STATE OF CONNECTICUT

DEPARTMENT OF ENERGY AND ENVIRONMENTAL PROTECTION

Guidance for Calculating the 95% Upper Confidence Level for Demonstrating Compliance with the Remediation Standard Regulations

May 30, 2014

Robert J. Klee, Commissioner

79 Elm Street, Hartford, CT 06106
www.ct.gov/deep/remediation
860-424-3705

TABLE OF CONTENTS

LIST OF ACRONYMS iii

DEFINITION OF TERMS iv

1. Introduction 1

1.1 Definition of 95 % Upper Confidence Level 1

1.2 Data Quality Considerations 2

1.3 Applicability 2

1.3.1 Soil 3

1.3.2 Groundwater 3

1.4 Document Organization 4

2. Developing a Data Set for a Release Area 4

2.1 Data Selection for 95% UCL Calculation for a Release Area 5

2.2 Non-Detect Soil Results in a Release Area Data Set 7

2.3 Quality Control Soil Sample Results in a Release Area Data Set 7

3. Developing a Data Set for a Groundwater Plume 7

3.1 Data Selection for 95% UCL Calculation for a Groundwater Plume 8

3.2 Non-Detect Results in a Groundwater Plume Data Set 8

3.3 Quality Control Results in a Groundwater Plume Data Set 8

4. Evaluating the Data Set 9

4.1 Distribution of COC Concentrations in the Environment 9

4.2 Appropriate Data Set Size 10

4.3 Statistical DQOs 10

4.3.1 Randomness of Data Set 10

4.3.2 Strength of Data Set 10

4.3.3 Skewness of Data Set 11

5. Statistical Calculation Methods 11

5.1 Data Distributions 12

5.2 Handling of Non-Detect Results in Statistical Calculations 12

6. Requests for Alternative Method for Demonstrating Compliance Using the 95% UCL 12

7. References 14

Example 1. Effect of Collection of Additional Samples after Delineation 16

Appendix A Laboratory Analytical Limits and How They Affect Non-Detect Results 20

Appendix B Summary of UCL Calculation Methods 21

Appendix C ProUCL Version 5.0 Calculation Method Decision Summary Tables[1] 22

LIST OF ACRONYMS

95% UCL	Ninety-five percent upper confidence level of the arithmetic mean
COC	Constituent of Concern
CSM	Conceptual Site Model
DEC	Direct Exposure Criteria
DPH	Department of Public Health
DQO	Data Quality Objective
GWPC	Groundwater Protection Criteria
KM	Kaplan-Meier
MDL	Method Detection Limit
PMC	Pollutant Mobility Criteria
QC	Quality Control
RL	Reporting Limit
ROS	Regression on Order Statistics
RSR	Remediation Standard Regulation
SCGD	Site Characterization Guidance Document
SWPC	Surface Water Protection Criteria

DEFINITION OF TERMS

Term	Definition
Analyte	An analyte means a substance being measured by a laboratory analytical procedure.
Gamma Distribution	Gamma distribution is an arrangement of data in which applying the gamma function transforms the data to have a normal distribution. The gamma function is an extension of the factorial function.
Lognormal Distribution	Lognormal distribution is an arrangement of data in which the logarithms of the data have a normal distribution.
Method Detection Limit	A method detection limit is a statistically-calculated result used to evaluate precision and accuracy of analytical results obtained by a given method process. (further details are provided in Appendix A)
Non-Detect	A non-detect is an analytical result that is below the level that could be detected or reliably quantified using a particular analytical method.
Non-Parametric	Non-parametric describes statistical methods that do not assume that the data set has any known distribution. Non-parametric methods make few assumptions about the underlying distribution; therefore, they can be applied to data sets with any distribution, including those that are unknown.
Normal Distribution	Normal distribution is an arrangement of data graphically represented as a bell-shaped frequency curve symmetrical about the mean.
Randomness	Randomness of the data set is the degree to which the introduction of bias has been reduced and the resulting data points are more likely to be independently and equally distributed within the population (e.g., release area, groundwater plume).
Reporting Limit	The reporting limit is the minimum concentration of an analyte that can be reliably quantified and reported by the laboratory using a specific laboratory analytical method during routine laboratory operating conditions. Reporting limits are determined by the laboratory, are above instrument detection limits, and are adjusted based on laboratory and sample conditions. (further details are provided in Appendix A)
Skewness	Skewness is the degree to which a data set is not in balance around the mean (asymmetrical or lopsided). Distributions with extreme values (outliers) above the mean have positive skew, and the distributions with outliers below the mean have negative skew.
Strength	Strength is a measure of the relationship between variables. The strength of the data set is directly related to the size of the data set. The larger the data set size, the stronger the data set and therefore the more reliable and robust the results of the 95% UCL estimate.

1. INTRODUCTION

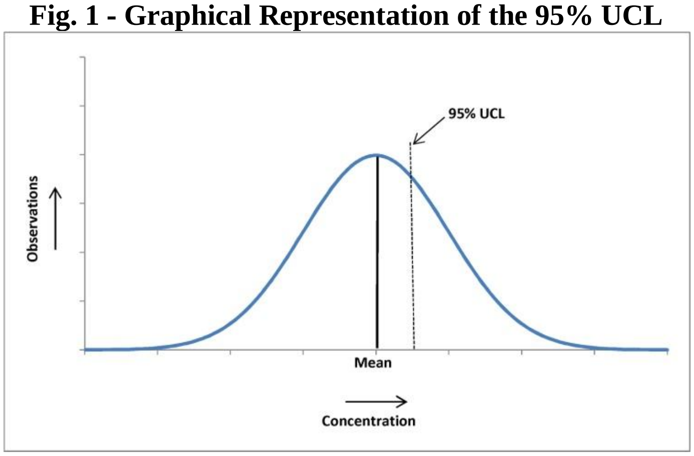
In accordance with the Connecticut Remediation Standard Regulations (RSRs), Sections 22a-133k-1 through 22a-133k-3 of the Regulations of Connecticut State Agencies, the ninety-five percent upper confidence level of the arithmetic mean (95% UCL) is a self-implementing option that may be used to demonstrate compliance with the direct exposure criteria (DEC), pollutant mobility criteria (PMC), groundwater protection criteria (GWPC), and the surface water protection criteria (SWPC).

A workgroup consisting of personnel from the Department of Energy and Environmental Protection (Department) and the Department of Public Health (DPH), along with environmental professionals from the private-sector, developed this guidance for Calculating the 95% Upper Confidence Level for Demonstrating Compliance with the Remediation Standard Regulations (Document) to guide the regulated community in performing the 95% UCL statistical calculation on soil and groundwater data sets to demonstrate compliance with certain RSR criteria.

1.1 Definition of 95 % Upper Confidence Level

The "ninety-five percent upper confidence level of the arithmetic mean" is defined in the RSRs as a value that, when repeatedly calculated for randomly drawn subsets of size *n* from a population, equals or exceeds the population arithmetic mean ninety-five percent of the time. The arithmetic mean is calculated by adding up all the numbers in a data set and dividing the result by the total number of data points. This is quite different than a geometric mean, which is calculated by multiplying the numbers in the data set, and taking the *n*th root of the result. In order to use the 95% UCL to demonstrate compliance with the RSRs, it is important to ensure that the 95% UCL is calculated for the arithmetic mean (intended to be used when the individual data points are independent of each other), since the use of the geometric mean (intended to be used when the individual data points are dependent on the previous data points) would be inappropriate.

Fig. 1 - Graphical Representation of the 95% UCL



1.2 Data Quality Considerations

Data quality objectives (DQOs) are specific goals developed to ensure that a sufficient quality and quantity of data are collected to make appropriate decisions. Prior to demonstrating compliance with applicable RSR criterion, the environmental professional is expected to have completed the characterization of the subject release area or groundwater plume in accordance with prevailing standards and guidelines, including the Department's *Site Characterization Guidance Document* (SCGD)[1]. Data collected according to DQOs related to the characterization of the site may or may not support the use of statistics to demonstrate compliance.

Prior to calculating the 95% UCL to demonstrate compliance, the environmental professional should ensure that the data set meets the underlying assumption of the statistical methods used in the calculation. For example, the statistical methods described in this Document for calculating the 95% UCL are based, in part, on the assumption of random sampling. During the investigation and characterization of sites, sampling is typically focused on identifying and delineating areas of contamination. Such sampling program will likely produce a biased data set with samples mostly collected for delineation, which represent the lower concentrations of the release area rather than being randomly distributed throughout the release area. This resulting data set would not be appropriate for use in a 95% UCL calculation without the collection of additional samples from the underrepresented sections of the release area. The environmental professional needs to ensure that the particular data set is of sufficient quality and quantity (as discussed in Section 4) to represent the subject release area or subject groundwater plume to use with these statistical methods. The collection of additional samples may be necessary to meet the underlying assumptions of the statistical methods.

1.3 Applicability

This section provides an overview of the provisions of the RSRs that provide the option to use a 95% UCL to demonstrate compliance. In the event of inconsistencies between this Document and the RSRs, the language in the regulations supercedes this Document. In addition, under certain circumstance Federal Regulations may also apply which may limit the applicability of this Guidance Document.

[1] The Department's Site Characterization Guidance Document (dated September 2007, revised December 2010) provides the Department's recommendations for the multi-phased approach to site characterization using conceptual site modeling.

1.3.1 Soil

The RSRs allow the calculation of a 95% UCL to demonstrate compliance with the DEC and the PMC for an individual release area.

1.3.1.1 *Direct Exposure Criteria*

Section 22a-133k-2(e)(1) of the RSRs provides the option to demonstrate compliance with the applicable DEC if the 95% UCL of all soil sample results from the subject release area is equal to or less than the applicable DEC. Accounting for any institutional controls implemented, only those soil sample results from locations and depths where the DEC apply should be used to calculate the 95% UCL for DEC compliance.

1.3.1.2 *Pollutant Mobility Criteria*

Section 22a-133k-2(e)(2)(A) of the RSRs provides the option to demonstrate compliance with PMC if the 95% UCL of at least 20 samples collected from the subject release area and above the water table is equal to or less than the applicable PMC. Accounting for any institutional controls implemented, only those soil sample results from locations and depths where the PMC apply should be used to calculate the 95% UCL for PMC compliance.

1.3.2 Groundwater

The RSRs also allow the calculation of a 95% UCL to demonstrate compliance with the GWPC and the SWPC for a groundwater plume.

1.3.2.1 *Groundwater Protection Criteria*

Section 22a-133k-3(g)(2)(B) of the RSRs provides the option to demonstrate compliance with the GWPC if the 95% UCL of at least twelve consecutive monthly sampling events from each well location within the subject groundwater plume is equal to or less than the GWPC.

1.3.2.2 *Surface-Water Protection Criteria*

Section 22a-133k-3(g)(2)(C) of the RSRs provides the option to demonstrate compliance with the SWPC if the 95% UCL of all sample results representative of the groundwater plume is equal to or less than the SWPC. Note that Section 22a-133k-3(g)(2)(A)(ii) requires a minimum of four sampling events which reflect seasonal variability on a quarterly basis, provided that all sampling events were performed within two years, for determining compliance with applicable criteria.

1.4 **Document Organization**

The remainder of this Document is organized as follows:

- Section 2 provides a discussion on developing a potential 95% UCL data set for a soil release area to demonstrate compliance with the RSRs;
- Section 3 provides a discussion on developing a potential 95% UCL data set for a groundwater plume to demonstrate compliance with the RSRs;
- Section 4 identifies factors to consider when evaluating whether a data set is appropriate for calculating a 95% UCL;
- Section 5 presents an overview of statistical calculation methods, including the recommendation to use United States Environmental Protection Agency's (EPA's) ProUCL software;
- Section 6 provides information on requesting an alternative method for demonstrating compliance with the RSRs related to the use of the 95% UCL provisions;

The text of this document is followed by an example illustrating the effects of the collection of additional samples after the completion of site characterization on the estimation of the 95% UCL. In addition, the following appendices are included in this Document:

- Appendix A presents an explanation of laboratory method detection limits and reporting limits and how they relate to non-detect results;
- Appendix B presents a summary of 95% UCL calculation methods; and
- Appendix C provides EPA ProUCL's (Version 5.0) recommended calculation methods for data sets with varying distribution, sample size, and skewness.

2. DEVELOPING A DATA SET FOR A RELEASE AREA

Many other state and federal guidance documents discuss estimating a 95% UCL over an exposure unit[2] whereas, the RSRs require the estimation of a 95% UCL for a release area. As defined in Section 22a-133k-1(a)(56) of the RSRs, a *Release Area* is "land area at and beneath which polluted soil is located as a result of a release." *Polluted Soil* is defined in Section 22a-133k-1(a)(50) of the RSRs as "soil affected by a release of a substance at a concentration above the analytical detection limit for such substance." For the purposes of this Document, the term "substances" is herein referred to as constituents of concern (COCs).

[2] An exposure unit is defined by USEPA as the geographical area within which a receptor is randomly exposed to contaminated media for a relevant exposure duration.

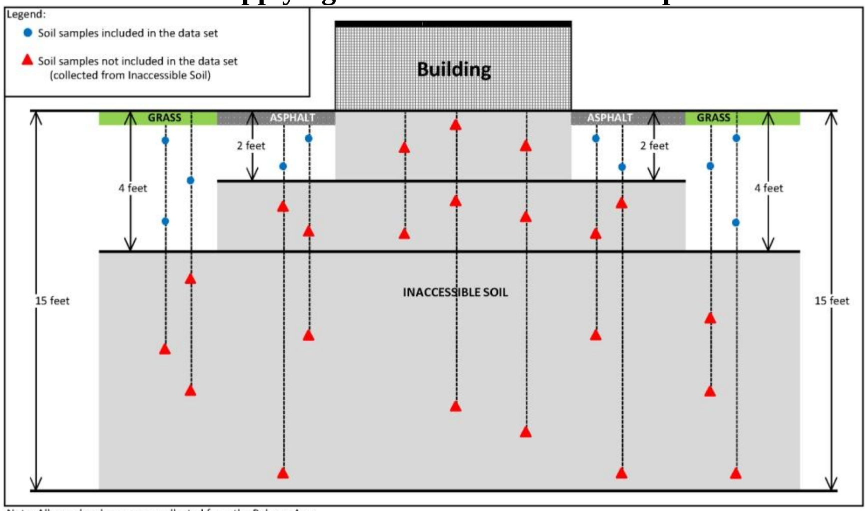
The data set for use in the 95% UCL calculation is obtained, wholly or in part, through the characterization of a Release Area. The SCGD describes the Conceptual Site Model (CSM) approach to investigation that is used to characterize the nature, degree, and extent of contamination associated with a Release Area. If a portion of the Release Area has been excavated, data collected from within the excavated area (either during characterization or remediation) cannot be included in the data set since this portion is no longer part of the Release Area. Confirmation samples collected within the remaining Release Area should be included in the data set for calculating the 95% UCL. The limits of a Release Area are defined by the extent of detectable evidence of COCs in soil associated with the same source. As described in the SCGD, multiple lines of evidence such as visual observations, mobile lab results, and other field screening results, can be used in conjunction with traditional fixed laboratory analytical results to define the extent of a Release Area. If supported by the CSM and the rationale documented by the environmental professional, a suite of chemically-related COCs such as petroleum hydrocarbons, chlorinated volatile organic compounds, or polycyclic aromatic hydrocarbons associated with the same source could also be utilized to delineate the Release Area.

2.1 Data Selection for 95% UCL Calculation for a Release Area

All results for soil sampling points located within the lateral and vertical limits of the Release Area comprise the data population for use in the 95% UCL calculation. However, the data set used must also be representative for the specific purpose of the statistical evaluation. For example, when applying the 95% UCL calculation to evaluate compliance with the PMC, soil samples collected at depths above the applicable seasonal water table within the limits of the Release Area would comprise the data set. In contrast, if a 95% UCL calculation is used to assess compliance with the DEC, the data set would consist of soil samples collected from within the Release Area limits to depths of less than or equal to 15-feet below grade, regardless of the depth of the water table.

As stated in Section 1.3.1, only those soil sample results from locations and depths where the soil criteria apply, taking into consideration any institutional controls implemented, should be used to calculate the 95% UCL for DEC compliance. Figure 2 illustrates a scenario where the DEC no longer applies to much of the release area due to use of the inaccessible soil exception pursuant to the RSRs. Only those samples collected where the DEC still applies are appropriate for use in the 95% UCL data set, which in this scenario are those collected from zero to two feet under the paved surfaces and zero to four feet under the open landscaped areas.

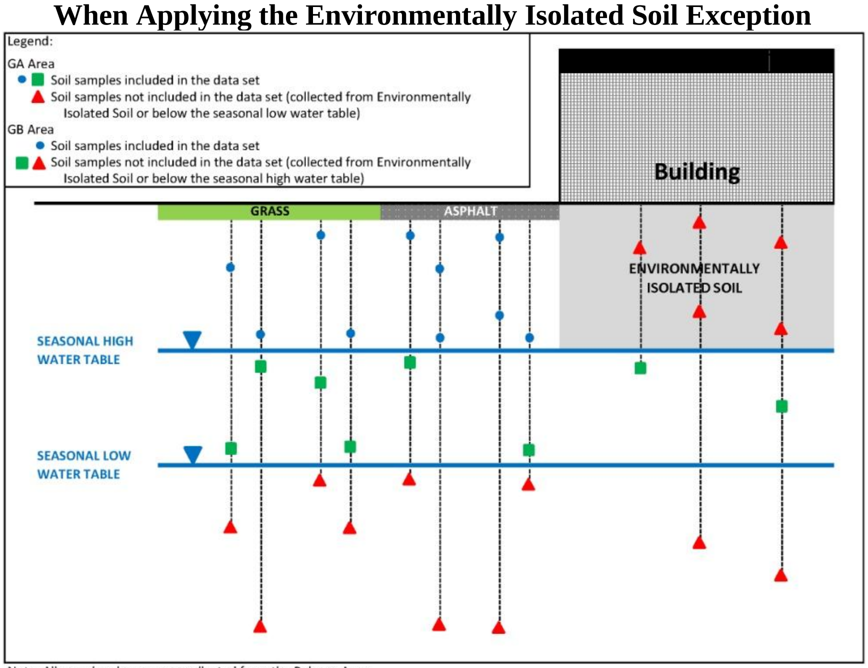
Fig. 2 - Release Area Sample Selection for DEC 95% UCL Calculation When Applying the Inaccessible Soil Exception



Note: All samples shown were collected from the Release Area

Figure 3 illustrates a scenario where the PMC no longer applies to some of the release area due to the use of the environmentally isolated soil exception pursuant to the RSRs. Only those samples collected where the PMC still applies are appropriate for use in the 95% UCL data set, which in this scenario are those collected outside the environmentally isolated soil area and above the seasonal low groundwater table in GA groundwater classification areas or seasonal high groundwater table in GB groundwater classification areas.

Fig. 3 - Release Area Sample Selection for PMC 95% UCL Calculation When Applying the Environmentally Isolated Soil Exception



Note: All samples shown were collected from the Release Area

2.2 Non-Detect Soil Results in a Release Area Data Set

Proper and effective delineation of a Release Area in accordance with the SCGD and a site-specific CSM will result in some non-detect analytical results. Non-detect results from within a Release Area may be a result of the variability and complexity of environmental systems (e.g., heterogeneity, temporal fluctuation, anisotropy, chemical properties, fate and transport, elevated reporting limits, etc.), and can be used when estimating a 95% UCL. However, non-detect results along the lateral and/or vertical extents of a Release Area should not be used when estimating a 95% UCL unless evidence can be provided to substantiate such a decision (e.g., presence of related COCs from the same release area, field observations of impacts, elevated reporting limits, etc.).

2.3 Quality Control Soil Sample Results in a Release Area Data Set

Duplicate (or replicate or split) samples for quality control (QC) purposes are collected to evaluate sample precision. Since a duplicate sample is considered to be identical to the parent sample, both sample results should not be included in the data set for the 95% UCL calculation, as that would impart a bias due to double counting. For the 95% UCL calculation, the higher of the two sample results should be used.

3. DEVELOPING A DATA SET FOR A GROUNDWATER PLUME

The purpose of this section is to guide the environmental professional in developing an appropriate data set for a Groundwater Plume when calculating a 95% UCL to demonstrate compliance with the SWPC. This section does not apply to the calculation of a 95% UCL to demonstrate compliance with the GWPC since the data set used is based on data collected from an individual well, not averaged over the Groundwater Plume, for GWPC compliance. As defined in Section 22a-133k-1(a)(28) of the RSRs, a *Groundwater Plume* is "ground water which has been polluted by a release and in which ground water one or more substances from such release is present at a concentration above the analytical detection limit." As previously noted in Section 2, for the purposes of this Document, these substances are referred to as COCs. In accordance with the RSRs, in certain cases a groundwater plume may be determined as analytical results of one or more COCs at concentrations greater than a documented background concentration. The calculation of a 95% UCL to demonstrate compliance with applicable groundwater criteria can only be applied to a steady-state or diminishing Groundwater Plume. A steady-state (or diminishing) Groundwater Plume is a plume where the extent and degree of the groundwater plume are not increasing over time, except as a result of either natural attenuation or seasonal variations.

3.1 Data Selection for 95% UCL Calculation for a Groundwater Plume

The quantity and quality of data needed to delineate a Groundwater Plume for the purpose of demonstrating compliance with the RSRs is a function of the CSM and the DQOs. The horizontal and vertical limits of a Groundwater Plume are defined by the extent of COCs detected in groundwater above either laboratory reporting limits or a documented background concentration. The data set used in the 95% UCL calculation may not be sufficient if obtained solely from monitoring wells that were installed to determine the nature, degree, extent, and temporal variation of a groundwater contaminant plume. Professional judgment should be used when evaluating if the quantity and quality of data used for the data set are sufficient to fully characterize the groundwater quality three-dimensionally and temporally.

3.2 Non-Detect Results in a Groundwater Plume Data Set

Within a steady-state or diminishing Groundwater Plume, there is a potential for non-detect results to be obtained due to the inherent variability and complexity of environmental systems (e.g., temporal fluctuation, anisotropism of the aquifer, chemical properties, fate and transport, elevated reporting limits, etc.). Non-detect results for the COCs inside the limits of a steady-state or diminishing Groundwater Plume may be valid for the determination of compliance with applicable groundwater criteria. However, non-detect results from monitoring well locations that define the limits of the Groundwater Plume should not be factored into the estimation of a 95% UCL, as these data points are functionally outside the Groundwater Plume.

3.3 Quality Control Results in a Groundwater Plume Data Set

As discussed earlier in Section 2.2, duplicate (or replicate or split) samples for QC purposes are collected to evaluate sample precision. Since a duplicate sample is considered to be identical to the parent sample, both sample results should not be included in the data set for the 95% UCL calculation, as that would impart a bias due to double counting. For the 95% UCL calculation, the higher of the two samples should be used.

4. EVALUATING THE DATA SET

In order to calculate an accurate and defensible 95% UCL, the Release Area or Groundwater Plume data set should be evaluated for the following:

- the representativeness of the nature and extent of the COC distribution throughout the Release Area or Groundwater Plume;
- the size of the Release Area or Groundwater Plume; and
- the statistical DQOs for the Release Area or Groundwater Plume.

If remediation by excavation of a portion of a Release Area has been completed, the environmental professional's evaluation of the data set should also include the confirmation samples collected from within the remaining Release Area. The following sections provide details on the representativeness, size, and statistical DQOs that should be considered when using professional judgment in establishing the data set.

4.1 Distribution of COC Concentrations in the Environment

Determining the distribution of COC concentrations is necessary for the development of a data set that is both representative of a Release Area or Groundwater Plume and appropriate for statistical analysis. For the purpose of calculating a 95% UCL, the data set needs to emulate what is in the Release Area or Groundwater Plume so that the 95% UCL calculated from the resulting distribution appropriately reflects site conditions. The remedial decision made based on the 95% UCL estimate is only as good as the data set that is utilized. The individual samples used for the 95% UCL calculation must be discrete and representative of the statistical population (i.e., soil depths, soil types, release mechanisms, and other characteristics) in the Release Area or Groundwater Plume. These concepts are further discussed in the example provided at the end of this Document.

Outliers are also an important concept when evaluating the distribution of COC concentrations in the environment and an analysis for statistical outliers can be performed as part of the evaluation of a 95% UCL data set. It is important to remember that any statistical outlier identified represents site conditions. If the statistical outlier is removed from the 95% UCL data set, the area with the elevated concentrations will need to be addressed through remediation.

4.2 Appropriate Data Set Size

An appropriate amount of data must be used in order to meet the underlying assumptions of the statistical methods and calculate an accurate and defensible 95% UCL. Although a 95% UCL can be estimated using small data sets, the result will likely not have the strength to provide appropriate compliance or remediation decisions. Based on research of statistical applications under ideal conditions, a minimum of ten (10) samples from each Release Area or Groundwater Plume should be used for the estimation of the 95% UCL. However, the RSRs require additional samples to demonstrate compliance with the PMC (20 samples) and the GWPC (12 samples).

> Obtaining the largest data set possible will provide an estimate of a 95% UCL that is closer to the true population mean. Simulation study results by EPA, ProUCL, and other statistical models suggest that it may take hundreds of samples to obtain optimal results. The Department recognizes that it may not be practical to collect this number of samples; however, it is up to the environmental professional to ensure that a sufficient number of samples are collected to achieve the desired DQOs and obtain a defensible 95% UCL.

4.3 Statistical DQOs

The statistical DQO process is used to determine the technical objectives and appropriate data set for a 95% UCL estimate. The randomness, strength, and skewness of the data set are important factors to consider when evaluating the appropriateness of the data set.

4.3.1 Randomness of Data Set

The calculation methods for estimating 95% UCLs are based on the assumption of random sampling. However as stated in Section 1.2, environmental investigations are typically focused on identifying and delineating areas of contamination. As a result, a certain amount of non-random bias towards the lower concentrations will be introduced into the data set. Stratified random sampling (i.e., random sampling within specified target areas) is one way to avoid excessive bias in the data set. Other sampling programs may be used to minimize bias as long as they are constructed properly (e.g., there is no collection of samples across different release areas). The environmental professional needs to ensure that the data set used for the 95% UCL calculation does not have an unacceptable amount of bias.

4.3.2 Strength of Data Set

The *strength of the data* set is an important concept when evaluating the appropriateness of the data set for use in estimating a 95% UCL. Strength of the data set is related to the size of the data set. Therefore, the data set size should be as large as possible to represent the distribution of concentration levels and to increase the strength of the data set and the reliability of the 95% UCL. Use of professional judgment, which includes evaluation of all appropriate detectable evidence of the COC and multiple lines of evidence, is required to ensure that the data set is of sufficient strength to obtain a meaningful 95% UCL. Additional samples may be necessary to increase the strength of the data set.

4.3.3 Skewness of Data Set

Skewness is a measure of the asymmetry of the distribution of the data set. Application of the appropriate 95% UCL calculation method for the degree of skewness is critical to obtaining accurate results. As the size, strength, and randomness of the data set is increased, the skewness of the data set also tends to decrease. Tables presenting the appropriate 95% UCL calculation methods for the degree of skewness in a specific data set are provided in Appendix C.

> Skewness can be measured in terms of the standard deviation of log-transformed data: the higher the standard deviation, the higher the skewness.

5. STATISTICAL CALCULATION METHODS

Based on a review of readily available calculation methods, the Department recommends the use of ProUCL for calculating a 95% UCL. ProUCL is a free software application available from EPA at:

http://www.epa.gov/osp/hstl/tsc/software.htm

ProUCL makes recommendations (based on data distribution, data set size, skewness, and percentage of non-detect observations) on how to obtain an accurate 95% UCL. In some cases, ProUCL may suggest more than one 95% UCL estimate. In these cases, the environmental professional should evaluate the data set and select the most appropriate 95% UCL.

If an alternate calculation method other than ProUCL is used, it will be necessary for the environmental professional to provide additional documentation regarding the calculation method and how it is applicable given the data set used.

Regardless of whether ProUCL or another calculation method is used, an evaluation of the data distribution and the method for handling of non-detect results are critical components of the calculation methods.

5.1 Data Distributions

The calculation method used should be appropriate for the distribution of the data set. The distribution of the data set is typically determined prior to selecting the calculation method for estimating a 95% UCL. The possible distributions evaluated by ProUCL include normal, lognormal, gamma, or unknown (non-parametric).

ProUCL utilizes various parametric and non-parametric methods for estimating a 95% UCL. Parametric methods incorporate assumptions based on the distribution of the data set (i.e., normal distribution), whereas non-parametric methods are valid for data from populations without a known distribution. Although ProUCL will typically estimate values using all of the available methods, there are instances where no value may be calculated for a particular method.

5.2 Handling of Non-Detect Results in Statistical Calculations

The Department views a non-detect result as an analytical result that is below the laboratory reporting limit (RL) and is not associated with the method detection limit (MDL). It is important for users of environmental laboratory data to have a clear understanding of the difference between an MDL and the RL. The MDL is an index of analytical low-level precision and accuracy, while the RL is an index of the reliability of the value reported. Appendix A provides further detail regarding MDL and RL. The use of non-detect results in developing the data set for a release area or groundwater plume are discussed Sections 2 and 3.

Historically, if the data set contained non-detect results, the substitution method was used to replace non-detect results with a set value, typically one-half the RL. Currently, the best practice is to use statistical methods to handle the non-detect results. ProUCL uses Regression on Order Statistics methods for known distributions and the Kaplan-Meier method for non-parametric data sets. These methods are based on replacing non-detect results with values generated to match the distribution of the rest of the data set. These methods can handle non-detect results with varying RLs. If ProUCL is not used, a similar method that is consistent with current best practices should be used when handling non-detect results.

6. REQUESTS FOR ALTERNATIVE METHOD FOR DEMONSTRATING COMPLIANCE USING THE 95% UCL

The RSRs require a specific number of samples in the data set to calculate the 95% UCL for demonstrating compliance with the PMC and GWPC. For the PMC, Sections 22a-133k-2(d)(3) and (5) of the RSRs allows the regulated community to request an alternative method of demonstrating compliance. Pursuant to Section 22a-133k-2(e)(2)(A), a minimum of twenty (20) samples are required to utilize the 95% UCL to demonstrate compliance with the PMC. If it can be shown that a data set with less than 20 samples is representative of site conditions given the size of the release area and the data set achieves the statistical DQOs, an alternative method of demonstrating compliance can be requested. The environmental professional should submit the request in accordance with Section 22a-133k-1(f) and Section 22a-133k-2(d)(1).

For the GWPC, Section 22a-133k-3(g) does not include a provision for requesting alternative methods for using the 95% UCL to demonstrate compliance.

7. REFERENCES

Connecticut Department of Environmental Protection (2000). Calculating the 95% Upper Confidence Level (UCL), from the DRAFT October 5, 2000 CTDEP Bureau of Water Management Site Characterization Guidance Document.

Florida Department of Environmental Protection (2005). FL UCL – 95% UCL calculation tool developed by FLDEP. Software, presentation, and review of ProUCL version 3 and 4 provided at: http://www.dep.state.fl.us/waste/categories/wc/pages/ProgramTechnicalSupport.htm

Michigan Department of Environmental Quality (2002). Sampling Strategies and Statistics Training Materials for Part 201 Cleanup Criteria. http://www.michigan.gov/documents/deq/deq-erd-stats-s3tm_250015_7.pdf

New Jersey Department of Environmental Protection (2008). New Jersey Department of Environmental Protection Guidance Document, Directions to Determine 95 Percent Upper Confidence Level of the Mean Using USEPA ProUCL Version 4.0 Software Package.
http://www.state.nj.us/dep/srp/guidance/rs/proucl.pdf

Ohio Environmental Protection Agency (2005). Guidance for Computing the 95% UCL of an Environmental Data Set.

Ohio EPA (2009). Closure Plan Review Guidance. http://epa.ohio.gov/portals/32/pdf/2008CPRG.pdf

Pennsylvania Department of Environmental Protection (2002). Pennsylvania Department of Environmental Protection Land Recycling Program Technical Guidance Manual, (Section IV – General Guidance).
http://files.dep.state.pa.us/EnvironmentalCleanupBrownfields/LandRecyclingProgram/LandRecyclingProgramPortalFiles/GuidanceTechTools/section_iv_january_2008_revision.pdf

United States Environmental Protection Agency (1992). Supplemental Guidance to RAGS: Calculating the Concentration Term. OSWER Number 9285.7-081.

US EPA (1997). The Lognormal Distribution in Environmental Applications. EPA/600/S-97/006.
http://www.epa.gov/esd/tsc/images/256cmb04lognorm.pdf

US EPA (2002). Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites. OSWER Number 9285.6-10. http://www.epa.gov/oswer/riskassessment/pdf/ucl.pdf

US EPA (2002). Estimation of the Exposure Point Concentration Term Using a Gamma Distribution. EPA/600/R-02/084. http://www.epa.gov/esd/tsc/images/289cmb02.pdf

US EPA (2006). On the Computation of a 95% Upper Confidence Limit of the Unknown Population Mean Based Upon Data Sets with Below Detection Limit Observations. EPA/600/R-06/022. http://www.epa.gov/esd/tsc/appendix/EPA600R-06-022.pdf

US EPA (2009). Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Unified Guidance. EPA/530/R-09-007. http://www.epa.gov/wastes/hazard/correctiveaction/resources/guidance/sitechar/gwstats/unified-guid.pdf.

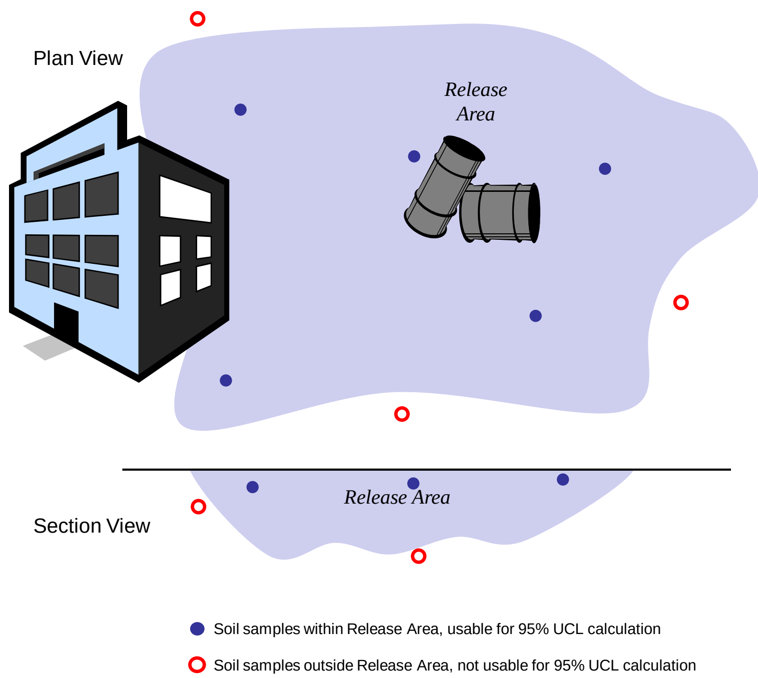
US EPA (2013). ProUCL Statistical Software for Environmental Applications and Documentation. Download software and obtain User Guide, Technical Guide, Fact Sheet, and archived webinars at http://www.epa.gov/esd/tsc/software.htm

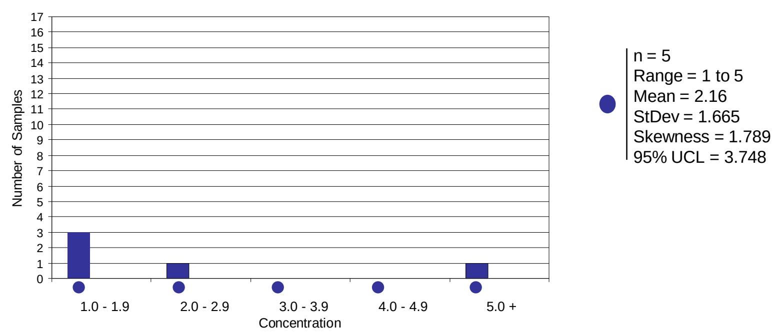
US EPA (2013). Title 40, Code of Federal Regulations: Protection of Environment, Part 136 – Guidelines Establishing Test Procedures for the Analysis of Pollutants, Appendix B – Definition and Procedure for the Determination of the Method Detection Limit.

Example 1. Effect of Collection of Additional Samples after Delineation

The following example illustrates the potential effects of collecting additional soil data after the completion of site characterization on the estimated 95% UCL for demonstrating compliance with the DEC.

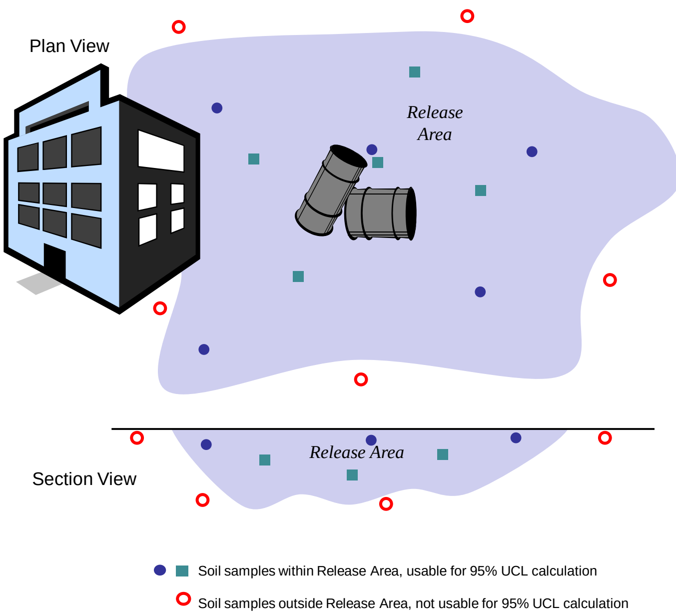
Typically, site characterization data is focused on identifying and delineating Release Areas. However, a data set generated from only characterization may not be representative of a Release Area and likely may not be sufficient for use in a 95% UCL calculation.

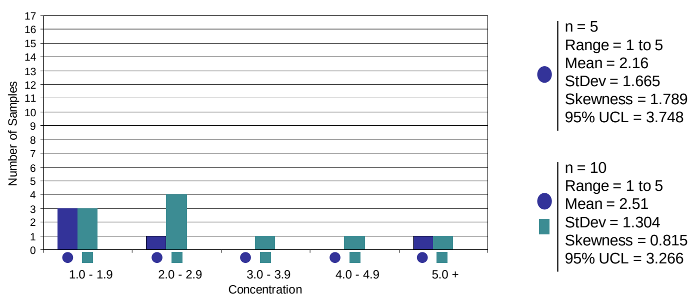




Because of the limited sampling presented above, the initial data set does not exhibit a defined distribution, has a high degree of skewness biased to the lower concentrations (i.e., delineation sample points), and likely will not result in an estimated 95% UCL that is representative of the Release Area.

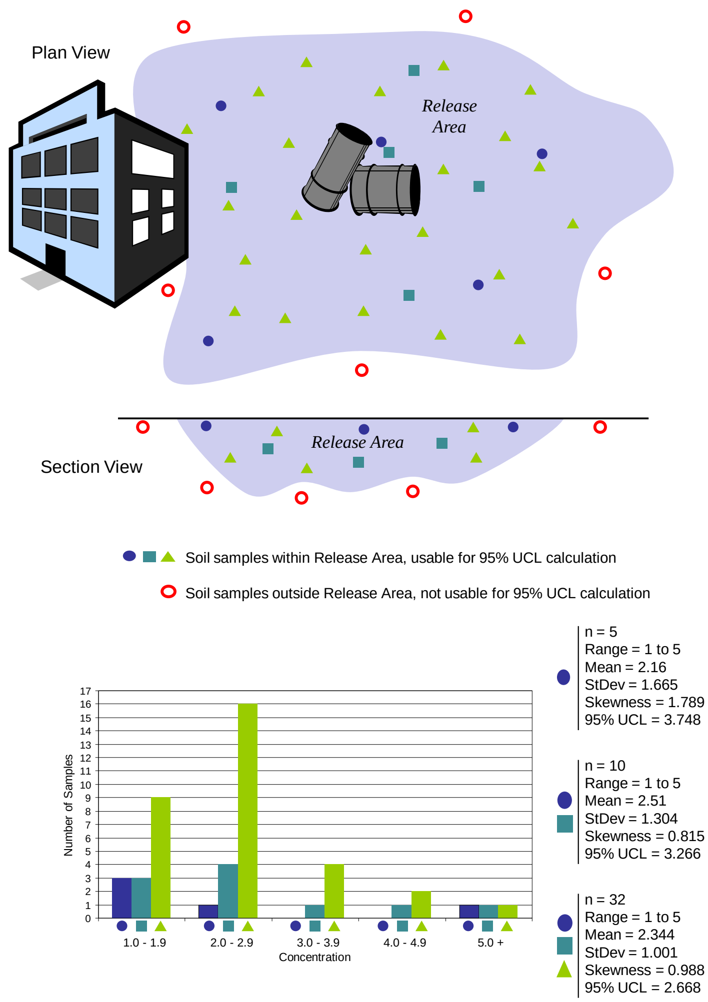
As more samples are collected, as shown below, focused on increasing the randomness of the data set and collecting data between the hot spot and edge of the Release Area, the data set approaches a defined distribution (lognormal).





Therefore, with 10 samples collected from the Release Area the mean has increased due to the reduction in bias to the delineation samples, however, the standard deviation and skewness decreased. The estimated 95% UCL also decreased, even though the mean has increased, due to the increased strength of the data set. Due to these factors, the statistical evaluation has resulted in a more meaningful and representative 95% UCL, which is more appropriate for decision-making purposes.

As shown below, with even more samples collected focused on increasing the randomness, the lognormal distribution becomes even more apparent.



Given that the data set more closely exhibits a lognormal distribution and the strength of the data set has increased, the estimated 95% UCL for the Release Area is substantially lower than when only 5 or 10 samples were collected.

The purpose of this example is to graphically demonstrate that the collection of more data to increase the randomness and decrease the skewness or bias in the data set results in a stronger data set which will lead to a more representative 95% UCL and therefore allow better remedial decisions to be made. The number of samples necessary and the evaluation on whether the generated data set is representative of the release area should be based on the principles discussed in this Document and professional judgment. In the example presented above, the collection of more data results in a lower 95% UCL, but it is possible, especially if the initial data set is strongly skewed with a bias to the lower concentrations, that the collection of more representative data will result in an increase in the 95% UCL. In addition, for a given data set the calculation methods used for known distributions (in this case, lognormal) have more built in assumptions which tend to calculate a lower 95% UCL than when those assumptions are not used (as in non-parametric methods). Therefore, it is very important that the appropriate method is used when calculating the 95% UCL for a given data set to obtain a representative 95% UCL.

Appendix A Laboratory Analytical Limits and How They Affect Non-Detect Results

Detection levels affecting reported non-detect results include: the Method Detection Limit (MDL) and laboratory Reporting Limit (RL). Laboratory calibration processes establish differing equipment detection levels for the MDL and RL.

The MDL is a theoretical limit of detection that is specific for the analyte, sample matrix, instrumentation, method and technician's skill. The MDL is the smallest amount of an analyte, if it is present, that a technician can reliably observe 99% of the time. It is a theoretical value determined by statistically extrapolating the point where a positive instrument response signal can be distinguished from instrument noise. Although not quantifiable, the MDL is derived from the lowest concentration of an analyte that can be reproducibly detected and distinguished from a concentration of zero.

The RL is established based on the MDL. Since the MDL cannot be reliably quantified, the RL is higher than the MDL. How much higher depends on the accepted level of reliability for the reported value. The protocol for determining the level of reliability is outlined in the Code of Federal Regulations, Title 40, Part 136, Appendix B. The reliability of the reported result increases with the ratio of RL/MDL; therefore, the reliability of laboratory data is closely tied to the RL and MDL. The laboratory establishes an RL for a given method based on the technician's ability to perform a particular analysis.

Only results reported above the RL are considered usable. Results below the RL (also referred to as non-detect) are reported as less than the reporting limit. The RL is established by the laboratory to provide quantifiably usable results and are typically 2.5 to 5 times the MDL.

Appendix B Summary of UCL Calculation Methods[3]

Method	Applicability	Advantages	Disadvantages	Reference
For Normal or Lognormal Distributions				
Student's *t*	means normally distributed, samples random	simple, robust if *n* is large	distribution of means must be normal	Gilbert 1987; EPA 1992
Land's H	lognormal data, small variance, large *n*, samples random	good coverage[1]	sensitive to deviations from lognormality, produces very high values for large variance or small *n*	Gilbert 1987; EPA 1992
Chebyshev Inequality (MVUE)	skewness and variance small or moderate, samples random	often smaller than Land	may need to resort to higher confidence levels for adequate coverage	Singh et al. 1997
Wong	gamma distribution	second order accuracy[2]	requires numerical solution of an improper integral	Schulz and Griffin 1999; Wong 1993
Nonparametric/Distribution-free Methods				
Central Limit Theorem - Adjusted	large *n*, samples random	simple, robust	sample size may not be sufficient	Gilbert 1987; Singh et al. 1997
Bootstrap t Resampling	sampling is random and representative	useful when distribution cannot be identified	inadequate coverage for some distributions; computationally intensive	Singh et al. 1997; Efron 1982
Hall's Bootstrap Procedure	sampling is random and representative	useful when distribution cannot be identified; takes bias and skewness into account	inadequate coverage for some distributions; computationally intensive	Hall 1988; Hall 1992; Manly 1997; Schultz and Griffin 1999
Jackknife Procedure	sampling is random and representative	useful when distribution cannot be identified	inadequate coverage for some distributions; computationally intensive	Singh et al. 1997
Chebyshev Inequality	skewness and variance small or moderate, samples random	useful when distribution cannot be identified	inappropriate for small sample sizes when skewness or variance is large	Singh et al. 1997; EPA 2000c

[1] Coverage refers to whether a UCL method performs in accordance with its definition.
[2] As opposed to maximum likelihood estimation, which offers first order accuracy.

[3] From: US EPA (2002). *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites,* OSWER Number 9285.6-10.

Appendix C ProUCL Version 5.0 Calculation Method Decision Summary Tables[1]

DECISION SUMMARY TABLES

Table C-1. Skewness as a Function of σ (or its *MLE*, $s_y = \hat{\sigma}$), *sd* of log(*X*)

Standard Deviation of Logged Data	Skewness
$\sigma < 0.5$	Symmetric to mild skewness
$0.5 \leq \sigma < 1.0$	Mild skewness to moderate skewness
$1.0 \leq \sigma < 1.5$	Moderate skewness to high skewness
$1.5 \leq \sigma < 2.0$	High skewness
$2.0 \leq \sigma < 3.0$	Very high skewness (moderate probability of outliers and/or multiple populations)
$\sigma \geq 3.0$	Extremely high skewness (high probability of outliers and/or multiple populations)

Table C-2. Summary Table for the Computation of a 95% UCL of the Unknown Mean, μ_1, of a Gamma Distribution

$\hat{k}$	Sample Size, n	Recommendation
$\hat{k} > 1.0$	$n >= 50$	Approximate gamma 95% UCL
$\hat{k} > 1.0$	$n < 50$	Adjusted gamma 95% UCL
$\hat{k} \leq 1.0$	$n < 15$	95% UCL based upon bootstrap-t or Hall's bootstrap method*
$\hat{k} \leq 1.0$	$n \geq 15$	Adjusted gamma 95% UCL if available, otherwise use approximate gamma 95% UCL

***** In case the bootstrap-t or Hall's bootstrap methods yield erratic, inflated, and unstable UCL values, the UCL of the mean should be computed using an adjusted gamma UCL.

Table C-3. Summary Table for the Computation of a 95% UCL of the Unknown Mean, μ_1, of a Lognormal Population

$\hat{\sigma}$	***Sample Size, n***	***Recommendation***
$\hat{\sigma} < 0.5$	For all *n*	Student's t, modified-t, or H-UCL
$0.5 \le \hat{\sigma} < 1.0$	For all *n*	H-UCL
$1.0 \le \hat{\sigma} < 1.5$	$n < 25$	95% Chebyshev *(Mean, Sd)* UCL
	$n \ge 25$	H-UCL
$1.5 \le \hat{\sigma} < 2.0$	$n < 20$	97.5% or 99% Chebyshev *(Mean, Sd)* UCL
	$20 \le n < 50$	95% Chebyshev *(Mean, Sd)* UCL
	$n \ge 50$	H-UCL
$2.0 \le \hat{\sigma} < 2.5$	$n < 20$	99% Chebyshev *(Mean, Sd)* UCL
	$20 \le n < 50$	97.5% Chebyshev *(Mean, Sd)* UCL
	$50 \le n < 70$	95% Chebyshev *(Mean, Sd)* UCL
	$n \ge 70$	H-UCL
$2.5 \le \hat{\sigma} < 3.0$	$n < 30$	99% Chebyshev (*Mean, Sd*)
	$30 \le n < 70$	97.5% Chebyshev *(Mean, Sd)* UCL
	$70 \le n < 100$	95% Chebyshev *(Mean, Sd)* UCL
	$n \ge 100$	H-UCL
$3.0 \le \hat{\sigma} \le 3.5$**	$n < 15$	Bootstrap-t or Hall's bootstrap method*
	$15 \le n < 50$	99% Chebyshev(*Mean, Sd*)
	$50 \le n < 100$	97.5% Chebyshev *(Mean, Sd)* UCL
	$100 \le n < 150$	95% Chebyshev *(Mean, Sd)* UCL
	$n \ge 150$	H-UCL
$\hat{\sigma} > 3.5$**	For all *n*	Use nonparametric methods*

***** In the case that Hall's bootstrap or bootstrap-t methods yield an erratic unrealistically large UCL value, UCL of the mean may be computed based upon the Chebyshev inequality: Chebyshev *(Mean, Sd)* UCL.

****** For highly skewed data sets with $\hat{\sigma}$ exceeding 3.0, 3.5, it is suggested the user pre-processes the data. It is very likely that the data consist of outliers and/or come from multiple populations. The population partitioning methods may be used to identify mixture populations present in the data set.

Table C-4. Summary Table for the Computation of a 95% *UCL* of the Unknown Mean, μ_1, Based Upon a Skewed Data Set (with all Positive Values) without a Discernible Distribution, Where $\hat{\sigma}$ is the *sd* of Log-transformed Data

$\hat{\sigma}$	*Sample Size, n*	*Recommendation*
$\hat{\sigma} < 0.5$	For all *n*	Student's t, modified-t, or H-UCL Adjusted CLT UCL, BCA Bootstrap UCL
$0.5 \le \hat{\sigma} < 1.0$	For all *n*	95% Chebyshev *(Mean, Sd)* UCL
$1.0 \le \hat{\sigma} < 1.5$	For all *n*	95% Chebyshev *(Mean, Sd)* UCL
$1.5 \le \hat{\sigma} < 2.0$	$n < 20$	97.5% Chebyshev *(Mean, Sd)* UCL
	$20 \le n$	95% Chebyshev *(Mean, Sd)* UCL
$2.0 \le \hat{\sigma} < 2.5$	$n < 15$	Hall's bootstrap method
	$15 \le n < 20$	99% Chebyshev *(Mean, Sd)* UCL
	$20 \le n < 50$	97.5% Chebyshev *(Mean, Sd)* UCL
	$50 \le n$	95% Chebyshev *(Mean, Sd)* UCL
$2.5 \le \hat{\sigma} < 3.0$	$n < 15$	Hall's bootstrap method
	$15 \le n < 30$	99% Chebyshev (*Mean, Sd*)
	$30 \le n < 70$	97.5% Chebyshev *(Mean, Sd)* UCL
	$70 \le n$	95% Chebyshev *(Mean, Sd)* UCL
$3.0 \le \hat{\sigma} \le 3.5$**	$n < 15$	Hall's bootstrap method*
	$15 \le n < 50$	99% Chebyshev(*Mean, Sd*) UCL
	$50 \le n < 100$	97.5% Chebyshev *(Mean, Sd)* UCL
	$100 \le n$	95% Chebyshev *(Mean, Sd)* UCL
$\hat{\sigma} > 3.5$**	For all *n*	99% Chebyshev (*Mean, Sd*) UCL

***** If Hall's bootstrap method yields an erratic and unstable UCL value (e.g., happens when outliers are present), a UCL of the population mean may be computed based upon the 99% Chebyshev *(Mean, Sd)* method.

****** For highly skewed data sets with $\hat{\sigma}$ exceeding 3.0 to 3.5, it is suggested that the user pre-processes the data. Data sets with such high skewness are complex and it is very likely that the data consist of outliers and/or come from multiple populations. The population partitioning methods may be used to identify mixture populations present in the data set.

Notes: Suggestions regarding the selection of a 95% UCL are provided to help the user to select the most appropriate 95% UCL. These recommendations are based upon the results of the simulation studies summarized in Singh, Singh, and Iaci (2002) and Singh and Singh (2003). For additional insight, the user may want to consult a statistician.

---

[1] Tables copied directly from US EPA (2013), *ProUCL Version 5.0.00 Technical Guide*, Appendix A.