

# **Appendix I      Use of Air Monitoring Data to Develop Estimates of Exposure Concentration (Data Analysis and Reduction)**

## **Table of Contents**

1.0	Introduction .....	<a href="#">1</a>
2.0	Data Treatment and Handling of Non-Detects .....	<a href="#">2</a>
3.0	Statistical Methods: Characterization of Concentration Data .....	<a href="#">4</a>
	References .....	<a href="#">8</a>



## 1.0 Introduction

This appendix discusses the process of air monitoring data analysis and reduction, the goals of which are to (1) extract and summarize air monitoring data needed for the risk assessment, (2) use the data to develop estimates of exposure concentration (EC), and (3) present the results of the air monitoring study in an informative and understandable format. In short, this Appendix describes how to take the refined air monitoring data sets developed according to the processes described in Appendix H and use them to develop estimates of exposure concentration. Standard computer software packages, such as Microsoft Excel® or the Statistical Analysis System,® may be used to generate summary statistics for each chemical and monitoring location. Summary statistics should include:

- The frequency of detection, or the proportion of total valid measurements collected which were present at or above the respective sample quantitation limit (SQL) and including detections marked with certain data qualifier (e.g., “J” values - see Appendix H);
- The range of concentrations detected (highest and lowest concentrations measured for each chemical at each monitoring site – including J values);
- The statistical description of the data (e.g., normally distributed, log-normally distributed), based on standardized statistical tests;
- The range of sample quantitation limits (SQLs); and
- An arithmetic mean value, the standard deviation, the median value (i.e., 50th percentile), and the 95th percentile upper confidence limit (95% UCL) of the arithmetic mean.

### Tentatively Identified Compounds (TICs)

As noted in Appendix H, TICs are chemicals identified in the laboratory, but which cannot be identified with complete accuracy. Given that there is not certainty as to their identify (and because, there often is no toxicity data for them), TICs are often assess only qualitatively in the risk assessment. The level of detail applied to TICs depends on their tentative identification (are they known toxic compounds), their concentration, known sources, and frequency of detection. Depending on the answers to these questions, the analyst may recommend that re-sampling be performed to try to more accurately determine the nature of the TICs.

The mathematical formulas and procedures for calculating these summary statistics are provided in Section 3 below.

Statistical analysis of air monitoring data may be conducted using standard methods such as those outlined in EPA’s *Guidance for Data Quality Assessment - Practical Methods for Data Analysis*.<sup>(1)</sup> This manual provides a detailed description of the formulae that should be used in estimating the parameters mentioned above, and reviews issues associated with data treatment (e.g., treatment of non-detects, use of J-qualified data). EPA’s *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*<sup>(2)</sup> is also an important reference to consider when evaluating air monitoring data for exposure assessments. Readers are encouraged to review both of these document prior to using monitoring data to calculate exposure concentrations.

## Data Qualifiers

Having obtained a monitoring result, it is necessary to assign a qualifier to it so decision-makers can understand the quality of the result and, hence, the role the result might play in decisions (a more complete discussion of data qualifiers is provided in Appendix H).

- **U Flag.** If the value is below the MDL, the result should be flagged as <MDL or as U or “undetected.” This indicates that it cannot be determined, within the limits described in the DQOs, that the compound is present in the sample. (Note, however, that some labs flag data below the SQL as U, even though they actually detect it. It is important to work through such details with a laboratory prior to analysis of samples.)
- **J Flag.** If the result is above the MDL, but less than the SQL, the result should be flagged as J or “estimated concentration.” This indicates that the compound was detected in the sample, and can be quantified, but not within the limits on accuracy described in the DQOs.
- **R Flag.** If there are significant problems with the sample (e.g. improper calibration, or extensive holding time, or very low recovery efficiency), the result should be flagged as R or “unusable.” This might occur, for example, if calibration procedures are judged inadequate. If the compound is of interest, and/or other results suggest the potential for significant concentration, then re-sampling or re-analysis is usually necessary.

Interpretation of these flagged results in the context of a risk assessment is described in Chapters 5 and 6 of the EPA document *Guidance for Data Usability in Risk Assessment, Part A* (Publication 9285.7-09A, Washington, DC, April 1992; available at

[www.epa.gov/superfund/programs/risk/datause/parta.htm](http://www.epa.gov/superfund/programs/risk/datause/parta.htm)). Another excellent source is Exhibits 5-4 and 5-5 in EPA’s *Risk Assessment Guidance for Superfund (RAGS) Part A*, available at [www.epa.gov/oerrpage/superfund/programs/risk/ragsa/index.htm](http://www.epa.gov/oerrpage/superfund/programs/risk/ragsa/index.htm). Appendix H of this volume also discusses this subject in some detail.

## 2.0 Data Treatment and Handling of Non-Detects

Calculation of summary descriptive statistics (arithmetic mean, standard deviation, median, 95% upper confidence limit) requires resolution of certain issues regarding the treatment of sampling data. Specifically, assumptions must be made regarding:

- Treatment of duplicate samples;
- Treatment of instances in which chemicals are not detected; and
- Use of measurements in which the identity of a chemical is certain but its concentration is estimated with some uncertainty (often reported as “J-qualified” data).

Duplicate samples refer to the simultaneous collection or analysis of multiple (usually two) samples under conditions that are kept as similar as possible. Field duplicates usually refer to separate samples collected side-by-side in the field, while laboratory duplicates involve separately analyzing portions of the contents of a single sample. Both types of duplicates serve the similar purpose of providing a sense of the reliability, reproducibility, and precision of measurements. Ideally, duplicate samples should yield the same results. Large differences in the results of duplicate measurements potentially indicate uncertainty in data quality.

In general, once it is clear that there are no issues with field duplicate samples, they should be treated as a single sample by simply averaging their results. In cases where a chemical is detected in one but not both duplicates (or the data is J-qualified), the chemical should be assumed to be present and the two values should be averaged using the procedure for handling non-detects as described below.

When a chemical is not detected in any sample at a monitor, that chemical can usually be removed from further consideration if there are no known problems with the method, the method meets DQOs, and there is no reason to suspect that the chemical should have been detected (e.g., there are no known sources, and the chemical was also not found at other monitors). In some instances, the monitoring methodology (or interferences by other substances) do not allow for the detection of a substance, even when it is present. The assessors must weigh these types of evidence when deciding to drop a chemical from further consideration.

Various procedures have been used in risk assessments to treat non-detects (i.e., samples in which the chemical concentration is not present at or higher than the sample quantification limit (SQL)), ranging from the assumption that the chemical is absent (i.e., the true concentration is zero) to the assumption that the chemical was present in a sample at a level infinitesimally beneath the SQL (i.e. very close to the SQL and so essentially equal to the SQL). Some algorithms differentiate assignment of values to non-detects based upon the frequency of a chemical's detection. For example, if a chemical is detected in almost all samples, a concentration equal to (or some fraction of) the analytical SQL is assigned to non-detects, but if the chemical is detected in few or no samples, a concentration of zero is assumed for non-detects. In general, the strategy described below may be used to address the issue of non-detects. References 1 and 2 provide more information on this subject and analysts are encouraged to become familiar with both of these documents prior to beginning data analysis. Also note that the generic upon which the procedure described below is based assumes approximately 30 or more samples collected over the course of a year are being averaged to develop an estimate of long term exposure concentration; however, air toxics monitoring sampling schemes usually collect samples on at least a one-in-six day schedule, giving the analyst approximately 60 or more samples to work with. Sampling frequencies are sometimes even greater.

**The MDL or the SQL: Which One Should I Use for Risk Assessment?**

When including non-detected data in the averaging processes described on this page, one may either include the non-detected sample as ½ the MDL or ½ the SQL. The MDL *is not appropriate* for this task because it is a statistical measure developed by each lab for each analytical instrument and can fluctuate from day to day. In other words, it is not a stable measure of true detection “limit.” In addition, many labs that actually do detect a chemical in a sample at levels less than the quantitation limit do not routinely report the detection because they cannot accurately quantitate its concentration). It is for these two reasons that ½ the SQL is used when including nondetected samples in the averaging process. This holds even when the lab in question routinely reports J-valued data.

- If less than 15% of the monitored concentrations of a given chemical at a given location are below the SQL, then a value equal to ½ of the respective SQL is assigned to these concentrations and these values are used in the calculation of summary statistics as described below.

- If greater than 90% of the monitored concentrations of a given chemical at a given location are less than the respective SQL, no estimation of the statistical descriptors is undertaken initially. If concentrations were only detected on a limited number of days (i.e., 1 to 3 days) then an investigation may be undertaken to assess the potential sources for these chemicals and the validity of the measurements. A knowledgeable statistician can help determine an appropriate method for developing summary statistics from such a data set, if appropriate.
- If between 15% and 90% of the monitored concentrations of a given chemical at a given location are greater than the respective SQL, then a value equal to ½ of the respective SQL is assigned to these concentrations and these values are used in the calculation of summary statistics as described below. For chemicals in this group that end up contributing significantly to risk, a knowledgeable statistician may reevaluate the data according to the procedures in appropriate guidance (e.g., those provided in references 1 and 2).

### **3.0 Statistical Methods: Characterization of Concentration Data**

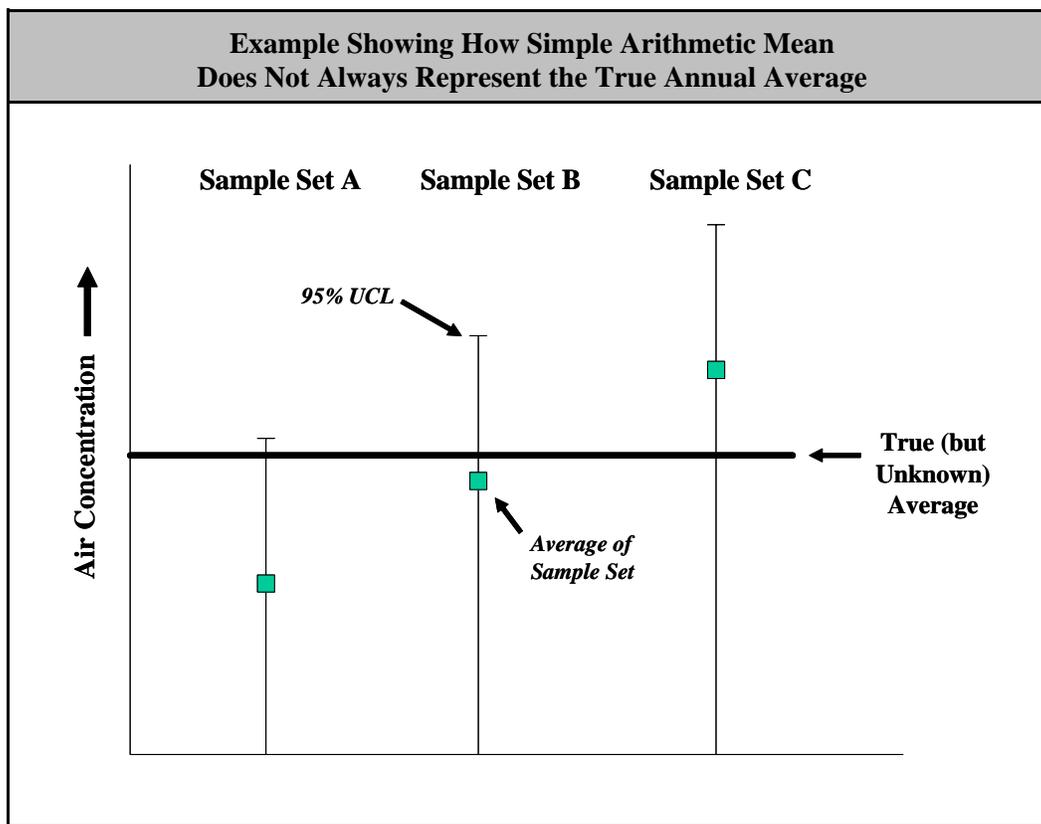
One method to estimate the long-term annual average concentration would be to calculate a simple arithmetic mean for each analyte/monitor combination. The arithmetic mean, or average is constructed from discrete sample measurements taken at the monitor over time. As noted previously, constraints on resources almost always place limits on the amount of sampling possible (e.g., air toxics samples usually cannot be collected every day). Instead, samples are usually collected roughly one out of every six days and in a manner to eliminate obvious sources of bias (e.g., samples are not uniformly collected on the same day of the week, or only on weekdays or only on weekends). In addition, collecting samples for a year allows for an evaluation of seasonal variability.

All factors being equal, one would expect the sampling results from such a monitoring program to contain equal probabilities of sampling on days when pollutant concentrations may have been relatively high as on days when pollutant concentrations may have been relatively low (or on days when meteorological conditions were conducive to high ground-level concentrations and days when they were not). Since samples are usually not collected every single day, however, one cannot be absolutely certain that all possible conditions were sampled equally. The arithmetic mean concentration is thus subject to uncertainty due to a number of factors, including:

- Daily variability in concentrations;
- The ability to measure only a finite number of instances from the distribution of concentrations over time; and
- Potential inaccuracy in individual measurements of concentrations.

This uncertainty produces a result in which the simple arithmetic mean of sampling results may underestimate, approach, or overestimate the true annual average. (The example below illustrates how three different monitoring data sets taken at the same monitor may result in an average concentration that underestimates, overestimates, or is close to the true long term average concentration.) Given this uncertainty in the use of the arithmetic mean concentration to describe “average” exposure concentration, the 95% Upper Confidence Limit of the mean (95% UCL) is commonly used as a public health protective estimate of the true annual average. Proceeding in this manner is likely to overestimate the true long-term average exposure;

however, this method virtually obviates the risk of underestimating the true exposure. EPA's Superfund program has routinely used this procedure to evaluate exposures at hazardous sites and this process has garnered long term acceptance as a public health protective approach, in light of the uncertainties.



### Distributional Analysis

To calculate the 95% UCL for a chemical data set from a monitor, it is necessary to understand its underlying statistical distribution, including whether the sampling results are normally or lognormally distributed. Once the analysis goes beyond these commonly understood distributional types, the level of statistical sophistication can increase substantially. EPA's Office of Air Quality Planning and Standards (OAQPS) has developed the following pragmatic strategy to evaluate the distribution of monitoring data sets; however, other approaches are available (see references 1 and 2). Specifically, EPA suggests the following procedure:

- Inspect each data set for normality using standard test procedures (e.g., Shapiro-Wilk Test, Komolgorov-Smirnoff Test, or Filibens Test). If the assumption of normality holds, then the summary descriptive statistics, including the 95% UCL, should be calculated as described below with the equations based on the statistical assumption of a normal distribution.
- If the data are not normally distributed, then they are *presumed to be lognormal* and are log-transformed by taking the natural logarithm of the measured concentrations. The assumption of normality is then used to test the transformed data. If the assumption of normality holds for the transformed data, the summary descriptive statistics, including he

95% UCL, are developed with the transformed data using the equations based on the statistical assumption of a lognormal distribution.

- If the transformed data are not lognormal, they may be treated initially as lognormal. For chemicals in this group that significantly contribute to risk, a knowledgeable statistician may reevaluate the data (e.g., according to the procedures suggested in References 1 and 2).

The use of this simple and pragmatic approach to data analysis allows most scientists and engineers with a basic background in statistics to perform these analyses without access to advanced statistical analysis resources. Presuming a data set is lognormally distributed generally results in a 95% UCL that is conservative and, thus, public health protective. Only those chemicals that the initial risk characterization identifies as being significant risk drivers would be reevaluated with more robust statistical procedures, depending on the needs of the risk manager.

### **STATISTICAL FORMULAS**

The following Exhibits provide the basic equations for developing the 95% UCL for chemical data sets that are either normally distributed (Exhibits 1 and 2) or lognormally – or presumed to be lognormally – distributed (Exhibits 3 and 4). The Student's *t* and *H* statistics that are needed to perform these calculations are available in Gilbert's 1987 book *Statistical Methods for Environmental Pollution Monitoring*.<sup>(3)</sup>

#### ***Normally Distributed Data Sets***

##### **Exhibit 1. Directions for Computing UCL for the Mean of a Normal Distribution – Student's *t***

Let  $X_1, X_2, \dots, X_n$  represent the  $n$  randomly sampled concentrations.

STEP 1: Compute the sample mean  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

STEP 2: Compute the sample standard deviation  $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$

STEP 3: Use a table of quantiles of the Student's *t* distribution to find the  $(1-\alpha)^{\text{th}}$  quantile of the Student's *t* distribution with  $n-1$  degrees of freedom. For example, the value at the 0.05 level with 40 degrees of freedom is 1.684. A table of Student's *t* values can be found in Gilbert (1987, page 255, where the values are indexed by  $p = 1-\alpha$ , rather than  $\alpha$  level). The *t* value appropriate for computing the 95% UCL can be obtained in Microsoft Excel<sup>®</sup> with the formula `TINV ((1-0.95)*2, n-1)`.

STEP 4: Compute the one-sided  $(1-\alpha)$  upper confidence limit on the mean

$$UCL_{1-\alpha} = \bar{X} + t_{\alpha, n-1} s / \sqrt{n}$$

**Exhibit 2. An Example Computation of UCL for a Normal Distribution – Student's  $t$**

25 VOC samples were collected from an air monitoring station and analyzed for a specific chemical. The values observed are 228, 552, 645, 208, 755, 553, 674, 151, 251, 315, 731, 466, 261, 240, 411, 368, 492, 302, 438, 751, 304, 368, 376, 634, and 810  $\mu\text{g}/\text{m}^3$ . It seems reasonable that the data are normally distributed, and the Shapiro-Wilk  $W$  test for normality fails to reject the hypothesis that they are ( $W = 0.937$ ). The UCL based on Student's  $t$  is computed as follows:

STEP 1: The sample mean of the  $n = 25$  values is  $\bar{x} = 451$

STEP 2: The sample standard deviation of the values is  $s = 198$

STEP 3: The  $t$ -value at the 0.05 level for 25-1 degrees of freedom is  $t_{0.05,25-1} = 1.710$

STEP 4: The one-sided 95% upper confidence limit on the mean is therefore:

$$95\% \text{ UCL} = 451 + (1.710 \times 198 / \sqrt{25}) = 519$$

**Lognormally Distributed Data**

**Exhibit 3. Directions for Computing UCL for the Mean of a Lognormal Distribution – Land Method**

Let  $X_1, X_2, \dots, X_n$  represent the  $n$  randomly sampled concentrations.

STEP 1: Compute the arithmetic mean of the log-transformed data  $\overline{\ln X} = \frac{1}{n} \sum_{i=1}^n \ln(X_i)$

STEP 2: Compute the associated standard deviation  $s_{\ln X} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\ln(X_i) - \overline{\ln X})^2}$

STEP 3: Look up the  $H_{1-\alpha}$  statistic for sample size  $n$  and the observed standard deviation of the log-transformed data. Tables of these values are given by Gilbert (1987, Tables A-10 and A-12) and Land (1975).

STEP4: Compute the one-sided  $(1-\alpha)$  upper confidence limit on the mean

$$UCL_{1-\alpha} = \exp(\overline{\ln X} + s_{\ln X}^2 / 2 + H_{1-\alpha} s_{\ln X} / \sqrt{n-1})$$

#### Exhibit 4. An Example Computation of UCL for a lognormal Distribution – Land Method

31 VOC samples were collected from an air monitoring stations and analyzed for a specific chemical. The values observed are 2.8, 22.9, 3.3, 4.6, 8.7, 30.4, 12.2, 2.5, 5.7, 26.3, 5.4, 6.1, 5.2, 1.8, 7.2, 3.4, 12.4, 0.8, 10.3, 11.4, 38.2, 5.6, 14.1, 12.3, 6.8, 3.3, 5.2, 2.1, 19.7, 3.9, and 2.8  $\mu\text{g}/\text{m}^3$ . Because of their skewness, the data may be lognormally distributed. The Shapiro-Wilk  $W$  test for normality rejects the hypothesis, at both the 0.05 and 0.01 levels, that the distribution is normal. The same test fails to reject at either level the hypothesis that the distribution is lognormal. The UCL on the mean based on Land's  $H$  statistic is computed as follows:

STEP 1: Compute the arithmetic mean of the log-transformed data  $\overline{\ln X} = 1.8797$

STEP 2: Compute the associated standard deviation  $s_{\ln X} = 0.8995$

STEP 3: The  $H$  statistic for  $n = 31$  and  $s_{\ln X} = 0.90$  is 2.31

STEP4: The one-sided 95% upper confidence limit on the mean is therefore:

$$95\% \text{ UCL} = \exp(1.8797 + 0.8995^2 / 2 + 2.31 \times 0.8995 / \sqrt{31 - 1}) = 14.4$$

It is statistically possible for the 95% UCL confidence limit of the mean to exceed the maximum measured concentration for a chemical. If this exceeding occurs, the maximum concentration of the chemical is commonly used in place of the 95<sup>th</sup> percentile upper confidence limit as the exposure concentration, with certain caveats (see reference 2).

#### References

1. U.S. Environmental Protection Agency. 1998. *Guidance for Data Quality Assessment – Practical Methods for Data Analysis*. EPA/QA-G9, QA97 Version, EPA 600/R96/O84, Washington, DC, January 1998; available at [www.epa.gov/swrust1/cat/epaqag9.pdf](http://www.epa.gov/swrust1/cat/epaqag9.pdf).
2. U.S. Environmental Protection Agency. 2002. *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*. Office of Emergency and Remedial Response, Washington, DC, December 2002. OSWER 9285.6-10, available at <http://www.epa.gov/superfund/programs/risk/ragsa/ucl.pdf>.
3. Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. John Wiley & Sons, New York, NY.