Designation: ~~D7048 – 04 (Reapproved 2010)~~ D7048 – 16

Standard Guide for Applying Statistical Methods for Assessment and Corrective Action Environmental Monitoring Programs[1]

This standard is issued under the fixed designation D7048; 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 The scope and purpose of this guidance is to present a variety of statistical approaches for assessment, compliance and corrective action environmental monitoring programs. Although the methods provided here are appropriate and often optimal for many environmental monitoring problems, they do not preclude use of other statistical approaches that may be equally or even more useful for certain site-specific applications.

1.2 In the following sections, ~~complete~~the details of select statistical procedures used in assessment and corrective action programs for environmental monitoring (soil, groundwater, air, surface water, and waste streams) are presented.

1.3 The statistical methodology described in the following sections should be used as guidance. Other methods may also be appropriate based on site-specific conditions or for monitoring situations or media that are not presented in this document.

1.4 This practice offers an organized collection of information or a series of options and does not recommend a specific course of action. This document cannot replace education, experience and professional judgements. Not all aspects of this practice may be applicable in all circumstances. This ASTM standard is not intended to represent or replace the standard of care by which the adequacy of a given professional service must be judged without consideration of a project's many unique aspects. The word Standard in the title of this document only means that the document has been approved through the ASTM consensus process.

1.5 *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 requirements prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

D653 Terminology Relating to Soil, Rock, and Contained Fluids
D5092 Practice for Design and Installation of Groundwater Monitoring Wells
D5792 Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives
D6250 Practice for Derivation of Decision Point and Confidence Limit for Statistical Testing of Mean Concentration in Waste Management Decisions
D6312 Guide for Developing Appropriate Statistical Approaches for Groundwater Detection Monitoring Programs

3. Terminology

3.1 *Definitions*—For definitions of common terms in this guid, see Terminology D653.

3.2 ~~*Definitions:*~~*Definitions of Terms Specific to This Standard:*

~~3.1.1 *assessment monitoring*—investigative monitoring that is initiated after the presence of a contaminant has been detected in groundwater above a relevant criterion at one or more locations. The objective of the program is to determine if there is a statistical exceedance of a standard or criteria at a Potential Area of Concern (PAOC) or at the groundwater discharging to surface water interface, and/or to quantify the rate and extent of migration of constituents detected in groundwater above applicable criteria.~~

[1] This guide is under the jurisdiction of ASTM Committee D18 on Soil and Rock and is the direct responsibility of Subcommittee D18.21 on Groundwater and Vadose Zone Investigations.

Current edition approved ~~July 1, 2010~~Oct. 1, 2016. Published ~~September 2010~~October 2016. Originally approved in 2004. Last previous edition approved in ~~2004~~2010 as ~~D7048~~D7048–04(2010).~~–04.~~ DOI: ~~10.1520/D7048-04R10.~~10.1520/D7048-16.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

~~3.1.2 *compliance monitoring*—as specified under 40 CFR 264.99, compliance monitoring is instituted when hazardous constituents have been detected above a relevant criterion at the compliance point during RCRA detection monitoring. Groundwater samples are collected at the compliance point, facility property boundary, and upgradient monitoring wells for analysis of hazardous constituents to determine if they are leaving the regulated unit at statistically significant concentrations above background.~~

3.2.1 *corrective action monitoring*—under ~~RCRA,~~ RCRA (in the United States), corrective action monitoring is instituted when hazardous constituents from a RCRA regulated unit have been detected at statistically significant concentrations between the compliance point and the downgradient facility property boundary as specified under 40 CFR 264.100. Corrective action monitoring is conducted throughout a corrective action program that is implemented to address groundwater contamination. At non-RCRA sites, corrective action monitoring is conducted throughout the active period of corrective action to determine the progress of remediation and to identify statistically significant trends in groundwater contaminant concentrations.

~~3.1.4 *detection limit, DL*—the true concentration at which there is a specified level of confidence (for example, 99 % confidence) that the true concentration is greater than zero.~~

~~3.1.5 *detection monitoring*—a program of monitoring for the express purpose of determining whether or not there has been a release of a contaminant to groundwater. Under RCRA, Detection Monitoring involves collection of groundwater samples from compliance point and upgradient monitoring wells on a semi-annual basis for analysis of hazardous constituents of concern, as specified under 40 CFR 264.98. Results are evaluated to determine if there is a statistically significant exceedance of the groundwater protection criterion and/or background. At non-RCRA sites, monitoring is conducted in a similar manner and results are compared to criteria to determine if there is a statistically significant exceedance.~~

~~3.1.6 *direct push sampling*—groundwater sampling conducted with a device that is temporarily pushed into the ground with a hydraulic system or with a hammer. After groundwater sample collection, the device is removed from the ground. Examples include Geoprobe®, Hydropunch® direct push, and environmental soil probe.~~

~~3.1.7 *false negative rate*—the rate at which the statistical procedure does not indicate contamination when contamination is present.~~

3.2.2 *false positive rate*—the rate at which the statistical procedure ~~indates~~indicates contamination when contamination is not present.

3.2.3 *lognormal distribution*—a frequency distribution whose logarithm follows a normal distribution.

3.2.4 *lower confidence limit, LCL*—a lower limit that has a specified probability (for example, 95 %) of including the true concentration (or other parameter). Taken together with the upper confidence limit, forms a confidence interval that will include the true concentration with confidence level that accounts for both tail areas (for example, 90 %).

3.2.5 *lower prediction limit, LPL*—a statistical estimate of the minimum concentration that will provide a lower bound for the next series of *k* measurements from that distribution, or the mean of *m* new measurements for each of *k* sampling locations, with specified level of confidence (for example, 95 %).

3.2.6 *nonparametric*—a term referring to a statistical technique in which the distribution of the constituent in the population is unknown and is not restricted to be of a specified form.

3.2.7 *nonparametric prediction limit*—the largest (or second largest) of *n* background samples. The confidence level associated with the nonparametric prediction limit is a function of *n*,*m* and *k*.

3.2.8 *normal distribution*—a frequency distribution whose plot is a continuous, infinite, bell-shaped curve that is symmetrical about its arithmetic mean, mode and median (which are numerically equivalent). The normal distribution has two parameters, the mean and variance.

3.2.9 *outlier*—a measurement that is statistically inconsistent with the distribution of other measurements from which it was drawn.

3.2.10 *parametric*—a term referring to a statistical technique in which the distribution of the constituent in the population is assumed to be known.

~~3.1.17 *quantification limit, QL*—a lower limit on the concentration at which quantitative determinations of an analyte's concentration in the sample can be reliably made during routine laboratory operating conditions. The QL is typically described quantitatively as the true concentration at which the signal to noise ratio of measured concentration or instrument response is 10:1. The signal to noise ratio is often determined by a percent relative standard deviation of 10 %.~~

3.2.11 *potential area of concern*—areas with a documented release or likely presence of a hazardous substance that could pose an unacceptable risk to human health or the environment.

~~3.1.19 *phase I environmental site assessment*—non-intrusive investigation that identifies PAOCs which may require further investigation in subsequent phases of work.~~

~~3.1.20 *phase II environmental site assessment, ESI*—intrusive survey to confirm or deny existence of a release into the environment at a PAOC at levels which may adversely impact public health or the environment.~~

3.2.12 *upper confidence limit, UCL*—an upper limit that has a specified probability (for example, 95 %) of including the true concentration (or other parameter). Taken together with the lower confidence limit, the UCL forms a confidence interval that will include the true concentration with confidence level that accounts for both tail areas.

3.2.13 *upper prediction limit, UPL*—a statistical estimate of the maximum concentration that will not be exceeded by the next series of *k* measurements from that distribution, or the mean of *m* new measurements for each of *k* sampling locations, with specified level of confidence (for example, 95 %) based on a sample of *n* background measurements.

3.3 *Symbols:* μ = the true population mean of a constituent

~~$\bar{x}$~~ $\bar{x}$ = the sample-based mean or average concentration of a constituent computed from *n* background measurements which differs from μ because of sampling variability, and other error

σ^2 = the true population variance of a constituent

s^2 = the sample-based variance of a constituent computed from *n* background measurements

s = the sample-based standard deviation of a constituent computed from *n* background measurements

~~$\bar{y}$~~ $\bar{y}$ = the mean of the natural log transformed data (also the natural log of the geometric mean)

s_y = the standard deviation of the natural log transformed data

n = the number of background (offsite or upgradient) measurements

k = the number of future comparisons for a single monitoring event (for example, the number of downgradient monitoring wells multiplied by the number of constituents to be monitored) for which statistics are to be computed

α = the false positive rate for an individual comparison (that is, one sampling location and constituent)

m = the number of onsite or downgradient measurements used in computing the onsite mean concentration

α^* = the site-wide false positive rate covering ~~all~~the sampling locations and constituents

t = the 100(1 − α) percentage point of Student's *t*-distribution on *n* − 1 degrees of freedom

H_L = the factor developed by Land (1971) (**1**)[3] to obtain the lower 100(α) % confidence limit for the mean of a lognormal distribution

H_U = the factor developed by Land (1971) (**1**) to obtain the upper 100(α) % confidence limit for the mean of a lognormal distribution

4. Summary of Guide

4.1 The guide is summarized as Figs. 1-7. These figures provides a flow-chart illustrating the steps used in computing the comparisons to regulatory or health based groundwater protection standard (GWPS) in assessment and corrective action environmental monitoring programs.

5. Significance and Use

5.1 The principal use of this standard is in assessment, compliance and corrective action environmental monitoring programs (for example, for ~~any~~a facility that could potentially contaminate groundwater). The significance of the guidance is that it presents a statistical method that allows comparison of groundwater data to regulatory and/or health based limits.

5.2 Of course, there is considerable ~~USEPA~~ support for statistical methods applied to detection, assessment and corrective action monitoring programs that can be applied to environmental ~~investigations. For example, the 90 % upper confidence limit (UCL) of the mean is used in SW846 (Chapter 9) for determining if a waste is hazardous. If the UCL is less than the criterion for a particular hazardous waste code, then the waste is not a hazardous waste even if certain individual measurements exceed the criterion. Similarly, in the USEPA Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Addendum to the Interim Final Guidance (1992)~~ sites. ~~(2), confidence intervals for the mean and various upper percentiles of the distribution are advocated for assessment and corrective action. Interestingly, both the 1989 and 1992 USEPA guidance documents (2, 3) suggest use of the lower 95 % confidence limit (LCL) as a tool for determining whether a criterion has been exceeded in assessment monitoring. The latest USEPA guidance in this area (that is, the draft USEPA Unified Statistical Guidance) calls for use of the LCL in assessment monitoring and the UCL in corrective action. In this way, corrective action is only triggered if there is a high degree of confidence that the true concentration has exceeded the criterion or standard, whereas corrective action continues until there is a high degree of confidence that the true concentration is below the criterion or standard. This is the general approach adopted in this guide, as well.~~

NOTE 1—For example, in the United States, the 90 % upper confidence limit (UCL) of the mean is used in USEPA's SW846 (Chapter 9) for determining if a waste is hazardous. If the UCL is less than the criterion for a particular hazardous waste code, then the waste is not a hazardous waste even if certain individual measurements exceed the criterion. Similarly, in the USEPA Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Addendum to the Interim Final Guidance (1992) (**2**), confidence intervals for the mean and various upper percentiles of the distribution are advocated for assessment and corrective action. Interestingly, both the 1989 and 1992 USEPA guidance documents (**2, 3**) suggest use of the lower 95 % confidence limit (LCL) as a tool for determining whether a criterion has been exceeded in assessment monitoring.

The latest guidance in this area calls for use of the LCL in assessment monitoring and the UCL in corrective action. In this way, corrective action is only triggered if there is a high degree of confidence that the true concentration has exceeded the criterion or standard, whereas corrective action continues

[3] The boldface numbers in parentheses refer to a list of references at the end of this standard.

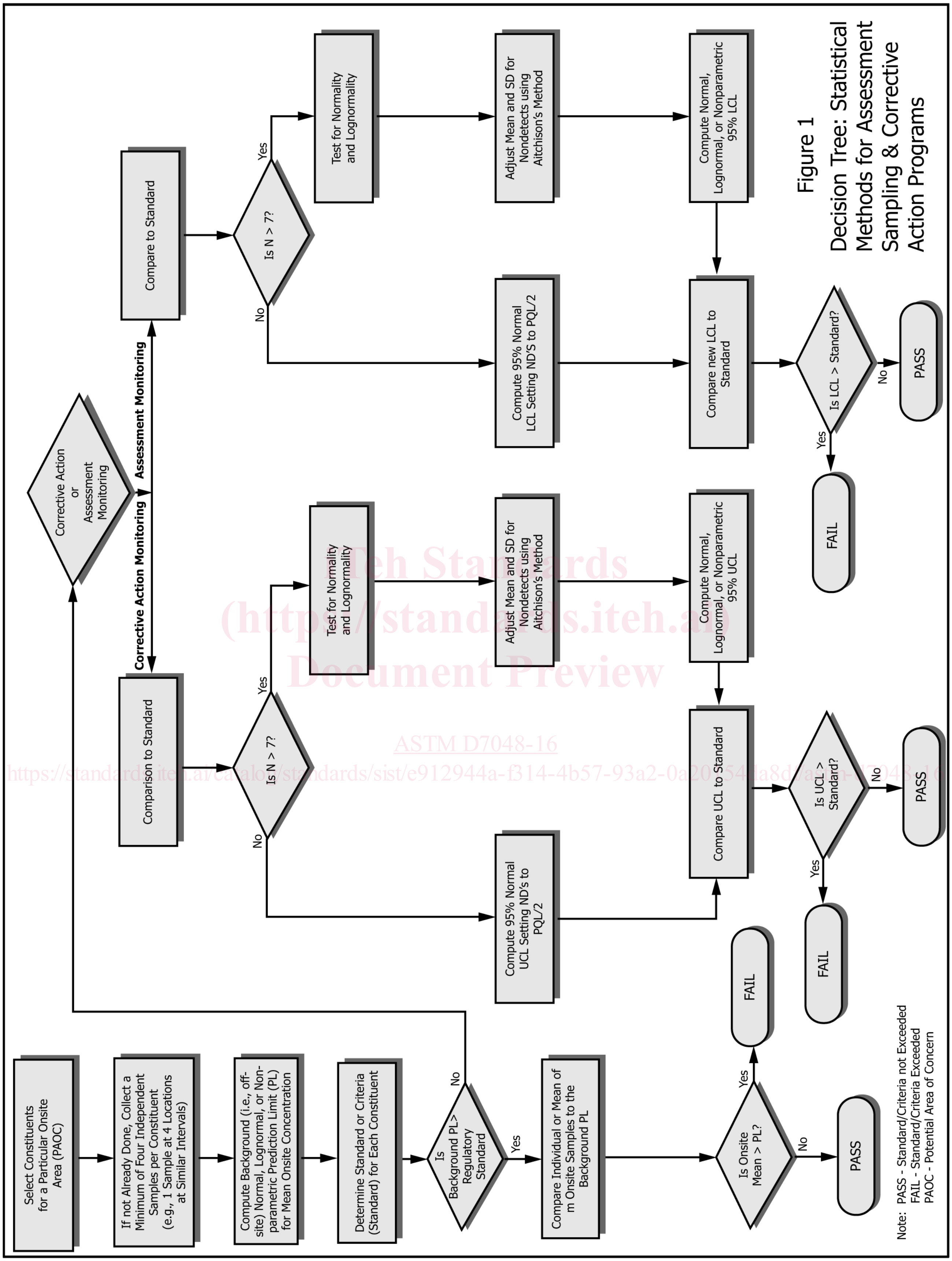

FIG. 1 Decision Tree—Statistical Methods for Assessment Sampling and Corrective Action Programs

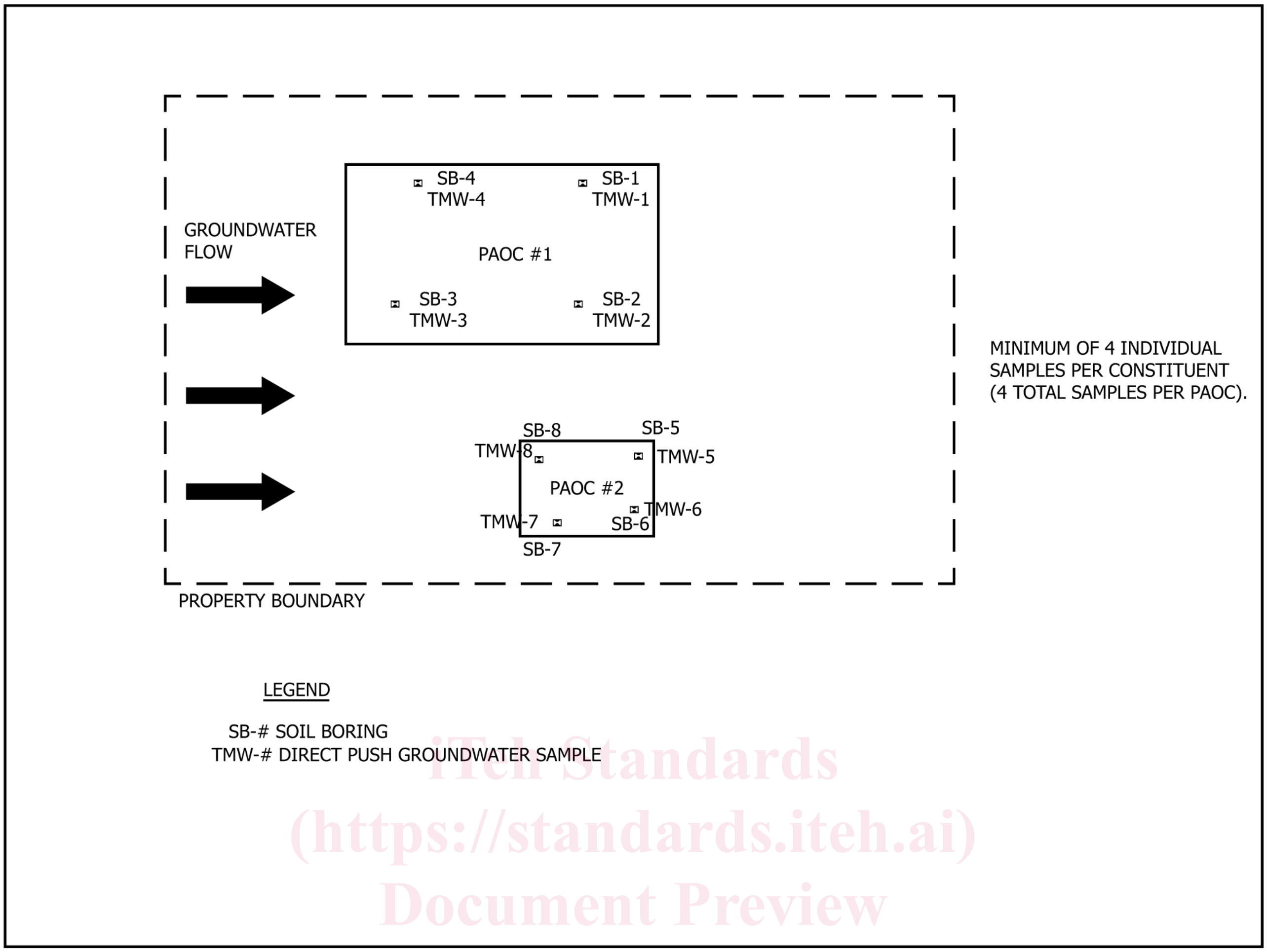

FIG. 2 Single PAOC Comparison to a Standard/Criteria

~~until there is a high degree of confidence that the true concentration is below the criterion or standard. This is the general approach adopted in this guide, as well.~~

5.3 There are several reasons why statistical methods are ~~essential~~needed in assessment and corrective action monitoring programs. First, a single measurement indicates very little about the true concentration in the sampling location of interest, and with only one sample ~~there is no way of knowing~~ it cannot be determined if the measured concentration is a typical or an extreme value. The objective is to compare the true concentration (or some interval that contains it) to the relevant criterion or standard. Second, in many cases the constituents of interest are naturally occurring (for example, metals) and the naturally existing concentrations may exceed the relevant criteria. In this case, the relevant comparison is to background (for example, off-site soil or upgradient groundwater) and not to a fixed criterion. As such, background data ~~must~~should be statistically characterized to obtain a statistical estimate of an upper bound for the naturally occurring concentrations so that it can be confidently determined if onsite concentrations are above background levels. Third, there is often a need to compare numerous potential constituents of concern to criteria or background, at numerous sampling locations. By chance alone there will be exceedances as the number of comparisons becomes large. The statistical approach to this problem can ~~insure that~~ decrease the potential for false positive ~~results are minimized.~~results.

5.4 Statistical methods for detection monitoring have been well studied in recent years (see Gibbons, 1994a, 1996, USEPA 1992 (**2, 4, 5**) and Practice D6312, formerly PS 64-96 authored by Gibbons, Brown and Cameron, 1996). Although equally important, statistical methods for assessment monitoring, Phase I and II ~~investigations,~~Investigations, on-going monitoring and corrective action monitoring have received less attention, (Gibbons and Coleman, 2001) (**6**).

5.5 The guide is summarized in Fig. 1, which provides a flow-chart illustrating the steps in developing a statistical evaluation method for assessment and corrective action programs. Fig. 1 illustrates the various decision points at which the general comparative strategy is selected, and how the statistical methods are to be selected based on site-specific considerations.

6. Procedure

6.1 In the following, the general conceptual and statistical foundations of the sampling program are described. Following this general discussion, media-specific details (that is, soil, groundwater, and waste streams) are provided.

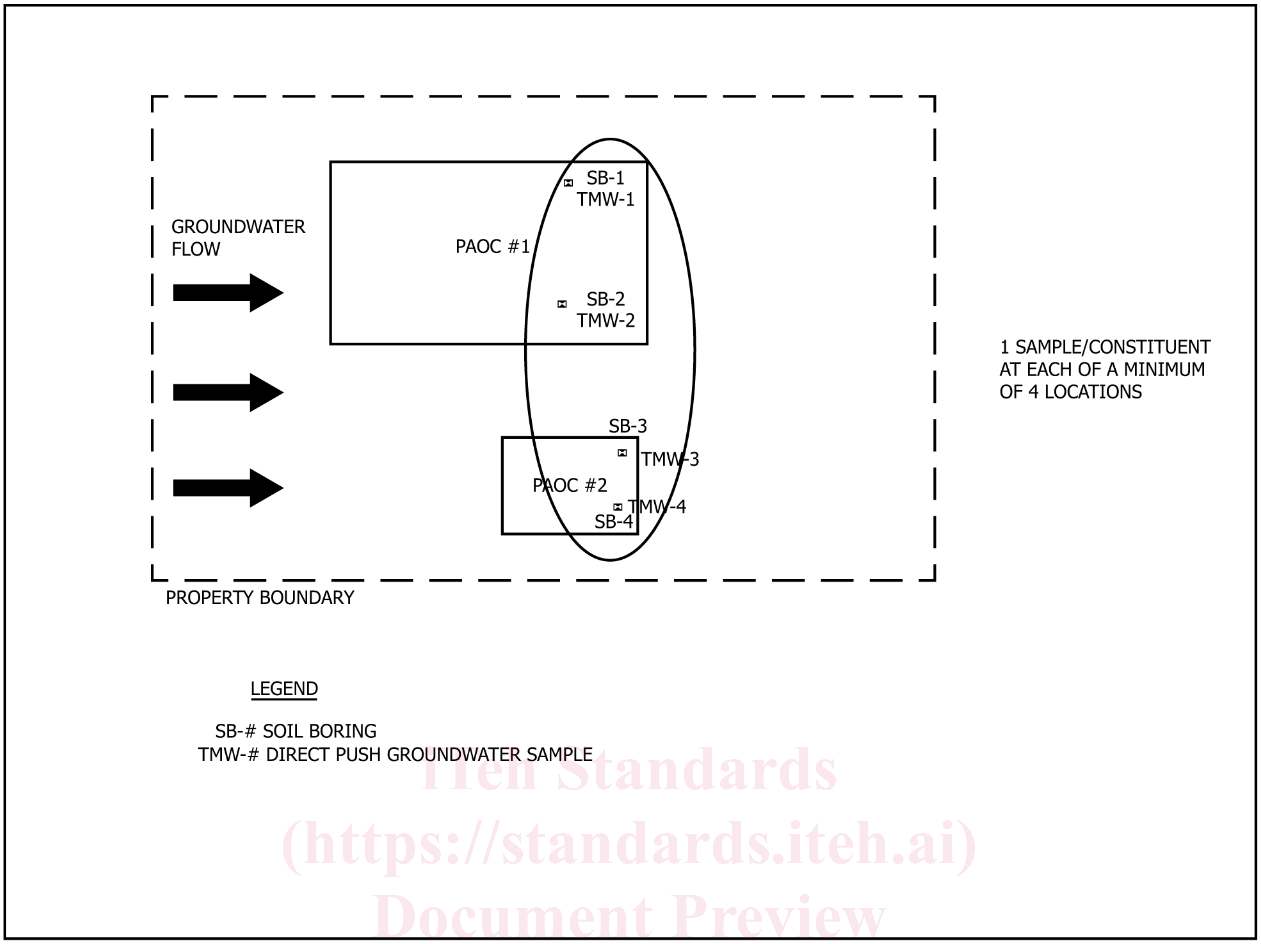

FIG. 3 Multiple PAOC Comparison to a Standard/Criteria

6.1.1 Identify relevant constituents for the specific type of facility, media (for example, ~~soil, groundwater etc.)~~soil and/or groundwater) and area of interest. A facility is generally comprised of a series of subunits or "source areas" that may have a distinct set of sampling locations and relevant constituents of concern (referred to as a PAOC). The subunit may consist of a single sampling point or collection of sampling points. In some cases, the entire site may comprise the area of interest and all sampling locations are considered jointly. The boundaries of the "source area" or "decision unit" should be defined. In ~~all~~most cases, the owner/operator should select the smallest ~~possible~~practical list of constituents that adequately characterize the source area in terms of historical use.

6.1.2 For each constituent obtain the appropriate regulatory criterion or standard (for example, maximum contaminant level, MCL) if one is available. The appropriate criterion or standard should be selected based on relevant pathways (for example, direct contact, ingestion, inhalation) and appropriate land use criteria (for example, commercial, industrial, residential).

6.1.3 For each constituent which may have a background concentration higher than the relevant health based criterion, set "background" to the upper 95 % confidence prediction limit (UPL) as described in the Technical Details section. The prediction limits are computed from ~~all~~ available data collected from background, or outside source areas that are unlikely to be contaminated, upstream, upwind or upgradient locations only. Henceforth, background refers to ~~any of~~ these types of offsite sources. The background data are first screened for outliers and then tested for normality and lognormality (see Technical Details section).

6.1.3.1 If the test of normality cannot be rejected (for example, at the 95 % confidence level), background is equal to the 95 % confidence normal prediction limit.

6.1.3.2 If the test of normality is rejected but the test of lognormality cannot be rejected, background is equal to the 95 % confidence lognormal prediction limit.

6.1.3.3 If the data are neither normal nor lognormal, or the detection frequency is less than 50 %, background is the nonparametric prediction limit. When we are interested in a single potentially impacted measurement, normal, lognormal, and nonparametric prediction limits are identical with respect to the parameter being compared (that is, an individual measurement). However, when the comparison to background is for an onsite/downgradient mean concentration, they differ in that the nonparametric prediction limit is for the median whereas the parametric prediction limits are for the mean. This limitation is unavoidable, so whenever ~~possible,~~practical, parametric prediction limits should be used. Note that, if the detection frequency is