.8 This curve and several others that illustrate the relationships between these factors are discussed in the example in Appendix X1.~~

~~6.7.2.9~~

6.7.2.8 *Control of Decision Errors*—While decision errors cannot be eliminated, their errors can be reduced by (*1*) reducing sampling and measurement errors ~~(sampling or analytical variabilities, or both)~~ or (*2*) increasing the number of samples taken. These issues relate to optimization of the study design and are covered in Step 7 (see 6.8 and Guide D 6311).

6.7.3 *Output*—The rational and acceptable errors for both the false positive and false negative errors for each decision from Step 1 must be documented.

6.7.4 *DQO Summary*:

6.7.4.1 *Purpose*:

(*1*) The purpose of this step is to present the results of the DQO process clearly and concisely, in a form usable for optimizing data collection design (6.8; Step 7). This presentation of the DQOs and the complete documentation of the outputs and logic from which they were derived is essential for the initiation of data collection design.

(*2*) The DQOs are derived from the outputs of all of the preceding steps in the DQO process. Each output is important. However, the uncertainty on the decision and the decision rules incorporate the decision, boundaries, and inputs required to generate a sampling design. Indeed, the uncertainties on the decisions, together with the respective decision rules, are the primary results of the DQO process for a particular problem.

6.7.4.2 *Activities*:

(*1*) Activities include the establishment of a framework in which the decision rule(s) and associated limits on decision error are expressed as the DQO(s) supported by the documented logic and outputs of the previous steps of DQO process development. Within this decision framework, the DQOs can be improved and refined through an iterative process that includes use of and further evaluation of the following:

(*a*) Problem statement,

(*b*) Possible decisions,

(*c*) Inputs,

(*d*) Definition of spatial and temporal boundaries,

(*e*) Development of decision rule(s), and

(*f*) Acceptance of limits on decision error.

(*2*) Establishment of the DQOs by integration of concise decision rule(s) with their associated limits on decision error and the documentation of the DQO process is critical in facilitating understanding of the risk of making the wrong decision by the decision makers.

6.7.4.3 *Outputs*:

(1) Primary outputs consist of clear and concise presentation of the DQO process and complete documentation of the logic involved in development of the decision rules and associated limits on decision errors.

(2) As a useful tool, the DQO process can be integrated graphically into a typical decision tree or logic flow diagram that clearly indicates actions to be taken as the result of implementation of the decision rule(s) (see Fig. X1.1). These diagrams and associated descriptive text are effective formats for use during the optimization of data collection design and are important elements in project work plans.

(3) For example, the following are DQO summaries from Appendix X1: To make the following decision for the "cadmium incineration waste problem" with a false positive error not to exceed 20 % and a false negative error not to exceed 10 %. If the mean cadmium concentration in the toxicity characteristic leaching procedure (TCLP) extract is equal to or >1 mg/L, then dispose of the fly ash load in a suitable landfill. If the mean cadmium concentration in the TCLP extract is <1 mg/L, then dispose of the fly ash load in a sanitary landfill.

6.7.4.4 *Application of Data Quality Objectives*:

(*1*) The DQOs are applied on a day-to-day basis by incorporating the decision errors into the ~~action level.~~ decision point. This makes the decision rule easier to use. To apply DQOs, statisticians apply statistical methods such as those used in the example in Appendix X1 to calculate ~~an action level~~ a decision point that takes into account the acceptable decision uncertainty.

(*2*) The applied DQOs from Appendix X1 are as follows:

(*a*) If the average concentration of cadmium is ≥0.87 mg/L, then dispose of the waste fly ash in a hazardous waste landfill; and

(*b*) If the average concentration of cadmium is <0.87 mg/L, then dispose of the waste fly ash in a sanitary landfill.

6.7.4.5 *Decision Tree Format*—In decision tree format, the DQOs are presented along with the actions and tasks that are required in the data collection design step (see Fig. 5).

6.8 *Step 7—Optimizing Data Collection Design*:

6.8.1 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 undertake to optimize the actual sampling design for data collection. Care should be taken to separate the factual material from the DQO team's assumptions or estimates, or both, 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.

6.8.2 *Purpose*:

6.8.2.1 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.

6.8.2.2 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 different types of decision errors with the resources available (time, money, and personnel) to resolve the problem.

6.8.2.3 Once sampling designs are developed, the sampling design alternatives and required resources for each should be presented to the decision makers. These alternatives allow for an understanding of the benefits and resource commitments to each sampling design. 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. This decision is the responsibility of the decision makers and requires that all DQO team members be involved. New members may be added if, in the opinion of the decision makers, their expertise is needed to develop acceptable DQOs.

6.8.3 *Activities*—The activities involved in the development of an optimal sampling design and chemical analyses are shown in Fig. 8.

6.8.3.1 *Summary of Information*—The data collectors should summarize any previous data and the outputs from the previous six steps of the DQO process. This allows data collectors to remain focused on the decision makers' needs in design optimization.

6.8.3.2 *Development of Sampling Design Alternatives*—Alternative sampling designs must be based on DQOs, which were developed with an understanding of measurement variability and the resources available for resolving the problem. Design alternatives must address the degree of representation of any one sample within the problem boundaries. This is accomplished by selecting from among the sampling designs those that best describe the system. These include, but are not limited to, random, sequential random, systematic, and stratified sampling designs.

(1) Probabilities of selecting an appropriate sample are related to the type of sampling design. An equal probability of selecting a sample implies a random sample design. Selecting unequal probabilities for sample selection implies a stratified sample design. The more heterogeneous the sampling units, the more likely unequal probabilities will be assigned to the sample. Furthermore, the more heterogeneous the waste site, the more useful historical or process information is in assessing the sampling design alternatives. The participation of a qualified statistician is critical in this process.

(2) Variability may also be introduced during sample handling and preparation procedures that may be necessary between field

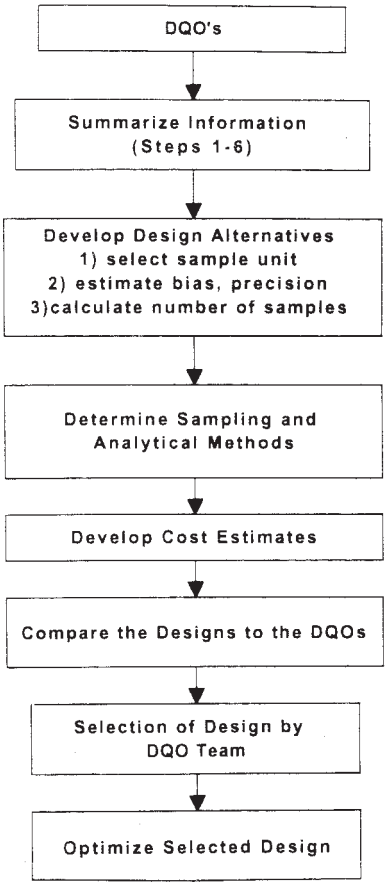

FIG. 8 Optimization of Sample Design

sampling and analytical methods. Consideration of the important factors impacting sample variability should occur during the design process.

6.8.3.3 *Determination of Analytical Chemistry Methods* —The alternative analytical chemistry methods as documented during the DQO process must be considered. Factors that affect selecting alternative methods include, but are not limited to, the following:

(1) Detection limits versus ~~action levels;~~ decision points;

(2) Matrix effects on detection limits, bias, and variability; and

(3) Sample amount available (volume or weight).

6.8.3.4 *For Each Sampling Design Alternative, Selection of the Sample Unit that Satisfies the DQOs*—Sampling units include drums, tanks, an area within a grid, a boring location on a grid, a depth interval in a boring, or any other appropriate defined physical unit from which material can be obtained. Different sampling units may and often will be appropriate for different materials or locations. The sampling unit may depend on logistical and resource issues, such as whether the material will be disposed by drum or truck or the amount of material that can be excavated.

6.8.3.5 *For Each Sampling Design Alternative, Calculation of the Optimal Number of Samples that Satisfies the DQOs*— ~~Typically, samples are collected from each sample unit for chemical analyses. Using~~ —Using the mathematical expressions for sampling design optimization, solve for the optimal number of samples that meet the uncertainty limits on the decision errors specified in the DQOs. Selection of the number of samples is an iterative process. Initial selection of the number of samples may be based on different project criteria (for example, budget, precision limits, and so forth). These initial calculations should be examined to determine whether they are adequate for the specified decision errors. In addition, preliminary sample designs may be required for better estimates of mean concentrations and measurement variability for optimal planning of larger sample designs.

6.8.3.6 *For Each Sampling Design Alternative, Development of Cost Estimates*—The estimates should relate the total cost of sampling and chemical analyses for alternative sampling designs. These cost functions may take into account such items as the cost of remediation or waste disposal by sample unit. This enables the decision makers to assess whether sampling and chemical analyses are more cost effective than proceeding with cleanup or disposal with minimal data collection.

6.8.4 *Outputs*—The list of sampling design alternatives is submitted to the decision makers for selection. After selection of the final sampling design, document the operational details and theoretical assumptions of the selected sampling design in a final sampling and chemical analyses plan. The documentation should include the sampling plan, sampling and analytical chemistry procedures, data assessment procedures, quality control requirements, and overall project quality assurance requirements.

7. Documentation of the Data Quality Objective Process

7.1 The following statements and information document the outputs of the specific DQO process used to develop the DQOs. The DQOs are meaningless if they are not connected with the specific problem and other qualifying information used to develop them.

7.2 DQO process documentation summaries can vary from problem to problem, but most will include information such as the following:

7.2.1 Facility name, location, and process;

7.2.2 List of decision makers, affiliations, and responsibilities for this project.

7.2.3 Statement of the problem.

7.2.4 Summary of logic for the decisions chosen for consideration. For each problem there must be at least one decision.

7.2.5 Information and inputs such as those given in 6.4.2. There should be appropriate inputs to allow generation of the data to make a decision. It may be useful to establish separate decisions for each matrix (that is, soil, sediment, and water).

7.2.6 Defined boundaries, which should be addressed for each decision. It may be useful to segregate the boundaries by matrix.

7.2.7 Decision rules, which should incorporate appropriate boundaries. The rules may be stated by matrix.

7.2.8 Limits on decision error. The rationale or assumptions upon which decision error estimates are based should be documented.

8. Keywords

8.1 data quality objectives; DQOs; project planning; waste analysis; waste testing

APPENDIX

(Nonmandatory Information)

X1. DQO CASE STUDY—CADMIUM-CONTAMINATED FLY ASH WASTE

X1.1 *Background:*

X1.1.1 A municipal waste incineration facility located in the Midwest routinely removes "fly ash" from its flue gas scrubber system and disposes it in a sanitary landfill. It was determined previously that the ash was nonhazardous under hazardous waste regulations. However, the incinerator has recently begun treating a new waste stream. As a result, a local environmental public interest group asked that the ash be retested and evaluated for hazardous waste compliance before it is disposed. The group is primarily concerned that the ash may contain hazardous levels of cadmium due to the new waste sources. The facility manager has agreed to test the ash and decided to use the DQOs process to help guide decision making throughout the project. Although not constrained by cost, the facility is interested in minimizing expenditures.

X1.1.2 The 40 CFR Part 261 RCRA toxicity characteristic criteria **(7)** for determining whether a solid waste is hazardous requires collection of a "representative portion" of the waste and performance of TCLP. During this process, the solid fly ash will be "extracted" or mixed in an acid solution for 18 h. The extraction liquid will then be subjected to tests for specific metals.

X1.1.3 Since the impact of this new waste stream is not known, a preliminary study was conducted to determine the variability of the concentration of the contaminants. Random samples were collected from the first 20 truckloads. Since process knowledge of the waste stream indicated that cadmium was the only toxicity characteristic (TC) constituent in the waste, these samples were analyzed individually for cadmium using TCLP. The results were expressed as the average concentration along with the standard deviation.

X1.2 *Data Quality Objective Development*—The following is an example of the outputs from each step in the DQO process.

X1.2.1 *Statement of the Problem*:

X1.2.1.1 *Identification of the DQO Team*—The plant manager assembled a DQO team consisting of himself and a representative of the current disposal facility staff. The two of them subsequently assembled the additional DQO team members.

(*1*) The decision makers on the DQO team included the incinerator owner and incineration plant manager, and a representative of the environmental public interest group, in which a representative of the community in which the ash is currently being disposed. The technical staff included a statistician, toxicologist, and chemist with sampling experience.

X1.2.1.2 *Statement of the Problem*—The problem is to determine whether any loads of fly ash are hazardous with cadmium under RCRA regulations using TCLP testing. If a load is hazardous, it must be disposed of in a RCRA landfill.

X1.2.2 *Identification of Possible Decisions*:

X1.2.2.1 *Decision*—Determine whether the concentration of cadmium in TCLP leachate from waste fly ash exceeds the regulatory RCRA standards.

X1.2.2.2 *Statement of the Actions that Could Result from the Decision*:

(*1*) If the average concentration of cadmium is greater than or equal to the ~~action level,~~ decision point, dispose of the waste fly ash in a RCRA landfill.

(*2*) If the average concentration of cadmium is less than the ~~action level,~~ decision point, dispose of the waste fly ash in a sanitary landfill.

X1.2.3 *Identification of Inputs to Decisions*—The DQO team identified the following inputs or information needed for the decision rules:

X1.2.3.1 *Preliminary Study Information* —Since the concern is with a new waste stream, the DQO team ordered a pilot study of the fly ash to determine the variability in the concentration of cadmium between loads of fly ash leaving the facility. They have determined that each load is fairly homogeneous. However, there is a high variability between loads due to the nature of the waste-stream. Most of the fly ash produced is not a RCRA hazardous waste and may be disposed of in a sanitary landfill. Because of this, the company has decided that testing each individual waste load before it leaves the facility would be the most economical. In that way, they could send loads of ash that exceeded the regulated cadmium concentrations to the higher-cost RCRA landfills and continue to send the others to the sanitary landfill.

(*1*) The study showed that the standard deviation of the cadmium concentration within a load was $S_w = 0.4$ mg/L, and the standard deviation of the cadmium concentration between loads was $S_b = 1.4$ mg/L. Sample and quality control data indicate that a normal distribution can be assumed.

X1.2.3.2 *Identification of Contaminants of Concern, Matrix, and Regulatory Limits*—The DQO team identified the following factors critical to the problem:

(1) Contaminants of Concern—Cadmium soluble in the TCLP extract.

(2) Sample Matrix—Fly ash.

(3) Regulatory Threshold—1 mg/L.

X1.2.3.3 *Specific Project Budget and Time Constraints*—The incinerator plant manager has requested that all stages of the

operation be performed in a manner that minimizes the cost of sampling, chemical analysis, and waste disposal. However, no formal cost constraints have been implemented.

(1) The environmental public interest group has threatened to file a lawsuit for violation of environmental regulations if testing does not proceed within a "reasonable time-frame."

(2) The waste does not pose a threat to humans or the environment while contained in the trucks. Additionally, since the fly ash is not subject to change, disintegration, or alteration, the chemical properties of the waste do not warrant any temporal constraints. However, in order to expedite decision making, the DQO team has placed deadlines on sampling and reporting. The fly ash waste will be tested within 48 h of being loaded onto waste hauling trailers. The analytical results from each sampling round should be completed and reported within five working days of sampling.

X1.2.3.4 *Identification of the Testing Methods*—In this case, 40 CFR Part 261, Appendix II specified the TCLP Method SW 846, Method 1311 **(8)**. The leachate must be analyzed by an appropriate method. Potential methods of characterizing the leachate for cadmium include, but are not limited to, SW 846, Methods 6010, 6020, 7130, or 7131.

X1.2.4 *Inputs to Be Determined*:

X1.2.4.1 *Method Validation and Quality Control (QC)*—The analytical method accuracy and precision and method detection limits in leachate from the fly ash matrix must be determined. The QC samples must be specified.

X1.2.4.2 *Identification of Sampling Procedure or Devices*—The following must be determined:

(1) Number of samples,

(2) Sampling methods for composite or grab samples of ash, and

(3) The QC requirements for sampling.

X1.2.5 *Definition of the Boundaries* —Define a detailed description of the spatial and temporal boundaries of the decision, characteristics that define the environmental media and objects or people of interest, and any practical considerations for the study.

X1.2.5.1 *Specification of the Characteristics that Define the Sample Matrix*—The fly ash should not be mixed with any other constituents except the water used for dust control.

X1.2.5.2 *Identification of Spatial Boundaries*—The variability between loads was greater than within a load; therefore, a decision will be made on each load. The waste fly ash will be tested after it has been deposited in the trailer used by the waste hauler. Separate decisions regarding the toxicity of the fly ash will be made for each load of ash leaving the incinerator facility. Each load of ash should fill the waste trailer at least 70 %. In cases in which the trailer is filled less than 70 %, the trailer must wait on-site until more ash is produced and can fill the trailer to the appropriate capacity.

X1.2.5.3 *Identification of Temporal Boundaries (Including the Time Frame Over Which the Study Should Be Conducted)* —The waste does not pose a threat to humans or the environment while contained in the trucks. However, in order to expedite decision making, the DQO team has placed deadlines for reaching a decision. The fly ash waste will be tested and a disposal decision made within 48 h of being loaded onto waste hauling trucks.

X1.2.6 *Development of Decision Rules* —The arithmetic mean of sample results will be compared to the ~~action level.~~ decision point.

X1.2.6.1 *Decision Rule*:

(*1*) If the average concentration of cadmium in a truck load is equal to or greater than the ~~action level,~~ decision point, then dispose of the waste fly ash in a RCRA landfill; or

(*2*) If the average concentration of cadmium in a truck load is less than the ~~action level,~~ decision point, then dispose of the waste fly ash in a sanitary landfill.

Note that the DQO team will decide that the ~~action level~~ decision point is less than the regulatory level in order to meet a 10 % false negative error for concentrations at the regulatory level of 1 mg/L.

X1.2.7 *Specification of Limits on Decision Errors*:

X1.2.7.1 The decision makers specify acceptable decision errors based on the consequences of making an incorrect decision. Both types of decision errors have negative consequences.

(1) False Positive Error (declaring the load hazardous when it is not)—If the true cadmium concentration is below 1 mg/L, but the average measured cadmium concentration is above the ~~action level,~~ decision point, the nonhazardous fly ash waste will be sent to a RCRA landfill. The consequence of a false positive error is that the company will have to pay additional cost to dispose of the waste with a cadmium concentration between the ~~action level~~ decision point and regulatory threshold at a RCRA facility as opposed to a less expensive method of disposal in a sanitary landfill.

(2) False Negative Error (declaring the load nonhazardous when it is hazardous)—If the true cadmium concentration is equal to or greater than 1 mg/L, but the average measured cadmium concentration is below the ~~action level,~~ decision point, the hazardous fly ash waste will be sent to a sanitary landfill. The consequence of a false negative error is that the fly ash waste may be disposed of in a manner that will be harmful to human health or the environment. Legal consequences and subsequent remedial costs are also possible consequences.

X1.2.7.2 The purpose of this stage of the process is to specify the probabilities of making incorrect decisions that are acceptable to decision makers. The DQO team must agree on which type of decision error is of greater concern, either a false positive error or false negative error.

X1.2.7.3 For this example, the DQO team is more concerned about a false negative error because of the increased liability due

to sending potentially hazardous waste to a sanitary landfill. The DQO team set a value for the false negative error of 10 % when the true concentration is 1 mg/L. The false negative error is a greater concern because of the perceived increased liability due to sending potentially hazardous waste to a sanitary landfill. This level is determined based on the comfort of the decision makers accepting the risk associated with calling a hazardous waste nonhazardous.

X1.2.7.4 *Data Quality Objective Summary* —Application of the DQOs on a day-to-day basis depends on (*1*) selecting the number of samples and (*2*) quantifying the ~~action level~~ decision point for the decision rule. The decision performance curves are used to visually compare the desired decision errors versus the possible true cadmium concentrations for different numbers of samples.

(1) The uncertainty for the DQOs can be quantified by calculating the ~~action level~~ decision point based on a false negative error of 10 % when the true cadmium concentration of a TCLP extract for a fly ash load has a value of the regulatory threshold (1 mg/L).

(2) To begin the early phases of design optimization, the DQO team determined how the environmental data should be summarized and used in the decision. The DQO team identified that the mean concentration of cadmium from each load would be compared to the ~~action level.~~ decision point. The background data indicated that a normal distribution can be used to calculate the ~~action level.~~ decision point. A normal distribution is an appropriate probability model for the preliminary data. A false negative error less than 50 % implies that ~~an action level~~ a decision point will be lower than the regulatory threshold.

(3) How the statisticians on the DQO team calculated the ~~action level~~ decision point for the project is shown as follows. The ~~action level~~ decision point is dependent on variables such as regulatory threshold, standard deviation, false negative error, and number of samples. Changing one variable will affect the value of the ~~action level.~~ decision point. Another iteration through the last DQO process steps must be made if any of these changes are made.

X1.2.7.5 *Concentration Range and ~~Action Level~~ Decision Point*—The DQO team examined the concentration data from the first 20 analyses and determined that a reasonable concentration range to examine was between 0.6 and 1.3 ~~µ~~ mg/L. The DQO team agreed that the ~~action level~~ decision point should be based on a 10 % false negative error at the regulatory threshold. This implies that the ~~action level~~ decision point will be less than the regulatory threshold. Paragraph X1.2.8 describes the calculations for several ~~action levels~~ decision points corresponding to different numbers of samples in the decision performance curve, using the standard deviation, the limits of error, and the desired false negative error. The decision performance curve will be calculated to determine the ~~action level~~ decision point and review the performance of the decision rule. To calculate the decision performance curve, decision makers use the following steps:

(*1*) *Step 1—Number of Samples:*

(*a*) Selecting the number of samples is always difficult because imperfect knowledge is available concerning the variability of the measurement process for the selected sample matrix. All calculations for the number of samples are approximations. Different methods can be used to determine the number of samples. For the cadmium example, an initial selection of the number of samples is determined by an estimation method that specifies the precision limits on determining the concentration in the TCLP extract. Another sample size method would be based on the decision performance curve that examines the effect of a different number of samples on the decision errors. This decision method for number of samples is investigated in X1.2.8. Another method would be to calculate the number of samples for specified values of the measurement standard deviation, ~~action level,~~ decision point, and false positive error and false negative error. This procedure is illustrated in Guides C 970 and C 1215.

(*b*) For the initial fly ash waste loads, chemists on the DQO team would like to verify that their instrument is calibrated for the proper concentration range. They want to estimate the true cadmium concentration in the TCLP extract with an uncertainty of ± 0.2 mg/L. In addition, the decision makers are willing to allocate resources to learn that the true cadmium concentration is in this interval with a confidence of 95 %. The number of samples for these precision limits can be approximated by a normal probability distribution. Another approximation to the number of samples could use an iterative method for a Student's *t*-distribution rather than the normal distribution. This more general assumption usually adds only two or three samples beyond the normal distribution used herein.

(*c*) The number of samples (*n*) is calculated by the following equations **(9, 10)**, with $L = 0.2$ mg/L, $\sigma = S_w = 0.4$ mg/L, and $\alpha = 0.05$ (or $Z_{\alpha/2} = 1.960$ for a 95 % confidence level):

$$n = \left(\frac{Z_{\alpha/2}\,\sigma}{L}\right)^2$$

$$n = \left(\frac{1.960 \times 0.4}{0.2}\right)^2 \approx 16 \qquad \text{(X1.1)}$$

where:

n = number of samples,
L = limit of error on the average (for example, 0.2 mg/L),
$1 - \alpha$ = probability level for the confidence interval for $\alpha = 0.05$, and then $1 - \alpha = 0.95$ confidence interval,
σ = standard deviation of the measurement process (for example, 0.4 mg/L), and
$Z_{\alpha/2}$ = $\alpha/2$ percentile point of normal probability distribution (for example, $Z_{\alpha/2} = Z_{0.025}$). Common normal percentile values are given in Table X1.1.

(*2*) *Step ~~2—Action Level~~ 2—Decision Point*—The ~~action level~~ decision point value for the decision rule is determined by controlling the false negative error established in the DQO process. The quantification of the ~~action level~~ decision point used a

TABLE X1.1 Common Normal Percentile Points

$Z_{0.20}$	$Z_{0.10}$	$Z_{0.05}$	$Z_{0.025}$	$Z_{0.01}$	$Z_{0.005}$
0.842	1.282	1.645	1.960	2.326	2.576

value of 0.10 (or 10 %) for the probability of the false negative error and 16 samples to determine the average cadmium concentration from the TCLP extracts. The probability calculations are based on an approximating normal probability distribution for the cadmium concentration measurements. This approximating normal probability assumes a mean = RT = 1.0 mg/L and a standard deviation = S_w = 0.4 mg/L. The 10 % percentile point for the standardized normal probability distribution is $Z_{0.10}$ = 1.282 (see Table X1.1). The probability (Pr) for the false negative error evaluated at RT is as follows:

$$\begin{aligned}&\Pr(\text{false negative error})\\ &= \Pr(\text{average} < AL \text{ when the true concentration}\\ &= RT)\\ &= 0.10\end{aligned}$$

or

$$\Pr(FN) = \Pr\left[\frac{\text{average} - \text{RT}}{S_w/\sqrt{n}} < \frac{\text{AL} - \text{RT}}{S_w/\sqrt{n}}\right] = 0.10,$$

$$\frac{AL - RT}{S_w/\sqrt{n}} = -Z_{0.10,}$$

$$\text{AL} = \text{RT} - \text{Z}_{0.10}\frac{S_w}{\sqrt{n}}.$$

$$\text{AL} = 1.0 \text{ mg/L} - (1.282)(0.4 \text{ mg/L})/4 = 1.0 \text{ mg/L} - 0.13 \text{ mg/L},$$
$$\text{AL} = 0.87 \text{ mg/L}. \quad \text{(X1.2)}$$

where:

AL = ~~action level,~~ decision point,
RT = regulatory threshold,
S_w = standard deviation of the measurement process estimated from a sufficient number of samples, and
$Z_{0.10}$ = tabulated 10 % percentile point from a standard normal distribution (see Table X1.1).

Therefore, the decision rule is as follows:

(*a*) If (average concentration of cadmium) ≥ 0.87 mg/L, the fly ash load is considered to be a RCRA waste and will be disposed of in a RCRA landfill; or

(*b*) If (average concentration of cadmium) < 0.87 mg/L, the fly ash load is not considered to be a RCRA waste and will be disposed of in a sanitary landfill.

X1.2.7.6 *Decision Tree Format*—Fig. X1.1 shows the decision tree format for the DQOs, along with the ~~action level~~ decision point and tasks that are required in the data-collection design step.

(*1*) *Step 3—True Concentration Corresponding to the False Positive Error:*

(*a*) Calculate the true concentration (θ mg/L < RT) that corresponds to a probability for the false positive error of 20 % using ~~an action level~~ a decision point of AL = 0.87 mg/L. This calculation again uses the approximating normal probability distribution

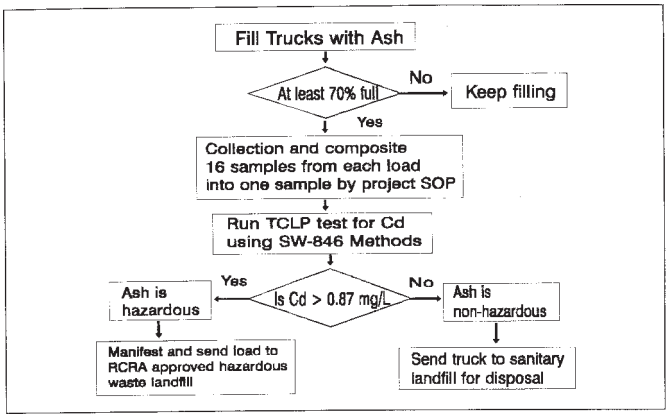

FIG. X1.1 Decision Tree for the Cadmium Example

for the cadmium concentration measurements. For the specified false positive error, the approximating normal probability assumes a mean = θ mg/L (to be determined), a standard deviation = S_w = 0.4 mg/L, and the number of samples = 16. The 20 % percentile point for the standardized normal probability distribution is $Z_{0.20}$ = 0.842 (see Table X1.1).

$$\text{Pr (false positive error)} = \text{Pr}\{\text{average} \geq \text{AL when the true concentration} = \theta < RT\} = 0.20$$

or

$$\text{Pr(FP)} = \text{Pr}\left[\frac{\text{average} - \theta}{S_w/\sqrt{n}} \geq \frac{\text{AL} - \theta}{S_w/\sqrt{n}}\right] = 0.20,$$

$$\frac{AL - \theta}{S_w/\sqrt{n}} = + Z_{0.20,}$$

$$\theta = AL - Z_{0.20}\frac{S_w}{\sqrt{n}}.$$

$$\theta = 0.87 \text{ mg/L} - (0.842)(0.4 \text{ mg/L})/4 = 0.87 \text{ mg/L} - 0.08 \text{ mg/L},$$
$$\theta = 0.79 \text{ mg/L}. \quad \text{(X1.3)}$$

where:

AL = ~~action level,~~ decision point,
RT = regulatory threshold, and
$Z_{0.20}$ = tabulated 20 % percentile point from a standard normal distribution (see Table X1.1).

(*b*) The decision performance curve would have a probability of taking an action (that is, sending fly ash waste to a RCRA landfill) of 0.20 at a true cadmium concentration of θ = 0.79 mg/L. The possible true cadmium concentration values in the interval (0.79 and 1.0 mg/L) represent values that cause the decision rule to send fly ash waste to a RCRA landfill even though the true concentration is below the regulatory threshold. This interval can be reduced by increasing the number of samples, changing the false negative error, or changing the false positive error.

(*2*) *Step 4—Drawing the Decision Performance Curve:*

(*a*) Draw the decision performance curve by using the standardized normal probability distribution. The standardized normal probability distribution is defined as a normal probability distribution with mean = 0 and standard deviation = 1.0. There are many tables and computer programs that can be used to calculate probabilities for a standardized normal random variable, *Z*. A normal random variable, *X*, with mean = μ and standard deviation = σ can be transformed to a standardized normal random variable by $Z = (X - \mu)/\sigma$.

$$\text{Prob (action)} = \text{Pr (average} \geq \text{AL when the true concentration} = \theta)$$

$$\text{Prob (action)} = 1.0 - \text{Prob}\left(Z \leq \frac{\text{AL} - \theta}{S_w/\sqrt{n}}\right),$$

$$\text{Prob (action)} = 1.0 - \text{Prob}\left(Z \leq \frac{0.87 - \theta}{0.1}\right). \quad \text{(X1.4)}$$

(*b*) Fig. X1.2 is a plot of the decision performance curve generated by calculating a Prob (action) value using the standard normal probability distribution for each possible true concentration value θ. The decision performance curve can frequently be drawn freehand if three pairs of (concentration and probability) values are determined: ((RT, 1 − Pr (false negative error)), (AL, 0.50), and (θ, Pr (false positive error)).

X1.2.8 *Optimizing Data Collection and Design*—The decision makers will select the lowest-cost sampling design that is expected to achieve the DQOs. The series of designs for sampling the fly ash waste will be generated by the statisticians on the DQO team. The choice of sampling plan will be decided by consensus.

X1.2.8.1 *Decision Performance Curve* —The decision performance curve in Fig. X1.2 plots the probability of taking action (disposing of the waste in a RCRA landfill) versus different possible values for the true concentration in the TCLP extract. The DQO process specified a probability of 0.10 for the false negative error when the true concentration is at the RT. This specified false negative error implies that the decision performance curve will have a probability of taking action equal to 0.90 when the true concentration is equal to RT. If the true concentration value is equal to the value of the ~~action level~~ decision point (0.87 mg/L), there is a probability of taking action of 0.50. The DQO team can also determine the true concentration for a specified false positive error from the decision performance curve.

(*1*) Fig. X1.2 shows three decision performance curves for three different numbers of samples (8, 16, and 24). All three decision performance curves meet the specified probability for the false negative error of 0.10 at a true concentration equal to RT. The

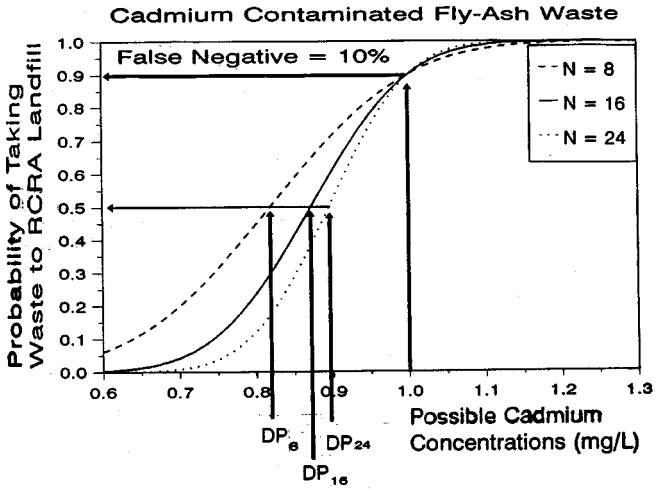

FIG. X1.2 Decision Performance Curves for Cadmium Example

purpose of these curves is to assess the effects of taking more or fewer samples on the ~~action level~~ decision point and the false positive error. This analysis can be used to update applying the decision rule. For example, the decision makers concluded that eight additional samples (that is, 24) does not improve the AL value and false positive error sufficiently to justify the increase in cost.

X1.2.8.2 *Implementation*— Cadmium concentration values from the TCLP extracts will be collected over a long time period because this waste stream is a continuous process. The decision makers will establish a QC program to monitor the cadmium concentration values for process changes. After every 30 fly ash loads, the process variability will be reestimated and new values for the number of samples and ~~action level~~ decision point will be considered. This strategy becomes part of the decision process.

X1.2.8.3 *Documentation of the Data Quality Objective Process*—The following statements and information document the outputs of the specific DQO process used to develop the above-stated DQOs. These objectives are meaningless if they are not connected with the specific problem and other qualifying information used in the DQO development.

(*1*) The DQO team required that the documentation be a concise summary of the following information:

(*a*) Facility name, location, and process;

(*b*) List of DQO team members, affiliations, and responsibilities for this project;

(*c*) Statement of the problem;

(*d*) Logic for the solutions chosen for consideration;

(*e*) Information and inputs required by the DQO team to make the decision, including sample matrix, preliminary study results, sampling methods required, and use of each input in reaching a decision.

(*f*) Defined boundaries;

(*g*) Decision logic in rule or decision tree format; and

(*h*) Assumptions made regarding the decision error and any information used to generate preliminary ~~action levels~~ decision points and the number of samples.

(*2*) All meetings held by the DQO team should be documented. The meeting minutes should include the attendees, information used to generate each step of the process, and rationale used to make final agreements on the decision logic, boundaries, inputs, and decision errors.

REFERENCES

(**1**) *Guidance for Planning for Data Collection in Support of Environmental Decision Making Using the Data Quality Objectives Process*, EPA QA/G-4, U.S. Environmental Protection Agency, Washington, DC, 1993.
(**2**) Blacker, S., and Maney, J., "The DQO Process: System DQO Planning Process and the Analytical Laboratory," *Environmental Testing and Analysis*, July/August 1993.
(**3**) Neptune, D., Blacker, S., Fairless, B., and Ryti, R., "Application of Total Quality Principles to Environmental Data Operations," *National Energy Division Conference*, American Society of Quality Control, 1990.
(**4**) *A Rationale for the Assessment of Errors in the Sampling of Soils*, EPA-600/4-90-013, Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Washington, DC, May 1990.
(**5**) *Methods for Evaluating the Attainment of Cleanup Standards Volume 1: Soils and Solid Media*, U.S. Environmental Protection Agency, Washington, DC, February 1989.
(**6**) *Characterizing Heterogeneous Wastes: Methods and Recommendations*, EPA 600/R-92/033, U.S. Environmental Protection Agency and Department of Energy Office of Technology, Washington, DC, February 1992.
(**7**) 40 Code of Federal Regulations (CFR), Part 261, 1995.
(**8**) *Test Methods for Evaluating Solid Waste*, 3rd ed., with Updates I, II, IIA, IIB, SW-846, Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC, January 1992.
(**9**) Cochran, W. G., *Sampling Techniques*, 3rd ed., John Wiley and Sons, Inc., New York, NY, 1977.
(**10**) Desu, M. M., and Raghavarao, D., *Sample Size Methodology*, Academic Press, San Diego, CA, 1990.