Statistical procedures for addressing nondetect results in environmental analyses

Steven W. Hinton

ABSTRACT: Censor data sets, those containing observations that are less than the analytical chemistry limit of detection (LOD) value, are becoming more common as regulatory agencies and dischargers strive to reduce concentrations of trace organic contaminants in treated effluent. The recently reissued "Technical Support Document For Water Quality Based Toxics Control," a frequently used reference during NPDES permit development, states that the mean and variance of censored data sets may be determined by assuming the data conform to what EPA has termed the "delta log-normal distribution" (D-LOG). The USEPA D-LOG statistical procedure along with two commonly used alternatives (Cohen's MLE and regression of normal order scores) are described, illustrated with example calculations and evaluated using Monte Carlo simulation experiments. Using minimal absolute bias as the evaluation criteria, Cohen's method [Technometrics 1: 217 (1959)] had superior predictive capability overall, particularly for real world data sets. For mean and standard deviation estimators, Cohen's method was at least 2.5 and 1.2 times less biased than those of the EPA recommended methodology, respectively.

KEYWORDS: Averages, computation, data, evaluation, simulation, statistical methods, statistics.

Data sets of chemical concentrations which contain observations that are less than the analytical chemistry limit of detection (LOD) value are becoming increasingly more common as regulatory agencies and dischargers strive to reduce the concentrations of trace organic contaminants in treated effluent. Observations quantified as <LOD or nondetect (ND) arise because the compound being measured is either absent or exists at such low concentrations that the chemist judges the analytical result unreliable. Data censoring by the chemist can occur for a variety of reasons including noncompliance with quality control/quality assurance criteria such as low signal-to-noise ratio or poor recovery of spiked compounds.

The presence of <LOD or ND observations in a data set makes estimation of the parent population's parameters (e.g., μ and σ) more involved because the lack of quantitation associated with the censored observations prohibits their direct incorporation into the simple straight-forward calculations of common statistics such as the arithmetic mean (i.e., $\bar{x} = \sum x_i/n$). For most environmental monitoring situations, the censored observations represent random values occurring between the LOD and zero. Although substitution of a fixed value for censored observations followed by common statistical calculations is an expedient (and common) way of converting a superficially cumbersome analysis into a simple one, such an approach provides biased estimates of the parent population characteristics and is not recommended since statistical methods with only modest computational burdens are available *(1, 2)*. Statistical method bias

Hinton is research engineer, NCASI, Tufts University, Civil Engineering Dept. Medford, MA 02155.

can be a particularly important issue when the decrease of an environmental variable is being monitored because the degree of bias increases with increasing numbers of censored observations in the data set, resulting in a distorted picture of the environmental variable's change.

The purpose of this paper is to describe three commonly used methods for estimating a parent population's mean and standard deviation from a data set that is left censored at a single value, and to quantify the bias and root mean square error (RMSE) of each estimator based on the results of Monte Carlo simulation experiments.

Background and theory

The methods described and evaluated in this report can be used to estimate the mean and variance of a normally distributed population (or one which can be made normal through transformation) from a data set containing both fully quantified measurements and left censored (i.e., nondetect or less than) observations, provided the fully quantified measurements are greater than a single threshold value which exceeds the highest detection limit value. Prior to introducing the calculation procedures, the concept of data censoring type and its role in method selection is introduced through a discussion of typical environmental monitoring situations occurring in regulatory related analyses of receiving water and effluent quality conditions.

Problem setting

The appropriate statistical method to use in estimating a population's mean and variance parameters from a censored data set depends in part on the numbers and magnitudes of censoring levels. **Table IA** shows four typical censored data sets from the chemical analysis of a biologically treated effluent for chlorophenols. A prudent first step for any censored data set analysis situation is to rank the observations from lowest to highest, as shown in **Table IB**. Ranking the observations permits ready determination of whether the analysis can be conducted assuming that all censored observations are less than and all fully quantified measurements are greater than a single threshold, permitting the use of "single-censoring-point" (SCP) statistical methods which can be applied with only the aid of a hand calculator and do not require burdensome iterative calculations.

In contrast, the analysis of data sets with fully quantified measurements intermixed with nondetected observations is most accurately performed by statistical methods which are tedious and employ iterative calculations that are typically performed by specialized computer software. Due to their most common application, such methods are often termed "multiple-censoring-point" (MCP) statistical methods, although a single censoring level can necessitate their use (e.g., data set 4 in Table IB). Except for as described below, the statistical methods described in this report should not be applied to analysis situations involving multiple detection limits.

Despite the apparent limitation imposed by single censoring threshold criteria, SCP statistical methods are applicable to many practical analysis situations using common sense and exploiting the MCP method's solution characteristics. The analysis of the data sets depicted in Table IB increase in difficulty from left to right. Data set 1 with only a

I. Typical water quality data sets

A. As Received

Set 1	*Set 2*	*Set 3*	*Set 4*
3	ND(11)	ND(50)	161
1	ND(10)	281	ND(100)
2	18	ND(500)	203
ND(1)	ND(11)	104	42
ND(1)	ND(12)	114	42
1	ND(12)	ND(50)	37
ND(1)	ND(12)	89	
ND(1)	16	83	
ND(1)	ND(12)	61	
	ND(12)	80	
	ND(10)	134	
	15	56	

B. Ordered for Analysis

Set 1	*Set 2*	*Set 3*	*Set 4*
ND(1)	ND(10)	ND(50)	37
ND(1)	ND(10)	ND(50)	42
ND(1)	ND(11)	56	42
ND(1)	ND(11)	61	ND(100)
ND(1)	ND(12)	80	161
1	ND(12)	83	203
1	ND(12)	89	
2	ND(12)	104	
3	ND(12)	114	
	15	134	
	16	281	
	18	ND(500)	

single censoring level is conceptually the easiest to analyze and unconditionally meets SCP criteria. Although data set 2 contains nondetected values at three censoring thresholds and could be analyzed with a MCP method, a valid analysis can be made with a SCP method by assuming a single censoring level of 12 since observations that are less than 10 or 11 are also less than 12; the error introduced by this approach is negligible, particularly when compared to the uncertainty associated with measurement near the LOD. When multiple censoring levels are present in a data set, a valid analysis with a SCP method is sometimes possible because "nondetected" observations with censoring thresholds greater than either the largest fully quantified measurement or sum of the mean plus two standard deviations, contribute very little to a MCP method's estimate of parent population parameters. Thus, the nondetected value that was less than 500 can be legitimately eliminated from data set 3 and an analysis conducted with a SCP method without imposing a significant calculation error since the <500 censoring threshold has a greater magnitude than the largest fully quantified measurement value of 281. Data set 4 cannot be analyzed adequately with a SCP method because: (a) the nondetected observation can not legitimately be eliminated due to the censoring level proximity to the population's median, and (b) assuming that the 37, 42, and 42 values were "nondetected" would necessitate the perilous action of estimating the parent population's parameters from only two data points when the original data set contained 5 fully quantified measurements. In general, a minimum of 3 or more fully quantified measurements *(3)* and less than 60% data censoring *(4)* are needed to obtain reasonably accurate estimates of parent population parameters.

Statistical method calculation procedures

At least seven generic types of statistical methods for left censored data analysis exist *(1)*. These include: (a) maximum likelihood estimators (MLE), (b) regression of normal order statistics, (c) USEPA D-Log procedure, (d) delta-log normal statistics, (e) balancing techniques, (f) graphical approaches, and (g) replacement/deletion. Due to availability of comprehensive descriptions *(1, 2)*, only the first three methods will be described and evaluated in this report since they are perceived to be the most widely used.

To simplify the presentation in the following sections, the methods are described and illustrated by reference to the analysis of normally distributed data, though this condition occurs infrequently in typical environmental data analysis. This presentation strategy was used because with the exception of the USEPA D-LOG procedure, statistical methods for censored data analysis assume normally distributed parent populations and are concisely illustrated using such data. However, it is frequently necessary to transform real environmental data before analysis in order to approximate normality; typically the logarithmic transformation of $Y_i = \ln(X_i)$ is used, although other transformations are certainly possible and perhaps preferable. When the log normal transformation is used prior to censored data set analysis, it is necessary to transform the analysis results back to the original scale of measurement using the following equations:

$$\overline{X} = \exp\overline{Y}\,[+ 0.5\, s_y^2] \qquad (1)$$

$$s^2_X = \exp[2\overline{Y} + 0.5\, s_Y^2]\,(\exp s_Y^2 - 1.0) \qquad (2)$$

Cohen's maximum likelihood estimator (MLE) method. A sample of n observations is subject to the restriction that full measurement of the variable is possible only if $y \geq y_0$, where y_0 is a known and fixed point of censoring. For typical water quality applications, y_0 is the LOD and it is assumed that the same LOD applies to each observation. Of a total of n observations, n_c observations have $y \leq y_0$ and are censored. The number of observations with $y > y_0$ is $k = n - n_c$. The fraction of censored data is $h = n_c/n$. Using the k fully measured observations, calculate a crude estimate of the mean and variance: $\overline{y} = \Sigma\, y_i/k$ and $\hat{s}^2 = \Sigma\,(y_i - \overline{y})^2/k$. Note that $\hat{s}^2$ is calculated using a denominator of k and not $k - 1$.

For the case where the left tail of the distribution is censored, this estimate of the mean will be too high and the estimate of the variance will be too low. These estimates are adjusted to get the maximum likelihood estimator (MLE) equations for the estimated mean, $\mathbf{m}_{MLE}$, and variance, $\mathbf{s}^2_{MLE}$:

$$\mathbf{m}_{MLE} = \overline{y} - \hat{\lambda}(\overline{y} - y_o) \qquad (3)$$

$$\mathbf{s}^2_{MLE} = \hat{s}^2 + \hat{\lambda}(\overline{y} - y_o)^2 \qquad (4)$$

The $\hat{}$ symbol indicates that the value is an estimate.

Cohen *(5)* provided tables of $\hat{\lambda}$ as a function of h and $\gamma = \hat{s}^2/(\overline{y} - y_o)^2$, which he calls the auxiliary estimation function. Haas and Scheff *(2)* developed a numerical approximation to these tables using a power series expansion, given below, which they state can be used to estimate $\hat{\lambda}$ to within 6% relative error of the tabulated values.

$$\ln\hat{\lambda} = 0.182344 - 0.3756/(\gamma+1) + 0.10017\gamma + 0.78079\omega - 0.00581\gamma^2 - 0.06642\omega^2 - 0.0234\gamma\omega + 0.000174\gamma^3 + 0.001663\gamma^2\omega - 0.00086\gamma\omega^2 - 0.00653\omega^3 \qquad (5)$$

where γ is defined above and $\omega = \ln[h/(1-h)]$.

For an example application of Cohen's method, consider the 27 data points in **Table II** including 23 measured and 4 censored values, with a

censoring threshold, $y_o = 6$. The measured values are from a normal distribution. The average and variance of the $k = 23$ observations with measurable values are $\bar{y} = \Sigma y_i/k = 7.9$ mg/L and $\hat{s}^2 = \Sigma(y_i - \bar{y})^2/k = 1.1078$. The proportion of censored data is $h = 4/27 = 0.1481$ and $\gamma = {}^2/(\bar{y} - y_o)^2 = 1.1078/(7.9 - 6.0)^2 = 0.30687$. The value of the adjustment factor $\hat{\lambda}$ is found by interpolating values from Cohen's tabular values to find $\hat{\lambda} = 0.20117$. The estimates of the mean and variance for this example are:

$$\mathbf{m}_{MLE} = \bar{y} - \hat{\lambda}(\bar{y} - y_o)$$
$$= 7.9 - 0.20117(7.9 - 6.0) = 7.52$$

$$\mathbf{s}^2_{MLE} = \hat{s}^2 + \hat{\lambda}(\bar{y} - y_o)^2$$
$$= 1.1078 + 0.20117(7.9 - 6.0)^2 = 1.8340.$$

Gilliom and Helsel *(6)* have used Cohen's method and several other approaches on data from different distributions and having different levels of censoring. They showed that while Cohen's method is not superior in every case, it does a very good job on most data sets, and is excellent when the data are distributed normally or log normally.

Details of the procedure's equations derivation and additional sample calculations are available *(5)*. A spreadsheet program and a FORTRAN program to calculate Cohen's estimators for normal or log normal data are also available *(1)*.

Regression of normal order statistics (RNOS). This method uses regression to fit a line to a probability plot of the fully measured values. First, the cumulative probabilities are determined by counting the number of samples in which the concentration does not exceed a specified concentration and using this to estimate the probability of the particular value being exceeded. To make such a probability statement, it is necessary to make some provision for having a finite number of observations. One way of doing this is to calculate the probability as $p_i = (i - \alpha)/(n + 1 - 2\alpha)$, where n is the sample size and α is a scaling parameter of magnitude 0, 0.375, or 0.5. For the small data sets ($n<25$) typically found in environmental analyses, a α value of 0.375 is recommended although $\alpha = 0.5$ is frequently used by statisticians, and in many engineering texts the cumulative probabilities are computed with $\alpha = 0$. For many analysis situations, there is no important difference between the probabilities computed using these three methods *(1)*.

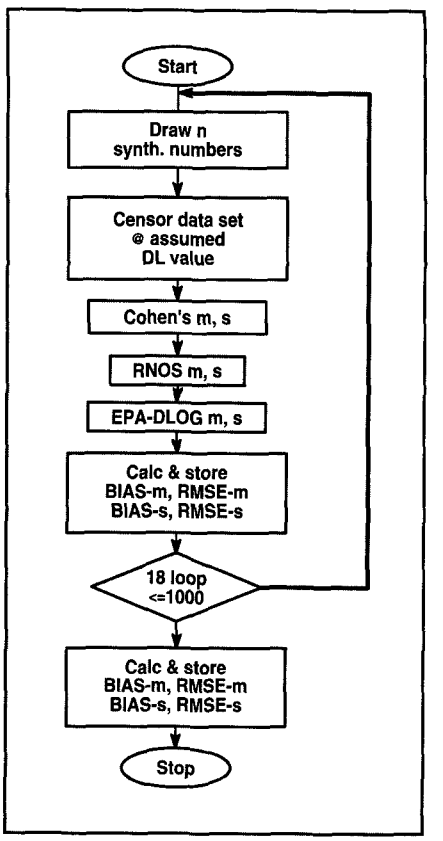
If the data are normally distributed, or have been transformed to make them normal, the probabilities, p_i, are converted to normal order scores, $R_i = F^{-1}(p_i)$, where F^{-1} is the inverse cumulative normal probability distribution and p_i is the plotting position of the i^{th} ranked observation. This is equivalent to rescaling the graph in terms of standard deviations instead of probabilities; (i.e., the standard deviation's magnitude is represented by a normal order statistic of 1.0). The normal order statistics are also known as rankits; these are tabulated in standard statistical tables for $n \leq 50$.

The regression model $Y_i = B_1 + B_2 R_i + e_i$ is fitted to the normal order scores of the noncensored portion of the data to estimate B_1 and B_2. The e_i are the discrepancies between the fitted line and the observed values. For normally distributed data, the coefficients B_1 and B_2 are, respectively, estimates of the mean ($\mathbf{m}_{RNOS}$) and standard deviation (s_{RNOS}) of the uncensored distribution. If the data have been transformed to obtain normality, then B_1 and B_2 estimate the mean and standard deviation in the transformed measurement scale. Rankits are symmetrical about zero so the 50th percentile corresponds to $R_i = 0$. For the normal distribution, the 50th percentile is the mean, and also the median. For the lognormal distribution, the median is the geometric mean and it must be transformed back into the original measurement scale to obtain the estimated mean.

II. Normally distributed example calculation data set

<6	<6	<6	<6	6.1	6.3	6.5	6.7	6.9	
7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0	
8.1	8.3	8.5	8.7	8.9	9.2	9.4	9.6	10.1	

1. Simulation study strategy



The RNOS technique applied to the example normal data in TableII yields the following fitted model:

$$Conc_i = 7.5456 + 1.3662 * (Rankit_i) \quad (6)$$

Thus, the estimated mean is $\mathbf{m}_{RNOS} = 7.5456$ and the estimated standard deviation is $s_{RNOS} = 1.3662$ (i.e., $s^2_{RNOS} = 1.8665$).

USEPA D-LOG procedure. The USEPA *(7)* has developed a variant of the statistical method for the delta-log normal distribution *(8)* herein termed the D-LOG procedure. The distinguishing characteristic of D-LOG procedure is that a certain proportion, δ, of all observations (i.e., fraction of censored values) are as-

sumed to be concentrated at D = LOD (instead of distributed below the LOD) with the remaining observations distributed log normally. With this assumption, the mean and variance estimators given by USEPA [7] are:

$$\mathbf{m}_{\text{D-LOG}} = \delta D + (1 - \delta) \exp(\bar{y} + 0.5s^2) \quad (7)$$

$$\mathbf{s}^2_{\text{D-LOG}} = (1 - \delta) \exp(2\bar{y} + s^2)[\exp s^2 - (1 - \delta)] + \delta(1 - \delta)D[D - 2\exp(\bar{y} + 0.5s^2)] \quad (8)$$

In the above, $\delta = (n\text{-}k)/n$, $\bar{y} = (\Sigma y_i)/k$, $s^2 = \Sigma(y_i - \bar{y})^2/(k - 1)$, $y_i = \ln(x_i)$, and summations are over the k fully quantifiable measurements of the n total observations. Consequently, a sample's estimated mean is the weighted combination of (a) the mean estimated for the fully quantifiable measurements assuming they are log normally distributed, and (b) the mean assigned for the nondetected observations assuming a point distribution occurs at the LOD (i.e., nondetected observations are assigned the LOD value but not transformed and treated as fully quantifiable measurements).

For consistency, the example data in Table II is used to illustrate an application of the D-LOG procedure. The 23 fully quantified measured values are used to compute $\delta = 0.1481$, $\bar{y} = 2.0580$, and $s^2 = 0.017731$ while the four values <6 in the example data are assumed to be concentrated at D = 6.0. The estimated population parameters are $\mathbf{m}_{\text{D-LOG}} = 7.6213$ and $\mathbf{s}^2_{\text{D-LOG}} = 1.4527$, which are in the original scale of measurement. These values are surprisingly similar to the MLE and RNOS estimates considering the data were normally distributed (i.e., not a mixture of variates from log normal and discrete distributions). The simulation study results presented subsequently suggest that this occurrence resulted from the relatively small amount of censoring of the data set.

Simulation study

Methodology

Using Monte Carlo simulation studies, the bias (i.e., $\Sigma[x_i - mu]/1000$) and RMSE (i.e., $[\Sigma[x_i - mu]^2/1000]^{0.5}$) of the mean and standard deviation estimators were determined for Cohen's method, the RNOS technique and the USEPA D-LOG procedure. Each simulation analyzed 1000 sets of n randomly drawn values from a distribution with known mean and standard deviation (SD). Five levels of censoring (5, 20, 40, 60, and 80 percentile); sample sizes of 10, 15, and 20; and four parent population characteristics were evaluated. Sampling distributions included log normal with coefficients of variation (CV) of 0.3 and 1.0 as well as two data-defined distributions created from an industrial effluent discharge record of 102 fully quantified chlorophenol measurements. Simulations were programmed in FORTRAN and executed on an IBM-PC compatible computer using the strategy outlined in **Fig. 1**.

Results and discussion

Table III shows the simulation results for the ideal test cases assuming log normally distributed populations with mean of 2.0 and CVs of 0.3 or 1.0. In these, the average bias (i.e., magnitude and direction of average prediction error) and RMSE (i.e., magnitude of prediction error variability) of the mean and SD estimators are related to the assumed detection limits expressed in terms of percentiles. For the sample sizes of 10, all estimators of the mean are positively biased (i.e., over estimate) at all levels of censoring and bias increases with increased censoring and population variability (i.e., increased CV). In general, the D-LOG estimator is more biased than the MLE estimator and the RNOS estimator for the CV = 0.3 condition; the difference in bias can be particularly large (2.5X+) when the censoring point is greater than the 40th percentile. The RMSEs of all mean estimators are similar for censoring less than approximately the distribution's 40th percentile, above which the D-LOG and RNOS RMSE become inferior to the MLE RMSE.

The SD estimators for sample sizes of 10 were both positively and negatively biased depending on method type and CV condition (Table III). For CV = 0.3 condition, all SD estimators were negatively biased with the absolute magnitude of the D-LOG bias always significantly greater than the MLE and RNOS bias. For the CV = 1.0 condition, the D-LOG estimator was negatively biased compared to the positive bias for the MLE and RNOS estimators; none of the estimators performed well (relative error >15%) for censoring above the distribution's 20th percentile, particularly the RNOS estimator.

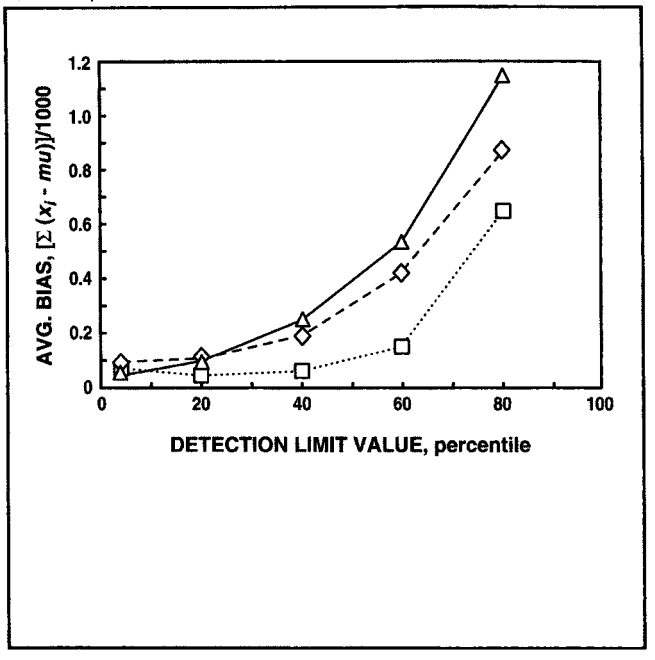
With the exception of the D-LOG SD estimator, increasing sample size has the effect of reducing the bias and RMSE of all estimators but the reductions in bias are proportionately greater for the MLE and RNOS estimators (Table III); the D-LOG SD estimator bias remained relatively constant or increased slightly with increasing sample size. Except for the slightly negative bias of the MLE estimator, the relative performance and superiority of the MLE mean estimator compared to the D-LOG and RNOS mean estimators was similar to that for sample sizes of 10. However, the bias of the RNOS mean estimator was less than that of D-LOG mean estimator when the censoring point was greater than the 20th percentile and $n \geq 15$. The relative performance of all SD estimators for $n \geq 15$ was similar to that observed for the $n = 10$ sampling size except the MLE estimator performed well for censoring up to the distribution's 60th percentile while the D-LOG and RNOS estimators performed well for censoring up to the distribution's 40th percentile .

III. Bias and RMSE for samples of 10,15 and 20 drawn from log normal distributions.

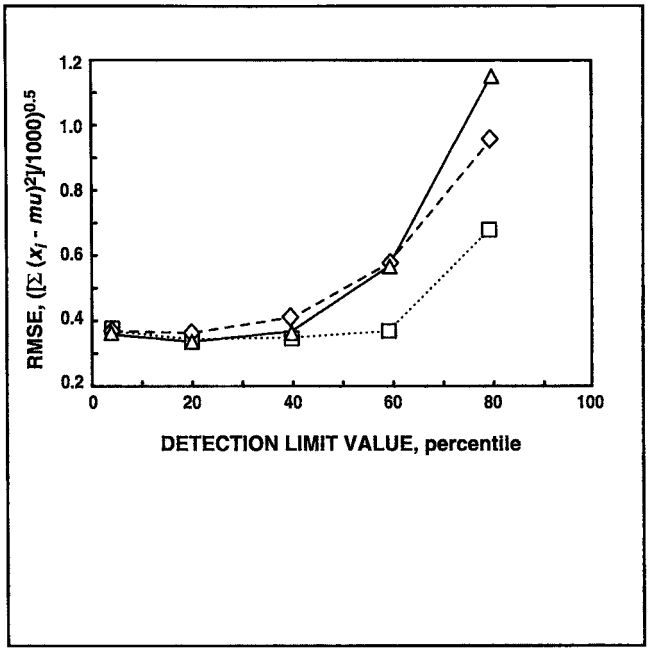
	Mean estimator						*Std. Dev. Estimator*					
	D-LOG Bias	*MLE Bias*	*RNOS Bias*	*D-LOG RMSE*	*MLE RMSE*	*RNOS RMSE*	*D-LOG Bias*	*MLE Bias*	*RNOS Bias*	*D-LOG RMSE*	*MLE RMSE*	*RNOS RMSE*
Log normal w/μ = 2, CV = 1 & *n* = 10												
5	0.7740	0.0945	0.1567	0.6665	0.6850	0.7378	-0.0191	0.2099	0.3466	1.239	1.511	1.916
20	0.0939	0.0701	0.1909	0.6435	0.6714	0.7695	-0.2270	0.1862	0.5239	1.172	1.614	2.792
40	0.2291	0.0775	0.2627	0.6591	0.6785	0.8386	-0.3406	0.3153	0.8880	1.280	2.333	5.127
60	0.5597	0.1452	0.4949	0.8231	0.6932	1.535	-0.4097	0.4991	15.13	1.765	4.318	262.0
80	1.5151	0.6260	2.1730	1.658	0.9281	14.20	-0.2287	1.416	3287	3.906	12.84	49865
Log normal w/μ = 2, CV = 0.3 & *n* - 10												
5	0.217	0.0118	0.0304	0.2090	0.2120	0.2138	-0.0397	-0.0134	-0.0030	0.1878	0.1971	0.2125
20	0.0577	0.0044	0.0385	0.2027	0.2136	0.2199	-0.0949	-0.0248	-0.0035	0.2063	0.2119	0.2424
40	0.1470	0.0030	0.0581	0.2254	0.2221	0.2317	-0.1701	-0.0292	-0.0060	0.2576	0.2360	.02853
60	0.3069	0.0347	0.1097	0.3377	0.2266	0.2910	-0.2585	-0.0478	-0.0110	0.3316	0.2734	0.3695
80	0.6469	0.2291	0.3500	0.6551	0.3047	0.4839	-0.3345	-0.0542	-0.0104	0.4029	0.3158	0.5859
Log normal w/μ = 2, CV = 1, & *n* = 15												
5	0.0350	0.0541	0.0968	0.5011	0.5183	0.5394	-0.1065	0.1115	0.1799	0.8608	1.048	1.190
20	0.0563	0.0417	0.1199	0.4851	0.5150	0.5525	-0.3033	0.1033	0.2445	0.8416	1.216	1.405
40	0.1932	0.0490	0.1641	0.5022	0.5183	0.5804	-0.4120	0.1666	0.4262	0.9407	1.521	2.249
60	0.5059	0.0680	0.2490	0.6696	0.5205	0.6540	-0.5247	0.2812	0.9775	1.116	2.340	10.82
80	1.368	0.3035	0.8675	1.426	0.5609	3.356	-0.5796	0.4816	425.2	1.777	4.127	11715
Log normal w/μ = 2, CV = 0.3, & *n* = 15												
5	0.0153	0.0062	0.0208	0.1678	0.1708	0.1714	-0.0378	-0.0104	-0.0043	0.1479	0.1547	0.1637
20	0.0527	0.0029	0.0275	0.1641	0.1701	0.1737	-0.0940	-0.0183	-0.0043	0.1719	0.1684	0.1852
40	0.1421	0.0033	0.0389	0.1968	0.1762	0.1864	-0.1686	-0.0224	-0.0050	0.2289	0.1937	0.2214
60	0.2969	0.0084	0.0645	0.3167	0.2002	0.2285	-0.2611	-0.0329	-0.0131	0.0398	0.2279	0.2664
80	0.6176	0.1104	0.2115	0.6225	0.2553	0.4007	-0.3642	-0.0497	-0.0234	0.4032	0.2590	0.3875
Log normal w/μ = 2, CV = 1 & *n* = 20												
5	0.0045	0.0230	0.0590	0.4410	0.4505	0.4668	-0.1465	0.0570	0.1179	0.7509	0.8728	0.9679
20	0.0300	0.0178	0.0790	0.4247	0.4509	0.4780	-0.3303	0.0601	0.1847	0.7539	0.9730	1.421
40	0.1700	0.0243	0.1083	0.4395	0.4529	0.4884	-0.4413	0.1026	0.2570	0.8314	1.130	1.396
60	0.4875	0.0322	0.1765	0.6202	0.4629	0.5481	-0.5401	0.2223	0.7141	1.013	1.984	5.967
80	1.318	0.1668	2.482	1.368	0.5012	57.44	-0.5958	0.6435	1.551	2.583	9.403	5E7
Log normal w/μ = 2, CV = 0.3 & *n* = 20												
5	0.0067	-.0018	0.0100	0.1494	0.1524	0.1523	-0.0384	-0.0108	-0.0046	0.1280	0.1302	0.1398
20	0.0452	-.0043	0.0140	0.1449	0.1523	0.1542	-0.0935	-0.0140	-0.0008	0.1557	0.1458	0.1593
40	0.1367	-.0035	0.0236	0.1817	0.1566	0.1632	-0.1681	-0.0165	-0.0033	0.2145	0.1654	0.1847
60	0.2930	-.0055	0.0411	0.3093	0.1895	0.2095	-0.2588	-0.0222	-0.0071	0.2968	0.1960	0.2290
80	0.6060	0.0499	0.1310	0.6099	0.2555	0.3719	-0.3742	-0.0446	-0.0118	0.4069	0.2406	0.4449

To test the robustness of the estimators under realistic analysis conditions when the underlying distribution is unknown, two empirical distributions created from actual industrial effluent discharge monitoring records were used. For simulations employing a 2,3,4,6-tetrachlorophenol (2346-TCP) data distribution, the same general performance trends as those described above for the log normal distributions were observed. The average bias of the D-LOG mean estimator was greater than the MLE and RNOS mean estimators when the censoring point was greater than the 20th percentile (**Fig.2**) while RMSE

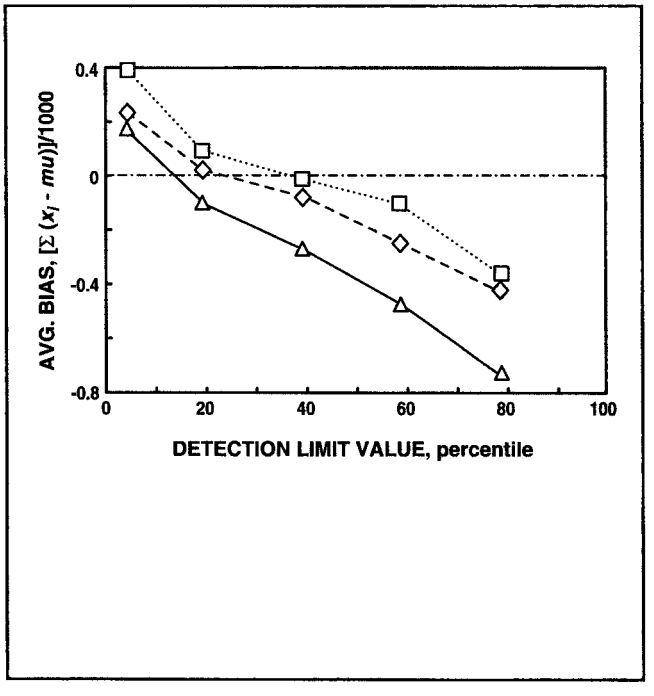
2. Bias of mean estimator for samples of 10 drawn from an empirical 2346-TCP distribution with $\mu = 2.1$ and $\sigma = 1.1$ (□, MLE; ◊, RNOS; Δ, DLOG).



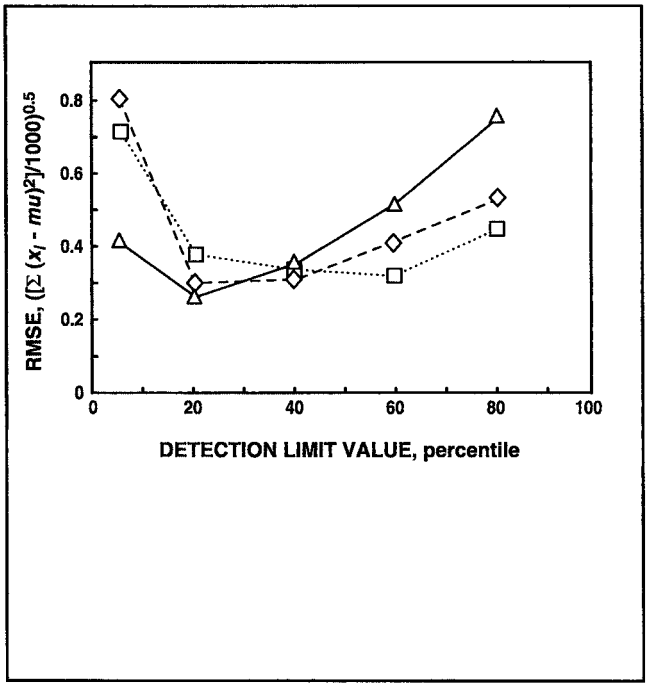
3. RMSE of mean estimator for samples of 10 drawn from an empirical 2346-TCP distribution with $\mu = 2.1$ and $\sigma = 1.1$ (□, MLE; ◊, RNOS; Δ, DLOG).



4. Bias of SD estimator for samples of 10 drawn from an empirical 2346-TCP distribution with $\mu = 2.1$ and $\sigma = 1.1$ (□, MLE; ◊, RNOS; Δ, DLOG).



5. RMSE of SD estimator for samples of 10 drawn from an empirical 2346-TCP distribution with $\mu = 2.1$ and $\sigma = 1.1$ (□, MLE; ◊, RNOS; Δ, DLOG).



for the D-LOG method was similar to the other mean estimators for censoring at or below the distribution's 40th percentile (**Fig.3**). When the SD estimator is applied to "realistic" data sets, its bias can vary from positive to negative (**Fig.4**); however, on an average absolute basis, the magnitude of the bias was greatest for the D-LOG procedure by a factor exceeding 1.5. The RMSE of the D-LOG SD estimator was favorable when the censoring point was less than the distribution's 20th percentile; however, for greater censoring, the D-LOG estimator performance degenerates compared to the MLE and RNOS estimators (**Fig. 5**). Simi-

lar relative performance among the three methods was observed for simulations employing a 4,5-dichlorocatechol data distribution with $\mu = 8.3$ and $\sigma = 5.1$; these are omitted for brevity.

Summary and conclusions

Three general approaches for estimating the mean and standard deviation of parent populations from censored data sets were describe and evaluated for bias and RMSE characteristics. These include: Cohen's maximum likelihood estimator (MLE) method, regression of normal order scores (RNOS) technique, and USEPA's D-LOG procedure. The MLE and RNOS approaches assume parent populations are normally distributed and this usually necessitates data transformation prior to analysis of typical environmental data sets. The D-LOG procedure assumes the parent population is a mixture of discrete point and log normally distributed values; is applied without compensation for the parent population's distributional characteristics.

The Monte Carlo simulation experiments support Gilliom and Helsel *(5)* assertion that although Cohen's MLE method is not superior in every case, it does a very good job on most data sets. In addition, the following specific conclusions are drawn:

1. For the sample sizes tested, all approaches appeared unreliable when the censoring threshold exceeded the distribution's 60 percentile value.
2. For estimating population means, all approaches either over estimate or are approximately unbiased. In general, Cohen's MLE method was superior to the RNOS and D-LOG approaches and its bias was at least 2.5 time less than that of the D-LOG procedure.
3. Increasing sample size from 10 to 15 or 20, decreased mean estimator bias for the MLE and RNOS approaches in contrast to the D-LOG procedure those bias increased or remained constant with increasing sample size.
4. Decreased population variability (i.e., smaller CVs), improves the bias and RMSE performance of either the mean or standard deviation estimators for all three approaches.
5. For estimating standard deviations, no approach was particularly satisfactory when population variability was large (CV = 1.0) and sample size was small (n = 10); however, with decreasing CV and increasing sample size, D-LOG bias was at least 1.2X (and frequently much) greater than MLE bias. ■

Literature cited

1. Berthouex, P. M., Estimating the Mean of Data Sets that Include Measurements Below the Limit of Detection, Technical Bulletin No. 621, National Council For Air and Stream Improvement, New York, 1991.
2. Haas, C. N. and Scheff, P. A., *Environ. Science Technol.* 24: 912(1990).
3. U.S. EPA, Development Document For Effluent Limitations Guidelines and Standards For The Organic Chemicals, Plastics and Synthetic Fibers Point Source Category, 10/87; U.S. EPA Office of Water, Washington, DC, 1987.
4. Hinton, S. W., *Environ. Science Technol.* 27: 2247(1993).
5. Cohen, A. C. Jr., *Technometrics* 1: 217(1959).
6. Gilliom, R. J. and Helsel, D. R., *Water Resources Research* 22: 135(1986).
7. U.S. EPA, Technical Support Document for Water Quality-based Toxics Control, 9/91, U.S. EPA Office of Water, Washington DC, 1991.
8. Owen, W. J. and DeRouen, T. A. *Biometrics*, 1980, 36, 707–719.

Jay P. Unwin, Joan D. Abbott, and Nancy Bartlett provided helpful discussions and comments during manuscript preparation. Erik P. Drake developed the initial version of the FORTRAN simulation code.

Received for review Dec. 22, 1993.

Accepted Jan. 6, 1994.

Presented at the TAPPI 1994 Environmental Conference.

CORRECTION

Testing Machines Inc. in Amityville, NY, promoted John L. Sullivan to president and COO, James Pryor to director of operations, and Richard Young to director of marketing and international trade. The company also named Jeffrey Bott technical director and Timothy Keefe director of sales. In error, *Tappi Journal*'s March 1994 edition (Vol.77, No. 3, p. 301), notes these promotions and appointments occuring at the company's Montreal, PQ office.