WATER RESEARCH 45 (2011) 3107–3118

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Application of robust statistical methods to background tracer data characterized by outliers and left-censored data

*Malcolm S. Field**

U.S. Environmental Protection Agency, National Center for Environmental Assessment (8623P), 1200 Pennsylvania, Ave., N.W. Washington, D.C. 20460, USA

ARTICLE INFO

Article history:
Received 10 November 2010
Received in revised form
26 January 2011
Accepted 13 March 2011
Available online 21 March 2011

Keywords:
Tracer-breakthrough curves
Outliers
Left-censored data
Mean
Maximum likelihood estimation
Median

ABSTRACT

Accurate analysis of tracer-breakthrough curves is dependent on the removal of measured background concentrations from the measured tracer recovery data. Background concentrations are commonly converted to a single mean background concentration that is subtracted from tracer recovery data. To obtain an improved estimate for the mean background concentration, a statically-robust procedure addressing left-censored data and possible outliers in background concentration data is presented. A maximum likelihood estimate and other robust methods coupled with outlier removal are applied. Application of statically-robust procedures to background concentrations results not only in better estimates for mean background concentration but also results in more accurate quantitative analyses of tracer-breakthrough curves when the mean background concentration is subtracted.

Published by Elsevier Ltd.

1. Introduction

Hydrologic tracer studies are influenced by background tracer concentrations that must be properly accounted for in every tracer test. The occurrence of background concentrations can be significant and results from a variety of naturally occurring and man-made substances. Even low background concentrations of the tracer can have profound effects on the detection and analysis of the collected samples and any subsequent numerical analyses of the tracer-test results. Unfortunately, although it is generally well recognized that background concentration measurements are important, basic tracer-test methods textbooks (e.g., Leibundgut et al., 2009; Käß, 1998; Gaspar, 1987a,b; Davis et al., 1985) only briefly mention background concentrations. Other technical reports also mention background-concentration measurements (e.g., Wilson et al., 1986; Mull et al., 1988; Kilpatrick and Wilson, 1989; Field, 2002, 2003; Schudel et al., 2002; Benischke et al., 2007), but do not emphasize the need to collect more than a single background measurement. Collecting only a single background measurement suggests no temporal variability in the flow system.

Background tracer concentrations are caused by a variety of factors. For fluorescent dyes, background signals may develop from natural substances such as dissolved organic carbon (DOM). Fluorescent properties associated with DOM develop from highly reactive, oxygen-rich functional groups that are accompanied by phenols, amines, alcohols, inorganic iron and aluminum species, and other organic acids (Brown, 2009, p. 15). Anthropogenic sources of fluorescence typically

* Tel.: +1 703 347 8601; fax: +1 703 347 8692.
E-mail address: field.malcolm@epa.gov.
0043-1354/$ – see front matter Published by Elsevier Ltd.
doi:10.1016/j.watres.2011.03.018

detected in background samples include optical brighteners used in laundry detergents, antifreeze colored with uranine dye, mineral oil, hydraulic fluids, polycyclic aromatic hydrocarbons, landfill leachate, some agricultural chemicals and pharmaceuticals, and wood preservatives (Brown, 2009, pp. 20–21), Other commonly-used tracing agents, such as salts, can be useful by increasing the electrical conductivity. However, most cations (e.g., Li^+, Na^+, K^+, Mg^{2+}, Ca^{2+}, and Sr^{2+}) are prone to ion exchange, but this is not necessarily so with anions (Benischke et al., 2007). Anions, such as Br^- and Cl^-, are significantly affected by such factors as variances in atmospheric concentrations, temporal, seasonal, and high and low precipitation, dry deposition, run-on and runoff processes, soil characteristics, evapotranspiration, and land-use changes (Guan et al., 2010; Hrachowitz et al., 2010). In addition, widespread application of NaCl by such activities as road salting (Leibundgut et al., 2009, p. 105) cause elevated background concentrations of Cl^-.

Because hydrologic tracer tests are adversely affected by high background concentrations, accurate and precise background concentration measurements are of critical importance. Hydrologic tracer tests consist of the measurement of tracer concentrations in water samples from which background concentrations are subtracted. However, separation of representative background concentrations from an injected tracer substance is difficult if not impossible after injection of the tracer substance. In general, after a tracer substance has been injected, separation of the sample signal from the background signal cannot be accomplished (Brandt, 1999, p. 169). According to Brandt the expected number of signals in an experiment is a Poisson distribution in which $\lambda = \lambda_S + \lambda_B$ where λ represents signal measurements. To obtain information about the sought after parameter for the number of signal events λ_S the parameter for the background events λ_B must be known. This is problematic for accurate tracer-breakthrough curve (BTC) analysis because of the near impossibility of separating the two signals.

Typically, for hydrologic tracer tests, several background measurements are taken prior to tracer injection of any tracer substance. After tracer injection, background concentration samples can no longer be collected because it is impossible to know if the measurement represents a background signal, or an injection signal. These multiple background measurements may also be problematic because of temporal differences between the time that the background measurements were taken and the time of tracer injection. In addition, there is no consensus method by which the measured background concentrations may be utilized when subtracting from measured tracer concentrations. Most commonly, the arithmetic mean of the background concentrations is calculated and this parameter is then subtracted from the measured tracer recovery data.

Unfortunately, the sample mean is a very non-robust statistic and as such is adversely affected by outliers (Daszykowski et al., 2007). Whether a result of instrumental errors, transcription errors, or some other cause, outliers are a serious problem. Incorporating one or more outliers in the calculation of a mean background concentration (MBC) leads to an excessively large estimate for the MBC. Subtracting an overly large MBC from measured tracer recovery data may actually result in some negative concentration values, an obvious impossibility. Resulting negative concentrations are usually converted to zero concentrations as a matter of convenience, not for any valid scientific or statistical reason.

In the statistical sense, the term, outliers, is an ill-defined concept without clear boundaries, but is useful when regarded as a continuous transition to ordinary observations (Hampel, 2001). Outliers may be a result of gross errors in analysis (e.g., transcription errors, laboratory-analysis errors, or computer-controlled-equipment problems) or they may be a result of anomalous field conditions such as an unusual spike in background concentrations. A common reaction to outliers according to (Hampel, 2001) is to subjectively or objectively reject outliers when, in principle, outliers should be given separate treatment.

An additional problem with the MBC occurs when some measured concentrations are greater than or equal to zero, but less than the method detection limit (DL). The DL is the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero when determined from analysis of a sample in a given matrix containing the analyte (USEPA, 2009). Because the DL is a statistical concept it is quite possible that a substance can be detected at concentrations below the DL (Ripp, 1996). According to Ripp, censoring data below some unspecified or non-statistical reporting limit severely biases data sets and restricts its usefulness. This censoring can lead to erroneous decisions when calculating sample means or mass balances. A number reported as less than with no corresponding information is very difficult to interpret, and often must be discarded even though it may have some validity. The lower the DL, the more likely the analyte of interest will be detected in a sample.

Measured concentration data that are below the detection limit (BDL) are commonly termed left-censored data (Antweilier and Taylor, 2008) or Type I censored data (Kroll and Stedinger, 1996), and have been a source of controversy for many years. Most recently, environmental regulators have struggled with development of an appropriate methodology for incorporating BDL environmental samples into site pollution assessments. For example, commonly-used approaches entail deleting the censored data (Helsel, 2005, p. 11) or multiplying the DL by some constant falling within the interval [0,1] the result of which is then substituted for all BDL data. Typically, the chosen constant takes a value of zero, 0.5, or 1.0 (El-Shaarawi and Esterby, 1992; Singh and Nocerino, 2001). Other substitution methods replace left-censored data with the DL multiplied by 0.75 as has been done with geochemical data or by $DL/\sqrt{2}$ as is sometimes done for air-quality data and industrial-hygiene chemistry (Hewett and Ganser, 2007; Helsel, 2005, p. 56). However, there is no scientific basis to support such a replacement methodology. Such data fabrication reportedly can produce seriously biased estimates (Gleit, 1985; Gilliom and Helsel, 1986; Haas and Scheff, 1990; Helsel, 2006), although a reportedly accurate substitution method was recently developed (Ganser and Hewett, 2010).

There will always be questions regarding the reliability of background concentration measurements and analysis methods, but appropriate statistical methods have the advantage of repeatability. Subjective methods for dealing with background measurements, based on experience and/or judgment may arguably be a justifiable approach in some circumstances by some experienced individuals, but such an approach may not be as defensible as a valid statistical approach. The purpose of this paper is to investigate and develop a statistically-valid approach for assessing the MBC subjected to outliers and BDL data as it affects the measured BTC.

2. Robust statistical methods

For every hydrologic tracer test, an adequate number of background samples are needed to properly characterize background variability. Obtaining a reasonably large background sample set makes various forms of data plotting possible, which allows trends in the data to be observed, and may help in determining appropriate statistical routines to employ.

The sample mean $\bar{x}_b$ and standard deviation s_b for background concentrations is calculated from, respectively,

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

and

$$s_b = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_b)^2}, \tag{2}$$

where n is the sample size and x_i is a sample measurement. Eqs. (1) and (2) consider all measured values equally so undue influence may be imparted by one or more extreme low or high values, especially for small sample sizes, unless varying weights are assigned to the measured values. Unfortunately, it is very difficult to assign adequate weights to the measurements if specific details about each measurement are lacking. When weights cannot be applied to the data, alternative robust statistical methods need to be employed.

A measure of the robustness of a statistical estimator against outliers is the breakdown value, which may be defined as an indicator of the smallest fraction of contaminants (errors) in a sample that causes the estimator to breakdown (i.e., to take on values that are arbitrarily bad or meaningless) (Hubert and Debruyne, 2009). The closer the breakdown value to 100%, the better. Another measure of robustness, the influence function, measures the effect of a small number of outliers (Rousseeuw et al., 2006). The influence function is defined as the influence of an infinitesimal proportion of bad observations on the value of the estimate (Cowell and Victoria-Feser, 1993). It is essentially the first derivative of an estimator (Hampel, 1974) that measures the effect or influence of a single observed value on an estimator of a particular parameter. The breakdown value for the sample mean and standard deviation is $1/n$, or 0% for large n. The influence function for the mean is unbounded and thus reflects the non-robustness of the sample mean (Rousseeuw et al., 2006). For these two reasons alone, the sample mean should be accepted as representative of background tracer concentrations only after careful consideration and assessment.

2.1. Outlier determination

Outliers in a set of data are observations that are far from the bulk of the data (Olive, 2008, p. 4). Because data outliers can be very problematic in data analysis it is critical that they be accurately identified and properly addressed.

Background data plotting can be advantageous for identifying outliers. Constructing a simple scatter plot, histogram, dot plot, boxplot, quantile plot, or stem and leaf plot may allow detection of possible outliers (Chambers et al., 1983, pp. 11–29). For example, a typical boxplot will include whiskers representing the data points at the tenth and ninetieth percentiles of the data. Any data plotted beyond these two whiskers may be data outliers. Unfortunately, robust statistical methods can only detect certain configurations of outliers and the ability to detect outliers rapidly decreases as the sample size and the number of predictors increase (Olive, 2008, p. 9). Because of the difficulty associated with outlier determination, improvements in detection methods continue to be investigated (e.g., Hubert and Van der Veeken, 2008).

If any outliers are believed to exist based on one or more of the data plotting methods, then appropriate statistical methods for dealing with the outliers need to be considered. It is possible to remove outliers based on personal experience and judgment, but such an action often lacks statistical validity. Alternative actions consider outlier accommodation and removal by statistical determination.

2.1.1. Outlier accommodation

Accommodating outliers generally is not a common practice, but some studies suggest that this may be a more reliable method for addressing outliers than is removal (AMC, 1989). The approach recommended by the (AMC, 1989) consists of a set of iterative calculations for estimating the population standard deviation σ_b alongside the population mean μ_b. At each iteration, pseudo-values $\tilde{x}_i$ are formed and their mean $\bar{x}_b$ and variance s_b^2 computed. This allows for basic calculations that are repeated until the values stabilize.

2.1.2. Outlier removal

There are several methods for statistically removing outliers when calculating a mean and standard deviation. The easiest but least acceptable method is to just arbitrarily remove any apparent outliers from a data set. A more acceptable method is to employ some statistical procedure to remove outliers after a scientific determination has been taken regarding the appropriateness of all of the data.

Statistical determination and removal of outliers can be accomplished using any one of several robust methods. One powerful method developed at NASA (Swaroop and Winter, 1971) identifies outliers by computing, for a given significance level, a critical value Δ^* for the data set and a positive real number Δ_i for each observation vector, x_i where

$$\Delta_i = \frac{(x_i - \bar{x})^T}{S}(x_i - \bar{x}), \tag{3}$$

$$\Delta_* = \frac{p(n-1)^2 F_{\alpha;p,n-p-1}}{n(n-p-1) + npF_{\alpha;p,n-p-1}}, \tag{4}$$

$$S = \frac{1}{n-1}\sum_{1}^{n}(x_i - \overline{x})(x_i - \overline{x})^T, \tag{5}$$

$F_{\alpha;\, p, n-p-1}$ is the significance level (either 5 percent or 1 percent) of the F-distribution with p and ($n-p-1$) degrees of freedom, p is the number of variables, and T is the transpose of the matrix. If $\Delta i > \Delta^*$, observation x_i is identified as an outlier and removed from further analysis.

Another common approach is to calculate the median *MED* and median absolute deviation *MAD* from, respectively (Olive, 2008, p. 27)

$$\text{MED} = \begin{cases} x_{[(n+1)/2]} & \text{if } n \text{ is odd} \\ \dfrac{x_{(n/2)} + x_{[n/(2+1)]}}{2} & \text{if } n \text{ is even} \end{cases}, \tag{6}$$

$$\text{MAD} = \underset{j=1,\ldots,n}{\text{MED}}\left|x_j - \underset{i=1,\ldots,n}{\text{MED}}\,(x_i)\right|, \tag{7}$$

and a robust standard deviation estimate MAD_E from (Burke, 2001)

$$\text{MAD}_\text{E} = 1.4826\ \text{MAD}. \tag{8}$$

Because MED measures the exact middle of a set of data it is insensitive to outliers. Replacing a single value by some arbitrary value does not result in much of a change in the MED. Significantly, the breakdown value of the median is 50%, which means that it can resist up to 50% outliers and its influence function is bounded. However, the median is a less efficient estimator than is the mean for the normal model (Rousseeuw et al., 2006) because the MED is adversely affected by heavy-tailed distributions close to the normal (AMC, 1989), which renders the MED and its associated estimators problematic.

A different approach when 15–50% of the data are non-detects is to calculate a trimmed mean $\overline{x}_t$ or winsorized mean $\overline{x}_w$ of the data as suggested by USEPA (2000, pp. 4-42–4-44) and Navy (1999, pp. 11–15). The trimmed mean $\overline{x}_t$ and associated standard deviation s_t are given by

$$\overline{x}_t = \frac{1}{n-2j}\sum_{i=j+1}^{n-j} x_i, \tag{9}$$

and

$$s_t = \sqrt{\frac{1}{n-2j-1}\sum_{i=j+1}^{n-j}(x_i - \overline{x}_t)^2}, \tag{10}$$

and the winsorized mean $\overline{x}_w$ and associated standard deviation s_w are given by (Irwin, 2005)

$$\overline{x}_w = \frac{1}{n}\left[(j+1)x_{(j+1)} + \sum_{i=j+2}^{n-j-1} x_i + (j+1)x_{(n-j)}\right], \tag{11}$$

and

$$s_w = \left[\frac{1}{n-1}(j+1)\left(x_{(j+1)} - \overline{x}_w\right)^2 + \sum_{i=j+2}^{n-j-1}(x_i - \overline{x}_w)^2 + (j+1)\left(x_{(n-j)} - \overline{x}_w\right)^2\right]^{0.5}, \tag{12}$$

where j represents the number of trimmed or winsorized observations.

The trimmed mean approach removes the least and greatest values and the descriptive statistics are calculated from what is left while the winsorized mean approach replaces the least and greatest values with each's nearest neighbors. Both approaches are robust and generally result in better approximations for the data mean and standard deviation, but the trimmed mean is also insensitive to small numbers of gross errors and is not adversely affected by heavy-tailed distributions close to the normal (AMC, 1989).

2.2. *Left-censored data*

Whether any outliers exist or not, most background concentration data sets contain a significant number of values that equal zero (a data floor) and a significant number of values that are greater than zero but less than the DL (i.e., left-censored data). Such concentrations result in positively skewed data. Measured zero concentrations are still real measurements and need to be included in any statistical calculations, but left-censored data require evaluation using appropriate statistical techniques.

A comprehensive study on censored data analysis methods (EFSA, 2010) reported that there are three appropriate methods for handling left-censored data: parametric maximum likelihood estimation (MLE) models, the log-probit regression method, and the non-parametric Kaplan–Meier method. In this study, only the MLE was evaluated for background concentration data.

2.2.1. *Maximum likelihood estimation*

The MLE is reported to be the gold standard for evaluating left-censored data (Hewett and Ganser, 2007; Ganser and Hewett, 2010) and has been shown to be perhaps the best method to choose (Kroll and Stedinger, 1996), despite some suggestions that it may not be quite as good as generally accepted (Thompson and Nelson, 2003). For example, although the MLE is considered the best of the methods usually applied to censored data, Helsel (2005, p. 13) reported that the MLE is only valid for data sets with ≥50 observations and is sensitive to outliers. However, the MLE is still preferred over other non-parametric methods where either the percent of censored data is small or the shape of the distribution can be well characterized. Ganser and Hewett (2010) recently showed that the MLE can be applied to sample sizes as small as $n = 3$ provided there are at least two uncensored values, but they also found that for data sets of $n \le 10$, the estimates for the sample geometric mean and geometric standard deviation can be severely biased, particularly when the actual percent censored data is large.

For the MLE to work, the data are assumed to follow some statistical distribution and the parameters of the distribution are estimated to compatibly work with the percentage of BDL data to best fit the distribution of the observed values above the DL. The estimated parameters are the ones that maximize the likelihood function (EFSA, 2010).

The MLE is commonly applied to the log-normal for concentration values that, when log-transformed approximates a normal distribution, the Weibull distribution, and the gamma distribution. Analyses described by the EFSA (2010) suggest that for left-censored data, a gamma distribution may inflict the least bias when the MLE method is used.

The sample geometric mean $\bar{x}_g$ and geometric standard deviation s_g are the values that maximize the likelihood function (LF) obtained from (Hewett and Ganser, 2007)

$$LF = \prod_{i=k+1}^{n} \mathrm{PDF}(\ln x_i | \ln \bar{x}_g, \ln s_g) \times \prod_{j=1}^{k} \mathrm{CDF}(\ln x_j | \ln \bar{x}_g, \ln s_g), \quad (13)$$

where k is the number of censored data, PDF is the probability density function, CDF is the cumulative distribution function, and the $\bar{x}_g$ and s_g are calculated from, respectively

$$\bar{x}_g = \left(\prod_{i=1}^{n} x_i\right)^{1/n} = \exp\left(\frac{1}{n}\sum_{i=1}^{n} \ln\ x_i\right), \quad (14)$$

and

$$s_g = \exp\left(\sqrt{\frac{\sum_{i=1}^{n}(\ln\ x_i - \ln\ \bar{x}_g)^2}{n-1}}\right). \quad (15)$$

3. Quantitative example

To demonstrate the effect of outliers and left-censored data on descriptive statistics of background concentration data, a synthetic set of background concentration data values (Fig. 1) were applied to a selected measured BTC (Fig. 2) that was slightly modified for this analysis. The synthetic background data consists of two apparent outliers, several instances of zero concentration, and eight left-censored concentration values based on a DL = 10 ng L^{-1}. Calculating the MBC from the synthetic data set should equal 0.01 μg L^{-1} (representing the background events signal λ_B) for subtraction from the measured BTC (representing the signal measurements λ) if the resulting BTC (representing signal events λ_S) is to be obtained.

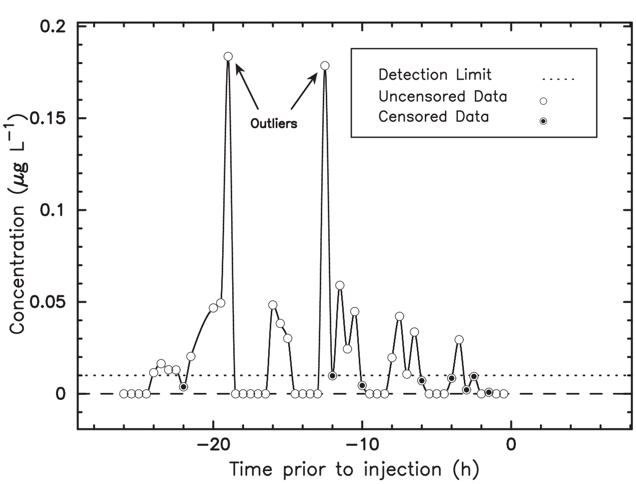

Fig. 1 – Data plot of pre-injection concentration data. Pre-injection background concentration data include two apparent outliers, several instances of zero concentration, and eight left-censored concentration values (concentrations greater than zero, but less than the method detection limit).

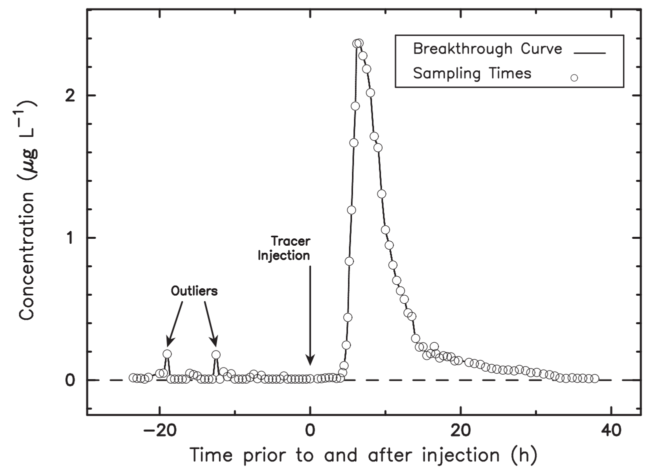

Fig. 2 – Data plot of combined pre-injection and post-injection concentration data. Synthetic pre-injection background concentration data include two apparent outliers. Note that zero time represents time of tracer injection.

The BTC depicted in Fig. 2 was developed on top of a constant background concentration equal to 0.01 μg L^{-1} (i.e., $\bar{x}_b$ = 0.01 μg L^{-1} and sb = 0.0 μg L^{-1}). Subtracting 0.01 μg L^{-1} from each measured tracer concentration results in the actual recovered tracer data. Evaluating the BTC shown in Fig. 2 using Qtracer2 (Field, 2002, pp. 25–30) resulted in a ~96% tracer-mass recovery. A perfect 100% tracer-mass recovery almost never occurs because of errors in discharge measurements, tracer sorption, and tracer decay (Field, 2002, pp. 25–30).

The combined plot of the pre-injection (synthetic background data) and post-injection data shown in Fig. 2 suggests the existence of the two apparent background data outliers. However, if the peak recovery concentration was significantly greater than that shown in Fig. 2 (e.g., ≥2× the peak), the existence of the outliers would be less apparent.

The basic statistical distribution for the synthetic background data is depicted in Fig. 3, which shows just how extreme the two outliers in the synthetic data set really are. No lower whisker was calculated because a data floor representing zero concentration measurements occur in the synthetic background data. The dashed line in Fig. 3 is the sample mean, $\bar{x}_b$ = 0.0192, which denotes a bias toward the stronger concentrations even though nearly half of the data equals zero concentration. The interquartile range (IQR = 0.0244), however, clearly shows that the median value (MED = 0.0042) tends to the lower end of the boxplot because of the influence of the number of data values equal to zero concentration.

3.1. *Descriptive statistics estimates*

Estimates for $\bar{x}_b$ and s_b were initially developed for the full data set along with the MED, MAD, and MAD_E. The trimmed and winsorized descriptive statistics were also calculated. All calculations were then repeated after outlier removal according to the method developed by Swaroop and Winter (1971) at the 1 percent significance level. The results for the

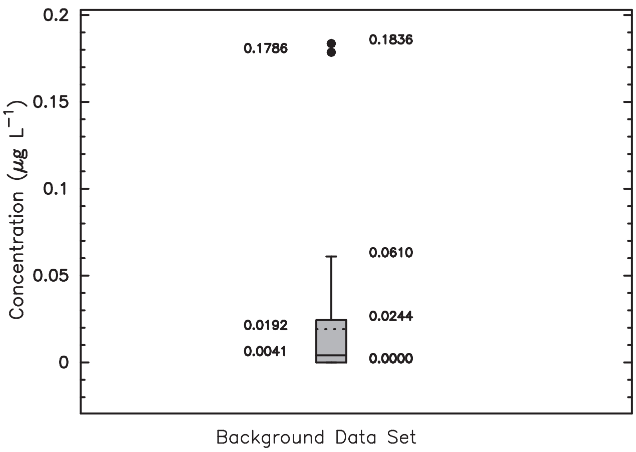

Fig. 3 – Boxplot of synthetic pre-injection background concentration data. Note the two apparent outliers. Dashed line represents the calculated mean background concentration.

descriptive statistics estimates are listed with outliers included in Table 1 and with outliers deleted in Table 2.

A performance index for the mean estimates may be determined using the root mean square error rMSE (modified from Helsel and Hirsch, 2002, p. 358).

$$rMSE = \sqrt{\frac{(\bar{x}_b - \mu_b)^2/\mu_b}{n}}, \quad (16)$$

where $\bar{x}_b$ represents the background estimate for all methods described and μ_b represents the true background concentration. The nearer rMSE to zero, the better the performance index.

Calculation of the rMSE for the various background estimates are also shown in Tables 1 and 2. However, a perhaps more useful performance index for the background estimates is subtraction of each estimate from the measured tracer BTC. Results of subtracting the background estimates from the measured BTC are shown in Table 3, the effects of which are listed in Table 4 and plotted in Fig. 4.

3.2. *Assessment of the estimation methods*

The significance of using estimates for the background mean that deviate from the expected 0.0100 μg L^{-1} is evident when the rMSE for each background estimate listed in Tables 1 and 2 are examined. On inspection of both Tables 1 and 2, it is initially apparent that removing outliers from the background data (Table 2) results in rMSE calculations that are closer to zero than when outliers are included (Table 1). Removal of outliers appears to be appropriate in this instance (the MLE estimate of 0.0100 μg L^{-1} results in a rMSE = 0.01), but such an action is still dependent on a careful assessment of the validity of the data outliers.

Background estimates ranged from a low of 0.0 μg L^{-1} for a median calculation for instances in which BDL data are deleted (or substitution of the DL multiplied by zero) for either instance of outliers inclusion or exclusion to a high of 0.1176 μg L^{-1} for an MLE with outliers included (Tables 1 and 2). This suggests that the median is strongly affected by a large number of background measurements equal to zero concentration while the MLE is strongly affected by outliers in the data. Methods resulting in a rMSE = 0.01 (i.e, background estimate = 0.0100 μg L^{-1}) occurs when the MLE is applied to the data after deleting the outliers, but also for instances when the DL × 1.0 is substituted for BDL data for the median estimate regardless of whether outliers are included or excluded. Application of the MLE to data that includes outliers results in a background estimate that represents more than a five-fold increase over the next largest background estimate (0.0217 μ g L^{-1}), which suggests that blind use of the MLE could result in extremely serious errors in background estimates. However, data fabrication methods, such as substituting various fractions of the DL for BDL data for MBC estimation, is unsupported by any scientific or statistical basis and cannot be assured to produce acceptable results when applied to other data sets.

If outliers are accommodated the estimated mean ranges from a low of 0.0071 μg L^{-1} with a standard deviation of 0.0062 μg L^{-1} (method A15 in AMC (1989)) to a high of 0.0119 μg L^{-1} with a standard deviation of 0.0164 μg L^{-1} (method H15 in AMC (1989)). Although neither estimate for the

Table 1 – Descriptive statistics for synthetic background concentration data with outliers included.[a]

Censoring Method	Analysis Method														
	Arithmetic Mean			Trimmed Mean			Winsorized Mean			Data Median[b]			MLE		
	$\bar{x}_b$	s_b	rMSE	$\bar{x}_t$	s_t	rMSE	$\bar{x}_w$	s_w	rMSE	MED	MADe	rMSE	$\bar{x}_b$	s_b	rMSE
<DL Included	0.0192	0.0374	0.0130	0.0110	0.0144	0.0014	0.0135	0.0130	0.0049	0.0042	0.0062	0.0083	0.1176	0.2516	0.1522
<DL Deleted	0.0217	0.0403	0.0181	0.0130	0.0167	0.0047	0.0151	0.0151	0.0079	0.0000	0.0000	0.0154	–	–	–
0.0 DL Substituted	0.0183	0.0377	0.0117	0.0099	0.0149	0.0002	0.0126	0.0136	0.0036	0.0000	0.0000	0.0141	–	–	–
0.5 DL Substituted	0.0191	0.0374	0.0128	0.0109	0.0144	0.0012	0.0134	0.0130	0.0048	0.0050	0.0074	0.0071	–	–	–
0.75 DL Substituted	0.0195	0.0372	0.0134	0.0114	0.0142	0.0019	0.0138	0.0129	0.0053	0.0075	0.0111	0.0035	–	–	–
1.0 DL Substituted	0.0199	0.0371	0.0140	0.0119	0.0141	0.0026	0.0142	0.0128	0.0059	0.0100	0.0148	0.0000	–	–	–
DL/$\sqrt{2}$ Substituted	0.0194	0.0373	0.0133	0.0113	0.0142	0.0018	0.0137	0.0129	0.0052	0.0071	0.0105	0.0041	–	–	–

a All background value estimates and standard deviation estimates have units of μg L^{-1}.
b Although the median cannot be construed to be the same as the mean, it is appropriate to compare the sample median MED with the true background mean μ_b here because the sample median is being considered for subtraction from the measured breakthrough curve.

Table 2 – Descriptive statistics for synthetic background concentration data with outliers deleted.[a]

Censoring Method	Analysis Method: Arithmetic Mean $\bar{x}_b$	s_b	rMSE	Trimmed Mean $\bar{x}_t$	s_t	rMSE	Winsorized Mean $\bar{x}_w$	s_w	rMSE	Data Median[b] MED	MADe	rMSE	MLE $\bar{x}_b$	s_b	rMSE
<DL Included	0.0124	0.0171	0.0035	0.0098	0.0133	0.0002	0.0119	0.0123	0.0028	0.0030	0.0045	0.0101	0.0100	0.0129	0.0000
<DL Deleted	0.0138	0.0184	0.0060	0.0109	0.0146	0.0013	0.0132	0.0132	0.0050	0.0000	0.0000	0.0158	–	–	–
0.0 DL Substituted	0.0115	0.0176	0.0021	0.0087	0.0138	0.0019	0.0110	0.0127	0.0014	0.0000	0.0000	0.0144	–	–	–
0.5 DL Substituted	0.0123	0.0171	0.0033	0.0097	0.0133	0.0005	0.0118	0.0122	0.0026	0.0050	0.0074	0.0072	–	–	–
0.75 DL Substituted	0.0127	0.0170	0.0039	0.0102	0.0131	0.0003	0.0122	0.0121	0.0032	0.0075	0.0111	0.0036	–	–	–
1.0 DL Substituted	0.0132	0.0169	0.0045	0.0107	0.0131	0.0001	0.0126	0.0120	0.0038	0.0100	0.0148	0.0000	–	–	–
DL/$\sqrt{2}$ Substituted	0.0127	0.0170	0.0038	0.0101	0.0131	0.0002	0.0122	0.0121	0.0031	0.0071	0.0105	0.0042	–	–	–

a All background value estimates and standard deviation estimates have units of μg L^{-1}.
b Although the median cannot be construed to be the same as the mean, it is appropriate to compare the sample median MED with the true background mean μ_b here because the sample median is being considered for subtraction from the measured breakthrough curve.

Table 3 – Number and location of data points less than zero after subtracting background concentration value.

Analysis Method	Outliers Included: Background Value, μg L^{-1}	Data Pts. Less Than Zero	Data Points Affected: Early Time Data Pts.	Late Time Data Pts.	Outliers Excluded: Background Value, μg L^{-1}	Data Pts. Less Than Zero	Data Points Affected: Early Time Data Pts.	Late Time Data Pts.
Data Mean								
<DL Included	0.0192	15	1–8	62–68	0.0124	7	1–3, 7	63, 67–68
<DL Deleted	0.0217	15	1–8	62–68	0.0138	8	1–3, 7	63, 65, 67–68
0.0 DL Substituted	0.0183	14	1–5, 7–8	62–68	0.0115	5	1–3	67–68
0.5 DL Substituted	0.0191	15	1–8	62–68	0.0123	7	1–3, 7	63, 67–68
0.75 DL Substituted	0.0195	15	1–8	62–68	0.0127	7	1–3, 7	63, 67–68
1.0 DL Substituted	0.0199	15	1–8	62–68	0.0132	8	1–3, 7	63, 65, 67–68
DL/$\sqrt{2}$ Substituted	0.0194	15	1–8	62–68	0.0127	7	1–3, 7	63, 67–68
Trimmed Mean								
<DL Included	0.0110	4	1–3	67	0.0098	0	–	–
<DL Deleted	0.0130	7	1–3, 7	63, 67–68	0.0109	3	1–2	67
0.0 DL Substituted	0.0099	0	–	–	0.0087	0	–	–
0.5 DL Substituted	0.0109	3	1–2	67	0.0097	0	–	–
0.75 DL Substituted	0.0114	5	1–3	67–68	0.0102	2	1–2	–
1.0 DL Substituted	0.0119	7	1–3, 7	63, 67–68	0.0107	2	1–2	–
DL/$\sqrt{2}$ Substituted	0.0113	5	1–3	67–68	0.0101	1	1	–
Winsorized Mean								
<DL Included	0.0135	8	1–3, 7	63, 65, 67–68	0.0119	7	1–3, 7	63, 67–68
<DL Deleted	0.0151	10	1–3, 7–8	63, 65–68	0.0132	8	1–3, 7	63, 65, 67–68
0.0 DL Substituted	0.0126	7	1–3, 7	63, 67–68	0.0110	4	1–3	67
0.5 DL Substituted	0.0134	8	1–3, 7	63, 65, 67–68	0.0118	6	1–3	63, 67–68
0.75 DL Substituted	0.0138	8	1–3, 7	63, 65, 67–68	0.0122	7	1–3	63, 67–68
1.0 DL Substituted	0.0142	8	1–3, 7	63, 65, 67–68	0.0126	7	1–3	63, 67–68
DL/$\sqrt{2}$ Substituted	0.0137	8	1–3, 7	63, 65, 67–68	0.0122	7	1–3, 7	63, 67–68
Data Median								
<DL Included	0.0042	0	–	–	0.0030	0	–	–
<DL Deleted	0.0000	0	–	–	0.0000	0	–	–
0.0 DL Substituted	0.0000	0	–	–	0.0000	0	–	–
0.5 DL Substituted	0.0050	0	–	–	0.0050	0	–	–
0.75 DL Substituted	0.0075	0	–	–	0.0075	0	–	–
1.0 DL Substituted	0.0100	0	–	–	0.0100	0	–	–
DL/$\sqrt{2}$ Substituted	0.0071	0	–	–	0.0071	0	–	–
Max. Likelihood Estimate								
Left-Censored Data	0.1176	32	1–10	47–68	0.0100	0	–	–

Table 4 – Effect of subtracting background value estimates on percent mass recovered and time of travel from the measured breakthrough curve.

	Outliers Included					Outliers Excluded				
		Negative Concentrations Retained		Negative Concentrations Converted to Zero			Negative Concentrations retained		Negative Concentrations Converted to Zero	
Analysis Method	Background Value, μg L^{-1}	% Mass Recovered	Mean Travel Time, h	% Mass Recovered	Mean Travel Time, h	Background Value, μg L^{-1}	% Mass Recovered	Mean Travel Time, h	% Mass Recovered	Mean Travel Time, h
Data Mean										
<DL Included	0.0192	93.8	9.86	94.2	9.90	0.0124	95.7	10.03	95.7	10.04
<DL Deleted	0.0217	93.2	9.79	93.7	9.86	0.0138	95.3	10.00	95.4	10.01
0.0 DL Substituted	0.0183	94.1	9.88	94.4	9.92	0.0115	95.9	10.06	96.0	10.06
0.5 DL Substituted	0.0191	93.9	9.86	94.3	9.90	0.0123	95.7	10.04	95.8	10.04
0.75 DL Substituted	0.0195	93.8	9.85	94.2	9.90	0.0127	95.6	10.03	95.7	10.03
1.0 DL Substituted	0.0199	93.7	9.84	94.1	9.89	0.0132	95.5	10.01	95.6	10.02
DL/$\sqrt{2}$ Substituted	0.0194	93.8	9.85	94.2	9.90	0.0127	95.6	10.03	95.7	10.03
Trimmed Mean										
<DL Included	0.0110	96.1	10.07	96.1	10.07	0.0098	96.4	10.10	96.4	10.10
<DL Deleted	0.0130	95.5	10.02	95.6	10.02	0.0109	96.1	10.07	96.1	10.07
0.0 DL Substituted	0.0099	96.4	10.10	96.4	10.10	0.0087	96.7	10.13	96.7	10.13
0.5 DL Substituted	0.0109	96.1	10.07	96.1	10.07	0.0097	96.4	10.10	96.4	10.10
0.75 DL Substituted	0.0114	96.0	10.06	96.0	10.06	0.0102	96.3	10.09	96.3	10.09
1.0 DL Substituted	0.0119	95.8	10.05	95.9	10.05	0.0107	96.2	10.08	96.2	10.08
DL/$\sqrt{2}$ Substituted	0.0113	96.0	10.06	96.0	10.06	0.0101	96.3	10.09	96.3	10.09
Winsorized Mean										
<DL Included	0.0135	95.4	10.01	95.5	10.01	0.0119	95.8	10.05	95.9	10.05
<DL Deleted	0.0151	95.0	9.97	95.1	9.98	0.0132	95.5	10.01	95.6	10.02
0.0 DL Substituted	0.0126	95.6	10.02	95.7	10.03	0.0110	96.1	10.07	96.1	10.07
0.5 DL Substituted	0.0134	95.4	10.01	95.5	10.01	0.0118	95.9	10.05	95.9	10.05
0.75 DL Substituted	0.0138	95.3	10.00	95.4	10.01	0.0122	95.8	10.04	95.8	10.04
1.0 DL Substituted	0.0142	95.2	9.99	95.3	10.00	0.0126	95.6	10.03	95.7	10.03
DL/$\sqrt{2}$ Substituted	0.0137	95.3	10.00	95.4	10.01	0.0122	95.8	10.04	95.8	10.04
Data Median										
<DL Included	0.0042	97.9	10.24	97.9	10.24	0.0030	98.3	10.27	98.3	10.27
<DL Deleted	0.0000	99.1	10.34	99.1	10.34	0.0000	99.1	10.34	99.1	10.34
0.0 DL Substituted	0.0000	99.1	10.34	99.1	10.34	0.0000	99.1	10.34	99.1	10.34
0.5 DL Substituted	0.0050	97.7	10.22	97.7	10.22	0.0050	97.7	10.22	97.7	10.22
0.75 DL Substituted	0.0075	97.0	10.16	97.0	10.16	0.0075	97.0	10.16	97.0	10.16
1.0 DL Substituted	0.0100	96.4	10.09	96.4	10.09	0.0100	96.4	10.09	96.4	10.09
DL/$\sqrt{2}$ Substituted	0.0071	97.1	10.17	97.1	10.17	0.0071	97.1	10.17	97.1	10.17
Max. Likelihood Estimate										
Left-Censored Data	0.1176	67.1	6.25	78.0	8.60	0.0100	96.4	10.09	96.4	10.09

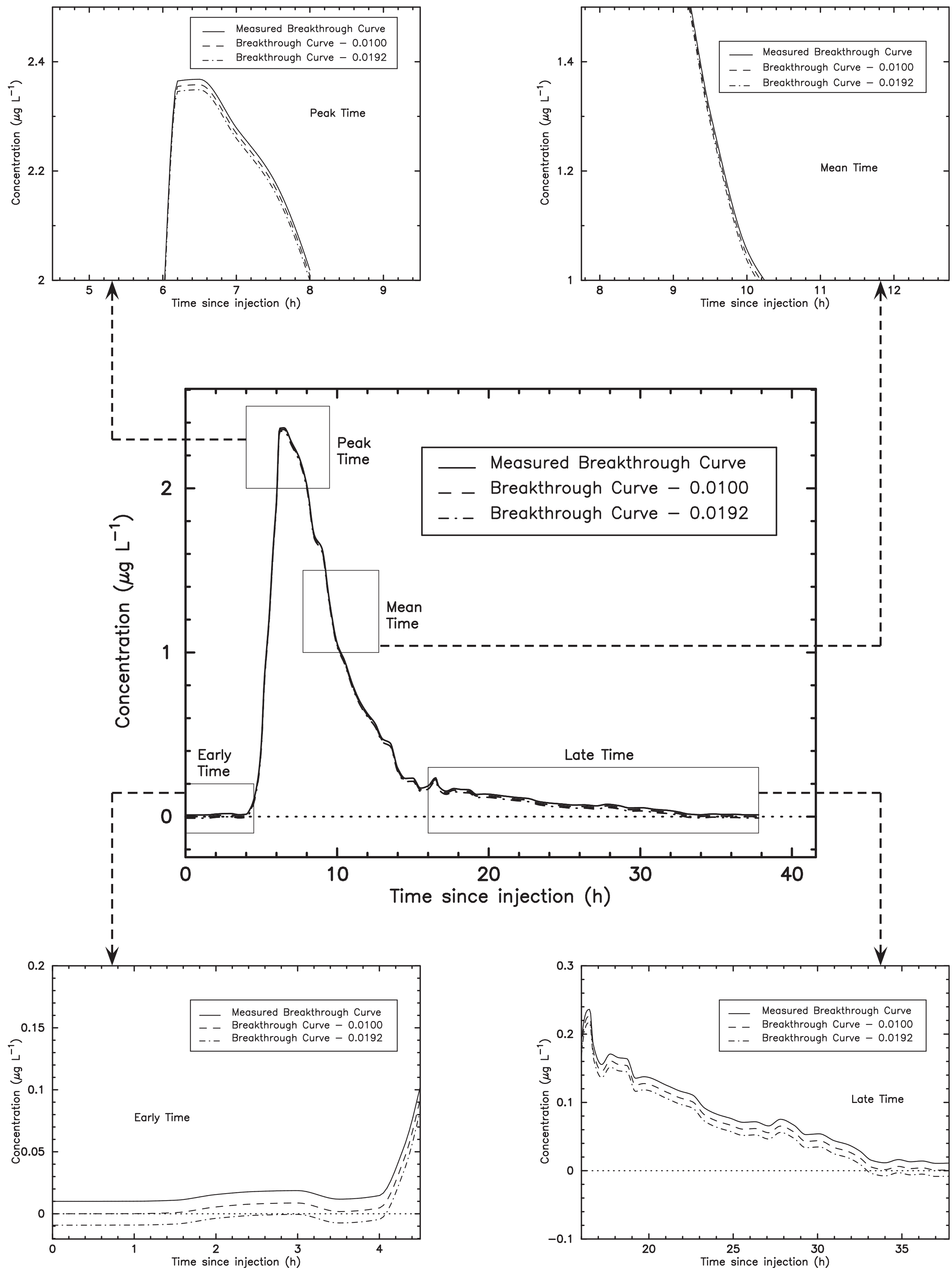

Fig. 4 – Effects of subtracting an arithmetic mean background value ($\bar{x}_b = 0.0192$ μg L^{-1}) and a maximum likelihood estimate (MLE = 0.0100 μg L^{-1}) from the measured breakthrough curve. Early time, peak time, mean time, and late-time plots are shown enlarged to emphasize the effect of subtracting an overly large background value ($\bar{x}_b = 0.0192$ μg L^{-1}) causing negative concentration values in the early time and late-time plots and the importance of correct the MLE estimate of 0.0100 μg L^{-1}.

mean is exactly 0.0100 μg L^{-1}, both are sufficiently close to 0.0100 μg L^{-1} to be considered acceptable. A rMSE = 0.0041 for the A15 estimate and 0.0027 for the H15 estimate implies that these two estimates underestimate the expected value of 0.0100 μg L^{-1}.

Removing the outliers results in an A15 = 0.0055 μg L^{-1} with a standard deviation of 0.0044 μg L^{-1} and an rMSE = 0.0065. A corresponding H15 = 0.0100 μg L^{-1} with a standard deviation of 0.0139 μg L^{-1} results in a rMSE equal to zero. This suggests that the method developed by the (AMC) is also affected by data outliers, but less so than other methods such as the sample mean and the MLE.

3.2.1. *Effect of background estimates on breakthrough curves*

Subtracting background estimates greater than 0.0100 μg L^{-1} results in a significant number of data points less than zero in both the early-time and late-time data (Table 3). For example, subtracting the sample mean using all the data ($\bar{x}_b = 0.0192$ μg L^{-1}) from the measured BTC causes the first eight data points and last seven data points to be less than zero for a total of 15 data points that are less than zero. Traditionally, these 15 negative concentration values would be converted to zero concentration. The net effect of retaining the negative concentrations or conversion to zero concentrations will be to influence basic numerical analyses and overall quantitative assessment of the tracer test. To emphasize this fact, the percent tracer-mass recovered and mean travel time for the example data set (Fig. 2) after subtracting the estimated background value are shown in Table 4. In this instance the traditional mean estimate of 0.0192 μg L^{-1} will result in 93.8% of the mass being recovered and a mean travel time of 9.86h whereas the correct values are 96.4% of the mass recovered and a mean travel time of 10.09 h.

Fig. 4 shows how some early- and late-time data points will be less than zero when the traditional mean value of 0.0192 μg L^{-1} is subtracted from the BTC. Fig. 4 also shows how it is possible to obtain an incorrect percent mass recovery when the early-, peak-, and late-time data are examined. The mean time may not appear greatly affected for this particular BTC, but mean travel time errors become more significant when tailing or right-skew is more severe. Overall differences may not appear extreme for this particular data set, but very large differences could occur with a different data set and different sample mean.

3.2.2. *Potential for calculation errors and incorrect site assessments*

The potential for a large difference in percent mass recovered and mean travel time becomes apparent when the background value estimated by the MLE with the outliers included is applied to the BTC. In this instance, only 67.1% of the mass is recovered, but now the mean travel time is estimated to take only 6.25 h. Such a poor mass recovery typically implies serious data collection and analysis errors, inaccurate discharge estimates, possible tracer migration to unmonitored locations, and/or the assumption that tracer transport is nonconservative. The apparent faster mean travel time also may result in an incorrect sampling frequency for possible pollutants. However, the real source of recovery error in this instance was caused by use of the MLE with inclusion of outliers so the implied physical problems associated with tracer-test design, implementation, and subsequent data collection is incorrect.

Application of the estimated transport parameters in a theoretical model might correct the apparent error, but such a correction should not be expected. Worse, if one or more of the parameters in error is very far from the true value, it is likely that the theoretical transport model will isolate a local minimum rather than continuing to the global minimum and the final assessment regarding solute-transport processes could be drastically wrong.

4. Conclusions

Accurate calculation of a single background concentration for subtraction from a measured tracer-breakthrough curve (BTC) is essential for the calculation of such basic parameters as percent tracer-mass recovered, mean travel time, and related solute-transport parameters. Typical use of the arithmetic mean can often lead to an excessively large background value for subtraction from the measured BTC if one or more outliers are included in the background data set. The creation of negative concentrations that are usually treated as zero concentrations in the resultant BTC routinely occur when the arithmetic mean is used as the background concentration value. Below detection limit data, conventionally known as left-censored data, can also adversely affect background value estimation.

Proper evaluation and possible removal of outliers can improve background value estimation. In addition, application of statistical techniques designed to address censored data have been shown to better represent estimates than the arithmetic mean. However, use of one or more of the common data fabrication methods in which some fraction of the method detection limit is substituted for data below the detection limit has no basis in fact scientifically or statistically. Application of statistical methods designed for censored data, such as the maximum likelihood estimation (MLE), have been shown in the past to generally provide better descriptive statistics than does the arithmetic mean. In particular, when applied to background data, the MLE was found to produce excellent estimates for the descriptive statistics after any outliers were first removed. However, although the MLE appears to have worked well with the synthetic background data set provided, a major problem with the MLE as applied to the background data set is the assumption that the data conforms to a log-normal distribution, which may not necessarily be the case. Although the MLE produced the expected estimate for the sample mean, in reality, the background concentration data appears to conform to a gamma distribution so use of the MLE method in this paper should be viewed with some skepticism. Future work should focus on developing an MLE for gamma-distributed data.

Alternatively, outliers may be accommodated during data analysis, but the possibility that the median value may adversely affect the analysis should be considered. It is still appropriate to carefully evaluate any potential outliers to determine if they should be included or excluded in any outlier accommodation analysis because the outliers may or may not represent actual data measurements. However, whether outliers are accommodated in background data analysis or not, methods that appropriately consider left-censored data in which numerous zero concentration values are included need to be addressed.

Disclaimer

The views expressed in this paper are solely those of the author and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

Acknowledgments

The author would like to thank Dr. John Fox of the National Center for Environmental Assessment of the U.S. Environmental Protection Agency for suggesting an approach for dealing with left-censored data with a data floor and Dr. Leonid Kopalev of the National Center for Environmental Assessment of the U.S. Environmental Protection Agency for his review and constructive criticisms of the manuscript. The author would also like to thank two anonymous reviewers for their comments and suggested improvements to the manuscript.

REFERENCES

AMC, 1989. Robust statistics—how not to reject outliers: Part 1. Basic concepts. Analytical methods committee. Analyst 114, 1693–1697.

Antweilier, R.C., Taylor, H.E., 2008. Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: I. Summary statistics. Environmental Science and Technology 42 (10), 3732–3738.

Benischke, R., Goldscheider, N., Smart, C., 2007. Tracer techniques. In: Goldscheider, N., Drew, D. (Eds.), Methods in Karst Hydrogeology, Vol. 26. Taylor & Francis, London, UK, pp. 148–170.

Brandt, S., 1999. Data Analysis: Statistical and Computational Methods for Scientists and Engineers, third edition. Springer-Verlag, New York.

Brown, T.L., 2009. Fluorescence characterization of karst aquifers in east Tennessee. Master's thesis, University of Tennessee, Knoxville, Tenn., 151 p.

Burke, S., 2001. Missing values, outliers, robust statistics & non-parametric methods. Statistical data analysis. LC–GC European Online Supplement 59, 19–24. URL. http://www.lcgceurope.com/lcgceurope/data/articlestandard/lcgceurope/502001/4509/article.pdf (accessed 15.07.10).

Chambers, J.M., Cleveland, W.S., Kleiner, B., Tukey, P.A., 1983. Graphical Methods for Data Analysis. Brooks/Cole Publishing Company, Pacific Grove, Calif.

Cowell, F.A., Victoria-Feser, M.P., 1993. Robustness properties of inequality measures: The influence function and the principle of transfers. Tech. Rep. DARP 01. Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD), London School of Economics and Political Science, London, U.K. http://eprints.lse.ac.uk/2141/1/Robustness_Properties_of_Inequality_Measures_The_Influence_Function_and_the_Principle_of_Transfers.pdf (accessed 05.01.11).

Daszykowski, M., Kaczmarek, K., Vander Heyden, Y., Walczak, B., 2007. Robust statistics in data analysis—a review: basic concepts. Chemometrics and Intelligent Laboratory Systems 85, 203–219. doi:10.1016/j.chemolab.2006.06.016.

Davis, S.N., Campbell, D.J., Bentley, H.W., Flynn, T.J., 1985. Ground Water Tracers. National Water Well Association, Worthington, Ohio.

EFSA, 2010. Management of left-censored data in dietary exposure assessment of chemical substances. European Food Safety Authority. EFSA Journal 8 (3), 96. doi:10.2903/j.efsa.2010.1557. URL. http://www.efsa.europa.eu/en/scdocs/doc/1557.pdf (accessed 17.07.10).

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

Field, M.S., 2002. The Qtracer2 Program for Tracer-breakthrough Curve Analysis for Hydrological Tracer Tests. Tech. Rep. EPA/600/R-02/001 and EPA/600/CR-02/001. U.S. Environmental Protection Agency, Washington, D.C. URL. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=54930 (accessed 21.07.10).

Field, M.S., 2003. Tracer-test Planning Using the Efficient Hydrologic Tracer-test Design (EHTD) Program. Tech. Rep. EPA/600/R-03/034 and EPA/600/CR-03/034. U.S. Environmental Protection Agency, Washington, D.C. URL. http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=66073, 175 pp., (accessed 12.01.11).

Ganser, G.H., Hewett, P., 2010. An accurate substitution method for analyzing censored data. Journal of Occupational and Environmental Hygiene 7 (4), 233–244. doi:10.1080/15459621003609713.

Gaspar, E., 1987a. Tracer Hydrology, Vol. I. CRC Press, Boca Rato, Fla.

Gaspar, E., 1987b. Tracer Hydrology, Vol. II. CRC Press, Boca Rato, Fla.

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

Gleit, A., 1985. Estimation for small normal data sets with detection limits. Environmental Science and Technology 19 (12), 1206–1213.

Guan, H., Love, A.J., Simmons, C.T., Hutson, J., Ding, Z., 2010. Catchment conceptualization for examining applicability of chloride mass balance method in an area with historical forest clearance. Hydrology and Earth System Sciences 14, 1233–1245.

Haas, C.H., Scheff, P.A., 1990. Estimation of averages in truncated samples. Environmental Science and Technology 24 (6), 912–919.

Hampel, F.R., 1974. The influence curve and its role in robust estimation. Journal of the American Statistical Association 69 (346), 383–393.

Hampel, F.R., 2001. Robust Statistics: a Brief Introduction and Overview. Tech. Rep. 94. Eidgenössische Technische Hochschule (ETH), CH-8092 Zürich, Switzerland. URL. ftp://129.132.148.131/sfs/pub/Research-Reports/94.pdf (accessed 22.10.10).

Helsel, D.R., 2005. Nondetects and Data Analysis: Statistics for Censored Environmental Data. John Wiley & Sons, New York.

Helsel, D.R., 2006. Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. Chemosphere 65, 2434–2439. doi:10.1016/j.chemosphere.2006.04.051.

Helsel, D.R., Hirsch, R.M., 2002. Statistical Methods in Water Resources, Techniques of Water-resources Investigations of the United States Geological Survey, Book 4, Hydrological Analysis and Interpretation. Tech. Rep. Chapter A3. U.S. Geological Survey, Washington, D.C. URL. http://pubs.usgs.gov/twri/twri4a3/pdf/twri4a3-new.pdf (accessed 22.07.10).

Hewett, P., Ganser, G.H., 2007. A comparison of several methods for analyzing censored data. The Annals of Occupational Hygiene 51 (7), 611–632. doi:10.1093/annhyg/mem045.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Malcolm, I.A., Scoups, G., 2010. Gamma distribution models for transit time estimation in catchments: physical interpretation of parameters and implications for time-variant transit time assessment. Water Resources Research 46 (W10536). doi:10.1029/2010WR009148.

Hubert, M., Debruyne, M., 2009. Breakdown value. Computational Statistics 1 (3), 296–302. doi:10.1002/wics.034. URL. http://onlinelibrary.wiley.com/doi/10.1002/wics.34/pdf (accessed 11.06.10).

Hubert, M., Van der Veeken, S., 2008. Outlier detection for skewed data. Journal of Chemometrics 22 (3/4), 235–246. doi:10.1002/cem.1123.

Irwin, M.E., 2005. Summary statistics in SAS. URL. http://www.markirwin.net/stat135/Lecture/Lecture27.pdf (accessed 19.07.10).

Käß, W., 1998. Tracing Technique in Geohydrology. A.A. Balkema, Rotterdam, The Netherlands.

Kilpatrick, F.A., Wilson Jr., J.F., 1989. Measurement of Time of Travel in Streams by Dye Tracing. Tech. Rep. Techniques of Water-Resources Investigations of the U.S. Geological Survey, Book 3, Chapter A9. U.S. Geological Survey, Washington D.C.

Kroll, C.N., Stedinger, J.R., 1996. Estimation of moments and quantiles using censored data. Water Resources Research 32 (4), 1005–1012.

Leibundgut, C., Małoszewski, P., Külls, C. (Eds.), 2009. Tracers in Hydrology. John Wiley & Sons Ltd., West Sussex, U.K.

Mull, D.S., Liebermann, T.D., Smoot, J.L., Woosley Jr., L.H., 1988. Application of Dye-tracing Techniques for Determining Solute-transport Characteristics of Ground Water in Karst Terranes. Tech. Rep. EPA/904/9-88-001. U.S. Environmental Protection Agency, Region IV, Atlanta, Ga.

Navy, 1999. Handbook for Statistical Analysis of Environmental Background Data. Tech. rep. Department of the Naval, Southwest Division, Naval Facilities Engineering Command, San Diego, Cal. https://portal.navfac.navy.mil/portal/page/portal/navfac/navfac_ww_pp/navfac_nfesc_pp/environmental/erb/lab/hndbk-sw.pdf (accessed 22.08.10).

Olive, D.J., 2008. Applied Robust Statistics. Preprint M-02–006 URL. http://www.math.siu.edu/olive/ol-bookp.htm (accessed 15.07.10).

Ripp, J., 1996. Analytical Detection Limit Guidance & Laboratory Guide for Determining Method Detection Limits. Tech. Rep. PUBL-TS-056–96. Laboratory Certification Program, Wisconsin Department of Natural Resources, Madison, Wis. URL. http://www.dnr.state.wi.us/org/es/science/lc/outreach/-publications/lod (accessed 07.01.11).

Rousseeuw, P.J., Debruyne, M., Engelen, S., Hubert, M., 2006. Robustness and outlier detection in chemometrics. Analytical Chemistry 36 (3), 221–242. doi:10.1080/10408340600969403.

Schudel, B., Biaggi, D., Dervey, T., Kozel, R., Müller, I., Ross, J.H., Schindler, U., 2002. Utilisation des traceurs artificiels en hydrogéologie: Guide pratique. Tech. Rep. BWGG-3-F. Groupe de travail Traçage de la Société suisse d'hydrogéologie (SSH), Berne, Switzerland. URL. http://www.bafu.admin.ch/publikationen/publikation/00439/index.html?lang=fr (accessed 11.01.11).

Singh, A., Nocerino, J., 2001. Robust Estimation of Mean and Variance Using Environmental Data Sets with below Detection Limit Observations. Tech. Repp. 01–062. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Las Vegas, Nev. URL. http://www.epa.gov/esd/tsc/images/robust_estim.pdf (accessed 13.07.10).

Swaroop, R., Winter, W.R., 1971. A Statistical Technique for Computer Identification of Outliers in Multivariate Data. Tech. Rep. NASA TN D-6472. NASA Flight Research Center, Washington, D.C. URL. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19710021895_1971021895.pdf (accessed 20.07.10).

Thompson, M.L., Nelson, K.P., 2003. Linear regression with Type 1 interval-and left-censored response data. Environmental and Ecological Statistics 10 (2), 221–230.

USEPA, 2000. Guidance for Data Quality Assessment: Practical Methods for Data Analysis (EPA QA/G-9, QA00 Update). Tech. Rep. EPA/600/R-96/084. U.S. Environmental Protection Agency, Washington, D.C. URL. http://www.clu-in.org/conf/tio/pasi_121603/g9-final.pdf (accessed 22.10.2010).

USEPA, 2009. Definition and procedure for the determination of the method detection limit—revision 1.11. 40 CFR Part 136, Appendix B. In: Code of Federal Regulations. U.S. Government, Washington, D.C., pp. 343–346. URL. http://edocket.access.gpo.gov/cfr_2009/julqtr/pdf/40cfr136AppB.pdf (accessed 07.01.11).

Wilson Jr., J.F., Cobb, E.D., Kilpatrick, F.A., 1986. Fluormetric Procedures for Dye Tracing. Tech. Rep. In: Techniques of Water-Resources Investigations of the U.S. Geological Survey, Vol. 3. U.S. Geological Survey, Washington D.C. Chapter A12.