EEDP-04-23
July 1995

Environmental Effects of Dredging Technical Notes

Guidelines for Statistical Treatment of Less Than Detection Limit Data in Dredged Sediment Evaluations

Purpose

This technical note provides recommendations for methods of handling less than detection limit data to permit statistical comparisons of sediment contaminant or bioaccumulation samples in dredged sediment evaluations. Ten censored data methods are evaluated; performance depends upon data characteristics such as equality of variances, type of frequency distribution, and the proportion of the data that is below detection limit.

Background

Regulatory evaluations of dredged sediments frequently require managers to assess contaminant concentrations in the sediments themselves, or in the tissues of organisms exposed to those sediments, as part of a tiered testing protocol (U.S. Environmental Protection Agency/U.S. Army Corps of Engineers (USEPA/USACE) 1991, 1994). A typical Tier III assessment, for example, includes comparison of contaminant bioaccumulation in organisms exposed to the dredged sediment(s) with bioaccumulation in organisms exposed to a reference sediment. Statistical procedures for performing such comparisons are described in detail in Appendix D of the Inland Testing Manual (USEPA/USACE 1994). However, most statistical protocols of the Inland Testing Manual cannot be applied directly in the common situation where some contaminant concentrations are reported only as less than some numerical detection limit (DL). The actual concentrations of these "censored" data are unknown and are presumed to fall between zero and the DL.

Previous studies (El-Shaarawi 1989; El-Shaarawi and Esterby 1992; Gaskin, Dafoe, and Brooksbank 1990; Gilliom and Helsel 1986; Gleit 1985; Haas and Scheff 1990; Helsel 1990; Helsel and Cohn 1988; Helsel and Gilliom 1986; Kushner 1976; Newman and others 1989; Porter and Ward 1991) have examined a variety of methods for handling data that include nondetects. Some of these studies identified methods that perform well in parameter estimation problems, for example, when a mean contaminant concentration must be estimated to determine compliance with air or water quality standards. Censored data methods recommended for estimation are based on maximum likelihood and regression procedures. However, there is no consensus on which censored data methods should be used when samples must be compared with each other, as in the Tier III bioaccumulation assessments mentioned above, and accurate parameter estimation is unnecessary. The most commonly used methods are the simplest techniques, namely deletion of nondetects or substitution of a constant such as zero, DL, or one-half DL (DL/2) for the unknown observations. Interim guidance in the draft Inland Testing Manual recommended substitution of DL/2 until statistically validated guidelines could be developed.

To address the need for censored data guidelines for sample comparisons in dredged sediment evaluations, a simulation study was conducted to assess the performance of 10 censored data methods. The study procedures and general results have been described elsewhere (Clarke 1994, 1995). The 10 censored data methods are described in this technical note, with recommendations regarding which method to use in specific situations.

Additional Information

For additional information, contact the author, Ms. Joan U. Clarke, (601) 634-2954, or the manager of the Environmental Effects of Dredging Programs, Dr. Robert M. Engler, (601) 634-3624.

Note: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.

Introduction

Monte Carlo simulations were conducted to evaluate the performance of 10 censored data methods using the statistical procedures recommended in the Inland Testing Manual (USEPA/USACE 1994, Appendix D). Specifically, this entailed comparison of one or more dredged sediments with a reference sediment using the Least Significant Difference (LSD) test on untransformed, log-transformed, or rankit-transformed data (refer to the decision tree, Figure D-5A,B of Inland Testing Manual).

Simulations were conducted using equal and unequal variances with several sample sizes, statistical population distributions, and numbers of sediments to be simultaneously compared with a reference. Censoring was imposed at a "detection limit" equivalent to 20, 40, 60, 80, or 95 percent of the reference sediment population for each set of simulations; uncensored data were also analyzed. Parameter specifications for the simulations are described in detail in Clarke (1995). The entire focus of the study was on small sample size, necessitated by the high cost of contaminant residue chemical analysis; equal and unequal sample sizes ranging from three to eight replicates were used in the simulations. A total of 335,000 simulations were performed. Simulation results were verified using 271 comparisons of actual chemical concentration data from sediment and tissue samples analyzed for several dredged material contaminant evaluation projects (Clarke 1995).

In the simulations and verifications, censored data methods were evaluated for power and for type I statistical error rate (α). Power is the probability of the statistical test (in this study, the LSD test) to detect true significant differences. Type I error rate is the probability of the statistical test to falsely detect as significant a difference that does not exist in the populations from which the samples were drawn. By convention, α is generally set to 0.05 in biological testing, that is, a false positive error rate of 5 percent or less is considered acceptable. Ideally, power should be about 95 percent, but this is frequently impossible due to fiscal or logistical constraints on the number of samples that can be collected or analyzed. Censored data methods should be chosen to maximize power, and if possible, minimize α. Although all methods can be expected to lose power as the amount of censoring increases, the best methods should minimize loss of power and inflation of α with increased censoring.

Censored Data Methods

Ten censored data methods amenable to simulations using SAS® (SAS Institute, Inc. 1988a,b,c) were chosen for evaluation:

- **DL.** Substitution of the detection limit for all nondetects.
- **DL/2.** Substitution of one-half the detection limit for all nondetects.
- **ZERO.** Substitution of zero for all nondetects.

When data are subsequently transformed to rankits, the above three methods produce the exact same results (assuming all uncensored observations in the sample are greater than DL), and are called **CONST** for substitution of any constant between 0 and DL.

- **UNIF.** Nondetects are replaced by ordered observations x_i ($i = 1, 2 \ldots nc$, where nc is the number of censored observations in the sample) between 0 and DL, where

  $$x_i = DL(i - 1)/(nc - 1)$$

  and $x_i = DL/2$ when $nc = 1$. This produces a uniform distribution symmetric around DL/2 (Gilliom and Helsel 1986).
- **UNIFR.** Replacement of nondetects by random numbers from a uniform distribution between 0 and DL. This may be done using a random numbers table or a random number generator such as the RANUNI function in SAS (SAS Institute, Inc. 1988c).

- **MLE NORM.** Maximum likelihood estimation of below-DL values assuming a normal distribution, using the SAS LIFEREG procedure (SAS Institute, Inc. 1988a).
- **MLE LOGN.** Maximum likelihood estimation of below-DL values assuming a lognormal distribution, using the SAS LIFEREG procedure (SAS Institute, Inc. 1988a).
- **MLE WEIB.** Maximum likelihood estimation of below-DL values assuming a Weibull distribution, using the SAS LIFEREG procedure (SAS Institute, Inc. 1988a).

In the three MLE methods, the $i = 1,2...nc$ censored observations are replaced by the values corresponding to the first nc of n evenly spaced percentiles of the MLE-generated distribution.

- **NR.** Substitution of estimated values from a normal distribution using linear regression of above-DL concentrations versus their rankits (Gilliom and Helsel 1986).
- **LR.** Substitution of estimated values from a lognormal distribution using linear regression of logarithms of above DL concentrations versus their rankits (Gilliom and Helsel 1986, Clarke 1992).

The regression equation calculated in these methods is used to extrapolate values for the censored observations. For LR, antilogs of the extrapolated values are used.

SAS program statements for the methods described above are provided in Appendix D of USEPA/USACE (1994) or can be obtained from the author. Several other censored data methods are available but were considered unsuitable for this study (Clarke 1995). In particular, deletion of censored data is not recommended as it results in excessive loss of information and power as the amount of censoring increases. Slymen, de Peyster, and Donohoe (1994) describe and recommend tobit analysis using the SAS LIFEREG procedure for comparing samples with values below DL in environmental studies. The authors present statistical justification for this method, but it could not be compared with the other methods described in this technical note due to the limitations of SAS LIFEREG output for conducting large numbers of simulations.

Considerations in Selecting the Best Censored Data Methods

Simulation results clearly indicate that no single censored data method works best in all situations. Before selecting a method for treatment of nondetects in contaminant evaluations, the investigator should determine, if possible, certain characteristics of the data. Are variances equal or unequal among the samples being compared? If variances are equal, what is the coefficient of variation (CV = standard deviation ÷ mean) of the combined samples? If variances are unequal, do they increase as sample means increase, or do they follow no particular pattern in relation to sample means (mixed variances)? When the samples are combined, are the residuals normally distributed, lognormally distributed (that is, do they pass the test of normality following log transformation), or nonnormally distributed? The type of data distribution and the variance characteristics appear to have the greatest influence upon the censored data methods. For the limited ranges considered in this study, sample size and number of treatments being compared seem to have less effect upon the censored data methods.

To determine type of data distribution and variance characteristics for censored samples, investigators can apply two or more of the censored data methods described above to obtain a range of possible variances and CVs. The revised data (both untransformed and log-transformed) can then be tested for normality and equality of variances using procedures such as those described in Appendix D of USEPA/USACE (1994).

When samples are severely censored, investigators may be able to make an educated guess concerning distribution and variance characteristics based on uncensored data for the same contaminant or on historical data from the same location. Of the 271 comparisons performed using real chemical data in the verification study, half had equal variances among the samples being compared, while 30 percent had mixed variances and 20 percent had variances proportional to the sample means. Sixty percent of the samples passed the Shapiro-Wilk's test of normality (USEPA/USACE 1994), 25 percent passed when data were log-transformed, and 15 percent failed. Nevertheless, in the absence of information to the contrary, it may be reasonable to assume a lognormal distribution for environmental trace chemical data (El-Shaarawi 1989; Gilliom, Hirsch, and Gilroy 1984; Kushner 1976; Newman and others 1989; Ott and Mage 1976; Porter and Ward 1991; Travis and Land 1990). A normal distribution is unlikely for contaminant concentration data when the CV exceeds 1, as such a distribution would include a fair amount of negative concentrations. For example, a normal distribution contains ≈17 percent negative values when the CV = 1 and ≈31 percent negative values when the CV = 2.

The next consideration should be the relative importance of power versus type I error rate (α) in the statistical comparisons. The censored data methods were compared based on power adjusted for α (that is, mean power minus mean α). The most powerful methods generally had α in the range of 0.05 to 0.10 for amounts of censoring up through 80 percent, but much higher α at 95-percent censoring. If it is crucial to maintain α at approximately 0.05 or less, it may be necessary to select somewhat less powerful methods in certain cases. In a number of situations, there are no suitably powerful methods with $\alpha \leq 0.05$.

When several methods had adjusted mean power within 0.05 of the uncensored data, priority was given to the simplest method(s). In order of increasing complexity, the censored data methods were constant substitution (DL, DL/2, ZERO), substitution from a uniform distribution (UNIF and UNIFR), regression techniques (NR and LR), and maximum likelihood techniques (MLE LOGN, MLE NORM, and MLE WEIB). In most situations, the simplest methods were also the most powerful.

Table 1. Recommended Censored Data Methods for Small Samples to Be Used in Statistical Comparisons						
Amount of Censoring, %	Variances	Distribution	Coefficient of Variation	Data Transformation[a]		
				Log	None[b]	Rankit
≤20	Equal	All	≤ 0.25	**DL**	**DL**	**CONST, UNIF**
		Normal	0.26 - 1	**DL**	**DL/2, UNIF, DL**	CONST, UNIF
		Lognormal, Nonnormal	0.26 - 1	***DL/2***, **DL**		CONST, UNIF
			>1	***DL/2***, ***DL***, UNIF		CONST, UNIF
	Increase as Means Increase	Normal		***ZERO***[c]	**LR**	—[d]
		Lognormal, Nonnormal		**DL**		CONST, UNIF
	Mixed	Normal		—[d]	**DL**	*CONST,* MLE NORM
		Lognormal, Nonnormal		***DL***[c]		*MLE WEIB*[c]
21 - 40	Equal	All	≤0.25	**DL**	DL	**CONST, UNIF**
		Normal	0.26 - 0.5	DL	**DL/2,** DL	CONST, UNIF
		Lognormal, Nonnormal	0.26 - 1	**DL/2**		CONST, UNIF
			>1	***DL/2***, DL		CONST, UNIF
	Increase as Means Increase	Normal		***ZERO,*** MLE NORM	DL	—[d]
		Lognormal, Nonnormal		**DL, DL/2**		**CONST, UNIF**
	Mixed	Normal		***DL/2***[c]	***ZERO,*** *DL/2*[c]	*CONST,* MLE WEIB
		Lognormal, Nonnormal		**DL**		CONST

(Continued)

[a] Method(s) in bold indicate most powerful transformation(s). Methods in italics have mean α between 0.06 and 0.10; nonitalicized methods have mean $\alpha < 0.06$.
[b] Untransformed data generally should not be used with lognormal or nonnormal distributions.
[c] All methods with acceptable power have $\alpha \geq 0.06$.
[d] All methods have unacceptably low power and/or high α.

Table 1. (Concluded)

Amount of Censoring, %	**Variances**	**Distribution**	**Coefficient of Variation**	**Data Transformation**[a] **Log**	**None**[b]	**Rankit**
41 - 60	Equal	All	≤0.25	DL/2	DL/2	**CONST**
		Normal	0.26 - 1	DL/2	**DL/2**	**CONST**
		Lognormal, Nonnormal	>0.25	**DL/2**		CONST
	Increase as Means Increase	Normal		**DL**	**DL**	CONST
		Lognormal		**DL/2**		**CONST**
		Nonnormal		**DL, DL/2**		CONST
	Mixed	Normal		—[d]	—[d]	*CONST*[c]
		Lognormal		**DL/2**		*CONST*[c]
		Nonnormal		**DL/2**		CONST
61 - 80	Equal	All	≤0.25	*DL/2*[c]	DL/2	**CONST**
		Normal	0.26 - 1	**DL/2**	**DL/2, ZERO**	**CONST**
		Lognormal, Nonnormal	0.26 - 0.5	***DL/2***[c]		**CONST**
			0.51 - 1	**DL/2**		CONST
		Lognormal	>1	***DL/2***[c]		***CONST***[c]
		Nonnormal	>1	—[d]		—[d]
	Increase as Means Increase	Normal		DL	**DL/2, ZERO**	CONST
		Lognormal		**DL/2**		**CONST**
		Nonnormal		**UNIF**		CONST
	Mixed	Normal		—[d]	—[d]	—[d]
		Lognormal, Nonnormal		***DL/2***[c]		***CONST***[c]

Recommendations for Censored Data Methods

Censored data methods recommended for various situations of equal and unequal variances, statistical frequency distributions, CVs, data transformations, and amounts of censoring are given in Table 1. When two or three methods are essentially equivalent in power, type I error rate, and simplicity, all are listed in the table in order of decreasing power. Method(s) highlighted in bold indicate the data transformation(s) having the highest adjusted power in a given situation. Methods in italics have mean α between 0.06 and 0.10; nonitalicized methods have mean $\alpha < 0.06$. When the recommended method has mean $\alpha \geq 0.06$, if possible, an alternative (although usually less powerful) method having lower α is given in the table. Situations in which all methods have unacceptably low power and/or high α are also indicated in the table. Methods having adjusted mean power within 0.05 of the most powerful method for a given censoring percentile and variance-distribution-CV combination and at least half the power of the uncensored data for that combination were considered to have acceptable power.

In most situations shown in Table 1, a single powerful method can be applied regardless of which data transformation, if any, might be needed. For example, when censoring is ≤20 percent, variances are equal, and CV is ≤0.25, DL should be substituted for all nondetects. The tests of assumptions in Appendix D of USEPA/USACE (1994) would then determine whether untransformed, log-transformed, or rankit-transformed data should be used in the statistical comparisons. Alternatively, UNIF could be used with rankits. These methods have approximately equal power. However, if censoring is between 40 and 60 percent, variances are equal, and CV is ≤0.25, CONST with rankits should be preferred, as the power of this combination exceeds that of any method with untransformed or log-transformed data. In cases when power is exceptionally low, especially when variances are unequal, a different method for each transformation may be required to maximize power.

Following is a discussion of the individual censored data methods and the situations in which they should or should not be used.

DL is generally the preferred method at low to moderate proportions of censoring, especially when the CV is low, or when variances are unequal and data are not normally distributed. In particular, DL performs better than all other methods and much better than the other simple substitution methods at ≤40 percent censoring when the CV is extremely low (≤0.25). In most cases DL should not be used with data that are highly censored (>60 percent censoring). DL has low power at ≤40 percent censoring with log transformation when data are normally distributed and variances increase with increasing means.

DL/2 generally begins to surpass DL in power as CV and censoring increase. DL/2 tends to have slightly higher α than DL when variances are equal. DL/2 should not be used when the CV is extremely low (≤0.25) and less than

40 percent of the data are censored. DL/2 also has low power and/or high α at ≤60 percent censoring when data are normally distributed and variances are unequal.

ZERO is recommended for use with untransformed, normally distributed data in a few situations. In general, ZERO should not be used with log-transformed data as this amounts to deletion of the censored data, resulting in low power and high α. One exception, in which ZERO proved to be the most powerful method with log-transformed data, was normal distribution at ≤40 percent censoring when variances increase as means increase. However, α in this case exceeds 0.05.

CONST is almost universally appropriate for rankit-transformed data, and is usually the most powerful method with rankits. In several situations CONST with rankits is equally or more powerful than the best-performing method with untransformed or log-transformed data. However, when data are normally distributed, variances increase with means, and censoring is ≤40 percent, all methods with rankits have unacceptably low power compared with log-transformed and untransformed data. Type I error rate is high for CONST with rankits when variances are mixed and data are normally distributed; in almost all other cases, α does not exceed 0.06.

UNIF is the most powerful method with log-transformed data at high amounts of censoring when data are nonnormal and variances increase as means increase. When used with rankits, UNIF is essentially equal in power to CONST in most situations. Type I error rate tends to be extremely low for UNIF, especially as censoring increases. Therefore, UNIF can be a suitable alternative to the most powerful method in some situations when low α is desired.

UNIFR is generally slightly less powerful, with slightly higher α, than UNIF. Power is low for most situations at 60 percent censoring or more. UNIFR is not the recommended method in any situation.

MLE NORM is recommended in two situations as an alternative to the most powerful method when low α is desired: with rankits at ≤20 percent censoring when variances are mixed and data are normal, and with log-transformed data at 21 to 40 percent censoring when variances increase with means and data are normal. MLE NORM has low power at 60 percent censoring or more, and also in many cases at 40 percent and even 20 percent censoring. MLE NORM should not be used with log-transformed data when the CV is high as this method may substitute negative concentrations for the nondetects.

MLE LOGN is not the most powerful method in any situation. Power is low when censoring exceeds 40 percent, and α tends to be high for log-transformed data in many cases at low amounts of censoring.

MLE WEIB is recommended for rankits as an alternative to CONST at 21 to 40 percent censoring when variances are mixed, data are normally distributed, and low α is required. MLE WEIB should also be used with rankits at $\leq$20 percent censoring when variances are mixed and data are not normally distributed. In most other cases MLE WEIB has less power than MLE LOGN, and is inappropriate for log-transformed data, or for any data when censoring exceeds 40 percent.

LR and **NR** appear to be inappropriate as censored data methods for statistical comparisons of small samples in most circumstances. Power is generally low even at 20 percent censoring, and declines precipitously as censoring increases. Conversely, α is generally high even at 20 percent censoring and increases dramatically as censoring increases, sometimes approaching 1. LR is recommended only for untransformed data at 20 percent censoring when variances are proportional to means and data are normally distributed.

The simple substitution (DL, DL/2, ZERO, CONST) and uniform distribution (UNIF, UNIFR) methods can be applied regardless of the amount of censoring. The MLE methods cannot be used when all observations in a sample are below DL. The regression methods (LR, NR) require at least three uncensored observations in each sample, and thus are inapplicable for small sample sizes when censoring exceeds about 20 percent.

Verifications

Verification results overwhelmingly support the simulation study conclusions that simple substitution or uniform distribution methods work best in most situations to prepare censored samples for statistical comparisons. In no case did the maximum likelihood or regression techniques have sufficient power in the verifications to be considered useful. Verification results favor the use of DL at 20 to 40 percent censoring when the CV is low ($\leq$0.25), and DL/2 otherwise. Although generally less powerful than DL/2, UNIF and UNIFR have low α and perform well at 20 to 40 percent censoring when log transformation is not used. ZERO also performs well, especially at 40 to 60 percent censoring, but should not be used with log transformation. No methods have sufficient power to be useful at 80 percent censoring except DL/2 when the CV is high (>0.75).

Summary

Simulation and verification results indicate that, in most cases, the sophisticated statistical techniques recommended for estimation problems involving censored data are unnecessary or even inappropriate for statistical comparisons of small, censored data samples. In general, the simple substitution methods work best to maintain power and control type I error rate in statistical comparisons. Recommended steps in selecting the best censored data method for a given situation are listed below.

For each contaminant for which some data are reported as nondetect or <DL:

- Determine proportion of data that are censored (all samples combined).
- Determine whether variances are equal or unequal among samples. If unequal, do the variances increase as means increase, or are the variances seemingly random (mixed)?
- Determine CV of combined samples.
- Determine whether combined sample residuals are distributed normally, lognormally, or nonnormally. If $CV \geq 1$, assume lognormal or nonnormal distribution.
- Refer to Table 1 to determine most appropriate method given the amount of data censoring, properties of variances, and type of statistical distribution. Where possible, preference should be given to methods in bold.
- If it is crucial to maintain α at approximately 0.05 or less, choose nonitalicized methods where available in Table 1.
- Apply selected method to censored data, then continue with tests of assumptions and statistical comparison procedures as outlined in USEPA/USACE (1994). Avoid a data transformation for which no method is given in Table 1 due to low power or excessively high α.
- Do not attempt statistical comparisons of severely censored samples in situations where no censored data methods are considered appropriate. In such cases, the probability of an erroneous outcome is high.

If it is impossible to determine characteristics of the variances or statistical distribution for censored data samples, use DL for up to 40 percent censoring or DL/2 for 40 to 80 percent censoring. An alternative, although somewhat less powerful in many situations, is to substitute any constant between 0 and DL, convert the data to rankits, and then follow the nonparametric decision procedures in Figure D-5B of USEPA/USACE (1994). Power loss using CONST with rankits, when compared with DL or DL/2 on untransformed or log-transformed data, is generally around 5 to 10 percent when variances are equal and data are lognormally or nonnormally distributed, <4 percent when variances are equal and data are normally distributed, up to 14 percent when variances are proportional to the means, and up to 6 percent when variances are mixed. No matter what technique is used, power will generally decline as censoring increases. Beyond 60 to 80 percent censoring, it is unlikely that any technique will perform acceptably.

It is quite possible that an evaluation including a number of sediments and contaminants would produce comparisons involving several different combinations of censoring proportions, variance characteristics, and data frequency distributions. Following the guidelines herein would result in the application of more than one censored data method to the project data. This is entirely acceptable when the censored data methods are selected for the purpose of maximizing power and minimizing type I error. *What is not acceptable is to try several censored data methods for the purpose of finding one that will produce a desired statistical comparison outcome.*

The simulation study did not address the performance of censored data methods in the common situation of multiple detection limits within a set of replicate observations. However, the simple substitution methods shown to work well in nearly all cases with single-detection limit censored samples can be applied without modification to multiple-detection limit censored samples.

References

Clarke, J. U. 1992. "Analyzing "less than" data: An example from sediment bioaccumulation testing," *Water Quality '92, Proceedings of the 9th Seminar,* 16-20 March 1992, San Antonio, TX, Miscellaneous Paper W-92-3, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, pp 178-191.

Clarke, J. U. 1994. "Evaluating methods for statistical analysis of less than detection limit data using simulated small samples; General results," *Dredging '94, Proceedings of the Second International Conference on Dredging and Dredged Material Placement,* Vol 1, E. C. McNair, Jr., ed., American Society of Civil Engineers, New York, pp 747-755.

Clarke, J. U. 1995. "Evaluating methods for statistical analysis of less than detection limit data using simulated small samples; Study description," *Water Quality '94, Proceedings of the 10th Seminar,* 14-18 February 1994, Savannah, GA, Miscellaneous Paper W-95-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS, pp 330-335.

El-Shaarawi, A. H. 1989. "Inferences about the mean from censored water quality data," *Water Resources Research* 25(4):685-690.

El-Shaarawi, A. H., and Esterby, S. R. 1992. "Replacement of censored observations by a constant: An evaluation," *Water Research* 26(6):835-844.

Gaskin, J. E., Dafoe, T., and Brooksbank, P. 1990. "Estimation of analytical values from sub-detection limit measurements for water quality parameters," *Analyst* 115:507-510.

Gilliom, R. J., and Helsel, D. R. 1986. "Estimation of distributional parameters for censored trace level water quality data; 1. Estimation techniques," *Water Resources Research* 22(2):135-146.

Gilliom, R. J., Hirsch, R. M., and Gilroy, E. J. 1984. "Effect of censoring trace-level water-quality data on trend-detection capability," *Environmental Science and Technology* 18(7):530-535.

Gleit, A. 1985. "Estimation for small normal data sets with detection limits," *Environmental Science and Technology* 19:1201-1206.

Haas, C. N., and Scheff, P. A. 1990. "Estimation of averages in truncated samples," *Environmental Science and Technology* 24(6):912-919.

Helsel, D. R. 1990. "Less than obvious," *Environmental Science and Technology* 24(12):1766-1774.

Helsel, D. R., and Cohn, T. A. 1988. "Estimation of descriptive statistics for multiply censored water quality data," *Water Resources Research* 24(12):1997-2004.

Helsel, D. R., and Gilliom, R. J. 1986. "Estimation of distributional parameters for censored trace level water quality data; 2. Verification and applications," *Water Resources Research* 22(2):147-155.

Kushner, E. J. 1976. "On determining the statistical parameters for pollution concentration from a truncated data set," *Atmospheric Environment* 10:975-979.

Newman, M. C., Dixon, P. M., Looney, B. B., and Pinder, J. E., III. 1989. "Estimating mean and variance for environmental samples with below detection limit observations," *Water Resources Bulletin* 25(4):905-916.

Ott, W. R., and Mage, D. T. 1976. "A general purpose univariate probability model for environmental data analysis," *Computers and Operations Research* 3: 209-216.

Porter, P. S., and Ward, R. C. 1991. "Estimating central tendency from uncensored trace level measurements," *Water Resources Bulletin* 27(4):687-700.

SAS Institute, Inc. 1988a. "SAS/STAT User's Guide, Release 6.03 Edition," Cary, NC.

SAS Institute, Inc. 1988b. "SAS Procedures Guide, Release 6.03 Edition," Cary, NC.

SAS Institute, Inc. 1988c. "SAS Language Guide for Personal Computers, Release 6.03 Edition," Cary, NC.

Slymen, D. J., de Peyster, A., and Donohoe, R. R. 1994. "Hypothesis testing with values below detection limit in environmental studies," *Environmental Science and Technology* 28(5):898-902.

Travis, C. C., and Land, M. L. 1990. "Estimating the mean of data sets with nondetectable values," *Environmental Science and Technology* 24(7):961-962.

U.S. Environmental Protection Agency/U.S. Army Corps of Engineers. 1991. "Evaluation of dredged material proposed for ocean disposal (Testing manual)," EPA-503/8-91/001, USEPA Office of Marine and Estuarine Protection, Washington, DC.

U.S. Environmental Protection Agency/U.S. Army Corps of Engineers. 1994. "Evaluation of dredged material proposed for discharge in waters of the U.S.: Testing manual (Draft)," U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.