Environmental Toxicology and Chemistry—Volume 37, Number 3—pp. 643–656, 2018
Received: 10 August 2017 | Revised: 19 September 2017 | Accepted: 21 November 2017

Critical Review

Toward Improved Analysis of Concentration Data: Embracing Nondetects

Niloofar Shoari* and Jean-Sébastien Dubé

Department of Construction Engineering, École de technologie supérieure, Montreal, Québec, Canada

Abstract: Various statistical tests on concentration data serve to support decision-making regarding characterization and monitoring of contaminated media, assessing exposure to a chemical, and quantifying the associated risks. However, the routine statistical protocols cannot be directly applied because of challenges arising from nondetects or left-censored observations, which are concentration measurements below the detection limit of measuring instruments. Despite the existence of techniques based on survival analysis that can adjust for nondetects, these are seldom taken into account properly. A comprehensive review of the literature showed that managing policies regarding analysis of censored data do not always agree and that guidance from regulatory agencies may be outdated. Therefore, researchers and practitioners commonly resort to the most convenient way of tackling the censored data problem by substituting nondetects with arbitrary constants prior to data analysis, although this is generally regarded as a bias-prone approach. Hoping to improve the interpretation of concentration data, the present article aims to familiarize researchers in different disciplines with the significance of left-censored observations and provides theoretical and computational recommendations (under both frequentist and Bayesian frameworks) for adequate analysis of censored data. In particular, the present article synthesizes key findings from previous research with respect to 3 noteworthy aspects of inferential statistics: estimation of descriptive statistics, hypothesis testing, and regression analysis. *Environ Toxicol Chem* 2018;37:643–656. © 2017 SETAC

Keywords: Left-censored; Nondetect; Exposure assessment; Environmental data analysis; Limit of detection

INTRODUCTION

Even with technical advances in chemical analysis protocols and laboratory instrumentations, there is still a threshold below which concentrations of chemicals cannot be differentiated from the background noise and thus are not precisely quantifiable. These concentrations are called "nondetects" or "left-censored" and lie between zero and the detection limit of the measuring instrument. Censored data are a pervasive problem in different disciplines of environment- and health-related research including risk assessment, environmental science and engineering (e.g., concentration data), ecotoxicology, dietary exposure assessment and occupational exposure studies (e.g., biomarker data), and immunology because these observations complicate statistical analysis of data.

Several techniques based on survival analysis methods have come to the rescue when left-censored observations appear; however, these techniques are often overlooked or used improperly for a number of reasons. One reason is that guidelines from regulatory agencies are ad hoc and often conflicting. Another reason is that it is sometimes difficult to identify the most efficient statistical methods for analyzing censored data. A third obstacle is the concern among some practitioners and researchers that implementation of appropriate methods in available statistical software is challenging. Many engineers, practitioners, and statisticians are not familiar with the problem of left-censored data or are not trained to employ the previously developed statistical methods. Finally, although a few software programs exist to accommodate left-censoring, it is rare for a single package to handle all of the commonly needed options. These issues combine to obstruct the widespread use of appropriate methods and consequently result in the arbitrary choice of a statistical technique. Even worse, nondetects are discarded or, more frequently, substituted with arbitrary constants prior to data analysis, introducing bias into the estimates of key statistics [1,2].

The present article aims at increasing the awareness of the scientific community, including practitioners and policymakers, about the importance of censored observations in statistical

This article includes online-only Supplemental Data.
* Address correspondence to niloofar.shoari.1@ens.etsmtl.ca
Published online 23 November 2017 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.4046

analysis of concentration data. We conducted a comprehensive review of relevant literature to outline the merits and limitations of different statistical methods within the context of estimation of descriptive statistics, hypothesis testing, and regression analysis. Recommendations regarding the best course of action have been made accordingly. Finally, to promote incorporating censored data in the analysis, the present article summarizes a number of statistical tools as well as key functions. Instructions on the use of available software in conjunction with discussions will increase the understanding and immediacy of appropriate methods for the analysis of left-censored data.

WHAT ARE LEFT-CENSORED DATA?

A distinctive feature of concentration data is the inevitable presence of left-censored observations as a result of analytical limitations. For a given analytical method, each laboratory defines a limit of detection (LOD) and a limit of quantitation (LOQ). Adopting the definitions from the International Conference on Harmonisation [3], the LOD is the lowest concentration of an analyte in a sample that can be detected but not necessarily quantified as an exact value. In other words, the LOD is the threshold concentration below which the results cannot be distinguished from a zero reading. Similarly, the LOQ is defined as the lowest amount of analyte in a sample that can be quantitatively determined with suitable precision and accuracy.

Multiple approaches to define the LOD and LOQ exist. Generally, these concepts are defined using an analytical calibration curve, which is the scatterplot of measures of analytes versus the known concentrations for a series of standard solutions. The LOD and the LOQ are not fixed values because they depend on many parameters such as the type of assay, sample weight, extraction capability, concentration of standard solutions, and amount of solution that is injected into an instrument [4]. For more details regarding the procedures for calculating the LOD and LOQ defined by other regulatory agencies, refer to International Union of Pure and Applied Chemistry [5] and Shrivastava and Gupta [6].

Some laboratories use the LOD as the reporting threshold, whereas others use the LOQ. Since the mathematical computation does not change whether the LOD or the LOQ is utilized, the general term detection limit (DL) is used throughout the present article, and any measurement less than the DL is considered a left-censored or nondetect observation. If only a single DL appears, data are referred to as singly censored, whereas when 2 or more distinct DLs are present, data are called multiply censored. Another data scenario is when a concentration measurement lies in the interval between the LOD and the LOQ, in which case interval censoring occurs. The concepts of left and interval censoring are illustrated in Figure 1.

WHY ACCOUNT FOR LEFT-CENSORED OBSERVATIONS?

Although a left-censored observation does not report an exact value of a chemical concentration, it still contains the information that the measurement falls somewhere between zero and the DL. It is crucial to highlight that left-censored concentrations do not necessarily indicate the absence of a chemical; rather, they indicate that the precision of the analytical instrument was too low to reliably quantify a concentration value. The importance of accounting for left-censored concentrations is more pronounced when DLs are higher, such as when dealing with historical concentration data, where analytical instruments were less powerful than instruments today. In addition, in the case of highly toxic contaminants such as dioxins and arsenic, even trace levels may pose risks to human health and the environment [7,8].

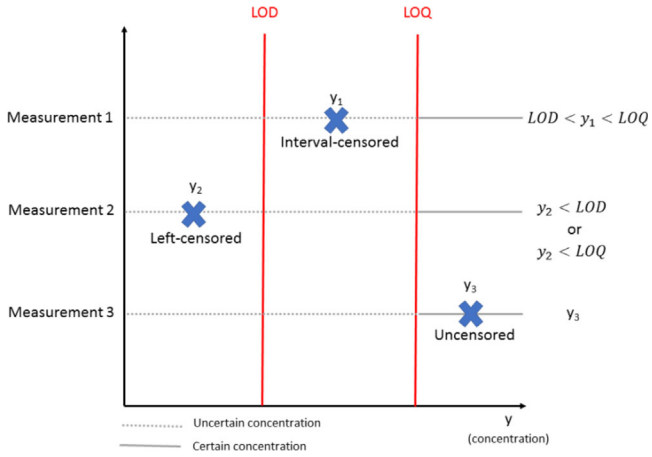

FIGURE 1: Scheme of left-censored, interval-censored, and uncensored observations. LOD = limit of detection; LOQ = limit of quantitation.

A wide range of management decisions are affected by left-censored observations because they can impact not only the estimation of statistical parameters but also the characterization of data distributions and inferential statistics (e.g., comparing the mean of 2 or more populations, the estimation of correlation coefficients, and the construction of regression models). There has been a proliferation of methodological and application reports that condemn substitution of censored values with constants while advocating the use of statistical methods tailored to account for left-censoring [1,2,9–11]. Some regulatory agencies, such as the European Food Safety Authority [4], have drawn attention to the importance of nondetects in a sound data analysis and established guidance documents regarding the best course of action to take in the presence of left-censored data, mainly in terms of amount of censoring and/or sample size.

Given the previous discussions and considering the efforts and expenses dedicated to environmental and health data collection and analysis, it seems worthwhile to investigate more sophisticated statistical methods in order to extract the maximum amount of reliable information from data containing left-censored observations.

CURRENT PRACTICE OF ANALYZING LEFT-CENSORED DATA

In current practice, nondetects are substituted with constants such as 0, DL, DL/2, and DL/$\sqrt{2}$ prior to data analysis. However,

data representativeness is deliberately diminished, and statistical results may be incorrectly interpreted because the uncertain measurements (i.e., nondetects) are treated as actually observed values. This approach in turn can potentially compromise human health and the environment and cause financial losses. Many studies based on computer-generated data have demonstrated that substitution of censored values introduces bias into estimates [1,2,10], produces significant errors for statistical group comparisons [12], and distorts regression coefficients [13]. Nevertheless, nondetects are still commonly substituted in many research articles because, as stated by Helsel [14], "there is an incredibly strong pull for doing something simple and cheap."

FREQUENTIST APPROACH FOR ANALYZING LEFT-CENSORED DATA

Instead of creating uncensored data by substituting censored values, alternative techniques that can effectively make use of censored observations should be employed. This section reviews the most common existing frequentist alternatives for the following inferential statistics: 1) estimation of statistics for describing the central tendency and spread, 2) comparison between groups, and 3) developing regression models.

Estimation of statistics for describing the central tendency and spread

Estimation techniques fall under two mainstream approaches: parametric and nonparametric. With respect to the parametric approaches, methods based on the maximum likelihood estimation (MLE) and regression on order statistics (ROS) require fitting a specific distribution to data. The customary distributional assumption is log-normal, although other possible options are gamma and, less frequently, Weibull distributions. Because of its many desirable asymptotic properties, when the model fit is good, the MLE method has been recognized as the gold standard, with best performance when the data set contains at least 30 to 50 observations [15]. More observations may be required if data are highly skewed and the censoring percentage is large [16]. In contrast to the MLE approach, where both censored and uncensored parts of data are assumed to follow the same distribution, the ROS method only requires distributional assumption for uncensored data. This method performs a regression on the scatter plot of uncensored observations versus quantiles of the assumed distribution.

Some problems associated with the preceding methods are that 1) they are sensitive to the distributional assumption, 2) they are impacted by the presence of outliers, and 3) if the lognormality assumption is used, the estimates of the mean and standard deviation (SD) suffer from back-transformation bias [2,17,18]. To solve these problems, robust MLE (rMLE) and robust regression on order statistics (rROS) have been proposed by Kroll and Stedinger [19] and Helsel and Cohn [20], respectively. In these methods, the censored observations are imputed, using some initial estimates obtained from fully parametric methods, and then combined with the observations above the DL, making it possible to use standard statistical methods. In the case that a gamma distribution is the underlying distribution of the rROS, the method is called gamma regression on order statistics (GROS) [21]. In addition to rMLE, rROS and GROS, the environmental literature reports other ad hoc imputation techniques that are less frequently used [22–24].

As a nonparametric method, the Kaplan-Meier technique is particularly useful in that it avoids making assumption about the distribution of data. This method was originally developed to treat right-censored data, where a response will occur at some point after the end of a study (e.g., in a reliability experiment, a product is still functioning and its failure is only known to be greater than the study time), and was further extended by Turnbull [25] to accommodate interval-censored data. Note that nondetects are special cases of interval-censored data because a concentration measurement is known to reside between 0 and DL. The Kaplan-Meier method estimates a cumulative distribution function that indicates the probability for an observation to be at or below a reported concentration. The estimated cumulative distribution function is a step function that jumps up at each uncensored value and is flat between uncensored values. The step size increases to the left after each left-censored value. An estimate for the mean value is obtained by calculating the area for $x > 0$ between the cumulative distribution function curve and a horizontal line at $y = 1$ (see Supplemental Data for an illustration). When the smallest observation of data is a nondetect, the cumulative distribution function remains undefined below the smallest uncensored observation (the Kaplan-Meier curve does not drop to zero), making it impossible to calculate the area delimited by the plot [26]. Statisticians have come up with different methods to correct for this issue by imputing a value for the cumulative distribution function at the undefined tail. For example, Efron's bias correction approach considers the smallest censored value to be detected, forcing the cumulative distribution function to drop to zero at the smallest censored value, and providing an upper bound for the estimated mean. The lower bound for the mean, on the other hand, can be computed by assuming that the undefined tail region takes the value of cumulative distribution function corresponding to the smallest censored observation (see Supplemental Data for a graphical illustration). If the estimated lower and upper bounds differ significantly, the uncertainty attributable to the undefined left tail of the cumulative distribution function is large, and Gillespie et al. [26] recommended presenting both lower and upper bounds. Current statistical programs do not offer an option for computing both the lower and upper bounds, but may provide one or the other. Given the upper bound, the lower bound can be obtained by subtracting the area of the uncertain rectangle, which is the smallest DL multiplied by the value of the cumulative distribution function at the smallest DL. If the lower bound is given, then the area just defined is added to give the upper bound. Users should familiarize themselves with the algorithms implemented in programs because different software packages adopt different procedures to calculate the mean based on the Kaplan-Meier estimates.

It should be noted that estimates of mean and SD are of great interest especially when the data distribution is symmetrical.

When data are highly skewed, estimating median, interquartile range, and outer percentiles is extremely useful and commonly used. The pros and cons of the parametric and nonparametric methods are summarized in Table 1.

Regulatory agencies' perspective

Regulatory agencies fall short in answering the question of how to analyze left-censored data because current norms are unclear and sometimes contradictory. Addressing the problem of left-censoring in water quality data, Appendix Q of the "Local Limits Development Guidance" [27] recognizes that substitution of censored data results in biased estimates and encourages use of the rROS technique. In another report, published by the Oak Ridge National Laboratory, it is recommended to use MLE as the first method of choice and Kaplan-Meier when the data distribution is hard to identify [28]. The US Environmental Protection Agency (USEPA) [29] advocates substitution of left-censored observations with 0, DL, or DL/2 when <15% of data are censored. However, their 15% cutoff value is simply based on judgment rather than any peer-reviewed publication [30]. When 15 to 50% of data are nondetects, trimmed and Winsorized mean and SD as well as Cohen's method are recommended. However, the latter method is a simplified version of MLE and is restricted to normally distributed data subject to a single DL. When a log-transformation is performed to obtain an approximately normal distribution, the Cohen's estimates would also suffer from back-transformation bias. In situations where >50% of data consist of nondetects, the USEPA guideline suggests using the test of proportions that estimates the frequency of uncensored observations exceeding a percentile larger than the censoring percent. If 65% of data are censored, for example, the 75th percentile can be considered.

In the field of food exposure assessment, the guideline provided by the European Food Safety Authority [4] suggests making the calculations twice, once for a lower bound by substituting nondetects with 0 and once for an upper bound by substituting nondetects with the DL. If the difference between the upper and lower bound of the estimated parameter is negligible, then substitution with the DL is recommended. When the difference is not negligible or the upper bound estimate is in the range of toxicological doses, alternative estimation techniques should be used. In that case, the choice of the optimal method depends on the sample size and censoring percent. If the sample size contains more than 50 observations, censoring percent is <50% and multiple DLs appear, the European Food Safety Authority guidelines suggest the nonparametric Kaplan-Meier or a parametric method. The guidelines also caution that if data are censored at a single DL and the smallest observation happens to be a nondetect, the estimated mean by the Kaplan-Meier method may suffer from bias attributable to corrections for the undefined tail of the cumulative distribution function (cdf) curve [31]. In the case of higher censoring percentages (50–80%) with over 50 observations, the European Food Safety Authority suggests adopting a parametric method under a distributional assumption based on the Akaike information criterion (AIC) or Bayesian information criterion (BIC). These recommendations are also referred to in the joint guidance prepared by the European Food Safety Authority, the United Nations Food and Agriculture Organization, and the World Health Organization.

Within the context of risk-based site characterization, the document "Guidance on Complex Human Health Detailed Quantitative Risk Assessment of Chemicals" provided by Health Canada [32] suggests substituting nondetects for scenarios of low censoring, although a threshold percentage for censoring is not mentioned. Furthermore, this document recommends using the rROS method for modest to large data sets and the MLE method only for large data sets without supporting their statement with a statistical basis. Table 2 summarizes 3 popular recommendations to analyze left-censored data, and it is notable that none of them include Bayesian estimation.

Making sense of conflicting findings

Despite numerous works to compensate for the lack of consistent guidelines and to identify the best course of action for the analysis of censored data, researchers are split over recognizing proper approaches. As a matter of fact, although some studies have assumed that chemical data are log-normally distributed and suggested the MLE method based on such an assumption [33,34], others have had some counter-arguments to the MLE method and advocated use of the nonparametric Kaplan-Meier technique [26,35,36]. In addition, many other researchers have favored imputation techniques such as rROS over the MLE and Kaplan-Meier methods [7,18]. An example illustrating such a discrepancy is reported as follows.

TABLE 1: Pros and cons of the maximum likelihood estimation, Kaplan-Meier, and robust regression on order statistics methods

Statistical method	Pros	Cons
Maximum likelihood	Supported by interesting asymptotic properties: consistency, efficiency, and asymptotic normality	Sensitive to distributional assumption Properties are valid for large enough sample sizes
Kaplan-Meier	Does not require data transformation Does not require distributional assumption about the shape of the data	Relies exclusively on data and their quality Requires manipulation when the smallest observation is a nondetect
Robust regression on order statistics	Robust against distributional misspecification and variations in data skewness	Predicted observations are treated as actual observations

TABLE 2: Recommended methods according to three guideline documents

	Censoring	Sample size	Recommended Methods
Helsel [40]	<50%	<50	Imputation or Kaplan-Meier/Turnbull
		>50	Imputation or Kaplan-Meier/Turnbull
	50–80%	<50	rMLE, rROS,
		>50	MLE multiple imputation
	>80%	<50	Report only percent above a meaningful threshold
		>50	May report high sample percentiles (90th, 95th)
US Environmental Protection Agency [29]	<15%	—	Substitution with 0, DL, DL/2, Cohen's method
	15–50%	—	Trimmed mean, Cohen's method, Winsorized mean, and standard deviation
	>50–90%	—	Test for proportions
European Food Safety Authority [4]	<50%	>50	Kaplan-Meier (for data sets with multiple DLs), Parametric method (for data sets with single and multiple DLs)
	>50%	>50	Parametric method (with the lowest AIC/BIC)

rMLE = robust maximum likelihood estimation; rROS = robust regression on order statistics; DL = detection limit; DL/2 = one-half detection limit; AIC/BIC = Akaike information criterion/Bayesian information criterion.

Example. A core element of exposure assessment studies is the estimation of descriptive statistics because these are used to determine exposure point concentration, a measure of average exposure of a receptor over an area in a certain period of time. Suppose we are interested in calculating the mean and SD of contaminant data obtained from the characterization of a brownfield site; for example, naphthalene data with 60 observations and 53% censoring and benzo[*a*]anthracene data with 62 observations and 27% censoring are considered. Table 3 illustrates how following different recommendations (as reported in Table 2) provides differing inferences. This table indicates that unclear and contradictory suggestions regarding optimal estimation techniques may push researchers to resort to the simplistic approach of substitution. It should be noted that the estimates of the median reported in Table 3 seem to be more robust to the choice of statistical method when compared with the estimates of the mean. Therefore, the median can be alternatively used as the measure of central tendency.

Such a discrepancy may have resulted from a number of shortcomings associated with previous studies. Skewness can significantly impact the performance of the estimation techniques. Indeed, conclusions derived for low-skewed distributions may not be valid for moderately and highly skewed ones [36]. Failure to account for this feature explains why MLE performs so well in the simulated low-skewed data but not necessarily with real chemical data. To scrutinize this observation, previous work based on an extensive simulation framework [16] showed that the MLE, rROS, GROS, and Kaplan-Meier estimators provide similar estimates of descriptive statistics if data distributions are low-skewed. With increasing distribution skewness, however, the log-normal MLE method provided inflated estimates and thus was considered unreliable. That study reported that this behavior is especially true when the censoring percentage is high and the sample size is not sufficiently large. The latter parameter (sample size) was shown to be dependent on the skewness itself and the underlying distribution of data, 2 parameters that are unknown in advance in real data analysis. The sensitivity to changes in data skewness was showed to be less significant when MLE employed a gamma distributional assumption [16,37].

TABLE 3: Estimates of the mean, standard deviation (SD), and median of 2 contaminants (naphthalene and benzo[*a*]anthracene) using different guidelines

	Naphthalene		Benzo[*a*]anthracene	
Guideline	Method used	(Mean, SD, median)	Method used	(Mean, SD, median)
Helsel [40]	Log-normal MLE	(4.02, 183.9, 0.09)	KM	(10.88, 41.09, 0.80)
			rROS	(10.87, 40.97, 0.80)
US Environmental Protection Agency [29]	Cohen's method	(NA, NA, NA)	Cohen's method	(35.02, 1604, 0.66)
	Winsorization	(0.37, 0.37, 0.1)	Winsorization	(3.27, 4.07, 0.80)
	Trimming	(0.25, 0.25, 0.1)	Trimming	(2.09, 2.69, 0.80)
European Food Safety Authority [4]	Log-normal MLE	(4.02, 183.9, 0.09)[a]	Log-normal MLE	(36.70, 2031, 0.66)
	Weibull MLE	(1.82, 8.92, 0.07)	Weibull MLE	(10.49, 39.49, 0.83)[b]
	Gamma MLE	(1.75, 4.70, 0.05)	Gamma MLE	(10.86, 24.37,1.10)

[a]The AIC values for different distributional assumptions do not significantly vary. However, as the AIC value for log-normal assumption is the smallest (185.0) among Weibull, and gamma assumptions (184.6 and 188.1, respectively), the log normality should be selected.
[b]The AIC values for different distributional assumptions do not significantly vary. However, as the AIC values for Weibull assumption is (339.8) among log-normal and gamma assumptions (342.9 and 345.8, respectively), the Weibull should be selected.
NA = not applicable because Cohen's method is not valid for data with multiple detection limits. Other abbreviations are defined in Table 2.

Another problem relates to the lack of investigations regarding the performance of parametric methods in situations where the selected distributional assumption does not represent real data. Although the Q-Q plots can be used as graphical tools to visually assess the fit of parametric models, formal model identification tests do not hold valid in the presence of nondetects, and the existing few ad hoc tests designed to check the distribution of censored data require statistical expertise. As a result, chemical data are assumed to be log-normally distributed by default, without performing tests to support this hypothesis. Thus, it is important to verify whether employing parametric methods under log-normality assumption is appropriate when the true underlying distribution of data is misspecified. A simulation study in this regard indicated that the log-normal MLE provides unreasonably large estimates and is unstable, whereas the MLE (gamma), rROS, and GROS methods are robust to distributional misspecification [16].

Finally, estimates based on left-censored data always suffer from some level of uncertainty because of random sampling error as well as the limited knowledge about the true value of nondetects (all we know is that the concentration falls somewhere between 0 and the DL). Although assessing the accuracy of different estimators has been the main direction of research, identifying estimators that minimize uncertainties is equally important. The frequentist way to express uncertainty is based on the 95% confidence interval and is typically obtained by bootstrapping. For example, using the MLE method followed by a bootstrapping technique, Zhao and Frey [38] estimated and compared the mean and its associated uncertainty for censored air toxic emission factors for log-normal, Weibull, and gamma distributional assumptions. In another study, uncertainty of the mean and SD estimates obtained from the MLE, rROS, GROS, and Kaplan-Meier methods was quantified for soil contamination data [39]. The authors showed that the MLE method using log-normal and Weibull distributional assumption provide the highest levels of uncertainty.

In consequence, the choice of the most suitable method depends on a number of factors such as sample size, distribution of data and skewness, number of uncensored observations, and more importantly, a combination of the aforementioned factors. When there is evidence that data follow a parametric distribution (for example, through the Q-Q plots), the MLE method is recommended. However, previous studies cautioned about the use of the MLE method based on log-normal distribution because it may give rise to inflated estimates, particularly when data are highly skewed and insufficient uncensored observations are available. The same criticism has been made for the Weibull distributional assumption, although the issue was not observed for MLE based on gamma distribution. Plotting the Kaplan-Meier cumulative distribution function curve is also beneficial to have an overview of the data. If the cumulative distribution function suggests a significant difference between the lower and upper bounds of the mean estimate and we cannot identify a parametric distribution to model data, estimates of the median and percentiles may be better alternatives for the measure of location.

Returning to the naphthalene and benzo[*a*]anthracene data discussed earlier, we plot the Q-Q plots for 3 potential candidate distributions: log-normal, Weibull, and gamma (Figure 2). The Q-Q plots suggest that the log-normality assumption would be acceptable because the data points fall approximately on a straight line. However, the curved right extremity of these Q-Q plots indicates that we might be dealing with a highly skewed distribution. This might justify why the estimates obtained by log-normal MLE (reported in Table 3) are significantly larger compared with other parametric distributions. Under this situation, Kaplan-Meier, rROS, and GROS provide better estimates of mean and SD, which for naphthalene are (1.80, 5.59), (1.75, 5.55), and (1.74, 5.55), respectively. For benzo[*a*]anthracene the pairs of estimates are (1.67, 6.12), (1.64, 6.09), and (1.62, 6.09), respectively.

Group comparisons

Many environmental and epidemiological studies involve hypothesis tests that compare levels of a chemical substance in 2 or more populations. The key assumption is that no trend can be attributed to concentration data. If data exhibit a temporal trend, for example, the detected difference might simply indicate the temporal variability within the populations rather than the actual between-population differences. Because of difficulties caused by concentrations below DLs, these observations are commonly substituted so that familiar statistical tests such as the parametric *t* test and the nonparametric Mann-Whitney test become applicable. However, this approach may lead to distorted inferences because it modifies the original distribution of data and affects the order of data ranking [10,40].

Suppose $x_1, x_2 \ldots x_n$ and $y_1, y_2 \ldots y_m$ represent independent samples of sizes n and m collected from groups 1 and 2, respectively. The techniques for comparison between groups 1 and 2 are classified into parametric and nonparametric procedures. Several ad hoc parametric procedures have been discussed in the literature. For example, Stoline [41] developed a parametric test procedure based on an expectation maximization algorithm for comparing the medians of 2 log-normal groups in the presence of censored data. The most commonly used parametric method, however, consists of assuming a specific distribution for data and consequently employing the maximum likelihood to estimate the slope of a censored regression model, where the slope is the difference between group means. It then conducts a hypothesis test on the significance of the slope parameter. The reliability of this method depends on the validity of the distributional assumption as well as the equal group variances condition. In practice, however, the aforementioned conditions are hardly verified, especially under small sample size (fewer than 50 observations) and high censoring percentage [12].

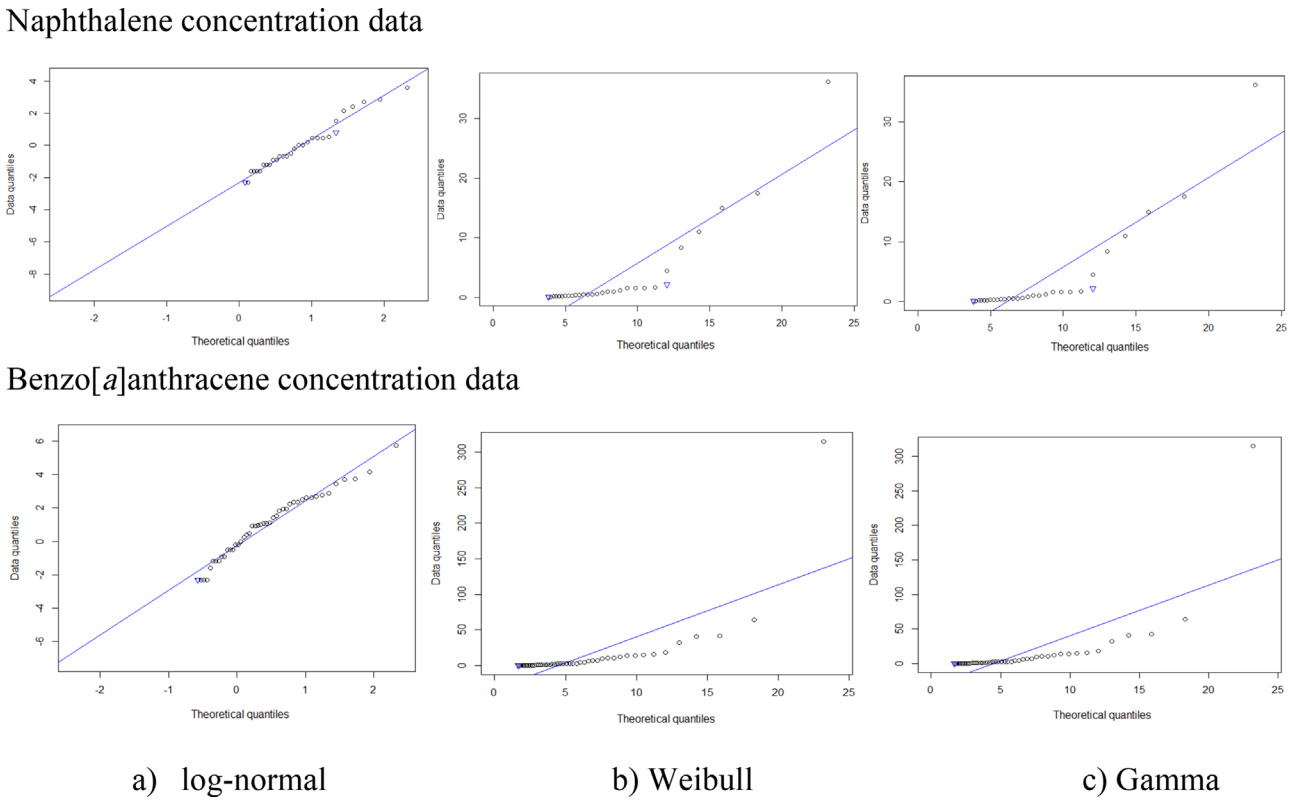

FIGURE 2: Visual model checking through Q-Q plots for naphthalene and benzo[a]anthracene.

Because most chemical data sets are small (e.g., fewer than 50 observations), highly skewed, and contain both nondetects and outliers, nonparametric comparison tests such as the Wilcoxon-Mann-Whitney, Gehan, and Peto-Prentice tests perform better than their parametric counterparts. Although these tests do not require knowledge about the precise distribution of data, they rely on the assumption that the shape of the distribution functions of 2 groups is similar. Using data ranks, this category of comparison tests identifies whether observations in a group are consistently larger (smaller) than those in another group. Generally, nonparametric tests start by pooling the observations from 2 groups into a single data set and ranking them from smallest to largest, while indicating which measurement belongs to group 1 and which belongs to group 2. Based on the rank of observations and a weight function, which differs from test to test, a test statistic is then calculated. If the calculated test statistic exceeds the critical value at a significance level $\propto$, the null hypothesis that there is no difference between the 2 groups should be rejected.

Despite the same principles in performing the Wilcoxon-Mann-Whitney, Gehan, and Peto-Prentice tests, each has advantages and disadvantages under specific circumstances. For example, the Wilcoxon-Mann-Whitney test has the limitation of being applicable to data sets subject to a single DL. If multiple DLs appear in data, the Gehan test can be alternatively used [42]. The Gehan statistic is frequently used by environmental scientists and is routinely calculated by many statistical software programs, but several studies have shown that the Peto-Prentice test is more powerful even when the censoring percentage is high and sample sizes in 2 groups are not equal [43]. The following example compares the conclusions obtained from the Peto-Prentice, Gehan, and *t* tests after censored data are substituted.

Example. At a brownfield site located in Montreal, naphthalene contamination levels in a soil layer mostly constituted of backfill material are compared with those in another layer of natural soil. Visually, the cumulative distribution function of the 2 groups displayed in Figure 3 suggests that the layer of natural soil is less contaminated than the backfill material layer. To perform a formal 2-group comparison test, we establish the null hypothesis that backfill concentrations do not differ from those of the natural soil; in other words, the cumulative distribution function of backfill data is equal to that of natural soil layer. The backfill layer data consist of 275 measurements of naphthalene, and the DLs of the measuring instrument(s) are 0.02, 0.05, 0.1, 0.2, and 0.3 mg/kg, resulting in 51% censoring. The natural soil layer has 91 measurements from which 72% are censored at DL = 0.1 mg/kg. The Peto-Prentice test based on an asymptotic variance yields a *p* value of 0.001, indicating that the null hypothesis should be rejected at the 0.05 significance level. This is strong evidence that the naphthalene concentration in the backfill layer is higher than natural soil levels. If the Gehan test is applied instead, one gets the same finding ($p = 0.0003$). In contrast, substituting nondetects with DL/2 followed by the *t* test yields a *p* value of 0.12, implying that naphthalene contamination in the natural soil layer is not significantly different from that in the backfill layer.

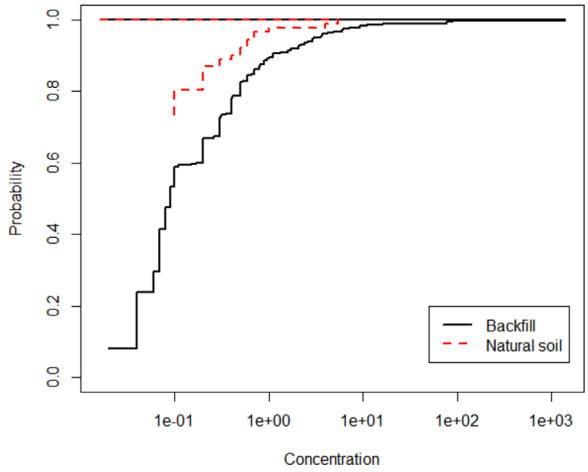

FIGURE 3: Comparison of naphthalene contamination between natural soil and backfill material layers in a brownfield site.

Regression models

Tobit regression, often called "censored" regression, has been widely used to model data with left-censored observations in environmental, medical, and epidemiological studies. Without the need to resort to substitution of nondetects, the Tobit model establishes a linear relationship between a response variable (e.g., concentration data) and predictors (e.g., population density) by estimating the parameters of the regression model using the MLE method. The residuals of the regression models (i.e., error term) are assumed to be independently and normally distributed with mean 0 and a constant variance σ^2. In the case that the normality assumption is violated, data transformation such as log-transformation of observations should be considered. Although the Tobit model was originally developed based on the normality assumption of the error term, other distributional assumptions can be alternatively used. The performance of the Tobit model is not satisfactory when assumptions related to normality and uniformity of errors are violated [44], sample size is small [45], and the censoring percentage is large (>30% according to Uh et al. [46]). Another popular parametric method to model censored data is to impute values for nondetects and combine them with uncensored data before analysis. Although simulation studies have shown that the multiple imputation method produces unbiased estimates of regression parameters [46,47], Lubin et al. [23] pointed out that this technique is not necessary when individual values for nondetects are not needed, in which case they recommended using a Tobit regression. Although there is a large literature on the modeling of left-censored data, there is limited research on dealing with left-censored covariates. Because this topic can be discussed in another article, we limit ourselves to referring the reader to a few relevant articles [48–50].

The essential assumption of most censored regression models is independence between observations, which may not necessarily be true when observations reside in groups. The dependence between observations is often induced by the data-collection process. In an epidemiological context, when measurements are repeatedly made over time within an individual, observations made for an individual are likely to be correlated. For example, for assessing the exposure of workers to air pollutants at a workplace, Peretz et al. [51] collected repeated measurements of inhalable particulates from a randomly selected number of workers. The measurements obtained from the same worker formed a group and thus were likely to be correlated. In an environmental context, concentration measures obtained over a region potentially exhibit correlation when compared with those obtained from different regions. An application example was reported by Bogner et al. [52], who collected soil samples from different areas of a forest situated approximately 50 m apart (called *plots*) and at different depths (called *horizons*) in order to investigate the significant factors affecting water flow in a forest soil. They showed that observations collected from the same plot (or horizon) were likely to be more related to each other than to the observations from different plots (or horizons).

Despite a well-documented literature on the application of censored regression models, the following challenges remain. First, applying standard censored regression models to data sets with dependent observations (similar to the previous examples) does not necessarily give consistent results because the independence assumption is not satisfied. As a matter of fact, biased estimates with erroneously narrow confidence intervals may be obtained [53], implying that a regression parameter is significant when actually it is not. Second, because of logistical and financial constraints, missing data might appear. For example, although several soil samples are collected at different depths in a site characterization study, not all of the samples undergo chemical analysis, leading to missing measurements at some depths.

A promising approach to account for data dependencies while accommodating left-censored and missing observations is mixed effects models. In addition to producing the least biased estimates of regression parameters, mixed effects models enable estimation of within- and between-group variance components, whereas simple regression models only provide a global variance. Some studies in the field of exposure assessment [4,9] and epidemiology [54,55] have benefited from mixed effects models. However, to the best of our knowledge, only a few studies in the field of environmental studies have exploited mixed effects models; some related works are Bogner et al. [52] and Shoari and Dubé [56].

BAYESIAN APPROACH FOR ANALYZING LEFT-CENSORED DATA

Bayesian data analysis is fundamentally based on a different paradigm from the previously mentioned classical statistics. The Bayesian approach uses some prior knowledge about probable values of parameters (prior distribution) along with the observed data (likelihood function) to produce the posterior distribution of the parameter of interest. Usually, a noninformative prior (e.g., uniform distribution) is assumed. In that case, the estimates

obtained by the Bayesian and MLE approaches can become comparable [57,58]. When some prior information is available based on previous related studies, informative priors are employed and more accurate and precise estimates are obtained [59]. Observations influence the posterior distribution through the likelihood function; left-censored observations contribute to the likelihood function by cumulative distribution functions, and the uncensored observations contribute by probability distribution function [60].

Some advantages of the Bayesian framework compared with its frequentist competitor are, first, that Bayesian statistics are flexible and capable of dealing with complex data (such as left-censored data) and complicated data structures (such as clustered data). Second, the inferences are exact for any sample size without relying on asymptotic approximation. Third, in the frequentist context the uncertainty is expressed in terms of a 95% confidence interval, which represents the interval containing the estimated parameter in 95% of similar repeated experiments. This occurs because the estimated parameter in frequentist statistics is a fixed value, and we typically resort to bootstrapping to derive conclusions about the uncertainty. Conversely, the Bayesian approach yields the distribution of the parameter of interest, enabling us to quantify the uncertainty with an actual probability statement. This probability statement is called the 95% credible interval, which is the interval containing the parameter with a probability of 0.95. Finally, recent advances in Bayesian methodology, in particular Markov Chain Monte Carlo simulations, and the availability of free software such as WinBUGS have made Bayesian data analysis computationally cheap.

The problem of Bayesian inference for left-censored data has been the subject of many studies. In the field of water quality assessment, Qian et al. [61] discussed the advantages of the Bayesian hierarchical modeling approach to estimate the mean concentration of chemicals occurring in the United States public drinking water system. In a later study, Bayesian methods were employed to estimate the mean and probabilities of exceeding quality standards for arsenic and fluoride concentrations in China's source waters [62]. Bayesian modeling was shown to be beneficial in investigating the spatial (at the regional scale) and seasonal variations in atrazine contamination in surface waters in the United States [63]. With respect to assessing adverse health impacts of exposure to contaminants in epidemiological studies, Bayesian analysis showed its merits in predicting exposure to one chemical in a mixture based on exposure to another chemical in the mixture even when the response and predictor were both left-censored [64]. Some other applications of Bayesian modeling with left-censored data in the context of risk assessment include those of Hayashi and Kashiwagi [65], Hickey et al. [66], and Paulo et al. [67].

DO NOT FIX YOUR DATA, FIX YOUR APPROACH

Substitution of nondetects with constants is a convenient and quick fix prior to data analysis, although it is prone to bias and uncertainty. Instead of manipulating data to make standard statistical methods work, one should employ survival analysis–based procedures that use information regarding the distribution of data, the proportion of data below the DL, and the numerical value of the DL(s) to provide a more accurate estimate of the statistics of interest. However, a major impediment to the application of more sophisticated methods relates to insufficient ability of users in many disciplines to implement and perform routine tests based on censored data. In addition, only a very few statistical packages include comprehensive procedures for left-censored data, and even these need improvements. This section promotes application of appropriate techniques for analyzing left-censored data by summarizing the most relevant statistical functions in the Minitab, R, ProUCL, SAS, JMP, and WinBUGS programs.

Minitab is a general-purpose commercial software package that not only provides many options under its reliability/survival command but also can be integrated with readily available macro. With respect to implementing the MLE for left-censored data, Minitab requires specifying the bounds of each observation in 2 columns. For example, for concentration data consisting of <1 and 2, the start point column is defined as 0 (because we cannot attribute a negative number to a concentration) and 2, whereas the endpoint column is represented by 1 and 2. This software allows user to specify a number of distributional assumptions for data and provides probability plots that can help to choose the appropriate distribution. Indeed, when data points fall approximately on a straight line on a probability plot, the assumed distribution would be a good choice when performing the parametric analysis. The Minitab procedure for performing the Kaplan-Meier method, on the other hand, requires 2 columns: one consisting of the numerical value for uncensored observations or the numerical value of the DLs (column c1) and the other representing the censoring indicator, which takes the value of 0 if censored and 1 if uncensored (column c2). Following the textbook by Helsel [40], the command %kmstats c1 c2 produces the estimates of the mean, SD, and percentiles. When computing the mean estimate from the Kaplan-Meier method, it should be noted that the commands executed in Minitab and the NADA package in R can lead to different values because different approaches to correct for the undefined tail problem are adopted. Minitab assumes that the smallest censored observation is a detected one, whereas NADA considers the area under the cumulative distribution function only up to the first uncensored observation.

Another common software that incorporates several macro commands for handling left-censored data is SAS. To obtain the MLE estimates, the PROC LIFEREG command fits a parametric distribution to data. When dealing with repeated measurements, the NLMIXED procedure, based on the MLE method, offers the possibility of including both fixed and random effects in models. More details and code examples are provided by Jin et al. [9]. The PROC ICLIFETEST procedure estimates the survival function (which is defined as 1 – the cumulative distribution function) in a nonparametric way and allows comparison of the survival functions among different populations. To execute this command, 2 variables are needed, an upper and a lower limit for each observation. The MODEL then reads as MODEL (lower, upper) = covariates. Some useful Minitab and SAS macro and their applications are reported in Table 4.

TABLE 4: Some useful Minitab and SAS commands and their applications

	Macro command	Application
Minitab	Cbox	Plots boxplots while illustrating the highest detection limit by a line
	KMSTATS	Computes descriptive statistics based on the Kaplan-Meier method
	Cros	Computes descriptive statistics using the robust regression on order statistics method based on the log-normal assumption
	BootKM	Computes the mean and bootstrap 95% confidence interval using the Kaplan-Meier method
	BootMLE	Computes the mean and bootstrap 95% confidence interval using the maximum likelihood method under log-normal assumption
	BootROS	Computes the mean and bootstrap 95% confidence interval using the regression on order statistics method assuming data are log-normal
	GW	Computes a Peto-Prentice test while plotting the boxplots and cumulative density functions of 2 groups
SAS	LIFEREG	Fits parametric cumulative distribution functions for log-normal, log-logistic, Weibull, generalized 3-parameter gamma, and non-log alternatives (e.g., normal and logistic distributions) to left- and interval-censored data Fits parametric multivariable regression models to left- and interval-censored data Produces probability plots with confidence bands to check goodness of fit of parametric distributions Estimates the reverse Kaplan-Meier (Turnbull) cumulative distribution function
	RELIABILITY	Constructs probability plots and fits parametric cumulative distribution functions with left- and interval-censored data Fits regression models with left- and interval-censored data
	NLMIXED	Fits mixed effects models using maximum likelihood
	ICLIFETEST	Estimates and plots the nonparametric cumulative distribution function with left- or interval-censored data and allows comparing cumulative distribution functions of 2 or more populations
	ICPHREG	Fits the semiparametric Cox proportional hazards regression model to left- and interval-censored data

The reliability and survival methods implemented in the JMP program can accommodate left-censoring. Similar to the data format for Minitab, each data point is represented by its lower and upper bounds in 2 columns; for a left-censored value, the lower bound is missing and the upper bound is the DL. Some important JMP platforms are reported in Table 5. Gillespie et al. [26] provide illustrations on how to implement nonparametric Kaplan-Meier in JMP, SAS, and Minitab.

The latest version of ProUCL program, ProUCL5.1 [21], is a user-friendly and freely available software that provides rigorous statistics and graphical tools to address typical issues related to site characterization including establishing background levels, comparing background and site concentration data, and determining outliers. The parametric procedures in ProUCL give users the flexibility of basing their computation on a normal, log-normal, and gamma distribution. This is an important feature because the default use of log-normal distribution for moderately to highly skewed concentration data yields unreasonably inflated estimates, especially in data scenarios with small sample sizes and a high proportion of nondetects [16,36]. Instead, such concentration data are better modeled by a gamma distribution, and the ProUCL developers have given particular attention to this aspect and incorporated statistical tools based on the gamma distribution such as the GROS method. Another unique feature of ProUCL that distinguishes it from other software is its ability to provide accurate estimates of decision statistics such as the upper confidence level that are adjusted for data skewness. Some other statistical tests of ProUCL are single-sample, 2-sample group comparison, trend, and goodness-of-fit tests. Details regarding the implementation of methods are provided in the user manual as well as a technical guide. Numerous application examples can be also found in a textbook by Ofungwu [68].

The R computing environment is popular open-source software that offers numerous statistical packages

TABLE 5: Some useful commands in JMP platform and their applications

JMP command	Applications
Analyze −> Reliability and Survival −> Survival	Computes the Kaplan-Meier survival estimates for one or more groups and provides the survival/failure plots. Shows exponential, Weibull, and log-normal diagnostic failure plots to graphically check the appropriateness of using these distributions for further modeling. Parameter estimates are available on request.
Analyze −> Reliability and Survival −> Fit Life by X	Fits linear regression models with a single covariate to left- or interval-censored outcomes; it has numerous options for distributions and transformations.
Analyze −> Reliability and Survival −> Life Distribution	Compares plots of multiple distributional fits for a given data set together with model comparison statistics such as AIC and BIC. Constructs Bayesian fits when all data are left-censored. Fits a mixture distribution to data and identifies observations belonging to a cluster. Compares different groups using a single specified distribution.
Analyze −> Reliability and Survival −> Fit Parametric Survival	Fits linear regression models to left- or interval-censored outcomes using Weibull, log-normal, exponential, Fréchet, and log-logistic distributions.

AIC and BIC abbreviations are defined in table 2

TABLE 6: List of some important R functions to analyze left-censored data

Package	Function	Application
Graphical plotting		
cg	boxplot	Create boxplots with nondetects shown as "V"[a]
cg	kmGraph	Plots cdf using the Kaplan-Meier method
EnvStats	cdfCompareCensored	For one data set, plots cdf obtained from the Kaplan-Meier method along with a theoretical cdf resulted from fitting a distribution to data. For two data sets, plots the two cumulative distribution functions based on the Kaplan-Meier method
EnvStats	qqPlotCensored	Produces Q-Q plot for left-censored data
NADA	cenboxplot	Create boxplots with the highest censoring point shown as a horizontal line
Zcompositions	splineKM	Plots cdf using the Kaplan-Meier method and a cubic smoothing spline fitted to it
Estimation of descriptive statistics		
EnvStats	elnormAltCensored	Computes mean and coefficient of variation of a lognormal distribution and estimates the confidence interval around them
EnvStats	egammaCensored	Computes shape and scale of a gamma distribution and estimates the confidence interval around them
EnvStats	enparCensored	Computes mean, standard deviation, and their confidence interval using the non-parametric Kaplan-Meier method
fitdistrplus	mledist	Computes the estimates of descriptive statistics by fitting an assumed distribution using maximum likelihood[b]
km.ci	km.ci	Computes confidence interval of the Kaplan-Meier curve
NADA	cenfit	Computes the cdf using the Kaplan-Meier method
NADA	cenmle	Computes the estimates of descriptive statistics using maximum likelihood under lognormal distribution
NADA	cenros	Computes the estimates of descriptive statistics using the robust regression on order statistics method (data are lognormal)
survival	survreg	Computes the estimates of descriptive statistics by fitting a survival regression model to datac
Group Comparison		
EnvStats	twoSampleLinearRankTest Censored	Performs different two-sample comparison test including the Peto-Prentice and Gehan tests, among others
interval	ictest	Performs different group comparison tests (e.g., Wilcoxon-Mann-Whitney, logrank) for interval censored data
NADA	cendiff	Computes the the Peto-Prentice test
Regression models		
AER	tobit	Fits a tobit model to data[c]
censReg	censReg	Performs classical as well as random effects tobit regression
flexmix	FLXMRlmmc	Fits a mixture of linear models with random effects
lmec	lmec	Fits linear mixed effects models
MCMCpack	MCMCtobit	Uses the Bayesian Markov Chain Monte Carlo (MCMC) method to estimate censored regression model
NADA	cenreg	Performs a censored regression to singly censored lognormal data by maximum likelihood

[a]Data frame should be first prepared with the *prepareCGOneFactorData*(.) function.
[b]Distributional assumptions can be normal, lognormal, exponential, poisson, cauchy, gamma, logistic, negative binomial, geometric, beta, and Weibull.
[c]Distributional assumptions can be Weibull, exponential, normal, logistic, lognormal, and loglogistic.

downloadable from the Comprehensive R Archive Network website. The *NADA* package in R is of particular relevance for scientists who are interested in analyzing left-censored data because it contains numerous parametric and nonparametric functions for routine procedures. In addition to straightforward descriptions of NADA functions in the Helsel book [40], several examples that are helpful in implementing the functions are reported elsewhere [68,69]. Table 6 lists a number of key functions contained in NADA as well as other packages that offer similar capabilities. Users should understand software capabilities before making implementation decisions. For example, the cenfit function of NADA for Kaplan-Meier estimation shifts over the probabilities associated with each observation by one value. Similarly, the upper and lower endpoints of the Kaplan-Meier curve plotted by Minitab are incorrect and need amendment. In addition, JMP falls short in providing options for nonparametric group comparisons.

WinBUGS is a versatile package designed to carry out Markov chain Monte Carlo computations for Bayesian data modeling. To model left-censoring in WinBUGS, we need to define 3 variables for each observation: the actual concentration data, $x[i]$, using *NA* when the observation is left-censored, and the lower and upper bounds of concentration values, $L[i]$ and $U[i]$, respectively. Assuming, for example, a log-normal distribution, the likelihood can be defined as

$$x[i] \sim dlnorm(mu, tau)I(L[i], U[i])$$

where the term $I(L[i], U[i])$ provides information on censoring and can be adapted to define left-, right-, and interval-censoring. When $x[i] = NA$, WinBUGS understands that there is a censored data point and $L[i]$ and $U[i]$ specify the interval including the censored observation. For example, in the case of log-normal data, these intervals can be defined as $L[i] = ln(0.001)$ and $U[i] = \ln(DL)$. When dealing with an uncensored observation, the term $I(L[i], U[i])$ is ignored. For a more detailed description of model implementation in WinBUGS, see Ntzoufras [70].

CONCLUSIONS

Although left-censored observations report only a range for a chemical concentration (i.e., between 0 and DL), the present article has discussed that those are real data and failing to acknowledge them in the analysis can obscure data patterns and distort inferences. Several statistical techniques exist for the analysis of censored data sets; however, researchers have rarely taken advantage of them and have simply resorted to substituting censored observations with arbitrary constants before data analysis. The present article has scrutinized available statistical techniques, under both frequentist and Bayesian frameworks, with respect to estimating descriptive statistics, comparing populations, and developing regression models, and offered recommendations toward a proper statistical analysis. Admittedly, employing more sophisticated techniques requires some statistical skills, but that does not justify the use of a bias-prone approach such as substituting censored values. It is indeed for this reason that we have included a summary of key functions and their implementations in popular software programs including Minitab, ProUCL, R, SAS, JMP, and WinBUGS.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.4046.

Acknowledgment—The present research was funded by the Canadian Research Chair on Contaminated Site Characterization and the Natural Sciences and Engineering Research Council Discovery Grant, awarded to J.-S. Dubé. The authors thank the anonymous reviewers for many helpful comments and suggestions.

Data availability—Data used in the present research and computation files are available upon request from the corresponding author (niloofar.shoari.1@ens.etsmtl.ca).

REFERENCES

[1] Singh A, Nocerino J. 2002. Robust estimation of mean and variance using environmental data sets with below detection limit observations. *Chemometr Intell Lab Syst* 60:69–86.

[2] Hewett P, Ganser GH. 2007. A comparison of several methods for analyzing censored data. *Ann Occup Hyg* 51:611–632.

[3] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonized tripartite guideline. Validation of analytical procedures: Text and methodology. 2005.

[4] European Food Safety Authority. 2010. Management of left-censored data in dietary exposure assessment of chemical substances. *EFSA J* 8:1–96.

[5] International Union of Pure and Applied Chemistry. 1997. Detection and quantification capabilities. In *Compendium of Analytical Nomenclature*, 3rd ed. Chapter 18, Section 18.4.3.7. Available from http://old.iupac.org/publications/analytical_compendium/

[6] Shrivastava A, Gupta VB. 2011. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles of Young Scientists* 2:21.

[7] Baccarelli A, Pfeiffer R, Consonni D, Pesatori AC, Bonzini M, Patterson DG, Bertazzi PA, Landi MT. 2005. Handling of dioxin measurement data in the presence of non-detectable values: Overview of available methods and their application in the Seveso chloracne study. *Chemosphere* 60:898–906.

[8] Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. 2012. *Heavy Metal Toxicity and the Environment. Molecular Toxicology*. Springer, Basel, Switzerland.

[9] Jin Y, Hein MJ, Deddens JA, Hines CJ. 2011. Analysis of lognormally distributed exposure data with repeated measures and values below the limit of detection using SAS. *Ann Occup Hyg* 55:97–112.

[10] Shoari N, Dubé JS, Chenouri S. 2016. On the use of the substitution method in left-censored environmental data. *Hum Ecol Risk Assess* 22:435–446.

[11] Chen H, Quandt SA, Grzywacz JG, Arcury TA. 2013. A Bayesian multiple imputation method for handling longitudinal pesticide data with values below the limit of detection. *Environmetrics* 24:132–142.

[12] Zhang D, Fan C, Zhang J, Zhang CH. 2009. Nonparametric methods for measurements below detection limit. *Stat Med* 28:700–715.

[13] Lyles RH, Fan D, Chuachoowong R. 2001. Correlation coefficient estimation involving a left censored laboratory assay variable. *Stat Med* 20:2921–2933.

[14] Helsel DR. 2010. Much ado about next to nothing: Incorporating nondetects in science. *Ann Occup Hyg* 54:257–262.

[15] Huston C, Juarez-Colunga E. 2009. Guidelines for computing summary statistics for data-sets containing non-detects. Department of Statistics and Actuarial Science, Simon Fraser University. [cited 2017

November 15]. Available from: http://bvcentre.ca/files/research_reports/08-03GuidanceDocument.pdf.

[16] Shoari N, Dubé JS, Chenouri S. 2015. Estimating the mean and standard deviation of environmental data with below detection limit observations: Considering highly skewed data and model misspecification. *Chemosphere* 138:599–608.

[17] Helsel D. 1990. Less than obvious—Statistical treatment of data below the detection limit. *Environ Sci Technol* 24:1766–1774.

[18] Huybrechts T, Thas O, Dewulfand J, Van Langenhove H. 2002. How to estimate moments and quantiles of environmental data sets with non-detected observations? A case study on volatile organic compounds in marine water samples. *J Chromatogr A* 975:123–133.

[19] Kroll CN, Stedinger JR. 1996. Estimation of moments and quantiles using censored data. *Water Resour Res* 32:1005–1012.

[20] Helsel D, Cohn TA. 1988. Estimation of descriptive statistics for multiply censored water quality data. *Water Resour Res* 24:1997–2004.

[21] Singh A, Singh AK. 2015. *ProUCL. Technical Guide-Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations*. Ver 5.1.00. US Environmental Protection Agency, Washington, DC.

[22] Lynn HS. 2001. Maximum likelihood inference for left-censored HIV RNA data. *Stat Med* 20:33–45.

[23] Lubin JH, Colt JS, Camann D, Davis S, Cerhan JR, Severson RK, Bernstein L, Hartge P. 2004. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Persp* 1:1691–1696.

[24] Ganser GH. Hewett P. 2010. An accurate substitution method for analyzing censored data. *J Occup Environ Hyg* 7:233–244.

[25] Turnbull BW. 1974. Nonparametric estimation of a survivorship function with doubly censored data. *J Am Stat Assoc* 69(345):169–173.

[26] Gillespie BW, Chen Q, Reichert H, Franzblau A, Hedgeman E, Lepkowski J, Adriaens P, Demond A, Luksemburg W, Garabrant DH. 2010. Estimating population distributions when some data are below a limit of detection by using a reverse Kaplan-Meier estimator. *Epidemiology* 21(Suppl. 4):S64–S70.

[27] US Environmental Protection Agency. 2004. Local limits development guidance appendices. EPA 833-R-04-002B. Washington, DC.

[28] Frome EL, Wambach PF. 2005. Statistical methods and software for the analysis of occupational exposure data with non-detectable values. ORNL/TM-2005/52. US Department of Energy, Springfield, VA.

[29] US Environmental Protection Agency. 2006. Data quality assessment: Statistical methods for practitioners. EPA QA/G-9S. Washington, DC.

[30] Helsel DR. 2006. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere* 65:2434–2439.

[31] Helsel DR. 2010. Summing nondetects: Incorporating low-level contaminants in risk assessment. *Integr Environ Assess Manag* 6:361–366.

[32] Health Canada. 2010. Federal contaminated site risk assessment in Canada, Part V: Guidance on complex human health detailed quantitative risk assessment for chemicals (DQRACHEM). Ottawa, ON, Canada.

[33] King GKK, Veber P, Charles S, Delignette-Muller ML. 2014. MOSAIC_SSD: A new web-tool for the species sensitivity distribution, allowing to include censored data by maximum likelihood. *Environ Toxicol Chem* 33:2133–2139.

[34] Shumway RH, Azari RS, Kayhanian M. 2002. Statistical approaches to estimating mean water quality concentrations with detection limits. *Environ Sci Technol* 36:3345–3353.

[35] She N. 1997. Analyzing censored water quality data using a non-parametric approach. *J Am Water Resour Assoc* 33:615–624.

[36] Singh A, Maichle R, Lee SE. 2006. On the computation of a 95% upper confidence limit of the unknown population mean based upon data sets with below detection limit observations. EPA/600/R-06/022. US Environmental Protection Agency, Las Vegas, NV.

[37] Schmoyeri RL, Beauchamp JJ, Brandt CC, Hoffman FO. 1996. Difficulties with the lognormal model in mean estimation and testing. *Environ Ecol Stat* 3:81–97.

[38] Zhao Y, Frey HC. 2006. Uncertainty for data with non-detects: Air toxic emissions from combustion. *Hum Ecol Risk Assess* 12: 1171–1191.

[39] Shoari N, Dubé JS. 2016. An investigation of the impact of left-censored soil contamination data on the uncertainty of descriptive statistical parameters. *Environ Toxicol Chem* 35:2623–2631.

[40] Helsel DR. 2012. *Statistics for Censored Environmental Data Using Minitab and R*. John Wiley & Sons, Hoboken, NJ, USA.

[41] Stoline MR. 1993. Comparison of two medians using a two-sample lognormal model in environmental contexts. *Environmetrics* 4:323–339.

[42] Murphy B, Morrison RD. 2007. *Introduction to Environmental Forensics*. Academic, Amsterdam, The Netherlands.

[43] Millard SP, Deverel SJ. 1988. Nonparametric statistical methods for comparing two sites based on data with multiple nondetect limits. *Water Resour Res* 24:2087–2098.

[44] Austin PC, Escobar M, Kopec JA. 2000. The use of the Tobit model for analyzing measures of health status. *Qual Life Res* 9:901–910.

[45] Fu P, Hughes J, Zeng G, Hanook S, Orem J, Mwanda OW, Remick SC. 2016. A comparative investigation of methods for longitudinal data with limits of detection through a case study. *Stat Methods Med Res* 25:153–166.

[46] Uh HW, Hartgers FC, Yazdanbakhsh M, Houwing-Duistermaat JJ. 2008. Evaluation of regression methods when immunological measurements are constrained by detection limits. *BMC Immunol* 9:59.

[47] Thompson ML, Nelson KP. 2003. Linear regression with type I interval- and left-censored response data. *Environ Ecol Stat* 10:221–230.

[48] Arunajadai S, Rauh V. 2012. Handling covariates subject to limits of detection in regression. *Environ Ecol Stat* 19:369–391.

[49] Bernhardt PW, Wang HJ, Zhang D. 2014. Flexible modeling of survival data with covariates subject to detection limits via multiple imputation. *Comput Stat Data Anal* 69. DOI: 10.1016/j.csda.2013.07.027.

[50] Atem FD, Qian J, Maye JE, Johnson KA, Betensky RA. 2017. Linear regression with a randomly censored covariate: Application to an Alzheimer's study. *J R Stat Soc Ser C Appl Stat* 66:313–328.

[51] Peretz C, Goren A, Smid T, Kromhout H. 2002. Application of mixed-effects models for exposure assessment. *Ann Occup Hyg* 46:69–77.

[52] Bogner C, Gaul D, Kolb A, Schmiedinger I, Huwe B. 2010. Investigating flow mechanisms in a forest soil by mixed-effects modelling. *Eur J Soil Sci* 61:1079–1090.

[53] Kreft I, De Leeuw J. 1998. *Introducing Multilevel Modeling*. Sage, Newbury Park, CA.

[54] Thiébaut R, Jacqmin-Gadda H. 2004. Mixed models for longitudinal left-censored repeated measures. *Comput Methods Programs Biomed* 74:255–260.

[55] Vaida F, Liu L. 2009. Fast implementation for normal mixed effects models with censored response. *J Comput Graph Stat* 18:797–817.

[56] Shoari N, Dubé JS. 2017. Application of mixed effects models for characterizing contaminated sites. *Chemosphere* 166:380–388.

[57] Busschaert P, Geeraerd AH, Uyttendaele M, Van Impe JF. 2011. Hierarchical Bayesian analysis of censored microbiological contamination data for use in risk assessment and mitigation. *Food Microbiol* 28:712–719.

[58] Moy GG, Vannoort RW. 2013. *Total Diet Studies*. Springer-Verlag, New York, NY, USA.

[59] Huynh T, Quick H, Ramachandran G, Banerjee S, Stenzel M, Sandler DP, Engel LS, Kwok RK, Blair A, Stewart PA. 2015. A comparison of the β-substitution method and a Bayesian method for analyzing left-censored data. *Ann Occup Hyg* 60:56–73.

[60] Gelman A. 2014. *Bayesian Data Analysis*, 3rd ed. CRC, Boca Raton, FL, USA.

[61] Qian S, Schulman A, Koplos J, Kotros A, Kellar P. 2004. A hierarchical modeling approach for estimating national distributions of chemicals in public drinking water systems. *Environ Sci Technol* 38:1176–1182.

[62] Wu R, Qian S, Hao F, Cheng H, Zhu D, Zhang J. 2011. Modeling contaminant concentration distributions in China's centralized source waters. *Environ Sci Technol* 45:6041–6048.

[63] Yun J, Qian SS. 2015. A hierarchical model for estimating long-term trend of atrazine concentration in the surface water of the contiguous US. *J Am Water Resour As* 51:1128–1137.

[64] Groth C, Banerjee S, Ramachandran G, Stenzel MR, Sandler DP, Blair A, Engel LS, Kwok RK, Stewart PA. 2017. Bivariate left-censored Bayesian model for predicting exposure: Preliminary analysis of worker exposure

during the Deepwater Horizon oil spill. *Ann Work Expo Health* 61:76–86.

[65] Hayashi TI, Kashiwagi N. 2011. A Bayesian approach to probabilistic ecological risk assessment: Risk comparison of nine toxic substances in Tokyo surface waters. *Environ Sci Pollut Res* 18: 365–375.

[66] Hickey GL, Kefford BJ, Dunlop JE, Craig PS. 2008. Making species salinity sensitivity distributions reflective of naturally occurring communities: Using rapid testing and Bayesian statistics. *Environ Toxicol Chem* 27:2403–2411.

[67] Paulo MJ, van der Voet H, Jansen MJW, ter Braak CFJ, van Klaveren JD. 2005. Risk assessment of dietary exposure to pesticides using a Bayesian method. *Pest Manag Sci* 61:759–766.

[68] Ofungwu J. 2014. *Statistical Applications for Environmental Analysis and Risk Assessment*. John Wiley & Sons, Hoboken, NJ, USA.

[69] Feigelson ED, Babu GJ. 2012. *Modern Statistical Methods for Astronomy: With R Applications*. Cambridge University Press, Cambridge, UK.

[70] Ntzoufras I. 2009. *Bayesian Modeling Using Winbugs*. John Wiley & Sons, Hoboken, NJ, USA.

