EPA/600/R-06/022
March 2006
www.epa.gov

On the Computation of a 95% Upper Confidence Limit of the Unknown Population Mean Based Upon Data Sets with Below Detection Limit Observations

Contract No. 68-W-04-005
Task Order No. 09

Prepared for

Chris Sibert

U.S. Environmental Protection Agency
Office of Research and Development
National Exposure Research Laboratory
Environmental Sciences Division
944 E. Harmon Ave.
Las Vegas, NV 89119

Prepared by

Anita Singh, Ph.D.
Robert Maichle and Sanghee E. Lee

Lockheed Martin Environmental Services
1050 E. Flamingo Road, E120
Las Vegas, NV 89119

U.S. Environmental Protection Agency
Office of Research and Development
Washington, DC 20460

211cmb06

Notice

Executive Summary

Environmental scientists often encounter trace level concentrations of contaminants of potential concern (COPC) when evaluating sample analytical results. Those low level analytical results cannot be measured accurately, and therefore are typically reported as less than a certain detection limit (DL) values. Type I, left-censored data arise when certain low values lying below the DL are ignored or unknown as they cannot be measured accurately. However, practitioners need to obtain reliable estimates of the population mean, μ_1, the population standard deviation, σ_1, and various upper limits, including the upper confidence limit (UCL) of the population mass or mean, the upper prediction limit (UPL), and the upper tolerance limit (UTL).

Several methods to estimate the population mean, and the standard deviation based upon left-censored data sets exist in the environmental literature. Some recommendations (Helsel (2005), page 78, USEPA (2000)) have also been made on how to compute those summary statistics based upon data sets with below detection limit observations. However, no specific guidance with justification is available in the statistical and environmental literature on how to compute an appropriate 95% UCL (UCL95) of the population mean or mass and other limits based upon left-censored data sets of varying degree of skewness. Most of the available estimation and UCL computation methods for left-censored data proposed and recommended in the literature have been discussed and evaluated in this report. The UCL95s are used in several environmental applications, including the estimation of the exposure point concentration (EPC) terms needed to assess risk due to average exposure by individuals over an area during a certain period of time. This report emphasizes the development of defensible statistical procedures to accurately compute a UCL95 of the population mass based upon left-censored data sets with varying censoring intensities. Distributions of varying degree of skewness, including mild, moderate, and high skewness, have been considered. Some data sets from Superfund sites have also been utilized to elaborate on the issues of distortion of estimates and of upper limits by: 1) the presence of a few outliers, and 2) the use of a lognormal distribution to accommodate those outlying observations.

The robustness of an estimation and UCL95 computation method needs to be demonstrated for distributions of varying degrees of skewness. Specifically, it should be noted that an estimation method (e.g., jackknife or percentile bootstrap methods on maximum likelihood estimates (MLE) or Kaplan-Meier (KM) estimates, robust regression on order statistic (ROS) of log-transformed data, or a robust MLE method) that yields reasonably good 95% UCLs (providing adequate coverage for the population mean) for symmetric or mildly skewed distributions may not perform well on a data set obtained from moderately or highly skewed distributions. It is observed that a UCL95 obtained using some of the methods (e.g., MLE, ROS on log-transformed data, robust MLE) listed above actually provides coverage much lower than the specified coverage (95%) for the population mass.

By definition, the UCL95 of a population mean, μ_1, assumes that one is dealing with a single statistical population with mean, μ_1. Throughout this report, it is implicitly assumed that the user is dealing with a data set collected from a single population, has preprocessed the data set; and has identified all of the potential outliers (if any) and multiple populations. In order to obtain meaningful and practical results, the procedures described in this report should be used on data sets that represent “single” (e.g., a background area, or an exposure unit), and “not mixture” populations. The intent of these statements is to familiarize the user with the underlying assumptions required by the various statistical methods, including the estimation methods based upon left-censored data sets. It should, however be noted that the mathematical UCL95 computation formulae and methods as described in this report can be used on any left-censored data set with or without the outliers. The user should keep in mind that the UCL95 based upon data sets with potential outliers or mixture populations may not be reliable and representative of the dominant population (e.g., an exposure area) under investigation.

The estimation methods as described in this report are applicable to data sets coming from a “single” statistical population such as a single contaminated or remediated area of the site, an unimpacted clean background, or reference population. The sampled data should represent a random sample from the area such as an exposure area (EA), a remediation unit (RU), or some other site area under study. This means that the data should be representative of the entire population (e.g., EA, RU) under study. A few outliers (e.g., representing contaminated locations, hot spots) in a full uncensored data set or in a left-censored data set may distort all classical statistics, including EM estimates, MLEs, restricted MLEs, regression estimates both in raw as well as log scale, and also the associated upper limits such as UCLs, UPLs, and UTLs.

The main objectives of this study are: 1) to evaluate and compare the performances of the various parametric and nonparametric UCL95 computation methods for left-censored data sets and 2) to make recommendations accordingly. Monte Carlo simulation experiments have been conducted to study and compare the performances of the various UCL95 computation methods based upon left-censored data sets covering a wide range of skewed distributions. It is anticipated that this document will represent a comprehensive tutorial-type report giving the details of the various UCL95 computation methods with recommendations needed to compute a meaningful and reliable estimate of the population mass in several environmental applications. The methods considered for the estimation of population mean and the standard deviation are:

- Maximum likelihood method (CMLE) (Cohen (1950, 1959, and 1991)),
- Bias-corrected MLE method (UMLE),
- Restricted MLE (RMLE) method (Perrson and Rootzen (1977)),
- Expectation Maximization (EM) method (Gleit (1985)),
- Delta (delta) method (USEPA (1991),
- Regression on order statistics (ROS) on raw data (Newman, Dixon, and Pinder (1990) and UNCENSOR 5.1 (2003)),
- Regression on order statistics on log-transformed data (Helsel (1990), Helsel (2005), and RPcalc 2.0 (2005)),
- Kaplan-Meier (KM) method (Kaplan-Meier (1958)),
- DL/2 substitution (DL/2) method, and
- Winsorization method (Gilbert (1987)).

Using the estimated mean and standard deviation, and the extrapolated NDs obtained using one of the estimation methods listed above, the 95% UCLs of the mean can be computed using the following methods:

- Tiku’s UCL method (Tiku (1967 and 1971)),
- Schneider’s approximate UCL method (Schneider (1986)),
- Ad hoc UCL methods using Student’s t-statistic as mentioned in UNCENSOR 5.1; (Millard (2002), Helsel (2005), and USEPA-UGD (2004)),
- Ad hoc UCL methods based upon Land’s H-statistic,

- Gamma UCL (Singh, Singh, and Iaci (2002)),
- Nonparametric Chebyshev inequality (Singh, Singh, and Engelhardt (1997)), and
- Bootstrap (percentile, standard bootstrap, bootstrap t, and bias-corrected accelerated (BCA)) methods (Efron and Tibshirani (1993) and ProUCL 3.0 User Guide (2004)).

All of the UCL95 computation methods listed above have been included in the simulation experiments. Normal, lognormal, and gamma distributions covering a wide range of skewed distributions were used in the simulation experiments to generate Type 1 left-censored data sets of various sizes including: 10, 15, 20, 25, 30, 35, 40, 50, 75, and 100. Left-censored samples of varying degree of censoring intensities, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, and 70% have been considered. The performances of the various UCL95 methods listed above have been evaluated by comparing their coverage probabilities.

In order to evaluate and compare the various UCL95 computation methods listed above, some commercial (e.g., MINITAB, SAS) and freely available software packages (UNCENSOR 5.1 (1993) and RPcalc 2.0 (2005)) were used. The main purpose of this evaluation was to compare and verify our results obtained using a SimCensor program (in ProUCL 4.0) with the results obtained using the software packages mentioned above. During this evaluation process, several numerical errors have been identified in the UNCENSOR 5.1 program. It is also noted that there seems to be some confusion in the environmental literature about the use of a lognormal distribution and the interpretation and derivation of the conclusions based upon the results and statistics computed in the transformed scale or in the original scale after using some back-transformation formula. In order to elaborate on these confusing issues, brief evaluations of UNCENSOR 5.1 and RPcalc 2.0 software packages have also been conducted; those comments are provided in Section 5.8.

Based upon the extensive Monte Carlo simulation experiments, recommendations have been made on how to compute appropriate 95% UCLs based upon left-censored symmetric and moderately skewed to highly skewed data sets of varying degrees of censoring intensity. The recommendations have been made as a function of the sample size, censoring intensity, and skewness measured in terms of the standard deviation of log-transformed data. The summary of the simulation results and findings are described in Section 8, and the recommendations are summarized in Section 9. The recommended UCL95 computation methods for the various parametric and nonparametric distributions have also been summarized in Table 9-1 of Section 9. It is observed that there is not an existing single UCL95 computation method that will work for data sets of all sizes and of varying skewness. Since all of these recommended 95% UCL values have been incorporated in the revised ProUCL 4.0 software package (still in progress), the users do not have to keep track of the various recommendations that have been made as a function of the sample size, censoring intensity, and skewness. In addition to the recommended UCL95 methods, some other UCL95 computation methods based upon MLE and ROS methods have also been included in ProUCL 4.0. After completion, the updated ProUCL 4.0 software package will be available at the NERL EPA Tech Support Center Web site given at: http://www.epa.gov/nerlesd1/tsc/tsc.htm. It should be noted that ProUCL 4.0 will have most of the statistical methods as described and used in EPA Guidance Documents (2002a, 2002b).

Table of Contents

Supplement

Acronyms and Abbreviations

A-D test	Anderson-Darling test
AOC	area of concern
BC	Box-Cox type transformation
BCA	bias-corrected accelerated method
BTV	background threshold value
cdf	cumulative distribution function
CI, C.I.	confidence interval
CMLE	Cohen's maximum likelihood estimate
Cohen's MLE (Tiku)	UCL based upon Cohen's maximum likelihood estimates using Tiku's method
Cohen's MLE (t)	UCL based upon Cohen's maximum likelihood estimates using Student's t-distribution cutoff value
COPC	contaminants of potential concern
CV	coefficient of variation
DDT	dichlorodiphenyltrichloroethane
df, d.f.	degrees of freedom
DL, L	detection limit
DL/2 (t)	UCL based upon DL/2 method using Student's t-distribution cutoff value
DOE	Department of Energy
EA	exposure area
EDF	Empirical Distribution Function
EM	Expectation Maximization method
EPC	exposure point concentration
FP-ROS (Land)	UCL based upon fully parametric ROS method using Land's H-statistic
G	gamma distribution
Gamma ROS (BCA)	UCL based upon Gamma ROS method using the bias-corrected accelerated percentile bootstrap method.
Gamma ROS (Appx.)	UCL based upon Gamma ROS method using the gamma approximate-UCL method

H-ROS (BCA)	UCL based upon Helsel's robust method using the bias-corrected accelerated percentile bootstrap method
H-ROS (%)	UCL based upon Helsel's robust method using the percentile bootstrap method
H-ROS (jackknife)	UCL based upon Helsel's robust method using the jackknife method
KM (z)	UCL based upon Kaplan-Meier estimates using standard normal distribution cutoff value
KM (t)	UCL based upon Kaplan-Meier estimates using the Student's t-distribution cutoff value
KM (%)	UCL based upon Kaplan-Meier estimates using the percentile bootstrap method
KM (BCA)	UCL based upon Kaplan-Meier estimates using the bias-corrected accelerated percentile bootstrap method
KM (Chebyshev)	UCL based upon Kaplan-Meier estimates using the Chebyshev theorem
KM (jackknife)	UCL based upon Kaplan-Meier estimates using the jackknife method
K-S test	Kolmogorov-Smirnov test
LN	lognormal distribution
ML	maximum likelihood
MLE	maximum likelihood estimate
N	normal distribution
ND	nondetect
OLS	ordinary least squares
pdf	probability density function
PLE	product limit estimate
PROP	Proposed Influence Function
QA	Quality Assurance
QC	Quality Control
Q-Q plot	quantile-quantile plot
RCRA	Resource Conservation and Recovery Act
RMLE	restricted MLE method
ROS	regression on order statistics
RU	remediation unit
SA	study area
sd	standard deviation
SE	standard error of the mean

SND	standard normal distribution
UCB	upper confidence bound
UCL	upper confidence limit
UCL95	95% upper confidence limit
UGD	Unified Guidance Document
UMLE	bias-corrected MLE method
UMLE (Tiku)	UCL based upon unbiased maximum likelihood estimates using Tiku’s method
UPL	upper prediction limit
USEPA	United States Environmental Protection Agency
UTL	upper tolerance limit

Acknowledgements

The authors wish to thank Ms. Jennifer Hubbard of USEPA for requesting the development and evaluation of appropriate UCL95 computations methods for data sets with below detection limit observations. Her request and support led to the initiation of the work as summarized in this report. The authors also want to thank Mr. Gareth Pearson, the former director of the Technical Support Center, Environmental Sciences Division-NERL, USEPA, Las Vegas, NV, for his support during the early stages of the project. Special thanks go to Professor Ashok Singh of UNLV for numerous technical discussions and valuable suggestions, and to his students: Mr. Pant, Mr. A. Khago, Ms. V. Hennesey, and Miss P. Saravanan for independently evaluating some of the UCL computation methods as considered in this report.

We greatly appreciate the technical review of the report conducted by Dr. Dennis Helsel, Dr. Nadine Adkins, and Mr. John Nocerino. The detailed comments, and several useful and valuable suggestions made by Dr. Helsel and by Mr. Nocerino, made the report more comprehensive and complete.

Finally, we are grateful to Mr. Chris Sibert and Dr. Brian Schumacher, Interim Directors of the Technical Support Center, Environmental Sciences Division-NERL, USEPA, Las Vegas, NV, for their continuing support for this project.

Section 1

Introduction

Processing the analytical results of environmental samples containing potentially hazardous chemicals is often complicated by the fact that some of those pollutants are present at low and trace levels. These "trace level" or "low level" contaminants cannot be measured reliably and are therefore reported as results lying numerically below a certain detection limit, DL (also denoted by L). The resulting data set thus obtained with below detection limit observations represents a Type 1 left-censored data set. However, since the presence of some of those contaminants (e.g., dioxin) in the various environmental media can pose a threat to human health and the environment even at trace level concentrations, these nondetects (NDs) should not be ignored or deleted from subsequent statistical analyses.

For exposure assessment and site characterization purposes, such as to determine mean contamination levels at various locations or areas of a contaminated site, it is desirable to obtain reliable estimates of population mean, standard deviation, and associated upper limits (e.g., upper confidence limits (UCLs), upper prediction limits (UPLs), and upper tolerance limits (UTLs)) using the left-censored data sets. Improperly computed estimates of these parameters and quantities can result in inaccurate estimates of cleanup standards, which in turn can lead to incorrect remediation decisions. In this report, several defensible methods have been developed and evaluated that can be used to compute upper limits (UCLs, UPLs, and UTLs) based upon Type 1 left-censored data sets.

Nondetects, or below detection limit observations, are inevitable in most environmental data sets. A myriad of parametric as well as nonparametric estimation methods exists in the statistical literature (e.g., see Cohen (1991), Gibbons and Coleman (2001), Schneider (1986), Gilliom and Helsel (1986), Singh and Nocerino (2002), Helsel (2005)) to estimate the population mean, μ_1, and the population standard deviation, σ_1, based upon data sets with the below detection limit observations. However, appropriate methods with specific guidance are lacking on how to accurately compute the various upper limits often used in many environmental applications. For example, in exposure and risk assessment applications, one needs to compute a 95% upper confidence limit (UCL95) of the mean based upon data sets with below detection limit observations. Background evaluation and comparison studies often require the computation of UCLs, UPLs, and UTLs based upon left-censored data sets. The main objective of this study is to research and develop the most appropriate UCL computation method(s) for the left-censored data sets. It is hoped that those methods would also be useful to compute defensible UPLs and other relevant background statistics.

Recent environmental literature (e.g., Millard (2002) and USEPA-UGD (2004)) cites the use of ad hoc "rule-of-thumb" type methods based upon the Student's t-statistic or Land's H-statistic to compute the 95% UCLs, 95% UPLs, and 95% UTLs. For example, it is noted that the Student's t-statistic (on Cohen's maximum likelihood estimates) is used to compute an upper prediction limit (pages 10-26, USEPA-UGD (2004)). However, it is noted that the distribution of the t-type statistic used to construct a UPL (pages 10-26, USEPA-UGD (2004)) based upon the maximum likelihood estimate (MLE) of the mean and the standard deviation based upon left-censored data set is not known. The MLEs of the population mean and the standard deviation based upon left-censored data sets are very different from the traditional mean and the standard deviation (*sd*) used in the definition of a typical Student's t-statistic. This "rule-of-thumb" method to compute UCLs, UPLs and UTLs is difficult to defend for moderately skewed to highly skewed data sets with standard deviation of the log-transformed data exceeding 0.75-1.0. This has been demonstrated via simulation experiments and results as summarized in Section 7 and Appendix C of the report. For skewed distributions, such as a lognormal distribution, the coefficient of variation (CV) and skewness are functions of the *sd* of log-transformed data; therefore, in this report, skewness is defined and measured in terms of the standard deviation, σ, of the log-transformed data as considered and used in the ProUCL 3.0 User Guide (2004).

Helsel (2005) first proposed the use of the percentile bootstrap method on the Kaplan-Meier (KM) method (1958) to compute a 95% UCL of the mean based upon left-censored data sets. The performances of the various ad hoc, parametric, and nonparametric UCL95 computation methods for left-censored data sets require detailed investigation before recommending their use for the various environmental applications. One of the objectives of this report is to determine which of the UCL95 computation methods (e.g., KM method, MLE methods, Chebyshev inequality, jackknife, and bootstrap methods) will provide approximately 95% coverage (at least roughly) for the population mean, μ_1, especially if the data sets are moderately skewed to highly skewed.

Simulation studies were conducted to evaluate the performances of the various UCL95 computation methods. Since several of the estimation methods (e.g., MLE, EM) considered can handle only a single detection (denoted by DL or L) limit case, only the single detection limit case has been evaluated in the simulation experiments as summarized in Sections 7 and 8. If multiple detection limits were present in a data set, then all detection limits are replaced by the maximum detection limit resulting in the single detection limit case. The KM method and regression on order statistics (ROS) method (Helsel, 2005) as incorporated in ProUCL 4.0 (currently under development) would be able to handle data sets with multiple detection limits.

In Section 8 of this report, the performances of the various UCL95 methods have been compared in terms of percent of coverages provided by the respective 95% UCLs to estimate the unknown population mean or mass. The numerical simulation results are summarized in Appendix C, whereas the graphical displays of coverage probabilities and average UCL95 values are given in Appendices A and B, respectively. A simulation program (SimCensor) was developed for this report and was used to compute the various 95% UCLs of the population mean based upon left-censored data sets from normal, lognormal, and gamma distributions. Based upon the results and findings of the simulation study as summarized in Section 9, several UCL95 computation methods will be available in the forthcoming ProUCL 4.0 software, which is expected to be available in early 2007.

The methods considered to estimate population mean and the standard deviation are:

- Maximum likelihood method (CMLE) (Cohen (1950, 1959, and 1991)),
- Bias-corrected MLE method (UMLE),
- Restricted MLE (RMLE) method (Perrson and Rootzen (1977)),
- Expectation Maximization (EM) method (Gleit (1985),
- Delta (delta) method (USEPA (1991)),
- Regression on order statistics (ROS) on raw data (Newman, Dixon, and Pinder (1990) and UNCENSOR 5.1 (2003)),
- Regression on order statistics on log-transformed data (Helsel (1990), Helsel (2005), and RPcalc 2.0 (2005)),
- Kaplan-Meier (KM) method (Kaplan-Meier (1958)),
- DL/2 substitution (DL/2) method, and
- Winsorization method (Gilbert (1987)).

Using the mean and the standard deviation, and the extrapolated NDs obtained using one of the estimation methods listed above, the 95% UCLs of the mean can be computed using the following methods.

- Tiku's UCL method (Tiku (1967 and 1971)),
- Schneider's approximate UCL method (Schneider (1986)),
- Ad hoc UCL methods using Student's t-statistic as mentioned in UNCENSOR 5.1 (Millard (2002), Helsel (2005), and USEPA-UGD (2004)),
- Ad hoc UCL methods based upon Land's H-statistic,
- Gamma UCL (Singh, Singh, and Iaci (2002)),
- Nonparametric Chebyshev inequality (Singh, Singh, and Engelhardt (1997)), and
- Bootstrap (percentile, standard bootstrap, bootstrap t, and bias-corrected accelerated (BCA)) methods (Efron and Tibshirani (1993) and ProUCL 3.0 User Guide (2004)).

All of the UCL95 computation methods listed above have been included in the simulation experiments. Normal, lognormal, and gamma distributions covering a wide range of skewed distributions were used in the simulation experiments to generate Type 1 left-censored data sets of various sizes including: 10, 15, 20, 25, 30, 35, 40, 50, 75, and 100. Left-censored samples of varying degree of censoring intensities, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, and 70% have been considered. The performances of the various UCL95 methods listed above have been evaluated by comparing their coverage probabilities.

1.1 Evaluation of the Existing Software Packages for Left-Censored Data Sets

While evaluating and comparing the various UCL95 computation methods listed above, some well-documented and freely available software packages (UNCENSOR 5.1 (1993) and RPcalc 2.0 (2005)) were also evaluated. The main purpose of this evaluation was to compare and verify our results obtained using the SimCensor program with the results obtained using the two software packages mentioned above. During this evaluation process, several numerical errors have been identified in the UNCENSOR 5.1 program. The detailed evaluation and comparisons are summarized in Sections 5 and 6. Since UNCENSOR 5.1 software has been referenced in the literature, including Manly (2001), Shumway, Azari, and Kayhanian (2002), USEPA (2000), and USEPA (1993 – SW-846, currently under revision), it is desirable that those errors be corrected.

In order to compare the estimates obtained using the ROS methods, the RPcalc 2.0 program (2005) was used on data sets considered in Sections 4 and 6. The ROS estimates obtained using the program developed for this report (SimCensor) and RPcalc 2.0 have been summarized in Table 5-1 of Section 5.8. For the robust ROS method, estimates of the mean and *sd* as produced by SimCensor are in close agreement with the estimates obtained using RPcalc 2.0 (2005). The minor differences in the estimates occur due to the fact that SimCensor (described in ProUCL 3.0 User Guide (2004)) and RPcalc 2.0 (described in Helsel (2005)) calculate the normal quantiles using slightly different methods. RPcalc 2.0 calculates the estimates based upon fully parametric ROS method on log-transformed data. The differences in the estimates obtained using the two methods (RPcalc 2.0 and equation (3-22)) can be significant. At present, it is not known which of these two back-transformation methods (from log scale to original scale) performs better. These issues are discussed in detail in Section 5.8.

1.2 Assumptions Needed for the Methods Discussed and Developed in This Report

The estimation methods as described in this report are applicable to data sets coming from a "single" statistical population such as a single contaminated or remediated area of the site, an unimpacted clean background or reference population. The sampled data set should represent a random sample from the area such an as exposure area (EA), a remediation unit (RU), or some other site area under study. This means that the data set should be representative of the entire population (e.g., EA, RU) of interest under study. A few outlying observations (e.g., representing contaminated locations, hot spots) in a full uncensored data set or in a left-censored data set may distort all classical statistics, including the EM estimates, MLEs, restricted MLEs, regression (intercept and slope) estimates both in raw as well as log scale, and also the associated upper limits such as UCLs, UPLs, and UTLs.

In practice, it is the presence of a few outlying observations (extreme high analytical values) or the presence of multiple populations that distorts the symmetry (and normality) of a data distribution under study. Such a mixture data set may have been obtained from two or more populations with significantly different mean concentrations such as the one coming from the clean background area and the other obtained from a contaminated part of the site. Unfortunately, many times such a mixture data set or a data set with a few low probability outlying observations can be incorrectly modeled by a lognormal distribution with the lognormal assumption hiding the outliers and contamination (Singh, Singh, and Engelhardt (1997) and Singh, Singh, and Iaci (2002)).

One can argue against "not using the outliers" while estimating the various environmental parameters such as the EPC terms and BTVs. An argument can be made that the outlying observations are inevitable and can be naturally occurring (not impacted by site activities) in some environmental data sets. Often, in groundwater applications, a few elevated values (occurring with lower probabilities) are naturally occurring, and as such may not be representing the impacted and contaminated monitoring well data values. Those data values originating from the groundwater studies may require separate investigation, and all interested parties should decide on how to deal with data sets that contain naturally occurring unimpacted outlying observations. The project team should come to an agreement whether to treat the outlying observations separately or to include them in the computation of the various statistics of interest such as the sample mean, standard deviation, and the associated upper limits. However, it should be noted that statistics (e.g., mean, UCL) based upon the data set with outliers would yield distorted and inflated estimates of the population mass or average. The distorted estimate (a UCL95) of the population mass often exceeds the largest data value in the data set. It is noted that in such cases, practitioners (e.g., USEPA (1992)) often use the maximum value in a data set as an estimate of the population mean or mass. It does not seem desirable to estimate the population mass or average for the entire site area (e.g., AOC, EA) based upon the distorted statistics or the maximum observed data value. The authors of the report recommend that the outliers be treated separately. The extreme observations coming from the tails of a data distribution often represent low probability observations. The main objective of using a statistical procedure is to model the majority of the data representing the main dominant population, and not to accommodate a few outliers that may yield inflated and impractical results. The cleanup and remediation decisions for a site should be based upon reliable statistics (and not distorted statistics) and the data set representing the dominant population. A few outlying observations coming from the tails of the data distribution should be separately investigated.

Preprocessing of data to identify outliers and multiple populations should be conducted to obtain accurate and reliable estimates of the environmental parameters considered in this report. The user may want to use informal graphical displays (e.g., quantile-quantile plots, histograms) and formal population partitioning methods (e.g., see Singh, Singh, and Flatman (1994)) to identify multiple populations (if any). The UCL95 computation methods as considered in this report or in any other related reference such as Helsel (2005) should be used separately on each of the identified sub-population.

It should be noted that the methods as investigated and described in this report may be used on any data set with or without the outliers. However, the authors of the report want to caution and alert the users that the distorted statistics based upon data sets with outliers may lead to incorrect conclusions and decisions. These issues are illustrated by some examples in Section 6. If deemed necessary, all identified outliers should be excluded (with consultation and agreement of all interested parties) from all further analyses, including the computation of UCLs, UPLs, and UTLs. All of those identified environmental outliers perhaps represented by their locations, time period, laboratory, or analytical method require further investigation. Reliable and defensible statistics can be obtained based upon the majority of a data set representing the main body of the dominant population under study.

1.3 Disposition of Outliers and Use of a Lognormal Distribution

In addition to nondetects, outliers are also inevitable in most environmental applications. The issue of handling and disposition of outliers is a confusing and controversial topic in environmental applications, including the estimation of population mean or mass. Depending upon the objective (e.g., estimation of a background level concentration, or an EPC term) of the study and data collection, all interested parties including the project team should decide about the disposition of outliers. All relevant statistics should be computed on data sets with and without the outliers, the results compared, and a decision made accordingly. That is, the project team should decide which of the mean statistic (with outliers or without outliers) represents a better and more accurate estimate of the population mass under consideration. This topic is further discussed in Section 4, and several examples have been considered to address some outlier-related issues.

Outliers typically represent observations from different population(s) perhaps representing hot spots and contaminated areas impacted by the site activities. A data set consisting of such observations may represent a mixture sample from one or more populations. Those outliers representing hot spots and impacted site areas obviously require separate investigation. Many times, a few extreme observations come from the tails of the data distribution under study with much lower probabilities than the rest of the data coming from the main dominant population (e.g., site, monitoring well). High outlying values contaminate the underlying left-censored or uncensored full data set from the population under study. In practice, it is the presence of a few extreme outliers that cause the rejection of normality of a data set (Barnett and Lewis (1994)).

The use of a lognormal distribution to estimate the population mean or mass based upon a data set with mixture samples or a few outliers often results in unrealistically large estimates of population mass of no practical merit. Therefore, instead of accommodating a few outliers by using a lognormal distribution, it is desirable to separately investigate the outliers, and compute the relevant statistics based upon the main body of the data set representing the dominant population. The objective is to compute a meaningful estimate of the population mass of the main dominant population, and not to accommodate a few extreme observations resulting in distorted and unrealistic estimate of the population mass and various other parameters.

If possible, it is recommended that robust or resistant procedures be used to compute the mean, standard deviation, and other statistics from left-censored data sets. Several robust and resistant estimation methods have been considered and evaluated by Singh and Nocerino (2002). Robustness and resistance of an estimator go hand-in-hand. In practice, a resistant (to outliers) estimator also represents a robust estimator as it often turns out to be insensitive to the distributional assumptions. The user should make an effort to identify all potential outliers using appropriate statistical methods (e.g., see Huber (1981), Rousseeuw and Leroy (1987), Barnett and Lewis (1994), Singh (1993), and Singh and Nocerino (1995)) before proceeding with the estimation of population mean, standard deviation, UCLs, UPLs, UTLs, and other summary statistics based upon left-censored data sets. The detailed discussion and use of those robust and resistant regression methods is beyond the scope of this report. Instead of computing and using distorted summary statistics (e.g., an inflated average for the entire EA), a few outliers, if any, should be studied separately. Some of the classical and robust outlier identification methods have been incorporated in the Scout software package developed by the U.S. Environmental Protection Agency (USEPA (2002)).

1.3.1 What Represents a Distorted Estimate of Population Mass or Mean?

A distorted value of a statistic represents an unrealistic and unstable number of no practical merits. In the present application of estimation of population mean or mass, the value (a number) of the statistic larger than the largest observation (or > 1.5 times or twice the largest value) in a data set can be considered as a distorted estimate of the population mean. An estimate of the population mean or mass should not exceed the largest observation in the data set used to estimate that mean or mass. However, these occurrences are quite common, especially when one estimates the population mean based upon a data set with a few potential outliers assuming a lognormal distribution accommodating those few outlying observations. In such cases, it is a common practice (USEPA (1992)) to use the maximum (which itself may be an outlier) observed value as an estimate of the population mass. Just as before, it is recommended that the maximum detected value should not be considered as an estimate of the population mean or mass. Therefore, whenever the population mass (mean) needs to be estimated, instead of computing distorted statistics based upon a data set with a few potential outliers, every effort should be made to compute a reliable estimate of the population mass (average) using the data set representing the dominant population (e.g., EA, RU).

There seems to be no general consensus on how to appropriately treat the outliers in the various decision-making (statistical analyses) processes. However, most of the environmental scientists do recognize that outliers when present distort all statistics of importance, and, therefore, it is important to be able to identify potential outliers in environmental data sets. Since the treatment and handling of outliers is a controversial and subjective topic, this report suggests that the outliers be treated on a site-specific basis using all existing knowledge about the site (e.g., EA, RU, area of concern (AOC), and monitoring well) under investigation. The treatment of outliers and disposition of outliers (include or not to include) should be a team decision based upon the knowledge of the experts involved with the site investigations. The project team should clearly state the objective of estimating the population mean (by point estimate or by a UCL95). Such an estimate should be representative of the average or mass of the dominant population. All interested parties should understand and determine the importance of including or not including a few low probability outliers in the estimation of the mass of the dominant population. Specifically, the project team should decide whether to compute a reasonable and defensible estimate of the population mean based upon the majority of the data, or to compute a population mass by accommodating a few low probability outlying observations (perhaps by using a transformation), which could lead to a distorted and inflated estimate.

One of the objectives of this report is to clearly specify the inherent assumptions needed to estimate the population parameters based upon left-censored data sets. As far as the use of the mathematical UCL95 computation formulae is concerned, those formulae and methods can be used on any left-censored data set with or without the outliers. It is not a requirement to delete or omit the outliers (occurring with lower probabilities) before using the estimation or UCL95 computation methods (e.g., KM (BCA) UCL, MLE, and ROS methods). But, an explanation should be provided if outliers are included in the estimation process.

Classical MLE (CMLE, RMLE, and UMLE) and EM estimates are distorted by outliers as illustrated by Examples 1 through 3 in Section 4. Also, the ROS estimates of the intercept and slope (Rousseeuw and Leroy (1987)), and hence, the mean, the standard deviation, and the extrapolated nondetects, are distorted by the presence of even a single outlier. These distortions increase rapidly with even a minor increase in the standard deviation, σ_y (= σ) of the log-transformed variable Y = Ln(X). The extrapolated NDs based upon ROS (raw data) often result in infeasible negative estimates of nondetects. The use of a log-transformation alone does not result in robust and resistant estimates of the intercept and slope. One of the advantages of using ROS on log-transformed data is that the extrapolated NDs cannot become negative. It is, however, noted that, contrary to the statement made in Shumway, Azari, and Kayhanian (2002), the extrapolated nondetects do become larger than the detected observed values even for well-behaved normally distributed data sets, as used in Example 1 of Section 4.

All interested parties should come to a consensus about the inclusion or exclusion of outliers in the computation of the summary statistics and UCL95 of the population mean. Based upon professional experience, the authors of the report recommend that summary statistics and all other related upper limits be computed based upon the main dominant population (e.g., AOC, EA) without including a few potential outliers, especially when the objective is to estimate the overall mean (mass) contaminant concentration of an AOC (or of an EA, RU, or a monitoring well) under investigation. In practice, such a mean statistic without using a few outliers represents the population "mass" or average better that other statistical methods based upon sample percentiles such as the median, 75th, and 90th percentiles. It is recommended that all relevant statistics be computed both with and without the outliers, and compare the results to evaluate the potential impact of outliers on the statistics used in the various decision-making processes.

Section 2

Mathematical Formulation and Distributions Considered

Censoring generally means that observations at one or both extremes (tails) are not available. In Type I censoring, the point of censoring (e.g., the detection limit, DL) is "fixed" *a priori* for all observations, and the number, *k*, of the censored observations varies. In Type II censoring, the number of censored observations, *k*, is fixed *a priori*, and the point(s) of censoring vary. For example, Type II right censoring (large values are not available) typically occurs in lifetime expectancy testing and reliability applications. In a life testing application, *n* items (e.g., electronic items, laboratory animals) are subjected to a lifetime expectancy testing experiment that terminates as soon as (*n*-*k*) of the *n* data values have been observed (failed/died). The lifetime of the remaining *k* living objects is unavailable or being censored.

In this report, we are concerned about Type 1 left-censoring UCL95 computation methods. The computation of the mean, standard deviation, and quantiles of normal and lognormal populations from censored samples has been studied by several researchers, including Cohen (1950, 1959), Perrson and Rootzen (1977), Gleit (1985), Schneider (1986), Gilliom and Helsel (1986), Kroll and Stedinger (1996), She (1997), Shumway, Azari, and Kayhanian (2002), and Singh and Nocerino (2002). These articles cover a myriad of procedures to estimate the sample mean and the standard deviation, including the simple substitution methods and likelihood procedures such as Cohen's maximum likelihood estimation (CMLE) procedure, Perrson and Rootzen's restricted MLE (RMLE) method, and regression on order statistics (ROS) methods (Gilliom and Helsel (1986), Newman, Dixon, and Pinder (1990) and Helsel (1990)). The simple substitution methods are the replacement of below detection limit data by zero, by half of the detection limit, DL/2, or by the detection limit, DL, itself. Helsel (2005) discusses the performances of several of the estimation methods as described in the above references. Based upon the findings of the various researchers, Helsel (2005) summarizes some recommendations to compute the summary statistics for left-censored data sets (page 78). It is noted that the recommendations as described by Helsel (2005) are for the computation of summary statistics such as the sample mean and the standard deviation and not for the computation of UCL95 based upon left-censored data sets. It is also noted that not much is known about some of the recommended methods such as the robust MLE method (Kroll and Stedinger (1996)). In this report, we evaluate the performance of the several UCL95 computation methods (Section 8) and make recommendations (in Section 9) accordingly. It is observed that an estimation method (e.g., MLE method, EM method) which may be considered a good method for estimation of the population mean and *sd* may not be good enough to compute a UCL95 for the population mass.

Using Monte Carlo simulation experiments, several researchers, including Gleit (1985), Gilliom and Helsel (1986), and Haas and Scheff (1990), Singh and Nocerino (2002), concluded that the data substitution methods (including the uniform generation of NDs in the interval (0, DL]) result in biased point estimates of the population mean. In practice, probably due to computational ease, these data substitution methods are commonly used in many environmental applications. Depending upon the sample size, *n*, and the censoring intensity, *k*, substitution of the censored values by DL/2 is one of the recommended estimation methods in some USEPA guidance documents (e.g., USEPA (2000), USEPA (2006)). Another method that the practitioners sometimes want to use is the random generation (uniformly distributed) of the *k* nondetects (NDs) in the interval (0, DL]. Singh and Nocerino (2002) concluded that this uniform generation method does not work well and produces biased estimates of the population mean and the standard deviation. Therefore, the uniform generation method and other proxy (e.g., substituting NDs by '0', or DL) methods have not been included in this study dealing with the computation of UCL95. Since the use of a lognormal model (e.g., Singh, Singh, and Engelhardt (1997)) results in unstable and unrealistically large UCLs as illustrated by examples in Section 6, the use of a lognormal distribution should be avoided. Emphasis is given to evaluate and compare the performances of the various nonparametric (e.g., bootstrap and Chebyshev on KM estimates) UCL95 computation methods that can be used on any left-censored data set-symmetric or skewed.

In order to thoroughly address the issue of UCL95 computations from left-censored data sets, several UCL95 computation methods, including the EM algorithm, MLE, UMLE, RMLE methods, regression on order statistics (ROS) on raw and log-transformed data (fully parametric and robust ROS), EPA delta method, KM method, and nonparametric jackknife and bootstrap (standard, bootstrap t, percentile bootstrap, and BCA bootstrap) methods, have been considered in the Monte Carlo simulation experiments as summarized in Section 7. Three distributions: a normal distribution, a lognormal distribution, and a gamma distribution have been included in the simulation study. Singh, Singh, and Iaci (2002) and Singh and Singh (2003) concluded that UCL95 computation methods, which perform well on symmetric and mildly skewed data sets, may not perform well on moderately skewed to highly skewed data sets. Therefore, in this report, several distributional parameters have been considered to cover a wide range of skewness from mildly skewed to highly skewed distributions. Type 1 left-censored data sets of various sizes for several computed censoring levels (10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, and 70%) have been generated from the three distributions (normal, lognormal, and gamma). The sample sizes evaluated are: 10, 15, 20, 25, 35, 40, 50, 75, and 100. Some fixed detection limit cases have also been considered.

In an effort to document all of the UCL95 computation methods in one report, most of the UCL95 methods for left-censored data sets cited and mentioned in the literature have been included (perhaps with some modification) in this report. It is important to address the issue of the appropriate treatment of outliers as data sets with a few high outliers that can be modeled by a lognormal distribution – thus hiding and accommodating potential contamination represented by outliers. As mentioned before, this may require some preprocessing of the data set using population partitioning (e.g., Singh, Singh, and Flatman (1994); Singh and Singh (1996)) and outlier identification (Barnett and Lewis (1994)) methods. These issues have been illustrated by using a couple of skewed data sets as discussed in Examples 6 and 7 in Section 6. These examples also demonstrate that in practice, the use of a lognormal distribution can yield "unstable and impractical" results of no practical merit. This is especially true for moderately skewed to highly skewed data sets of smaller sizes (e.g., Singh, Singh, and Iaci (2002)).

2.1 Distributions Considered

Three distributions, normal, gamma, and lognormal (Johnson, Kotz, and Balakrishnan (1994)), have been included in the present simulation study. Singh and Singh (2003) used bootstrap and Monte Carlo simulation experiments to compare the performance of the various UCL computation methods for full uncensored data sets generated from normal, lognormal, and gamma distributions. In this study, we compare the performances of the various UCL95 computation methods for left-censored data set.

2.1.1 The Normal Distribution

Let X be a continuous random variable (e.g., concentration of COPC) that follows a normal distribution, $N(\mu,\sigma^2)$ with mean, μ, and variance, σ^2. The probability density function of X is given by the following equation:

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp[-(x-\mu)^2/2\sigma^2]; \qquad -\infty < x < \infty$$

As mentioned before, several maximum likelihood estimation (MLE) methods are available in the literature to estimate sample mean and the standard deviation for left-censored data set. Some of those are described later in Section 3.

2.1.2 The Gamma Distribution

Singh, Singh, and Iaci (2002) studied the gamma distribution to model positively skewed environmental data sets. For full data sets, the use of a gamma distribution results in reliable and stable 95% UCL values. It is, therefore, desirable to test if an environmental data set follows a gamma distribution. If a skewed data set follows a gamma model, then a 95% UCL of population mean may be computed using a gamma distribution. For details of gamma goodness-of-fit tests, estimation of gamma parameters, and computation of a 95% UCL of the mean based upon a gamma distribution, refer to Singh, Singh, and Iaci (2002). In this study, gamma distribution has been used to compute UCL95 based upon left-censored data sets.

A continuous random variable, X (e.g., concentration of COPC), is said to follow a gamma distribution, G(k, θ), with parameters $k > 0$ and $\theta > 0$, if its probability density function is given by the following equation:

$$f(x;k,\theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-x/\theta}; \qquad x > 0$$
$$= 0; \qquad \textit{otherwise}$$

The parameter, k, is the shape parameter, and θ is the scale parameter. Many positively skewed data sets follow a lognormal as well as a gamma distribution. A gamma distribution can be used to model positively skewed environmental data sets. In addition to jackknife and bootstrap methods, a UCL95 method based upon linear regression on order statistics (ROS) of (n-k) pairs (gamma quantiles versus ordered detected values) has also been described in the following sections. If NDs are present, it is desirable to use an available goodness-of-fit test (e.g., as in ProUCL 3.0 and all later versions) on the detected data supplemented by a gamma quantile-quantile (Q-Q) plot.

2.1.3 The Lognormal Distribution

Environmental data are often (by default) modeled by a lognormal distribution. Singh, Singh, and Iaci (2002) and Singh and Singh (2003) compared the performances of the various UCL95 computation methods for full (uncensored) data sets obtained from lognormal and gamma distributions. In this report, similar methods have been used to compare the performances of the various UCL95 computation methods for Type 1 left-censored data sets.

Let $x_1, x_2, ..., x_n$ be a random sample from a lognormal population, LN(μ, σ^2), where the natural logarithm, y, of data is normally distributed, N(μ,σ^2), with the mean, μ, and the variance, σ^2. Let $\bar{y}$ and s_y be the sample mean and sample sd, respectively, of the log-transformed data, $y_i = \log(x_i)$; $i = 1, 2, ..., n$. The mean, μ_1, of the lognormal population in the original x-scale is given by $\mu_1 = \exp(\mu + 0.5\sigma^2)$, the variance is given by $\sigma_1^2 = [\exp(2\mu + \sigma^2)][(\exp(\sigma^2) - 1)]$, and the median is given by M = exp(μ). The sample mean, $\bar{y}$, and the sample standard deviation (*sd*), s_y, are given by:

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

2.2 UCL of the Mean, *μ₁*, of a Lognormal Model Based Upon Land's Method (Full Data Set)

The use of Land's H-UCL on MLEs obtained using log-transformed left-censored data set has been mentioned as one of the potential UCL95 computation method in the literature. This method may also be used (not a recommended method though) on the full data set obtained by extrapolating the NDs based upon ROS on MLEs (Kroll and Stedinger (1996)), or ROS on log-transformed left-censored data set. For full data sets, a (1 – α) 100% UCL for the mean, μ_1, based upon Land's H-statistic (Land (1971)) is given by:

$$\text{UCL} = \exp\left(\bar{y} + 0.5 s_y^2 + s_y H_{1-\alpha} \big/ \sqrt{n-1}\right),$$

where $\bar{y}$ and s_y^2 are the sample mean and variance of the log-transformed data. For left-censored data sets, these estimates are replaced by the MLEs (Kroll and Stedinger (1996)) or ROS estimates based upon log-transformed data sets. Technically, on the data set obtained using ROS on log-transformed data (which assumes a lognormal distribution for both detected and nondetected data), the H-statistic can be used to compute a UCL95 based upon the full data set obtained using the detected and extrapolated nondetects in the log scale. The 95% H-UCL given by the above equation does provide the specified 95% coverage (Singh and Singh (2003)) for the lognormal population mean, μ_1. However, the practical merit of the associated H-UCL is quite questionable as it may result in unacceptably high UCL values. This is especially true for samples of small size (e.g., $n < 50$) with values of s_y exceeding 1.5-2.5 (e.g., Singh, Singh, Engelhardt (1997) and ProUCL 3.0 User Guide (2004)). A similar behavior is observed for the UCL95 based upon the USEPA (1991) delta log method, and the H-UCL95 computed using the MLEs or the full data set with NDs estimated using the ROS on log-transformed data.

The Monte Carlo results, as summarized in this report (Section 8), suggest that, just as for the case of full data sets, the H-statistic-based UCL does provide approximate 95% coverage of the population mean, but may yield unrealistic, unstable, and inflated UCL values, especially when the sample size is small (e.g., n < 50-70) and the skewness is high with the standard deviation of log-transformed data exceeding 1, 1.25, and so on. It should also be noted that for larger samples (e.g., >100), the H-UCL sometimes results in a value smaller than the sample mean of the detected data-which is again questionable. Therefore, it is again recommended to avoid the use of a lognormal distribution when computing the UCL95. It is preferable to use nonparametric methods to compute a UCL95 of the mean. A description of some of the available estimation methods for left-censored data sets is given in Section 3 and several nonparametric UCL95 computation methods based upon resampling bootstrap and jackknife methods are described in Section 5.

Section 3

Estimation of Population Mean and Variance Based Upon Left-Censored Data Sets

This section provides a description of the various methods available to estimate the population mean, and the standard deviation based upon left-censored data sets. It has been implicitly assumed that the data set under consideration has been obtained from a "*single*" parametric (e.g., normal, lognormal, and gamma) or nonparametric population. This assumption is needed for the validity of the use of a UCL95 as an estimate of the mean of the population under study, such as a study area (SA), area of concern (AOC), remediation unit (RU), exposure unit (EU), or an exposure area (EA).

It is suggested to avoid the use of transformations (Singh, Singh, Engelhardt (1997) and Singh, Singh, and Iaci (2002)) on the raw data sets to achieve symmetry (approximate normality). Typically, the parameter (hypothesis of interest) in the transformed space is not of interest to make remediation and cleanup decisions. Many times, the practitioners do not know how to interpret the transformed results or back-transform the results in the original scale. For example, the program UNCENSOR 5.1 can compute the CI of the population mean in the log scale, but does not provide any guidance to a user on how to interpret, back-transform, or use those intervals to estimate the population mass. Shumway, Azari, and Kayhanian (2002) used Box-Cox (BC) type transformation on skewed data sets to achieve symmetry. They used a mildly skewed (almost normal) lognormal distribution, LN(μ_x = 2.77, $sd = \sigma_x$ = 0.56), to illustrate their procedures. Any of the available estimation methods (e.g., MLE, EM, and ROS) will work equally well on the raw data sets obtained from such mildly skewed populations. There is no need to use a BC type transformation on data sets obtained from such a mildly skewed population. For such mildly skewed data sets with σ_x = 0.56 (and σ_y = 0.2 as can be seen from Tables 3-1 and 3-2), estimates and UCL95 based upon raw data will do just equally well. It is desirable to demonstrate the gains and advantages (in terms of bias and percent coverages achieved by UCL95) of using a transformation in the estimation and UCL computation methods.

3.1 Impact of Skewness on Estimation Methods and Robustness of Methods Used on Log-Transformed Data Sets

In order to support some of the statements made above about skewness and avoiding the use of transformations such as a log-transformation to achieve symmetry, a couple of tables (Table 3-1 and Table 3-2) have been constructed showing the relationships between sd, σ_y of log-transformed variable, Y = log(X), and CV and skewness of the lognormal variable, X, in the original scale. In the following, the subscript y is used for log-transformed variable, Y = log(X), and subscript x is used for variable, X, in the original raw scale. Also, note that CV and skewness of a lognormal variable, X, are functions of sd, σ_y of log-transformed variable, Y. In this report, skewness for a lognormal and for any other skewed distribution is defined in terms of standard deviation, σ_y, of the log-transformed variable, Y. Based upon the results (as incorporated in ProUCL 3.0) of an extensive simulation study, Singh and Singh (2003) concluded that for mildly skewed lognormal or other distributions (with sd, σ_y of log-transformed variable < 0.5), the difference between a UCL95 based upon Student's t-statistic (assuming a normal distribution) or any other parametric (Land's H-UCL) or nonparametric method (e.g., bias-corrected accelerated (BCA) bootstrap) is not of any practical significance.

Therefore, for such mildly skewed distributions with CV = 0.6, as cited and used in the Colorado state document (Colorado Water Quality Control Division (WQCD) (2003)), or with *sd*, $\sigma_y = 0.2$ (corresponding to $\sigma_x = 0.56$ as used in Shumway, Azari, and Kayhanian (2002)), there is no need to use a transformation to achieve symmetry using a Box-Cox (BC) type transformation. As mentioned before, for such low values of skewness, all parametric and nonparametric methods on raw data as well as on transformed data will yield similar and comparable results (use ProUCL 3.0). For values of standard deviation, σ_y, exceeding 1, the estimates and the UCL95 change drastically with a minor increase in standard deviation, σ_y of Y. Thus, the observations made and derived for mildly skewed distributions (as in Colorado WQCD (2003) and Shumway, Azari, and Kayhanian, (2002)) cannot be generalized and determined to be robust for moderately skewed to highly skewed distributions with σ_y of log-transformed data exceeding 1.0. The conclusions about the robustness of ROS and EM methods as derived in Shumway, Azari, and Kayhanian (2002) need to be demonstrated for higher values of CV, skewness, and σ_y.

Table 3-1. Relationship between Parameters of the Distributions of *X* and *Y*

***Y* = log (*X*)**		***X* is Lognormal**			
μ_y	σ_y	μ_x	σ_x	CV_x	$Skewness_x$
2	0.25	7.62	1.94	0.25	0.78
2	0.50	8.37	4.46	0.53	1.75
2	0.75	9.79	8.51	0.87	3.26
2	1.00	12.18	15.97	1.31	6.18
2	1.50	22.76	66.31	2.91	33.47
2	2.00	54.60	399.72	7.32	414.36
2	2.20	83.10	930.79	11.20	1,439.03
2	2.40	131.63	2,341.21	17.79	5,679.99
2	2.60	217.02	6,370.42	29.35	25,380.48
2	2.80	372.41	18,766.02	50.39	128,103.03
2	3.00	665.14	59,870.45	90.01	729,551.38

Table 3-2. Relationship between Parameters of the Distributions of *X* and *Y*

***Y* = log(*X*)**		***X* is Lognormal**			
μ_y	σ_y	μ_x	σ_x	CV_x	$Skewness_x$
5	0.25	153.12	38.89	0.25	0.78
5	0.50	168.17	89.63	0.53	1.75
5	0.75	196.62	170.85	0.87	3.26
5	1.00	244.69	320.75	1.31	6.18
5	1.50	457.14	1,331.83	2.91	33.47
5	2.00	1,096.63	8,028.53	7.32	414.36
5	2.20	1,669.03	18,695.36	11.20	1,439.03
5	2.40	2,643.87	47,024.40	17.79	5,679.99
5	2.60	4,359.01	127,953.23	29.35	25,380.48
5	2.80	7,480.09	376,925.61	50.39	128,103.03
5	3.00	13,359.73	1,202,530.08	90.01	729,551.38

For clarification, relationships between standard deviation, coefficient of variation (CV), and skewness of a lognormal variable, *X*, and its transformed variable, $Y = \log(X)$, are summarized in Tables 3-1 and 3-2 as follows. CV and skewness of a lognormal variable, *X*, only depend on the standard deviation, σ_y, of the log-transformed variable, *Y*. A quick review of the following tables reveals that values of CV = 0.6 (as used in Colorado WQCD (2003)), CV < 1, or CV < 2 represent only mildly skewed distributions. UCL95 computation methods for low values of CV < 1, 2 behave in a very significantly different manner (e.g., in terms of coverage probabilities) than the UCL95 methods for higher values of CV exceeding 2, 3, and so on. Values of CV = 3, 4, and higher are very common for lognormally distributed environmental data sets. Conclusions, results, and robustness of methods for distributions with CV < 1 cannot be generalized to all lognormal distributions with CV exceeding 2.0, and so on. Most of the examples and simulation results as considered in the environmental literature (e.g., Shumway, Azari, Kayhanian (2002), Kroll and Stedinger (1996), and Colorado WQCD (2003)) deal with mildly skewed distributions with low values of CV, such as 0.6, 1.0.

From Tables 3-1 and 3-2, it is easy to note that for a minor increase in σ_y, such as from 1.0 to 1.5, the skewness increases from 6.18 to 33.47; for σ_y increasing from 2.0 to 2.2, the skewness increases from 414.36 to 1439! As mentioned before, the estimated values of σ_y = 1.5, 2.0, and 2.5 for log-transformed data are quite common for data sets originated from environmental applications. The robustness of the methods (e.g., robust ROS on log-transformed data, robust MLE method, and EM method followed by jackknife method) found to be effective in the environmental literature (e.g., Shumway, Azari, and Kayhanian (2002)) should be demonstrated to be robust and effective for moderately skewed to highly skewed data distributions with σ_y exceeding 1.0, 2.0, and so on. Once again, it is reiterated that the conclusions derived for low values of σ_y (< 0.5, 1) cannot be generalized for moderately skewed to highly skewed data sets with σ_y exceeding 1.0.

For a mildly skewed lognormal variable, $X \sim LN(2.77, \sigma_x = 0.56)$, as considered in Shumway, Azari, and Kayhanian (2002), the *sd*, σ_y of log-transformed variable, $Y = \log(X)$ is about 0.20 (< 0.5). For such a distribution, all estimation methods for left-censored data sets (e.g., MLE, EM, ROS on raw and ROS on log-transformed data) will yield almost similar results as also determined by Shumway, Azari, and Kayhanian (2002). Since the data are only mildly skewed, any of the MLE methods (RMLE, EM) on the raw scale can be used. There is no need to complicate the process by using a BC type transformation and introducing an unknown amount of transformation bias in the estimates. Moreover, the results of the simulation experiment conducted for a mildly skewed lognormal distribution with σ_x = 0.56 cannot be generalized to other heavily skewed (with *sd*, σ_y of log-transformed data exceeding 1, 1.25) lognormally distributed data sets often encountered in environmental populations.

It is also noted that the robustness of ROS on log-transformed data has been studied only for mildly skewed distributions with low CV values (Gilliom and Helsel (1986) and Shumway, Azari, and Kayhanian (2002)). As mentioned before, the results and conclusions obtained for mildly skewed distributions with low CV values should not be generalized to moderately and highly skewed distributions, especially for the lognormal distribution with standard deviation of log-transformed data exceeding 1, 1.25, and so on. In order to truly establish and demonstrate the robustness of ROS estimation method on log-transformed distribution or other skewed distributions, some moderately and highly skewed distributions with *sd*, σ_y of log-transformed data exceeding 1, 1.25, 2, and 2.5 should be considered. Following Helsel's November 2005 review comments on an earlier version of this report, Monte Carlo experiments have been used on several distributions covering a wide range of skewness to evaluate the performance of the UCL95 computation method based upon robust ROS followed by bootstrap methods.

It is also well known that the estimates of the back-transformed parameters (from transformed space based upon a BC type transformation) in the original space may suffer from an unknown amount of transformation bias (e.g., back-transformation of ROS estimates to original scale). Many times, the transformation bias can be quite large (for highly skewed data sets) and unreasonable, leading to incorrect decisions. Moreover, it is not always possible to use transformed parameters and assign physical meaning to them. In several environmental applications, it is the mean parameter that is used to make cleanup and remediation decisions. Specifically, the mean contaminant concentrations are used in site characterization, exposure and risk assessment, and background evaluation studies.

3.2 Jackknife Method to Compute a UCL95 Based Upon Sample Mean of Full Data Set

The details about the computation of a jackknife UCL are given in Section 5. Helsel reviewed an earlier version of this report in November 2005. Following Shumway, Azari, and Kayhanian (2002), Helsel recommended the use of the jackknife method on the full data set obtained using the robust ROS method to compute a UCL95. The jackknife method is a useful resampling technique used to reduce bias in an estimator. However, it is noted that a jackknife UCL95 of the population mean based upon sample mean and the standard deviation of a full data set is equivalent to Student's t-UCL95 (Dudewicz and Misra (1988) and ProUCL 3.0 User Guide (2004)). These methods, such as Student's t-UCL95 = jackknife UCL95 on robust ROS, robust MLE (Kroll and Stedinger (1996)), or an EM method, provide adequate coverage to the population mean only for mildly skewed data sets with *sd* of log-transformed variable < 0.5 (Singh and Singh (2003) and ProUCL 3.0 User Guide (2004)). The coverage of the mean provided by Student's t-UCL95 (hence by jackknife UCL) deteriorates (decreases) fast with increases in skewness. Therefore, the UCL results as summarized in Shumway, Azari, and Kayhanian (2002) cannot be generalized for moderately skewed to highly skewed distributions with *sd*, σ_y of log-transformed data exceeding 1. The jackknife UCL95 as suggested in Singh, Singh, and Engelhardt (1997) based upon an estimator other than the sample mean, such as the median or the minimum variance unbiased estimator (MVUE) of the mean of a lognormal distribution, will be different from the Student's t-UCL95. This fact is clearly stated in the jackknife section of the ProUCL 3.0 User Guide (2004).

3.2.1 Helsel Robust ROS on Log-Transformed Data Followed by Jackknife Method

It is noted that the EM method (Shumway, Azari and Johnson (1989), Gleit (1985), and Singh and Singh (2002)) is equivalent to a substitution and fill-in method (such as the substitution by DL/2 method), where the fill-in values are based upon the conditional expectation restricted to fall below the detection limit, DL. Similarly, Helsel's robust ROS method and ROS based upon robust MLE (Kroll and Stedinger (1996)) method yields full data sets obtained by extrapolating the nondetected values. Thus, the use of the jackknife UCL95 method on the full data set obtained using the ROS or EM methods will simply yield a Student's t-UCL95. As mentioned before, a Student's t-UCL95 (or equivalently jackknife UCL95 based upon sample mean) does not provide adequate coverage (ProUCL 3.0 User Guide (2004) and Singh and Singh (2003)) to population mean of moderately skewed to highly skewed (e.g., when $\sigma_y > 1.0$) data sets. Therefore, for highly skewed data distribution with σ_y>1.0, the EM method, robust ROS method, or ROS on MLEs (Kroll and Stedinger (1996)) method followed by the jackknife UCL95 method may not be a reasonable method to use, especially when the skewness is high. This was the reason that the authors of this report did not consider robust ROS followed by the jackknife UCL95 method (and also bootstrap methods) in the simulation experiments to compute a UCL95 of the population mean. Once one has obtained a full data set based upon robust ROS, MLE ROS method, or the EM method, ProUCL 3.0 can be used to compute an appropriate UCL95 based upon full data set with extrapolated NDs.

3.2.2 Helsel Robust ROS on Log-Transformed Data Followed by Bootstrap Methods

In his November 2005 review, Helsel also suggested the use of bootstrap methods on the full data set obtained using his robust ROS method on log-transformed data. Bootstrap methods to compute UCL95 on full data sets obtained using robust ROS on log-transformed data, ROS on MLEs, or an EM method are already available in ProUCL 3.0. For the sake of direct comparison of the coverage probabilities, the authors have conducted additional simulations (Sections 7 and 8) to include bootstrap UCL95 methods on the robust ROS method as suggested by Helsel in November 2005. The graphical displays of coverage probabilities and average UCLs for some of the new simulation results have been included in Appendices D and E. Our recommendations based upon the older (Appendices A, B, and C) as well as newer (Appendices D and E) simulation results have been summarized in Sections 8 and 9.

3.3 Classical and Robust Estimation of the Mean and the Standard Deviation

This section is an introductory section and describes various methods to estimate the population mean, and the standard deviation based upon left-censor data sets. The robust versions of the various estimation methods have been described. The corresponding classical estimates can be obtained by simply replacing each of the n weights, w_i, i: = 1, 2, …, n, by 1. Formally, let $x_1, x_2, ..., x_n$ be a random sample from a normal population, $N(\mu,\sigma^2)$, with k of the nondetects, $x_1, x_2, ..., x_k$, lying numerically below the detection limit, DL, which is denoted by L in this section. The normality assumption is needed for the various MLE methods (CMLE, UMLE, and RMLE) and the EM method. Let φ and Φ be the probability density function (pdf) and cumulative distribution function (cdf) of the standard normal distribution (SND). The logarithm of the likelihood function is given as follows:

$$\ln L(x,\mu,\sigma) = k\ln\Phi(Z) - n_0 \ln\sigma - \sum_{k+1}^{n}(x_i-\mu)^2 \big/ 2\sigma^2 + \text{constant} \qquad (3\text{-}1)$$

where $n_0 = (n-k)$ and $Z = (L-\mu)/\sigma$, $\Phi(Z)$ representing the probability that an observation is less than DL (= L in the formula of Z). The mean, $\bar{x}_o$, and variance, s_o^2, obtained using the $(n-k)$ detected data values are:

$$\bar{x}_o = \frac{\sum_{i=k+1}^{n} x_i}{n-k} \text{ and } s_o^2 = \frac{\sum_{i=k+1}^{n}(x_i-\bar{x}_o)^2}{n-k}. \qquad (3\text{-}2)$$

Note that in equation (3-2), the denominator of sample variance is $(n-k)$ and not $(n-k-1)$. The program UNCENSOR 5.1 uses the factor $(n-k-1)$ in the denominator of sample variance. In SimCensor, the sample variance formula (3-2) has been used in the derivation of Cohen's MLE (CMLE) and various other MLE estimates. Due to this difference, the statistics based upon the detected observations obtained using UNCENSOR 5.1 and our development program, SimCensor, are slightly different. It should be, however, noted that it is the sample variance given by (3-2) which has been used in the various MLE equations (e.g., Schneider (1986)); therefore, the correct denominator of the sample variance should be $(n-k)$.

A brief description of some of the procedures to estimate population parameters for left-censored samples is given below. The robust versions of those procedures are described in the following. It should be noted that the robust and resistant procedures perform better than their classical counter parts in the presence of outliers (Singh and Nocerino (2002)). It is easy to obtain the classical estimates from the following robust estimates by replacing each of the n weights, w_i, i: = 1, 2, …, n, by 1. The simulations results as described in Sections 7, 8, and 9 are based upon the various classical UCL 95 methods (Section 5) for left-censored data sets.

The robust estimate formulae, as presented below, are called the robust M-estimation procedures based upon the notion of the influence function (Hampel (1974)). The influence functions are used to assign reduced weights to the outlying (contaminating) observations. For fully uncensored data sets, several robust procedures exist in the literature for the estimation of population mean and variance (Huber (1981), Rousseeuw and Leroy (1987), Staudte and Sheather (1990), and Singh and Nocerino (1995)). For left-censored data sets, Singh and Nocerino (2002) studied the performances of the various robust estimation (of the mean and the standard deviation) procedures based upon the PROP influence function. Those robust methods are included in this report to illustrate the influence of outliers on the various estimates as shown in Examples 2 and 3 of Section 4.

For left-censored data sets, in order to identify and subsequently assign reduced weights to the outliers that may be present in the right tail of a data set, robust sample mean, $\bar{x}_o^*$, and *sd*, s_o^* using the (n – k) detected values need to be obtained first. These values are then used in the various estimation methods, such as MLE, UMLE, RMLE, and the EM method, to obtain robust estimates of the population mean and *sd*. The PROP influence function and the corresponding iteratively obtained sample mean and *sd* based on (n – k) detected data are given as follows:

$$\begin{aligned}\psi(d_i) &= d_i &&; d_i \le d_0 \\ &= d_0 \exp[-(d_i - d_o)] &&; d_i > d_0\end{aligned} \tag{3-3}$$

$$\bar{x}_o^* = \sum_{i=k+1}^{n} w_1(d_i)x_i \Big/ \sum_{i=k+1}^{n} w_1(d_i)\,;\; s_o^{*2} = \sum_{i=k+1}^{n} w_2(d_i)(x_i - \bar{x}_o^*)^2 / v. \tag{3-4}$$

Here, $d_i^2 = (x_i - \bar{x}_o^*)^2 / s_o^{*2}$; $i = k + 1, k + 2, \ldots, n$, and d_0^2 is the α*100% critical value from the scaled beta distribution, $(n-k-1)^2 \beta(1/2,(n-k-2)/2)/(n-k)$ of the distances, d_i^2. The weights are given by $w_1(d_i) = w_i = \psi(d_i)/d_i$ and $w_2(d_i) = w_1^2 = w_1^2(d_i)$, with $v = (w\text{sum2} - 1)$, and $w\text{sum1} = \sum w_1(d_i)$, and $w\text{sum2} = \sum w_2(d_i)$.

3.4 Cohen's MLE (CMLE) and Unbiased MLE (UMLE) Methods

Cohen's MLEs for the mean and variance are obtained by solving the following equations:

$$\hat{\mu}_{MLE} = \bar{x}_o - (\bar{x}_o - L)\lambda(g,h) \text{ and } \hat{\sigma}^2_{MLE} = s_o^2 + (\bar{x}_o - L)^2 \lambda(g,h), \tag{3-5}$$

where $g = s_o^2/(\bar{x}_o - L)^2$, $h = k / n$, and L = DL. These ML estimates have been computed by using numerical methods rather than the using the look-up tables developed by Cohen. The estimates of μ and σ given by equation (3-5) are biased. For a Type II censored data set from a normal population, Saw (1961) tabulated the first-order bias correction terms, which were simplified by Schneider (1986). The bias correction terms are given as follows:

$$Bias_{\hat{\mu}} = -\exp[2.692 - 5.439(n-k)/(n+1)] \text{, and} \quad (3\text{-}6)$$

$$Bias_{\hat{\sigma}} = -[0.312 + 0.859(n-k)/(n+1)]^{-2} . \quad (3\text{-}7)$$

In practice, the bias corrections given by formulas (3-6) and (3-7) are also used for Type I censored data. The bias-corrected MLE denoted by UMLE are given as follows:

$$\hat{\mu}_{UMLE} = \hat{\mu}_{MLE} - \frac{\hat{\sigma}_{MLE} Bias_{\hat{u}}}{(n+1)} \text{ and } \hat{\sigma}_{UMLE} = \hat{\sigma}_{MLE} - \frac{\hat{\sigma}_{MLE} Bias_{\hat{\sigma}}}{(n+1)} . \quad (3\text{-}8)$$

The corresponding robust ML and UML estimates of μ and σ are obtained simply by using the robust estimates, $\bar{x}_o^*$ and s_o^*, in place of $\bar{x}_o$ and s_o in (3-5) and (3-8), respectively.

3.4.1 Difference between MLE Method and Cohen's MLE Method

During the 1950s, Cohen (1950, 1959) derived the maximum likelihood (ML) equations for censored samples and prepared tables (due to unavailability of computers and software programs) of the constants needed to compute the MLEs of μ and σ as given by equation (3-5). However, today, instead of using those look-up tables, one can easily use a personal computer to solve the ML equations iteratively using a suitable numerical method such as the Newton-Raphson method (Faires and Burden (1993)). In the examples and simulation experiments as conducted in this report, the MLE estimates of the mean and the standard deviation based upon left-censored data sets have been computed using the numerical Newton-Raphson method. Cohen's look-up tables of constants were not used which was not even possible for the simulation study as described in Section 7.

Some authors (Helsel (2005), and Shumway, Azari, and Kayhanian (2002)) have differentiated between the ML estimates obtained using the numerical Newton-Raphson method and Cohen's ML estimates based upon the look-up tables. In this report, we make no such distinction. It is understood that with the availability of fast personal computers, most of the users are using numerical methods to compute the critical values and constants needed for various statistical methods including Cohen's MLE method. Also, it will not be possible to conduct the simulation experiments as described in this report without directly using (as coded in SimCensor) the numerical methods to compute the MLE (CMLE) estimates. In this report, we have used both CMLE as well as MLE to represent the MLE estimates obtained using the Newton-Raphson method.

3.5 Expectation Maximization (EM) Algorithm

Dempster, Laird, and Rubin (1977) developed the EM algorithm to maximize the likelihood function based upon censored and missing data. The iterative EM algorithm works on the detected values assuming that no observations were censored. At the initial iteration, using the (*n-k*) detected data values, one could start with some convenient estimates for μ and σ, such as the sample mean and *sd*, or a simple one-step robust pair represented by the median and MAD/0.6745. The iterations are defined as successively maximizing the expectation of the conditional likelihood function of the complete data. Gleit (1985) used this procedure for left-censored samples and found it to possess a lower mean square error (MSE) than the various other substitution and likelihood procedures. For the single DL case, the estimates of μ and σ at the $(j+1)^{\text{th}}$ iteration are given as follows (Shumway, Azari, and Johnson (1989); also see Section 3.3):

$$\hat{\mu}_{j+1} = [\sum_{i=k+1}^{n} x_i + \sum_{i=1}^{k} E_j(x_i \mid x_i \le L)]/n, \tag{3-9}$$

$$\hat{\sigma}^2_{j+1} = [\sum_{i=k+1}^{n} (x_i - \hat{\mu}_j)^2 + \sum_{i=1}^{k} E_j((x_i - \mu_j)^2 \mid x_i \le L)]/(n-1), \tag{3-10}$$

$$E_j(x_i \mid x_i \le L) = \hat{\mu}_j - \hat{\sigma}_j[\varphi(Z)/\Phi(Z)], \text{ with } Z = (L - \mu_j)//\sigma_j, \tag{3-11}$$

$$E_j((x_i - \mu_j)^2 \mid x_i \le L) = \hat{\sigma}^2_j(1 - Z[\varphi(Z)/\Phi(Z)]). \tag{3-12}$$

Note that the EM method is an iterative substitution method, where at each iteration, all of the nondetects are replaced by the "same" conditional expected value as given by equation (3-11). In the presence of outliers, the conditional expected value given by (3-11) gets distorted (e.g., becomes negative as can be seen in Example 2 of Section 4, or even become greater than DL), and may result in inadequate estimates given by (3-9) and (3-10). Typically, contaminant concentrations are nonnegative and substituting a negative value for the nondetects will be inappropriate, resulting in biased estimates. In these cases, the NDs may be replaced by zero, or half of the detection limit, DL/2. In this report, and in the examples and the simulation results discussed in Sections 7 and 8, whenever the conditional expected value became negative, it was replaced by DL/2, and whenever the EM fill-in value became greater than DL, it was replaced by DL. This modified method is called the "EM Check" method in the examples and the simulation results, as summarized in Appendix C. As shown in Examples 2 and 3, the robust EM estimation procedure takes care of this problem by assigning reduced weights (see Section 3.3) to the outlying observations. The robust EM estimates at the $(j+1)^{\text{th}}$ iteration are given as follows:

$$\hat{\mu}_{j+1} = [\sum_{i=k+1}^{n} w_i x_i + \sum_{i=1}^{k} E_j(x_i \mid x_i \le L)]/(w\text{sum}1 + k), \text{ and} \tag{3-13}$$

$$\hat{\sigma}^2_{j+1} = [\sum_{i=k+1}^{n} w_i^2 (x_i - \hat{\mu}_j)^2 + \sum_{i=1}^{k} E_j((x_i - \mu_j)^2 \mid x_i \le L)]/(w\text{sum}2 + k - 1) \tag{3-14}$$

3.6 Restricted Maximum Likelihood (RMLE) Method

Perrson and Rootzen (1977) obtained the restricted likelihood estimates by simplifying the ML equations. The likelihood function can be written as follows:

$$L(x,\mu,\sigma) = [\Phi(Z)]^k (2\Pi\sigma^2)^{-(n-k)/2} \exp\left(-[\sum_{i=k+1}^{n} (y_i + Z\sigma)^2 / 2\sigma^2]\right), \tag{3-15}$$

where $y_i = x_i - L$; $i := k+1, k+2, \ldots, n$. The random variable, $(n - k)$, representing the number of detected values above DL (= L), can be expressed as a binomial random variable with pdf given as follows:

$$P \text{ ((No. of observation lying above } L) = r) = \frac{n!}{r!(n-r)!}[1 - \Phi(Z)]^r [\Phi(Z)]^{n-r} \tag{3-16}$$

where $r = 0, 1, 2, \ldots, n$. An estimate of $\Phi(Z)$, the probability that an observation lies below L, is k/n. Thus, for $0 < k < n$, an estimate, $\lambda_{k/n}$, of Z is given by $\hat{Z} = \lambda_{k/n} = \Phi^{-1}(k/n)$. Substituting $\lambda_{k/n}$ for Z in (3-15) and maximizing the resulting restricted likelihood function yields the following closed-form estimates of μ and σ (Perrson and Rootzen (1977)).

$$\hat{\sigma}_{RML} = \frac{1}{2}[c + (c^2 + \frac{4}{(n-k)}\sum_{i=k+1}^{n} y_i^2)^{1/2}], \text{ and } \hat{\mu}_{RML} = L - \lambda_{k/n}\hat{\sigma}_{RML}, \quad (3\text{-}17)$$

where $c = \lambda_{k/n}\sum_{i=k+1}^{n} y_i/(n-k)$. The estimates given by (3-17) are biased, which can be corrected as follows. For left-censored samples, $E[\bar{x}_o] = \mu + \sigma\alpha$ and $E[s_o^2] = \sigma^2[1 + (\alpha Z - \alpha^2)]$, where $\alpha = \varphi(Z)/(1 - \Phi(Z))$, and the bias-corrected RMLEs are given as follows:

$$\hat{\mu}_{BRML} = \bar{x}_o - \hat{\alpha}\hat{\sigma}_{BRML} \text{ and } \hat{\sigma}_{BRML} = [s_o^2 - (\hat{\alpha}\lambda_{k/n} - \hat{\alpha}^2)\hat{\sigma}_{RML}^2]^{1/2}, \quad (3\text{-}18)$$

where $\hat{\alpha} = \varphi(\lambda_{k/n})/(1 - k/n)$.

The robust RMLEs are obtained by assigning reduced weights (see Section 3.3) to each of the outlying observations (if any) in the right tail of the data set. The bias-corrected robust RMLEs (Singh and Nocerino (2002)) are given by:

$$\hat{\mu}^*_{BRML} = (\sum_{i=k+1}^{n} w_i x_i / w\text{sum1}) - \hat{\alpha}\hat{\sigma}_{BRML} \text{ and} \quad (3\text{-}19)$$

$$\hat{\sigma}^*_{BRML} = [(\sum_{i=k+1}^{n} w_i^2 x_i^2 / w\text{sum2}) - (\sum_{i=k+1}^{n} w_i x_i / w\text{sum1})^2 - (\hat{\alpha}\lambda_{k/n} - \hat{\alpha}^2)\hat{\sigma}_{RML}^2]^{1/2} \quad (3\text{-}20)$$

3.7 Regression Methods to Estimate Mean, Standard Deviation, and UCL95 of the Mean Based on Left-Censored Data Sets

Several, regression on order statistics (ROS) methods have been cited and used in the environmental literature when dealing with left-censored data sets. The regression methods are parametric in nature as they involve extrapolation of the nondetects based upon certain distributional assumptions about the detected observations. In this process, in addition to nondetects, the detected data are also assumed to follow a certain distribution (normal, lognormal, or gamma). The slope and intercept of the regression line are computed based upon the quantiles obtained using the assumed distribution for the detected observations. Three distributions, normal, lognormal, and gamma have been considered in this report.

It should be noted that most ROS methods as described in this report first extrapolate the nondetects based upon the statistics (e.g., regression line, slope, and intercept) obtained using the detected observations. Obviously, in order to be able to compute reliable estimates of 1) nondetects, and 2) the resulting UCL95 with adequate coverage for the mean, enough detected observations should be available. If the use of general Data Quality Objectives (DQOs) (e.g., USEPA (2000)) is not possible (e.g., when the data might have been collected without using a statistical sampling design), every effort should be made to obtain a representative sample with about 10 detected observations. For accurate and reliable results, whenever possible, more (larger than 10) detected observations should be used when the percentage of

NDs becomes greater than 40%, 50%, and so on. It should be noted that, the use of a minimum of 10 to 15 detected observations is desirable to compute a UCL95 or any other statistics based upon resampling bootstrap methods.

It is assumed that the (*n*-*k*) detected observations come from a well-known parametric distribution, such as a normal, a lognormal, or a gamma distribution. However, it is not easy to verify the distribution of left-censored data sets, especially when a large percentage of observations are being censored (nondetected). An ad hoc simple method to verify the data distribution is based upon the quantile-quantile (Q-Q) plot of the (*n*-*k*) detected observations supplemented with the available goodness-of-fit test (e.g., as in ProUCL 3.0) statistics computed using the detected observations. This ad hoc goodness of fit procedure will be available in ProUCL V 4.0 that is currently under development, and hopefully will be available for public release by the end of 2006 or early 2007.

For a Q-Q plot of a left-censored data set, the *k* nondetects are estimated (extrapolated) by fitting a regression line to the detected raw (normal, gamma) or log-transformed data. The linear regression fit is obtained by using an ordinary least squares (OLS) method to the (*n*-*k*) pairs, $(q_{(i)}, x_{(i)})$; $i := k + 1, k + 2, \ldots, n$, where $x_{(i)}$ are the ordered detected raw or log-transformed values arranged in ascending order. The *n* quantiles, $q_{(i)}$, are computed based upon the distributional (normal, gamma) assumptions. Any regression method based upon this procedure is called the ROS method in the environmental literature. The available ROS methods are:

1) Use of regression on only the (*n*-*k*) detected data (Newman, Dixon, and Pinder (1989). This method is available in UNCENSOR 5.1 (2003)) to estimate the mean and *sd*.

2) ROS on the raw detected data with extrapolated NDs obtained in the original raw scale using a normal distribution or a gamma distribution.

3) Fully parametric ROS: ROS on log-transformed data with extrapolated NDs obtained in a log scale, the mean and *sd* computed using $n = k + (n\text{-}k)$ data points in log scale, and then back-transforming the mean and *sd* in the original units assuming a lognormal distribution. Note that the estimates thus obtained often suffer from a significant amount of transformation bias. This ROS method is often incorrectly called Helsel's ROS method or robust ROS method on log-transformed data. For example, the program UNCENSOR 5.1 (2003) also incorrectly calls this method as Helsel's robust ROS method. This should be corrected in UNCENSOR 5.1 to avoid confusion. Even though both ROS methods operate upon log-transformed data set, there are some differences between the fully parametric ROS method and Helsel's robust ROS method (#4 below). These differences are further illustrated by examples discussed in Section 6.

4) Robust ROS on log-transformed data (known as Helsel's robust ROS): Helsel's robust ROS method on log-transformed data sets is a well documented and recommended method in many state and EPA guidance documents (e.g., Colorado WQCD (2003), California's Ocean Plan (2005), Helsel (2005), USEPA (1993 – SW-846)). In robust ROS on log-transformed data, the *k* NDs are extrapolated in log scale; then those *k* NDs are transformed (via exponentiation) back to the original scale. Finally, the sample mean and the standard deviation are computed directly in the raw scale using the full data set thus obtained (Gilliom and Helsel (1986)). The estimates of the mean and the standard deviation thus obtained are called the robust estimates, and this method is called the robust ROS method on log-transformed data. It is noted that the estimates in the original scale obtained using this process do not suffer from back-transformation bias. But the resulting statistics may represent biased estimates of the mean and *sd* as many times the

extrapolated NDs become greater than the detection limit, DL (L) and the detected observations. This can be seen in Examples 1-3 in Section 4.

Another ROS method is known as the robust ROS MLE method (Kroll, C.N. and J.R. Stedinger (1996)). This use of this method has been suggested in the literature (Helsel (2005)) to compute summary statistics. In this hybrid method, MLEs are computed using log-transformed data. Using the regression model as given by equation (3-21) below, the MLEs of the mean (used as intercept) and *sd* (used as slope) in the log scale are used to extrapolate the NDs in the log scale. Just like in the robust ROS method, all of the NDs are transformed back in the original scale by exponentiation. This results in a full data set in the original scale. One may then compute the mean and *sd* using the full data set. The estimates thus obtained are called robust ROS ML estimates (Kroll and Stedinger (1996)). However, the performance of such a hybrid estimation method is not well known. Moreover, for higher censoring levels, the MLE methods sometimes behave in an unstable manner.

It is also observed that in the environmental literature (e.g., Colorado WQCD (2003), California's Ocean Plan (2005), RPcalc 2.0 (2005), Shumway, Azari, and Kayhanian (2002), and USEPA (1993-SW-846)) any estimation, or UCL, UTL computation method based upon robust ROS estimates is typically called Helsel's robust method or robust ROS method. It should be noted that prior to this report, not many studies have evaluated the performances of the UCL95 methods based upon a robust ROS method covering a wide range of skewed distributions. In this report, we use the same terminology, and any UCL computation method such as bootstrap method based upon the above robust ROS estimates of the mean and *sd* will be called Helsel's robust ROS method.

As suggested by Helsel (November 2005 review of this report), it is advised that the users make a note of the differences between the two ROS methods as described above. In order to avoid confusion about the appropriate definition and use of the robust ROS method on log-transformed data, some researchers (e.g., Helsel (2005)) have suggested avoiding the use of the fully parametric version of ROS on log-transformed data. Following Helsel's recommendation supplemented with our examples (Section 6) and simulation results of Section 8, the fully parametric ROS on log-transformed data will not be available in ProUCL 4.0. A detailed description of the various regression methods is given as follows.

3.7.1 ROS on Detected Raw Data – Assumes a Normal Distribution

The ordinary least squares (OLS) regression is obtained by fitting a linear model to the detected data (perhaps after a suitable transformation) and the hypothetical normal quantiles. In other words, it is assumed that the k censored observations, $x_1, x_2, ..., x_k$, follow the zero-to-detection limit (DL) portion of a normal (or transformed such as log-transformed) distribution. A least squares regression line is obtained using the (n-k) pairs, $(q_{(i)}, x_{(i)})$; $i := k + 1, k + 2, \ldots, n$, where $x_{(i)}$ are the k detected values arranged in ascending order. Then n quantiles, $q_{(i)}$, are obtained using an appropriate normal probability statement, such as $P(Z \le q_{(i)}) = (i - 3/8) / (n + 1/4)$, $i := 1, 2, \ldots, n$ (Johnson and Wichern (1988)). The OLS regression line fitted to the last (n-k) pairs $(q_{(i)}, x_{(i)})$; $i := k + 1, k + 2, \ldots, n$, corresponding to the detected values is given by:

$$x_{(i)} = \mathrm{a} + \mathrm{b}q_{(i)};\ i := k + 1, k + 2, \ldots, n. \quad (3\text{-}21)$$

For full data sets, Barnett (1976) used the intercept and the slope of the regression line to estimate the population mean and the standard deviation. Newman, Dixon, and Pinder (1989) followed a similar approach, and used the intercept and the slope of the OLS line given by (3-21) to estimate population mean, μ, and the standard deviation, σ, from left-censored data sets. Singh and Nocerino (2002) noted that the use of this method using only (*n-k*) detected values results in a biased estimate of the mean and the standard deviation. This method has been incorporated in the UNCENSOR 5.1. A 95% UCL of the mean for this method may be obtained (not a recommended method) using the Student's t-statistic with (*n* – *k*-1) degrees of freedom (*df*).

Note: This method assumes normality of the data set, completely ignores the *k* ND values, and yields biased estimates. This method has been illustrated by Example 1 of Section 4. It is also noted that this method does not perform well (in terms of the coverage for the population mean); therefore, the graphical displays of the simulation results for this method have not been included in Appendices A, B, D, and E.

3.7.2 ROS on Raw Data, Extrapolate k NDs – Assumes a Normal Distribution

This ROS method uses the *k* extrapolated values of nondetects obtained using the model given by (3-21). This approach estimates the population mean, and the standard deviation using the (*n-k*) detected values and the *k* extrapolated nondetects obtained assuming a normal distribution. However, the following should be noted:

- There is no guarantee that the *k* extrapolated NDs will lie below the detection limit, DL, contrary to the statement given in Shumway, Azari, and Kayhanian (2002) on page 3346. The extrapolated NDs often exceed the DL and the detected observations as can be seen in Example 1.

- For ROS on raw data, the extrapolated NDs sometime result in infeasible negative values. This is especially true when a few outliers may be present in the data set. This is illustrated by Examples 2 and 3 of Section 4.

Using the full data set thus obtained, one can compute the sample mean and the standard deviation. The sample mean based upon infeasible extrapolated NDs will yield a biased estimate of the population mean. A 95% UCL of the mean can be obtained using Student's t-distribution with (*n*-1) *df*, as this method assumes that the data set follows a normal distribution. Alternatively, on the full data set with extrapolated NDs, one can use ProUCL 3.0 to compute a UCL95 of the population mean provided the extrapolated nondetects represent feasible estimates of the nondetect observations.

3.7.3 ROS on Log-Transformed Data – Assumes a Lognormal Distribution

The ROS method on log-transformed detected data represents a parametric method as the quantiles used to estimate the slope, intercept, and nondetects (log scale) are based upon the assumption of a lognormal distribution. The estimates of the mean and *sd* (in the original scale) based upon the regression on log-transformed data can be obtained in several ways. Two of those methods (fully parametric ROS and robust ROS) have been considered in this report. The program RPcalc (2005) computes these estimates in the original scale by using yet another method, as discussed in Section 5 of the report. Let *Org* stand for the data in the original unit and *Ln* stand for the data in the natural log-transformed unit. Using equation (3-21) on the log-transformed detected data, the nondetects in transformed log-units are obtained by extrapolation corresponding to the first *k* normal quantiles. Once the nondetects have been estimated, the sample mean, standard deviation (and the associated UCLs) can be computed using one of the following two methods on full data sets: fully parametric ROS on log-transformed data with extrapolated NDs in log scale; and robust ROS on log-transformed data with NDs obtained in the original units by exponentiating the nondetects.

It is noted that both estimation methods are parametric methods as they both are based upon the assumption of a lognormal distribution of the data set. Due to the similarities between these two ROS methods, there seems to some confusion about their use in the environmental literature. The differences between these methods have been described in Section 3.7.3.3.

3.7.3.1 Fully Parametric ROS on Detected Log-Transformed Data

The mean, $\hat{\mu}_{Ln}$, and *sd*, $\hat{\sigma}_{Ln}$, are computed in log scale using a full data set obtained by combining the (*n-k*) detected log-transformed data values and the *k* extrapolated nondetect (log-transformed) values. Note that some of those extrapolated values can be larger than the DL and the detected values, contrary to the statement made by Shumway, Azari, and Kayhanian (2002). Assuming, lognormality, El-Shaarawi (1989) estimated μ and σ by back-transformation using the following equations as one of the several ways of computing these estimates. Note that these estimates suffer from a significant amount of transformation bias as can be seen in examples discussed in Section 6. The estimates given by (3-22) are neither unbiased nor have the minimum variance (Gilbert (1987)). Therefore, it is recommended to avoid the use of the fully parametric ROS method to compute UCL95 and various other limits.

$$\hat{\mu}_{Org} = \exp(\hat{\mu}_{Ln} + \hat{\sigma}^2_{Ln}/2) \text{ and } \hat{\sigma}^2_{Org} = \hat{\mu}^2_{Org}(\exp(\hat{\sigma}^2_{Ln}) - 1) \quad (3\text{-}22)$$

SimCensor back-transforms the estimates of the mean and *sd* based upon equation (3-22). It is noted that UNCENSOR 5.1 uses some other unverifiable method to back-transform the estimates of the mean and *sd* from log scale to the original scale, and, as mentioned before, the program RPcalc 2.0 (2005) uses yet another method to compute the estimates of the mean and *sd* in the original scale. Then using those estimates in original scale, recomputes estimates in the log scale. This is further illustrated in Section 5. It has not been established which one (if any) of these methods may yield the most accurate estimates.

<u>Note:</u> It is noted that the use of back-transformation equations from the log scale to the original scale as given by (3-22) often results in unrealistically elevated estimates of the population mean and the standard deviation. This is illustrated in Section 6 by using some skewed left-censored data sets. It is, therefore, recommended to avoid the use of the well-documented back-transformation formula given by (3-22) above. It is also noted that, other than (3-22), there does not exist any other well-documented or recommended method available in the literature, which can be used to back-transform estimates of the mean and *sd* from transformed space (e.g., MLE, UMLE estimates in log scale) to original scale.

So far as the computation of a UCL95 is concerned (which is the objective here), one may be tempted to compute a 95% UCL of the population mass based upon Land's H statistic using the sample mean and variance of the log-transformed data obtained using ROS on log-transformed data. However, it is well known, depending upon the data skewness, such a UCL95 based upon Land's H-statistic can be unrealistically large. Therefore, an obvious approach is to use one of the nonparametric methods (e.g., bootstrap or Chebyshev methods) available in ProUCL 3.0 on the full data set obtained after back-transforming the extrapolated NDs to original units by exponentiation. Depending upon the data skewness, ProUCL 3.0 provides several alternative methods and picks the most appropriate method to compute an appropriate UCL95 of the mean. Note that the UCL95 based upon this approach (using ProUCL 3.0 on extrapolated full data set) will be the same as the UCL95 computed using the full data set in the original scale obtained using Helsel's robust ROS method described below.

3.7.3.2 Robust ROS on Detected Log-Transformed Data (also Known as Helsel's Robust Method)

In robust ROS on log-transformed data, the NDs are extrapolated in the same manner as in the fully parametric ROS method described in Section 3.7.3.1 above. However, the extrapolated nondetects are back-transformed first in the original units before computing the mean, standard deviation, and other relevant summary statistics. This results in a full data set in the original units. One can, then, compute the sample mean and the standard deviation based upon the full data set ($n = k + (n-k)$) obtained in the original scale. It is noted that, even though the extrapolated NDs cannot become negative, they can exceed the detection limit and the detected observations (Examples 1-3), which in turn results in biased estimates (Singh and Nocerino (2002)) of the population mean or mass. One may use ProUCL 3.0 on the full robust ROS data set to compute a UCL95 of the population mass.

The jackknife UCL95 of the population mean based upon the sample mean (using a full robust ROS data set) is not mentioned here as the use of the jackknife method on the full data set obtained using robust ROS (or any other method such as the EM method, and robust ROS on MLE estimates) is equivalent to Student's t-UCL95 computed based upon the sample mean and the sample standard deviation. This fact is mentioned in ProUCL 3.0 User Guide (2004). It is well known (Singh and Singh (2003), ProUCL 3.0 User Guide (2004)) that, for moderately skewed to highly skewed data sets (with *sd* of log-transformed data > 0.5), the Student's t-UCL95 (and equivalently jackknife UCL95 based upon sample mean) does not provide adequate coverage (much lower than 0.95) to the population mean.

It is noted that the robustness of the ROS method on log-transformed data has been determined based upon limited simulation studies for mildly skewed distributions represented by low values of the standard deviation, σ_y (e.g., < 1.0) of log-transformed variable, *Y* or low values of CV (Gilliom and Helsel (1986), Shumway, Azari, and Kayhanian (2002)). These studies do not cover moderately skewed to highly skewed distributions (with σ_y exceeding 1, 1.5) that are inevitable in many environmental applications. It is desirable that the robustness of the robust ROS method on log-transformed data or the robustness of ROS on MLE (Kroll and Stedinger (1996)) method be evaluated and demonstrated for skewed data sets with σ_y exceeding 1.0.

As suggested by Helsel (November 2005 review of this report), additional simulations were performed to evaluate the performance of a UCL95 based upon a robust ROS estimation method for moderately skewed to highly skewed distributions. The results of our additional simulation experiments as summarized in this report (Appendices D and E) suggest that the 95% UCLs (e.g., obtained using jackknife and bootstrap methods on a robust ROS full data set) based upon a robust ROS method do not provide adequate coverage to the population mass for moderately skewed to highly skewed data sets. Depending upon the skewness and sample size (as observed in earlier simulation results as graphed in Appendices A and B), one should use the BCA bootstrap or Chebyshev UCL method on KM estimates to compute an appropriate UCL95. Detailed recommendations are given in Sections 8 and 9.

Note: Helsel (2005) extended his robust ROS on log-transformed estimation method for data sets with multiple detection limits. His proposed method can be used to estimate the sample mean, sample *sd*, and the nondetect observations in the original scale. To date, not much is known about the performances of such estimates. Helsel's (2005) method to extrapolate NDs when multiple detection limits are present in a data set will be available in ProUCL 4.0. It is noted that, once the NDs have been estimated resulting in a full data set, one may want to use ProUCL 3.0 to compute the most appropriate UCL95 based upon a full data set including the detected values and the extrapolated nondetects.

3.7.3.3 Differences and Similarities between Fully Parametric ROS and Helsel's Robust ROS Methods to Estimate Population Mean and the Standard Deviation

- Both methods are parametric methods, as the slope, intercept, and the NDs are estimated based upon the assumption of a lognormal distribution.

- The main difference between the two ROS estimation methods on log-transformed data is the way the mean and the standard deviation are being computed. The fully parametric ROS method yields estimates that may suffer from transformation bias. More than one method exists to perform back-transformation from log scale to original scale (e.g., given by equation (3-22), incorporated in UNCENSOR 5.1 and in RPcalc 2.0). However, it is not known which one of the back-transformation estimation methods performs the best (in terms of bias and MSE). On the other hand, the robust ROS (Helsel) method computes the mean and the standard deviation in the original units avoiding the transformation bias.

- So far as the computation of a UCL95 is concerned, there should not be any difference between the UCL95s obtained using the fully parametric ROS method or Helsel's ROS robust method, provided they are computed appropriately on the full data set (estimated NDs + detected data) transformed into the original scale. The user should have a good understanding of the objective (computing a UCL95) and how to compute it (e.g., use ProUCL 3.0) appropriately.

3.7.3.4 Influence of Outliers on ROS Methods

Singh and Nocerino (2002) demonstrated that classical MLE methods and the various ROS approaches (on raw or log-transformed data) do not perform well in the presence of outliers. The estimates obtained using the classical methods in the original or log-transformed scale get distorted by outliers. This results in distorted estimates of intercept (population mean) and slope (*sd*), which gives rise to infeasible extrapolated nondetects. For example, the estimated nondetects can become negative (when dealing with raw data), larger than DL, and even larger than some of the observed values (e.g., $x_{(k)}$). The use of such extrapolated NDs results in biased estimates of the population mean and *sd*. Conclusions derived using distorted statistics and UCL95 can be incorrect and misleading. In these situations, subjective checks may be provided to modify the regression method: negative estimates of NDs may be replaced by DL/2, and the estimated nondetects greater than DL may be replaced by DL itself. The mean and variance are computed using the replacement values. Singh and Nocerino (2002) considered this method in their simulation study and concluded that the modified regression method also yields biased estimates of population mean and variance. Therefore, the modified ROS method on raw data has not been included in the simulation experiments as discussed in Section 7.

Kroll, C.N. and J.R. Stedinger (1996) also cautioned the readers about the influence of outliers on MLE estimates, their robust ROS method on MLEs. Some of these outlier-related issues including the distortions of statistics by outliers are illustrated in Section 4.

3.8 ROS on Left-Censored Gamma Distributed Data

Many positively skewed data sets follow a lognormal as well as a gamma distribution. Singh, Singh, and Iaci (2002) noted that gamma distributions are better suited to model positively skewed environmental full data sets. For full data sets, it is observed that the use of a gamma distribution results in reliable, practical, and stable UCL95 values. Also, note that in order to use a gamma distribution, there is no need to transform the data and back-transform the resulting statistics. If a left-censored data set follows a gamma distribution (can be verified using a Q-Q plot and goodness-of-fit tests as incorporated in ProUCL 3.0), then those NDs can be extrapolated using the regression model (3-21) based upon (*n-k*) pairs given by ((*n-k*) higher order gamma quantiles, ordered (*n-k*) detected observations)). However, one has to estimate the gamma parameters before computing the gamma quantiles and extrapolating the NDs. This may have some effect on the adequacy and accuracy of the estimated gamma quantiles, and consequently on the accuracy of the extrapolated NDs. Just like all other distributions, outliers, when present, can distort all statistics including slope, intercept, extrapolated NDs, mean, *sd*, and UCL95. The details of this process can be found in the ProUCL 3.0 User Guide (2004). A brief description of the computation of gamma quantiles is given as follows.

3.8.1 Quantile-Quantile (Q-Q) Plot for a Gamma Distribution

Let $x_1, x_2, ..., x_n$ be a random sample from the gamma distribution, G(k, θ). One should not get confused with k, the shape gamma parameter, which is different from k, the number of NDs as used above. Let $x_{(1)} \le x_{(2)} \le ... \le x_{(n(}$ represent the ordered sample. In order to compute gamma quantiles, one needs to estimate the gamma parameters, k and θ. Let $\hat{k}$ and $\hat{\theta}$ represent the maximum likelihood estimates (*MLEs*) of k and θ, respectively. It should be noted that, initially, these MLE estimates of k and θ are computed based upon the (*n-k*) detected observations. For details of the computation of *MLEs* of k and θ, refer to Singh, Singh, and Iaci (2002). Just like all other ROS methods, in order to be able to compute reliable estimates of the nondetects and the resulting UCL95 with adequate coverage for the mean, enough (about 10 or more) detected observations should be available. The Q-Q plot is obtained by plotting the scatter plot of pairs $(x_{0i}, x_{(i)})$ $i :=$ k+1, 2, …, *n,* where k = number of nondetects. The n quantiles, x_{0i}, are given by the equation, $x_{0i} = z_{0i}\hat{\theta}/2$; $i := 1, 2, \ldots, n$, where the quantiles z_{0i} (already ordered) are obtained by using the inverse chi-square distribution and are given as follows:

$$\int_0^{z_{0i}} f(\chi^2_{2\hat{k}})d\chi^2_{2\hat{k}} = (i-1/2)/n; \quad i := 1, 2, \ldots, n$$

In the above equation, $\chi^2_{2\hat{k}}$ represents a chi-square random variable with $2\hat{k}$ degrees of freedom (*df*). The program, PPCHI2 (Algorithm AS91), as given in Best and Roberts (1975), has been used to compute the inverse chi-square percentage points, z_{0i}, as given by the above equation. This is an informal graphical method to test for a gamma distribution. A linear pattern displayed by the scatter plot of the bulk of the data may suggest an approximate gamma distribution. For example, a high value (e.g., 0.95 or greater) of the correlation coefficient of the linear pattern may suggest an approximate gamma distribution (provided no obvious jumps and breaks of significant magnitude are present in the Q-Q plot) of the data set under study. On this Q-Q plot, points well separated from the bulk of data may represent outliers. Also, obvious breaks and jumps of significant magnitude in the gamma Q-Q plot suggest the presence of multiple populations or outliers.

After fitting a linear regression model (3-21) to the (*n-k*) pairs, (gamma quantiles, detected data), one can extrapolate k NDs for left-censored data set. This will yield a full data set of size $n = k + (n-k)$. A 95% UCL of the mean for the gamma distribution then can be computed using UCL computation methods as described in Singh, Singh, and Iaci (2002), and incorporated in ProUCL 3.0. This method has also been included in our simulation experiments to compare the performances of the various UCL95 methods. Both approximate and adjusted gamma UCL95 methods (as in ProUCL 3.0) have been included in the simulation study.

3.9 EPA Delta Lognormal Method – Assumes Lognormality

The USEPA (1991) proposed the use of the delta lognormal method for the estimation of the population mean and the standard deviation based upon left-censored data sets. This method has been included in the UNCENSOR 5.1 program. However, it is noted that the procedure used to compute a UCL95 as included in UNCENSOR 5.1 may be incorrect. The delta lognormal method has also been included in the simulation experiments as considered in this report. It is noted that, since the delta lognormal method is based upon the lognormal assumption, the resulting UCL95 often becomes unrealistically large of no practical merit. Some examples illustrating these issues are discussed in Section 6.

This method assumes that a certain proportion, $\delta = k/n$, of data values are at or below the detection limit, L (or DL), and that the (n-k) observations (with proportion = 1 - δ) above the detection limit, L, are assumed to follow a lognormal distribution. The delta lognormal distribution models the data as a mixture of two distributions: k nondetects are modeled by a discrete distribution all taking values at L with probability δ, and by a lognormal distribution for (n-k) observations above the L (or DL). Thus, the entire data set is assumed to follow a delta lognormal distribution with a detection limit at L (or DL). The estimates of the mean and *sd* (in the original scale denoted by x's) obtained using such a hybrid distribution are given by:

$$\hat{\mu}_x = \delta L + (1-\delta)\exp(\hat{\mu}_y + 0.5\hat{\sigma}_y^2)$$

$$\hat{\sigma}_x^2 = [(1-\delta)\exp(2\hat{\mu}_y + \hat{\sigma}_y^2)][\exp(\hat{\sigma}_y^2) - (1-\delta)] + [\delta(1-\delta)L][L - 2\exp(\hat{\mu}_y + 0.5\hat{\sigma}_y^2)]$$

where, $y_i = \ln(x_i)$; $k = 1 \leq n$, $k < n$,

$\hat{\mu}_y = \sum y_i \big/ (n-k)$; $k = 1 \leq n$, $k < n$,

$\hat{\sigma}_y^2 = \sum (y_i - \hat{\mu}_y)^2 \big/ (n-k-1)$; $k = 1 \leq n$, $k < n$, and

$\delta = k/n$.

It is noted that no guidance has been provided in the literature on how to compute a 95% UCL of the mean. For the delta lognormal method, the UNCENSOR 5.1 computes a 95% confidence interval (CI) for the population mean based upon some undocumented method. We could not duplicate the UCL97.5 (upper end of a 95% CI) result as reported by UNCENSOR 5.1. Following a similar procedure as described above to compute the sample mean, one may use Land's (1971) H-statistic to compute a UCL95 for such a distribution. A 95% UCL based upon the delta lognormal method is given as follows:

$$\text{UCL} = \delta L + (1 - \delta)\ (\text{H-UCL}),$$

where H-UCL is computed using Land's H-statistic based only upon (n-k) detected values. One can use ProUCL 3.0 to compute an H-UCL based upon (n-k) detected observations. The simulation results as summarized in Section 8 and Appendix C suggest that this method does not perform well as it also yields unrealistically large UCL95 (just like Land's H-UCL on full data set). These issues are also illustrated by some exampled in Section 6.

3.10 Nonparametric Winsorization Method

Some practitioners (e.g., Gilbert (1987)) and USEPA guidance documents (USEPA (2000)) suggest the use of the Winsorization (Dixon and Tukey (1968)) method to compute the mean, the standard deviation, and a UCL95 from left-censored data sets when dealing with "symmetric" distributions. It is noted that this method does not perform well (yields biased estimates) for skewed data sets, as can be seen from the simulation results presented in Appendix C. In this method, values at both ends are replaced by other intermediate values. For a data set of size *n* with *k* (< *n*) nondetects, the procedure is described as follows:

- Replace all of the *k* NDs by the next largest value. Note that all NDs are replaced by the same value.

- Since this method deals with symmetric distributions, the same number of the largest *k* values is replaced by the next smallest value. This will result in (*n*-2*k*) unmodified values in the middle.

- Based upon the modified data of size *n*, compute the sample mean, $\bar{x}_w$ (called the Winsorized mean). Similarly, compute the Winsorized *sd*, s_w, using the following equation.

$$s_w = \frac{s(n-1)}{\nu - 1}$$

This s_w represents an approximate unbiased estimator of the population standard deviation, *σ*. Here, ν = (*n*-2*k*) denotes the number of unmodified values in the middle of the data set of size, *n*. For normally distributed data sets, a 100(1 – α) UCL of the mean, *μ*, is given by the following equation.

$$\text{UCL} = \bar{x}_w + t_{(1-\alpha),(\nu-1)} \frac{s_w}{\sqrt{n}}$$

The simulation study as summarized in this report suggests that this method does not perform well (as expected) when dealing with data sets from asymmetrical distributions. It is observed that for normally distributed data sets, the results obtained using this method are comparable with results obtained using the various other MLE methods. It is noted that, by definition of Winsorization, this method could not be used on data sets with censoring intensity exceeding 50%.

3.11 Nonparametric Kaplan-Meier (KM) Method

The Kaplan-Meier (1958) estimation method, also known as the product limit estimation (PLE) method (Bechtel Jacobs Company (2000)), is based upon a statistical distribution function estimate, like the sample distribution function, except that this method adjusts for censoring. The KM method is quite popular in survival analysis (dealing with right-censored data – such as dealing with terminally ill patients) and various medical applications. Helsel (2005) is one of the first few researchers recommending the use of the KM method when dealing with left-censored environmental data sets. Following, Helsel's recommendation and its well-established success in medical applications, the authors of the report included the KM estimation method in their simulation study as summarized in this report. A brief description of the estimation of sample mean, and SE of the sample mean based upon the KM method is given as follows. For details, refer to Kaplan and Meier (1958) and the report prepared by Bechtel Jacobs Company for DOE (2000). It should be pointed out that the KM estimation method has an added advantage over other methods as it can be used on data sets with multiple detection limits.

Let $x_1, x_2, ..., x_n$ (detection limits or actual measurement) represent *n* data values obtained from samples collected from an area of concern (AOC), and let $x'_1 < x'_2, \ldots < x'_n$ denote the n' distinct values at which detects are observed. That is, n' ($\le n$) represents distinct observed values in the collected data set of size n. For $j = 1,\ldots, n'$, let m_j denote the number of detects at x'_j and let n_j denote the number of $x_i \le x'_j$. Also, let $x_{(1)}$ denote the smallest x_i. Then

$$
\begin{aligned}
\tilde{F}(x) &= 1, && x \ge x'_{n'} \\
\tilde{F}(x) &= \prod_{j \text{ such that } x'_j > x} \frac{n_j - m_j}{n_j}, && x'_1 \le x \le x'_{n'} \\
\tilde{F}(x) &= \tilde{F}(x'_1), && x_{(1)} \le x \le x'_1 \\
\tilde{F}(x) &= 0 \text{ or undefined}, && 0 \le x \le x_{(1)}
\end{aligned}
$$

Note that in the last equality statement of $\tilde{F}(x)$ above, $\tilde{F}(x) = 0$ when $x_{(1)}$ is a detect, and is undefined when $x_{(1)}$ is a nondetect. The estimation of the population mean using the KM method is given as follows:

$$\hat{\mu} = \sum_{i=1}^{n'} x'_i[\tilde{F}(x'_i) - \tilde{F}(x'_{i-1})], \text{ where, } x_o = 0.$$

Using the KM method as described above, an estimate of the standard error (SE) of the mean can be obtained by using the following equation:

$$\hat{\sigma}^2_{SE} = \frac{n-k}{n-k-1} \sum_{i=1}^{n'-1} a_i^2 \frac{m_{i+1}}{n_{i+1}(n_{i+1} - m_{i+1})}$$

Here, k = number of observations below the detection limit and

$$a_i = \sum_{j=1}^{i} (x'_{j+1} - x'_j)\tilde{F}(x'_j), \; i: = 1, 2, \ldots, n'\text{-}1.$$

As mentioned before, some researchers, specifically Helsel (2005), have suggested that the KM method perhaps is the most appropriate method to compute the sample mean and SE for left-censored data sets. Helsel (2005) felt that the percentile bootstrap method on the KM estimate of the mean should be appropriate to compute a 95% UCL of the population mean. Following his recommendations, the KM method has been included in the simulation experiments as summarized in this report. The percentile bootstrap approach along with other approaches, including the Chebyshev inequality, jackknife and bias-corrected accelerated (BCA) bootstrap methods, have been included in the simulation study as discussed in Section 7. All our observations and recommendations have been summarized in Sections 8 and 9.

Using the KM estimates of the mean and the SE of the mean, some investigators have mentioned the use of the normal distribution-based z cutoff value (Helsel (2005)) or a Student's t-distribution-based cutoff value (Bechtel (2000)) to compute a 95% UCL of the mean. Specifically, using a t cutoff value, a 95% UCL of the mean based upon the KM estimates is given by the following equation:

$$\text{UCL95} = \hat{\mu} + t_{0.95,(n-1)}\sqrt{\hat{\sigma}^2_{SE}}\ .$$

In the simulation experiments as discussed in Section 7, a 95% UCL of the mean based upon the KM (or PLE) method has been computed using: 1) the normal approximation based upon standard normal critical values, $z_{\alpha,}$ and Student's t-critical value (Appendix D and E); 2) several of the bootstrap methods, including the percentile bootstrap method and the bias-corrected accelerated (BCA) bootstrap method; and 3) the Chebyshev inequality. As expected, it is noted that the approximate KM-UCL95 based upon the normal approximation (using z or t distributions) does not provide adequate coverage to the mean of non-normal skewed populations. It should be pointed out that in the simulation study as considered in this report, only a single detection limit case for the KM method (as most of the other methods considered can handle only the single detection limit case) has been considered. However, in ProUCL 4.0 (under development), the KM method along with the robust ROS method (Helsel (2005)) on log-transformed data will be able to handle data sets with multiple detection limits.

Some examples illustrating the estimation of NDs, mean, and the standard deviation using the methods discussed in this section are considered next. A few examples have been used to illustrate the influence of outliers on the computation of relevant statistics. These examples also demonstrate the differences in the estimates obtained using the two versions of ROS method on log-transformed data sets.

Section 4

Examples Illustrating the Estimation of the Mean and the Standard Deviation for Left-Censored Data Sets

Robust and resistant methods (Singh and Nocerino (2002)) have also been considered to demonstrate the distortions of the various statistics and estimates in the presence of outliers, and why it is important not to accommodate a few outlying observations with low probability of occurrence in the computation of the estimates of population mean or mass and the standard deviation. The outliers are special and require separate investigation. It should be pointed out that, in general, the robustness and resistance of an estimator go hand in hand (e.g., Rousseeuw and Leroy (1987), and Singh and Nocerino (1995)). Typically, it is the presence of a few outliers or multiple populations in a data set that affects the normality of the data set under study. The robust or resistant computations have been performed using the Censor (Singh and Nocerino (2002)) program, as robust or resistant estimation methods based upon influence function (e.g., PROP influence function) approaches are not available in any other software package. In the following examples, the substitution values for the ROS (raw or log-transformed) method are identified or marked by an (*), and the substitution value for the EM method is marked by (**).

4.1 Example 1: Sulfate Data Set without Outliers

This well-behaved left-censored data set is taken from the USEPA RCRA guidance document (1992). The detection limit is set at 1450. The data with 3 nondetects and 21 detected sulfate values, are: <1450, <1450, <1450, 1850, 1760, 1710, 1575, 1475, 1780, 1790, 1780, 1790, 1800, 1800, 1840, 1820, 1860, 1780, 1760, 1800, 1900, 1770, 1790, and 1780. The sample mean and *sd* obtained using 21 detected data are 1771.91 and 92.702, respectively. The estimates of the mean and the standard deviation (raw data) obtained using some of the methods are summarized in the Table 4-1. Note that the ROS method used in Table 4-1 assumes the normality of the raw data set.

Table 4-1. Raw Classical or Robust Results (without Outliers), *n* = 24, *k* = 3, and DL = 1450

Method	DL/2	DL	CMLE	UMLE	RMLE	ROS*	EM**
Mean	1641.04	1731.67	1724.0	1724.94	1725.55	1751.36	1723.66
Sd	364.09	138.92	153.65	159.39	144.37	103.21	157.80
	*(1571.91, 1613.25, 1637.46)		** (1385.97)				

For the fully parametric ROS (FP-ROS) on log-transformed data: Mean = 1751.68, *sd* = 107.146, and for Helsel's robust ROS method: Mean = 1751. 46, and *sd* = 103. It is noted that for this well-behaved mildly skewed data set, the differences between the estimates obtained using the ROS method on raw data set and the two ROS methods on log-transformed data are not that significant. Also note that the substitution by DL/2 method resulted in a biased estimate of the mean with the highest variability. All of the MLE and the EM methods resulted in fairly similar estimates. For this mildly skewed well-behaved data set, any of the MLE methods can be used to estimate the population mean and *sd*.

The ROS method on raw data set resulted in three (3) estimated nondetects (denoted by *) each of which is larger than DL, and some even exceed the detected observations. The sample mean and *sd* thus obtained represent biased estimates of the population mean and *sd* (Singh and Nocerino, 2002). Ideally, the extrapolated nondetects are supposed to be less than the DL. The EM method resulted in the same conditional expected substitution value = 1385.97 for each of the three nondetects (marked by (**)). The use of the slope and intercept obtained using ROS on raw (ignoring the extrapolated nondetects as in UNCENSOR 5.1) detected value yields 1751.36 and 92.15 as biased estimates of the mean and the standard deviation. The ROS on raw data with extrapolated NDs, and the two ROS methods, Helsel and FP-ROS on log-transformed data, yield similar (but biased) results with lower standard deviations than those obtained using the MLE methods. It is noted that the ROS methods on the raw or log-transformed data resulted in extrapolated NDs higher than the detection limit and also higher than some of the detected observations. This is the reason that the ROS methods yield higher mean and lower variance.

4.2 Example 2: Sulfate Data Set with Outliers in Original-Scale

In order to illustrate how the presence of outliers distorts the various estimates, three arbitrarily chosen outliers, 7000, 8000, and 11,000, are added to the data set of Example 1. The relevant classical (traditional) and robust statistics (Singh and Nocerino (2002)) for the raw data set are summarized in Table 4-2. The classical sample mean and *sd* based upon the detected 24 data values with outliers are 2633.75 and 2410.35. Using equation (3-21), the intercept and slope for the left-censored data with outliers are 2216.51 and 2061.25. The classical ROS on detected raw data resulted in infeasible negative values for the extrapolated nondetects (marked by *). The classical EM method also resulted in a negative distorted value = -254.79 marked by (**) in Table 4-2. It is noted that using the robust (based upon PROP function) EM method, the estimates of nondetects are = 1385.97 (marked by ** in the Robust column), which are identical to the classical EM estimate without the three outliers as given in Table 4-1.

Table 4-2. Raw Results with 3 Outliers, *n* = 27, *k* = 3, and DL = 1450

	Classical		Robust / Resistant (α = 0.01, 0.05)	
Method	Mean	*sd*	Mean	*sd*
MLE	2317.64	2437.60	1729.83	147.41
UMLE	2329.79	2516.74	1730.56	152.19
RMLE	2204.61	2609.06	1731.14	139.17
ROS*	2216.51	2574.10	*Classical estimates = (-1899.83, -995.304, -469.177)	
EM**	2312.80	2491.23 **(-254.79)	1723.66	157.80 **(1385.97)

It should be noted that no robust or resistant estimates of nondetects were computed for the ROS method, as those robust regression procedures (e.g., Rousseeuw and Leroy (1987)) are not well-studied for left-censored data sets.

All classical estimates, including the extrapolated NDs, got distorted by outliers. It is noted that the robust or resistant results for MLE, RMLE, and EM methods are in close agreement with or without the outliers, as can be seen by comparing Tables 4-1 and 4-2. The estimates obtained using a log-transformation are discussed in Example 4.3 below.

Note: In the absence of the availability of suitable robust and resistant methods (which are not easily available in software packages), it is desirable that one preprocesses the data and make sure that one is dealing with a single statistical population without the outliers. If outliers are present in a data set, then those outliers, in consultation with all interested parties and the project team, should be treated and investigated separately. Once again, the objective is to compute a reliable estimate of population mass based upon the majority (instead of trying to accommodate a few outliers by using a log-transformation) of the data representing the dominant population. The project team should decide which of the estimates of the mean (with outliers or without the outliers) is representative of the mass of the population under consideration. The team should decide about the disposition of outliers. It is desirable to compute the relevant statistics with and without the outliers, compare the results, and make a decision depending upon the objectives of the study.

4.3 Example 3: Sulfate Data Set with Outliers in the Log Scale

The computations for Example 2 are repeated here based upon the log-transformation of the data set. In practice, a lognormal distribution is used as a default model. This is especially true when a few outliers may be present or the data set is skewed. The estimates based upon log-transformed data are given below in Table 4-3. The classical mean and *sd* for the detected log-transformed data are 7.675 and 0.537. In the following, all back-transformation results are obtained using equation (3-22). The outliers distorted the estimates of the mean and *sd* for all of the methods, including the fully parametric ROS on log-transformed data and Helsel's robust ROS method. As mentioned earlier, the log-transformation alone cannot produce robust or resistant estimates, especially in the presence of outliers. The methods used have to be robust as well as resistant to outliers. One of the advantages of using the log-transformed data is that the substitution values for the nondetects cannot become negative (as in Example 2) for the ROS and EM methods. Just as in Example 2, all classical estimates based upon log-transformed data got distorted by the three outlying observations as can be seen in Table 4-3. The substitution values are marked by an * for the ROS method, and by ** for the EM method.

Table 4-3. Classical Estimates for Log-Transformed Data with Outliers, n = 27, and k = 3

	Log-Transformed		**Back-Transformed**		
Method	Mean	*sd*	Mean	*sd*	Extrapolated value
MLE	7.59	0.56	2314.35	1393.69	
UMLE	7.59	0.57	2344.59	1456.60	
RMLE	7.58	0.58	2315.67	1476.96	
ROS*	7.58	0.58	2311.29	1461.96	*(6.62, 6.83, 6.95)
EM**	7.59	0.57	2327.96	1438.29	**(6.92)

Note that the ROS method as included in Table 4-3 represents the FP-ROS on log-transformed data (Helsel (2005)) with back-transformed estimates obtained using equation (3-22). For the Helsel robust ROS method: Mean = 2441.60, and *sd* = 2334.066 as given in Table 5-1. It is noted that all ROS estimates (robust or FP-ROS) were distorted by the outliers. Depending upon the objective of the study, the project team should decide which of the mean statistic (with outliers or without outliers) is more representative of the population mass.

Note: It is recommended to avoid the use of a lognormal distribution that tends to hide contamination by accommodating a few outliers or multiple populations. Several more examples have been discussed in

Section 6 supporting the statement: "Avoid the use of a lognormal distribution." It is suggested to use alternative models (e.g., gamma distribution), or nonparametric bootstrap and KM estimation methods to obtain more reliable, stable, and defensible estimates of the parameters of interest such as the population mass.

The PROP robust or resistant estimates (Singh and Nocerino (2002)) with α = 0.05 on the log-transformed data are given in Table 4-4. The robust estimates of the mean and *sd* based upon the detected data are 7.48 and 0.0086, respectively. The results summarized in Table 4-4 are in close agreement with the robust results obtained using the data in the original scale (Table 4-2), and with the estimates obtained using the classical procedure without the outliers (Table 4-1).

Table 4-4. Robust Estimates for Log-Transformed Data with Outliers, *n* = 27, and *k* = 3

	Log-Transformed		**Back-Transformed**		
Method	Mean	*sd*	Mean	*sd*	
MLE	7.45	0.09	1731.10	155.89	
UMLE	7.45	0.09	1732.34	161.09	
RMLE	7.45	0.08	1731.72	146.78	
EM**	7.45	0.10	1725.57	166.57	**(7.24)

Section 5

UCL Computation Methods for Left-Censored Data Sets

As the title of the report suggests, the main objective of the present study is to evaluate and compare the performances of the various UCL95 computation methods for Type 1 left-censored data sets. Several methods and recommendations are available in the literature (as described in Section 3) on how to obtain a point estimate of the population mean based upon left-censored data sets. However, to date, no clear-cut and specific guidance is available on how to compute an appropriate UCL95 of the mean based upon left-censored data sets. Actually, there seems to be some confusion about the computation of an appropriate UCL95 based upon left-censored data sets. For an example, it is noted that on page 78 of Helsel (2005), some recommendations are provided on how to estimate the population mean and the standard deviation. Those recommendations are not for the computation of other relevant statistics, including the upper limits such as UCL95 and UPL95. Several potential UCL computation methods are listed (without specific recommendations) in Chapter 6 of Helsel (2005). Several UCL95 computation methods (e.g., Tiku's method on MLEs, bootstrap and jackknife on KM method, and ROS methods), including some ad hoc methods and methods listed in Helsel (2005), have been considered and evaluated in this report. Recommendations have been made based upon the findings of the simulation results as summarized in Appendices A, B, C, D, and E of this report.

Extensive simulation experiments covering a wide range of skewed distributions have been conducted to evaluate and compare the performances of the various potential UCL95 computation methods. It is noted that the various ROS methods, EPA delta lognormal method, and the MLE methods, CMLE, UMLE, and EM, depend upon distributional assumptions that are often hard to justify or verify. The MLE methods are iterative and sometimes do not converge properly, especially for smaller sample sizes and larger censoring intensities. Therefore, the use of distribution-free nonparametric UCL methods based upon the KM method, Chebyshev inequality, resampling bootstrap, and jackknife methods provide viable alternatives to compute UCL95 based upon left-censor data sets. Based upon our simulation study as summarized in Section 8, it is noted that the nonparametric methods not only perform better (e.g., coverage probabilities) than their parametric counterparts, but also yield meaningful and practical results (estimates of the mean, *sd*, and of EPC terms (UCL95)). Several of these methods will be available in the forthcoming ProUCL 4.0.

Some UCL95 methods based upon EPA delta lognormal method and ROS methods have been discussed earlier. Several other parametric and nonparametric potential UCL95 methods cited in the literature are listed as follows:

- Ad hoc UCL methods obtained using Student's t-statistic as mentioned in Helsel (2005), Millard (2002), EPA-Unified Guidance Document (2004)
- Ad hoc UCL method based upon Land's H-statistic (Helsel (2005))
- UCLs based upon the Chebyshev Inequality (Singh, Singh, and Engelhardt (1997))
- UCLs based upon Tiku's approximation method (Symmetrical censoring, Tiku (1971))
- UCLs based upon Schneider's approximation (for non-symmetric censoring) (Schneider (1986))
- Bootstrap UCL methods (standard, bootstrap t, percentile bootstrap, and BCA bootstrap)

Since the performances (e.g., coverage probabilities) of the UCL computation methods listed above are not well known and well established, most of those methods, new and old, have been included in the simulation study as summarized in Sections 7 and 8.

5.1 Ad hoc UCL95 Computation Method Based Upon Student's t-Distribution

Several documents (e.g., Helsel (2005), Millard (2002), USEPA-UGD (2004), and UNCENSOR 5.1 (2003)) mention the use of Student's t-statistic as one of the potential method to compute a UCL95 for left-censored data sets. Specifically, Cohen's MLE (CMLE), unbiased MLE (UMLE), or EM estimates of the mean and the standard deviation are used to compute Student's t-statistic-based UCL95 of the population mean. One such UCL95 based upon Cohen's MLE, denoted by CMLE(t) method (in the simulation results) is given as follows:

$$\text{UCL95} = \hat{\mu}_{MLE} + t_{0.95,(n-1)}\sqrt{(\hat{\sigma}^2_{MLE}/n)}$$

Similar UCL95 equations can be developed for RMLE, UMLE, EM, and various ROS methods. Some of these UCL95 methods have been included in the simulation experiments as described in Section 7. It is noticed that for normally distributed left-censored data sets with low censoring intensities, such as lower than 20%, the UCL95 based upon CMLE(t) ad hoc method does provide about 95% coverage to the population mean. For normally distributed data sets, the coverage provided by this CMLE(t) UCL method decreases very slowly as the censoring intensity (percentage of NDs) increases. It is also noted that for normally distributed data sets, the MLE (Tiku) and UMLE (Tiku) methods provide about 95% coverage to the population mean for all censoring levels from 10% to 70%. However, as expected, these UCL methods do not provide adequate coverage to the population mean or mass of skewed data distributions.

In UNCENSOR 5.1 output results based upon log-transformed data, it is noted that 95% confidence intervals (CIs) are provided in log scale based upon Student's t-statistic for the CMLE, RMLE, and UMLE methods. It is not clear how one will interpret and use such CIs to derive conclusions about the population means (which is the main objective here) in the original scale. At best, such a CI represents (after transforming the end points by exponentiation) a CI for the population median and not the population mean. The difference between the two parameters can be huge for skewed data sets. Moreover, the back-transformed estimates of the end points in original scale will suffer from an unknown amount of transformation bias.

5.2 (1 – α)100% UCL of the Mean Based Upon the Chebyshev Inequality

The Chebyshev-type inequality (as used in ProUCL 3.0) can also be used to compute a UCL95 of the mean for left-censored data sets. The two-sided Chebyshev inequality (Hogg and Craig (1978)) for a random variable, *X*, with finite mean and the standard deviation, μ_1 and σ_1, is given by:

$$P\left(-k\sigma_1 \le x - \mu_1 \le k\sigma_1\right) \ge 1 - 1/k^2$$

Using this probability statement, an approximate (1 – α)100% UCL (ProUCL 3.0, 2004) of the mean, μ_1, can be obtained by using the following equation:

$$\text{UCL} = \bar{x} + \sqrt{((1/\alpha) - 1)}\, s_x \big/ \sqrt{n}\ .$$

The above UCL equation can be used to compute a UCL95 based upon any of the estimation methods, including the MLE and KM methods listed above. Specifically, in order to compute such UCLs, instead of using the classical sample mean and *sd* (or SE), one uses the mean and *sd* (or SE) obtained using any of the methods such as MLE, EM, or KM described earlier in Section 3. For mildly skewed left-censored data sets, a UCL95 based upon Chebyshev inequality as described here tends to yield conservative estimates (with coverage possibly larger than 95%) of the population mean. Specifically, it is noted that the UCL95 based upon Chebyshev inequality (with KM estimates) yields a reasonable UCL of the mean. However, just like for full-uncensored highly skewed data sets (ProUCL 3.0), a higher confidence coefficient may be needed (such as 97.5% or 99%) to compute a UCL95 based upon Chebyshev inequality for highly skewed left-censored data sets. This topic is discussed further in Section 8.2.5.

Note: It is noted that the Chebyshev UCL computation method as described here is a nonparametric method as it does not require any distributional assumptions about the population under study.

5.3 UCL95 Based Upon Tiku's Method (Symmetrical Type II Censoring)

For symmetrical Type II censoring, Tiku (1971) suggested the use of a Student's t-distribution with ($n - k$-1) degrees of freedom. In practice, this method is also used for Type 1 censoring. The method can be used on any of the MLE methods (e.g., CMLE, RMLE, and UMLE). The UCL95s based upon Tiku's approximation method (using any of the MLEs) have been included in the simulation experiments as described in Section 7. Due to symmetrical censoring, MLE estimates of the mean and the standard deviation are independent (Schneider (1986)), and Student's t-statistic can be used to construct a UCL95.

A (1 – α)100% UCL of the mean, as proposed by Tiku (1971), is given by:

$$\mathrm{UCL}_{Tk} = \hat{\mu}_{MLE} + t_{\alpha,(n-k-1)}\sqrt{Variance(\hat{\mu}_{MLE})}$$

The above equation can also be written as follows:

$$\mathrm{UCL}_{Tk} = \hat{\mu}_{MLE} + t_{\alpha,(n-k-1)}\hat{\sigma}_{MLE}\ Gam_{11}\Big/\sqrt{(n-k-1)}\ .$$

Here Gam_{11} is computed using the Fisher's information matrix, I (Schneider (1986)). This UCL is denoted by Tiku method in Appendices A, B, and C. Tiku's approximate method is simple and performs better (in terms of coverage probabilities) than Schneider's approximation as discussed in Section 5.4.

5.4 UCL95 Based Upon Schneider's Method (Non-symmetrical Type II Censoring)

In this setting, when the censoring is non-symmetrical, the MLEs of the mean and variance are no longer independently distributed. One also has to consider the covariance between the MLE estimates of the mean and variance, which are derived from the Fisher's information matrix, I. The method as described in Schneider (1986) is an approximate method and uses the critical values of a standard normal distribution to compute the approximate UCL95 of the mean. It is noted that this approximate method does not provide adequate coverage to the population mean, especially when the censoring intensity increases or the data distribution is not symmetric. The details can be found in Schneider (1986). The following equations may be used for any of the MLE methods including: CMLE, UMLE, and RMLE. The asymptotic variance and covariance matrix for $\hat{\mu}_{MLE}$ and $\hat{\sigma}_{MLE}$ is given by:

$$\text{ASCV}(\hat{\mu}_{MLE}, \hat{\sigma}_{MLE}) = \text{ASCV} = \frac{\hat{\sigma}^2}{nDET}\begin{pmatrix} J_{22} & -J_{12} \\ -J_{12} & J_{11} \end{pmatrix},$$

Here, *DET* is the determinant of the matrix and is given by the following equation:

$$DET = J_{11}J_{22} - J_{12}J_{12},$$

$J_{11} = 1 + z_1(z_2 + E)$, $J_{12} = z_1(1 + E(z_2 + E))$, $J_{22} = 2 + EJ_{12}$,

where $E = \dfrac{DL - \hat{\mu}_{MLE}}{\hat{\sigma}_{MLE}}$, $z_1 = \dfrac{\varphi(E)}{1 - \Phi(E)}$ and $z_2 = \dfrac{\varphi(-E)}{1 - \Phi(-E)}$, and, as usual, $\varphi(E)$ and $\Phi(E)$ are the

probability density function (pdf) and cumulative distribution function (cdf) of a standard normal distribution (SND). Some other equations used are given as follows:

$$Gam_{11} = J_{22} / DET,\ Gam_{12} = -J_{12} / DET, \text{ and } Gam_{22} = J_{11} / DET.$$

Note that $V(\hat{\mu}) = \dfrac{\hat{\sigma}^2}{v} Gam_{11}$ is the variance of the MLE estimate of the mean (e.g., CMLE, UMLE). This variance (or the SE) has been used in the derivation of UCL95 based upon Tiku's method. The approximate UCL for asymmetrical censoring is given as follows:

$$\text{UCL} = \hat{\mu}_{MLE} + C_1\hat{\sigma}_{MLE},$$

$$\text{where } C_1 = -A + \sqrt{A^2 + B},\ A = \frac{-Gam_{12}z_\alpha^2}{(n - z_\alpha^2 Gam_{22})}, \text{ and } B = \frac{Gam_{11}z_\alpha^2}{(n - z_\alpha^2 Gam_{22})}.$$

As usual, z_α is the upper percentile of a SND used to compute a (1 – α)100% UCL of the mean. All of the equations listed above have been implemented in the SimCensor program to estimate the mean, *sd*, and to compute the UCL95 of the mean for left-censored data sets.

5.5 Jackknife UCL Computation Method for Left-Censored Data Set

Let $x_1, x_2, ..., x_n$ be a random sample (left-censored with *k* NDs, and (*n-k*) detects) of size *n* from a population with an unknown parameter, θ (*e.g.*, $\theta = \mu_1$), and let $\hat{\theta}$ be an estimate of θ, which is a function of all *n* observations. For example, the parameter, θ, could be the population mean, and a reasonable choice for the estimate, $\hat{\theta}$, is the Cohen's MLE, UMLE, RMLE, or KM mean.

In the jackknife approach, *n* estimates of θ are computed by deleting one observation at a time (Dudewicz and Misra (1988)). Specifically, for each index, *i*, denote by $\hat{\theta}_{(i)}$, the estimate of θ (computed similarly as $\hat{\theta}$ such as CMLE) when the i^{th} observation is omitted from the original sample of size *n*, and let the arithmetic mean of these *n* estimates (e.g., of the mean, median, or of percentile) be given by:

$$\tilde{\theta} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)}$$

A quantity known as the i^{th} "pseudo-value" is defined by:

$$J_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}$$

The jackknife estimator of θ is given by the following equation:

$$J(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} J_i$$

If the original estimate $\hat{\theta}$ is biased, then, under certain conditions, part of the bias is removed by the jackknife method, and an estimate of the standard error (SE) of the jackknife estimate, $J(\hat{\theta})$, is given by:

$$\hat{\sigma}_{J(\hat{\theta})} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(J_i - J(\hat{\theta}))^2}$$

Consider the t-type statistic given by:

$$t = \frac{J(\hat{\theta}) - \theta}{\hat{\sigma}_{J(\hat{\theta})}}$$

The t-type statistic given above has an approximate Student's t-distribution with (*n*-1) degrees of freedom. A jackknife (1 – α)100% UCL for θ is given as follows:

$$\text{UCL} = J(\hat{\theta}) + t_{\alpha,n-1}\hat{\sigma}_{J(\hat{\theta})}$$

If the sample size, *n*, is large, then the upper α^{th} *t*-quantile in the above equation can be replaced by the corresponding upper α^{th} standard normal quantile, z_α.

5.5.1 Jackknife UCL Method Based on Sample Mean of a Full Data Set – as Obtained Using Helsel ROS Method or ROS on ML Estimates

This special jackknife UCL needs some discussion and clarification. As noted earlier, the jackknife UCL based upon the sample mean of a full data set (given or obtained after extrapolating *k* NDs) as obtained using a robust ROS method, or ROS on robust MLEs (Kroll and Stedinger (1996)) or EM method (Gleit (1985), Shumway, Azari, and Johnson (1989)) is equivalent to Student's t-UCL of the mean (Dudewicz and Misra (1988) and ProUCL 3.0 User Guide (2004)) based upon that full data set. The jackknife UCL95 (on full data sets), as suggested in Singh, Singh, and Engelhardt (1997), based upon an estimator other than the sample mean, such as the median or the minimum variance unbiased estimator (MVUE) of the mean of a lognormal distribution, will be different from the Student's t-UCL95. This fact has been acknowledged in the jackknife section of the ProUCL 3.0 User Guide (2004).

The EM method (Shumway, Azari and Johnson (1989), and Gleit (1985)) is equivalent to a substitution and fill-in method (such as the substitution by DL/2 method), where the fill-in values (all fill-in values are equal) are based upon the conditional expectation restricted to fall below the detection limit, DL. Similarly, the use of robust ROS and the use of robust ROS based upon MLEs (Kroll and Stedinger (1996)) yield full data sets obtained by extrapolating the nondetected values. The use of the jackknife UCL95 method on a full data set obtained using the ROS or EM methods will simply yield a Student's t-UCL95 based upon that full data set with extrapolated NDs. Therefore, the use of EM method, robust ROS method, or ROS on MLEs followed by the jackknife UCL95 is not needed as those UCLs can be computed simply by using Student's t-statistic on the sample mean and the standard deviation of the full data set obtained using the ROS or EM methods.

It is well known that Student's t-UCL95 or equivalently the jackknife UCL95 based upon the sample mean does not provide adequate coverage (ProUCL 3.0 User Guide (2004) and Singh and Singh (2003)) to population mean of moderately skewed to highly skewed (e.g., when $\sigma_y > 1.0$) data sets. This was the reason that the authors of this report did not include Helsel's robust ROS followed by the jackknife UCL95 method (or bootstrap methods) in the simulation experiments to compute a UCL95 of the population mean. Moreover, once a full data set based upon Helsel's ROS, ROS on robust MLEs, or EM method has been obtained, one can simply use ProUCL 3.0 to compute a UCL95, as most of the available UCL95 methods for full data sets are included in ProUCL 3.0. Depending upon the sample size and data skewness, ProUCL 3.0 computes and recommends the most appropriate UCL95 for the unknown population mean.

Note: It is noted that the jackknife UCL computation method as described here is a nonparametric method as it does not require any distributional assumptions about the population under study.

5.6 Bootstrap on Censored Data Sets

In this report, we compare the performances (coverages) of the four bootstrap methods to compute UCL95 based upon left-censored data sets. The four bootstrap methods included in the study are: standard bootstrap, bootstrap t-method, percentile bootstrap method, and the bias-corrected accelerated (BCA) bootstrap method (Efron and Tibshirani (1993), Manly (1997)). These methods have been used on several estimation methods including the KM method and the various likelihood methods (CMLE, RMLE, and UMLE). For full data sets, the bootstrap procedures (Efron (1982) and Efron and Tibshirani (1993)) have been recommended for the computation of UCL95 for the means of skewed distributions (ProUCL 3.0 User Guide (2004)).

The bootstrap procedures require no assumptions regarding the statistical distribution (e.g., normal, lognormal, gamma) of the underlying population and can be applied to a variety of situations, including the left-censored data sets. These methods are specifically useful when: the exact distributions of the statistics used (e.g., Cohen's MLE, RMLE) are not known; or the critical values of the test statistics are not available; the test statistics and their distributions depend upon the unknown number of nondetects, k, which might be present in a data set.

Let $x_1, x_2, ..., x_n$ be a random sample of size n from a population with an unknown parameter, θ (e.g., θ = mean, μ_1). The sample is left-censored with k observations below the detection limit, DL, and $(n - k)$

observations above the detection limit. Let $\hat{\theta}$ be an estimate of θ, which is a function of k nondetected and $(n - k)$ detected observations. For example, the parameter, θ, could be the population mean, μ, and a reasonable choice for the estimate, $\hat{\theta}$, might be the Cohen's MLE, Helsel's robust ROS mean, or KM estimate of the mean.

The bootstrap procedure on a censored data set is similar to the general bootstrap technique used on full-uncensored data sets. The only difference is that an indicator variable, I (taking only two values: 0 and 1), is used when dealing with left-censored data sets (e.g., see Efron (1981) and Barber and Jennison (1999)). The indicator variable, I, is associated with the detection status of the selected observations, x_i ; i: = 1, 2, ..., n, in a bootstrap sample. The indicator variable, I, takes on a value = 1 when a detected value is selected and $I = 0$ if a nondetected value is selected in a bootstrap sample. In this setting, the detection limit is fixed at DL and the number of nondetects may vary from bootstrap sample to bootstrap sample. There may be k_1 nondetects in the first bootstrap sample, k_2 nondetects in the second sample, ..., and k_N nondetects in the N^{th} bootstrap sample. Since the sampling is conducted with replacement, the number of nondetects, k_i, i: = 1, 2, ..., N, can take any value from 0 to n, inclusive. This is typical of a Type I left-censoring bootstrap process.

A suitable parametric (CMLE, RMLE) or nonparametric (Winsorized, KM method) estimation method is used on each of the N left-censored bootstrap samples. The following two steps are common to the four bootstrap methods considered in this report.

Step 1. Let $(x_{i1}, x_{i2}, ..., x_{in})$ represent the i^{th} sample of size n with replacement from the original left-censored data set $(x_1, x_2, ..., x_n)$. Note that an indicator variable (as mentioned above) is tagged along with each sample taking values 1 (if a detect) and 0 (if a nondetect is selected). Compute an estimate of the population mean (e.g., Cohen's MLE of the mean, KM mean, ROS) using the i^{th} bootstrap sample, i: = 1, 2, ..., N.

Step 2. Repeat Step 1 independently N times (e.g., $N = 2000$), each time calculating a new estimate. Denote these estimates (e.g., CMLE, KM means, and ROS means) by $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N$. The bootstrap estimate of the population mean is given by the arithmetic mean, $\bar{x}_B$, of the N estimates $\bar{x}_i$ (e.g., N CMLEs or N KM means). The bootstrap estimate of the standard error is given by:

$$\hat{\sigma}_B = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\bar{x}_i - \bar{x}_B)^2} \ .$$

In general, a bootstrap estimate of θ may be denoted by $\bar{\theta}_B$ (instead of $\bar{x}_B$). The estimate, $\bar{\theta}_B$, is the arithmetic mean of the N bootstrap estimates (e.g., KM mean, or CMLE mean) given by $\hat{\theta}_i$, *i:=1,2,..., N*. The N bootstrap estimates are computed in the similar way as the original estimate, $\hat{\theta}$ (e.g., KM mean, or CMLE mean), of the parameter, θ. Note that if the estimate, $\hat{\theta}$, represents the KM estimate of, θ, then $\hat{\theta}_i$ (denoted by $\bar{x}_i$ in the above paragraph) also represents the KM mean based upon the i^{th} bootstrap sample. The difference, $\bar{\theta}_B - \hat{\theta}$, provides an estimate of the bias of the estimate, $\hat{\theta}$.

After these two steps, a bootstrap procedure (BCA, Boot-t) is similar to a conventional bootstrap procedure used on a full data set as described in ProUCL 3.0 User Guide (2004). For clarification, those bootstrap UCL computation methods for left-censored data sets are described as follows.

<u>Note:</u> In bootstrap methods, resamples are generated with replacement from the same original left-censored data set. In this process, there is a positive chance that all (or most) values in a bootstrap resample are equal. This is true when one is dealing with small data sets. In order to avoid such situations (with all values in a bootstrap sample to be the same), it desirable to have at least 10 (preferably more) detected observations in a left-censored data set.

5.6.1 UCL of the Mean Based Upon Standard Bootstrap Method

The bootstrap estimate of *SE* of an estimate, $\hat{\theta}$ (e.g., KM mean, CMLE), is given by:

$$\hat{\sigma}_B = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\hat{\theta}_i - \bar{\theta}_B)^2}\ .$$

A (1 – α)100% standard bootstrap UCL for θ (population mean) based upon the estimate, $\hat{\theta}$ (e.g., KM mean), is given as follows:

$$\text{UCL} = \hat{\theta} + z_\alpha \hat{\sigma}_B$$

Here, z_α is the usual upper α^{th} critical value (quantile) of the standard normal distribution. It is observed that the standard bootstrap method does not adequately adjust for skewness and the UCL given by the above equation often fails to provide the specified (1 – α)100% coverage to the population mean of skewed (e.g., lognormal and gamma) distributions.

5.6.2 UCL of the Mean Based Upon Bootstrap t-Method

Another variation of the bootstrap method, called the "bootstrap *t*" is a nonparametric procedure, which estimates the quantiles of the associated t-statistic (pivotal quantity). Here $\hat{\theta}$ (or $\bar{x}$) is an estimate (e.g., KM mean, MLE mean) of the population mean based upon the left-censored data and s_x is the associated sample standard deviation (e.g., CMLE or KM estimate of standard deviation) computed using the original left-censored data set. Let $\bar{x}_i$ and $s_{x,i}$ represent the corresponding estimates (e.g., using CMLE or KM method) of the mean and the standard deviation computed from the i^{th} (with *i:=1,2, ..., N*) bootstrap resample of the original left-censored data obtained following the procedures as described in Steps 1 and 2 above.

The *N* pivotal quantities, $t_i = \sqrt{n}(\bar{x}_i - \bar{x})/s_{x,i}$ are computed and sorted, yielding ordered *N* quantities, $t_{(1)} \le t_{(2)} \le \ldots \le t_{(N)}$. The estimate of the lower α^{th} quantile of the pivotal quantity, *t*, is given by $t_{\alpha,B} = t_{(\alpha N)}$. For example, if *N* = 1000 bootstrap samples are generated, then the 50th ordered value, $t_{(50(}$, would be the bootstrap estimate of the lower 0.05th quantile of the pivotal t-statistic. Then a (1 – α)100% UCL of the mean based upon the bootstrap t-method is given as follows:

$$\text{UCL} = \bar{x} - t_{(\alpha N)}\frac{s_x}{\sqrt{n}}$$

It should be noted that in the above UCL equation, $\bar{x}$ and s_x are not the simple sample mean and the standard deviation. They actually represent the estimates of population mean and the standard deviation for the chosen estimation method (e.g., KM method, CMLE method) for left-censored data sets. Typically, for skewed data sets (e.g., gamma, lognormal), the 95% UCL based upon the bootstrap t-method performs better than the 95% UCLs based upon the simple percentile and the BCA percentile methods. However, it should be pointed out that the bootstrap t-method sometimes results in unstable and erratic UCL values, especially in the presence of outliers (Efron and Tibshirani (1993)) or when the data set may appear to look skewed (perhaps due to the presence of contaminated observations). Therefore, the bootstrap t-method should be used with caution. In case this method results in erratic unstable UCL values (e.g., unusually larger than the Chebyshev UCL); the use of an appropriate Chebyshev (e.g., 95%, 97.5%) inequality-based UCL is recommended.

5.6.3 Percentile Bootstrap Method

Bootstrap resampling of the original left-censored data set is used to generate the bootstrap distribution of the unknown population mean as described in Steps 1 and 2 above. Just as before in Section 5.6.2, here also, $\bar{x}_i$, represents an estimate such as the KM mean, or CMLE of the population mean based upon the i^{th} bootstrap sample. Note that $\bar{x}_i$ is not a simple sample mean, and is actually computed based upon the chosen method (e.g., KM method) from the i^{th} resampling (with i: = *1, 2, ..., N*) of the original left-censor data. These *N*, $\bar{x}_i$, i: = *1, 2, ..., N*, are arranged in ascending order as $\bar{x}_{(1)} \le \bar{x}_{(2)} \le \ldots \le \bar{x}_{(n)}$. The (1 – α)100% UCL of the population mean, μ, is given by the value that exceeds the (1 – α)100% of the generated mean values. The 95% UCL of the mean is the 95th percentile and is given by:

$$95\% \text{ Percentile} - \text{UCL} = 95\text{th}\% \, \bar{x}_i \, ; \, i := 1, 2, \ldots, N$$

For example, when $N = 1000$, a simple 95% percentile-UCL (e.g., based upon KM method) is given by the 950th ordered mean value given by $\bar{x}_{(950)}$.

5.6.4 Bias-Corrected Accelerated (BCA) Percentile Bootstrap Procedure

The BCA bootstrap method is also a percentile bootstrap method, which adjusts for bias in the estimate (Efron and Tibshirani (1993) and Manly (1997)). The performance of this method for skewed distributions (e.g., lognormal and gamma) is not well studied. In this report, we study and compare its performance (in terms of coverage probabilities) with parametric methods and other bootstrap methods. It is observed that, for skewed data sets, this method does represent a slight improvement (in terms of coverage probability) over the simple percentile method. However, for moderately skewed to highly skewed data sets with the *sd* of log-transformed data >1, this improvement is not adequate enough and yields UCLs with a coverage probability much lower than the specified coverage of 0.95. The BCA upper confidence limit of the intended (1 – α) coverage for a selected method (e.g., CMLE, KM, and Chebyshev) is given by the following equation:

$$(1-\alpha)100\% \ UCL_{PROC} = \text{BCA-UCL} = \bar{x}_{PROC}^{\alpha_2},$$

where $\bar{x}_{PROC}^{\alpha_2}$ is the $\alpha_2 100^{th}$ percentile of the distribution of the $\bar{x}_{PROC}$; i: = 1, 2, ..., N and *PROC* is one of the many (e.g., CMLE, KM, DL/2, and Chebyshev) mean estimation methods included in this simulation study. For example, when N = 2000, $\bar{x}_{PROC}^{\alpha_2}$ = $(\alpha_2 N)^{th}$ ordered statistic of $\bar{x}_{PROC}$; i: = 1, 2, ..., N, and is given by the order statistic, $\bar{x}_{PROC,(\alpha_2 N)}$.

Here α_2 is given by the following probability statement:

$$\alpha_2 = \Phi\left[\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1-\hat{\alpha}(\hat{z}_0 + z^{(1-\alpha)})}\right]$$

Here, $\Phi(Z)$ is the standard normal cumulative distribution function and $z^{(1-\alpha)}$ is the $100(1-\alpha)^{th}$ percentile of a standard normal distribution. For example, $z^{(0.95)} = 1.645$, and $\Phi(1.645) = 0.95$. Also, $\hat{z}_0$ (bias correction) and $\hat{\alpha}$ (acceleration factor) are given as follows:

$$\hat{z}_0 = \Phi^{-1}\left[\frac{\#(\bar{x}_{PROC,i} < \bar{x}_{PROC})}{N}\right], \ i: = 1,\ 2,\ ...,\ N$$

$\Phi^{-1}(x)$ is the inverse function of a standard normal cumulative distribution function, e.g., $\Phi^{-1}(0.95) = 1.645$, and $\#(\bar{x}_{PROC,i} < \bar{x}_{PROC})$ represents the number of times, $\bar{x}_{PROC,i}$ is less than $\bar{x}_{PROC}$ with *i: = 1, 2, ..., N.* $\hat{\alpha}$ is the acceleration factor and is given by the following equation:

$$\hat{\alpha} = \frac{\sum(\bar{x}_{PROC} - \bar{x}_{-i,PROC})^3}{6[\sum(\bar{x}_{PROC} - \bar{x}_{-i,PROC})^2]^{1.5}}$$

The summations in the above equation are being carried from $i = 1$ to $i = n$, the sample size. $\bar{x}_{PROC}$ and $\bar{x}_{-i,PROC}$ are respectively the *PROC* mean (e.g., KM mean) based upon all n observations, and the *PROC* mean based upon (n-1) observations without the i^{th} observation, i: = 1, 2, ..., n.

5.7 Bootstrap Methods to Compute UCLs Based Upon Full Data Sets – Obtained Using ROS Methods

In his November 2005 review of this report, Helsel suggested the use of bootstrap methods on the full data set obtained using his robust ROS method on log-transformed data. We want to point out that the bootstrap methods to compute UCL95 on full data sets as obtained using Helsel's robust ROS, ROS on robust MLE, or EM method on left-censored data sets represent conventional bootstrap estimation methods to be used on full data sets. One can use ProUCL 3.0 to compute such bootstrap 95% UCLs.

However, for the sake of direct comparisons (e.g., with KM methods) of the coverage probabilities, the authors considered bootstrap UCL95 methods on robust ROS method. Additional simulations were conducted to perform such an investigation and comparisons. Some of those new simulation results have been graphed in Appendices D and E. Our recommendations based upon the older (Appendices A, B, and C) as well as newer (Appendices D and E) simulation results have been summarized in Sections 8 and 9.

5.8 QA/QC of the Procedures and Algorithms Used in the Report and as Incorporated in the SimCensor Program

Program SimCensor has been developed to study and evaluate the various methods used to estimate the mean and variance and to compute the UCL95 for Type I left-censored data sets. Most of the old or new parametric or nonparametric methods available in the literature have been incorporated into the SimCensor program. Based upon the performances of the methods considered, several of those methods will be available in ProUCL 4.0. It should be noted that different methods used on the same data set often yield different estimates and UCL95 values. In order to decide which of the methods perform better than the other methods, an extensive Monte Carlo simulations study has been conducted as summarized in Sections 7 and 8.

The accuracy of the methods included in the SimCensor program used for estimation of the mean, variance and computation of UCL95 has been verified (whenever available in other software packages) by using the procedures available in the existing software programs. These programs include UNCENSOR 5.1, RPcalc 2.0, SAS, and MINITAB. SAS and MINITAB have been used mainly to compare the computations for the Kaplan-Meier (KM) method. The KM estimates obtained by using SimCensor, SAS, and MINITAB are in complete agreement. Computations for some of the bootstrap methods on KM estimates have also been independently verified by Ms. V. Hennesey and Miss P. Saravanan of UNLV.

SAS and MINITAB deal with right-censored data sets. Therefore, the left-censored data sets from environmental applications need to be flipped (values are subtracted from a selected value larger than all of the values in the data set, which needs to be adjusted back after the calculations) before computing the desired statistics using SAS or MINITAB. In SimCensor (and in ProUCL 4.0), the algorithm for the KM method has been implemented directly for left-censored data sets (Bechtel and Jacob (2000)). Therefore, the user does not have to flip the data before using the KM estimation method.

In addition to estimating the mean and standard error (SE) of the mean, the KM method as incorporated in SimCensor can also compute an estimate of the population variance. It is noted that a direct KM estimate of the variance will be very useful to compute UPLs and UTLs based upon left-censored data sets. Typically, the KM estimate of the population variance is not available in the SAS and MINITAB software packages. As a result, practitioners (e.g., Helsel (2005)) often suggest the use of the “rule-of-thumb” type estimate by using the equation: $sd = SE\sqrt{n}$. Based upon the limiting testing on some data sets, it is observed that the difference between the two standard deviations thus obtained is not very significant. However, this statement cannot be generalized for all data sets of varying skewness and censoring intensity.

5.8.1 Discussion of UNCENSOR 5.1

So far as UNCENSOR 5.1 (2003) is concerned, several discrepancies and errors have been observed. It should be pointed out that earlier versions (e.g., UC3.0) were producing several incorrect values, some of which have been corrected in UNCENSOR 5.1.

- First of all, it is pointed out that in UNCENSOR 5.1, for the computation of the variance of the detected data values, the factor ($n - k$-1) should be replaced by (n-k) as used in the literature (Cohen (1991) and Schneider (1986)).

- There is no guidance given in UNCENSOR 5.1 on how to compute the UCL95 in the raw scale after transforming the data in the log scale and computing the UCLs in the log scale based upon Student's t-statistic.

- For log-transformed data, UNCENSOR 5.1 generates a 97.5% UCL (actually a two-sided 95% confidence interval of the mean) based upon Student's t-statistic for estimates obtained using CMLE, UMLE, and RMLE methods. The program does not back-transform the end points of the CI in log scale to original scale. Moreover, after a back-transformation in the original scale, such an interval will be at best a CI for the population median and not for the population mean. In practice, this can be confusing to a typical user, as these ideas (how to back-transform) and the appropriate use of lognormal distribution are not clear to most practitioners. Furthermore, the back-transformed end points of a CI will suffer from an unknown (not well established) amount of transformation bias.

- Many times, UNCENSOR 5.1 yields a 95% CI interval that does not include the sample mean. This is illustrated by examples discussed in Section 6.

- It is also noted that the back-transformation formula, as given in the UNCENSOR 5.1 manual to compute the estimates of the mean and the standard deviation in the original scale, does not seem to yield the same results as we get by using the SimCensor program (using equation (3-22) of Section 3) or by using other programs such as MINITAB. Some of these discrepancies have been mentioned in Section 6.

- The bias-corrected MLE method and delta lognormal method often result in erratic and unrealistically large estimates as illustrated by examples in Section 6.

- Helsel's robust ROS has been incorrectly labeled. The method used in UNCENSOR 5.1 actually represents the FP-ROS method as discussed in Section 3 of this report.

5.8.2 Discussion of RPcalc 2.0

In Sections 4 and 6, we have used RPcalc 2.0 (2005) to compute estimates of the mean and *sd* based upon fully parametric ROS (variable denoted by Ln(*X*-new) in RPcalc 2.0) and robust ROS (variable denoted by *X*-new in RPcalc 2.0) methods on a log-transformed data set. It is noted that in the program, RPcalc 2.0, variable *X*-new represents the new data set in raw scale obtained after extrapolation of NDs based upon ROS method, and LN(X-new) represents the corresponding log-transformed values.

For the robust ROS method, estimates of the mean and *sd* as computed by SimCensor are in close agreement with the estimates obtained using RPcalc 2.0 as can be seen in columns (6) and (7) of Table 5-1. The minor insignificant differences in the estimates occur due to the fact that SimCensor (as described in ProUCL 3.0 User Guide (2004)) and RPcalc 2.0 (as described in Helsel (2005)) calculate the normal quantiles using slightly different methods.

However, we do not agree with the way RPcalc 2.0 calculates the estimates of the mean and *sd* of the log-transformed variable, Ln(*X*-new), based upon the ROS method on log-transformed data. There are other estimation methods (Gilbert (1987) and El-Shaarawi (1989)) also available to compute estimates of the mean and *sd* in the original-scale for the fully parametric ROS method. The differences in the estimates obtained in the original scale obtained using the two methods (RPcalc 2.0 and equation (3-22)) can be enormously large. It is not known (studied) which of the two methods yields better (in terms of bias) estimates. Also note that the program, UNCENSOR 5.1, uses yet another back-transformation method (we could not duplicate their results) to obtain estimates of the mean and *sd* in the original scale from log scale. These observations give one more reason to avoid the use of a log-transformation on environmental data sets.

For the comparison sake, Table 5-1 with 8 columns summarizes the estimates of the mean and the standard deviation for all of the examples (1 through 3 in Section 4 and 4 through 7 in Section 6) considered in this report. The estimates as summarized in Table 5-1 are obtained using SimCensor and RPcalc 2.0. It should be noted that FP stands for fully parametric. Therefore, fully parametric ROS on raw stands for ROS estimates in the *original scale* based upon the normal distribution assumption (given in column 3), FP-ROS on log stands for ROS estimates based upon the lognormal assumption in the log scale (column 4 of Table 5-1), and FP-ROS on log (back-transformed) means estimates in the *original scale* obtained using equation (3-22), as given in column (5) of Table 5-1.

Table 5-1. Estimates for FP-ROS and Robust (Helsel's) ROS Methods Obtained Using SimCensor and RPcalc 2.0 Programs

		SimCensor				**RPcalc 2.0**	
		FP-ROS on Raw	FP-ROS on log Y	FP-ROS on log (back-transformed)	Helsel's ROS	Helsel's ROS (*X*-new)	Helsel's ROS Y=(Ln(*X*-new))
Example 1 (Sulfate, 24)	Mean	1751.3592	7.4664	1751.6814	1751.4647	1750.1006	7.4656
	Sd	103.2066	0.0611	107.1462	102.9931	104.8811	0.0599
Example 2, 3 (Sulfate, 27)	Mean	2216.5071	7.5773	2311.2887	2441.6013	2437.5324	7.4728
	Sd	2574.1029	0.5801	1461.9600	2334.0660	2336.7274	0.8073
Example 4 (Manly)	Mean	-2.0775	-0.8271	2.0406	3.9853	3.9841	-0.2638
	Sd	24.0534	1.7552	9.3012	20.2760	20.2762	1.8144
Example 5 (Blood Lead)	Mean	-0.1044	-4.7909	0.0489	0.0361	0.0347	-4.1407
	Sd	0.1823	1.8828	0.2835	0.0668	0.0674	1.2495
Example 6 (4,4' DDT,11)	Mean	0.3877	-3.7514	7.5667	1.1650	1.1650	-0.9835
	Sd	4.1191	3.3986	2438.0786	3.4365	3.4365	1.5074
Example 6 (4,4' DDT,10)	Mean	0.0593	-4.2077	0.4213	0.1316	0.1315	-2.8204
	Sd	0.3363	2.5858	11.9180	0.2588	0.2589	1.2585
Example 7 (Aroclor 1254)	Mean	893.3271	3.2019	102019.01	1271.7664	1271.7652	6.1778
	Sd	3479.3181	4.0819	423427416	3105.5848	3105.5853	1.3931

The corresponding estimates of the sample mean and sample standard deviation of the log-transformed variable *Y* = Ln(*X*-new) in the log scale as given in the last column (8) of Table 5-1 have been obtained using RPcalc 2.0. The estimates of the mean and *sd* of *Y* = Ln(*X*-new) as given in columns (4) and (8) are significantly different. The back-transformed estimates of the mean and *sd* in the original scale based upon these two sets of estimates will be significantly different.

It is noted that the results as given in column (4) are commonly used in the literature as estimates of the mean and *sd* of the log-transformed variable, *Y*=Ln(*X*-new). These estimates in column (4) are obtained by using the full data set (in log scale) of size, *n,* with *k* extrapolated NDs (in log scale) and (*n*-*k*) log-transformed detected values. The MVUEs (assuming a lognormal distribution) of the parameters of *Y*=Ln(*X*-new) as given in column (4) are often used to compute various upper limits including UPLs and UTLs for lognormally distributed, variable *X* (here, *X*-new).

Initially, we could not figure out how the estimates in log scale as given in column (8) were computed. It seems like the developers of RPcalc 2.0 have used the following well-known equations representing the relationship between the means and the standard deviations of variables, *X* and *Y*= Ln(*X*).

$$\text{Mean of } X \text{ (raw)} = \mu_1 = \exp(\mu + 0.5\sigma^2)$$
$$\text{Variance of } X \text{ (raw)} = \sigma_1^2 = [\exp(2\mu + \sigma^2)][\exp(\sigma^2) - 1]$$

RPcalc 2.0 calculates μ and σ^2 for variable *Y* = Ln(*X*-new) by solving the above equations for μ and σ^2 given as follows:

$$\mu = \ln(\mu_1) - 0.5\sigma^2$$
$$\sigma^2 = \ln\left((\sigma_1^2 / \mu_1^2) + 1\right)$$

There is no disagreement in computing the "parameters" (mean and variance) of *Y* based upon the "parameters" of *X*, and vice versa. However, the performance of the estimates of the mean and variance of *Y* = Ln(*X*-new) based upon the sample mean and sample variance (which are not even MLEs) of *X* =*X*-new is not known. The estimation method as used in RPcalc 2.0 is an entirely new and different way of estimating the mean and variance of the *Y* = Ln(*X*-new). It is based upon the simple sample mean and sample variance of the *X*-new variable that is assumed to be lognormally distributed. It should be noted that the sample mean and sample variance of *X*-new as used by RPcalc 2.0 to estimate μ and σ are not the MVUE (not even MLE) of μ_1 and σ_1, even when the distributional assumptions are met. In the literature (e.g., Gilbert (1987), El-Shaarawi (1989), Helsel (2005), ProUCL 3.0 User Guide (2004)), the sample arithmetic mean and sample variance of the log-transformed variable, *Y* values are used as estimates of the population mean and variance of *Y* (= Ln(*X*-new) here). The performance of the estimates as given in column (8) of Table 5-1 is not well studied.

As a side note for groundwater monitoring applications, the authors want to point out their disagreement with the methods used to calculate the upper confidence bound (UCB) based upon fully parametric and Helsel's robust ROS estimates as incorporated in RPcalc 2.0 (2005). Specifically, the assumptions (e.g., normal UCB computation) made may not be justifiable. For example, it is noted that RPcalc 2.0 computes a 95%-95% UCB (UTL; that is, an upper 95% tolerance bound to contain at least 95% of the population (95^{th} percentile) with at least 95% confidence) assuming a normal distribution for the lognormally distributed (assumed) variable, *X*-new! That is, a normal distribution is used for an assumed lognormal variable *X*-new. The program was developed for groundwater applications. It is desirable that the developers of RPcalc 2.0 double-check the assumptions made and, if deemed necessary, make the appropriate corrections. The detailed discussion on the topic of the computation of a UCB (95%-95% upper tolerance limit) is beyond the scope of this report; therefore it will not be discussed in the rest of the report.

From the discussion and observations presented in this section, it is recommended to avoid the use of a lognormal distribution, and use nonparametric methods when a raw data set cannot be modeled by a well-known statistical distribution.

Section 6

Examples Illustrating UCL95 Computations for Left-Censored Data

In this section, we discuss the computation of the UCL95 of the population mean or mass using the various methods as described in Section 5. Several examples have been considered to illustrate and address the various statistical issues when computing an appropriate UCL95 based upon left-censored data sets. Specifically, two data sets from the literature, and a few data sets from Superfund sites, have been considered to illustrate the impact of outliers and the use of lognormal model on the various UCL computation methods. The numerical results have been obtained using two programs: 1) UNCENSOR 5.1 (2003), a well cited, and freely available software package, and 2) SimCensor developed to conduct the simulation study to assess and compare the performances of the various UCL95 methods.

6.1 Example 4 (Manly Data Set)

This data set is taken from the book by Manly (2001). The data set has 75 observations with 20 observations below the single detection limit = 0.24. The full data set is: 0.24, 1.33, 0.28, 0.47, 18.40, 168.6, 0.25, 0.25, 0.48, 0.26, 5.56, 0.29, 0.31, 0.33, 3.29, 0.33, 0.34, 0.37, 0.25, 2.59, 0.39, 0.40, 0.28, 0.43, 6.61, 0.48, 0.49, 0.51, 0.51, 0.38, 0.92, 0.60, 0.61, 0.43, 0.75, 0.82, 0.85, 0.94, 1.05, 1.10, 0.54, 1.53, 1.19, 1.22, 0.62, 1.39, 1.39, 1.52, 0.33, 1.73, 2.35, 2.46, 1.10, 51.97, 2.61, and 3.06.

It is noted that there may be at least one potential outlier = 168.6 in the data set. This outlier obviously will distort all estimates and computations. One may (and should) want to compute the estimates without the outlier. In order to save time and some space, no attention is paid to this potential outlier, 168.6. The influence of an outlier on the various estimates is considered in Example 6 in greater detail. Based upon the simulation results as discussed in Section 8 (not paying attention to potential occurrence of outliers), an appropriate UCL95 = 9.72 for this data set is obtained by using the BCA bootstrap method on the KM estimates.

6.1.1 Estimates Obtained Using UNCENSOR 5.1 on Raw Data

Censored Values (Using 75-20=55 Detected Data Points)	
Mean:	5.40982
Variance:	555.62800
Detection Limit:	0.24000
No. Below DL:	20
Sample Size:	75

Cohen's Maximum Likelihood	
Mean:	-1.28911
Variance:	590.26027
Std. Dev.:	24.29527
95% Confidence Interval on Mean:	[-6.46077, 3.30129]

EM Iterative (Gleit)	
Mean:	-0.92193
Variance:	517.19819
Std. Dev.:	22.74199

Bias-Corrected MLE	
Mean:	-1.28795
Variance:	605.36503
Std. Dev.:	24.60417
95% Confidence Interval on Mean:	[-6.52537, 3.36082]

RMLE-One-Step	
Mean:	-4.61163
Variance:	795.67765
Std. Dev.:	28.20776
95% Confidence Interval on Mean:	[-10.61613, 0.71801]

Winsorization	
Mean:	0.59520
Variance:	0.62272
Std. Dev.:	0.78913
95% Confidence Interval on Mean:	[0.41003, 0.78037]

Note:

- It is noted that the sample mean based upon the detected data set is about 5.41, and none of the 95% CI listed above contains this mean of detected observations.
- All MLE and EM methods resulted in a negative estimate of the population mean, even though no sample observation is negative.

6.1.2 Estimates Obtained Using UNCENSOR 5.1 on Log-Transformed Data

Statistics Based Upon 55 Log-Transformed Detected Observations	
Mean:	-0.05513
Variance:	1.69986
Detection Limit:	0.24000
No. Below DL:	20
Sample Size:	75
Back-Transformed Values-Detected data	
Mean:	2.16944
Variance:	17.90691

Cohen's MLE Method	
Mean:	-0.70167
Variance:	2.58691
Std. Dev.:	1.60839
95% Confidence Interval on Mean:	[-1.05387, -0.35935]
Back-Transformed Values	
Mean:	1.74018
Variance:	27.21618
Std. Dev.:	5.21691

EM-Iterative (Gleit)	
Mean:	-0.68436
Variance:	2.34396
Std. Dev.:	1.53100
Back-Transformed Values	
Mean:	1.57627
Variance:	17.94207
Std. Dev.:	4.23581

Bias-Corrected MLE (UMLE)	
Mean:	-0.70160
Variance:	2.65311
Std. Dev.:	1.62884
95% Confidence Interval on Mean:	[-1.05827, -0.35492]
Back-Transformed Values	
Mean:	1.79621
Variance:	30.72427
Std. Dev.:	5.54295

RMLE-One-Step	
Mean:	-0.72948
Variance:	2.78682
Std. Dev.:	1.66938
95% Confidence Interval on Mean:	[-1.09502, -0.37417]
Back-Transformed Values	
Mean:	1.86187
Variance:	37.02247
Std. Dev.:	6.08461

Winsorization	
Mean:	-0.70967
Variance:	1.84646
Std. Dev.:	1.35884
95% Confidence Interval on Mean:	[-1.02852, -0.39082]
Back-Transformed Values	
Mean:	1.21004
Variance:	6.50642
Std. Dev.:	2.55077

Helsel's Robust[1]	
Mean:	-0.82700
Variance:	3.08000
Std. Dev.:	1.75499
Back-Transformed Values	
Mean:	1.94170
Variance:	51.27795
Std. Dev.:	7.16086

EPA Delta Log	
Mean:	1.68760
Variance:	16.84146
Std. Dev.:	4.10384
95% Confidence Interval on Mean:	[0.24000, 40.37513]

[1]This is actually the FP-ROS method on log-transformed data

Note:

- The method labeled as Helsel's method in the UNCENSOR 5.1 is not the robust ROS method as proposed and recommended by Helsel (1990 and 2005). The Helsel robust method in UNCENSOR 5.1 actually is the fully parametric ROS on log-transformed data (Helsel (2005)), where the estimates of the mean and the standard deviation are obtained by using the back-transformation formula (equation (3-22)) from log scale to original scale. Practitioners often tend to forget to differentiate between the two ROS methods used on log-transformed data. These methods are different and often result in significantly different results as illustrated in this section.

- It is noted that the CI listed above is given in the log scale. No attempt has been made to transform the interval end points in the original scale. For a typical user, it is not clear how one will interpret and use such an interval in log scale to make remediation, cleanup, or risk assessment decisions which have to be made using the estimates of population mean (and not the median) in the original raw scale.

- The CI obtained after back-transforming (by simple exponentiation) the end points of the CI obtained using log-transformed data will not be a CI for the mean. At best, it may represent a CI for population median provided the end points are properly transformed.

- To the best of our knowledge, no well-studied and tested method (in terms of stability and coverage probabilities) is available in the literature that can be used to back-transform the end points of a CI in log scale to original scale.

- It is noted that the program, UNCENSOR 5.1, produces a 95% two-sided confidence interval for the population "mean." Therefore, a UCL obtained using UNCENSOR 5.1 represents a 97.5%

UCL (and not a 95% UCL), whereas the UCLs obtained using the SimCensor program represent 95% UCLs of the population mean.

- The 97.5% UCL as given by the EPA delta method seems to be in error obtained using some undocumented method.

- In UNCENSOR 5.1, it is observed that the estimates obtained after back-transformation from log scale to original scale have been obtained by some "unknown" formula. They did not use the formula (Gilbert, (1987) and El-Shaarawi (1989)) as given in the references and help section of UNCENSOR 5.1 (same as equation (3-22) above).

- Often, the bias-corrected MLE method yields incorrect values (both in the raw scale and the log scale) as can be seen by the blood data set discussed in Example 5.

- It is well known that when the percent censoring is large (e.g., > 40%), as also recognized in UNCENSOR 5.1, the MLE (e.g., CMLE, UMLE, RMLE, and EM) methods yield unreasonable estimates, such as negative estimates of the mean. This observation suggests that the use of MLE methods for the estimation of the population mean and *sd* should be avoided for data sets with censoring levels exceeding 40%.

6.1.3 Estimates Obtained Using SimCensor on Raw Data

Input File:	**MANLY-EX.TXT**		
Detection Limit:	**0.2400**		
Number of Observations (ND + Detects):	**75**		
Number of Nondetects:	**20**		
Mean of Detected Data:	**5.4098**		
Standard Deviation of Detected Data:	**23.3565**		
Method	**Mean**	**Std. Dev.**	**SE of the mean**
DL / 2	3.9992	20.2732	
Cohen's MLE	-1.2384	24.0810	
Restricted MLE	-4.6141	28.0284	
Unbiased_MLE	-1.1471	24.4442	
EM	-1.2923	24.3088	
EM Check	3.9992	24.2489	
Winsorization	0.5952	0.7891	
Kaplan-Meier	4.0339	-.--	2.3460
EPA Delta Log	1.6876	4.1038	
Regress Detected	-2.0775	16.9433	
FP-ROS (Raw)	-2.0775	24.0534	
Helsel's ROS	3.9853	20.2760	
FP-ROS (Log)	2.0406	9.3012	

UCLs Obtained Using Tiku's Method	
Cohen's MLE UCL:	3.5085
RMLE UCL:	0.9670
Unbiased MLE UCL:	3.6701
EM UCL:	3.5001
EM Check UCL:	8.7625

UCLs Obtained Using Schneider's Approximation Method	
Cohen's MLE UCL:	2.5232
Unbiased MLE UCL:	2.6716
EM UCL:	2.5048
EM Check UCL:	7.8381

Various Ad Hoc Methods Using Student's t with (*n*-1) Degrees of Freedom:	
Cohen's MLE UCL:	3.3933
Unbiased MLE UCL:	3.5545
EM UCL:	3.3833
EM Check UCL:	8.6632
DL/2 UCL:	7.8985

Nonparametric Methods on Raw Data	
Kaplan-Meier UCL (Z-cutoff):	7.8935
Winsorized UCL:	0.7493

UCLs Obtained Using Standard Bootstrap Method	
DL/2 UCL:	7.8695
Cohen's MLE UCL:	1.6652
Unbiased MLE UCL:	1.7321
Kaplan-Meier UCL:	8.1638
EM UCL:	1.4287
EM Check UCL:	7.8726

UCLs Obtained Using Bootstrap t-Method	
DL/2 UCL:	30.9891
Cohen's MLE UCL:	0.7141
Unbiased MLE UCL:	0.8852
Kaplan-Meier UCL:	31.3381
EM UCL:	0.9853
EM Check UCL:	41.8488

Note:

- As pointed out before, the data set has a few outliers; therefore, as expected, the bootstrap t-method resulted in inflated UCL95 values.

UCLs Obtained Using Percentile Bootstrap Method	
DL/2 UCL:	8.4888
Cohen's MLE UCL:	1.7980
Unbiased MLE UCL:	1.8350
Kaplan-Meier UCL:	8.3853
EM UCL:	2.0959
EM Check UCL:	8.3291

UCLs Obtained Using BCA Bootstrap Method	
DL/2 UCL:	8.3155
Cohen's MLE UCL:	4.9115
Unbiased MLE UCL:	4.3995
Kaplan-Meier UCL:	9.7196
EM UCL:	3.5304
EM Check UCL:	9.1065

UCLs Obtained Using Chebyshev Inequality	
Cohen's UCL:	10.8821
Unbiased MLE UCL:	11.1562
Kaplan-Meier UCL:	14.2597
EM UCL:	10.9429
EM Check UCL:	16.2042

Gamma UCLs	
Approximate Estimated Gamma UCL:	7.6252
Adjusted Estimated Gamma UCL:	7.7262

UCLs Obtained Using Jackknife Method	
DL/2 UCL:	7.8985
Cohen's MLE UCL:	0.9171
Unbiased MLE UCL:	0.8324
Kaplan-Meier UCL:	7.9319
EM UCL:	0.9669
EM Check UCL:	7.8985

ROS on Raw Data Using Student's t-statistic	
ROS UCL:	2.5489

UCLs Obtained Using Land's H-Statistic	
Log DL/2 UCL:	2.1232
Delta UCL:	1.8144
FP-ROS UCL on Log-transformed Data:	3.8260

Note:

The sample size, *n* is large = 75, and the *sd* of the detected log-transformed data is moderate = 1.29 (as given below); in such cases, the use of H-UCL results in UCL values which are often smaller than the sample mean value (mean of detected data = 5.41 here)!

6.1.4 Estimates Obtained Using SimCensor on Log-Transformed Data

		Input File:	**MANLY-EX.TXT**	
		Detection Limit:	**0.2400**	
		Log of Detection Limit:	**-1.4271**	
		Number of Observations (ND + Detects):	**75**	
		Number of Nondetects:	**20**	
		Mean of Detected Log Data:	**-0.0551**	
		Standard Deviation of Detected Log Data:	**1.2919**	
Log-Transformed			**Back-Transformed**	
Method	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
DL / 2	-0.23071	1.15168	1.5411	2.5636
Cohen's MLE	-0.69928	1.59772	1.7808	6.1282
Restricted MLE	-0.72970	1.66017	1.9124	7.3421
Unbiased MLE	-0.69322	1.62182	1.8625	6.6837
EM	-0.70294	1.61403	1.8214	6.4481
EM Check	-0.42099	1.60462	2.3784	8.2831
Winsorization	-0.70967	1.35884	1.2381	2.8603
Kaplan-Meier	-0.41010	SE of the mean is 0.14604		
EPA Delta Log	-.--	-.--	1.6876	4.1038
Regress Detected	-0.82713	1.74699	2.0115	9.0307
ROS	-0.82713	1.75521	2.0406	9.3012
Helsel's ROS	-.--	-.--	3.9853	20.2760
FP-ROS (Log-trans.)	-.--	-.--	2.0406	9.3012

UCLs Obtained Using Ad Hoc Methods Based Upon Land's H-Statistic	
Cohen's UCL:	3.0465
RMLE UCL:	3.3893
Unbiased MLE UCL:	3.2294
EM UCL:	3.1444
EM Check UCL:	4.0843
Log DL/2 UCL:	2.1232
Delta UCL:	1.8144
FP-ROS UCL on Log-transformed Data:	3.8260

6.1.5 Summary of Results for Example 4

- The main objective of the present study is to evaluate and compare the various UCL95 computation methods; therefore, no attention was paid to the magnitudes of the observations, especially the outlying observations. Based upon the simulation results as given in Section 8, for this left-censored data set (assuming no outliers), the most appropriate UCL95 = 9.72, which is obtained by using the BCA bootstrap method on the KM estimates (BCA (KM)).

6.2 Example 5 (Blood Pb Data Set)

This blood lead data set has been used and discussed by Helsel (2005). The problem and the data set are originated from a study conducted by Golden *et al.* (2003). The data are: 0.02, 0.02, 0.02, 0.02, 0.02, 0.018644068, 0.02, 0.02, 0.02, 0.033962264, 0.02, 0.02, 0.02, 0.02, 0.106060606, 0.174074074, 0.02, 0.023529412, 0.01372549, 0.268965517, 0.049152542, 0.015517241, 0.02, 0.024528302, 0.02, 0.17704918, and 0.014285714. The data set is of size 27 with DL = 0.02. There are some detected data (e.g., 0.018644068) below the DL. However, for the present study, all observations below DL = 0.02 have been considered as nondetects. This yields 19 (> 50%) observations below the DL = 0.02. The results obtained using UNCENSOR 5.1 and SimCensor programs are summarized as follows.

6.2.1 Estimates Obtained Using UNCENSOR 5.1 on Raw Data

Censored Values (8 detected values)	
Mean:	0.10717
Variance:	0.00830
Detection Limit:	0.02000
No. Below DL:	19
Sample Size:	27

Maximum Likelihood	
Mean:	-0.06621
Variance:	0.02341
Std. Dev.:	0.15301
95% Confidence Interval on Mean:	[-0.12390, -0.01181]

EM Iterative (Gleit)	
Mean:	-0.01979
Variance:	0.00928
Std. Dev.:	0.09634

Bias-Corrected MLE	
Mean:	6.19152
Variance:	0.02468
Std. Dev.:	0.15709
95% Confidence Interval on Mean:	[6.13230, 6.24738]

RMLE One-Step	
Mean:	-0.06486
Variance:	0.02432
Std. Dev.:	0.15595
95% Confidence Interval on Mean:	[-0.12366, -0.00941]

Winsorization	
Mean:	0.02245
Variance:	0.00001
Std. Dev.:	-0.00366
95% Confidence Interval on Mean:	[0.03140, 0.01351]

6.2.2 Estimates Obtained Using UNCENSOR 5.1 on Log-Transformed Data

Censored Values (*n-k* detected values)	
Mean:	-2.61116
Variance:	0.94193
Detection Limit:	0.02000
No. Below DL:	19
Sample Size:	27
Back-Transformed Values	
Mean:	0.11478
Variance:	0.01731

Maximum Likelihood	
Mean:	-4.96641
Variance:	4.00578
Std. Dev.:	2.00145
95% Confidence Interval on Mean:	[-5.73162, -4.22137]
Back-Transformed Values	
Mean:	0.04277
Variance:	0.02692
Std. Dev.:	0.16408

EM Iterative (Gleit)	
Mean:	-4.36651
Variance:	1.60086
Std. Dev.:	1.26525
Back-Transformed Values	
Mean:	0.02690
Variance:	0.00202
Std. Dev.:	0.04497

Bias-Corrected MLE	
Mean:	76.88986
Variance:	4.22228
Std. Dev.:	2.05482
95% Confidence Interval on Mean:	[76.10425, 77.65477]
Back-Transformed Values	
Mean:	1.66349658386112544E34
Variance:	4.55321630870951488E69
Std. Dev.:	6.74775244708155648E34

RMLE One-Step	
Mean:	-4.91228
Variance:	3.80884
Std. Dev.:	1.95163
95% Confidence Interval on Mean:	[-5.65844, -4.18578]
Back-Transformed Values	
Mean:	0.04149
Variance:	0.02279
Std. Dev.:	0.15098

Helsel's Robust[2]	
Mean:	-4.79204
Variance:	3.54859
Std. Dev.:	1.88377
Back-Transformed Values	
Mean:	0.04182
Variance:	0.02002
Std. Dev.:	0.14148

EPA Delta Log	
Mean:	0.04893
Variance:	0.00840
Std. Dev.:	0.09167
95% Confidence Interval on Mean:	[0.02000, 0.41122]

[2]**It is actually the FP-ROS on log-transformed data**

Note:

For both raw and log-transformed data, the mean, *sd*, and 95% CI obtained using the bias-corrected MLE (UMLE) method are unreasonable and incorrect.

6.2.3 Estimates Obtained Using SimCensor on Raw Data

Input File:	**BLOOD_PB.TXT**		
Detection Limit:	**0.0200**		
Number of Data (ND + Detects):	**27**		
Number of Nondetects:	**19**		
Mean of Detected Data:	**0.1072**		
Standard Deviation of Detected Data:	**0.0852**		
Method	**Mean**	**Std. Dev.**	**SE of the Mean**
DL / 2	0.0388	0.0654	
Cohen's MLE	-0.0625	0.1485	
Restricted MLE	-0.0649	0.1526	
Unbiased MLE	-0.0459	0.1655	
EM	-0.0737	0.1622	
EM Check	0.0388	0.1415	
Winsorization	0.0235	-0.0000	
Kaplan-Meier	0.0483	-.--	0.0124
EPA Delta Log	0.0489	0.0917	
Regress Detected	-0.1044	0.1875	
ROS	-0.1044	0.1823	
Helsel's ROS	0.0361	0.0668	
FP-ROS (Log-trans.)	0.0489	0.2835	

UCLs Obtained Using Tiku's Method	
Cohen's MLE UCL:	0.0331
RMLE UCL:	0.0334
Unbiased MLE UCL:	0.0560
EM UCL:	0.0315
EM Check UCL:	0.1213

UCLs Obtained Using Schneider's Approximation Method	
Cohen's MLE UCL:	-0.0239
RMLE UCL:	-0.0252
Unbiased MLE UCL:	-0.0039
EM UCL:	-0.0314
EM Check UCL:	0.0753

Various Ad Hoc Methods Using Student's t with (*n*-1) Degrees of Freedom	
Cohen's MLE UCL:	-0.0137
RMLE UCL:	-0.0148
Unbiased MLE UCL:	0.0084
EM UCL:	-0.0205
EM Check UCL:	0.0852
DL/2 UCL:	0.0603

Nonparametric Methods on Raw Data (Z-Cutoff)	
Kaplan-Meier UCL:	0.0686
Winsorized UCL:	N/A[3]

UCLs Obtained Using Standard Bootstrap Method	
DL/2 UCL:	0.0590
Cohen's MLE UCL:	0.0222
RMLE UCL:	0.0185
Unbiased MLE UCL:	0.0235
Kaplan-Meier UCL:	0.0749
EM UCL:	6074.9116[4]
EM Check UCL:	0.0595

UCLs Obtained Using Bootstrap t-Method	
DL/2 UCL:	0.0754
Cohen's MLE UCL:	0.0047
RMLE UCL:	0.0048
Unbiased MLE UCL:	0.0031
Kaplan-Meier UCL:	0.0704
EM UCL:	0.0053
EM Check UCL:	0.0739

[3]Censoring > 50%
[4]Bootstrap resamples may have too many NDs and repetitive values

UCLs Obtained Using Percentile Bootstrap Method	
DL/2 UCL:	0.0624
Cohen's MLE UCL:	0.0085
Unbiased MLE UCL:	0.0179
ROS UCL:	0.0091
Kaplan-Meier UCL:	0.0788
EM Check UCL:	0.0618

UCLs Obtained Using BCA Bootstrap Method	
DL/2 UCL:	0.0647
Cohen's MLE UCL:	0.0341
Unbiased MLE UCL:	0.1619
Kaplan-Meier UCL:	0.1362
EM Check UCL:	0.0639

UCLs Obtained Using Chebyshev Inequality	
Cohen's UCL:	0.0621
RMLE UCL:	0.0632
Unbiased MLE UCL:	0.0929
Kaplan-Meier UCL:	0.1022
EM UCL:	0.0623
EM Check UCL:	0.1575

Gamma UCLs	
Approximate Estimated Gamma UCL:	0.1466
Adjusted Estimated Gamma UCL:	0.1633

UCLs Obtained Using Jackknife Method	
DL/2 UCL:	0.0603
Cohen's MLE UCL:	0.0277
RMLE UCL:	0.0227
Unbiased MLE UCL:	0.0055
Kaplan-Meier UCL:	0.0678
EM UCL:	0.0478
EM Check UCL:	0.0603

ROS on Raw Data Using Student's t-static	
ROS UCL:	-0.0445

Regression on Log Data Using Land's H-Statistic	
Log DL/2 UCL:	0.1751
Delta UCL:	0.0440
Fully Parametric ROS UCL:	0.1957

6.2.4 Estimates Obtained Using SimCensor on Log-Transformed Data

Input File:	**BLOOD_PB.TXT**			
Detection Limit:	**0.0200**			
Log of Detection Limit:	**-3.9120**			
Number of Observations (ND + Detects):	**27**			
Number of Nondetects:	**19**			
Mean of Detected Log Data:	**-2.6112**			
Standard Deviation of Detected Log Data:	**0.9078**			
	Log-Transformed		**Back-Transformed**	
Method	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
DL / 2	-2.15013	0.58867	0.1385	0.0891
Cohen's MLE	-4.93297	1.96075	0.0493	0.3331
Restricted MLE	-4.91240	1.92190	0.0466	0.2919
Unbiased MLE	-4.71441	2.18591	0.0978	1.0614
EM	-5.08704	2.14865	0.0621	0.6217
EM Check	-3.52658	1.94670	0.1956	1.2861
Winsorization	-3.74950	0.00000	0.0235	0.0000
Kaplan-Meier	-3.41222	SE of the mean is 0.14756		
EPA Delta Log	-.--	-.--	0.0489	0.0917
Regress Detected	-4.79085	1.93249	0.0537	0.3436
ROS	-4.79085	1.88280	0.0489	0.2835
Helsel's ROS	-.--	-.--	0.0361	0.0668
FP-ROS (Log-trans.)	-.--	-.--	0.0489	0.2835

Various Ad Hoc UCLs Based on Land's H-Statistic on Log-transformed Data	
Cohen's UCL:	0.2191
RMLE UCL:	0.1967
Unbiased MLE UCL:	0.6025
EM UCL:	0.3619
EM Check UCL:	0.8533
Log DL/2 UCL:	0.1751
Delta UCL:	0.0440
FP-ROS UCL:	0.1957

6.2.5 Which UCL95 to Use?

- The data set is a nonparametric data set as it does not follow any of the well-known distributions as incorporated in ProUCL 3.0.
- The data set seems to be moderately skewed with *sd* of log-transformed data as 0.9 (only detected data) and about 1.9 are based upon MLE methods.
- The censoring intensity is about ~ 67% (= 19/27).
- After consulting the recommendations described in Sections 8 and 9, one may use a 95% UCL based upon KM (BCA) method to estimate the exposure point concentration term. The 95% KM (BCA) UCL = 0.1362.

Since several computational errors and discrepancies have been identified in the UNCENSOR 5.1 program, we only use the SimCensor program on the rest of the examples. It is suggested that the developers of the program, UNCENSOR 5.1, fix all errors as mentioned and discussed in this report. Also, as concluded earlier, the UCL values obtained using the various ad hoc methods (e.g., Student's t on estimates obtained using log-transformed data) are not appropriate for skewed data distributions; those UCL results will not be reported in the rest of this section.

6.3 Example 6 (4,4'–DDT Data Set From a Superfund Site)

This example is considered to demonstrate the influence of a few outliers on the various UCL values and to illustrate how the use of a lognormal distribution accommodates outliers and yield unrealistically large values of no practical merit. These issues need to be addressed as the use of a lognormal distribution is very common on data sets with a few outliers or collected from mixture populations. The data set of 4,4'-DDT concentrations comes from a Superfund site. The data set of size 11 with 2 nondetects is given by: <0.002, <0.002, 0.00215, 0.00425, 0.0185, 0.0215, 0.0305, 0.08, 0.35775, 0.8, and 11.5. All of the data values (more than 90%), except the one value = 11.5, are less than 1.0. Therefore, an estimate of the population mass should also be about < 1.0. The data set has an obvious outlier = 11.5, and using the Shapiro-Wilk goodness-of-fit test, a lognormal model accommodates this outlier leading to the conclusion that the data set may have come from a single lognormal population. The use of a lognormal model resulted in unrealistic estimates and UCL values, as summarized below. For this example, the results have been computed with and without the outlier, 11.5.

6.3.1 Estimates Obtained Using SimCensor on Raw Data with Outlier, 11.5

	Input File:	**LOG-44DDT.TXT**	
	Detection Limit:	**0.0200**	
	Number of Observations (ND + Detects):	**11**	
	Number of Nondetects:	**2**	
	Mean of Detected Data:	**1.4239**	
	Standard Deviation of Detected Data:	**3.5712**	
Method	**Mean**	**Std. Dev.**	**SE of the Mean**
DL / 2	1.1651	3.4365	
Cohen's MLE	0.6675	3.7187	
Restricted MLE	0.3742	4.1177	
Unbiased MLE	0.7450	4.0573	
EM	0.6318	3.9515	
EM Check	1.1652	3.9451	
Winsorization	0.1122	0.2654	
Kaplan-Meier	1.1654	-.--	1.0478
EPA Delta Log	2.1344	91.1078	
Regress Detected	0.3877	3.5134	
FP-ROS (Raw)	0.3877	4.1191	
Helsel's ROS	1.1650	3.4365	
FP-ROS (Log-trans.)	7.5667	2438.0786	

Note:

- The *sd* based upon the EPA delta lognormal method and the fully parametric ROS on log-transform method are really inflated and hence of no practical merit.

UCLs Obtained Using Tiku's Method	
Cohen's MLE UCL:	2.6754
RMLE UCL:	2.5940
Unbiased MLE UCL:	2.9359
EM UCL:	2.7641
EM Check UCL:	3.3071

UCLs Obtained Using Schneider's Approximation Method	
Cohen's MLE UCL:	2.1705
RMLE UCL:	2.0068
Unbiased MLE UCL:	2.3862
EM UCL:	2.2218
EM Check UCL:	2.8028

Ad Hoc UCL Methods Using Student's t with (*n*-1) Degrees of Freedom	
Cohen's MLE UCL:	2.6997
RMLE UCL:	2.6245
Unbiased MLE UCL:	2.9622
EM UCL:	2.7912
EM Check UCL:	3.3211
DL/2 UCL:	3.0431

Nonparametric Methods on Raw Data (Z-cutoff)	
Kaplan-Meier UCL:	2.8892
Winsorized UCL:	0.2677

UCLs Obtained Using Standard Bootstrap Method	
DL/2 UCL:	2.7664
Cohen's MLE UCL:	2.0919
RMLE UCL:	1.8613
Unbiased MLE UCL:	2.2220
ROS UCL:	2.0357
Kaplan-Meier UCL:	2.9908
EM UCL:	2.0636
EM Check UCL:	2.8665

UCLs Obtained Using Bootstrap t-Method	
DL/2 UCL:	37.8245
Cohen's MLE UCL:	19.9861
RMLE UCL:	12.0291
Unbiased MLE UCL:	23.7025
ROS UCL:	10.8885
Kaplan-Meier UCL:	37.3964
EM UCL:	18.4816
EM Check UCL:	37.2462

UCLs Obtained Using Percentile Bootstrap Method	
DL/2 UCL:	3.1507
Cohen's MLE UCL:	2.4434
RMLE UCL:	2.3712
Unbiased MLE UCL:	2.5991
ROS UCL:	2.2603
Kaplan-Meier UCL:	3.2185
EM UCL:	2.8392
EM Check UCL:	3.1776

UCLs Obtained Using BCA Bootstrap Method	
DL/2 UCL:	3.2137
Cohen's MLE UCL:	2.1125
Unbiased MLE UCL:	2.9006
Kaplan-Meier UCL:	4.1978
EM UCL:	2.2068
EM Check UCL:	2.2926

UCLs Obtained Using Chebyshev Inequality	
Cohen's MLE UCL:	5.5549
RMLE UCL:	5.7860
Unbiased MLE UCL:	6.0772
Kaplan-Meier UCL:	5.7327
EM UCL:	5.8251
EM Check UCL:	6.3500

Gamma UCLs	
Approximate Estimated Gamma UCL:	8.2255
Adjusted Estimated Gamma UCL:	11.9472

UCLs Obtained Using Jackknife Method	
DL/2 UCL:	3.0431
Cohen's MLE UCL:	1.7505
Unbiased MLE UCL:	1.8246
Kaplan-Meier UCL:	3.0430
EM UCL:	1.7305
EM Check UCL:	3.0431

ROS on Raw Data Using Student's t-statistic	
ROS UCL:	2.6388

UCLs Based Upon Land's H-Statistic	
Log DL/2 UCL:	133.1102
Delta UCL:	113.2992
Fully Parametric ROS on Log-transformed Data UCL:	68611.6019

6.3.2 Recommended UCL Based Upon Statistics Computed Using the Outlier

- If, for a moment, it is assumed that this data set does not have any outliers, then the most appropriate UCL is given by the 99% UCL obtained using the Chebyshev inequality on KM estimates.

- Consulting the recommendations summarized in Sections 8 and 9, a 99% KM (Chebyshev) UCL is needed as the *sd* of log-transformed data is quite high = 2.55 (using detected data) and the estimated *sd* of log-transformed data is > 3.0 (based upon other MLE methods).

- However, since the data set consists of an outlier, 11.5, the use of an inflated 99% KM (Chebyshev) UCL based upon an inflated estimate of skewness (with *sd* of log data = 2.55) will not be appropriate as an estimate of the population mass with more than 90% of the observations < 1.0.

- This *sd* (and hence the skewness) can be reduced significantly by not including the outlier in the data set and the computations. Therefore, as before, it is recommended to preprocess a data set, and identify all outliers and multiple populations. All outliers need separate investigation. If more than one population is present (perhaps with consultation and agreement of all parties) in the data set, then a separate UCL95 should be computed for each of the identified population.

6.3.3 Estimates Obtained Using SimCensor on Log-Transformed Data with Outlier, 11.5

Input File:	**LOG-44DDT.TXT**			
Detection Limit:	**0.0200**			
Log of Detection Limit:	**-6.2146**			
Number of Observations (ND + Detects):	**11**			
Number of Nondetects:	**2**			
Mean of Detected Log Data:	**-2.6953**			
Standard Deviation of Detected Log Data:	**2.5487**			
	Log-Transformed		**Back-Transformed**	
Method	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
DL / 2	-2.77017	2.42366	1.1816	22.2542
Cohen's MLE	--3.65776	3.14378	3.6107	505.4961
Restricted MLE	-3.67375	3.18558	4.0560	648.1496
Unbiased MLE	-3.59231	3.42999	9.8761	3542.3449
EM	-3.68731	3.35198	6.8934	1897.7518
EM Check	-3.33514	3.34279	9.5066	2537.9067
Winsorization	-3.71063	3.49365	10.9373	4890.2110
Kaplan-Meier	-3.32199	SE of the mean is 0.85108		
EPA Delta Log	-.--	-.--	2.1344	91.1078
Regress Detected	-3.75144	3.58147	14.3251	8738.2348
ROS	-3.75144	3.39859	7.5667	2438.0786
Helsel's ROS	-.--	-.--	1.1650	3.4365
FP-ROS (Log-trans.)	-.--	-.--	7.5667	2438.0786

Note:

- Obviously, the results summarized above cannot be used in the decision-making process. The standard deviations as given above (after back-transformation) are inflated due to the use of a lognormal distribution on a data set with an outlier.
- Compare the differences in the estimates based upon raw data and the log-transformed data after back-transformation.
- **AVOID** the use of a lognormal model as its use often yields unrealistic and meaningless statistics and results.
- Whenever possible, use nonparametric methods to compute UCL95 for left-censored data.

6.4 Example 6 (4,4'–DDT Data Set without Outlier, 11.5)

The various estimates of the population mass (mean) based upon the data set without the outlier, 11.5, are summarized as follows.

6.4.1 Estimates Obtained Using SimCensor on Raw Data without Outlier, 11.5

Detection Limit:	**0.0020**		
Number of Observations (ND + Detects):	**10**		
Number of Nondetects:	**2**		
Mean of Detected Data:	**0.1643**		
Standard Deviation of Detected Data:	**0.2646**		
Method	**Mean**	**Std. Dev.**	**SE of the Mean**
DL / 2	0.1317	0.2588	
Cohen's MLE	0.0922	0.2859	
Restricted MLE	0.0770	0.3098	
Unbiased MLE	0.0996	0.3155	
EM	0.0887	0.3067	
EM Check	0.1317	0.3058	
Winsorization	0.0321	0.0619	
Kaplan-Meier	0.1319	-.--	0.0830
EPA Delta Log	0.2223	1.9236	
Regress Detected	0.0593	0.3299	
FP-ROS (Raw)	0.0593	0.3363	
Helsel's Regress	0.1316	0.2588	
FP-ROS (Log-trans.)	0.4213	11.9180	

UCLs Obtained Using Tiku's Method	
Cohen's MLE UCL:	0.2602
RMLE UCL:	0.2582
Unbiased MLE UCL:	0.2848
EM UCL:	0.2685
EM Check UCL:	0.3129

UCLs Obtained Using Schneider's Approximation Method	
Cohen's MLE UCL:	0.2175
RMLE UCL:	0.2103
Unbiased MLE UCL:	0.2376
EM UCL:	0.2220
EM Check UCL:	0.2693

UCLs Based Upon Ad Hoc Methods Using Student's t with (*n*-1) *df*	
Cohen's MLE UCL:	0.2579
RMLE UCL:	0.2566
Unbiased MLE UCL:	0.2824
EM UCL:	0.2665
EM Check UCL:	0.3090
DL/2 UCL:	0.2817

Nonparametric Methods on Raw Data (Z-cutoff)	
Kaplan-Meier UCL:	0.2684
Winsorized UCL:	0.0716

UCLs Obtained Using Standard Bootstrap Method	
DL/2 UCL:	0.2566
Cohen's MLE UCL:	0.2285
RMLE UCL:	0.2063
Unbiased MLE UCL:	0.2250
Kaplan-Meier UCL:	0.2705
EM UCL:	0.2264
EM Check UCL:	0.2571

UCLs Obtained Using Bootstrap t-Method	
DL/2 UCL:	1.3265
Cohen's MLE UCL:	0.8586
RMLE UCL:	0.7143
Unbiased MLE UCL:	0.9147
Kaplan-Meier UCL:	1.3458
EM UCL:	0.7974
EM Check UCL:	1.2174

UCLs Obtained Using Percentile Bootstrap Method	
DL/2 UCL:	0.2676
Cohen's MLE UCL:	0.2484
RMLE UCL:	0.2341
Unbiased MLE UCL:	0.2385
Kaplan-Meier UCL:	0.2867
EM UCL:	0.2473
EM Check UCL:	0.2801

UCLs Obtained Using BCA Bootstrap Method	
DL/2 UCL:	0.2761
Cohen's MLE UCL:	0.2322
RMLE UCL:	0.2248
Unbiased MLE UCL:	0.2306
Kaplan-Meier UCL:	0.3186
EM UCL:	0.2072
EM Check UCL:	0.2914

UCLs Obtained Using Chebyshev Inequality	
Cohen's MLE UCL:	0.4863
RMLE UCL:	0.5040
Unbiased MLE UCL:	0.5344
Kaplan-Meier UCL:	0.4935
EM UCL:	0.5114
EM Check UCL:	0.5532

Gamma UCLs	
Approximate Estimated Gamma UCL:	0.8630
Adjusted Estimated Gamma UCL:	1.2652

UCLs Obtained Using Jackknife Method	
DL/2 UCL:	0.2817
Cohen's MLE UCL:	0.2268
RMLE UCL:	0.2065
Unbiased MLE UCL:	0.2255
Kaplan-Meier UCL:	0.2816
EM UCL:	0.2286
EM Check UCL:	0.2817

ROS on Raw Data Using Student's t-statistic	
ROS UCL:	0.2542

UCLs Based Upon Land's H-Statistic	
Log DL/2 UCL:	3.4247
Delta UCL:	2.5838
Fully Parametric ROS on Log-transformed UCL:	152.1860

6.4.2 Recommended UCL Based Upon Statistics Computed without the Outlier, 11.5

- This *sd* (and hence the skewness) reduced significantly by not including the outlier in the data set and the computations. The *sd* of the log-transformed data without the outlier is about 1.9 (detected observations).

- The censoring intensity is about 18%. Consulting the recommendations made in Sections 8 and 9, for a sample of size 10, % censoring ~18%, and *sd* of log-transformed data ~ 1.9, an appropriate estimate of the population mass can be computed using a 99% Chebyshev (KM) UCL.

- Note that a 95% Chebyshev (KM) UCL is based upon the full data set with outlier = 5.73, and the corresponding 95% Chebyshev (KM) UCL without the outlier = 0.4935.

- **Recommended Estimate of the Population Mass:** For this data set, the most appropriate estimate of the population mass is given by the 99% Chebyshev (KM) UCL based upon the data set without the outlier, 11.5.

Note: All of the estimation methods (e.g., Chebyshev inequality on KM estimates) to compute 95%, 97.5%, and 99% UCL of the population mean will be available in ProUCL 4.0, which is expected to be released by the fall of 2006 or early 2007 (after comments and reviews).

6.4.3 *Estimates Obtained Using SimCensor on Log-Transformed Data*

Detection Limit:	**0.0020**			
Log of Detection Limit:	**-6.2146**			
Number of Observations (ND + Detects):	**10**			
Number of Nondetects:	**2**			
Mean of Detected Log Data:	**-3.3375**			
Standard Deviation of Detected Log Data:	**1.8963**			
	Log-Transformed		**Back-Transformed**	
Method	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
DL / 2	-3.29142	1.79052	0.1848	0.8993
Cohen's MLE	-4.18937	2.45911	0.3117	6.4022
Restricted MLE	-4.19209	2.46628	0.3164	6.6148
Unbiased MLE	-4.12617	2.71363	0.6413	25.4648
EM	-4.21884	2.64727	0.4893	16.2602
EM Check	-3.91288	2.63793	0.6482	21.0169
Winsorization	-4.27846	2.80997	0.7186	37.2361
Kaplan-Meier	-3.89842	SE of the mean is 0.68749		
EPA Delta Log	-.--	-.--	0.2062	2.1183
Regress Observed	-4.20770	2.73295	0.6230	26.0767
FP-ROS (Log-transform)	-4.20770	2.58580	0.4213	11.9180
Helsel's Regress	-.--	-.--	0.1316	0.2588
FP-ROS (Log-trans.)	-.--	-.--	0.4213	11.9180

Various Ad Hoc Methods Using Land's H-Statistic	
Cohen's MLE UCL:	65.2618
RMLE UCL:	68.2755
Unbiased MLE UCL:	412.8559
EM UCL:	232.5458
EM Check UCL:	295.3939

UCLs Based Upon Land's H-Statistic	
Log DL/2 UCL:	3.4247
Delta UCL:	2.5838
FP-ROS on Log-transformed Data UCL:	152.1860

6.4.4 Comparison of Results with and without Outlier, 11.5

- As mentioned before, the main objective of computing a UCL95 is to estimate the population mean (mass) based upon the majority of the data set representing the dominant population (e.g., AOC, EA, RU). Inclusion of a single outlier (= 11.5) resulted in distorted and unrealistic estimates. This is especially true when the statistics were computed based upon log-transformed data.

- It is also noted that the UCL95 obtained using Land's H-statistic on MLE estimates obtained using the log-transformed data without the outlier still are elevated. Based upon these observations, it is recommended not to use ad hoc UCL computation methods based upon Land's H-statistic and MLE estimates obtained using log-transformed data.

6.5 Example 7 (Aroclor 1254 Data Set From a Superfund Site)

This is another data set of a larger size, n = 53, representing Aroclor 1254 concentrations from a Superfund site. The data values are given by: 0.02, 0.02, 0.02, 0.02, 0.02, 0.1, 0.11, 0.185, 0.2, 0.21, 0.43, 0.6, 0.64, 1, 1.5, 2, 2, 3.15, 3.3, 5.9, 6.6, 16, 20, 21, 21.5, 35, 41, 44, 46, 48, 49, 140, 160, 210, 220, 310, 330, 360, 880, 924, 1300, 1300, 1600, 1600, 1700, 3400, 3900, 4400, 5000, 5400, 6600, 8300, and 19000. The detection limit is 0.02. The various statistics available in SimCensor are given as follows. The data set has some potential outliers (e.g., 8300, 19000). This data set can also be modeled by a lognormal distribution. As expected, methods based upon the lognormal model yield unreasonable and meaningless estimates and UCL values.

6.5.1 Estimates Obtained Using SimCensor on Raw Data

Detection Limit:	**0.0020**
Number of Observations (ND + Detects):	**53**
Number of Nondetects:	**5**
Mean of Detected Data:	**1404.2380**
Standard Deviation of Detected Data:	**3203.4921**

Method	Mean	Std. Dev.	SE of the Mean
DL / 2	1271.7637	3105.5859	
Cohen's MLE	1057.9383	3278.5122	
Restricted MLE	904.1001	3504.8772	
Unbiased MLE	1065.0644	3330.9423	
EM	1055.2397	3313.7903	
EM Check	1271.7637	3312.4700	
Winsorization	851.0175	1865.4543	
Kaplan-Meier	1271.7722	-.--	427.0125
EPA Delta Log	22695.5871	11640702.1138	
Regress Detected	893.3271	2837.4059	
FP-ROS (Raw)	893.3271	3479.3181	
Helsel's ROS	1271.7664	3105.5848	
FP-ROS (Log-trans.)	102019.0074	423427416.1488	

UCLs Obtained Using Tiku's Method	
Cohen's MLE UCL:	1754.7188
RMLE UCL:	1646.0303
Unbiased MLE UCL:	1772.8459
EM UCL:	1759.3141
EM Check UCL:	1978.9716

UCLs Obtained Using Schneider's Approximation Method	
Cohen's MLE UCL:	1685.1526
RMLE UCL:	1568.6211
Unbiased MLE UCL:	1702.0409
EM UCL:	1688.8176
EM Check UCL:	1911.1992

Various Ad Hoc Methods Using Student's t with (*n*-1) Degrees of Freedom	
Cohen's MLE UCL:	1812.1148
RMLE UCL:	1710.3488
Unbiased MLE UCL:	1831.3017
EM UCL:	1817.5315
EM Check UCL:	2033.7517
DL/2 UCL:	1986.1609

Nonparametric Methods on Raw Data (Z-cutoff)	
Kaplan-Meier UCL:	1974.2979
Winsorized UCL:	1282.0006

UCLs Obtained Using Standard Bootstrap Method	
DL/2 UCL:	1963.4956
Cohen's MLE UCL:	1704.5589
RMLE UCL:	1520.4657
Unbiased MLE UCL:	1702.1442
Kaplan-Meier UCL:	2005.0762
EM UCL:	1683.7905
EM Check UCL:	1993.2583

UCLs Obtained Using Bootstrap t-Method	
DL/2 UCL:	2608.4037
Cohen's MLE UCL:	2072.1922
RMLE UCL:	1885.2621
Unbiased MLE UCL:	2168.7201
Kaplan-Meier UCL:	2682.3231
EM UCL:	2166.6755
EM Check UCL:	2683.8457

UCLs Obtained Using Percentile Bootstrap Method	
DL/2 UCL:	2032.2509
Cohen's MLE UCL:	1777.6320
RMLE UCL:	1640.5575
Unbiased MLE UCL:	1782.7380
Kaplan-Meier UCL:	2123.7534
EM UCL:	1761.5033
EM Check UCL:	1984.0694

UCLs Obtained Using BCA Bootstrap Method	
DL/2 UCL:	1988.9692
Cohen's MLE UCL:	1640.4284
RMLE UCL:	1528.8293
Unbiased MLE UCL:	1837.4074
Kaplan-Meier UCL:	2110.6929
EM UCL:	1714.9544
EM Check UCL:	1883.3406

UCLs Obtained Using Chebyshev Inequality	
Cohen's MLE UCL:	3020.9173
RMLE UCL:	3002.6131
Unbiased MLE UCL:	3059.4355
Kaplan-Meier UCL:	3133.0763
EM UCL:	3039.3413
EM Check UCL:	3255.0747

Gamma UCLs	
Approximate Estimated Gamma UCL:	2600.9884
Adjusted Estimated Gamma UCL:	2655.4368

UCLs Obtained Using Jackknife Method	
DL/2 UCL:	1986.1609
Cohen's MLE UCL:	1680.2545
RMLE UCL:	1480.3708
Unbiased MLE UCL:	1682.7124
Kaplan-Meier UCL:	1986.1677
EM UCL:	1679.7116
EM Check UCL:	1986.1609

ROS on Raw Data Using Student's t-statistic	
ROS UCL on Raw Data:	1693.6963

Regression on Log Data Using Land's H-Statistic	
Log DL/2 UCL:	853531.3852
Delta UCL:	153553.3415
FP-ROS (Log-transform) UCL:	4072707.4758

6.5.2 Estimates Obtained Using SimCensor on Log-Transformed Data

Input File:	**LOG-CLOR.TXT**
Detection Limit:	**0.0200**
Log of Detection Limit:	**-3.9120**
Number of Observations (ND + Detects):	**53**
Number of Nondetects:	**5**
Mean of Detected Log Data:	**3.9385**
Standard Deviation of Detected Log Data:	**3.4818**

	Log-Transformed		**Back-Transformed**	
Method	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
DL / 2	3.38238	3.77045	35973.1529	43954757.7851
Cohen's MLE	3.02058	4.39648	322903.0923	5085402560.6239
Restricted MLE	3.03660	4.32404	239254.1758	2747512477.9080
Unbiased MLE	3.03013	4.46679	445185.6543	9574433367.5194
EM	3.01749	4.44579	400321.5539	7840422495.5001
EM Check	3.19786	4.44347	474517.1291	9197989531.6861
Winsorization	3.29619	4.62869	1212964.5656	54472860413.1406
Kaplan-Meier	3.34969	SE of the mean is 0.52507		
EPA Delta Log	-.--	-.--	22695.5871	11640702.1138
Regress Detected	3.20194	4.09042	105629.2730	453926320.8632
ROS	3.20194	4.08191	102019.0074	423427416.1488
Helsel's ROS	-.--	-.--	1271.7664	3105.5848
Regress Transform	-.--	-.--	102019.0074	423427416.1488

Note:

- This data set is known to have outliers (e.g., 19,000) that should be removed before computing any statistics. The use of a lognormal model accommodated the outliers, resulting in elevated, unstable, and impractical UCL values.
- Consulting the recommendations as summarized in Section 9, the most appropriate estimate of population mass (in terms of coverage and practical merit) is given by a UCL based upon Chebyshev inequality and KM estimates. The confidence coefficient to be used depends upon the skewness. For highly skewed data sets, a higher (e.g., 97.5%, 99%) confidence coefficient may have to be used to estimate the population mass.
- The higher is the skewness, the higher is the confidence coefficient.
- A 95% Chebyshev (KM) UCL is given by = 3133.08. Note that this UCL is inflated as it is based upon a data set with potential outliers (e.g., 19,000).

- Outliers have inflated the *sd* of log-transformed data and hence, the skewness. Just as in Example 6, the exclusion of a few outliers (8,300, 19,000) will reduce the skewness.

- It is likely that the data set represents a mixture sample with a few potential outliers (e.g., 19,000, 8,300).

- One may want to preprocess the data set, and recompute the estimate(s) of the population mass accordingly perhaps without the outlier, 19,000 (and also without 8,300).

Section 7

Description of the Simulation Experiments

Singh and Nocerino (2002) and many others (e.g., Gilliom and Helsel (1986), Gleit (1985), and Haas and Scheff (1990)) compared the performances of the various parametric (e.g., MLE and ROS) procedures to estimate the population mean and the standard deviation for Type 1 left-censored data sets. The performance measures used were the bias and mean square error (MSE). Bias of an estimate, $\hat{\theta}$, of a parameter, θ, is defined as its departure ($\hat{\theta}$-θ) from the parameter. For a simulation experiment with N iterations, Bias = $\sum_{i=1}^{N}(\hat{\theta}_i - \theta)/N$ and MSE = $\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2/N$, where $\hat{\theta}_i$ is an estimate (computed in the same manner as $\hat{\theta}$) of θ, obtained from the sample generated at the i^{th} iteration, i: = 1, 2, ..., N. Those studies concluded that the various substitution methods (replacement by 0, DL/2, and DL), regression method (ROS on raw data), and generation of NDs from a uniform distribution, U(0, DL], do not perform well as they yield biased estimates of the population mean and the standard deviation.

In this report, extensive simulation experiments covering a wide range of skewed distributions have been conducted to evaluate and compare the performances of the various parametric and nonparametric UCL95 computation methods. The performance measure used is the coverage percentages (probabilities) for the population mean achieved by the various UCL95 computation methods.

The substitution methods are simple but do not perform well in most cases as they yield estimates with a larger bias and MSE than those obtained using the MLE methods and other nonparametric methods. Therefore, those proxy methods (e.g., replacement of NDs by "0," DL, or values obtained using the uniform distribution in (0, DL]) are not included in this study. However, since the DL/2 substitution method and ROS on log-transformed data method are the most commonly cited and used methods, these methods have been included in the simulation experiments. Helsel reviewed an earlier version of this report in November 2005 and suggested including the robust ROS on log-transformed data method in the simulation experiments. Specifically, he suggested evaluating the performances of the UCL95 based upon the robust ROS method followed by jackknife and bootstrap methods. Those results are graphed in the two newly added appendices, D and E of the report.

The simulation study, as considered in this report, is probably the first extensive study to compare the performances of the various UCL95 computation methods, including the nonparametric KM method and various bootstrap and jackknife methods on left-censored data sets covering a wide range of skewness. One of the problems associated with such studies is that these methods are computer-intensive and very time-consuming. For example, on a fast personal computer, each simulation experiment (for a fixed or a computed detection limit) with 5,050 iterations (and 1,000 bootstrap resamples) took about 30 hours on the average. Some distributions, such as the lognormal distribution, LN(5, 2) with detection limits based upon higher censoring intensities (e.g., 50% or 60%) took even longer than 30 hours on a fast personal computer.

Two detection limit cases as discussed in the literature (e.g., Haas and Scheff (1990) and Gleit (1985)) have been considered. These cases are: 1) the fixed detection limit case in which DL stays fixed for samples of all sizes; and 2) the computed detection limit (% censoring) case where DL (as percentiles) is computed based on the distribution and the censoring intensity. Type 1 left-censored data sets have been generated from three distributions: normal, lognormal, and gamma. Generation of normal and lognormal deviates is well documented and established in the literature. The simulation process used to generate gamma-distributed data sets is similar to the one as described in Singh, Singh, and Iaci (2002) and Singh and Singh (2003). Values below the detection limit, DL (fixed or computed), are considered as nondetects. Several combinations of distribution and fixed DL have been considered.

7.1 Fixed Detection Limit Case

Depending upon the distribution, the detection limit (DL) has been fixed at a certain value, such as 10, 20. In this setting, the number of NDs represents a random variable. The following combinations of distribution and fixed detection limit have been considered.

	Distribution	**Censoring Intensity**	**Fixed Detection Limit, DL**
1	N(100, 30)	0.62%	DL Fixed at 25
2	N(100, 30)	4.77%	DL Fixed at 50
3	LN(5, 1)	5.50%	DL Fixed at 30
4	LN(5, 1.5)	14.33%	DL Fixed at 30
5	LN(5, 2)	21.18%	DL Fixed at 30
6	G(0.75, 100)	34.67%	DL Fixed at 25
7	G(2, 30)	20.33%	DL Fixed at 25
8	G(3, 20)	13.06%	DL Fixed at 25

7.2 Computed Detection Limit Case

In this setting, the detection limit, DL, changes with the censoring intensity. For a normal distribution, the "*computed*" detection limit, DL, for α100% censoring intensity is given by the equation, $DL = \mu + \sigma z_\alpha$, where z_α, the critical value of the standard normal distribution is given by $P(Z \le z_\alpha) = \alpha$. For a normal distribution, N(5, 2), with mean, 5 and *sd*, 2, and censoring intensities of 30%, and 60%, DL is given by $5 + 2z_{0.30} \sim 5 - 2 * 0.525 = 3.95$ and $5 + 2z_{0.6} \sim 5 + 2 * 0.255 = 5.51$, respectively. Similarly, the DL for the lognormal distribution is computed by exponentiation. Specifically, for a lognormal distribution, LN(5, 2), a 30% detection limit is computed by using the equation:

$$\exp(5 + 2*z_{0.30}) = \exp(5 - 2*0.525) = \exp(3.95) \sim 52$$

The α100% detection limit (percentile) for a gamma distribution can be computed using the relationship between a chi-square distribution and a gamma distribution. Specifically, the relationship between a gamma random variable, $X = G(k, \theta)$, and a chi-square variable, Y, is given by $X = Y\theta/2.0$, where Y follows a chi-square distribution with $2k$ degrees of freedom. Thus, the percentiles of a chi-square distribution (which are programmed in ProUCL 3.0 and SimCensor) can be used to determine the percentiles of a gamma distribution. In practice, the shape parameter, k, is replaced by its MLE. Thus, once a α100% percentile, $y_{(\alpha)}$, of a chi-square distribution with $2k$ degrees of freedom is obtained, the α100% percentile (a DL with α100% censoring intensity) for a gamma distribution can be obtained by using $x_\alpha = y_\alpha \theta / 2.0$.

7.3 Distributions, Sample Sizes, and Censoring Intensities Considered

For the computed detection limit case, several parameters have been considered covering a wide range of skewness. It is necessary to cover a wide range of skewness, because a method that performs well by providing adequate coverage for the population mean of a mildly skewed distribution may not perform well for moderately skewed to highly skewed distributions. The detection limits are computed following the procedure as described above for normal, lognormal, and gamma distributions. Seven combinations of distributional parameters have been considered. The seven distributions considered are: N(100, 30)-symmetric, LN(5, 0.75)-mild skewness, LN(5, 1.5)-moderate-to-high skewness, LN(5, 2) high skewness, G(0.5, 100)-high skewness, G(0.75, 100)-high skewness, and G(2, 100)-moderate-to-mild skewness.

It should be noted that for the four bootstrap methods, the number of bootstrap runs, *B*, used was set at *B* = 1000. The bootstrap procedure for censored data sets (Efron (1981)), as described in the bootstrap section, has been used on left-censored data sets. That is, for bootstrap resamples, an indicator variable, *I*, was used to keep track of detect or nondetect status of an observation. For all of the simulation experiments (fixed DL or computed DL), the sample sizes, *n*, considered are: 10, 15, 20, 25, 30, 35, 40, 50, 75, and 100. For the computed detection limit cases, the censoring levels (intensities) considered are: 10%, 15%, 20%, 25%, 40%, 50%, 60%, and 70%. Initially, 5,050 simulation iterations were used for each combination of the distribution, sample size, and censoring level.

Some convergence problems were observed with the MLE methods, EM method, and some bootstrap methods (e.g., for small sample sizes) based upon the ML estimates. Specifically, for higher censoring levels, the MLE and the EM methods do not converge properly and sometimes yield absurd and unreliable values (e.g., yielding negative MLEs of the mean even when all sample values are > 0). This is especially true when dealing with low detected values, as can be seen in the examples dealing with the Manly (2001) data set in Section 6.1 and the Blood Lead (Helsel (2005)) data set in Section 6.2.

Also, as mentioned before, a data set with potential outliers (e.g., Examples 6 and 7) or a mixture sample can be easily modeled by a lognormal distribution. Just as for the full data sets (e.g., Singh, Singh, and Iaci (2002)), the UCL95 computed from a left-censored data and the lognormal distribution assumption can be unrealistically large. In such cases, the bootstrap t-method also yields unrealistic UCL 95 results. It is well known that bootstrap t-estimates are not stable in the presence of outliers (Efron and Tibshirani (1993) and ProUCL 3.0 User Guide (2004)). These observations give more reasons to use nonparametric methods, such as the KM method, rather than parametric MLE or ROS methods when dealing with left-censored skewed data sets. However, it should be noted that in order to use bootstrap methods on left-censored data sets, the sample size should not be very small. When the sample size is small (e.g., $n < 10$-15) or the censoring level is high (e.g., > 40-50%), a bootstrap sample can result in all (or most) values equal to NDs, or all (or most) observations having the same detected numerical value. Statistics computed based upon such samples are meaningless. In order to take care of some of these problems (e.g., when a bootstrap resample has no detected value or has very few detected values such as < 4), some guidelines and necessary steps were taken to generate an appropriate number of detected observations in bootstrap samples. These steps are listed as follows.

7.4 Steps Used in Data Generation and Bootstrap Resampling Methods

First of all, when the number of detects in a data set is small (such as less than 8-10), it is desirable not to use statistical methods (including those described in this report). In order to get reasonable and reliable estimates, it is recommended to have at least ten (10) observations in a data set; more than 10 observations are preferable. The minimum sample size considered in the simulation experiments is 10. It should be noted that the sample of size of 10 might not be adequate enough to compute reliable UCL95 for higher censoring levels (e.g., >20-25%).

Secondly, in order to calculate meaningful and reliable statistics, for the simulation experiments, it is assumed that a data set has at least four detected values (more detected values are preferred). Estimation of the mean, standard deviation, and computation of a UCL95 (and of other limits) based upon data sets with less than 4-5 detected values cannot be considered reliable for both parametric (e.g., the example discussed on page 59 in Helsel (2005)) as well as nonparametric data sets. For a bootstrap method, it may not even be possible to compute a reliable UCL95 as the bootstrap samples may consist of all ND values or all values may be equal to the same detected value. The following steps were taken to avoid such situations and samples.

- Generated samples with less than four detected values were rejected. The number of total iterations, 5,050 (or 10,101 iterations in later experiments), is decreased by the number of rejected samples. This is done separately for each sample size considered.

- MLE methods do not perform well on small data sets. Therefore, samples with all nondetects or with less than four detected values (e.g., when $k > n$-3) were not considered in the simulation experiments. Also, whenever a bootstrap sample had all nondetects, or too few detected values (< 4-5), that sample was not included in the calculations; instead, another bootstrap resample was drawn. The occurrences of these situations are time-consuming.

- As noted earlier, sometimes the MLE methods do not converge well and yield incorrect or negative values for the mean, *sd*, and UCL95. When a MLE method fails, especially while using a bootstrap t or a standard bootstrap method, a flag was used for those failed bootstrap methods.

- The simulation program (SimCensor) automatically rejects samples (out of 5,050) having less than five values greater than the detection limit, DL. For example, for a sample of size 10 with 60% NDs, the average number of NDs will be 6. Bootstrapping such samples or using MLE methods on such samples may yield unstable and unreliable estimates of the mean, *sd*, and UCL95. In the simulation experiments, several counts and percentages were computed to keep track of the rejected samples (all summarized in Appendix C) before computing the appropriate statistics.

- Due to above-mentioned limitations with bootstrap and MLE methods, in some cases (especially for skewed data sets), the final number of simulated samples used may be smaller than the 5,050 number of generated samples for each combination of parameters, sample size, distribution, and censoring intensity. All of these statistics are included in the simulation results, as given in Appendix C.

- It should be noted that a rejected value (obtained from the rejected sample) was not used in the summation to compute the average UCL values or to compute the coverage percentages (e.g., number of times, UCL > population mean). The fail count for that UCL method was incremented each time a sample was rejected. The average UCL and the coverage probabilities were computed appropriately by adjusting for the number of iterations, 5,050 minus the number of rejected samples.

- In order to avoid an endless bootstrap loop (for highly skewed data sets perhaps with a few potential outliers), if the rejected number of bootstrap samples (for a single generated sample) exceeded the number of bootstrap samples (1,000 here) used, that sample was rejected (from 5,050) for the UCL and coverage calculations, and the fail count was adjusted accordingly. This is one of the reasons that the process takes longer to get all of the results for skewed distributions, such as a lognormal, LN(5, 2) distribution with censoring intensities higher than 50% or 60%.

7.5 Methods Selected for Graphical Comparisons

After working through Examples 6 and 7 of Section 6 and carefully and thoroughly reviewing the simulation results based upon 5,050 simulation runs (Appendix C) for each combination of distribution, sample size, and censoring level, several methods, such as the EM method, ROS on raw data, EPA delta lognormal method, Winsorization method, and the Bootstrap t-Method, have not been included in the graphical displays and comparisons of the UCL95 computation methods as summarized in Appendices A and B. Methods listed as follows have been selected to graphically represent and compare the coverage performances (percent coverage of the population mean by UCL95) and average UCL values. Interested readers may want to review the numerical simulation results for all methods, as summarized in Appendix C.

Normal Distribution	**Lognormal Distribution**	**Gamma Distribution**
KM (z)	KM (z)	KM (z)
DL/2 (t)	DL/2 (t)	DL/2 (t)
Cohen's MLE (Tiku)	Gamma ROS (Appx.)	Gamma ROS (Appx.)
Cohen's MLE (t)	KM (%)	KM (%)
UMLE (Tiku)	KM (BCA)	KM (BCA)
KM (%)	KM (Chebyshev)	KM (Chebyshev)
KM (BCA)	FP-ROS (Land)	KM (jackknife)

Since only 7 methods were selected for graphical comparisons, a smaller simulation study was conducted using 10,101 iterations (instead of 5,050) for the 7 selected methods for each combination of distribution, sample size, and the detection limit. The main reasons to undertake such a step are: 1) the use of a larger number of iterations, 10,101 (instead of 5,050) yields more stable statistics; 2) the number of rejected samples decreased considerably as the troublesome methods (e.g., EM method, bootstrap t, standard bootstrap methods) causing failures and rejections were not included in this simulation study; and 3) finally, it was much easier to manage and deal with the smaller number of spreadsheets with all of the results in one file. In the earlier simulation runs, separate UCL and percent coverage files were created for each UCL95 computation method (e.g., BCA bootstrap, Chebyshev). The compiling and extraction of useful data for the graphical displays of the selected methods (listed above) from the earlier version of the output files became very tedious and time-consuming. The time was saved by putting all of the results in one output file for the simulation experiment of 10,101 iterations. This effort reconfirms the adequacy of the results obtained in two sets of simulation experiments. The results of the two simulation experiments (given in Appendix C and graphical displays in Appendices A and B) are in good agreement.

The graphical comparisons for the selected methods (based upon 10,101 iterations) listed above are given in Appendix A (percent coverage by UCLs) and Appendix B (average UCL values). Based upon the graphical comparisons, several observations are made, which are summarized in Section 8. For all interested readers, Appendix C has all of the numerical results obtained using 5,050 simulation runs for all of the methods considered for each combination of sample size, distribution, and % censoring intensity.

7.6 Additional Simulation Runs to Evaluate Helsel's Robust ROS Method

An earlier version of this report was peer reviewed during November-December 2005. In the November 2005 review of the report, Helsel suggested evaluating the performance of the UCL95 based on the robust ROS on log-transformed data. He specifically suggested comparing the performance of the UCL95 method on Helsel's robust ROS followed by jackknife (as used in Shumway, Azari, and Kayhanian (2002)) and bootstrap methods with other methods as considered in this report. It is noted that on a full data set obtained using robust ROS on log-transformed data, one can use ProUCL 3.0 to compute a UCL95. ProUCL 3.0 has both jackknife and bootstrap methods to compute a UCL95 based upon a full data set. Depending upon the sample size and data skewness (Singh and Singh (2003)), the ProUCL 3.0 program selects the most appropriate UCL method to compute an appropriate UCL95.

However, in order to address Helsel's comments and suggestions, additional simulations were conducted to evaluate and directly compare the performance of Helsel's robust ROS (or equivalently robust ROS) UCL95 method with other methods, such as the KM UCL95 method. The main objective of this additional simulation study is to evaluate the robustness of the robust ROS method, and also to determine if the UCLs based upon the robust ROS method provide adequate coverage for the population mean of skewed data distributions. Simulation experiments using 10,101 iterations were conducted for the same combinations of sample size, distributions, and censoring intensities as used in the earlier simulation study described in Section 7.3. It should be noted that normal and symmetric distributions are not considered in this experiment as there are other better well established UCL computation methods available (e.g., see Sections 8.2.1 and 9.1) for normally distributed data sets. Also, steps as described in Section 7.4 were used during data generation and bootstrap resampling methods. The findings of the additional simulation results are summarized in Section 8.3. The methods considered in the additional simulation experiments are listed as follows.

Lognormal Distribution	**Gamma Distribution**
KM (t)	KM (t)
H-ROS (jackknife)	H-ROS (jackknife)
Gamma ROS (Appx.)	Gamma ROS (Appx.)
KM (%)	KM (%)
KM (BCA)	KM (BCA)
KM (Chebyshev)	KM (Chebyshev)
H-ROS (%)	KM (jackknife)
	H-ROS (BCA)
	H-ROS (%)

As noted earlier, most of the existing work (e.g., Kroll, C.N. and J.R. Stedinger (1996), Shumway, Azari, and Kayhanian (2002)) evaluated the robust ROS on log-transformed data or ROS on robust MLE method for mildly skewed cases with *sd* of log-transformed data = 0.2 (with *sd* of raw data = 0.56). Those results cannot be generalized to moderately skewed and highly skewed distributions. In this report, the robust ROS on log-transformed data has been evaluated for highly skewed distributions. It should be noted that a UCL95 based upon robust ROS followed by a jackknife procedure is similar to Student's t-UCL95. Therefore, the robust ROS method followed by the jackknife method is not included in most of the additional simulation experiments. It is also well known that when skewness is high (e.g., *sd* of log-transformed data > 1.0), Student's t-UCL95 (and, therefore, robust ROS followed by jackknife) does not provide adequate coverage to the population mean (Singh and Singh (2003) and ProUCL 3.0 User Guide (2004)).

Section 8

Summary of the Simulation Results

The main objectives of this report are: 1) to evaluate and compare the performances of the various parametric and nonparametric UCL95 (and not of point estimates of the mean and *sd*) computation methods based upon left-censored data sets; and 2) to make appropriate recommendations for the computations of UCL95 often used in various environmental applications, including risk assessment and background evaluations. It needs to be pointed out that in this process of computing and estimating the population mass or mean, EPC terms, and background statistics (e.g., UPLs, UTLs), the objective is to compute accurate and meaningful statistics based upon the "majority" of the data representing the main dominant population. It is not desirable to compute and use distorted estimates (often exceeding the maximum value in a data set) of mass or mean of the entire population by accommodating a few potential outlying observations (if any) occurring with lower probabilities. The inclusion of a few outliers distorts the statistics of interest for the entire population (e.g., study area, EA, RU, AOC). Those few extreme observations need separate investigation. This is specifically true when one is trying to establish background threshold values based upon the various upper limits.

Just as in earlier studies, it is again recommended to avoid the use of transformations (Singh, Singh, Engelhardt (1997) and Singh, Singh, and Iaci (2002)) on raw data to achieve symmetry (approximate normality), even when the log-transformed data appear to be normally distributed (Examples 6 and 7). Typically, the parameter (hypothesis of interest) in the transformed space (median) is not of interest to make remediation and cleanup decisions in the original scale (mean). Moreover, the practitioners often do not know how to interpret and use the transformed results or back-transform the results to the original scale, and draw conclusions based upon the statistics thus obtained. This issue was discussed in detail in Section 5. The estimates of the back-transformed parameters (from transformed space) in the original space suffer from a significant and unknown amount of transformation bias. The transformation bias can be unrealistic for skewed data sets with a few outliers, as illustrated in Example 6. Even though equation (3-22) is a well-documented equation in the environmental literature (Gilbert (1987), El-Shaarawi (1989), Singh and Nocerino (2002)), it is recommended not to use equation (3-22) to back-transform estimates from log scale to original scale, as illustrated in Section 6.

The question now arises – how one should back-transform results from a log scale (or any other transformed scale) to the original scale? Unfortunately, no defensible guidance is available in the environmental literature to address this question. Moreover, the back-transformation formula will change from transformation to transformation (e.g., BC-type transformations), and the bias introduced by such transformations will stay unknown. In many cases, the derivation of a back-transformation formula (e.g., for a BC-type transformation) may be quite involved and complicated. This in turn may force the user to report the results and estimates in the transformed space. However, the estimates obtained in transformed space may not be of much use as the remediation and cleanup decisions have to be made in the original raw scale. Therefore, in the cases when a data set in the "raw" scale cannot be modeled by a parametric distribution, it is desirable to use nonparametric methods rather than testing or estimating some parameter in the transformed space. The use of nonparametric methods will spare the user from: 1) developing back-transformation formulae associated with the various transformations and 2) assessing bias in estimates thus obtained.

8.1 General Observations Based Upon the Simulation Results

Some observations based upon the graphical displays (Appendices A, B, D, and E) and simulations results as presented in Appendix C are described as follows:

- *Coverage increases with censoring level:* It is observed that, except for the DL/2 (t) UCL method, the coverage of the population mean by the various UCL95 (e.g., KM (BCA), KM (%)) increases as the censoring intensity increases from 10% to 70% (figures in Appendix A).

- *Coverage increases with sample size:* It is noted that the coverage for the population mean increases for all methods except for the DL/2 (t) method as the sample size increases. This is true for all distributions and censoring levels.

- *DL/2* (*t*) *method does not perform well for all sample sizes:* It is noted that the coverage for the mean by the UCL95 based upon DL/2 (t) decreases as the sample size increases for all censoring levels.

- *DL/2* (*t*) *method does not perform well for low as well as high censoring levels:* The coverage for the population mean, provided by the DL/2 (t), decreases as the censoring intensity increases. This is true for all symmetric and skewed distributions. The performance of the UCL95 based upon the DL/2 (t) method is the worst (coverage much smaller than 95%) for symmetric distributions.

  - Thus, contrary to the general rule of thumb, it should be noted that the DL/2 (t) does not perform well even for low censoring levels (Figures 2a-2c, 3a-3c, and so on), such as 10%, 20%, and 30%. The coverage deteriorates fast as the censoring level increases (Figures 2d-2f, 3d-3f, and so on, Appendix A).

- *The H-statistic-based UCL95:* The H-statistic-based UCL95, computed by using the k extrapolated NDs and the (n-k) detected values obtained using ROS on log-transformed data (fully parametric ROS = FP-ROS), does provide approximately 94-95% coverage for the population mean (Figures 3, 3a, 3b, 3i, 8a, and 8i of Appendix A).

  - However, for moderately skewed to highly skewed data sets (with *sd* of the log-transformed data >1, 1.5, 2.0), just like in the case of full data sets, the H-UCL based upon k extrapolated NDs obtained using ROS on log-transformed data (FP-ROS method) and (n-k) observed values behave in an erratic and unstable manner. That is, Land's H-UCL results in impractically large UCL values. This can be seen in Figures 5, 6a-6i, 7a-7i, and 8a-8i of Appendix B.

- *Bootstrap t and standard bootstrap methods:* From the simulation results summarized in Appendix C, it is noted that the bootstrap t and standard bootstrap methods on MLE and EM estimates often yield erratic and impractically large UCL95 values. Therefore, coverages and UCL95 values based upon those two bootstrap methods have not been graphed in Appendices A and B.

- *The delta lognormal method:* The use of the delta lognormal method (USEPA (1991), Hinton (1993)) yields unstable UCL95 values, as can be seen from the results presented in Appendix C. Therefore, this method was not included in the graphical displays of Appendices A and B.

- *For symmetric distributions, Tiku's method performs better than the Schneider UCL method:* It is noted that Tiku's (1971) approximate UCL95 method performs better (in terms of coverages provided by respective UCL95 values) than the Schneider (1986) UCL95 method. Therefore, coverages and UCL95 for Schneider's method have not been graphed in Appendices A and B.

- *The Winsorization UCL method works only for symmetric distribution:* UCL95 based on Winsorization method provides adequate coverage to the mean only for symmetrical distributions (Appendix C). The coverage provided by this method is far from 0.95 for asymmetrical distributions.

- *UCL methods based upon KM estimates perform better than other UCL methods:* UCLs based upon the KM method (KM (z), KM (t), KM (%), KM (BCA)) seem to perform better than the various other methods as can be seen from the various figures included in Appendices A, B, D, and E.

- *MLE UCL methods work only for symmetric distributions for censoring levels lower than 40%:* The various parametric MLE methods (CMLE, UMLE, and RMLE) provide 95% coverage to the population mean only for normally distributed data sets. These MLE methods do not work well for skewed distributions including lognormal and gamma distributions. This can be seen from the various results as presented in Appendix C. This observation also suggests that ROS on MLE method (which is based upon log-transformed data) may not perform well for skewed data sets.

- *Convergence problems associated with iterative MLE methods:* It is noted that the MLE estimation methods (CMLE, RMLE, and UMLE) and the EM method sometimes fail to converge, and as a result yield unreliable estimates of the population mean. This is true even for normally distributed data sets, especially for data sets with higher censoring intensities (e.g., > 40%). Therefore, coverages and average UCL95 values have been graphed (Appendices A and B) only for a few of those methods such as UMLE and CMLE methods.

- *Bootstrap methods on MLE and EM methods do not perform well:* For higher censoring levels, it is observed that the bootstrap UCL95 based upon MLE and EM estimation methods become unrealistic and unstable. Therefore, the bootstrap results obtained using MLE and EM methods have not been graphed in Appendices A and B.

- *The jackknife UCL method works only for symmetric distributions:* For symmetric distributions, the jackknife method used on the KM estimates seems to work well as it provides adequate 95% coverage for the population mean.

  - However, the coverage by the jackknife UCL95 for the mean deteriorates (decreases) for skewed distributions, especially for lognormal distributions with standard deviation of log-transformed data exceeding 1.

  - The KM (jackknife) results have been graphed for the gamma distribution as can be seen in Figures 9a, 9b, ..., 9i, 13a, ..., 13i, and 14 of Appendices A and B. It can be seen that for gamma distribution, the KM (jackknife) method does not provide adequate 95%

coverage to the population mean for censoring levels < 40% (e.g., Figures 9a, 9b, ..., 9i in Appendix A).

Thus, based upon the above observations, percent coverages (Appendix A and D) and the UCL95 averages (Appendix B and E) have been graphed only for the following methods.

Table 8-1. Methods Included in Graphical Displays

Methods	Explanation
Cohen's MLE (Tiku)	UCL based upon Cohen's maximum likelihood estimates using Tiku's method.
Cohen's MLE (t)	UCL based upon Cohen's maximum likelihood estimates using Student's t-distribution cutoff value.
DL/2 (t)	UCL based upon DL/2 method using Student's t-distribution cutoff value.
FP-ROS (Land)	UCL based upon fully parametric ROS method using Land's H-statistic.
Gamma ROS (BCA)	UCL based upon Gamma ROS method using the bias-corrected accelerated percentile bootstrap method.
Gamma ROS (Appx.)	UCL based upon Gamma ROS method using gamma approximate-UCL method.
H-ROS (BCA)	UCL based upon Helsel's robust method using the bias-corrected accelerated percentile bootstrap method.
H-ROS (%)	UCL based upon Helsel's robust method using the percentile bootstrap method.
H-ROS (jackknife)	UCL based upon Helsel's robust method using the jackknife method.
KM (z)	UCL based upon Kaplan-Meier estimates using standard normal distribution cutoff value.
KM (t)	UCL based upon Kaplan-Meier estimates using Student's t-distribution cutoff value.
KM (%)	UCL based upon Kaplan-Meier estimates using the percentile bootstrap method.
KM (BCA)	UCL based upon Kaplan-Meier estimates using the bias-corrected accelerated percentile bootstrap method.
KM (Chebyshev)	UCL based upon Kaplan-Meier estimates using the Chebyshev theorem.
KM (jackknife)	UCL based upon Kaplan-Meier estimates using the jackknife method.
UMLE (Tiku)	UCL based upon unbiased maximum likelihood estimates using Tiku's method.

8.2 Observations Based Upon the Graphical Displays of Appendices A and B

The performance of the various estimation and UCL95 computation methods for left-censored data sets depends upon several things such as the sample size, skewness, censoring intensity, and data distribution. Using the graphical displays of Appendices A and B, detailed discussion of the observations made and conclusions derived for each of the distribution included in the simulation study are summarized as follows.

8.2.1 Normal Distribution

The normal distribution is symmetric and is the easiest distribution to compute the UCL95 based upon left-censored data sets. In this report, the normal distribution, N(100, 30), for the various censoring levels, 10%, 15%, ... , up to 70%, has been considered. Figures 1a (% censoring = 0.62%), 1b (% censoring = 4.8%), 2a (% censoring = 10%), and 2b, ..., 2i (% censoring = 70%) of Appendix A represent the coverages provided by the various UCL95 methods as a function of the sample size, while the corresponding figures in Appendix B represent the average UCL values for the various censoring intensities and sample sizes. From these figures, the following observations have been made.

- *Coverage for the mean by a UCL95 increases with sample size and censoring intensity:* As mentioned before, for all UCL methods (except for the DL/2 (t) UCL method), the coverages provided by the various UCL95 methods increase as the censoring intensity and sample sizes increase. This can be seen in Figures 1a, 1b, 2a, ..., 2i of Appendix A. Coverages provided by the KM UCL95 methods, including the KM (z), KM (t), KM (%), and KM (BCA) methods, are at or above 0.95 coverage for all censoring levels and sample sizes. Therefore, any of these methods can be used to compute a UCL95.

- *Do not use DL/2* (*t*) *method to compute a UCL:* The coverage to the population mean provided by the DL/2 (t) UCL method decreases steadily as the sample size increases and the censoring intensity increases. Even, for normally distributed data sets, it is observed that, for censoring levels as low as 5% or 10%, the DL/2 (t) UCL95 does not provide the specified coverage (of 95%) for samples of sizes 20 and larger, as can be seen in Figures 1b, 2a,…, 2i of Appendix A. Therefore, the use of the DL/2 (t) UCL should be avoided, even for low censoring levels, such as 10% and 15%. This is contrary to the general recommendation (conjecture) made to use the DL/2 method for censoring levels up to 20%, USEPA (2000) and SW-846 (USEPA (1993)). Obviously, the UCL averages for the DL/2 (t) method become even smaller than the population mean, 100, for larger censoring intensities, as can be seen in Figures 1a, 1b, 2a, ..., 2i of Appendix B.

- *Do not use ad hoc UCL methods based upon Cohen's MLE* (*t*) *or other MLE and EM* (*t*) *methods:* It is noted that for the ad hoc UCL95 method based upon Cohen's MLE (t), the coverage for the population mean decreases gradually as the censoring intensity increases. The coverage provided by Cohen's MLE (t) falls below 90% when the censoring level reaches 50% or higher.

- *MLE methods do not behave properly for higher censoring levels:* Cohen's MLE (Tiku) and UMLE (Tiku) perform well for lower censoring levels. However, many times, the parametric MLE methods, such as Cohen's MLE, UMLE, and the EM method, do not converge properly. This is especially true when the censoring intensity becomes large (e.g., > 40%), as can be seen in Figures 2g, 2h, and 2i (Appendix B). In such cases, it is preferable to use nonparametric UCL95 methods based upon the KM estimation method.

- *Use UCL95 computation methods based upon the KM estimation method:* From Figures 1a, 1b, 2a, ..., 2i of Appendix A, it is noted that the KM (BCA) UCL method provides slightly higher than 95% coverage for all censoring levels and sample sizes. Therefore, for symmetric distributions, the most appropriate methods to compute a UCL95 are the KM (z), KM (t), and KM (%) methods.

8.2.1.1 Recommended UCL95 Methods for Normal (Approximate Normal and Symmetric) Distributions

- *For normal and approximate normal* (*e.g., symmetric*) *distributions:* The most appropriate UCL95 computation methods are KM (z), KM (t), or KM (%) methods. These methods perform equally well for all of the censoring levels and sample sizes.

- *MLE methods:* Even though the MLE (Tiku) and UMLE (Tiku) perform well at least for censoring levels lower than 40%, their use is not recommended here. The main reason behind this recommendation is the availability of several equally good nonparametric UCL computation methods (KM (z), KM (t), or KM (%)) that perform well for all censoring levels considered in this report.

Note: These recommended methods for normally distributed data sets will be available in the ProUCL 4.0 software package. Additionally, some parametric UCL methods, such as CMLE (Tiku) and UMLE (Tiku), will also be available in ProUCL 4.0.

8.2.2 Gamma Distribution

The gamma distribution can be used to model asymmetric (skewed) distributions (Singh, Singh, and Iaci (2002)). Several gamma distributions have been considered in the simulation experiments, as summarized in this report. These are: G(0.5, 100)-highly skewed with 10%, 15%, ..., 70% censoring (Figures 9a, …, 9i); G(0.75, 100)-skewed with 34.67% censoring (fixed DL = 25, Figure 10), and 10%, 15%, ..., 70% censoring (Figures 11a, ...,11i); G(2, 30)-moderately skewed with fixed DL = 25 = 20.33% censoring (Figure 12), and 10%, 15%, ..., 70% censoring (shown in Figures 13a, ...,13i); and G(3, 20)-mildly skewed with fixed DL = 25 with 13.06% censoring (Figure 14). The graphs for the percent coverages by the UCL95 values are given in Appendix A, and the corresponding graphs for average 95% UCLs are given in Appendix B. From these figures, the following observations have been made.

- *Coverages for population mean by UCL95s increase with sample size and censoring intensity:* Just as in the case of normal distribution, it is observed that, for the gamma distribution, the coverages for the population mean provided by the various UCL95 methods (except for the DL/2 (t) method) increase as the censoring intensity and the sample sizes increase, as can be seen in Figures 9a, …, 9i, 10, ..., 13a, ..., 13i, and 14 in Appendix A.

- *Do not use DL/2* (*t*) *method to compute a UCL:* The DL/2 (t) method does not work for any of the gamma distributions (mildly skewed to highly skewed). The coverage provided by DL/2 (t) decreases gradually as the censoring level increases. It is not recommended to use the DL/2 (t) method for any of the censoring levels, including the low censoring levels of 10%, 20%, and so on (e.g., Figures 13d, 13e, 3f, 13g, and so on).

- *Use the KM* (*Chebyshev*) *UCL for highly skewed gamma* (*k = 0.5, 0.75, < 1*) *distributions with censoring levels < 30% for all sample sizes:* For highly skewed gamma, G(k, θ), distributed data sets (with estimated shape parameter, k = 0.5, 0.75, < 1) with censoring levels lower than < 30%, none of the methods, except for the KM (Chebyshev) and the estimated approximate gamma UCL95, provide 95% coverage for the population mean (Figures 9a, …, 9i, 11a, …,11i). Since both methods provide roughly the same coverage, the nonparametric KM (Chebyshev) method is preferred (as in practice, it may be hard to verify distributional assumptions on real data sets) to compute UCL95 for censoring levels < 30%. It should be noted that these two methods may yield "conservative UCL95" (a UCL providing at least (often higher) 95% coverage for the population mean) values; that is, the actual coverages provided by these two UCL methods tend to be higher than the specified coverage of 0.95.

- *Use the KM* (*BCA*) *UCL for highly skewed gamma* (*k = 0.5, 0.75, < 1*) *distributions with censoring levels in the interval, 30-50%, for samples of all sizes:* From Figures 9e, 9f, 9g, 11e, 11f, and 11g, it is observed that the UCL95% based upon the KM (BCA) method starts providing about 95% coverage for the population mean.

- As the censoring level approaches 50% and becomes larger than 50%, even the KM (%), KM (z), and KM (t) methods, and the KM (jackknife) method, start providing about 95% coverage for the population mean (Figures 9g, 9h, 9i, 11g, 11h, and 11i). Therefore, for censoring levels exceeding 50%, the KM (z) or KM (t) UCL95 method may be used to compute UCL95 for samples of all sizes. Note that the KM (t) UCL95 method has been included in the new simulation experiments, as graphed in Appendices D and E.

- *For moderately skewed gamma distributions, G(k, θ), with shape parameter $1 < k \leq 2$:*

  - For censoring level ≤ 10% (Figure 13a), one may use the KM (Chebyshev) UCL for samples of all sizes.

  - For censoring levels in the interval: 10-20%, one can use the KM (BCA) UCL method for samples of all sizes.

  - For censoring levels between 25% and 40%, one can use the KM (%) method.

  - For censoring levels ≥ 40%, any of the KM UCL95 methods, such as KM (%), KM (z) and KM (t), can be used, as can be seen in Figures 13f, 13g, 13h, and 13i.

- *For mildly skewed gamma distributions, G(k, θ), with $k > 2$ (Figure 14):*

  - Use the KM (BCA) method for lower censoring levels (≤ 20%),

  - Following a similar pattern as for the case when the shape parameter ≤ 2, for censoring levels > 20%, one can use the KM (%) method, and

  - For censoring ≥ 40%, one can use the KM (z) or KM (t) UCL computation method for samples of all sizes.

Note: Observations and findings about the use of the robust ROS (Helsel's method) method on gamma distributed left-censored data sets to compute appropriate UCL95 values are summarized in Section 8.3.

8.2.2.1 Recommended UCL95 Methods for Gamma Distributions

- *For highly skewed gamma distributions with shape parameter, $k \leq 1$:*

  - Use the nonparametric KM (Chebyshev) method to compute a UCL95 for censoring levels < 30%,

  - Use the nonparametric KM (BCA) method to compute a UCL95 for censoring levels in the interval [30%, 50%), and

  - Use the nonparametric KM (z) or KM (t) method to compute a UCL95 for censoring levels ≥ 50%.

- *For moderately skewed gamma distributions, G(k, θ), with shape parameter, $1 < k \leq 2$:*

    - For censoring level ≤ 10%, use the KM (Chebyshev) UCL method,
    - For censoring levels in the interval (10%, 25%), use the KM (BCA) UCL method,
    - For censoring levels in the interval [25%, 40%), use the KM (%) UCL method, and
    - For censoring levels ≥ 40%, use KM (z) or KM (t) UCL95 method.

- *For mildly skewed gamma distributions, G(k, θ), with $k > 2$:*

    - Use the KM (BCA) UCL method for censoring levels lower than ≤ 20%,
    - For censoring levels in the interval (20%, 40%), use the KM (%) UCL computation method, and
    - For censoring levels ≥ 40%, use the KM (z) or KM (t) UCL computation method.

Note: The recommended UCL methods for gamma distributed data sets will be available in the ProUCL 4.0 software package.

8.2.3 Lognormal Distribution

It is noted that the UCL95 for the various methods, including the conservative KM (Chebyshev) UCL, based upon a lognormal model does not perform well, as can be seen in Figures 5, 6a, …, 6i, 7, and 8a, ..., 8i in Appendix A. It is noted that most of the available methods (including the robust ROS method) fail to provide the specified 95% coverage to the population mean. This is especially true for moderately skewed to highly skewed data distributions with the standard deviation, σ, of the log-transformed variable, Y exceeding 1.0. It is also noted that several of the methods included in this study assume the lognormality of the data distribution. These methods are the ROS on log-transformed data (both fully parametric version and robust ROS on log-transformed data) and the EPA delta lognormal methods.

For fully parametric ROS on log-transformed data and the EPA delta log method, Land's H-statistic-based UCL95 has been computed. It is observed that just like for the full uncensored data sets (Singh, Singh, and Iaci (2002) and the ProUCL 3.0 User Guide (2004)), the UCL95 based upon Land's H-statistic (1971) does provide the approximate 95% coverage for the mean of the lognormal distributions considered (e.g., in Figures 5, 6a, …, 6i, 7, and 8a, ..., 8i, Appendix A), but the resulting H-UCL95 values are unstable and of no practical merit (e.g., Figures 5, 6a, …, 6i, 7, and 8a, ..., 8i in Appendix B), especially when the *sd*, $\hat{\sigma}$, of the log-transformed data starts exceeding 1.0. A similar pattern is observed for the various ad hoc H-UCL values based upon the various MLE and EM estimates obtained using log-transformed data sets.

Moreover, as mentioned before, the parametric MLE and EM methods sometimes fail to converge and yield unreliable values for higher (e.g., > 40%) censoring levels. Therefore, none of these methods based upon the lognormal assumption have been included in the graphical displays provided in Appendices A and B. Based upon all of these observations, in addition to the fully parametric ROS UCL95 (based upon the H statistic) and Gamma ROS Appx. UCL95 (based upon an approximate gamma UCL obtained using estimated gamma parameters), only the nonparametric methods (e.g., bootstrap and Chebyshev inequality) using the KM estimates have been considered and graphed in Appendices A and B.

Just like for uncensored full data sets (ProUCL 3.0 User Guide (2004)), the recommendations on how to compute a UCL95 based upon left-censored data sets from lognormal and other skewed nonparametric distributions have been made as a function of $\hat{\sigma}$, the *sd* of log-transformed data. It should be noted that both the skewness and CV of a lognormal distribution are functions of $\hat{\sigma}$, the *sd* of log-transformed data. The higher $\hat{\sigma}$, the higher is the skewness. These relationships between skewness, CV, and $\hat{\sigma}$ have been illustrated earlier in Tables 3-1 and 3-2 of Section 3.1. It is pointed out that as $\hat{\sigma}$ increases (even a small increase, such as from 2.0 to 2.2), the coverages provided by the various UCL95 methods start decreasingly rapidly. Also, the H-UCL based upon the FP-ROS method on log-transformed data starts behaving in an unstable manner, yielding unrealistically large UCLs (Figures 6a to 6i, 7, and 8a to 8i, Appendix B). A similar setting (in terms of $\hat{\sigma}$) was used in the ProUCL 3.0 User Guide (2004) to make recommendations for the computation of a UCL95 based upon uncensored full data sets obtained from skewed lognormal or nonparametric distributions. Using the graphical displays of Appendices A and B, the following observations have been made.

- *Coverages for population mean by various UCL95s increase with sample size and censoring intensity:* It is observed that, for the lognormal distribution, the coverages for the population mean provided by the various UCL95 methods increase as the censoring intensity and the sample sizes increase as shown in Figures 4, 5, 6a, …, 6i, 7, and 8a, ..., 8i, in Appendix A.

- *Do not use the DL/2* (*t*) *UCL method:* The UCL95 based upon the DL/2 (t) method does not provide the specified 95% coverage for the population mean of skewed lognormal or other skewed distributions for any censoring level (low or high) and sample size.

- *For mildly skewed data sets with* $\hat{\sigma} \leq 1$:

  – For lower censoring levels (≤ 20%), one can use the KM (Chebyshev) UCL95 for samples of sizes less than 50-70. This may yield slightly conservative (but stable) UCL values providing a higher than 95% coverage for the population mean or mass.

  – For lower censoring levels (≤ 20%), one can use the KM (BCA) UCL95 for samples of sizes greater than 50-70.

  – For censoring levels in the interval (20%, 40%), other UCL95 methods, such as the KM (BCA) method, can be used for samples of all sizes.

  – For censoring levels ≥ 40%, most of the methods (except the DL/2 (t) method) provide at least 95% coverage for the population mean. Therefore, one can use the KM (%), KM (t), or KM (z) method to compute a UCL95.

- *For moderately skewed to highly skewed data sets with $\hat{\sigma}$ in (1, 1.5]:*

  – For censoring levels (≤ 50%), even the KM (Chebyshev) UCL95 method fails to provide the specified 95% coverage for the population mean. This is especially true for smaller sample sizes (e.g., < 40). For such highly skewed data sets, a higher value for the confidence coefficient (e.g., 0.975 or 0. 99) may be used to achieve the specified (~0.95) amount of coverage by the UCL. Typically a higher confidence level (~0.975) is used for smaller samples, and a smaller confidence coefficient (~0.95) is used for larger samples.

  – For smaller sample sizes, $n < 40$, and censoring levels ≤ 50%, one may want to use a 97.5% KM (Chebyshev) UCL to estimate the population mass, EPC term, or some other threshold for the population mass.

  – For larger sample sizes, $n \geq 40$, and censoring levels ≤ 50%, one may want to use a 95% KM (Chebyshev) UCL to estimate the population mass or some other threshold value.

  – For censoring levels > 50%, one may want to use a KM (BCA) UCL95 to compute a UCL of the mean based upon left-censored data sets for samples of all sizes (Figures 6h and 6i in Appendix A).

- *For highly skewed data sets with $\hat{\sigma}$ in the interval (1.5, 2]:*

  – The H-UCL for fully parametric ROS on log-transformed data becomes very unstable and erratic (Figures 8a, 8b, ..., 8i). It is noted that for all censoring levels (≤ 70%), even the KM (Chebyshev) UCL95 method fails to provide the 95% coverage for the population mean.

  – For smaller sample sizes, $n < 40$, and censoring levels < 50%, one may want to use a 99% KM (Chebyshev) UCL to estimate the population mass, or EPC terms. In some cases, this may yield conservative, but stable, estimates of the population mass.

  – For sample sizes, $n \geq 40$, and censoring level< 50%, one can use a 97.5% KM (Chebyshev) UCL to estimate the population mass.

  – For sample sizes less than 40-50, and censoring levels ≥ 50%, use a 97.5% KM (Chebyshev) UCL.

  – For sample sizes ≥ 40-50, and censoring levels ≥ 50%, use a 95% KM (Chebyshev) UCL.

- *For extremely highly skewed data sets with $\hat{\sigma}$ exceeding 2:*

  – The UCL95 computation recommendation pattern, as described above, can be generalized to more highly skewed data sets with $\hat{\sigma}$ >2.0. For such highly skewed (see Tables 3-1 and 3-2) distributions, for lower sample sizes (e.g., <50-60), one may simply use a 99% KM (Chebyshev) UCL to estimate the population mean, EPC term, and other relevant threshold values. For sample sizes greater than 50-60, one may use a 97.5% KM (Chebyshev) UCL as an estimate of the population mean or mass.

Note: These methods and recommendations for the lognormal distribution will be available in the ProUCL 4.0 software package.

8.2.4 Nonparametric UCL95 Computation Method for Left-Censored Data Sets

It is noted that most of the recommended UCL computation methods for a lognormal distribution, as described in Section 8.2.3, are not based upon the assumption of a lognormal distribution. Therefore, those UCL computation methods (as functions of $\hat{\sigma}$) described in Section 8.2.3 can also be used on nonparametric data sets. Such nonparametric data sets do not follow any of the well-known parametric distributions.

- *For symmetric or approximately symmetric nonparametric data distributions:* Depending upon the symmetry or approximate symmetry of nonparametric data distributions, one may use the same UCL computation methods as for the data sets coming from a normal or an approximate normal distribution/population. These methods are described in Section 8.2.1.

- *Skewed nonparametric data distributions:* As mentioned before, it is noted that most of the recommended UCL computation methods for a lognormal distribution as described in Section 8.2.3 do not assume the lognormality of the data set. Therefore, for skewed distribution-free data sets (nonparametric distributions), the UCL computation methods as described in Section 8.2.3 can be used.

Note: These methods and recommendations for nonparametric data distributions will also be available in the ProUCL 4.0 software package.

8.2.5 Choosing a Confidence Coefficient of a UCL for Highly Skewed Data Sets

As summarized earlier, for extremely highly skewed data sets, an appropriate and stable estimate of the population mean or mass (e.g., EPC term) is given by a UCL based upon Chebyshev inequality and KM estimates. The confidence coefficient to be used depends upon the skewness and sample size of the data set under study. For highly skewed data sets, a higher (e.g., > 95%) confidence coefficient may have to be used to estimate the EPC term or the population mass. Depending upon skewness (e.g., *sd* of log-transformed data = 2, 3, or 4) and the sample size, one may have to use a 99% or a 97.5% KM (Chebyshev) UCL to estimate the EPC term and other relevant threshold values. As the skewness becomes higher, the value of the confidence coefficient becomes higher.

8.3 Simulation Results for Helsel's Robust ROS Method on Log-Transformed Data Sets

Following Helsel's review comments and suggestions (November 2005), additional simulations were conducted to evaluate the performance of the robust ROS method on log-transformed data to compute an appropriate UCL95 as an estimate of the population mass. Several UCL computation methods, including KM method (based upon Student's t-statistic, BCA and percentile bootstrap methods), robust ROS method followed by bootstrap methods (percentile and BCA) and Chebyshev UCL on KM estimates, have been included in the additional simulation experiments as described in this section. The graphical displays of the percent coverages provided by these UCL methods are given in Appendix D. The graphs of the corresponding average UCL95 values are given in Appendix E.

Since the robust ROS method assumes a lognormal distribution for the detected as well as nondetected observations, only skewed distributions (gamma and lognormal) have been considered. The UCL computation methods for normal distributions are given in Section 8.2.1. In order to evaluate the robustness of the robust ROS method on log-transformed data, several distributions covering a wide range of skewness (mild, moderate, high, extremely high), as described in Tables 3-1 and 3-2, have been considered. Special attention is given to the robust ROS method based upon percentile and BCA bootstrap methods. The two bootstrap UCL methods on the robust ROS method have been denoted by H-ROS (%) and H-ROS (BCA).

8.3.1 ROS UCL Method for Gamma Distribution

The robust ROS method on log-transformed data has been used on the various gamma distributions. The gamma distributions considered include: G(0.5, 100), G(0.75, 100), and G(2, 100). The coverage probabilities for the various censoring levels from 5% to 70% are displayed in Figures 4a-4h for the G(0.5, 100) distribution, in Figures 5a-5h for the G(0.75, 100) distribution, and in Figures 6a-6h for the mildly skewed gamma distribution, G(2, 100), Appendix D. The corresponding graphs for various average UCL95 values are given in Appendix E. The following observations have been made.

- *Coverages for population mean by UCL95s increase with sample size and censoring intensity*: As noted earlier in Appendix A, it is observed that the coverages for the population mean provided by the various UCL95 methods, including H-ROS (BCA) and H-ROS (%) methods, increase as the censoring intensity and the sample sizes increase. This can be seen in Figures 4a, ..., 4h, 5a, ..., 5h, and 6a-6h, Appendix D.

- *Bootstrap UCL methods on data sets obtained using robust ROS method*: It is noted that the performance (in terms of coverage percentage for the population mean) of Helsel's ROS method followed by percentile and BCA bootstrap methods falls in between the various other UCL methods already discussed in Section 8.2.2. None of the robust ROS methods (% and BCA) perform better than the KM (Chebyshev) and KM (BCA) methods as recommended in Section 8.2.2. This implies that the recommendations to compute appropriate UCLs for gamma distributed left-censored data sets as described in Section 8.2.2 do not change.

    - For highly skewed gamma distributions (e.g., G(0.5, 100), G(0.75,100)), just like all of the other UCL methods (except Chebyshev UCL), both the H-ROS (%) UCL and H-ROS (BCA) methods fail to provide adequate coverage for the population mean. This is especially true for censoring intensity lower than 30%. For lower censoring levels, it is recommended to use KM (Chebyshev) UCL.

    - It is noted that the H-ROS (BCA) UCL method performs better (in terms of coverage probabilities) than the H-ROS (%) method.

    - It is also noted that the H-ROS (%) method provides a much lower coverage for the population mean than the various other UCL methods, as can be seen in Figures 4a-4h and 5a-5h, Appendix D. This is especially true for smaller sample sizes and lower censoring intensities. The coverage for the H-ROS (%) UCL method improves as the sample size increases and the skewness decreases (Figures 6a-6h).

- *The H-ROS (%) and H-ROS (BCA) robust UCL methods:* These two robust ROS methods followed by bootstrap methods do not perform better than the KM (BCA) UCL method as recommended earlier in Section 8.2.2.

8.3.1.1 Recommended UCL Based Upon Robust ROS on Gamma Distribution

Based upon the results as summarized in Section 8.3.1, the recommendations to compute appropriate UCLs for gamma distributed left-censored data sets as described in Section 8.2.2 do not change. The UCLs based upon the robust ROS method (on log-transformed data) do not perform better than the UCL computation methods as recommended in Section 8.2.2.

8.3.2 Helsel ROS UCL Method for Lognormal Distribution

The lognormal distributions considered include: LN(5, 0.75), LN(5, 1.5) and LN(5, 2). The coverage probabilities for the various censoring levels from 5% to 70% are displayed in Figures 1a-1h for the LN(5, 0.75) distribution, in Figures 2a-2h for the LN(5, 1.5) distribution, and in Figures 3a-3h for the highly skewed lognormal distribution, LN(5, 2), Appendix D. The corresponding UCL95 graphs for the various lognormal distributions are given in Appendix E. The following observations have been made.

- From the figures in Appendix D, it is easy to observe that, for the various lognormal distributions considered, the coverages for the population mean provided by the various UCL95 methods, including H-ROS (BCA) and H-ROS (%) methods, increase as the censoring intensity and the sample sizes increase. This can be seen in Figures 1a, ..., 1i, 2a, ..., 2i, and 3a-3i, Appendix D.

- None of the robust ROS methods (based upon % and BCA bootstrap methods) perform better than the KM (Chebyshev) UCL method, as recommended in Section 8.2.3. The KM (Chebyshev) UCL method provides the highest coverage for all censoring levels and sample sizes.

- It is noted that the H-ROS (%) UCL method provides the lowest coverage (much lower than 95%) for the population mean for all of the censoring levels and most sample sizes, and the H-ROS (BCA) UCL method provides higher coverage to the population mean than the H-ROS (%) method for all of the censoring levels and sample sizes.

- From the figures presented in Appendix D, it is clear that the H-ROS (%) method cannot be recommended to compute a UCL for the population mean or mass based upon left-censored data sets.

- For highly skewed lognormal distributions such as LN(5, 1.5) and LN(5, 2), just like all other UCL methods, the H-ROS (BCA) UCL and H-ROS (%) methods also fail to provide adequate coverage (~ 95%) for the population mean.

    – For censoring levels exceeding 50-60%, the robust ROS method followed by the bootstrap % and bootstrap BCA methods provide the lowest coverages to the population mean, as can be seen in Figures 2f-2h and 3g-3h.

    – As the censoring intensity exceeds 50-60% (e.g., Figures 2f-2h, and 3g-3h), the coverage provided by the H-ROS (BCA) UCL method starts deteriorating rapidly. The coverage provided by the KM (BCA) UCL method becomes better (larger) than the H-ROS (BCA) UCL method.

- For highly skewed distributions (e.g., Figures 2a-2e, and 3a-3f), it is noted that for censoring levels lower than 40-50%, the coverage provided by the H-ROS (BCA) UCL method is a little larger than the KM (BCA) UCL method. However, for both methods, the coverages provided by the respective UCLs are much lower than 95%.

- Just like in Section 8.2.3, for lower censoring levels, it is recommended to use the KM (Chebyshev) UCL. The use of the confidence coefficient associated with the Chebyshev UCL will depend upon the skewness, sample size, and censoring level, as discussed in Section 8.2.5. Higher than 0.95 confidence coefficients may be used to compute a KM (Chebyshev) UCL for highly skewed distributions with *sd*, σ, of log-transformed data exceeding 1, 1.5, and so on.

- As discussed and recommended in Section 8.2.5, for lower sample sizes (<50-60) the use of the Chebyshev UCL is recommended; for higher sample sizes, one can use the KM (BCA) UCL as an estimate of the population mass.

8.3.2.1 Recommended UCL Method for Helsel's ROS on the Lognormal Distribution

- It is noted that for all of the skewed gamma and lognormal distributions considered, the performance (in terms of coverage percentage for the population mean) of Helsel's ROS method followed by percentile and BCA bootstrap methods is not better than the KM (Chebyshev) and KM (BCA) UCL methods as recommended in Sections 8.2.2 and 8.2.3.

- In order to keep the recommendations simple, it is not desirable to recommend the use of the robust ROS method on log-transformed data followed by the jackknife and bootstrap methods to compute a UCL as an estimate of the population mass.

- However, for comparison sake, the robust ROS method on log-transformed data followed by bootstrap methods will be available in ProUCL 4.0. Actually, it is already available in ProUCL 3.0. Specifically, one can simply use the bootstrap methods (and also the jackknife method), as available in ProUCL 3.0, on the full data set obtained using robust ROS on log-transformed data.

Section 9

Summary and Recommendations

This section summarizes the recommendations based on the results of the examples discussed in Section 6, and the graphical and numerical simulation results as summarized in Appendices A, B, C, D, and E. For convenience, the various recommended UCL95 computation methods have been tabulated in Table 9-1 as functions of the sample size, skewness, and censoring intensity.

9.1 General Recommendations

- In practice, it is not easy to verify (perform goodness-of-fit) the distribution of a left-censored data set. Therefore, in this study, emphasis is given on the use of nonparametric UCL95 computation methods.

- Most of the parametric MLE methods assume that there is only one detection limit. But in practice, a left-censored data set often has multiple detection limits. For such methods, the KM method can be used. The ProUCL 4.0 will have these estimation methods for left-censored data sets with multiple detection limits, including the KM method, and also the robust ROS method as described in Helsel (2005).

- For reliable and accurate results, it is suggested that the user should make sure that the data set under study represents a single statistical population (e.g., background reference area, or an AOC) and not a mixture population (e.g., clean and polluted site areas).

- It is recommended to identify all the potential outliers and study them separately. Decisions about the disposition of outliers should be made by all interested members of the project team. Several references are available in the literature to properly identify outliers (Rousseeuw and Leroy (1987) and Singh and Nocerino (1995)) and to partition a mixture sample into component sub-samples (Singh, Singh, and Flatman (1994)). A full chapter will be devoted to population partitioning methods to be included in the Background Guidance Document for CERCLA Sites (EPA (2002)) currently under revision by the NERL-Technical Support Center, EPA Las Vegas.

- Avoid the use of transformations (to achieve symmetry) while computing the upper limits for various environmental applications, as all remediation, cleanup, background evaluation decisions, and risk assessment decisions have to be made using statistics in the original scale. Also, it is more accurate and easier to interpret the results computed in the original scale. The results and statistics computed in the original scale do not suffer from transformations bias.

- Specifically, avoid the use of a lognormal model even when the data appear to be lognormally distributed. Its use often results in incorrect and unrealistic statistics of no practical merit (Examples 6 and 7). Several variations of estimation methods (e.g., robust ROS and FP-ROS on log-transformed data, delta lognormal method) on log-transformed data have been developed and used by the practitioners. Several other variations are discussed in Section 5.8.2. This has caused some confusion among the users of the statistical methods dealing with environmental data sets. The proper use of a lognormal distribution (e.g., how to properly back-transform UCL of mean in the log scale to obtain a UCL of mean in original scale) is not clear to many users, which in turn may result in the incorrect use and computation of an estimate (= UCL95) of the population mean.

- The parameter in the transformed space may not be of interest to make cleanup decisions. The cleanup and remediation decisions are often made in original raw scale; therefore, the statistics (e.g., UCL95) computed in transformed space need to be back-transformed in the original scale. It is not clear to a typical user how to back-transform results in log scale or any other scale obtained using a BC-Type transformation to original raw scale. Moreover, the transformed results often suffer from a significant amount of transformation bias. This was illustrated in Examples 6 and 7 of Section 6.

- It is recommended to avoid the use of equation (3-22) to back-transform estimates from log scale to original scale as illustrated by Section 6. The question now arises – how one should back-transform results from a log-space (or any other transformed space) to the original space? Unfortunately, no defensible guidance is available in the environmental literature to address this question. Moreover, the back-transformation formula will change from transformation to transformation (BC-Type transformations), and the bias introduced by such transformations will remain unknown.

  Therefore, in cases when a data set in the "raw" scale cannot be modeled by a parametric distribution, it is desirable to use nonparametric methods rather than testing or estimating some parameter in the transformed space.

- As noted in Section 8.3.2.1, for the various parametric (gamma and lognormal) and nonparametric skewed distributions, the performance (in terms of coverage percentage for the population mean) of the robust ROS method followed by percentile or BCA bootstrap methods is not better than the KM (Chebyshev) and KM (BCA) UCL methods, as recommended in Sections 8.2.2 and 8.2.3. It is observed that, for left-censor data sets of all sizes and various censoring levels, the robust ROS UCL (both % as well as BCA bootstrap methods) fail to provide adequate coverage for the population mean of highly skewed distributions.

- On page (78) of Helsel (2005), the use of the robust ROS MLE method (Kroll, C.N. and J.R. Stedinger (1996)) has been suggested to compute summary statistics. In this hybrid method, MLEs are computed using log-transformed data. Using the regression model as given by equation (3-21) of Section 3, the MLEs of the mean (used as intercept) and *sd* (used as slope) in the log scale are used to extrapolate the NDs in the log scale. Just like in robust ROS method, all of the NDs are transformed back in the original scale by exponentiation. This results in a full data set in the original scale. One may then compute the mean and *sd* using the full data set. The estimates thus obtained are called robust ROS ML estimates (Helsel (2005) and Kroll and Stedinger (1996)). However, the performance of such a hybrid estimation method is not well known. Moreover, for higher censoring levels, the MLE methods sometimes behave in an unstable manner, especially when dealing with moderately skewed to highly skewed data sets (e.g., with $\sigma > 1.0$).

  – It should be noted that the performance of this hybrid method is unknown.

  – It is not known why this method is called a robust method.

  - The stability of the MLEs obtained using the log-transformed data is doubtful, especially for higher censoring levels.

  - The BCA and (% bootstrap) UCLs based upon this method will fail to provide the adequate coverage for the population mean for moderately skewed to highly skewed data sets.

- If one really wants to use a robust ROS method (Helsel's method), or a robust ROS on MLE method, one can use ProUCL 3.0 (has jackknife, bootstrap, and various other methods) to compute the summary statistics, and a UCL95 to estimate the population mass based upon the full data set obtained using robust ROS or MLE ROS methods. However, the use of these methods is not recommended for highly skewed data sets, as described in Section 8.

- The maximum censoring level considered in the present simulation study is 70%. For data sets having a larger % of nondetects (e.g., 80%, 90%, or 99% nondetects), statistical estimates may not be reliable. Decisions about the use of an appropriate method (e.g., using proportion of NDs) should be made by the risk assessors and regulatory personnel on a site-specific basis. The use of nonparametric methods based upon the proportion of NDs is recommended in such cases (USEPA (2000); Helsel (2005)) with % censoring exceeding 70-80%.

- The DL/2 (t) UCL method does not provide adequate coverage (for any distribution and sample size) for the population mean, even for censoring levels as low as 10%, 15%. This is contrary to the conjecture and assertion (e.g., EPA (2000)) often made that the DL/2 method can be used for lower ($\leq$ 20%) censoring levels.

- The coverage provided by the DL/2 (t) method deteriorates fast as the censoring intensity increases.

- The KM method is a preferred method as it can handle multiple detection limits. Moreover, the various nonparametric UCL95 methods (KM (BCA), KM (z), KM (%), KM (t)) based upon the KM estimates provide good coverages for the population mean.

- For a symmetric distribution (approximate normality), several UCL95 methods provide good coverage (~95%) for the population mean, including the Winsorization mean, Cohen's MLE (t), Cohen's MLE (Tiku), KM (z), KM (t), KM (%) and KM (BCA).

Specific recommendations for the various distributions considered in this report are described as follows.

9.2 Recommended UCL95 Methods for Normal (Approximate Normal) Distribution

- *For normal and approximately normal* (*e.g., symmetric or with sd,* $\hat{\sigma} < 0.5$) *distribution*: The most appropriate UCL95 computation methods for normal or approximately normal distributions are the KM (t) or KM (%) methods. For symmetric distributions, both of these methods perform equally well on left-censored data sets for all censoring levels and sample sizes.

9.3 Recommended UCL95 Methods for Gamma Distribution

- *For highly skewed gamma distributions, G(k, θ), with shape parameter, k ≤ 1:*
  - Use the nonparametric KM (Chebyshev) UCL95 method for censoring levels < 30%,
  - Use the nonparametric KM (BCA) UCL95 method for censoring levels in the interval [30%, 50%), and
  - Use the nonparametric KM (t) UCL95 method for censoring levels ≥ 50%.
- *For moderately skewed gamma distributions, G(k, θ), with shape parameter, 1< k ≤ 2:*
  - For censoring level < 10%, use the KM (Chebyshev) UCL95 method,
  - For higher censoring levels [10%, 25%), use the KM (BCA) UCL95 method,
  - For censoring levels in [25%, 40%), use the KM (%) UCL95 method, and
  - For censoring levels ≥ 40%, use the KM (t) UCL95 method.
- *For mildly skewed gamma distributions, G(k, θ), with k > 2:*
  - Use the KM (BCA) UCL95 method for lower censoring levels (< 20%),
  - For censoring levels in the interval [20%, 40%), use the KM (%) UCL95, and
  - For censoring ≥ 40%, use the KM (t) UCL95 computation method.

9.4 Recommended UCL95 Methods for Lognormal Distribution

- *For mildly skewed data sets with $\hat{\sigma}$ ≤ 1:*
  - For censoring levels (< 20%) and sample of sizes less than 50-70, use the KM (Chebyshev) UCL95,
  - For censoring levels (< 20%) and samples of sizes greater than 50-70, use the KM (BCA) UCL95,
  - For censoring levels in the interval [20%, 40%) and all sample sizes, use the KM (BCA) UCL95, and
  - For censoring level ≥ 40%, use the KM (%) or KM (t) UCL95 method.

- *For data sets with $\hat{\sigma}$ in the interval (1, 1.5]:*

  – For censoring levels < 50% and samples of sizes < 40, use the 97.5% KM (Chebyshev) UCL,

  – For censoring levels < 50% and samples of sizes ≥ 40, use the 95% KM (Chebyshev) UCL, and

  – For censoring levels ≥ 50%, use the KM (BCA) UCL95 for samples of all sizes.

- *For highly skewed data sets with $\hat{\sigma}$ in the interval (1.5, 2]:*

  – For sample sizes < 40, and censoring levels < 50%, use the 99% KM (Chebyshev) UCL,

  – For sample sizes ≥ 40 and censoring levels < 50%, use the 97.5% KM (Chebyshev) UCL,

  – For samples of sizes < 40-50 and censoring levels ≥ 50%, use the 97.5% KM (Chebyshev) UCL, and

  – For samples of sizes ≥ 40-50, and censoring levels ≥ 50%, use the 95% KM (Chebyshev) UCL.

- *Use a similar pattern for more highly skewed data sets with $\hat{\sigma}$ > 2.0, 3.0:*

  – For extremely highly skewed data sets, an appropriate estimate of the EPC term (in terms of adequate coverage) is given by a UCL based upon the Chebyshev inequality and KM estimates. The confidence coefficient to be used will depend upon the skewness. For highly skewed data sets, a higher (e.g., > 95%) confidence coefficient may have to be used to estimate the EPC.

  – As the skewness increases, the confidence coefficient also increases.

  – For such highly skewed (see Tables 3-1 and 3-2) distributions (with $\hat{\sigma}$ > 2.0, 3.0), for lower sample sizes (e.g., < 50-60), one may simply use a 99% KM (Chebyshev) UCL to estimate the population mean, EPC term, and other relevant threshold (e.g., UPL, percentiles) values.

  – For sample sizes greater than 60, one may use a 97.5% KM (Chebyshev) UCL as an estimate of the population mean or mass.

9.5 Recommended UCL95 Methods for Nonparametric Distributions

- *For symmetric or approximately symmetric distribution-free, nonparametric data sets with $\hat{\sigma}$ < 0.5:* Use the same UCL computation methods as for the data sets coming from a normal or an approximate normal (symmetric) population. These methods are summarized in Section 9.2.

- *For skewed distribution-free, nonparametric data sets with $\hat{\sigma} \geq 0.5$:* Most of the recommended UCL computation methods for a lognormal distribution, as described in Section 9.4, do not assume the lognormality of the data set. Therefore, the UCL computation methods, as described in Section 9.4, can be used on skewed nonparametric data sets that do not follow any of the well-known parametric distributions.

Note: All of the methods as recommended in Sections 9.2, 9.3, 9.4, and 9.5 will be available in ProUCL 4.0. Additionally, some UCL95 computation methods based upon MLE and ROS methods have also been incorporated in ProUCL 4.0 for interested users and scientists. The updated version of ProUCL 4.0 is scheduled for release by spring of 2007.

Note: In Table 9-1, phrase "All *n*" represents only valid (e.g., $n > 3$) and recommended ($n > 8$-10) values of the sample size, *n*. As mentioned throughout the report, it is not desirable to use statistical methods on data sets of small sizes (e.g., with $n < 8$-10). However, it should be noted that the sample size requirements and recommendations ($n > 8$-10) as described in this report are not the limitations of the methods considered in this report. One of the main reasons for the recommendation that the sample size should be at least 8-10 is that the estimates and UCLs based upon small data sets, especially with many below detection limit observations (e.g., 30%, 40%, 50%, and more), may not be reliable and accurate enough to draw conclusions for the various environmental applications. Moreover, it should be noted that in order to be able to use bootstrap resampling methods, it is desirable to have a minimum of 10-15 observations (e.g., $n > 10$-15). Therefore, phrase "All *n*" in Table 9-1 should be interpreted as that the sample size, *n* is least 8-10. The software, ProUCL 4.0, will provide appropriate warning messages when a user will try to use a method on data sets of small sizes.

Table 9-1. Recommended UCL95 Computation Methods for Left-Censored Data Sets

Skewness	Sample Size	% ND	95% KM (t)	95% KM (%)	95% KM (Chebyshev)	97.5% KM (Chebyshev)	99% KM (Chebyshev)	95% KM (BCA)
Normal or Approximate Normal (with $\hat{\sigma}$ < 0.5) Distributions								
$\hat{\sigma} < 0.5$	All *n*	> 0%	●	●				
Gamma Distribution								
$\hat{k} \leq 1$	All *n*	< 30%			●			
	All *n*	[30%, 50%)						●
	All *n*	≥ 50%	●					
$1 < \hat{k} \leq 2$	All *n*	< 10%			●			
	All *n*	[10%, 25%)						●
	All *n*	[25%, 40%)		●				
	All *n*	≥ 40%	●					
$\hat{k} > 2$	All *n*	< 20%						●
	All *n*	[20%, 40%)		●				
	All *n*	≥ 40%	●					

Table 9-1. Continued

Skewness	Sample Size	% ND	95% KM (t)	95% KM (%)	95% KM (Chebyshev)	97.5% KM (Chebyshev)	99% KM (Chebyshev)	95% KM (BCA)
Lognormal Distribution								
$\hat{\sigma} \leq 1.0$	*n* ≤ 50-70	< 20%			●			
	n > 50-70							●
	All *n*	[20%, 40%)						●
	All *n*	≥ 40%	●	●				
$1 < \hat{\sigma} \leq 1.5$	*n* < 40	< 50%				●		
	n ≥ 40				●			
	All *n*	≥ 50%						●
$1.5 < \hat{\sigma} \leq 2.0$	*n* < 40	< 50%					●	
	n ≥ 40					●		
	n < 40-50	≥ 50%				●		
	n ≥ 40-50				●			
$\hat{\sigma} > 2.0, 3.0$	*n* < 50-60	> 0%					●	
	n ≥ 60					●		
Nonparametric – Symmetric or Approximate Symmetric								
$\hat{\sigma} < 0.5$	All *n*	> 0%	●	●				
Nonparametric – Moderately Skewed to Highly Skewed								
$0.5 \leq \hat{\sigma} \leq 1.0$	*n* ≤ 50-70	< 20%			●			
	n > 50-70							●
	All *n*	[20%, 40%)						●
	All *n*	≥ 40%	●	●				
$1 < \hat{\sigma} \leq 1.5$	n < 40	< 50%				●		
	n ≥ 40				●			
	All *n*	≥ 50%						●
$1.5 < \hat{\sigma} \leq 2.0$	n < 40	< 50%					●	
	n ≥ 40					●		
	n < 40-50	≥ 50%				●		
	n ≥ 40-50				●			
$\hat{\sigma} > 2.0, 3.0$	*n* < 50-60	> 0%					●	
	n ≥ 60					●		

References

Barber, S. and Jennison, C. 1999. *Symmetric Tests and Confidence Intervals for Survival Probabilities and Quantiles of Censored Survival Data*. University of Bath, Bath, BA2 7AY, UK.

Bechtel Jacobs Company, LLC. 2000. *Improved Methods for Calculating Concentrations used in Exposure Assessment*. Prepared for DOE. Report # BJC/OR-416.

Best, D.J. and Roberts, D.E. 1975. *The Percentage Points of the Chi-square Distribution*. Applied Statistics, 24, pp. 385-388.

Barnett, V. 1976. *Convenient Probability Plotting Positions for the Normal Distribution*. Appl. Statist., 25, No. 1, pp. 47-50, 1976.

Barnett, V. and Lewis T. 1994. *Outliers in Statistical Data*. Third edition. John Wiley & Sons Ltd., UK.

California's Ocean Plan. 2005. *Amendment of the Water Quality Control Plan for Ocean Waters Of California.* EPA, California. State Water Resources Control Board, Sacramento, California. Available at http://www.waterboards.ca.gov/plnspols/oplans.html.

Cohen, A.C., Jr. 1950. *Estimating the Mean and Variance of Normal Populations from Singly Truncated and Double Truncated Samples*. Ann. Math. Statist., Vol. 21, pp. 557-569.

Cohen, A.C., Jr. 1959. *Simplified Estimators for the Normal Distribution When Samples Are Singly Censored or Truncated*. Technometrics, Vol. 1, No. 3, pp. 217-237.

Cohen, A.C., Jr. 1991. *Truncated and Censored Samples*. 119, Marcel Dekker Inc. New York, NY, 1991.

Colorado Water Quality Control Division (WQCD). 2003. *Determination of the Requirement to Include Water Quality Standards-Based Limits in CDPS Permits Based on Reasonable Potential. Procedural Guidance*. Colorado WQCD-Permits Unit, Denver, Colorado. Available at http://www.cdphe.state.co.us/wq/PermitsUnit/rpguide.pdf.

Dempster, A.P., Laird, N.M., and Rubin, D.B. 1977. *Maximum Likelihood from Incomplete Data via the EM Algorithm*. Journal of the Royal Statistical Society, Ser. B39, pp. 1-38.

Dixon, W.J. and J.W. Tukey. 1968. *Approximate Behavior of the Distribution of Winsorized-T* (*Trimming/Winsorization*). Technometrics 10: pp. 83-98.

Dudewicz, E.D. and Misra, S.N. 1988. *Modern Mathematical Statistics*. John Wiley, New York.

Efron, B. 1981. *Censored Data and Bootstrap*. Journal of American Statistical Association, Vol. 76, pp. 312-319.

Efron, B. 1982. *The Jackknife, the Bootstrap, and Other Resampling Plans*. Philadelphia: SIAM.

Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York.

El-Shaarawi, A.H. 1989. *Inferences about the Mean from Censored Water Quality Data*. Water Resources Research, 25, pp. 685-690.

Faires, J. D., and Burden, R. L. 1993. *Numerical Methods*. PWS-Kent Publishing Company, Boston, USA.

Gibbons, R.D. and Coleman, D.E. 2001. *Statistical Methods for Detection and Quantification of Environmental Contamination*. John Wiley, New York.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Gleit, A. 1985. *Estimation for Small Normal Data Sets with Detection Limits*. Environmental Science and Technology, *19,* pp. 1206-1213, 1985.

Gilliom, R.H. and Helsel, D.R. 1986. *Estimations of Distributional Parameters for Censored Trace Level Water Quality Data 1. Estimation Techniques*. Water Resources Research, 22, pp. 135-146.

Golden, N.H., B.A. Rattner, J.B. Cohen, D.J. Hoffman, E. Russek-Cohen, and M.A. Ottinger 2003. *Lead Accumulation in Feathers of Nestling Black-Crowned Night Herons* (*Nycticorax nycticorax*) *Experimentally Treated.* In the field. Environmental Toxicology and Chemistry, Vol. 22. pp. 1517-1525.

Haas, C.H. and Scheff, P.A., 1990. *Estimation of Averages in Truncated Samples*. Environmental Science and Technology, 24, pp. 912-919.

Hampel, F.R. 1974. *The Influence Curve and Its Role in Robust Estimation*. Journal of American Statistical Association, *69,* pp. 383-393, 1974.

Helsel, D.R. 1990. *Less Than Obvious, Statistical Treatment of Data Below the Detection Limit*. ES&T Features Environmental Sci. Technol., Vol. 24, No. 12, pp. 1767-1774.

Helsel, D.R. 2005. *Nondetects and Data Analysis*. Statistics for Censored Environmental Data. John Wiley and Sons, New York.

Hinton, S.W. 1993. *Δ Log-Normal Statistical Methodology Performance.* ES&T Environmental Sci. Technol., Vol. 27, No. 10, pp. 2247-2249.

Hogg, R.V. and Craig, A.T. 1978. *Introduction to Mathematical Statistics*, New York: Macmillan Publishing Company.

Huber, P.J. 1981. *Robust Statistics*. John Wiley, New York.

Johnson, N.L., Kotz, S., and Balakrishnan, N. 1994. *Continuous Univariate Distributions, Vol. 1*. Second Edition. John Wiley, New York.

Johnson, R.A., and Wichern, D.W. 1988. *Applied Multivariate Statistical Analysis*. Prentice Hall.

Kaplan, E.L. and Meier, O. 1958. *Nonparametric Estimation from Incomplete Observations*. Journal of the American Statistical Association, Vol. 53. 457-481.

Kroll, C.N. and J.R. Stedinger. 1996. *Estimation of Moments and Quantiles Using Censored Data*. Water Resources, Vol. 32. pp. 1005-1012.

Land, C.E. 1971. *Confidence Intervals for Linear Functions of the Normal Mean and Variance*. Annals of Mathematical Statistics, 42, pp. 1187-1205.

Land, C.E. 1975. *Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance*. In Selected Tables in Mathematical Statistics, Vol. III, American Mathematical Society, Providence, R.I., pp. 385-419.

Manly, B.F.J. 1997. *Randomization, Bootstrap, and Monte Carlo Methods in Biology*. Second Edition. Chapman Hall, London.

Manly, B.F.J. 2001. *Statistics for Environmental Science and Management*. Chapman & Hall/CRC.

Millard, S.P. 2002. *EnvironmentalStats for S-PLUS*. Second Edition. Springer.

Newman, M.C., Dixon, P.M., and Pinder, J.E. 1990. *Estimating Mean and Variance for Environmental Samples with Below Detection Limit Observations*. Water Resources Bulletin, Vol. 25, No. 4, pp. 905-916.

Perrson, T. and Rootzen, H. 1977. *Simple and Highly Efficient Estimators for A Type I Censored Normal Sample*. Biometrika, 64, pp. 123-128.

ProUCL 3.0. 2004. *A Statistical Software*. National Exposure Research Lab, EPA, Las Vegas, Nevada, October 2004. The software ProUCL 3.0 can be freely downloaded from the EPA Web site: http://www.epa.gov/nerlesd1/tsc/tsc.htm.

Rousseeuw, P.J. and Leroy, A.M. 1987. *Robust Regression and Outlier Detection*. John Wiley, New York.

RPcalc 2.0 2005. *Reasonable Potential Analysis Calculator*, EPA, California. State Water Resources Control Board, Sacramento, California. Available at http://www.swrcb.ca.gov/plnspols/oplans/.

Saw, J.G. 1961. *The Bias for the Maximum Likelihood Estimates of Location and Scale Parameters Given A Type II Censored Normal Sample*. Biometrika, 48, pp. 448-451.

Schneider, H. 1986. *Truncated and Censored Samples from Normal Populations*. Vol. 70, Marcel Dekker Inc., New York, 1986.

Scout. 2002. *A Data Analysis Program*, Technology Support Project. USEPA, NERL-LV, Las Vegas, NV.

SimCensor. 2005. *A Program Developed to Perform Simulation Studies as Summarized in This Report*.

She, N. 1997. *Analyzing Censored Water Quality Data Using a Non-Parametric Approach.* Journal of the American Water Resources Association 33, pp. 615-624.

Shumway, A.H., Azari, A.S., Johnson, P. 1989. *Estimating Mean Concentrations Under Transformation for Environmental Data with Detection Limits*. Technometrics, Vol. 31, No. 3, pp. 347-356.

Shumway, R.H., R.S. Azari, and M. Kayhanian. 2002. *Statistical Approaches to Estimating Mean Water Quality Concentrations with Detection Limits*. Environmental Science and Technology, Vol. 36, pp. 3345-3353.

Singh, A. 1993. *Omnibus Robust Procedures for Assessment of Multivariate Normality and Detection Of Multivariate Outliers*. In Multivariate Environmental Statistics, Patil, G.P. and Rao, C.R., Editors, pp. 445-488, Elsevier Science Publishers.

Singh, A., A.K. Singh, and G.T. Flatman. 1994. *Estimation of Background Levels of Contaminants*. Journal of Mathematical Geology 26(3):361.

Singh, A. and Nocerino, J.M. 1995. *Robust Procedures for the Identification of Multiple Outliers*. Handbook of Environmental Chemistry, Statistical Methods, Vol. 2.G, pp. 229-277. Springer Verlag, Germany.

Singh, A.K., Singh, A., and Engelhardt, M. 1997. *The Lognormal Distribution in Environmental Applications*. Technology Support Center Issue Paper, 182CMB97, EPA/600/R-97/006.

Singh, A.K., Singh, A., and Engelhardt, M. 1999. *Some Practical Aspects of Sample Size and Power Computations for Estimating the Mean of Positively Skewed Distributions in Environmental Applications*. EPA/600/S-99/006.

Singh, A. and Nocerino, J.M. 2002. *Robust Estimation of the Mean and Variance Using Environmental Data Sets with Below Detection Limit Observations,* Vol. 60, pp. 69-86.

Singh, A., Singh, A.K., and Iaci, R.J. 2002. *Estimation of the Exposure Point Concentration Term Using a Gamma Distribution,* EPA/600/R-02/084, October 2002.

Singh, A. and Singh, A.K. 2003. *Estimation of the Exposure Point Concentration Term* (*95% UCL*) *Using Bias-Corrected Accelerated* (*BCA) Bootstrap Method and Several Other Methods for Normal, Lognormal, and Gamma Distributions*. Draft EPA Internal Report.

Singh, A. 2004. *Computation of an Upper Confidence Limit* (*UCL*) *of the Unknown Population Mean Using ProUCL Version 3.0*. Part I. Download from: www.epa.gov/nerlesd1/tsc/issue.htm.

Staudte, R.G. and Sheather, S.J. 1990. *Robust Estimation and Testing*. John Wiley, New York.

Tiku, M.L. 1967. *Estimating Mean and the Standard Deviation from a Censored Normal Sample*. Biometrika 54. pp. 155-165.

Tiku, M.L. 1971. *Estimating the Mean and the Standard Deviation from Two Censored Normal Samples*. Biometrika 58. pp. 241-243.

UNCENSOR 5.1. 2003. *A Statistical Program for Left-censored Data Sets*. University of Georgia. Savannah River Ecology Laboratory.

USEPA. 1991. *Technical Support Document for Water Quality Based Toxics Control*. Office of Water Enforcement and Permits, Washington, D.C. March 1991.

USEPA. 1992. *Statistical Analysis of Ground-water Monitoring Data at RCRA Facilities*. Addendum to Interim Final Guidance. Washington, D.C.: Office of Solid Waste. July 1992.

USEPA. 1993. SW-846, Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. Third Edition. Available at SW-846 on-line.

USEPA. 2002a. *Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites*. EPA 540-R-01-003-OSWER 9285.7-41. September 2002.

USEPA. 2002b. *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*. OSWER 9285.6-10. December 2002.

USEPA. 2004. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities. Unified Guidance Document* (*UGD*). Volumes I, II, and III. Peer Review Draft. Office of Solid Waste. September 2004.

USEPA. 2006. *Data Quality Assessment: Statistical Methods for Practitioners,* EPA QA/G-9S. EPA/240/B-06/003. Office of Environmental Information, Washington, D.C. Download from: http://www.epa.gov/quality/qs-docs/g9s-final.pdf.