Integrated Environmental Assessment and Management — Volume 6, Number 3—pp. 361–366

Summing Nondetects: Incorporating Low-Level Contaminants in Risk Assessment

*Dennis R Helsel**

Practical Stats, 9278 Lark Sparrow Dr., Highlands Ranch, Colorado 80126, USA

(Submitted 7 September 2009; Returned for Revision 26 October 2009; Accepted 20 November 2009)

ABSTRACT

Low-level contaminants often are present below the detection or reporting limits of a laboratory, resulting in values reported as a nondetect or less-than. How can these values be summed along with detected concentrations to obtain a total, particularly when weighting factors such as toxic equivalence factors (TEFs) are used? The most common method employed by environmental scientists for summing nondetects along with detected values is to substitute one-half the detection limit for each nondetect. This substitution allows the least precise measurements, data with high detection limits, to have a strong influence on the resulting total amount. Substitution methods have repeatedly been shown to provide substandard results in studies over the last 2 decades. Here an alternative, the Kaplan–Meier (KM) method used throughout the fields of medical and industrial statistics, is used to obtain the total. KM estimates are far less affected by the least precise data than are estimates computed using substitution. No assumptions about the distribution of data (whether they follow a normal or other distribution) need be made. Direct application of KM to computation of toxicity equivalence concentrations (TECs) is shown. Integr Environ Assess Manag 2010;6:361–366. © 2009 SETAC

Keywords: Nondetects Detection limit Toxicity Dioxin TEC

Original Research

INTRODUCTION

Low-level contaminants often are present below the detection or reporting limits of a laboratory, resulting in values reported to the user as a nondetect or less-than. Statisticians call data sets with nondetects ''censored data'', because they include values known only to exceed or to be lower than some threshold. Although methods for dealing with censored data are routine in the fields of medical and industrial statistics, they have only recently been applied to the environmental sciences (Helsel 2005a). In this paper, a method commonly used to compute means for censored data, the Kaplan–Meier (KM) procedure, is used to compute a sum of values for data that include nondetects.

One of the simpler needs in data analysis is to sum a series of numbers. This may be done to estimate the yearly total mass of a contaminant entering a water body. Twelve monthly measured values are summed to produce the total. A second, more complicated application is in performing ecological risk assessments (USEPA 1998). A simple, numerical measure of the effects on organisms from exposure to a suite of congeners of polychlorinated biphenhyls (PCBs), dioxins, and furans is needed. Toxicity equivalent concentrations (TECs) are typically calculated to estimate the general toxicity of a sample to classes of organisms (birds, fish, mammals, humans) by assuming that toxicities of individual chemical congeners are additive (USEPA 2001). TECs are a critical component in issuing fish consumption advisories to protect human health, for example, so their computation may have significant environmental and economic consequences.

* To whom correspondence may be addressed: dhelsel@practicalstats.com
Published on the Web 12/29/2009.
DOI: 10.1002/IEAM.31

Chemical congeners have differing toxicities to organisms, so each dioxin or furan congener is ''normalized'' to the toxicity level of the most toxic congener, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, or TCDD, using a toxic equivalent weighting factor (TEF) of relative toxicity (USEPA 2001). TEFs were developed by consensus of panels of scientists for each class of organism (Van den Berg et al. 1998). TCDD has a TEF of 1, whereas less toxic congeners have TEFs closer to 0. Measured concentrations are multiplied by the TEF to obtain the TEC for that congener, the contribution of that congener to the total TCDD-equivalent toxicity for the sample. The total TEC is the sum of the individual congener TEC values in the sample. At times, congener concentrations are below their detection limits, and the issue at hand is how to use these nondetects in the summing process when computing a total TEC.

The method most commonly used today by environmental scientists for summing data that include nondetects is to assign one-half the detection limit to each nondetect. The limitations of substitution when estimating a mean for censored data have been shown in numerous simulation studies (Gleit 1985; Helsel and Cohn 1988; Singh and Nocerino 2002). Kaplan–Meier was found to better estimate the mean of censored data than substitution in a recent study (Antweiler and Taylor 2008), in which it was pointed out that substituting one-half the detection limit is most problematic when laboratory detection limits are multiplied by a constant before use as reporting limits. When values measured below a detection limit are reported as less than a higher quantitation limit, while data between detection and quantitation limits are reported as single numbers (even though qualified), insider censoring (Helsel 2005b) can result, a bias elevating estimates of calculated mean and percentiles. Substitution can result in significant errors when subsequent hypothesis testing is performed (Helsel 2006). A much

more exhaustive list of references evaluating methods for computing descriptive statistics and hypothesis tests with censored data is given elsewhere (Helsel 2005; Antweiler and Taylor 2008).

Current USEPA draft guidance for computing toxicity equivalents is silent on how to incorporate nondetect data, other than the fact that 0 and the detection limit can be substituted and the range of possible TEC values reported (USEPA 2008). When the range of possible values is wide, this method is not very helpful, leading to the common but simplistic substitution of one-half of each detection limit to obtain a single total TEC. An example of using substitution when computing a TEC is shown in Table 1. Data are one of several example sediment conditions provided for a 2008 workshop on statistical treatment of sediments in northwest estuaries. No specific location was provided for the data. One-half the detection limit, or 0.3, is multiplied by the TEF of 0.01, and the resulting toxic equivalent concentration of 0.03 is summed along with the TECs for other congeners to produce the total TEC for this sample. There are at least 2 problems with substitution of one-half the detection limit before computing a sum such as this. The first problem is that the least precise measurements, data with high detection limits, will often have a strong influence on the resulting total TEC. For example, suppose a less precise method had been used for analysis of 1,2,3,7,8-PeCDD and instead of a measured 0.18 the laboratory had reported a value of <1. One-half of this or 0.5 would have been used to compute the TEC for this (toxic) congener, and the total TEC would have increased by 0.41, or by 12%, over the current TEC. This increase is caused only by falsely translating a loss in precision (higher detection limit) into a higher concentration by using substitution.

A second failure of substitution is that invasive data, data alien to the measured values actually present in the sample, are produced. These invasive values representing the pattern of one-half of detection limit determinations can take over and choke out any patterns in the original measured data. For example, if correlations were computed between congener concentrations to look for relationships among the compounds, false correlations could be produced between two congeners that both had high detection limits and, therefore, high substituted numbers. The actual congener concentrations in the samples might not have shown any correlation, but the pattern of high laboratory detection limits produced consistently high substituted values. That consistency resulted in a false correlation. Examples of not finding correlations that were actually present between congeners, because of the substitution of fabricated values for one congener, can also easily be imagined.

Table 1. Toxicity equivalence concentration calculations using substitution of one-half the detection limit and Kaplan–Meier (KM)

Compound	Concentration	One-half DL	Toxic equivalence factors	Toxicity equivalence concentrations ½ DL	Toxicity equivalence concentrations KM
1,2,3,4,6,7,8-HpCDD	25	—	0.01	0.25	0.25
1,2,3,4,6,7,8-HpCDF	1.8	—	0.01	0.018	0.018
1,2,3,4,7,8,9-HpCDF	<0.56	0.28	0.01	0.003	<0.006
1,2,3,4,7,8-HxCDD	0.26	—	0.1	0.026	0.026
1,2,3,4,7,8-HxCDF	<0.6	0.3	0.1	0.03	<0.06
1,2,3,6,7,8-HxCDD	2.1	—	0.1	0.21	0.021
1,2,3,6,7,8-HxCDF	0.33	—	0.1	0.033	0.033
1,2,3,7,8,9-HxCDD	0.77	—	0.1	0.077	0.077
1,2,3,7,8,9-HxCDF	0.37	—	0.1	0.037	0.037
1,2,3,7,8-PeCDD	0.18	—	1	0.18	0.18
1,2,3,7,8-PeCDF	0.24	—	0.03	0.007	0.007
2,3,4,6,7,8-HxCDF	<0.14	0.07	0.1	0.007	<0.014
2,3,4,7,8-PeCDF	<0.8 *<5.0**	0.4 *2.5**	0.3	0.12 *0.75**	<0.24 *<1.5**
2,3,7,8-TCDD	1.7	—	1	1.7	1.7
2,3,7,8-TCDF	5.1	—	0.1	0.51	0.51
OCDD	220	—	0.0003	0.066	0.066
OCDF	44	—	0.0003	0.013	0.013
Sum	—	—	—	3.29 *3.92**	3.21 *3.26**

*Result of increasing the detection limit for one congener from 0.8 to 5.

AN ALTERNATE METHOD FOR SUMMING DATA WITH NONDETECTS

The sample mean is usually computed by summing values in a data set and dividing the total by the number of observations, n (Eqn. 1). The sum and the sample mean are the same phenomenon; the mean is a sum standardized by the number of values summed. Reversing the equation, the sum equals the mean multiplied by n. For data with nondetects, the mean can be estimated using a reliable method that does not involve substitution, and the total is then computed by multiplying by the mean by n. The reliable method used here for censored data is the Kaplan–Meier (KM) procedure.

$$mean_x = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1)$$

KM is the standard procedure in medical and industrial statistics (survival analysis) for estimating a mean of censored data (Klein and Moeschberger 2003). For those applications the censoring is on the upper side (right-censored data, or greater-thans). It was recommended for use in environmental studies by Helsel (2005a) for left-censored less-than values and was found to be the most reliable method for computing the 95% upper confidence limit on the mean (UCL95) of concentration data in a large simulation study (Singh 2006). It is a nonparametric method and therefore does not involve transformations to normality or assumptions of any specific distributional shape. It involves only counting numbers of data above, at, and below each detected observation and so is easy enough to implement in MS Excel. An Excel spreadsheet for computing the KM estimate of the mean and other descriptive statistics is freely available for download at http://www.practicalstats.com/nada/ . It should be noted that, in the Singh (2006) study, the one-half substitution method for computing the UCL95 was cited as working poorly for percentages of nondetects as low as 10%. Because the mean and the sum are the same phenomenon, substitution of one-half the detection limit is also not expected to produce reliable estimates of a total sum. Based on the Singh (2006) study, KM is expected to provide a better estimate of the sum and of a confidence interval for the sum than does substitution. In particular, the occasional presence of high detection limits warrants use of a procedure that does not translate poor precision (a high detection limit) into a higher sum, as does substitution.

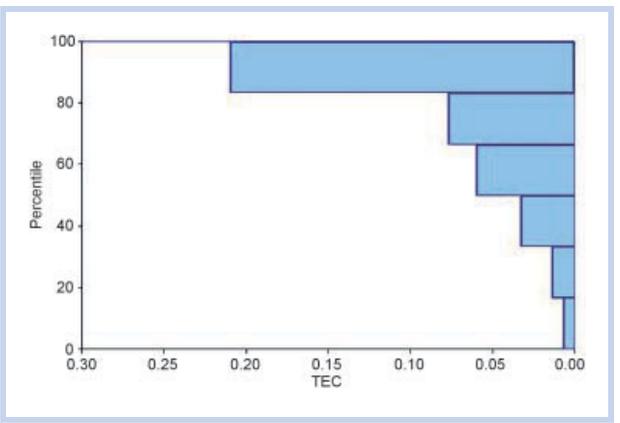

Figure 1. Kaplan–Meier method for estimating the mean without nondetects (see Table 2). The mean equals the total area inside the bars.

How KM is computed

Kaplan–Meier estimates a sum in the process of estimating the mean. It computes the area under the cumulative distribution function (cdf) of a set of data. Figure 1 illustrates the process for an example data set of 6 congener TEC values (0.21, 0.077, 0.06, 0.033, 0.014, 0.0072), ignoring for the moment that 2 of these values are nondetects. The data are also given in Table 2. Each observation is assigned percentiles $1/n$ apart from each other, and so they have percentiles at 5/6, 4/6, 3/6, 2/6, 1/6, and 0 to form the cdf. Looking at the colored rectangles in Figure 1 that make up this area, the height of each rectangle is $1/n$, or 0.16667 for $n = 6$. The area of each rectangle is 1/6 times the data value, so the area under the cdf curve equals the mean, 0.067. The data are plotted from right to left, seemingly backward from typical plots, because commercial survival analysis software assumes censored values are right-censored greater-thans instead of left-censored nondetects, so the plots come out backwards (Helsel 2005a). If the height of each rectangle were set to be 1 rather than $1/n$, the area equals the sum of the six numbers, 0.4012, and the histogram is a picture of the sum itself. The computation process using the percentile values of 0 to 5/6 illustrates that the mean is simply a scaled version of the sum.

Table 2. Percentiles for 6 observations with and without censoring as computed by Kaplan–Meier

Concentration	Toxic equivalence factors	Toxicity equivalence concentrations	Percentile ignoring the less than symbol (Figure 1)	Percentile accounting for nondetects (Figure 2)
2.10	0.10	0.2100	0.833333	0.833333
0.77	0.10	0.0770	0.666667	0.666667
<0.60	0.10	0.0600	0.500000	—
0.33	0.10	0.0330	0.333333	0.444444
<0.14	0.10	0.0140	0.166667	—
0.24	0.03	0.0072	0.000000	0.000000

These data are illustrated in Figures 1 and 2.

Kaplan–Meier computes percentiles by determining how many observations, detects, and nondetects are above, at, and below each detected observation. The process starts at high values and goes down the data set. For the Figure 1 data, in which all values are assumed to be detects, there are 6 observations equal to and below the highest detect, 1 observation at that value, and 5 below. The proportion of data below this detected observation is 5/6. The percentile for the highest detect is 5/6, or 0.83. For the second highest detect, there are 5 values at or below 0.077, with 4 below it. The proportion of data below this detected observation is 4/5 times the proportion below the previous, higher detected observation. Therefore, its percentile is 4/5 times the previous percentile of 0.83, or 0.667. Similarly, the third-highest observation (considering all 6 to be detects) has 4 observations at and below, with 3 below. Its percentile is 3/4 the previous percentile of 0.667 and so equals 0.5. The fourth-highest observation has 3 observations at and below, with 2 below. Its percentile is then $2/3 \cdot 0.5$, or 0.333. The fifth-highest observation has 2 values at and below, with 1 below. Its percentile equals $1/2 \cdot 0.333$, or 0.167. The lowest observation is assigned a percentile of 0. Without nondetects, each observation has a percentile of $(i - 1)/n$, where i is the rank of the observation from lowest to highest. Plotting positions other than $(i - 1)/n$ may be used (Stedinger et al. 1993), but this one is adopted here. The choice of plotting position will not affect the value computed for either the mean or the sum.

Now recognize that two observations in the list (0.21, 0.077, *0.06, 0.033, *0.014, 0.0072) are actually nondetects. KM computes percentiles only for detected observations, but the number and position of nondetects influences the percentile calculated for detected observations. For the highest observation of 0.21, there are still 6 observations at and below it, with 5 below, so its percentile is 5/6, just as it was when the nondetect designation was ignored (Table 2). This is appropriate, insofar as it is clear that the 2 nondetects at <0.06 and <0.014 are both below a detected 0.21. The second-highest detected observation is also as before, and so has the same percentile at 0.667. The third highest value is a nondetect at less than 0.06. Its position relative to all values below 0.06 cannot be known, so a percentile is not calculated for it. However, its influence shows in the calculation for the next lower value, a detected 0.033. This observation has 3 values that are known to be at or below it, with 2 observations (1 detect and 1 nondetect) below it. Its percentile is therefore calculated as 2/3 the previous percentile of 0.667, or 0.444. This is higher than the percentile assigned to the same observation when the 2 nondetects were treated as detected values, because there is some chance that the <0.06 lies below this detected 0.033. The lowest detected observation lies at a percentile of 0, as before. The calculated percentiles for both situations are shown in Table 2. Figure 2 shows the resulting histogram and area (i.e., KM mean) when the nondetects are recognized. With these KM percentiles, the rectangles in Figure 2 have unequal heights corresponding to the unequal difference in percentile values between the detected observations. These unequal differences reflect the information available in the nondetects, their relative positions in regard to some detected observations, even though they do not have a known single value as do detected observations. The KM mean for the 4 detects and 2 nondetects is 0.058, which when multiplied by $n = 6$ observations results in a sum of 0.35 for this data set. This KM estimate of the sum of the congener TECs lies between the estimates that result when 0 (0.320) and the detection limit (0.394) are substituted for all nondetects. It was obtained without substitution of any values for the nondetects and without assuming that the 6 observations follow any specific distributional shape.

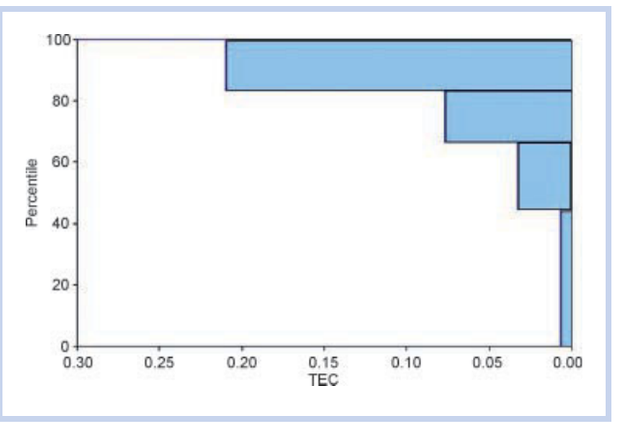

Figure 2. Kaplan–Meier method for estimating the mean with 2 nondetects (see Table 2). The mean again equals the total area inside the bars.

Considerations when the highest or lowest values are nondetects

Kaplan–Meier is a nonparametric or distribution-free procedure and as such will not use an external model to estimate how far below the lowest detection limit a nondetect value might lie. When the lowest value in a data set is detected, and the lowest nondetect is above that, KM performs as advertised. This may often be the case with TEC computations, but not when only concentrations are considered, without weighting factors. Efron's bias correction (Klein and Moeschberger 2003) is a commonly used adjustment to KM that always considers the lowest value to be detected. In this way, the mean and sum are computed using a value at the lowest detection limit for the lowest value. In applications to chemical data such as TECs, this is the highest that the value might be, and KM will not guess at a lower value. If the correction is not used, the lowest nondetect will be set to a value equal to the lowest detected value, above the lowest detection limit. This assignment is not applicable to left-censored chemical data, though it may sometimes be for the right-censored data of survival analysis. For left-censored chemical data including TEC applications, Efron's correction should therefore always be used.

Note that, if the highest value in a data set is a nondetect, as if there had been a seventh point at <0.5, the high nondetect does not enter into the calculations at all. There is no way to determine for any of the detected observations whether this high nondetect is above or below the detected value. Therefore, the high nondetect has zero information content in this situation. When all detections are below 0.5, a <0.5 has the same information as a <50 or <500; the limits are simply above the highest detected value. Less precise values (high individual TECs) have less influence on the outcome of the KM estimate, instead of having a large effect as when substituting one-half the detection limit. This property of the

KM estimate is far more appropriate from an information theory, and common sense, point of view than is substitution. However, if the congener that produced this high nondetect is of great concern, such as TCDD or any of the highest-toxicity congeners with TEFs close to 1, ignoring its contribution may not be acceptable. In this situation, the sample must be reanalyzed using a lower detection limit before reliable estimates of the total TEC can be made using any calculation method.

An example use

Table 1 presents concentrations, TEFs, and congener TECs for 17 dioxin and furan congeners. Four of the 17 concentrations are nondetects. With the KM procedure, the detection limit values for nondetects are multiplied by the congener's TEF so that the congener TEC for 1,2,3,4,7,8-HxCDF in the right-hand column is expressed as <0.06 rather than as a detected 0.03 resulting from substitution of one-half the detection limit. Because the KM procedure takes nondetects at face value, the 4 nondetect TECs are combined with the 13 detected values to produce an estimated mean of 0.189. Multiplying by $n = 17$, the KM sum equals 3.21, obtained without substituting any values for nondetects.

The potential effect of a high nondetect on both procedures is seen in Table 1, in which the detection limit for 1 congener has been altered and shown in italics. Suppose the 2,3,4,7,8-PeCDF concentration is reported as <5 rather than as <0.8. The higher detection limit could result from a number of causes: use of a different laboratory, interference from a higher total organic carbon content, or because a different protocol was used within the laboratory, among other causes. Substitution treats this value as if it were a detected 2.5, increasing the total TEC by 19%, from 3.29 to 3.92. In contrast, with KM, the total TEC increased by only 1.3%. This small change results from considering the TEC for this congener to be <1.5 instead of <0.24, shifting the percentiles computed for detected values below the new, higher detection limit of 1.5.

WHEN NOT TO USE KM FOR SUMMING DATA WITH NONDETECTS

Although substitution does not provide a satisfactory alternative to KM for summing data with nondetects, other methods are available in specific cases. If strong correlations exist between congeners in a series of samples so that the concentrations of one congener can be reliably predicted from others, the correlation can be used to predict values for concentrations measured as below the detection limit. This procedure has been used when insufficient amounts of sample prevented concentrations from being measured for some congeners (Cook et al. 2003). The resulting sum will be more accurate than with the use of KM if the estimated concentrations are close to the unknown, true concentrations for the congeners. Criteria for how strong a correlation should be to produce estimates with this method versus KM are not known.

Two other situations exist in which KM should not be used. The first is when all less-than values have only 1 threshold. In this situation, the KM estimate of the mean will equal that of substituting the threshold value for nondetects (Helsel 2005a). This is unlikely in the case of computing total TECs, because the thresholds for different congeners are computed by multiplying the reporting limit by the TEF weighting factor, which differs for different congeners. However, it may occur more commonly without weighting, when the reporting limit itself is the threshold. Studies of chemical concentrations using 1 laboratory over a short period of time return data that may have only 1 detection limit. KM is a nonparametric method and as such will not use an external model other than the data itself to estimate how far below the lowest detection limit a nondetect value might lie. Although this is not generally a problem for data with multiple detection limits as usually found in TEC calculations, KM will give the highest possible value for the mean, and sum, when used for computing sums of data having only one censoring threshold.

The second situation is when a high nondetect value, higher than all TECs from detected concentrations, occurs for one of the highest-toxicity congeners with TEFs close to 1. In this situation, no calculation procedure can give a reliable estimate of the total TEC. KM and any other statistics-based (as opposed to substitution) procedure will ignore this high nondetect, because it has no information content. A lower detection limit must be implemented before reliable estimates of the total TEC can be made using any calculation method. All that can be done in this situation is to substitute the detection limit in order to provide a worst-case value for the total TEC, realizing that the true total may be far lower.

CONCLUSIONS

Kaplan-Meier is the standard procedure in survival analysis for computing the mean of right-censored data. It has been shown in simulation studies to be one of the best methods for computing the mean and confidence intervals for left-censored environmental concentration data. The mean and the sum are the same phenomenon, simply presented on different scales. Methods such as KM that reliably estimate a mean can also reliably estimate a sum. The KM method should be more widely used to compute the sum of censored data in risk analysis and other environmental applications. It is easy to compute with available software or by hand. It is the nonparametric maximum likelihood procedure and so theoretically is optimal for computing a mean without assuming any specific distribution (Klein and Moeschberger 2003). It avoids the pitfalls of substitution and in most situations provides a practical solution to a vexing problem.

Acknowledgment—This work was partially funded through Avocet Consulting for presentation at a workshop sponsored by the USEPA, Region 10 Superfund Program. The author greatly appreciates the comments by the two anonymous reviewers, which improved the manuscript.

REFERENCES

Antweiler R, Taylor H. 2008. Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: I. Summary statistics. *Environ Sci Technol* 42:1201–1206.

Cook P, Robbins J, Endicott D, Lodge K, Guiney P, Walker M, Zabel E, Peterson R. 2003. Effects of aryl hydrocarbon receptor-mediated early life stage toxicity on lake trout populations in Lake Ontario during the 20th Century. *Environ Sci Technol* 37:3864–3877.

Gleit A. 1985. Estimation for small normal datasets with detection limits. *Environ Sci Technol* 19:1201–1206.

Helsel DR. 2005a. Nondetects and data analysis: Statistics for censored environmental data. Hoboken (NJ). John Wiley & Sons, 250 p.

Helsel D. 2005b. Insider censoring: Distortion of data with nondetects. *Hum Ecol Risk Assess* 11:1127–1137.

Helsel D. 2006. Fabricating Data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere* 65:2434–2439.

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

Klein JP, Moeschberger ML. 2003. Survival analysis: Techniques for censored and truncated data, 2nd ed. New York (NY). Springer, 536 p.

Singh A. 2006. On the computation of a 95% upper confidence limit of the unknown population mean based upon data sets with below detection limit observations. Washington DC: USEPA. EPA/600/R-06/022.

Singh A, Nocerino J. 2002. Robust estimation of mean and variance using environmental datasets with below detection limit observations. *Chemom Intell Lab Syst* 60:69–86.

Stedinger J, Vogel R, Foufoula-Georgiou E. 1993. Frequency of extreme events. In: Maidment D, editor. Handbook of hydrology. New York (NY): McGraw-Hill. Chapter 18.

[USEPA] US Environmental Protection Agency. 1998. Guidelines for ecological risk assessment: Washington DC: USEPA. EPA/630/R-95/002F.

[USEPA] US Environmental Protection Agency. 2001. Workshop report on the application of 2,3,7,8-TCDD toxicity equivalence factors to fish and wildlife. Washington DC: USEPA. EPA/630/R-01/002

[USEPA] US Environmental Protection Agency. 2008. Framework for application of the toxicity equivalence methodology for polychlorinated dioxins, furans, and biphenyls in ecological risk assessment. Washington DC: USEPA. EPA 10/R-08/004.

Van den Berg M, Birnbaum L, Bosveld BTC, Brunstrom B, Cook PM, Feeley M, Giesy JP, Hanberg A, Hasegawa R, Kennedy SW, Kubiak T, Larsen JC, van Leewen FXR, Djien Liem AK, Nolt C, Peterson RE, Poellinger RE, Safe S, Schrenk D, Tillitt D, Tysklind M, Waern F, Zacherewski T. 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environ Health Perspect* 106:775–792.