APPENDIX H
Censored Data

H-1. Introduction. Laboratories report analytical data in two ways, as censored or uncensored. An environmental testing laboratory reports a result as "non-detected" or "ND" when the result is below some numerical reporting threshold. The non-detect is typically reported as "< *X*" (e.g., or "*X* U") where *X* is some numerical value. This is called a "censored result," and the value of *X* is called the "censoring limit." Results reported as "detected" are "uncensored" results. Typically, uncensored results are numerical values in concentration units that are greater than either the critical value or the censoring limit. Unfortunately, different environmental laboratories use different types of censoring limits and reporting conventions. There is no standard industry practice regarding how to establish the censoring limit for non-detections. To exacerbate matters, as discussed previously, there is no standard terminology for the censoring limit. Reporting conventions differ from laboratory to laboratory. Some laboratories refer to the censoring limit as the DL, while other laboratories refer to this value as the "reporting limit" (RL).

H-1.1. Before evaluating censored data, it is important to understand the nature of the censoring limit being used, that is, to understand how it is being defined for a particular set of data. To confidently report a non-detect at the censoring limit, the censoring limit must be equal to or greater than the detection limit (as this quantity is defined by the IUPAC); ideally, non-detects should be reported as "< DL" (larger values are undesirable for statistical evaluations and smaller values are undesirable for the minimization of minimize false negatives). For normally distributed data, in general, the censoring limit should be at least two times greater than the reported critical value. However, it is not uncommon for laboratories to report non-detects to values as low as the MDL (where false negatives cannot be reliably reported). The censoring limit is often the laboratory's practical quantitation limit (PQL), which may also be simply called the QL. Under these circumstances, a laboratory reports numerical results greater than the QL as quantitatively reliable values. A result less than the QL may be reported as detection, consisting of a numerical value with a "data qualifier" or "flag" if the result is greater than the critical value (e.g., the method detection limit), or the result may be reported as a non-detect as "< QL." For example, if QL = 10 ppm, MDL = 1 ppm, and a result of 5 ppm is measured, the laboratory may report the result as either "5 J" or as "< 10." The reporting of the result as "5 J" indicates that the analyte is present, but the concentration of 5 ppm is a highly estimated value (i.e., is not quantitatively reliable). If the result were reported as "< 10," the result would be a censored value (indicating the concentration of analyte is no greater than 10 ppm). The J-qualifier is typically applied when the analyte is believed to be present at some concentration less than the QL. (Detection, quantitation, and RLs are discussed in detail in Appendix G.)

H-1.2. When measurement data are reported as "ND," the exact concentration of the chemical is unknown, but lies somewhere between zero and the censoring limit. No quantitative information is available for a non-detect (except that the result is less than the censoring limit) because no estimate is provided to quantify how much smaller the result is than the censoring limit. Although useful for data reporting and presentation, censored data complicate statistical analyses and data interpretation. Qualitative results cannot be used because statistical calculations require numerical values rather than attributes. For example, the inequality "< 10 ppm" cannot be substituted into the equation to calculate the sample mean although a value of 5 could be substituted for a result reported as "5 J."

H-1.3. Statistical literature, Federal standards, and USEPA guidance advocate the use of uncensored measured concentrations for statistical calculations. Uncensored data give rise to more accurate estimates of mean and standard error than censored data, which result in more accurate data interpretation and more reliable conclusions. However, under these circumstances, numerical values (even negative values) would be reported for each sample regardless of the magnitude of the concentrations relative to the DLs. Unfortunately, in practice, censored data are typically reported for environmental applications because uncensored data are often unavailable or difficult to obtain, especially for prior sampling events (e.g., some laboratory instruments are incapable of reporting uncensored values). Requesting uncensored data may also increase analytical laboratory costs because uncensored data are not routinely reported, but it can be done at a reasonable cost for select analytical methods (e.g., typically, for metal analyses).

H-1.4. As censored data are commonly reported for environmental testing, the next Paragraph presents a variety of strategies for treating censored data. Some are recommended, while others should be used with greater caution. Gilbert (1987) and Gibbons (1994) contain more information on dealing with censored data.

H-2. Overview of Strategies for Treating Censored Data. There are several possible approaches for treating censored data. Four general strategies are listed below and then described in more detail:

- The censored values can be ignored (omitted from the statistical calculations).
- Proxy values (e.g., the censoring limit, one-half the censoring limit, or zero) can be substituted for the NDs to obtain numerical values for computations (e.g., for the mean and variance).
- Statistical quantities such as the mean and variance can be adjusted based upon the proportion of NDs by making certain distribution assumptions.
- Nonparametric methods can be used.

No single approach can be used for all data sets and all data quality objectives. The characteristics of the data set and its end use must be taken into account when selecting the most appropriate approach.

H-2.1. *Approach 1*. The first approach, omitting the NDs from the data set, is typically undesirable as it decreases the total number of data points and the reliability of the statistical evaluations. In addition, the NDs often provide valuable information about the environmental population of interest. For example, a set of NDs that are all less than some risk-based decision limit provides valuable information about the site. This approach is potentially viable only under select circumstances and for select data quality objectives. For example, if there are a large number of samples for a study area, the censoring limit is small relative to some risk-based decision limit to which monitoring is being performed. In this case, if a statistical evaluation using only the set of detections were to indicate that contamination is present at concentrations significantly less than the decision limit, then the omission of the NDs would probably not affect decision-making.

H-2.2. *Approach 2*.

H-2.2.1. The second approach is called the "substitution method." Proxy or surrogate values are assigned to all the NDs. One approach for assigning proxy values is to assume that any value between zero and the censoring limit is equally probable and substitute one-half the censoring limit (midpoint of the range of possible values) for each ND. Other common proxy values are zero or the censoring limit itself. However, assigning proxies requires assumptions about the distribution of NDs. For example, assuming that all values less than the censoring limit are equally likely is equivalent to assuming a uniform probability distribution for all possible measurements between zero and the censoring limit. Assuming that all non-detects are equal to a fixed proxy value can bias the estimated standard deviation for the data set, particularly when a substantial number of results are NDs (see ASTM D-4210-89 for further discussion of this topic). For example, substituting the censoring limit could result in a sample mean that is biased high, and substituting zero could result in a mean that is biased low. Substituting one half the censoring limit may not bias the mean, but often adversely affects the estimate of the standard deviation. Biasing such summary statistics may result in erroneous conclusions about project objectives. In general, it is undesirable to assign proxy values when a significant portion of the data set (e.g., more than 15%) contains censored values.

H-2.2.2. As noted previously, laboratories often report uncensored data below the censoring limit as estimated positive detections (commonly indicated as J-flagged values). Using these uncensored data for statistical computations (not necessarily for data reporting) prevents the need to assign proxy concentrations based on arbitrary algorithms (EPA 9285.7-09A, Gilbert, 1987). While measurements below the censoring limit may not indicate the presence of target analytes as reliably as measurements above the limit, in many cases uncensored measurements are still better estimates of contaminant concentration than any proxy that might be applied. Generally, this approach allows data users and decision-makers to better characterize site conditions. Censored data are always relevant for determining the presence or absence of a contaminant at a site, as long as appropriate qualitative identification criteria have been satisfied.

H-2.3. *Approach 3*. The third approach entails adjusting the average and standard deviation instead of estimating proxy values for each ND result. However, to do this, it is also necessary to make assumptions about the data distributions (such as, all NDs vary in a manner similar to results above the censoring limit—maximum likelihood estimation procedure and the probability plotting method—or Cohen's method, which assumes a normal distribution). Adjustment methods provide accurate results only when the distribution assumptions are valid; otherwise, elevated estimates of the average and standard error could result. Usually, adjustment methods should be used when 15 to 50% of the values of the data results are censored.

H-2.4. *Approach 4*. A nonparametric approach should be considered when a significant portion of the data set consists of censored values. This approach typically involves ordering the data values (from smallest to largest) and replacing the data values with the corresponding rank number. The NDs are then treated as tied ranks and would be replaced by some common mid-rank value. Though not generally recommended, according to EPA guidance, if the DLs are not the same, then the NDs, instead of being treated as tied values, would be ranked according to their numerical estimates (EPA 68-W0-0025). The advantage of a non-parametric approach over the strategy of assigning proxy values is that no distribution assumptions are made. However, a larger number of data points are required for nonparametric methods to achieve the same level of confidence as parametric methods. Furthermore, though non-parametric methods can tolerate a greater proportion of NDs than parametric methods, non-parametric methods will not be viable if there are many NDs. For example, the median (refer to Appendix D) could not be determined from a data set that consists of more than 50% NDs.

H-2.5. *Complicating Factors*. For most projects, uniform numerical censoring limits will be available; however, there are instances when this is not the case. A laboratory can provide sample-specific detection limits or critical values (i.e., limit adjusted by the sample-specific dilution factor, soil moisture, or other analytical adjustments) that vary from sample to sample. In this case, use the sample-specific limits to establish the proxy values. As there is no standard nomenclature or well-established conventions for generating censoring limits in the environmental testing industry, it is recommended that the project chemist be consulted to establish the nature of the censoring limits being reported.

H-2.5.1. Censored results are sometimes reported as "ND" without the associated censoring limit. When censoring limits are not provided with data, this information can usually be obtained by contacting the laboratory if the analyses are current. If this information is not available, it might be viable to estimate a censoring limit based upon the lowest reported concentration, such as the lowest J-flagged result. Because J-flagged results are, by definition, concentrations that exceed the critical value, the minimum result represents a value that is closest to the critical value. A chemist should be consulted to examine the J-flagged values to determine if there are anomalous values that would set proxies at inappropriate levels. For example, an examination of the J-flagged results may indicate that there may be, in effect, two different censoring levels—one for "dirty" samples and one for "clean" samples. The project chemist might want to consider the issues of aliquot sizes and dilution conformity, among others factors, prior to making a final recommendation.

H-2.5.2. When using a nonparametric method to address NDs, ranking the data is often problematic when there are multiple censoring limits. For example, in general, it cannot be concluded that “< 10” represents a value that is greater than “< 1.” The most appropriate approach for addressing multiple censoring limits depends upon the nature of the parametric test being used. One approach consists of setting all of the non-detects to the largest censoring limit and treating these as tied values. Detected values less than the largest censoring limit (i.e., detection limit) must also be censored to the highest detection limit and treated as ties. This approach is not optimal because information is lost when all of the results are censored to the highest detection limit. However, the approach is statistically valid, simple to implement, and could be adequate for a large data set. It should also be noted that, rather than treating the NDs as ties, it is a common practice to rank the NDs according to their numerical estimates (EPA 68-W0-0025). Although this approach is used in this document (to be consistent with EPA guidance), it is not necessary appropriate.

H-2.6. *Overview Summary*. Some general guidelines are presented in Table H-1 based on the percentage of NDs. Substitution methods are recommended when less than 15% of the data are NDs. Adjustment or nonparametric methods should be considered when more than 15% of the results are censored. If more than 50% of the data set’s concentrations are NDs, it is recommended that nonparametric methods be used instead of adjustment methods.

Table H-1.
Guidelines for Analyzing Data with NDs

Percentage of NDs	Paragraph	Proxy Definition/Statistical Analysis Method
< 15	H-3	Replace NDs with one-half censoring limit or a very small number
15–50	H-4	Trimmed mean, Cohen’s or Atchison’s adjustment, Winsorized mean, and standard deviation or non-parametric methods
> 50 – 90	H-5	Use tests for proportions
> 90	H-6	Use tests based on Poisson distribution

H-2.6.1. OSWER 9285.7-41/EPA 540-R-01-003 recommends a substitution method for censored results that is not recommended herein. The EPA suggests that a proxy value for NDs, based on one-half the censoring limit or on a random value between zero and the censoring limit, be used. According to the document, the censoring limit should be equal to the “sample-specific quantitation limit” and the method may be used so long as fewer than 50% of the data set’s concentrations are NDs. However, as per Table H-1, *it is recommended that proxy values not be used when more than 15% of the results are reported as NDs*. Using proxy values can bias the results of the statistical evaluations. The data user should verify that the “sample-specific quantitation limit” (SQL) is an appropriate censoring limit and adequately addresses false negatives as discussed in Appendix G. False negatives will not be minimized at the SQL when this limit is essentially a sample-specific MDL.

H-2.6.2. Although guidelines in Table H-1 are usually adequate, they should be implemented cautiously. Professional judgment is critical. In particular, the use of proxy values for a substitution approach should be evaluated in terms of the data quality objectives of the project. If the censoring limits are greater than or near project decision levels, then this approach may not be appropriate.

H-2.6.3. In Table H-1, all of the suggested procedures for analyzing data with NDs depend on the percentage of data below the censoring limit. For relatively small amounts below the censoring limit, replacing the NDs with a small number and proceeding with the usual analysis may be satisfactory. For moderate amounts of data below the censoring limit, a more detailed adjustment is appropriate. In situations where relatively large amounts of data below the censoring limit exist, one may need only to consider whether a certain proportion of the samples display values greater than some threshold values. The interpretation of small, moderate, and large amounts of data below the censoring limit is subjective. Table H-1 provides guideline percentages to assist the user in evaluating their particular situation; however, it should be recognized that these percentages are not rigid rules, but should be based on judgment.

H-2.6.4. In addition to the percentage of samples below the censoring limit, sample size influences which procedures should be used to evaluate the data. For example, the case where the result for 1 sample out of 4 is not detected should be treated differently from the case where the results for 25 samples out of 100 are not detected. It is recommended that the data analyst consult a statistician for the most appropriate way to evaluate data containing values below the detection level.

H-2.6.5. The remaining portion of this Appendix describes in detail the various methods outlined above. Case studies and examples are also presented.

H-3. Substitution Methods for Less than 15% NDs. If small proportions, 15% or fewer, of the observations are NDs, these may be replaced with a small number, the DL, DL/2, or a random value between the DL and zero (see EPA 540-R-01-003). After the non-detected values have been given a proxy value, then the usual statistical analysis may be performed. If simple substitution of values below the DL is proposed when more than 15% of the values are reported as not detected, consider using nonparametric methods or a test of proportions to analyze the data.

H-3.1. As a simplified case study showing the magnitude of effect on simple statistics attributable to different proxy concentrations, consider the data in Table H-2 for sodium in surface soil at a site. This table presents summary statistics for sodium data when 3 results of the 21 samples analyzed are not detected and 8 types of proxy concentrations have been used to represent these non-detected results. These proxies are the DL, RL, ½DL, ½RL, and a random number selected in four different ways as described in the table.

H-3.2. Summary statistics, in particular the average and standard deviation, are affected by the choice of proxy concentration. Proxy concentrations were developed based on the sample-specific DL and the project RL to illustrate how they are affected by the limit used for estimation. In this case study, a concentration not detected is reported as < DL. The DL is more appropriate to use to estimate a proxy than the RL, because it is the closest value at which the non-detected concentration may have occurred. If a concentration was > DL, but still < RL, the concentration is reported as a detect. Hypothetically, had only the RL been available and no DL had been provided, an alternative method to determine a proxy concentration would be to select the lower of the RL and the minimum detected result. Then, the proxy value would be at least below all of the detected concentrations.

H-3.3. As a basis for comparison, the summary statistics were also calculated using only the positively detected results in column 1 of Table H-2. In this instance, it is expected that the calculated average concentration would be higher than the true average, and the calculated standard deviation would be lower than the true standard deviation. When the RL is used to create proxy values (columns 7, 9, and 11), the average is higher and the standard deviation is lower than the associated summary statistics when the DL is used (columns 6, 8, and 10). Of all the cases when the RL is used to create a proxy, the case when a random number between zero and the RL is used (column 11) tends to have estimates for the average and standard deviation that are similar to the cases when the DL is used. This may be related to the fact that when a simple substitution such as the RL or ½RL is used, the variability is reduced because the proxy concentrations do not account for the inherent variation among concentrations. Proxy values are consistently the same number, whereas a proxy value based on a random number varies. It is also interesting to note that, in general, the summary statistics are similar for cases using random numbers as proxy values, no matter if the proxy value was based on the DL, RL, or the lowest detected result.

H-4. Methods for 15 to 50% NDs. Adjustment methods for treating NDs are commonly applied when NDs compose 15 to 50% of the data set. These various methods have their strengths and weaknesses, and they are presented first. Cohen's method is probably the most frequently used. A brief outline of a non-parametric procedure follows the discussion of adjustment methods.

H-4.1. *Cohen's Method*. Cohen's method provides adjusted estimates of the sample mean and standard deviation that accounts for data below the detection level when data are normally distributed. The adjusted mean and standard deviation can then be used in the parametric test described in Appendix L (EPA 600/R-96/084 QA/G-9). This method requires knowing the censoring level, the percent of NDs, and either the arithmetic mean and standard deviation of the data (if the data are normally distributed) or the arithmetic mean and standard deviation of the log-transformed data (if the data are log-normally distributed). The data must also be evaluated for normality (Appendix F). For Cohen's method, the distribution is tested on the entire data set: positive detections and censored data. If the distribution testing fails to be normal or lognormal, Atchison's method (described below) may be more appropriate. Once the data distribution has been determined to be normal, the proxy concentrations themselves are essentially irrelevant when computing the adjusted mean and standard deviation.

H-4.1.1. Cohen's adjustment is a theoretically attractive method for handling cases with between 15 and 50% NDs. Conceptually, the method considers the detected results to be the top *X*% of an assumed distribution (normal, lognormal). The mean and standard deviation are then computed by filling in the bottom *Y*% of the assumed distribution (i.e., by assuming that the NDs represent the lower tail of the assumed distribution). These are referred to as the adjusted mean and adjusted standard deviation. This method appears to be a reasonable method for handling NDs and is attractive because it does not require the use of proxy concentrations (after normality has been determined).

H-4.1.2. There are, however, several practical difficulties encountered when applying this method, as follows.

H-4.1.2.1. Because there are no tests for how reliably the top *X*% of the data represent the top *X*% of a normal or lognormal distribution, there is a high degree of reliance on subjective judgment in selecting the appropriate distribution. So, the dilemma remains whether it is more appropriate to determine the distribution based on just the detected values or whether it is more appropriate to determine the distribution based on detects and proxy concentrations representing the NDs.

H-4.1.2.2. With Cohen's method, the sample size is effectively reduced because estimates are based only on the detected results. Estimates that are based on a small number of results are highly sensitive to the degree of uncertainty. This is particularly true when a lognormal distribution is assumed, and there is a high proportion of ND results. This leads to poorer estimates of the standard deviation, which can substantially impact calculations.

H-4.1.2.3. The method assumes that a single censoring level applies to all ND results. This is not always true (for example, if some NDs are for diluted samples and others are not), and the selection of the censoring level used in the calculations can have a substantial effect on the outcome.

Table H-2.
Case Study, Sodium in Surface Soil: Summary Statistics Using Various Substitution Methods for Proxy Values

		1	2	3	4	5	6	7	8	9	10	11
Sampling Event	Sample	Result (mg/kg)	DL (mg/kg)	RL (mg/kg)	1/2DL as Proxy	1/2RL as Proxy	DL as Proxy	RL as Proxy	Random Number between 0 and DL as Proxy	Random Number between 0 and RL as Proxy	Random Number between 0 and lower of min. result and DL as Proxy	Random Number between 0 and lower of min. result and RL as Proxy
A	SS-010	ND	50	500	25	250	50	500	18.4	68.4	4.4	7.3
A	SS-020	ND	50	500	25	250	50	500	24.5	272.0	38.1	56.3
A	SS-030	1710										
A	SS-040	1860										
A	SS-050	2150										
A	SS-060	ND	50	500	25	250	50	500	13.9	47.8	48.1	28.6
B	SB01	750										
B	SB02	2430										
B	SB03	1160										
B	SB04	66										
B	SB05	140				Positive detections for columns 2 through 11 are the same as reported in column 1. They are omitted to highlight the DL, RL, and various proxy values.						
B	SB06	89										
B	SB07	120										
B	SB08	60										
B	SB09	107										
B	SB10	170										
B	SB11	180										
B	SB12	310										
B	SB13	71										
B	SB14	88										
B	SB15	61										
Summary Statistics												
25th Percentile		88.25			66	89	66	89	66	71	66	66
Median		155			120	180	120	180	120	140	120	120
75th Percentile		1057.5			750	750	750	750	750	750	750	750
Average		640.1			552.2	584.4	555.8	620.1	551.4	567.2	553.0	553.1
Standard Deviation		828.9			795.4	776.9	792.9	765.8	796.0	786.8	794.9	794.8

DL = Sample-specific Detection Limit
RL = Project Reporting Limit

H-4.1.3. Because this method requires knowing the censoring level, which is sometimes not reported and sometimes differs from one sample to another when it is reported, the following recommendations should be followed when Cohen's method is used.

H-4.1.3.1. If the censoring level (DL) is reported, and is the same for all non-detected results, use this value.

H-4.1.3.2. If the censoring level (DL) is reported, but is not consistent across all non-detected results, it is preferable to use the minimum censoring level among the NDs if there is justification to do so.

H-4.1.3.3. If the censoring level is not reported, then the values used to compute the proxy concentrations is the lesser of the RL and the minimum detected result.

H-4.1.3.4. If the RL and minimum detected results are the same for all non-detects, use that value.

H-4.1.3.5. If the RL values are not consistent across all non-detected results, use the minimum value among the NDs if there is justification to do so.

H-4.1.4. Because of the unrealistically elevated summary statistics that result when Cohen's method is applied, this method should be used with caution. Using Cohen's method is not recommended in more complicated evaluations, such as those required for an analysis of variance. Despite the limitations, there may be specific instances where it is applicable. In these cases, the results should be examined carefully to ensure that the conclusions are reasonable. The computational details of Cohen's method are presented in Paragraph H-4.2, and an example is given in Paragraph H-4.3.

H-4.2. *Directions for Cohen's Method*. Let $x_1, x_2 \ldots x_m, \ldots, x_n$ represent the n data points with the first m values representing the data points above the DL. Thus, there are $(n - m)$ data points below the DL.

H-4.2.1. Verify the distribution of the data to determine if they follow a normal or lognormal distribution (Appendix F). If they follow a normal distribution, then the raw data should be used for the following calculations. If the data follow a lognormal distribution, then the log-transformed data should be used for the following calculations.

H-4.2.1.1. Compute the sample mean $\bar{x}_d$ from the data above the DL:

$$\bar{x}_d = \frac{1}{m}\sum_{i=1}^{m} x_i \, .$$

H-4.2.1.2. Compute the sample variance s_d^2 from the data above the DL:

$$s_d^2 = \frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{m}.$$

H-4.2.1.3. Compute

$$h = \frac{n-m}{n}$$

and

$$\gamma = \frac{s_d^2}{(\bar{x} - DL)^2}.$$

H-4.2.1.4. Use h and γ in Table B-3 of Appendix B to determine $\hat{\lambda}$. For example, if h = 0.4 and γ = 0.30, then $\hat{\lambda}$ = 0.6713. If the exact value of h and γ do not appear in the table, use double linear interpolation (Paragraph H-4.4) to estimate $\hat{\lambda}$.

H-4.2.1.5. Estimate the corrected sample mean, $\bar{x}$, and sample variance, s^2, to account for the data below the DL as follows:

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

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

H-4.2.2. If these estimates are based on the log-transformed data, then they can be transformed back to the original units to estimate the mean and variance of the lognormal distribution. For example, if n is large, the mean and variance (of the untransformed data set) can be calculated as follows:

$$\bar{x}_{Ln} = \exp\left(\bar{x} + \frac{s^2}{2}\right) \text{ and } s_{Ln}^2 = \bar{x}^2\left[\exp\left(s^2\right) - 1\right].$$

H-4.3. *Example—Application of Cohen's Method*. Of the groundwater analyses for benzene at Site A in monitoring well MW03, 10 of the 15 sample results are positive detections and

5 of the 15 sample results are NDs. Table H-3 presents the benzene concentrations, the DL, and natural log of the concentrations.

Table H-3.
Benzene Concentrations, the DL, and Natural Log of the Concentrations

Sampling Event	Result, *X* (µg/L)	DL (µg/L)	*Ln*(*X*) (*Ln*[µg/L])
29-Jan-98	ND	0.0605	ND
18-Apr-98	1.78	0.0375	0.5766
15-Jul-98	ND	0.0375	ND
18-Oct-98	2.31	0.0375	0.8372
18-Apr-99	7.24	0.0469	1.980
18-Jul-99	1.85	0.0759	0.6152
20-Oct-99	0.308	0.0759	–1.178
1-Apr-00	2	0.0504	0.6931
17-Jul-00	0.143	0.0353	–1.945
16-Oct-00	0.235	0.0353	–1.448
17-Jan-01	ND	0.0641	ND
4-May-01	0.759	0.0401	–0.2758
28-Jul-01	0.222	0.0401	–1.505
5-Nov-01	ND	0.0465	ND
31-Jan-02	ND	0.0465	ND

H-4.3.1. The total number of samples $n = 15$, the number of detects $m = 10$, and the number of non-detects $n - m = 5$.

H-4.3.2. The distribution of the positive detections was determined by the Shapiro-Wilk test (Appendix F) to be lognormal. The distribution of the entire data set, including NDs set to the proxy concentration equal to the DL, was also tested and evidence of a lognormal distribution was found. Thus, the Cohen's adjustment may be used.

$$\bar{x}_d = \frac{1}{10}\sum_{i=1}^{10} Ln(x_i) = -0.1650.$$

$$s_d^2 = \frac{\sum_{i=1}^{10}(Ln(x_i) - \bar{x}_d)^2}{10} = 1.683.$$

$$h = \frac{n-m}{n} = \frac{5}{15} = 0.3333 \ .$$

H-4.3.3. The DLs vary for the NDs. The lowest DL associated with the NDs will be used in these calculations. So, $DL = \text{Ln}(0.0375) = -3.283$.

$$\gamma = \frac{1.683}{(-0.1650 - (-3.283))^2} = 0.1731 \ .$$

H-4.3.4. Using $h = 0.3333$ and $\gamma = 0.1731$ in Table B-3 of Appendix B and double linear interpolation (see Paragraph H-4.4 for details), $\hat{\lambda} = 0.5020$,

$$\bar{x} = -0.1650 - \{0.5020 \times [-0.1650 - (-3.283)]\} = -1.730$$

and

$$s^2 = 1.683 + \{0.5020 \times [-0.1650 - (-3.283)]^2\} = 6.563 .$$

H-4.3.5. Though n is relatively small, for the purposes of illustration, the corrected sample mean and variance for the lognormal distribution (based on the original units) are calculated as discussed in Paragraph H-4.2.

$$\bar{x}_{Ln} = \exp\left(-1.730 + \frac{6.563}{2}\right) = 4.719$$

and

$$s_{Ln}^2 = (-1.730)^2 [\exp(6.563) - 1] = 2117.$$

H-4.4. *Double Linear Interpolation*. The details of the double linear interpolation are provided to assist in the use of Table B-3 of Appendix B. Suppose the desired value corresponds to γ = 0.1731 and h = 0.3333 from Paragraph H-4.3. The values $\hat{\lambda}$ from Table B-3 for interpolation are:

γ	H = 0.30	h = 0.35
0.15	0.4330	0.5296
0.20	0.4422	0.5403

H-4.4.1. There are 0.05 units between 0.30 and 0.35 on the h scale, and 0.0333 units between 0.30 and 0.3333. Therefore, the value of interest lies (0.0333/0.05)1000% = 66.6% of the distance along the interval between 0.30 and 0.35. To linearly interpolate between tabulated values on the h axis for γ = 0.15, the range between the values must be calculated, 0.5296 – 0.4330 = 0.0966; the value that is 66.6% of the distance along the range must be computed, 0.0966 × 0.666 = 0.06434; and then that value must be added to the lower point on the tabulated values, 0.4330 + 0.06434 = 0.4973. Similarly for γ = 0.20, 0.5403 – 0.4422 = 0.0981, 0.0981 × 0.666 = 0.06533, and 0.4422 + 0.06544 = 0.5075. So,

γ	h = 0.30	h = 0.3333	h = 0.35
0.15	0.4330	0.4973	0.5296
0.20	0.4422	0.5075	0.5403

H-4.4.2. On the γ-axis there are 0.0231 units between 0.15 and 0.1731, and there are 0.05 units between 0.15 and 0.20. The value of interest (0.1731) lies (0.0231/0.05)100% = 46.2% of the distance along the interval between 0.15 and 0.20, so 0.5075 – 0.4973 = 0.0102, 0.0102 × 0.462 = 0.004712. Therefore, $\hat{\lambda} = 0.4973 + 0.004712 = 0.5020$.

H-4.5. *Atchison's Method*. Previous adjustments to the mean and variance assumed that the data values really were present, but could not be recorded or seen as they were below the DL. In other words, if the DL had been substantially lower, the data values would have been recorded. There are cases, however, where the data values are below the DL because they are actually not present, the contaminant or chemical of concern being entirely absent. The investigator may have reason to believe that the contaminant is absent, but is unable to prove it is below the analytical DLs. Such data sets are actually a mixture—partly the assumed distribution (for example, a normal distribution) and partly a number of real zero values. Atchison's method is used in this situation to adjust the mean and variance for the zero values. It should also be noted that Atchison's method differs from Cohen's method, in that, for Atchison's method, a normality test is performed for the detected results only.

H-4.5.1. Atchison's method for adjusting the mean and variance of the values above the DL works quite well provided the percentage of NDs is between 15 and 50% of the total number of values. Care must be taken when using Atchison's adjustment because the mean is reduced and variance increased. With such an effect, it may become very difficult to use the adjusted data for tests of hypotheses or for predictive purposes.

H-4.5.2. As a diagnostic tool, Atchison's adjustment can lead to an evaluation of the data to determine if two populations are being sampled simultaneously: one population being represented by a normal distribution, the other being simply blanks. In some circumstances, such as investigating a hazardous site, it may be possible to relate the position of the sample through a posting plot and determine if the target population has not been adequately stratified. Directions for Atchison's method are contained in Paragraph H-4.6, and an example is contained in Paragraph H-4.7.

H-4.6. *Directions for Atchison's Method to Adjust Means and Variances*. Let $x_1, x_2, \ldots, x_m, \ldots, x_n$ represent the data points where the first m values are above the DL and the remaining (n – m) data points are below the DL.

H-4.6.1. Using the data above the detection level, verify this subset of data follows a normal distribution.

H-4.6.2. Using the data above the detection level, compute the sample mean,

$$\bar{x}_d = \frac{1}{m}\sum_{i=1}^{m} x_i$$

and the sample variance,

$$s_d^2 = \frac{\sum_{i=1}^{m}(x_i - \bar{x}_d)^2}{m-1}.$$

H-4.6.3. Estimate the corrected sample mean,

$$\bar{x} = \frac{m}{n}\bar{x}_d$$

and the sample variance,

$$s^2 = \frac{m-1}{n-1}s_d^2 + \frac{m(n-m)}{n(n-1)}\bar{x}_d^2.$$

H-4.7. *Example for Atchison's Method to Adjust Means and Variances*. Atchison's method will be used to adjust the mean and standard deviation of the groundwater concentrations for benzene at Site A and well MW03, presented in Paragraph H-4.3.

H-4.7.1. So, $n = 15$, $m = 10$, and $n - m = 5$.

H-4.7.2. According to Paragraph H-4.3, the detected results from this data set follow a lognormal distribution; so, the log-transformed data will be used to adjust the mean and variance. The sample mean and variance based on just the data above the detection level are

$$\bar{x}_d = -0.1650$$

and

$$s_d^2 = 1.683.$$

H-4.7.3. The corrected sample mean and variance (in the log-scale) are:

$$\bar{x} = \frac{10}{15} \times (-0.1650) = -0.1100$$

$$s^2 = \frac{10-1}{15-1} \times 1.683 + \frac{10(5)}{15(15-1)} \times (-0.1650)^2 = 1.088 \ .$$

H-4.8. *Selecting Between Atchison's Method or Cohen's Method*. To determine if a data set is better adjusted by Cohen's method or Atchison's method, a simple graphical procedure using a normal probability plot can be used. Directions for this procedure are given in Paragraph H-4.9, and an example is contained in Paragraph H-4.10.

H-4.9. *Directions for Selecting Between Atchison's Method or Cohen's Method*. Let $x_1, x_2, \ldots, x_m, \ldots, x_n$ represent the data points with the first *m* values above the DL and the remaining *n-m* data points below the DL.

H-4.9.1. Use Paragraph H-4.3 to construct a Normal Probability Plot using all the data, but only plot the values above the detection level. This is called the Censored Plot.

H-4.9.2. Use Paragraph H-4.3 to construct a Normal Probability Plot using only those values above the detection level. This is called the Detects Only Plot.

H-4.9.3. 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 Atchison's method to estimate the sample mean and variance.

H-4.10. *Example for Selecting Between Cohen's Method or Atchison's Method*.

H-4.10.1. This comparison will be made with the groundwater concentrations for benzene at Site A and well MW03, based on the log-transformed data presented in Paragraph H-4.3.

H-4.10.2. Using Paragraph H-4.3, we constructed normal probability plots based on the log-transformed data, as the data seem to follow a lognormal distribution based on the Shapiro-Wilk test. The Figure H-1 shows these plots. The Censored Plot was developed with just the detected results. The Detects Only Plot was developed with all of the data (using the DL as a proxy value), but only the detected results were plotted. The Detects Only Plot appears to fit a line better than the Censored Plot, so Atchison's Method seems to be the more appropriate method to estimate the sample mean and variance.

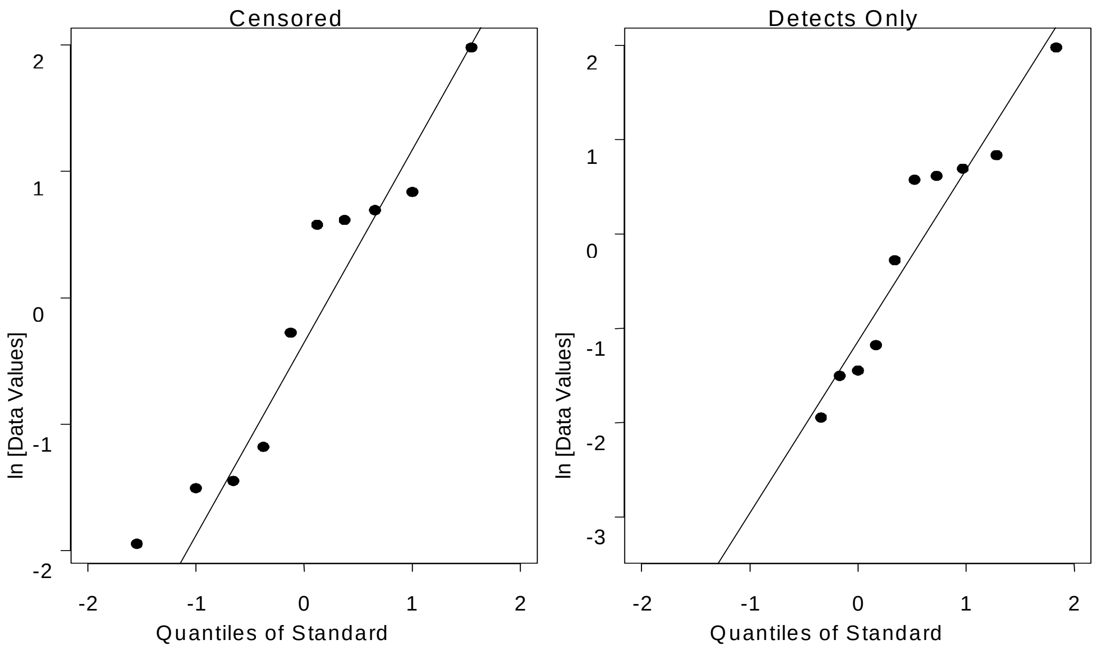

Figure H-1. Example of selecting between Atchison's method and Cohen's method.

H-4.11. *Trimmed Mean*. Trimming discards the data in the tails of a data set to develop an unbiased estimate of the population mean. This method is considered useful when the data set is generally symmetric, and there are concerns about outlier data that might be mistakes or otherwise unexplainable.

H-4.11.1. For environmental data, NDs usually occur in the left tail of the data, so trimming the data can be used to adjust the data set to account for NDs when estimating a mean. Developing a p100% trimmed mean involves trimming p100% of the data in both the lower and the upper tail. Note that p must be between 0 and 0.5 as p represents the portion deleted in both the upper and the lower tail. After np of the largest values and np of the smallest values are trimmed, there are $n(1 - 2p)$ data values remaining where n represents the original number of samples.

H-4.11.2. The proportion trimmed depends on the total sample size (n), as a reasonable number of samples must remain for analysis. For approximately symmetrical distributions, a 25% trimmed mean (the mid-mean) is a good estimator of the population mean. However, environmental data are often skewed (asymmetrical), and in these cases a 15% trimmed mean may be a better estimator of the population mean. It is also possible to trim the data only to replace the NDs. For example, if 3% of the data are below the DL, a 3% trimmed mean could be used to estimate the population mean. Directions for developing a trimmed mean are contained in Paragraph H-4.12, and an example is given in Paragraph H-4.13. A trimmed variance is rarely calculated and is of limited use.

H-4.12. *Directions for Developing a Trimmed Mean*. Let $x_1, x_2, \ldots, x_n$ represent the *n* data points. To develop a *p*100% trimmed mean ($0 < p < 0.5$):

H-4.12.1. Let *j* represent the integer part of the product *np*. For example, if $p = 0.25$ and $n = 17$, $np = (0.25)(17) = 4.25$, so $j = 4$.

H-4.12.2. Delete the *j* smallest values of the data set and the *j* largest values of the data set.

H-4.12.3. Compute the arithmetic mean of the remaining $n - 2j$ values,

$$\bar{x} = \frac{1}{n-2j} \sum_{i=j+1}^{n-j-1} x_i \; .$$

This value is the estimate of the population mean.

H-4.13. *Example for Developing a Trimmed Mean*. For simplicity, a 100*p*% trimmed mean ($0 < p < 0.5$) will be estimated using the benzene data presented in the example in Paragraph H-4.3. As 5 out of 15 of the data are NDs, a 33.3% trimmed mean will be calculated.

$$n = 15$$

$$p = 0.333$$

$$np = 15 \times 0.333 = 5$$

$$j = 5 \, (\text{the integer part of } np) \; .$$

So, the 5 NDs and the 5 largest values of the data set will be removed, and the remaining samples will be used to estimate the average:

$$\bar{x} = \frac{1}{5}(0.308 + 0.143 + 0.235 + 0.759 + 0.222) = 0.3334.$$

H-4.14. *Winsorized Mean and Standard Deviation*. Winsorizing replaces data in the tails of a data set with the next most extreme data value. For environmental data, NDs usually occur in the left tail of the data. Winsorizing can be used to adjust the data set to account for NDs, and the mean and standard deviation can then be computed on the new data set. Directions for Winsorizing data (and revising the sample size) are contained in Paragraph H-4.15, and an example is in Paragraph H-4.16.

H-4.15. *Directions for Developing a Winsorized Mean and Standard Deviation*. Let $x_1, x_2, \ldots, x_m \ldots, x_n$ represent the *n* data points and *m* represent the number of data points above the DL, and hence *n* – *m* below the DL.

H-4.15.1. List the data in order from smallest to largest, including NDs. Label these points $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ such that $x_{(1)}$ is the smallest, $x_{(2)}$ is the second smallest, …., and $x_{(n)}$ is the largest.

H-4.15.2. Replace the *n* – *m* non-detects with $x_{(m+1)}$ and replace the *n* – *m* largest values with $x_{(n-m)}$.

H-4.15.3. Using the revised data set, compute the sample mean, $\bar{x}$, and the sample standard deviation, *s*:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}\ .$$

H-4.15.4. The Winsorized mean $\bar{x}_w$ is equal to $\bar{x}$. The Winsorized standard deviation is

$$s_w = \frac{s(n-1)}{2m-n-1}\ .$$

H-4.16. *Example for Developing a Winsorized Mean and Standard Deviation*. A Winsorized mean and standard deviation will be estimated using the groundwater concentrations for benzene at Site A and well MW03. Table H-4 presents these concentrations ordered from smallest to largest, where the NDs are considered the lowest concentrations. The five NDs are replaced by the smallest detected result (the 6^{th} highest result) of 0.143, and the highest five detected results are replaced with the 10^{th} highest result of 0.759.

Table H-4.
Groundwater Concentrations for Benzene at Site A and Well MW03

Sampling Event	Detected Result (µg/L)	DL (µg/L)	Revised Data (NDs replaced with smallest detected result)
15-Jul-98	ND	0.0375	0.143
05-Nov-01	ND	0.0465	0.143
31-Jan-02	ND	0.0465	0.143
29-Jan-98	ND	0.0605	0.143
17-Jan-01	ND	0.0641	0.143
17-Jul-00	0.143	0.0353	0.143
28-Jul-01	0.222	0.0401	0.222
16-Oct-00	0.235	0.0353	0.235
20-Oct-99	0.308	0.0759	0.308
04-May-01	0.759	0.0401	0.759
18-Apr-98	1.78	0.0375	0.759
18-Jul-99	1.85	0.0759	0.759
01-Apr-00	2.00	0.0504	0.759
18-Oct-98	2.31	0.0375	0.759
18-Apr-99	7.24	0.0469	0.759

$n = 15$, $m = 10$, $n - m = 5$.

H-4.16.1. Using the revised data set, we find the sample mean to be $\bar{x} = 0.4118$; this value is also the Winsorized mean. Using the revised data set, we find the the sample standard deviation to be $s = 0.2970$.

H-4.16.2. The Winsorized standard deviation is

$$s_w = \frac{s(n-1)}{2m-n-1} = \frac{0.2970(15-1)}{(2\times 10)-15-1} = 1.0395.$$

H-4.17. *Nonparametric Procedure*. Another procedure that may be used, when the percent of NDs is between 15 and 50%, is a nonparametric analysis. First, all the data values need to be ordered and then replaced by their ranks. The NDs are then treated as tied values and replaced by their mid-ranks. The ranking procedure and adjustments for tied ranks are routinely performed for non-parametric tests such the Wilcoxon rank sum test.

H-5. 50 to 90% NDs. If more than 50% of the data are below the DL but at least 10% of the observations are quantified, tests of proportions may be used to test hypotheses using the data. 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 DL. For example, if 67% of the data are below the

DL, consider switching the parameter of interest to the 75th percentile. Then, the test of proportion can be applied to test the hypothesis concerning the 75th percentile. It is important to note that tests of proportions may not be applicable for composite samples. In this case, the data analyst should consult a statistician before proceeding with analysis.

H-6. Greater than 90% NDs. The Poisson distribution can be used when 90% or more of the data is non-detected. In this instance, the detected results would be considered the “rare events” as modeled by the Poisson distribution. The Poisson model describes the behavior of a series of independent events over a large number of trials, where the probability of occurrence is low but stays constant from trial to trial. This model represents a counting process where each particle or molecule of contamination is counted separately but cumulatively, so that the counts for detected samples with high concentrations are larger than counts for samples with smaller concentrations. So, the Poisson model maintains the magnitude of detected concentrations. For example, a detected result with a concentration of 100 ppb would have a Poisson count of 100. Counts for non-detected results can be taken as zero or half the DL. The Poisson model is a distribution, like a normal distribution, that can be used to derive summary statistics such as prediction limits and tolerance limits. See Appendix E for a description of the Poisson distribution.

H-7. Recommendations.

H-7.1. If the degree of censoring (the percentage of data below the DL) 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. Using nonparametric analyses is another approach for analyzing such data.

H-7.2. 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 is above (or below) the standard, or to test whether an upper quantile of the population distribution is above the numerical standard.