NSF Standard(s) Impacted: 40, 245, 350, possibly others

Background:
Provide a brief background statement indicating the cause and nature of concern, the impacts identified relevant to public health, public understanding, etc, and any other reason why the issue should be considered by the Committee. Reference as appropriate any specific section(s) of the standard(s) that are related to the issue.

Guidance is needed to assist certifiers on how they should handle chemical concentration data lower than the analytical method detection limit.

As researchers increasingly investigate trace substances in the world’s soil, air and water, they repeatedly find concentrations that are lower than the limits deemed reliable enough to report as a numeric value. Data sets containing values below the limit of detection are known as ‘censored data sets’. Such data sets are encountered regularly in the environmental contamination field.

A nondetectable result from an analytical laboratory analysis indicates the lowest concentration that can reliably be distinguished from zero, but is not quantifiable with acceptable precision, making the reported concentration an estimate for the parameter in question. A related parameter is the quantitation limit, which is the lowest concentration that can not only be detected, but also quantified with a specified degree of precision. At or above the quantitation limit, the parameter or analyte is both proven present and measured reliably. On the one hand, it would be unjustified and overly optimistic to assume that nondetect sample results are assumed absent (zero) for that parameter. It would also be highly conservative and inconsistent with the use of the best available science to assign the method detection limit as the value of all nondetect sample results. Such a method would produce a mean concentration that is biased high.

As effluent limitation standards become more stringent and equipment manufacturers respond to the opportunity to provide products meeting these standards, the likelihood of encountering nondetectable results will only increase.

The goal of an adopted protocol for handling data below method detection limits is to make decisions that are transparent, accountable and scientifically defensible regarding the reported assessments of the performance of products. The perfect time to make decisions on the appropriate handling of non-detect results is before (a priori) data are reported by the laboratory.

Recommendation:
Clearly state what action is needed: e.g., recommended changes to the standard(s) including the current text of the relevant section(s) indicating deletions by use of strike-out and additions by highlighting or underlining; e.g., reference of the issue to a Task Group for detailed consideration; etc.

The Joint Wastewater Committee should discuss the best way to address values less than the method detection limit at their next in-person meeting. A common, relatively simple approach would be to substitute half the method detection limit for all non-detect data. Statistical researchers (see supplementary material) caution that when non-detect data is greater than 15% of the entire data set that substitution (with zero, method detection limit or half method detection limit) may not be appropriate. Instead, more involved adjustment such as Cohen’s Method, Aitchison’s Method, Double linear interpolation or test of proportions should be conducted instead. If the Joint Committee wishes to explore the second path in detail, the proposer recommends NSF International engages the services of a knowledgeable statistician. Alternatively, the USEPA regularly evaluates datasets having nondetect
results and may be able to assist in determining how to apply its customary data analysis methods from the Superfund program to wastewater applications. In either case, an approved written protocol is needed.

**Supplementary Materials (photographs, diagrams, reports, etc.):**
If not provided electronically, the submitter will be responsible to have sufficient copies to distribute to committee members.

EPA (2016) Data quality assessment: statistical methods for practitioners. EPA QA/G-9s (cover and pages 130-136 (4.7 Values below detection limits) provided


*I hereby grant NSF International the non-exclusive, royalty free rights, including non-exclusive, royalty free rights in copyright; in this item and I understand that I acquire no rights in any publication of NSF International in which this item in this or another similar or analogous form is used.*

Signature*: Kevin M. Sherman
Company: Presby Environmental, a subsidiary of Infiltrator Water Technology, Inc.
Telephone Number: 603-631-5452   E-mail: kevin.sherman@presbyeco.com
Is this a revision of a previous Issue Paper (if yes put original issue number): __No____
Submission Date: August 8, 2019

**Please submit to: Joint Committee Secretariat, Jason Snider at jsnider@ NSF.org**

*Type written name will suffice as signature*
Title: NON-DETECT DATA IN ENVIRONMENTAL INVESTIGATIONS

Author(s): J. Wendelberger
K. Campbell

Submitted to: American Statistical Association
Toronto, Canada
August 14 - 18, 1994
1. INTRODUCTION

Environmental data frequently contain values that are below detection limits. Values that are below detection limits are reported as being less than some reported limit of detection, rather than as actual values. Detection limits may vary for sample and reference datasets or even for individual observations. The implications of these nondetects in the summary, analysis, and interpretation of environmental data are explored. Alternative approaches for handling nondetect data are examined, with practical considerations being a key element in the selection of appropriate methodology.

Two general approaches may be used in the statistical analysis of data that contain nondetect values: 1) The nondetect values may be replaced using one of a variety of replacement methods, or 2) statistical techniques may be employed which can handle nondetect data. One of the most commonly used replacement methods is to substitute each nondetect value by half of its detection limit. Other commonly used replacement values are zeros or the detection limits. To avoid clumping of replaced values in cases where there are several nondetect values that share a common detection limit, values may be spaced evenly from 0 to the detection limit or according to some specified probability distribution.

This paper will review selected work that has been conducted on the handling of nondetects...
and examine specific environmental decisions in the presence of nondetect values.

2. NON-DETECT VALUES

Currie (1984, 1968) has noted that there are many different definitions of nondetect values in the scientific literature and has attempted to clarify the statistical error structure implied by the presence of nondetect values. Lambert, Peterson and Terpenning (1991) examined the nature of nondetects using extensive information on the actual measured values associated with nondetect values and showed that common reporting practices can throw out valuable information. In practice, the exact nature of the nondetect values may not be known. Information of the type used by Lambert et al may not be available. Instead, the statistician is left with the task of trying to make decisions in the situation where some of the values are only known to be less than specified values. In some cases, clear-cut decisions may be possible from the existing information. However, in borderline cases, additional values and additional information about the nature of the nondetect values may be worth obtaining to make better-informed decisions. Porter, Ward and Bell (1988) discuss the merits of having actual measured values rather than values recorded as nondetects.

In actual practice, the source of the chemical analysis measurements may be sufficiently far removed that the exact nature of the nondetect values is unknown. Instead, the analyst only knows that the measured value is less than or equal to the specified detection limit. EPA (1989, 1992a, 1992b) describes some of the different types of nondetect values that can occur with analytical data and makes some limited suggestions about the treatment of this type of data.

In the statistical literature, the term censored data is applied to data for which values above or below a certain threshold are not able to be measured. Cohen (1959) has proposed an approach...
for computing maximum likelihood estimators for the mean and variance of normally distributed data in the presence of censored values.

Several authors have examined censoring specifically in the context of environmental data. Gleit (1985) discusses estimation for small normal data sets with detection limits and compares several estimators via simulation. Hinton (1993) examines the performance of the delta lognormal method for estimating the mean of environmental data. This method assumes that the underlying distribution of values is a mixture of nonzero lognormal values and zeroes. Newman and Dixon (1990) describe a computer program with 5 different methods for estimating the mean and standard deviation from normal and lognormal data. Owen and DeRouen (1980) compare the robustness to deviations from model assumptions of two different methods for estimating the mean for lognormal data containing zeroes and left-censored values. Shumway, Azari and Johnson describe a method for estimating mean concentrations for environmental data with nondetects that uses transformations of the data. Travis and Land (1990) discuss the use of log-probit analysis to estimate the mean from a lognormal distribution with nondetects. Gillom (1984) looks at the effect of nondetects on the ability to detect trends.

In the following sections, specific environmental decisions will be considered where nondetects may need to be addressed.

3. DATA SUMMARY

Environmental investigations can result in the collection of large volumes of chemical analysis data. A single environmental sample may be divided up and analyzed for hundreds of different radiological or chemical constituents. Depending on the size of the site under investigation, the number of samples collected may range from less than 10 to several hundred.
An important step in the reporting and analysis of large quantities of data is to summarize the individual values. The presence of nondetect values can make summarization a non-trivial task, because common sample statistics such as the median, mean, standard deviation, maximum and minimum are not clearly defined for the case where nondetect values are present.

To convey meaningful information, the summary must either provide separate information for detect and nondetect data, follow a specified procedure for replacement of nondetects by proxy values prior to the computation of sample statistics, or must use statistical techniques to provide estimates of sample quantities based on the full dataset. Helsel (1990) and Helsel and Cohn (1988) discuss the estimation of descriptive statistics for data with nondetects.

How the data is intended to be used may help determine the most appropriate method of handling the nondetect values. As urged by EPA (1989), "Do not simply omit the nondetected results..." The authors have encountered cases where individuals have simply omitted all nondetects from a data summary. Clearly, this practice can seriously bias any summary statistics computed from the data. The method of handling nondetects may also depend on what type of constituent is being examined. When a number of closely related constituents are being considered, the replacement of nondetects may be achieved from a multiple constituent standpoint. For example, for polycyclic aromatic hydrocarbons (PAHs), one method that has been proposed is to set all nondetects to zero if there are no detected PAHs in a given sample and to set all nondetects to their corresponding detection limits if one or more of the PAHs are present above a detection limit.

4. COMPARISONS TO FIXED VALUES

In some cases, individual values or sample means must be compared to specified fixed
levels to make a decision about whether further action is required at a site. At Los Alamos National Laboratory, a large number of sites are currently being investigated to determine whether contamination is present. In order to focus available resources on sites that require remediation, an initial screening is performed. Samples are collected at the sites and analyzed for a large number of constituents. The maximum values for each constituent are then compared to risk-based Screening Action Levels (SALs) to determine whether additional evaluation is required. In most cases, the detection limits are well below the SALs so nondetect values do not impact screening decisions. However, in some cases, the detection limits may be near the SALs or even above the SALs making comparisons difficult. When a simple comparison can not be made, alternative evaluation methods may be used. For example, historical knowledge of the site and the frequency of detected values in the dataset may be considered. Alternatively, comparisons to relevant background levels or risk assessment methodology, which are described in the next two sections, may be required.

5. COMPARING POPULATIONS

Non-detect values can pose an especially difficult problem when the goal is to compare two different populations. For example, data collected from an area suspected of possible contamination might be compared to data collected from some suitable background area that is similar, but known to be uncontaminated. Since the values sampled from the two populations may involve different censoring mechanisms and different limits of detection, care must be taken in order to arrive at valid conclusions. Gilbert and Simpson (1990, 1992) discuss several methods for comparing sample and reference populations and provide some advice on dealing with nondetects.
One simple approach is to apply standard parametric techniques with nondetects replaced by proxy values. A preferred approach may be to use nonparametric approaches. Hipel (1988) and Helsel (1987) discuss the advantages of using nonparametric techniques for the assessment of environmental data. Some nonparametric techniques can handle nondetect values directly. Others may require replacement of the nondetect values, but the exact values will have less impact than for parametric approaches. In other cases, the nonparametric techniques can handle nondetect values if certain assumptions may be made, such as all of the nondetect values being less than all of the values above detection limits.

The Wilcoxon test is a nonparametric test that may be used to compare samples from two populations. (See for example, Conover (1980).) The Wilcoxon test is effective for detecting location shifts of a distribution. The Wilcoxon test uses the ranks obtained from the dataset formed by combining the two samples. The test statistic is the sum of the ranks corresponding to one of the samples. For moderate to large sample sizes, a normal approximation may be used. Gehan (1965) proposed an extension of the Wilcoxon statistic for comparing two samples to handle censored data in the singly-censored case where all values are censored at the same value. Efron (1967) has proposed an alternative test statistic for the two sample problem with censored data.

When there are nondetects with multiple detection limits present, a multiple detection limit approach is required. Millard and Deverel (1988) discuss methods for comparing two sites when there are multiple detection limits and provide a comparison of various techniques based on a simulation study.

Another approach that may be especially appropriate for environmental applications is a
quantile test developed by Johnson, Verrill and Moore (1987). This method examines the upper tail behavior of two populations by computing the probability that \( k \) out of the \( n \) largest values from the combined data set of \( n_1 + n_2 \) values would come from one of the populations if the two populations had the same distribution. This method will detect differences in the upper tail of the distribution. In environmental applications, it is often these extreme values which are of most interest. Since nondetect values tend to occur in the lower end of the distribution, this method can be used in the presence of nondetect data, provided the nondetect values are not among the \( n \) largest values. Even if some of the nondetect values are in the set of \( n \) largest values, the method could still be used if a replacement method was used to provide proxy values for the nondetects.

6. RISK ASSESSMENT

Risk assessment involves summarization of contaminant concentrations over areas called exposure units that represent physical area sizes encountered by individuals under a given land use scenario. The risk assessment process uses sample information about average contaminant concentrations to determine whether a particular area poses a human health risk. The 95% upper confidence limit on the mean is typically used as a reasonable maximum concentration for an exposed individual. (See EPA (1989).)

To examine the impact of nondetect values on risk assessment, consider the formula used in computing the 95% upper confidence limit on the mean. This formula involves not only the sample mean, but also the sample standard deviation. Replacement of nondetect values by their detection limits is generally regarded as conservative when estimating the mean because it will provide a mean estimate that is greater than or equal to the mean that would be obtained if the
nondetect values were known. However, this replacement method will produce a smaller than desired standard deviation. The overall impact of replacing nondetects by their detection limits will depend on the relative magnitudes of the differences in the resulting mean and standard deviation from their true values induced by the replacement of the nondetects. If a simple replacement method is used, care should be taken that the resulting values do not result in an underestimate of the reasonable maximum values present at the site.

7. SUMMARY

Non-detect values are frequently encountered in the analysis of environmental data. The manner in which the nondetect values are handled should depend on the type of decision to be made and the magnitude and frequency of the nondetect values. If the nondetects are small in magnitude or low in frequency, the method of handling the nondetects will probably have minimal impact on the final outcome of the analysis. However, if the detection limits are close to important decision values, or if the frequency of nondetects is high, the treatment of the nondetect values can greatly influence resulting decisions.

ACKNOWLEDGEMENTS

The authors thank Larry Ticknor of the Los Alamos National Laboratory Statistics Group for helpful discussion and for bringing some of the references to our attention.

REFERENCES


27. Travis, C. C. and Land, M. L. (1990), "Estimating the Mean of Data Sets with Nondetectable
Data Quality Assessment: Statistical Methods for Practitioners

EPA QA/G-9S
Transformations may also make the analysis of data easier by changing the scale into one that is more familiar or easier to work with.

Once the data have been transformed, all statistical analysis should be performed on the transformed data. Rarely should an attempt made to transform the data back to the original form because this can lead to biased estimates. For example, estimating quantities such as means, variances, confidence limits, and regression coefficients in the transformed scale typically leads to biased estimates when transformed back into original scale. However, it may be difficult to understand or apply results of statistical analysis expressed in the transformed scale. Therefore, if the transformed data do not give noticeable benefits to the analysis, it is better to use the original data. There is no point in working with transformed data unless it adds value to the analysis.

4.7 VALUES BELOW DETECTION LIMITS

Data generated from chemical analysis may fall below the detection limit (DL) of the analytical procedure. These measurement data are generally described as non-detects rather than as zero or not present and the appropriate limit of detection is usually reported. In cases where measurement data are described as non-detects, the concentration of the chemical is unknown although it lies somewhere between zero and the detection limit. Data that includes both detected and non-detected results are called censored data in the statistical literature.

There are a variety of ways to evaluate data that includes values below the detection limit. However, there are no general procedures that are applicable in all cases. Some general guidelines are presented in Table 4-4. Although these guidelines are usually adequate, they should be implemented cautiously.

<table>
<thead>
<tr>
<th>Approximate Percentage of Non-Detects</th>
<th>Section</th>
<th>Statistical Analysis Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 15%</td>
<td>4.7.1</td>
<td>Replace non-detects with 0, DL/2, DL, Cohen’s Method.</td>
</tr>
<tr>
<td>15% - 50%</td>
<td>4.7.2</td>
<td>Trimmed mean, Cohen’s Method, Winsorized mean and standard deviation.</td>
</tr>
<tr>
<td>&gt; 50% - 90%</td>
<td>4.7.3</td>
<td>Tests for proportions (Section 3.2.1.5)</td>
</tr>
</tbody>
</table>

All of the suggested procedures for analyzing data with non-detects depend upon the amount of data below the detection limit. For relatively small amounts below detection limit values, replacing the non-detects with a small number and proceeding with the usual analysis may be satisfactory depending on the purpose of the analysis. For moderate amounts of data below the detection limit, a more detailed adjustment is appropriate. In situations where relatively large amounts of data fall below the detection limit, one may need only to consider whether or not the chemical was detected above some level. Table 4-4 provides percentages to
assist the user in evaluating their particular situation. However, it should be recognized that these percentages are not hard and fast rules.

In addition, sample size influences which procedures should be used to evaluate the data. For example, the case where 1 sample out of 4 is not detected should be treated differently from the case where 25 samples out of 100 are non-detects. Therefore, this guidance suggests that the data analyst consult a statistician for the most appropriate way to evaluate data containing values below the detection level.

4.7.1 Approximately less than 15% Non-detects - Substitution Methods

If a small proportion of the observations are non-detects, then these may be replaced with a small number, usually DL/2, and the usual analysis performed. Alternative substitution values are 0 (see Aitchison’s Method below) or the detection limit. It should be noted that Cohen’s Method (section 4.7.2.1) will also work with small amounts of non-detects.

4.7.1.1 Aitchison’s Method

Later adjustments to the mean and variance assume that the data values really were present but could not be recorded since they were below the detection limit. However, there are cases where the data values are below the detection limit because they are actually zero, i.e., the contaminant or chemical of concern being entirely absent. Such data sets typically contain a mixture of zero values and present, but nondetected values. Aitchison’s Method is simply adjustment formulas for the mean and variance if 0 values are substituted for non-detects. Directions for Aitchison’s method are contained in Box 4-29 with an example in Box 4-30.

---

**Box 4-29: Directions for Aitchison’s Method to Adjust Means and Variances**

Let $X_1, X_2, \ldots, X_m, X_{m+1}, \ldots, X_n$ represent the data points where the first $m$ values are above the detection limit and the remaining $n-m$ data points are below the detection limit.

**COMPUTATIONS:** Using the data above the detection level, compute the sample mean and sample variance:

$$
\bar{X}_a = \frac{1}{m} \sum_{i=1}^{m} X_i \quad \text{and} \quad s^2_a = \frac{1}{m-1} \left[ \sum_{i=1}^{m} X_i^2 - \frac{1}{m} \left( \sum_{i=1}^{m} X_i \right)^2 \right].
$$

Compute the adjusted sample mean and sample variance, $\bar{X} = \frac{m}{n} \bar{X}_a$ and $s^2 = \frac{m-1}{n-1} s^2_a + \frac{m(n-m)}{n(n-1)} \bar{X}_a^2$. 

---

EPA QA/G-9S

131

February 2006
Box 4-30: An Example of Aitchison’s Method

The following data consist of 10 Methylene Chloride samples: 1.9, 1.3, <1, 2.0, 1.9, <1, <1, <1, 1.6, 1.7. There are 6 values above the detection limit and 4 below, so m = 6 and n = 4. Aitchison’s method will be used to estimate the mean and sample variance of this data.

**COMPUTATIONS:** Compute the mean and variance for the 6 values above the detection limit

\[
\bar{x}_d = 1.733 \quad \text{and} \quad s_d^2 = 0.0667
\]

The adjusted sample mean and sample variance are: \( \bar{x} = \frac{6}{10} \cdot 1.733 = 1.04 \) and

\[
s^2 = \frac{6-1}{10-1} \cdot 0.0667 + \frac{6 \cdot 4}{10 \cdot (10-1)} \cdot 1.733^2 = 0.8382
\]

### 4.7.2 Between Approximately 15% - 50% Non-detects

#### 4.7.2.1 Cohen’s Method

Cohen’s method provides adjusted estimates of the sample mean and standard deviation that accounts for data below the detection level. The adjusted estimates are based on the statistical technique of maximum likelihood estimation of the mean and variance so that the non-detects are accounted for. Care has to be taken when using the adjusted mean and variance in statistical tests. If the percentage of data below the detection limit is relatively small (e.g., less than 20%), the significance level and power of the statistical test are approximately correct. As the proportion of data below detection increases, power declines and the true significance level increases dramatically. This is mainly attributable to the lack of independence between the adjusted mean and adjusted variance. If more than 50% of the observations are not detected, Cohen’s method should not be used. In addition, this method requires that the data without the non-detects be normally distributed and that the detection limit is always the same. Directions for Cohen’s method are contained in Box 4-31 with an example in Box 4-32.

#### 4.7.2.2 Selecting Between Aitchison’s Method and Cohen’s Method

Cohen’s underlying model is that the population contains a normal distribution, but we cannot see the values below the censoring point. Aitchison’s underlying model is that the population consists of a proportion following a normal distribution together with a proportion of values at zero. The difference in concepts becomes relevant depending on the types of inferences made. For example, in estimating upper quantiles, the analyst may use only the normal portion for the statistics, adjusting the quantile to account for the estimated proportion at zero. If a confidence interval for the mean was required a simple substitution of zero for all data below detection would suffice. To determine if a data set is better adjusted by Cohen’s method or Aitchison’s method, a simple graphical procedure using a Normal Probability Plot (Section 2.3.5) can be used. Directions for this procedure are given in Box 4-34 with an example in Box 4-35.
**Box 4-31: Directions for Cohen’s Method**

Let $X_1, X_2, \ldots, X_m, X_{m+1}, \ldots, X_n$ represent the data points where the first $m$ values are above the detection limit (DL) and the remaining $n-m$ data points are below the detection limit.

**COMPUTATIONS:** Using the data above the detection level, compute the sample mean and sample variance:

$$
\bar{X}_d = \frac{1}{m} \sum_{i=1}^{m} X_i \quad \text{and} \quad s_d^2 = \frac{1}{m-1} \left( \sum_{i=1}^{m} X_i^2 - \frac{1}{m} \left( \sum_{i=1}^{m} X_i \right)^2 \right).
$$

Compute $h = \frac{n-m}{n}$ and $\gamma = \frac{s_d^2}{(\bar{X}_d - DL)^2}$. Use $h$, $\gamma$, and Table A-11 to determine $\lambda$. If the exact values of $h$ and $\gamma$ do not appear in the table, use double linear interpolation (Box 4-33) to estimate $\lambda$.

Estimate the corrected sample mean, $\bar{X}$, and sample variance, $s^2$:

$$
\bar{X} = \bar{X}_d - \lambda(\bar{X}_d - DL)
$$

$$
s^2 = s_d^2 + \lambda(\bar{X}_d - DL)^2.
$$

---

**Box 4-32: An Example of Cohen’s Method**

Sulfate concentrations (mg/L) were measured for 24 data points with 3 values falling below the detection limit of 1450 mg/L. The 24 values are:

1850, 1760, <1450, 1710, 1575, 1475, 1780, 1790, 1780, <1450, 1790, 1800, <1450, 1800, 1840, 1820, 1860, 1780, 1760, 1800, 1900, 1770, 1790, 1780.

Cohen’s Method will be used to adjust the sample mean and sample variance for use in a t-test to determine if the mean is greater than 1600 mg/L.

**COMPUTATIONS:** The sample mean and sample variance of the $m = 21$ values above the detection level are

$$
\bar{X}_d = 1771.9 \quad \text{and} \quad s_d^2 = 8593.69.
$$

The values of $h$ and $\gamma$ are: $h = \frac{24 - 21}{24} = 0.125$ and $\gamma = \frac{8593.69}{(1771.9 - 1450)^2} = 0.083$. Table A-11 does not contain the exact entries for $h$ and $\gamma$, double linear interpolation was used to estimate $\lambda = 0.149839$ (see Box 4-33).

The adjusted sample mean and sample variance are:

$$
\bar{X} = \bar{X}_d - \lambda(\bar{X}_d - DL) = 1771.9 - 0.149839 \cdot (1771.9 - 1450) = 1723.67 = \bar{X}
$$

$$
s^2 = s_d^2 + \lambda(\bar{X}_d - DL)^2 = 8593.69 + 0.149839 \cdot (1771.9 - 1450)^2 = 24119.95 = s^2
$$

EPA QA/G-9S

February 2006
Box 4-33: Example of Double Linear Interpolation

The details of the double linear interpolation are provided to assist in the use of Table A-11. The desired value for $\hat{y}$ corresponds to $\gamma = 0.083$ and $h = 0.125$ from Box 4-32. The values from Table A-11 used for interpolation are:

\[
\begin{array}{c|cc}
\gamma & c_1 = 0.10 & c_2 = 0.15 \\
\hline
r_1 = 0.05 & x_{11} = 0.11431 & x_{12} = 0.17925 \\
\hline
r_2 = 0.10 & x_{21} = 0.11804 & x_{22} = 0.18479 \\
\end{array}
\]

We first interpolate between columns:

\[
y_1 = x_{11} + \frac{c - c_1}{c_2 - c_1} (x_{12} - x_{11}) = 0.11431 + \frac{0.125 - 0.10}{0.15 - 0.10} (0.17925 - 0.11431) = 0.14678
\]

\[
y_2 = x_{21} + \frac{c - c_1}{c_2 - c_1} (x_{22} - x_{21}) = 0.11804 + \frac{0.125 - 0.10}{0.15 - 0.10} (0.18479 - 0.11804) = 0.151415
\]

Now we interpolate between the rows:

\[
\hat{y} = y_1 + \frac{r - r_1}{r_2 - r_1} (y_2 - y_1) = 0.14678 + \frac{0.083 - 0.05}{0.10 - 0.05} (0.151415 - 0.14678) = 0.149839
\]

Box 4-34: Directions for Selecting Between Cohen’s Method or Aitchison’s Method

Let $X_1, X_2, \ldots, X_m, \ldots, X_n$ represent the data points with the first $m$ values are above the detection limit (DL) and the remaining $n-m$ data points are below the DL.

**STEP 1:** Use Box 2-17 to construct a Normal Probability Plot of all the data but only plot the values belonging to those above the detection level. This is called the Censored Plot.

**STEP 2:** Use Box 2-17 to construct a Normal Probability Plot of only those values above the detection level. This called the Detects only Plot.

**STEP 3:** If the Censored Plot is more linear than the Detects Only Plot, use Cohen’s Method to estimate the sample mean and variance. If the Detects Only Plot is more linear than the Censored Plot, then use Aitchison’s Method to estimate the sample mean and variance.

4.7.3 Greater than Approximately 50% Non-detects - Test of Proportions

If more than 50% of the data are below the detection limit but at least 10% of the observations are quantified, then the best option is a test of proportions. Thus, if the parameter of interest is a mean, consider switching the parameter of interest to some percentile greater than the percent of data below the detection limit. For example, if 67% of the data are below the DL, consider switching the parameter of interest to the 75th percentile. Then the method described in 3.2.1.5 can be applied to test a hypothesis concerning the 75th percentile. It is important to note that the tests of proportions may not be applicable for composite samples. In this case, the data analyst should consult a statistician before proceeding with analysis.
Box 4-35: Example of Determining Between Cohen’s Method and Aitchison’s Method

Readings of Chlorobenzene were obtained from a monitoring well:

<1, <1, 1.2, 1.25, 1.3, 1.45, 1.35, 1.55, 1.6, 1.85, 2.1

Step 1: Using the directions in Box 2-17 the following is the Censored Plot:

STEP 2: Using the directions in Box 2-17 the following is the Detects only Plot:

STEP 3: Since the Censored Plots is more linear than the Detects Only Plot, Cohen’s Method should be used to estimate the sample mean and variance.

4.7.4 Greater than Approximately 90% Non-detects

If very few quantified values are found, a method based on the Poisson distribution may be used as an alternative approach. However, with a large proportion of non-detects in the data, the data analyst should consult with a statistician before proceeding with analysis.

4.7.5 Recommendations

If the number of sample observations is small (n < 20), Cohen’s and other maximum likelihood methods can produce biased results since it is difficult to assure that the underlying distribution is appropriate and the solutions to the likelihood equation are statistically consistent only if the number of samples is large. Additionally, most methods will yield estimated parameters with large estimation variance, which reduces the power to detect import differences from standards or between populations. While these methods can be applied to small data sets, the user should be cautioned that they will only be effective in detecting large departures from the null hypothesis.

If the degree of censoring is relatively low, reasonably good estimates of means, variances and upper percentiles can be obtained. However, if the rate of censoring is very high (greater than 50%) then little can be done statistically except to focus on some upper quantile of
the contaminant distribution, or on some proportion of measurements above a certain critical level that is at or above the censoring limit.

When the numerical standard is at or below one of the censoring levels and a one-sample test is used, the most useful statistical method is to test whether the proportion of a population that is above (below) the standard is too large, or to test whether and upper quantile of the population distribution is above the numerical standard. Table 4-5 gives some recommendation on which statistical parameter to use when censoring is present in data sets for different sizes of the coefficient of variation.

<table>
<thead>
<tr>
<th>Approximate Coefficient of Variation (CV)</th>
<th>Approximate Proportion of Data Below the Detection Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large: CV &gt; 1.5</td>
<td>Mean or Upper Percentile</td>
</tr>
<tr>
<td>Medium: 0.5 &lt; CV &lt; 1.5</td>
<td>Mean or Upper Percentile</td>
</tr>
<tr>
<td>Small: CV &lt; 0.5</td>
<td>Mean or Median</td>
</tr>
</tbody>
</table>

Table 4-5. Guidelines for Recommended Parameters for Different Coefficient of Variations and Censoring

When comparing two data sets with different censoring levels (i.e., different detection limits), it is recommended that all data be censored at the highest censoring value present and a nonparametric test such as the Wilcoxon Rank Sum Test (Section 3.3.2.1.1) used to compare the two data sets. There is a corresponding loss of statistical power but to a certain extent this can be minimized through the use of large sample sizes.

4.8 INDEPENDENCE

When data are truly independent, the correlation between data points is by definition zero and the selected statistical tests attains the desired decision error rates (given the appropriate assumptions have been satisfied). When correlation exists, the effectiveness of statistical tests is diminished. Environmental data are particularly susceptible to correlation problems due to the fact that such environmental data are collected under a spatial pattern or sequentially over time.

If observations are positively correlated over time or space, then the effective sample size for a test tends to be smaller than the actual sample size—i.e., each additional observation does not provide as much 'new' information because its value is partially determined by the value of adjacent observations. This smaller effective sample size means that the degrees of freedom for the test statistic is smaller, or equivalently, the test is not as powerful as originally thought. In addition to affecting the false acceptance error rate, applying the usual tests to correlated data tends to result in a test whose actual significance level is larger than the nominal error rate.