Designation: D 6233 – 98

Standard Guide for Data Assessment for Environmental Waste Management Activities[1]

This standard is issued under the fixed designation D 6233; 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 guide covers a practical strategy for examining an environmental project data collection effort and the resulting data to determine if they will support the intended use. It covers the review of project activities to determine conformance with the project plan and impact on data usability. This guide also leads the user through a logical sequence to determine which statistical protocols should be applied to the data.

1.1.1 This guide does not establish criteria for the acceptance or use of data but instructs the assessor/user to use the criteria established by the project team during the planning (data quality objective process), and optimization and implementation (sampling and analysis plan) process.

1.2 *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:*

D 4687 Guide for General Planning of Waste Sampling[2]

D 5088 Practice for Decontamination of Field Equipment Used at Nonradioactive Waste Sites[3]

D 5283 Practice for Generation of Environmental Data Related to Waste Management Activities: Quality Assurance and Quality Control Planning and Implementation Activities[2]

D 5792 Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives[2]

D 5956 Guide for Sampling Strategies for Heterogeneous Wastes[2]

D 6044 Guide for Representative Sampling for Management of Waste and Contaminated Media[2]

3. Terminology

3.1 *Definitions of Terms Specific to This Standard:*

3.1.1 *bias, n*—a systematic error that is consistently negative or consistently positive.

3.1.2 *characteristic, n*—a property of items in a sample or population which can be measured, counted, or otherwise observed.

3.1.3 *composite sample, n*—a physical combination of two or more samples.

3.1.4 *confidence limit, n*—an upper and/or lower limit(s) within which the true value is likely to be contained with a stated probability or confidence.

3.1.5 *continuous data, n*—data where the values of the individual samples may vary from minus infinity to plus infinity.

3.1.6 *data quality objectives (DQOs), n*—DQOs are 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).

3.1.7 *data quality objective process, n*—a quality management tool based on the scientific method and developed to facilitate the planning of environmental data collection activities.

3.1.8 *discrete data, n*—data made up of sample results that are expressed as a simple pass/fail, yes/no, or positive/negative.

3.1.9 *heterogeneity, n*—the condition of the population under which all items of the population are not identical with respect to the parameter of interest.

3.1.10 *homogeneity, n*—the condition of the population under which all items of the population are identical with respect to the parameter of interest.

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

3.1.12 *representative sample, n*—a sample collected in such a manner that it reflects one or more characteristics of interest (as defined by the project objectives) of a population from which it is collected.

3.1.13 *sample, n*—a portion of material which is taken from a larger quantity for the purpose of estimating properties or composition of the larger quantity.

3.1.14 *sampling design error, n*—error which results from the unavoidable limitations faced when media with inherently

[1] This guide is under the jurisdiction of ASTM Committee D34 on Waste Management and is the direct responsibility of Subcommittee D34.01.01 on Planning for Sampling.

Current edition approved Feb. 10, 1998. Published June 1998.

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

[3] *Annual Book of ASTM Standards*, Vol 04.09.

variable qualities are measured and incorrect judgement on the part of the project team.

3.1.15 *subsample*, *n*—a portion of a sample that is taken for testing or for record purposes.

4. Significance and Use

4.1 This guide presents a logical process for determining the usability of environmental data for decision making activities. The process describes a series of steps to determine if the enviromental data were collected as planned by the project team and to determine if the *a priori* expectations/assumptions of the team were met.

4.2 This guide identifies the technical issues pertinent to the integrity of the environmental sample collection and analysis process. It guides the data assessor and data user about the appropriate action to take when data fail to meet acceptable standards of quality and reliability.

4.3 The guide discusses, in practical terms, the proper application of statistical procedures to evaluate the database. It emphasizes the major issues to be considered and provides references to more thorough statistical treatments for those users involved in detailed statistical assessments.

4.4 This guide is intended for those who are responsible for making decisions about environmental waste management projects.

5. General Considerations

5.1 This guide provides general guidance about applying numerical and other techniques to the assessment of data resulting form environmental data collection activities associated with waste management activities.

5.2 The environmental measurement process is a complex process requiring input from a variety of personnel to properly address the numerous issues related to the integrity of the sample collection and measurement process in sufficient detail. Table 1 lists many of the topics that are common to most environmental projects. A well-executed project planning activity (see Guide D 4687, Practices D 5088, D 5283, and D 5792) should consider the impact of each of these issues on the reliability of the final project decision. The data assessment process must then evaluate the actual performance in these areas versus that expected by the project planners. Significant deviations from the *a priori* performance level of any one or combination of these issues may impact the reliability of the project decision and necessitate a reconsideration of the decision criteria by the project decision makers.

5.3 Appropriate professionals must assess the project planning documents and completed project records to determine if the project findings match the conceptual model and decision logic. In areas where the findings don't match, the assessors must document and report their findings and, if possible, the potential impact on the decision process. Items subject to numerical confirmation are compared to the project plan and any discrepancies and their potential impact noted.

5.4 Effective quality control (QC) programs are those that empower the individuals performing the work to evaluate their performance and implement real-time corrections during the sampling or measurement process, or both. When quality control processes (including documentation) are properly implemented, they result in data sets (see Fig. 1) that are generated by in-control processes or out-of control processes that were not amenable to corrective action but whose details are explained by the project staff conducting the work. Good QC programs lead to reliable data that are seldom called into question during the assessment process. However, in cases where the absence of staff responsibility or authority to self-monitor and correct deficiencies at the working level is missing, the burden of assuring data integrity is placed on the quality assurance (QA) function. The data assessment process must determine the location (working level or QA level) where effective quality control occurs (detection of error and execution of corrective action) in the data collection process and focus on how well the QC function was executed. As a general rule, if the QC function is not executed in real-time and thoroughly documented by the staff performing the work, the more likely the data assessor will be to find questionable data.

5.5 In addition to addressing the issues listed in Table 1, the data assessment process must search for unmeasurable factors whose impact cannot be detected by the review of the project records against expectations or numerical techniques. These are the types of things that are controlled by effective quality

TABLE 1 Information Needed to Evaluate the Integrity of the Environmental Sample Collection and Analysis Process

General Project Details	• Site History • Process Description • Waste Generation Records • Waste Handling/Disposal Practices • Sources of Contamination • Conceptual Site Model • Potential Contaminants of Concern • Fate and Transport Mechanisms • Exposure Pathways • Boundaries of the Study Area • Adjacent Properties
Sampling Issues	• Sampling Strategy • Sample Location • Sample Number • Sample Matrix • Sample Volume/Mass • Discrete/Composite Samples • Sample Representativeness • Sampling Equipment, Containers and Preservatives
Analytical Issues	• Laboratory Sub-sampling • Sample Preparation Methods • Analytical Method • Detection Limits • Matrix Interferences • Bias • Holding Times • Calibration • Quality Control Results • Contamination • Reporting Requirements • Reagents/Supplies
Validation and Assessment	• Data Quality Objectives • Chain of Custody • Action Level • Completeness • Laboratory Audit Results • Field and Laboratory Records • Level of Uncertainty in Reported Values

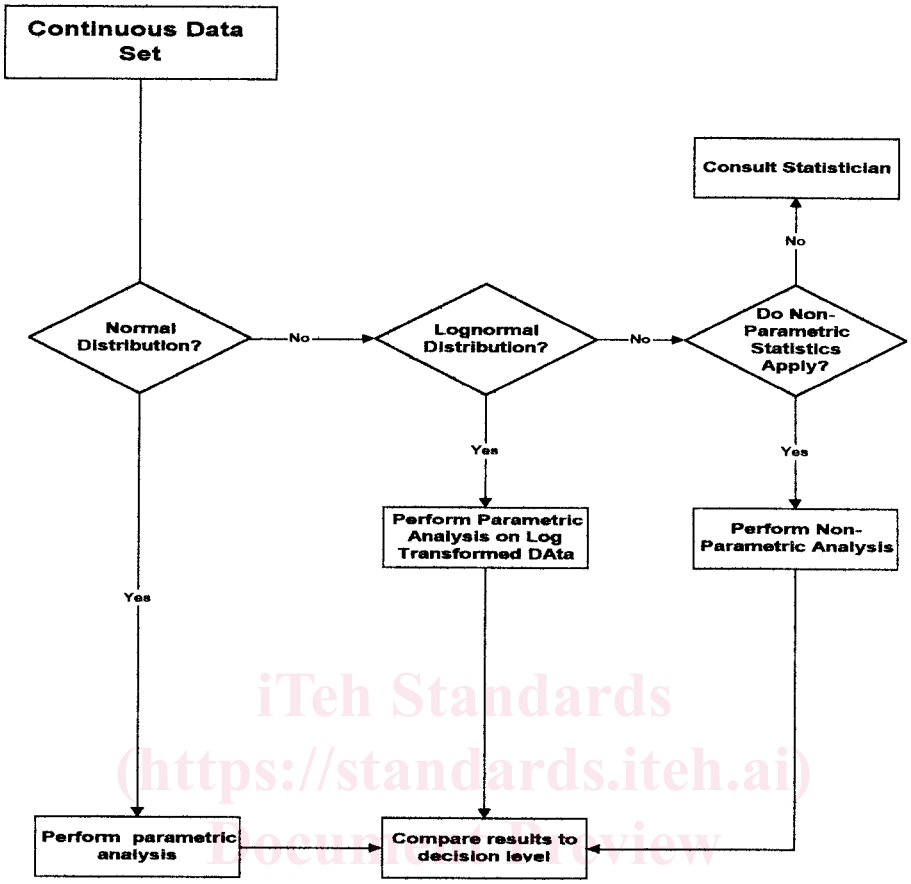

FIG. 1 General Strategy for Assessment of Continuous Data Sets

TABLE 2 Examples of Unmeasurable Factors Affecting the Integrity of Environmental Data Collection Efforts

• Biased Sampling/Subsampling	• Incorrect Dilutions
• Sampling Wrong Area or Material	• Incorrect Documentation
• Sample Switching (Mis-labeling)	• Matrix-Specific Artifacts
• Misweighing/Misaliquoting	

assurance programs, standard operating procedures, documentation practices, and staff training. Historically, efforts have been focused on the control of data collection errors through data review and the quality control process but little emphasis has been placed on the detection and evaluation of immeasurable errors using the quality assurance process. These unmeasurable sources of error are often the greatest source of uncertainty in the data collected for environmental projects. Examples of unmeasurable factors are given in Table 2.

5.6 Once the data assessment process has determined the degree to which the actual data collection effort met the expectations of the planners, the assessment process moves into the next phase to determine if the data generated by the effort can be verified and validated and whether it pass statistical tests for useability. These issues are discussed in the next sections.

6. Sources of Sampling Error

6.1 Sample collection may cause random or systematic errors. Random error affects the data by increasing the imprecision, whereas systemic error biases the data. The data assessment process should examine the available sampling records to determine if errors were introduced by improper sampling. A discussion of some of the more common sources of error follow.

6.1.1 *Random Error*:

6.1.1.1 Flaws in the sampling design which result in too few quality control samples being taken in the field can result in undetected errors in the sampling program. Adequate numbers of field QC samples (for example, field splits, co-located samples, equipment rinsate blanks, and trip blanks) are necessary to assess inconsistencies in sample collection practices, contaminated equipment, and contamination during the shipment process.

6.1.1.2 Variations (heterogeneity) in the media being sampled can cause concentration and property differences between and within samples. Field sampling and laboratory sub-sampling records should be examined to determine if heterogeneity was noted. This can explain wide variations in field and/or laboratory duplicate data.

6.1.1.3 Samples from the same population (including co-located samples) can be very different from each other. For example, one sample might be taken from a hot spot that was not visually obvious while the other was taken outside the

perimeter of the hot spot. If data from areas of high concentration is contained in data sets consisting primarily of uncontaminated material, statistical outlier analysis might suggest the sample data should be omitted from consideration when evaluating a site. This can cause serious decision errors. Prior to declaring the data point(s) to be outliers, it is important for the assessor to examine the QC records from the analysis yielding the suspect data. If the QC data indicates the system was in control and review of the raw sample data reveals no handling or calculation errors, the suspect data should be discussed in the assessor's report but it should not be discounted. The site history and operating records may hold clues to the possible existence of hot spots.

6.2 *Systematic Error*:

6.2.1 Flaws in the sampling design that result in sampling of inappropriate locations can result in significant bias in the data. The samples collected from such a flawed plan will not be representative of the population and can result in incorrect decisions. The assessor should review the sampling plan for signs of potential bias and discuss their findings in the final report.

6.2.2 Sampling tools and equipment can deselect certain parts of a sample based on the physical properties (density, particle size, multi-phasic materials, particle geometry, etc.). If the sample is biased because of some physical characteristic, then any constituent that is distributed in the material based on that characteristic, will be incorrectly reported. Both field and laboratory sampling equipment can introduce this type of bias.

6.2.3 Incorrect sampling procedures can cause losses of certain constituents of a sample such as volatile organics. Failure to control the loss of of constituents that exist in the gaseous state often comprises the collection of unsaturated media for volatile compound characterization. Deterioration of the sample can also occur after collection due to improper storage and transportation. For example, samples left standing in sunlight or in a hot vehicle can undergo photochemical reactions or lose volatile constituents.

6.2.4 Interactions between the sample and the material of the sampling equipment or container, or both, are potential sources of positive or negative bias.

6.2.5 Inappropriate preservation of the sample can cause a shift in chemical equilibria, loss of target analytes, or degradation, or all of these. For example, when analyzing a water sample for dissolved metals, addition of nitric acid to a water sample containing suspended solids might dissolve metals from the solids, resulting in an incorrect high concentration being reported. Failure to preserve water samples intended for organic analysis may allow significant biological alteration of the sample.

6.2.6 The time of day and prevailing weather conditions when samples are collected can affect the sample. For example, strong winds can blow dust that can contaminate the samples. Cool mornings or evening can lead to higher retention of volatile components in near-surface soil samples compared to the samples collected in the heat of the day.

6.2.7 The above examples only serve to illustrate the need for an experienced professional to review the sampling activities and to place the resulting analytical data in the proper context of the sampling activity. Such assessments add materially to the usability of the data.

7. Sources of Analytical Error

7.1 Variation in the analytical process may cause random or systematic error. Random error affects the data by increasing the imprecision, whereas systematic error increases the bias of the data. The data assessment process should examine the available analytical records to determine if errors were introduced in the data by the analytical process. Analytical results can also be impacted by sample matrix effects. Discussion of some of the more common sources these types of error follow.

7.1.1 *Random Error*:

7.1.1.1 Random errors in the analytical process are often uncontrollable and unobserved. They are usually distributed between positive and negative error and tend to cancel out and so have little effect. However, for any one measurement, random error can be significant.

7.2 *Systematic Error*—The bias resulting from systematic error can be either positive or negative but it affects all results in a data set(s) the same way. Sources of systematic error are most often associated with sample preparation or analysis. Incomplete digestion or insufficient reaction time during sample preparation are examples that can produce negatively biased results during the preparation process. Improperly calibrated instruments, incorrect standards, dirty detectors, and leaking sample introduction systems are examples of instrumental problems that cause systematic error. They are most often detected when reference samples and laboratory control samples fail to produce the expected results.

7.3 *Sample Matrix Effects*:

7.3.1 The sample matrix can introduce either systematic or random error in analytical results. Consistently high or low results (systematic error) can be obtained when the matrix contains a non-target constituent that interferes with the accurate measurement of the target analyte. The interfering substance must be uniformly distributed in the matrix to produce consistent deviations from the true value. If the interference is non-uniformly distributed in the matrix, the error will appear as a random error.

7.3.2 The relationship between the sample matrix and the analytical method can result in an important class of matrix errors. When the method selected is not appropriate to the matrix, errors may result. One of the most common types of mismatches of method and matrix is using methods designed for water analysis to analyze soils. Another is the use of methods designed for the analysis of naturally occurring materials, such as groundwater or soils, for the analysis of waste materials.

7.3.3 Most sample matrix and method selection errors can be detected by examining the results of matrix spike quality control samples where known amounts of the target analyte(s) are introduced into the sample before analysis. Spike results should be evaluated to determine the presence of any matrix effect. For certain types of analyses, simple dilution of the sample and re-analysis will demonstrate matrix effects when the second result, corrected for the dilution factor, is not consistent with the initial result.

8. Assessment of Environmental Data Sets

8.1 Data are usually verified and validated prior to comparing the results of environmental analysis to some decision level by suitable statistical processes. Data verification determines whether the laboratory carried out all steps required by the sampling and analysis plan or a contract, or both. After data is verified, it is validated. Validation examines the available laboratory data to determine whether an analyte is present or absent in a sample and the degree of overall uncertainty associated with the reported value. After data has been validated, it is normally compared to a decision level using suitable statistical techniques to determine the appropriate course of action.

8.2 The verification process compares the laboratory data package to a list of required data. These requirements are generated by two separate activities. The first is the contract for analytical services between the project and the laboratory and the second is the project sampling and analysis plan with its accompanying quality assurance project plan (QAPP) developed by project and laboratory staff. These two activities determine, *a priori*, the procedures the laboratory must use to produce data of known quality and the content of the analytical data package. Verification compares the material delivered by the laboratory against these requirements and produces a report that identifies those requirements which were not met (called *exceptions*). Verification exceptions normally identify:

8.2.1 Required steps not carried out by the laboratory (that is, incomplete analysis of all samples, lack of proper signatures, etc.),

8.2.2 Procedures not conducted at the required frequency (that is, too few blanks, duplicates, etc.),

8.2.3 Procedures which did not meet pre-set acceptance criteria (poor laboratory control sample recovery, unacceptable duplicate precision, etc).

8.3 The validation process begins with a review of the verification report or the laboratory data package, or both, to rapidly screen the areas of strength and weakness of the data set (tests of quality control). It continues with objective evaluation of sample data to confirm the presence or absence of an analyte (tests of detection) and to establish the statistical uncertainty (precision) of the measurement process for the analyte (test of uncertainty). Each data point is then qualified as to its integrity and dependability in the contest of all available laboratory data.

8.4 Examples of some important data project information that must be examined during the assessment of data are given in Table 1. Examples of some of the shortcomings that can occur are shown in Table 3. Some important characteristics of the data set that are frequently determined when examining quality control sample performance are given in Table 4. Data points not meeting the quality control criteria should be flagged and the magnitude and direction of any bias should be documented and made available for reference during the statistical evaluation processes that follow.

8.5 If project quality requirements are not met, further data assessment should not be undertaken until the data limitations are discussed with the project team. Data assessment cannot overcome basic design/execution flaws in the data collection process. Many times however, the project team can evaluate the problem and establish revised data quality objectives (different project expectations and new data requirements) factoring in the realities of the data collection effort which can then be used as the basis for data assessment.

TABLE 3 Common Data Requirements and Potential Shortcomings

Data Requirement	Potential Shortcomings
Number of samples	• Too few samples may have been collected or analyzed to be representative of the target population. • Too few samples were collected to narrow the estimate of the dispersion (variance, standard deviation, coeffiecient of variation, etc) of the measured results to acceptable levels.
Location of samples	• Samples were collected from the wrong locations due to error or inaccessibility.
Analyte/method	• Incorrect choice of analyte/method for the sample matrix
Quality control	• Measurement system not calibrated • Contamination found in field, trip, or method blanks • Method performance on reference samples unsatisfactory • Calculation errors
Method sensitivity	• Failure to meet minimum detectable limits
Method precision	• Failure to achieve satisfactory duplicate results for analysis of field samples due to sample characteristics or other analytical problems
Method bias	• Failure to demonstrate method performance on reference materials or analytical standards • Failure to demonstrate satisfactory target analyte spike/surrogate recoveries in field sample analysis
Interferences	• Presence of unanticipated materials/analytes in field samples that render accurate analysis suspect
Action level	• Not provided

9. Statistical Evaluation of Data Sets

9.1 The US EPA *Guidance for Data Quality Assessment, QA/G-9*(1)[4] is a good source for information on the following statistical approaches to data assessment.

9.1.1 *Continuous Data*:

9.1.1.1 Continuous data are data where the values of the individual samples may vary form zero to any maximum value. Examples of continuous data are the concentration of a constituent in soil or the percent moisture in an environmental sample. This is the type of information most frequently collected in environmental waste management projects. It is normally used to establish a statistical characteristic of the target population which is then compared to a decision level resulting in an action. This is referred to as the "decision rule" and normally takes the form:

[4] The boldface numbers given in parentheses refer to a list of references at the end of the text.

TABLE 4 Information Derived From Quality Control Samples[A]

Type of QC Sample	Type of Information										
	Precision			Bias			Contamination				
	Sampling	Splitting	Preparation and Analysis	Spiking	Field/ Shipping/ Storage	Laboratory	Containers and Preservatives	Field Environment	Equipment	Cross-Contamination	Laboratory
Replicates											
Splits, field		X	X								
Collocated, field	X		X								
Splits, laboratory		X	X								
Spikes											
Field				X	X	X					
Laboratory, matrix				X		X					
Blanks											
Trip							X			X	
Field							X	X		X	
Equipment							X	X	X	X	
Method											X

[A] Can be assessed using numerical techniques.

If (characteristic of the population) (method of comparison) (action level), then (action). Otherwise, (alternate action).

where the items in parentheses are determined by the project team on a project-specific basis. Two examples are:

If (the average concentration of mercury in the top 15 cm of soil over the site) (is greater than) (100 mg/kg), then (excavate the top 30 cm of soil and dispose of in a RCRA landfill). Otherwise, (no remediation is required).

and:

If (less than one half the randomly selected waste oil drums have an average organic halide concentration) (of less than 500 ppm), then (composite the contents of all drums and use it as boiler fuel). Otherwise, (send all drums to a RCRA treatment and disposal facility).

9.1.2 Before beginning the statistical interpretation of a continuous data set, plots of the data should be constructed to guide the statistical interpretation of the data that follows. Examples of the types of plots that can be constructed are:

9.1.2.1 Concentration versus time, and

9.1.2.2 Concentration versus location in two or three dimensions as appropriate

9.1.2.3 These types of plots provide a picture of the distribution of the parameter of interest and permit the identification of strata as a function of time or location. Plots also identify data points which are abnormally high or low with respect to the surrounding data. These are potential outliers and they can be more rigorously evaluated by the verification and validation process to determine whether there is an analytically-related explanation. This information will identify random or stratified data sets and outliers or QC-failed data prior to statistical evaluation.

9.1.3 *Normally Distributed Data*:

9.1.3.1 Once the data evaluation described above have been completed, statistical techniques should be used to evaluate the data against the decision criteria. The key steps in the sequence to evaluate continuous data are shown in Fig. 2.

9.1.3.2 The first step is to determine if the data are normally distributed. That is, are there an approximately equal number of values that are less than and greater than the mean and is the range of values approximately equal on either side of the mean (See Fig. 2). This property of normal distribution is a reasonable model of the behavior of certain random phenomena and can be used to approximate many kinds of data.

9.1.3.3 There are several graphical techniques that can be applied to determine if data are normally distributed. Among them are: stem- and leaf- diagrams, histogram/frequency plots, box and whiskers plots, ranked data plots, quantile plots, and, normal probability plots (quantile-quantile plots).

9.1.3.4 The use of plots to determine if data are normally distributed involves a subjective decision on the part of the individuals making the assessment. This is easy when the data are very non-normal but more difficult as the data approach normal distribution. There are series of formal numerical methods to test for normal distribution. The Shapiro - Wilkes test can be applied to data sets of less than 50 samples. For larger size sample sets (up to 1000 data points), Fillben's Statistic is frequently used. Both methods are difficult to implement by hand because of the large number of calculations required but are readily accomplished by computer programs.

9.1.4 Once the normal distribution of the data is shown, the straightforward calculation of the statistical quantities used in the project decision rule can be performed. For example, the two-sided confidence limits for the mean (that is a parametric population characteristic) can be performed. This allows the data user to determine the interval in which the true mean is expected to be found with specified confidence. The mean lead level, interval and confidence are frequently expressed as:

the level of lead in the soil is $X \pm x$ *at the* 95 % *confidence level*

9.1.4.1 The width of the interval, $\pm 2x$, can be calculated for varying degrees of confidence (selected by the data user) to meet project-specific tolerable error rates for making the