United States
Environmental Protection
Agency

Office of Water
(4303)

EPA 821-B-97-008
January 1998

EPA

Statistical Support Document for Proposed Effluent Limitations Guidelines and Standards for Industrial Waste Combustors

Statistical Support Document for Proposed Effluent Limitations Guidelines and Standards for Industrial Waste Combustors

Prepared for:

U.S. Environmental Protection Agency
Office of Water, Engineering and Analysis Division
401 M Street SW
Washington, DC 20460

Prepared by:

Science Applications International Corporation
Environmental and Health Sciences Group
Health and Environment Studies and Systems Division
11251 Roger Bacon Drive
Reston, VA 20190

January 14, 1998

1. INTRODUCTION

This document describes the statistical development of numerical effluent limitations guidelines and standards proposed for the Industrial Waste Combustor Subcategory of the Waste Combustors Point Source Category. Topics include data source identification, data conventions, the modified delta-lognormal distribution, and a procedure for developing percentile estimates from the delta-lognormal distribution.

2. DISTRIBUTIONAL ANALYSES SUPPORTING THE DEVELOPMENT OF NUMERICAL LIMITATIONS FOR BEST AVAILABLE TECHNOLOGY

This chapter describes the statistical analyses that support the development of the proposed effluent limitations guidelines and standards for the Industrial Waste Combustor (IWC) Industry. This chapter also provides an overview of the statistical analyses and describes the sources of data and the modified delta-lognormal distribution that was used to derive the proposed limitations.

2.1 Overview of Effluent Limitations Guidelines and Standards Statistical Analyses

The statistical analyses used in the development of the effluent limitations guidelines and standards are based on the following assumptions: (1) individual effluent measurements for pollutants to be limited are approximately delta-lognormal in probability distribution; (2) on a long-term average basis, good engineering practice will allow appropriately designed and well-operated wastewater treatment systems to perform at least as well as the observed performance of the system whose data were used to develop the limitations; (3) an allowance for the observed process variability will allow for the normal process variation associated with both waste combustion and a well-designed and operated treatment system; and (4) process variation within certain classes of pollutants, such as metals, are approximately equal.

For the two options listed in **Table 2-1**, EPA developed potential limitations for the following pollutants:

- 4 Classical Pollutants [Carbon Oxygen Demand (COD), Total Dissolved Solids (TDS), Total Organic Carbon (TOC), Total Suspended Solids (TSS)]
- 17 Metals (Aluminum, Antimony, Arsenic, Boron, Cadmium, Chromium, Copper, Iron, Lead, Manganese, Mercury, Molybdenum, Selenium, Silver, Tin, Titanium, and Zinc)

Table 2-1.
IWC Effluent Limitations Guidelines and Standards Options

Option	Technology
A	Primary Precipitation Liquid/Solid Separation Sludge Dewatering Secondary Precipitation Liquid/Solid Separation Sludge Dewatering
B	Option A+ Sand Filtration

2.2 Data Sources

A listing of the data used to support limitations development is included as Appendix A. The data used to calculate the proposed limitations for Options A and B were derived from the EPA Sample Control Center (SCC) physical sampling database. This database contains the measurement results of intensive sampling efforts at 17 sites between 1993 and 1995.

2.3 Description of Data Conventions

This section describes the types of data in the IWC analytical database and the procedures for data aggregation.

2.3.1 Data Review

The analytical sampling data in the SCC database were thoroughly reviewed by EPA. During this review, the integrity of each sample was assessed to ensure that all specifications of the sampling protocol were met. The reviewers determined that some samples should be excluded from the analyses. These samples were flagged in the database in a field labeled "SCC Qualifier" (see Appendix A). Samples with flags of "EXCLUDE" or "DETECTED" (a value was detected but the concentration value was not recorded) were set to missing values.

An engineering review of the database was also conducted and a few additional data values were excluded from the analyses for the reasons summarized in the record for the proposed rule-making.

2.3.2 Data Types

The IWC analytical database (from the SCC) contains the following two different types of samples delineated by certain qualifiers in the database:

- **Non-censored (NC):** a measured value, i.e., a sample measured above the level at which the detection decision was made.

- **Non-detect (ND):** samples for which analytical measurement did not yield a concentration above the sample-specific level at which the detection decision was made. For these samples, the level associated with the detection decision is reported.

Depending on the pollutant, the decision to call a measurement NC or ND was made either at the Minimum Level (ML) or at the Instrument Detection Limit (IDL). For all metals, the detection decision was made at the IDL. The IDL "refers to the smallest signal above background noise that an instrument can detect reliably."[1] The ML refers to the "lowest acceptable calibration point."[2] The term *detection limitation* will be used in the following text where it is possible to use either the IDL or the ML. The phrase *detection decision* indicates a choice between reporting a measurement result or treating a measurement result in the same fashion as when the analyte is not present.

Non-detected values were used as reported in analyzing the data.

---

[1]Keith, L.H., W. Crummett, J. Deegan, R.A. Libby, J.K. Taylor, G. Wentler (1983). "Principles of Environmental Analysis," *Analytical Chemistry*, Volume 55, Pages 2210-2218.

[2]U.S. EPA (1980). *Method 1624, Volatile Organic Compounds by Purge and Trap Isotope Dilution GCMS*, EPA Effluent Guidelines Division (WH-552), Washington, DC 20460.

2.3.3 Data Aggregation

Data aggregation for the IWC analytical data was performed due to the identification of field duplicates within the data. Field duplicates are defined as one or more samples collected for a particular sampling point at approximately the same time, assigned different sample numbers, and flagged as duplicates for a single episode number.

Data aggregation was performed for field duplicates. When all of the duplicates in a set were non-censored, detected samples, the arithmetic average of the duplicates was straightforward. However, when one or more of the duplicates was censored (that is, non-detect), the following methods were used to control the combination. (Note that the value of ND is the instrument detection limit for the non-detected sample and the value of NC is the observed concentration value.) **Table 2-2** outlines the methods for combining field duplicate samples for the statistical analyses.

Table 2-2.
Method for Averaging Field Duplicate Samples

If observations are:	Label of "aggregate"	Value of "aggregate" is:
Both ND values	ND	Maximum (ND_1, ND_2)
NC and ND with NC value > ND value	NC	(NC+ ND)/2
NC and ND with NC value ≤ ND value	ND	ND value
Both NC values	NC	$(NC_1+NC_2)/2$

NC = non-censored values
ND = non-detected values

2.4 Statistical Methodology

2.4.1 Overview of Methodology and Applicability to the IWC Effluent Limitations Guidelines and Standards Database

2.4.1.1 Basic Overview of Delta-lognormal Distribution

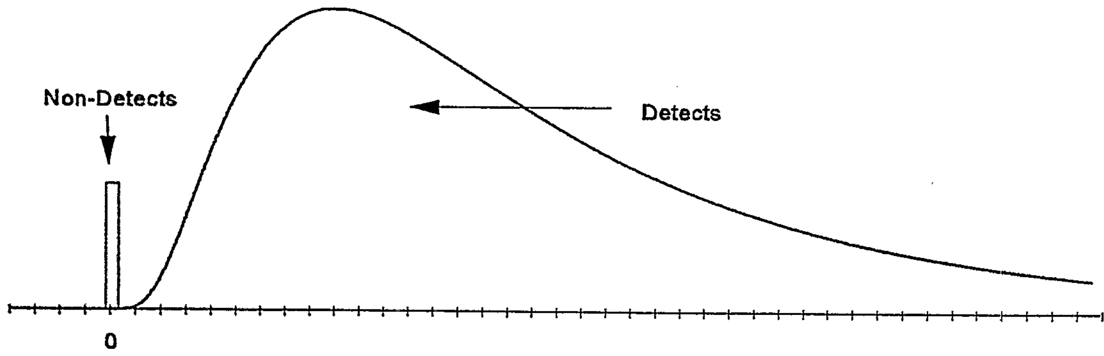
The classical delta-lognormal model is displayed in **Figure 2-1**. In this adaptation of the simple lognormal density (see Crow and Shimizu[3]), the model is expanded to include zero amounts by grouping together all positive amounts and fitting them to a lognormal density. All zero amounts are then segregated into another group of measurements representing a discrete distributional "spike" at zero. The resulting mixed distribution,

---

[3]Crow, E.L., Shimizu, K. (1988) *Lognormal Distributions: Theory and Applications,* Marcel Dekker, Inc., New York, NY 10016.

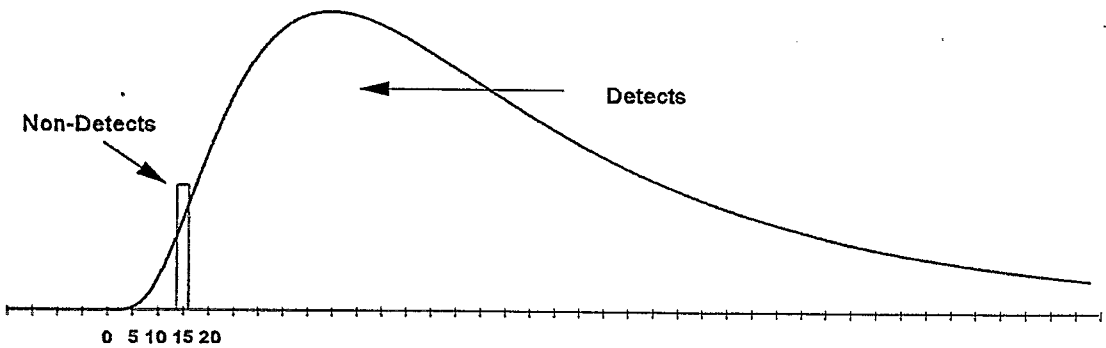
combining a continuous density portion with a discrete-valued spike, is known as the delta-lognormal distribution. The delta in the name refers to the percentage of the overall distribution contained in the spike at zero, that is, the percentage of zero amounts.

Figure 2-1.



Kahn and Rubin[4], 1989, further adapted the classical delta-lognormal model ("adapted model") to account for non-detect measurements in the same fashion that zero measurements were handled in the original delta-lognormal. The actual values of non-detects are not known, though each non-detect is assumed to have a concentration somewhere between zero and the reported detection limit. Instead of zero amounts and non-zero (positive) amounts, the data consisted of non-detects and detects. Rather than assuming that non-detects represented a spike of zero concentration, these samples were allowed to have a single positive value, usually equal to the level at which the detection decision is made (see **Figure 2-2**). Since each non-detect was assigned the same positive value, the distributional spike in this adapted model was located not at zero, but at the detection limitation. This adapted model was used in developing limitations for the Organic Chemicals, Plastics, and Synthetic Fibers (OCPSF) and pesticides manufacturing rulemaking.

Figure 2-2.



---

[4]Kahn, H.D., Rubin, M.B. (1989). "Use of Statistical Methods in Industrial Water Pollution Control Regulations in the United States," Environmental Monitoring and Assessment, Volume 12, Page 129-148.

In the adapted delta-lognormal model, the delta again referred to those measurements contained in the discrete spike, this time representing the proportion of non-detect values observed within the data set. By using this approach, computation of estimates for the population mean and variance could be done easily by hand, and non-detects were not assumed to follow the same distributional pattern as the detected measurements. The adapted delta-lognormal model can be expressed mathematically as:

$$Pr\ (U \le u) = \begin{cases} (1-\delta)\ \Phi\left[(\log(u) - \mu)/\sigma\right] & if \quad 0 < u < D \\ \delta + (1-\delta)\ \Phi\left[(\log(D) - \mu)/\sigma\right] & if \quad u = D \\ \delta + (1-\delta)\ \Phi\left[(\log(u) - \mu)/\sigma\right] & if \quad u > D \end{cases} \tag{2.1}$$

where δ represents the true proportion of non-detects (or the probability that any randomly drawn measurement will be a non-detect), D equals the Minimum Level value of the discrete spike assigned to all non-detects, $\Phi(\cdot)$ represents the standard normal cumulative distribution function, and μ and σ are the parameters of the lognormal density portion of the model. This model assumes that all non-detected values have a single detection limit D.

It is also possible to represent the adapted delta-lognormal model in another mathematical form; one in which it is particularly easy to derive formulas for the expected value (i.e., long-term average [LTA]) and variance of the model. In this case, a random variable distributed according to the adapted delta-lognormal distribution can be represented as the stochastic combination of three other independent random variables. The first of these variables is an indicator variable, I_u, equal to one when the measurement u is a non-detect and equal to zero when u is a detected value. The second variable, X_D, represents the value of a non-detect measurement (discrete). In the adapted delta-lognormal, this variable is always a constant ND equal to the concentration value assigned to each non-detect (i.e., equal to D in the adapted delta-lognormal model). In general, however, X_D need not be a constant, as will be seen below in the modified delta-lognormal model. The final random variable, X_C, represents the value of a detected measurement, and is distributed according to a lognormal distribution (continuous) with parameters μ and σ^2.

Using this formulation, a random variable from the adapted delta-lognormal model can be written as:

$$U = I_u X_D + (1-I_u)X_C \tag{2.2}$$

and the expected value of U is then derived by substituting the expected value of each quantity in the right-hand side of the equation. Because the variables I_u, X_D, and X_C are mutually independent, this leads to the expression

$$E(U) = \delta E(X_D) + (1-\delta)E(X_C) = \delta D + (1-\delta)\exp(\mu + 0.5\sigma^2) \tag{2.3}$$

where again δ is the probability that any random measurement will be non-detect and the exponentiated expression is the familiar mean of a lognormal distribution. In a similar fashion, the variance of the adapted delta-lognormal model can be established by squaring the expression for U above, taking expectations, and subtracting the square of E(U) to get:

$$Var(U) = E(U^2) - [E(U)]^2 = \delta Var(X_D) + (1-\delta)Var(X_C) + \delta(1-\delta)[E(X_D) - E(X_C)]^2. \tag{2.4}$$

Since, in the adapted delta-lognormal formulation, X_D is a constant, this expression can be reduced to the following:

$$Var(U) = (1-\delta)\exp(2\mu+\sigma^2)[\exp(\sigma^2)-(1-\delta)] + \delta(1-\delta)D[D-2\exp(\mu+0.5\sigma^2)]. \qquad \textbf{(2.5)}$$

In order to estimate the adapted delta-lognormal mean and variance from a set of observed sample measurements, it is necessary to derive sample estimates for the parameters δ, μ, and σ. δ is typically estimated by the observed proportion of non-detects in the data set. μ and σ are estimated using the logged values of the detected samples where μ is estimated using the arithmetic mean of the logged detected measurements and σ is estimated using the standard deviation of these same logged values. Non-detects are not included in the calculations. Once the param
eter estimates are obtained, they are used in the formulas above to derive the estimated adapted delta-lognormal mean and variance.

To calculate effluent limitations, it is also necessary to estimate upper percentiles from the underlying data model. Using the delta-lognormal formulation above in equation (2.1), letting U_α represent the $100*\alpha^{th}$ percentile of random variable U, and adopting the standard notation of z_s for the s^{th} percentile of the standard normal distribution, an arbitrary delta-lognormal percentile can be expressed as the following:

$$U_\alpha = \begin{cases} \exp(\mu+\sigma\, z_{\alpha/1-\delta}) & \text{if} \quad (1-\delta)\Phi((\log(D)-\mu)/\sigma) \geq \alpha \\ D & \text{if} \quad \delta+(1-\delta)\Phi((\log(D)-\mu)/\sigma) \geq \alpha \\ \exp(\mu+\sigma\, z_{\alpha-\delta/1-\delta}) & \text{if} \quad \delta+(1-\delta)\Phi((\log(D)-\mu)/\sigma) < \alpha \end{cases} \qquad \textbf{(2.6)}$$

The daily maximum limitations are established on the basis of an estimated upper 99^{th} percentile from the underlying data model, so that 0.99 would be substituted for α in the above expression. To derive the daily variability factor (VF) for the 99^{th} percentile based on the adapted delta-lognormal model, divide $U_{.99}$ in the expression above by the previous formula for the LTA, namely $U_{.99}/E(U)$.

2.4.1.2 Motivations for Modifications to the Adapted Delta-Lognormal Model

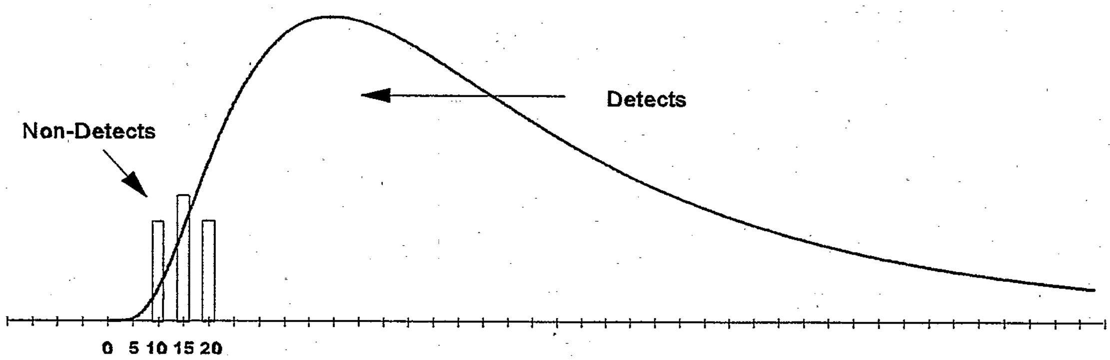
While the adapted delta-lognormal model has been used successfully for years by EPA in a variety of settings, the model makes two key assumptions about the observed data that are not fully satisfied within the IWC analytical database. First, the discrete spike portion of the adapted delta-lognormal model is a fixed, single-valued probability mass associated (typically) with all the non-detect measurements. If all non-detect samples in the IWC database had roughly the same reported detection limit, this assumption would be adequately satisfied. However, reported detection limits in the IWC analytical data vary. Because of this variation in detection limits, a single-valued discrete spike may not adequately represent the set of non-detect measurements observed in the IWC database and a modification to the model was considered.

In addition, the adapted delta-lognormal model sets all non-censored values below the detection limitation (D) to the Minimum Level of the analytical method. (For example, if the Minimum Level for Aluminum was 10 ppq, then any non-censored samples reported below 10 ppq were set to 10 ppq.) Several instances occurred in the IWC analytical data where a non-censored value was reported below the Minimum Level of the analytical method, though this did not occur with any of the data used to develop the proposed effluent limitations.

2.4.1.3 Modification of the Discrete Spike

To appropriately modify the adapted delta-lognormal model for the observed IWC database, a modification was made to the discrete, single-valued spike representing non-detect measurements. Because non-detect samples have varying detection limits, the spike of the delta-lognormal model has been replaced by a discrete distribution made up of multiple spikes. Each spike in this modification is associated with a distinct detection limit observed in the IWC database. Thus, instead of assigning all non-detects to a single, fixed value, as in the adapted model, non-detects can be associated with multiple values depending on how the detection limits vary, as seen in **Figure 2-3.**

Figure 2-3.



Hence, the discrete "delta" portion of the modified model is estimated in a way similar to the adapted delta-lognormal distribution, only now multiple spikes are constructed, linked to the distinct detection limits observed in the data set. In the adapted model, the parameter δ is estimated by computing the proportion of non-detects. In the modified model, δ again represents the proportion of non-detects, but is divided into the sum of smaller fractions, δ_i, each representing the proportion of non-detects associated with a particular and distinct detection limit. Thus it can be written:

$$\delta = \sum_i (\delta_i). \quad \textbf{(2.7)}$$

If D_i equals the value of the i^{th} smallest distinct detection limit in the data set, and let the random variable X represent a randomly chosen non-detect sample, then the discrete distribution portion of the modified delta-lognormal model can be mathematically expressed as:

$$Pr(X_D \le x) = \sum_{i:\, D_i \le x} \delta_i. \quad \textbf{(2.8)}$$

The mean and variance of this discrete distribution, unlike the discrete spike of the adapted delta-lognormal, can be computed such that the variance of the modified spike is non-zero using the following formulas:

$$E(X_D) = \frac{1}{\delta}\sum_i \delta_i D_i \quad \text{and} \quad Var(X_D) = \frac{1}{\delta^2}\sum_i \sum_{i<j} \delta_i \delta_j (D_j - D_i)^2. \tag{2.9}$$

It is important to recognize that, while replacing the single discrete spike in the adapted delta-lognormal distribution with a more general discrete distribution of multiple spikes increases the complexity of the model, the discrete portion with multiple spikes plays a role in limitations development identically parallel to the single spike case and offers flexibility for handling multiple observed detection limits.

2.4.2 Estimation Under the Modified Delta-Lognormal Model

Once the basic modification to the adapted delta-lognormal distribution is made, it is possible to fit a wide variety of observed effluent data sets to the modified model. Multiple detection limits for non-detects may now be handled. The same basic framework can be used even if there are no non-detect values.

Combining the discrete portion of the model with the continuous portion, the cumulative probability distribution of the modified delta-lognormal model can be expressed as follows, where D_n denotes the largest distinct detection limit observed among the non-detects, and where the first summation is taken over all those values D_i that are less than u.

$$Pr(U \le u) = \begin{cases} \sum_{i:D_i \le u} \delta_i + (1-\delta)\Phi\left[(\log(u) - \mu)/\sigma\right] & \text{if} \quad u < D_n \\ \delta + (1-\delta)\Phi\left[(\log(u) - \mu)/\sigma\right] & \text{if} \quad u \ge D_n \end{cases} \tag{2.10}$$

Again combining the discrete and continuous portions of the modified model, the expected value of the random variable U can be derived as a weighted sum of the expected values of the discrete and continuous lognormal portions of the distribution. This follows because the modified delta-lognormal random variable U can be expressed again as a combination of three other independent variables, that is,

$$U = I_u X_D + (1 - I_u) X_C \tag{2.11}$$

where, this time, X_D represents a random non-detect from the discrete portion of the model, X_C represents a random detected measurement from the continuous lognormal portion, and I_u is an indicator variable signaling whether any particular random measurement is detected or not. Then the expected value and variance of U have forms somewhat similar to the standard delta-lognormal model, namely

$$E(U) = \sum_i \delta_i D_i + (1 - \delta)\exp(\mu + 0.5\,\sigma^2) \tag{2.12}$$

$$Var(U) = \frac{\sum_i \sum_{k<j} \delta_i \delta_j (D_i - D_j)^2}{\delta} + (1-\delta)\exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1) + \delta(1-\delta)\left[\frac{\sum_i \delta_i D_i}{\delta} - \exp(\mu + 0.5\sigma^2)\right]^2 \qquad \textbf{(2.13)}$$

where the D_i equals individual detection limits for the non-detects, the δ_i are the corresponding proportions of not detected values with detection limit D_i, and $\delta = \Sigma\delta_i$.

2.4.2.1 Estimation of Long-Term Averages

Long-term averages were calculated for each sampling episode sample location separately. For the purposes of estimating these long-term averages (equal to the expected value in the equation (2.14)),it was necessary to divide the IWC data sets into two groups based on their size (number of samples) and the type of samples in the subset. Thus, the computations differed for each group:

Group 1: Less than 2 detected samples (NC) or less than 4 total samples.

Group 2: Two or more non-censored samples (NC) and 4 or more total samples.

For Group 1, the long-term averages were calculated as the arithmetic average of the samples, since the sample sizes for either the discrete portion or the continuous lognormal portion of the data were too small to allow distributional assumptions to be made. Specifically, Group 1 contained all data subsets with all non-detects or only one detect. Detection limits were substituted as the values associated with non-detectable samples.

For Group 2, the long-term averages were calculated using the formula for E(U) in equation (2.14). μ and σ parameters were estimated as the mean and variance of the logged NC values.

Appendix B presents summary statistics by analyte for each option and sampling episode combination.

2.4.2.2 Estimation of Variability Factors, Percentiles, and Limitations

After determining estimated long-term average values for each pollutant for each sample point location, EPA developed 1-day variability factors (VF1) for each pollutant and either 4-day or 20-day monthly average variability factors (VF4 and VF20) dependent on the assumed frequency of monitoring as outlined in **Table 2-3.** Appendix C presents estimated daily maximum limitations, monthly average limitations, and the associated variability factors for each option that are calculated using pollutant-specific variability factors. The estimation methodology is presented below.

Table 2-3.
Assumed Monitoring Frequencies

Pollutant Category	Frequency of Monitoring
Metals	Monthly (VF1, VF4)
COD	Monthly (VF1, VF4)
TDS	Weekly (VF1, VF4)
TSS	Daily (VF1, VF20)

Similar to the calculations for the long-term averages, the data were divided into the same two computation groups based on the number and type of samples in each data subset:

Group 1: Less than 2 detected (NC) samples or less than 4 total samples. Upper percentiles and variability factors could not be computed using the modified delta-lognormal methodology.

Group 2: Two or more non-censored samples (NC) and 4 or more total samples. The estimates of the parameters for the modified delta-lognormal distribution of the data were calculated empirically in the log-domain. Upper percentiles and variability factors were calculated using these estimated parameters.

2.4.2.2.1 Estimation of Facility-Specific 1-Day Variability Factors and 99th Percentiles

The 1-day variability factors are a function of the long-term average, E(U), and the 99th percentile. An iterative approach was used in finding the 99th percentile of each data subset using the modified delta-lognormal methodology by first defining $D_0=0$, $\delta_0=0$, and $D_{k+1}=\infty$ as boundary conditions where D_i equals half of the i^{th} smallest detection limit, and δ_i is the associated proportion of non-detects at the i^{th} detection limit. A cumulative distribution function, p, for each data subset was computed as a step function ranging from 0 to 1. The general form, for a given value c, is:

$$p = \sum_{i=0}^{m}\hat{\delta}_i + (1-\hat{\delta})\,\Phi\left[\frac{\log(c)-\hat{\mu}}{\hat{\sigma}}\right], \quad D_m \le c < D_{m+1}, \; m=0,1,\ldots k \qquad (2.14)$$

where Φ is the standard normal cumulative distribution function. The following steps were completed to compute the estimated 99th percentile of each data subset:

1. k values of p at $c=D_m$, $m=1,\ldots k$ were computed and labeled p_m.

2. The smallest value of m, such that $p_m \ge 0.99$, was determined and labeled as p_j. If no such m existed, steps 3 and 4 were skipped and step 5 was computed instead.

3. Computed $p^* = p_j - \delta_j$.

4. If $p^* < 0.99$, then $P_{99} = D_j$,
   else if $p^* \geq 0.99$, then

$$\hat{P}99 = \exp\left[\hat{\mu} + \Phi^{-1}\left[\frac{\left(0.99 - \sum_{i=0}^{j-1}\hat{\delta}_i\right)}{(1-\hat{\delta})}\right]\hat{\sigma}\right] \qquad \textbf{(2.15)}$$

5. If no such m exists, such that $p_m \geq 0.99$ (m=1,...k), then

$$\hat{P}99 = \exp\left[\hat{\mu} + \Phi^{-1}\left[\frac{0.99-\hat{\delta}}{(1-\hat{\delta})}\right]\hat{\sigma}\right] \qquad \textbf{(2.16)}$$

The daily variability factor, VF1, was then calculated as

$$VF1 = \frac{\hat{P}99}{\hat{E}(U)} \qquad \textbf{(2.17)}$$

Appendix C displays long-term averages and 1-day variability factors by analyte for each option and sampling episode combination.

2.4.2.2.2 Estimation of Facility-Specific 4-Day Variability Factors and 95th Percentiles of 4-Day Means

For all but TSS, it was necessary to calculate a variability factor for monthly averages based on the distribution of 4-day averages, because EPA is considering proposing that these pollutants be monitored weekly (approximately four times a month). In order to calculate the 4-day variability factor (VF4), the assumption was made that the approximating distribution of $\bar{U}_4$, the sample mean for a random sample of four independent concentrations, is also derived from this modified delta-lognormal distribution, with the same mean as the distribution of the concentrations. The mean of this distribution of 4-day averages is:

$$E(\bar{U}_4) = \delta_4 E(\bar{X}_4)_D + (1-\delta_4)E(\bar{X}_4)_C \qquad \textbf{(2.18)}$$

where $(X_4)_D$ denotes the mean of the discrete portion of the distribution of the average of four independent concentrations, (i.e., when all observations are not detected) and $(X_4)_C$ denotes the mean of the continuous lognormal portion of the distribution.

First, it is assumed that the probability of detection (δ) on each of the 4 days is independent of that on the other days, since the samples to be used for compliance monitoring are not taken on consecutive days, and no correlation is expected to exist such that $\delta_4 = \delta^4$.

Also, since $E(\bar{X}_4)_D = E(X_D)$ then

$$E(\bar{U}_4) = \delta^4 \sum_{i=1}^{k} \frac{\delta_i D_i}{\delta} + (1-\delta^4)\exp(\mu_4 + 0.5\sigma^2_4) \tag{2.19}$$

and since $E(\bar{U}_4) = E(U)$, then

$$\mu_4 = \log\left[\frac{E(U) - \delta^3 \sum_{i=1}^{k} \delta_i D_i}{(1-\delta^4)}\right] - 0.5\sigma^2_4. \tag{2.20}$$

The expression for σ^2_4 was derived from the following relationship:

$$Var(\bar{U}_4) = \delta_4 Var((\bar{X}_4)_D) + (1-\delta_4)Var((\bar{X}_4)_C) + \delta_4(1-\delta_4)[E(\bar{X}_4)_D - E(\bar{X}_4)_C]^2. \tag{2.21}$$

Since

$$Var((\bar{X}_4)_D) = \frac{Var(X_D)}{4}, \quad E(\bar{X}_4)_D = E(X_D), \quad \text{and} \quad \delta_4 = \delta^4 \tag{2.22}$$

then,

$$Var(\bar{U}_4) = \delta^4 \frac{Var(X_D)}{4} + (1-\delta^4)Var((\bar{X}_4)_C) + \delta^4(1-\delta^4)[E(X_D) - E(\bar{X}_4)_C]^2. \tag{2.23}$$

This further simplifies to:

$$\begin{aligned} Var(\bar{U}_4) = {} & \frac{\delta^4 \sum_{i=1}^{k} \sum_{i<j, j=1}^{k} \delta_i \delta_j (D_i - D_j)^2}{4\delta^2} + (1-\delta^4)\exp(2\mu_4 + \sigma^2_4)[\exp(\sigma^2_4) - 1] \\ & + \delta^4(1-\delta^4)\left[\sum_{i=1}^{k} \frac{\delta_i D_i}{\delta} - \exp(\mu_4 + 0.5\sigma^2_4)\right]^2 \end{aligned} \tag{2.24}$$

and furthermore

$$\exp(\sigma^2_4) - 1 = \frac{\left[Var(\bar{U}_4) - \frac{\delta^2 \sum_{i=1}^{k} \sum_{i<j, j=1}^{k} \delta_i \delta_j (D_i - D_j)^2}{4} - \delta^2(1-\delta^4)\left[\sum_{i=1}^{k} \delta_i D_i - \delta\exp(\mu_4 + 0.5\sigma^2_4)\right]^2\right]}{(1-\delta_4)\exp(2\mu_4 + \sigma^2_4)}. \tag{2.25}$$

Then from (2.21) above,

$$\exp(\mu_4 + 0.5\sigma^2_4) = \frac{(E(\bar{U}_4) - \delta^3 \sum_{i=1}^{k} \delta_i D_i)}{(1 - \delta^4)} = \frac{(E(U) - \delta^3 \sum_{i=1}^{k} \delta_i D_i)}{(1 - \delta^4)}, \quad \text{since } E(\bar{U}_4) = E(U) \tag{2.26}$$

and letting

$$\eta = E(U) - \delta^3 \sum_{i=1}^{k} \delta_i D_i, \quad \text{then,} \quad \exp(\mu_4 + 0.5\sigma^2_4) = \frac{\eta}{(1 - \delta^4)}. \tag{2.27}$$

Furthermore,

$$\sigma^2_4 = \log\left[1 + \frac{\left[Var(\bar{U}_4) - \frac{\delta^2 \sum_{i=1}^{k} \sum_{i<j=1}^{k} \delta_i \delta_j (D_i - D_j)^2}{4} - \delta^2(1 - \delta^4)\left(\sum_{i=1}^{k} \delta_i D_i - \frac{\delta\eta}{(1 - \delta^4)}\right)^2\right]}{\frac{(1 - \delta^4)\eta^2}{(1 - \delta^4)^2}}\right]. \tag{2.28}$$

Since $Var(\bar{U}_4) = Var(U)/4$ and by rearranging terms,

$$\sigma^2_4 = \log\left[1 + \frac{(1 - \delta^4)\, Var(U)}{4\eta^2} - \frac{(1 - \delta^4)\delta^2 \sum_{i=1}^{k} \sum_{i<j=1}^{k} \delta_i \delta_j (D_i - D_j)^2}{4\eta^2} - \frac{\delta^2\left[\sum_{i=1}^{k} \delta_i D_i (1 - \delta^4) - \delta\eta\right]^2}{\eta^2}\right]. \tag{2.29}$$

Thus, estimates of μ_4 and σ_4 were derived by using estimates of $\delta_1, \ldots \delta_k$ (sample proportion of non-detects at observed detection limits $D_1, \ldots D_k$), μ (mean of logged values), and σ^2 (MLE log variance) in the equations above.

In finding the estimated 95th percentile of the average of four observations, four non-detects (not all at the same detection limit) can generate an average that is not necessarily equal to $D_1, D_2, \ldots,$ or D_k. Consequently more than k discrete points exist in the distribution of the 4-day averages. For example, the average of four non-detects at k=2 detection limits, are at the following discrete points with the associated probabilities:

i	D^*_i	δ^*_i
1	D_1	δ_1^4
2	$(3D_1 + D_2)/4$	$4\delta_1^3\delta_2$
3	$(2D_1 + 2D_2)/4$	$6\delta_1^2\delta_2^2$
4	$(D_1 + 3D_2)/4$	$4\delta_1\delta_2^3$
5	D_2	δ_2^4

In general, when all four observations are not detected, and when k detection limits exist, the multinomial distribution can be used to determine associated probabilities, that is,

$$Pr\left[\bar{U}_4 = \frac{\sum_{i=1}^{k} u_i D_i}{4}\right] = \frac{4!}{u_1!\,u_2!\dots u_k!}\prod_{i=1}^{k}\delta_i^{u_i}. \qquad (2.30)$$

The number of possible discrete points, k^*, for k=1,2,3,4, and 5 are given below:

k	k^*
1	1
2	5
3	15
4	35
5	70

To find the estimated 95th percentile of the distribution of the average of four observations, the same basic steps (described in Section 2.4.2.2.1) as used for the 99th percentile of the distribution of daily observations were followed with the following changes:

1. Change P_{99} to P_{95}, and 0.99 to 0.95.
2. Change D_m to D_m^*, the weighted averages of the detection limits.
3. Change δ_i to δ_i^*.
4. Change k to k^*, the number of possible discrete points based on k detection limits.
5. Change the estimates of δ, μ, and σ to estimates of δ^4, μ_4, and σ_4, respectively.

Then, the estimate of the 95th percentile 4-day mean variability factor is:

$$VF4 = \frac{\hat{P}95}{\hat{E}(U)}, \qquad since \quad E(\bar{U}_4) = E(U). \qquad (2.31)$$

Appendix C displays long-term averages and 4-day variability factors by analyte for each option and sampling episode combination.

2.4.2.2.3 Estimation of Facility-Specific 20-Day Variability Factors and 95th Percentiles of 20-Day Means

Since TSS is proposed to be monitored daily, the monthly average limitation was based on 20 days of sampling. However, the data used to calculate the 20-day variability factors for TSS cover only 5 days of sampling of daily measurements. Therefore, at this time EPA does not have sufficient data to examine in detail and incorporate any autocorrelation between concentrations for TSS measured on adjacent days. The autocorrelation of TSS is further discussed in the preamble to the proposed regulation.

It is assumed that the concentrations for TSS are independent of one another, and

$$E(\bar{U}_{20}) = E(U) \qquad and \qquad V(\bar{U}_{20}) = \frac{V(U)}{20} \qquad (2.32)$$

where E(U) and V(U) are calculated as in equations (2.3) and (2.4). Finally, since $\bar{U}_{20}$ is approximately normally distributed by the Central Limit Theorem, the estimate of the 95th percentile of a 20-day mean and the corresponding 20-day average variability factor (VF20) are approximately

$$\hat{P}95_{20} = \hat{E}(\bar{U}_{20}) + \Phi^{-1}(0.95) * (\hat{V}(\bar{U}_{20}))^{\frac{1}{2}} \qquad \textbf{(2.33)}$$

and

$$VF20 = \frac{\hat{P}95_{20}}{E(\bar{U}_{20})} = \frac{\hat{P}95_{20}}{E(U)} \qquad \textbf{(2.34)}$$

where $\Phi^{-1}(0.95)$ is the 95th quantile of the standard normal distribution.

As noted in **Table 2-3**, EPA assumed 20-day variability factors for TSS. See Appendix C for the TSS 20-day facility-specific variability factors.

3. Estimation of Pollutant-Specific and Group-Level Variability Factors Resulting in Proposed Daily Maximum and Monthly Average Numerical Limitations

This chapter describes the estimation of variability factors by pollutant ("pollutant-specific") and by group ("group-level"). Each group contained pollutants that were chemically similar. The pollutant-specific and group-level variability factors were then used to develop limitations.

3.1 Estimation of Pollutant-Specific Variability Factors

After the facility-specific variability factors were estimated for a pollutant, the pollutant-specific variability factor was calculated. The pollutant-specific daily variability factor was the mean of the facility-specific daily variability factors for that pollutant in the option. Likewise, the pollutant-specific monthly variability factor was the mean of the facility-specific monthly variability factors for that pollutant in the option. Appendix D displays the pollutant-specific long-term averages and variability factors calculated as described above.

3.2 Estimation of Group-Level Variability Factors

After the pollutant-specific variability factors were estimated as described in section 2.1, group-level variability factors were calculated for metals. These metal pollutants were considered to be chemically similar.

The group-level daily variability factor was the median of the pollutant-specific daily variability factors for the pollutants within the group. Similarly, for the monthly variability factors, the group-level monthly variability factor was the median of the pollutant-specific monthly variability factors for the pollutants within the group. Appendix E displays the group-level long-term averages and variability factors calculated as described above.

3.3 Estimation of Potential Daily Maximum and Monthly Average Limitations

For metals, potential daily maximum and monthly average limitations for each pollutant within each option were set equal to the product of the pollutant-specific long-term average and the option group-level variability factor. Appendix F presents potential daily maximum and monthly average limitations for each option that are calculated using group-level variability factors.

APPENDIX A

RAW DATA LISTINGS

Appendix A. Raw Data Listings

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
A	Classicals	CHEMICAL OXYGEN DEMAND (COD)	C-004	410.1	UG/L	4646	05	26507	09/20/94		918,000.0	
							05	26515	09/21/94		197,000.0	
							05	26523	09/22/94		36,000.0	
							D05	26524	09/22/94		28,000.0	
							05	26533	09/23/94		29,000.0	
							05	26540	09/24/94		18,000.0	
A	Classicals	TOTAL DISSOLVED SOLIDS	C-010	160.1	UG/L	4646	05	26507	09/20/94		36,800,000.0	
							05	26515	09/21/94		36,300,000.0	
							05	26523	09/22/94		17,100,000.0	
							D05	26524	09/22/94		37,800,000.0	
							05	26533	09/23/94		42,800,000.0	
							05	26540	09/24/94		41,200,000.0	
A	Classicals	TOTAL ORGANIC CARBON (TOC)	C-012	415.1	UG/L	4646	05	26507	09/20/94			10,000.0
							05	26515	09/21/94			10,000.0
							05	26523	09/22/94			10,000.0
							D05	26524	09/22/94			10,000.0
							05	26533	09/23/94			10,000.0
							05	26540	09/24/94			10,000.0
A	Classicals	TOTAL SUSPENDED SOLIDS	C-009	160.2	UG/L	4646	05	26507	09/20/94		22,000.0	
							05	26515	09/21/94		6,000.0	
							05	26523	09/22/94		21,000.0	
							D05	26524	09/22/94		27,000.0	
							05	26533	09/23/94			4,000.0
							05	26540	09/24/94		11,000.0	
A	Metals	ALUMINUM	7429905	1620	UG/L	4646	05	26507	09/20/94		’267.0	
							05	26515	09/21/94		214.0	
							05	26523	09/22/94		201.0	
							D05	26524	09/22/94	B	187.0	
							05	26533	09/23/94	B	143.0	
							05	26540	09/24/94	B	165.0	

Appendix A. Raw Data Listings

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
A	Metals	ANTIMONY	7440360	1620	UG/L	4646	05	26507	09/20/94		353.0	
							05	26515	09/21/94		381.0	
							05	26523	09/22/94		303.0	
							D05	26524	09/22/94		339.0	
							05	26533	09/23/94		394.0	
							05	26540	09/24/94		456.0	
A	Metals	ARSENIC	7440382	1620	UG/L	4646	05	26507	09/20/94	B	9.80	
							05	26515	09/21/94	B	9.00	
							05	26523	09/22/94			20.0
							D05	26524	09/22/94			20.0
							05	26533	09/23/94	B	1.60	
							05	26540	09/24/94	B	3.80	
A	Metals	BORON	7440428	1620	UG/L	4646	05	26507	09/20/94		1,960.0	
							05	26515	09/21/94		1,640.0	
							05	26523	09/22/94		1,560.0	
							D05	26524	09/22/94		1,490.0	
							05	26533	09/23/94		1,670.0	
							05	26540	09/24/94		1,730.0	
A	Metals	CADMIUM	7440439	1620	UG/L	4646	05	26507	09/20/94		65.4	
							05	26515	09/21/94		95.5	
							05	26523	09/22/94		62.8	
							D05	26524	09/22/94		59.7	
							05	26533	09/23/94		8.60	
							05	26540	09/24/94			5.00
A	Metals	CHROMIUM	7440473	1620	UG/L	4646	05	26507	09/20/94			10.0
							05	26515	09/21/94			10.0
							05	26523	09/22/94			10.0
							D05	26524	09/22/94			10.0
							05	26533	09/23/94			10.0
							05	26540	09/24/94			10.0

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
A	Metals	COPPER	7440508	1620	UG/L	4646	05	26507	09/20/94		30.1	
							05	26515	09/21/94	B	12.5	
							05	26523	09/22/94		32.3	
							D05	26524	09/22/94		31.2	
							05	26533	09/23/94	B	10.6	
							05	26540	09/24/94			9.00
A	Metals	IRON	7439896	1620	UG/L	4646	05	26507	09/20/94		2,070.0	
							05	26515	09/21/94		1,210.0	
							05	26523	09/22/94		4,890.0	
							D05	26524	09/22/94		4,030.0	
							05	26533	09/23/94		1,010.0	
							05	26540	09/24/94		1,220.0	
A	Metals	LEAD	7439921	1620	UG/L	4646	05	26507	09/20/94			48.0
							05	26515	09/21/94			48.0
							05	26523	09/22/94			48.0
							D05	26524	09/22/94			48.0
							05	26533	09/23/94	B	49.4	
							05	26540	09/24/94			45.0
A	Metals	MANGANESE	7439965	1620	UG/L	4646	05	26507	09/20/94		571.0	
							05	26515	09/21/94		489.0	
							05	26523	09/22/94		489.0	
							D05	26524	09/22/94		480.0	
							05	26533	09/23/94		505.0	
							05	26540	09/24/94		537.0	
A	Metals	MERCURY	7439976	1620	UG/L	4646	05	26507	09/20/94			2.00
							05	26515	09/21/94			2.00
							05	26523	09/22/94		5.80	
							D05	26524	09/22/94		4.60	
							05	26533	09/23/94			2.00
							05	26540	09/24/94			2.00

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
A	Metals	MOLYBDENUM	7439987	1620	UG/L	4646	05	26507	09/20/94		676.0	
							05	26515	09/21/94		575.0	
							05	26523	09/22/94		479.0	
							D05	26524	09/22/94		487.0	
							05	26533	09/23/94		586.0	
							05	26540	09/24/94		571.0	
A	Metals	SELENIUM	7782492	1620	UG/L	4646	05	26507	09/20/94			20.0
							05	26515	09/21/94	B	24.7	
							05	26523	09/22/94		186.0	
							D05	26524	09/22/94	B	42.1	
							05	26533	09/23/94		66.7	
							05	26540	09/24/94	B	22.5	
A	Metals	SILVER	7440224	1620	UG/L	4646	05	26507	09/20/94			5.00
							05	26515	09/21/94			5.00
							05	26523	09/22/94		28.5	
							D05	26524	09/22/94		26.4	
							05	26533	09/23/94			5.00
							05	26540	09/24/94			5.00
A	Metals	TIN	7440315	1620	UG/L	4646	05	26507	09/20/94		32.8	
							05	26515	09/21/94		44.1	
							05	26523	09/22/94			30.0
							D05	26524	09/22/94			30.0
							05	26533	09/23/94			29.0
							05	26540	09/24/94			29.0
A	Metals	TITANIUM	7440326	1620	UG/L	4646	05	26507	09/20/94			3.00
							05	26515	09/21/94		7.00	
							05	26523	09/22/94	B	3.20	
							D05	26524	09/22/94	B	3.70	
							05	26533	09/23/94			3.00
							05	26540	09/24/94			3.00

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
A	Metals	ZINC	7440666	1620	UG/L	4646	05	26507	09/20/94		173.0	
							05	26515	09/21/94		103.0	
							05	26523	09/22/94		148.0	
							D05	26524	09/22/94		158.0	
							05	26533	09/23/94		79.4	
							05	26540	09/24/94		96.7	
B	Classicals	CHEMICAL OXYGEN DEMAND (COD)	C-004	410.1	UG/L	4646	06	26508	09/20/94		888,000.0	
							06	26516	09/21/94		291,000.0	
							06	26525	09/22/94		53,000.0	
							D06	26526	09/22/94		48,000.0	
							06	26534	09/23/94		18,000.0	
							06	26541	09/24/94		42,000.0	
B	Classicals	TOTAL DISSOLVED SOLIDS	C-010	160.1	UG/L	4646	06	26508	09/20/94		37,400,000.0	
							06	26516	09/21/94		33,700,000.0	
							06	26525	09/22/94		36,900,000.0	
							D06	26526	09/22/94		34,400,000.0	
							06	26534	09/23/94		43,100,000.0	
							06	26541	09/24/94		40,900,000.0	
B	Classicals	TOTAL ORGANIC CARBON (TOC)	C-012	415.1	UG/L	4646	06	26508	09/20/94			10,000.0
							06	26516	09/21/94			10,000.0
							06	26525	09/22/94			10,000.0
							D06	26526	09/22/94			10,000.0
							06	26534	09/23/94			10,000.0
							06	26541	09/24/94			10,000.0
B	Classicals	TOTAL SUSPENDED SOLIDS	C-009	160.2	UG/L	4646	06	26508	09/20/94			4,000.0
							06	26516	09/21/94			4,000.0
							06	26525	09/22/94		5,000.0	
							D06	26526	09/22/94			4,000.0
							06	26534	09/23/94		11,000.0	
							06	26541	09/24/94			4,000.0

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
B	Metals	ALUMINUM	7429905	1620	UG/L	4646	06	26508	09/20/94	B	164.0	
							06	26516	09/21/94	B	110.0	
							06	26525	09/22/94	B	181.0	
							D06	26526	09/22/94	B	127.0	
							06	26534	09/23/94	B	182.0	
							06	26541	09/24/94	B	192.0	
B	Metals	ANTIMONY	7440360	1620	UG/L	4646	06	26508	09/20/94		362.0	
							06	26516	09/21/94		260.0	
							06	26525	09/22/94		274.0	
							D06	26526	09/22/94		373.0	
							06	26534	09/23/94		411.0	
							06	26541	09/24/94		372.0	
B	Metals	ARSENIC	7440382	1620	UG/L	4646	06	26508	09/20/94	B	4.80	
							06	26516	09/21/94			2.00
							06	26525	09/22/94			20.0
							D06	26526	09/22/94			20.0
							06	26534	09/23/94		3.80	
							06	26541	09/24/94	B	10.0	
B	Metals	BORON	7440428	1620	UG/L	4646	06	26508	09/20/94		1,890.0	
							06	26516	09/21/94		1,730.0	
							06	26525	09/22/94		1,640.0	
							D06	26526	09/22/94		1,630.0	
							06	26534	09/23/94		1,700.0	
							06	26541	09/24/94		1,700.0	
B	Metals	CADMIUM	7440439	1620	UG/L	4646	06	26508	09/20/94		11.4	
							06	26516	09/21/94			5.00
							06	26525	09/22/94		56.9	
							D06	26526	09/22/94		58.5	
							06	26534	09/23/94		20.6	
							06	26541	09/24/94			5.00

Appendix A. Raw Data Listings

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
B	Metals	CHROMIUM	7440473	1620	UG/L	4646	06	26508	09/20/94			10.0
							06	26516	09/21/94			10.0
							06	26525	09/22/94			10.0
							D06	26526	09/22/94			10.0
							06	26534	09/23/94			10.0
							06	26541	09/24/94			10.0
B	Metals	COPPER	7440508	1620	UG/L	4646	06	26508	09/20/94			9.00
							06	26516	09/21/94	B	9.00	
							06	26525	09/22/94	B	19.9	
							D06	26526	09/22/94	B	9.30	
							06	26534	09/23/94			9.00
							06	26541	09/24/94			9.00
B	Metals	IRON	7439896	1620	UG/L	4646	06	26508	09/20/94	B	72.1	
							06	26516	09/21/94		110.0	
							06	26525	09/22/94		182.0	
							D06	26526	09/22/94	B	98.0	
							06	26534	09/23/94		147.0	
							06	26541	09/24/94		173.0	
B	Metals	LEAD	7439921	1620	UG/L	4646	06	26508	09/20/94			48.0
							06	26516	09/21/94			48.0
							06	26525	09/22/94			48.0
							D06	26526	09/22/94			48.0
							06	26534	09/23/94			45.0
							06	26541	09/24/94			45.0
B	Metals	MANGANESE	7439965	1620	UG/L	4646	06	26508	09/20/94		599.0	
							06	26516	09/21/94		548.0	
							06	26525	09/22/94		551.0	
							D06	26526	09/22/94		546.0	
							06	26534	09/23/94		505.0	
							06	26541	09/24/94		525.0	

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
B	Metals	MERCURY	7439976	1620	UG/L	4646	06	26508	09/20/94			2.00
							06	26516	09/21/94			2.00
							06	26525	09/22/94			2.00
							D06	26526	09/22/94			2.00
							06	26534	09/23/94			2.00
							06	26541	09/24/94			2.00
B	Metals	MOLYBDENUM	7439987	1620	UG/L	4646	06	26508	09/20/94		659.0	
							06	26516	09/21/94		626.0	
							06	26525	09/22/94		458.0	
							D06	26526	09/22/94		464.0	
							06	26534	09/23/94		590.0	
							06	26541	09/24/94		564.0	
B	Metals	SELENIUM	7782492	1620	UG/L	4646	06	26508	09/20/94			20.0
							06	26516	09/21/94			20.0
							06	26525	09/22/94			20.0
							D06	26526	09/22/94	B	29.3	
							06	26534	09/23/94	B	45.5	
							06	26541	09/24/94			20.0
B	Metals	SILVER	7440224	1620	UG/L	4646	06	26508	09/20/94			5.00
							06	26516	09/21/94			5.00
							06	26525	09/22/94			5.00
							D06	26526	09/22/94			5.00
							06	26534	09/23/94			5.00
							06	26541	09/24/94			5.00
B	Metals	TIN	7440315	1620	UG/L	4646	06	26508	09/20/94		39.5	
							06	26516	09/21/94			30.0
							06	26525	09/22/94			30.0
							D06	26526	09/22/94			30.0
							06	26534	09/23/94			29.0
							06	26541	09/24/94			29.0

Appendix A. Raw Data Listings

Option	Pollutant Type	Analyte Name	CAS Registry Number	Laboratory Method Used	Unit of Measurement for Amount	SCC Episode Number	Sample Point	SCC Sample Number	Sample Date	SCC Data Qualifier	Analyte Concentration	Detection Limit
B	Metals	TITANIUM	7440326	1620	UG/L	4646	06	26508	09/20/94			3.00
							06	26516	09/21/94	B	4.30	
							06	26525	09/22/94	B	4.90	
							D06	26526	09/22/94	B	3.90	
							06	26534	09/23/94			3.00
							06	26541	09/24/94		19.2	
B	Metals	ZINC	7440666	1620	UG/L	4646	06	26508	09/20/94			11.0
							06	26516	09/21/94			11.0
							06	26525	09/22/94		66.1	
							D06	26526	09/22/94			11.0
							06	26534	09/23/94		35.4	
							06	26541	09/24/94		24.8	

APPENDIX B

SUMMARY STATISTICS

Appendix B. Summary Statistics

(Concentration Units = ug/l)

Option	Category	Unit	Analyte	Cas_NO	Sample Point	Total Obs	Obs (ND)	Minimum (ND)	Maximum (ND)	Minimum (NC)	Maximum (NC)	Minimum (ND,NC)	Maximum (ND,NC)	Average (ND,NC)	Standard Deviation (ND,NC)
A	Classicals	UG/L	COD	C-004	05	5	0			18,000.0	918,000.0	18,000.0	918,000.0	239,000.0	387,000.0
			TDS	C-010	05	5	0			27,450,000.0	42,800,000.0	27,450,000.0	42,800,000.0	36,900,000.0	5,980,000.0
			TOC	C-012	05	5	5	10,000.0	10,000.0			10,000.0	10,000.0	10,000.0	0.0
			TSS	C-009	05	5	1	4,000.0	4,000.0	6,000.0	24,000.0	4,000.0	24,000.0	13,400.0	9,150.0
A	Metals	UG/L	ALUMINUM	7429905	05	5	0			143.0	267.0	143.0	267.0	197.0	47.8
			ANTIMONY	7440360	05	5	0			321.0	456.0	321.0	456.0	381.0	50.4
			ARSENIC	7440382	05	5	1	20.0	20.0	1.60	9.80	1.60	20.0	8.84	7.13
			BORON	7440428	05	5	0			1,525.0	1,960.0	1,525.0	1,960.0	1,710.0	161.0
			CADMIUM	7440439	05	5	1	5.00	5.00	8.60	95.5	5.00	95.5	47.2	39.2
			CHROMIUM	7440473	05	5	5	10.0	10.0			10.0	10.0	10.0	0.0
			COPPER	7440508	05	5	1	9.00	9.00	10.6	31.7	9.00	31.7	18.8	11.2
			IRON	7439896	05	5	0			1,010.0	4,460.0	1,010.0	4,460.0	1,990.0	1,440.0
			LEAD	7439921	05	5	4	45.0	48.0	49.4	49.4	45.0	49.4	47.7	1.62
			MANGANESE	7439965	05	5	0			484.5	571.0	484.5	571.0	517.0	36.4
			MERCURY	7439976	05	5	4	2.00	2.00	5.20	5.20	2.00	5.20	2.64	1.43
			MOLYBDENUM	7439987	05	5	0			483.0	676.0	483.0	676.0	578.0	68.5
			SELENIUM	7782492	05	5	1	20.0	20.0	22.5	114.0	20.0	114.0	49.6	40.9
			SILVER	7440224	05	5	4	5.00	5.00	27.4	27.4	5.00	27.4	9.49	10.0
			TIN	7440315	05	5	3	29.0	30.0	32.8	44.1	29.0	44.1	33.0	6.41
			TITANIUM	7440326	05	5	3	3.00	3.00	3.45	7.00	3.00	7.00	3.89	1.75
			ZINC	7440666	05	5	0			79.4	173.0	79.4	173.0	121.0	39.9

Appendix B. Summary Statistics
(Concentration Units = ug/l)

Option	Category	Unit	Analyte	Cas_NO	Sample Point	Total Obs	Obs (ND)	Minimum (ND)	Maximum (ND)	Minimum (NC)	Maximum (NC)	Minimum (ND,NC)	Maximum (ND,NC)	Average (ND,NC)	Standard Deviation (ND,NC)
B	Classicals	UG/L	COD	C-004	06	5	0			18,000.0	888,000.0	18,000.0	888,000.0	258,000.0	369,000.0
			TDS	C-010	06	5	0			33,700,000.0	43,100,000.0	33,700,000.0	43,100,000.0	38,200,000.0	3,830,000.0
			TOC	C-012	06	5	5	10,000.0	10,000.0			10,000.0	10,000.0	10,000.0	0.0
			TSS	C-009	06	5	3	4,000.0	4,000.0	4,500.0	11,000.0	4,000.0	11,000.0	5,500.0	3,080.0
B	Metals	UG/L	ALUMINUM	7429905	06	5	0			110.0	192.0	110.0	192.0	160.0	31.9
			ANTIMONY	7440360	06	5	0			260.0	411.0	260.0	411.0	346.0	57.1
			ARSENIC	7440382	06	5	2	2.00	20.0	3.80	10.0	2.00	20.0	8.12	7.28
			BORON	7440428	06	5	0			1,635.0	1,890.0	1,635.0	1,890.0	1,730.0	95.4
			CADMIUM	7440439	06	5	2	5.00	5.00	11.4	57.7	5.00	57.7	19.9	22.1
			CHROMIUM	7440473	06	5	5	10.0	10.0			10.0	10.0	10.0	0.0
			COPPER	7440508	06	5	3	9.00	9.00	9.00	14.6	9.00	14.6	10.1	2.50
			IRON	7439896	06	5	0			72.1	173.0	72.1	173.0	128.0	38.7
			LEAD	7439921	06	5	5	45.0	48.0			45.0	48.0	46.8	1.64
			MANGANESE	7439965	06	5	0			505.0	599.0	505.0	599.0	545.0	35.1
			MERCURY	7439976	06	5	5	2.00	2.00			2.00	2.00	2.00	0.0
			MOLYBDENUM	7439987	06	5	0			461.0	659.0	461.0	659.0	580.0	75.6
			SELENIUM	7782492	06	5	3	20.0	20.0	24.6	45.5	20.0	45.5	26.0	11.1
			SILVER	7440224	06	5	5	5.00	5.00			5.00	5.00	5.00	0.0
			TIN	7440315	06	5	4	29.0	30.0	39.5	39.5	29.0	39.5	31.5	4.50
			TITANIUM	7440326	06	5	2	3.00	3.00	4.30	19.2	3.00	19.2	6.78	6.98
			ZINC	7440666	06	5	2	11.0	11.0	24.8	38.5	11.0	38.5	24.2	13.0

APPENDIX C

FACILITY-SPECIFIC LONG-TERM AVERAGES, VARIABILITY FACTORS, AND POTENTIAL LIMITATIONS

Appendix C. Facility-Specific Long-term Averages, Variability Factors, and Estimated Limitations

Estimates Calculated Based on Assumed Underlying Delta-Lognormal Distribution

SCC Data Only (Concentration Units = ug/l)

Option	Category	Analyte	Cas_NO	SamplePoint	Total Obs	Obs (ND)	Average (ND,NC)	Standard Deviation (ND,NC)	Estimated Long Term Average	Estimated Standard Deviation	Estimated Daily Maximum Limit	Daily VF	Estimated Monthly Average Limit	Monthly VF
A	Classicals	COD	C-004	05	5	0	239,000.0	387,000.0	306,000.0	1,150,000.0	3,640,000.0	11.9	1,090,000.0	3.55
		TDS	C-010	05	5	0	36,900,000.0	5,980,000.0	37,000,000.0	6,500,000.0	54,700,000.0	1.48	42,600,000.0	1.15
		TSS	C-009	05	5	1	13,400.0	9,150.0	14,300.0	12,100.0	58,700.0	4.10	18,800.0	1.31
A	Metals	ALUMINUM	7429905	05	5	0	197.0	47.8	198.0	48.1	336.0	1.70	240.0	1.21
		ANTIMONY	7440360	05	5	0	381.0	50.4	382.0	50.1	513.0	1.34	424.0	1.11
		ARSENIC	7440382	05	5	1	8.84	7.13	9.52	8.23	32.3	3.39	17.2	1.81
		BORON	7440428	05	5	0	1,710.0	161.0	1,710.0	158.0	2,110.0	1.23	1,840.0	1.08
		CADMIUM	7440439	05	5	1	47.2	39.2	62.3	107.0	484.0	7.76	160.0	2.57
		COPPER	7440508	05	5	1	18.8	11.2	19.6	13.5	68.4	3.49	32.2	1.64
		IRON	7439896	05	5	0	1,990.0	1,440.0	2,030.0	1,350.0	6,910.0	3.40	3,280.0	1.62
		MANGANESE	7439965	05	5	0	517.0	36.4	518.0	35.9	607.0	1.17	548.0	1.06
		MOLYBDENUM	7439987	05	5	0	578.0	68.5	579.0	69.4	759.0	1.31	638.0	1.10
		SELENIUM	7782492	05	5	1	49.6	40.9	53.5	53.9	265.0	4.95	104.0	1.95
		TIN	7440315	05	5	3	33.0	6.41	33.2	7.01	57.3	1.73	39.3	1.19
		TITANIUM	7440326	05	5	3	3.89	1.75	4.03	2.26	13.1	3.25	6.14	1.53
		ZINC	7440666	05	5	0	121.0	39.9	122.0	41.1	248.0	2.03	159.0	1.30
B	Classicals	COD	C-004	06	5	0	258,000.0	369,000.0	351,000.0	1,190,000.0	4,000,000.0	11.4	1,210,000.0	3.44
		TDS	C-010	06	5	0	38,200,000.0	3,830,000.0	38,200,000.0	3,830,000.0	48,000,000.0	1.26	41,400,000.0	1.08
		TSS	C-009	06	5	3	5,500.0	3,080.0	5,840.0	4,420.0	24,300.0	4.16	7,460.0	1.28
B	Metals	ALUMINUM	7429905	06	5	0	160.0	31.9	161.0	35.7	262.0	1.62	192.0	1.19
		ANTIMONY	7440360	06	5	0	346.0	57.1	347.0	61.1	513.0	1.48	399.0	1.15
		ARSENIC	7440382	06	5	2	8.12	7.28	8.27	8.78	16.6	2.01	16.2	1.96
		BORON	7440428	06	5	0	1,730.0	95.4	1,730.0	93.4	1,960.0	1.13	1,810.0	1.04
		CADMIUM	7440439	06	5	2	19.9	22.1	22.0	28.9	137.0	6.20	49.3	2.24
		COPPER	7440508	06	5	3	10.1	2.50	10.3	3.12	22.4	2.18	13.1	1.28
		IRON	7439896	06	5	0	128.0	38.7	130.0	45.7	271.0	2.08	171.0	1.31
		MANGANESE	7439965	06	5	0	545.0	35.1	545.0	34.7	631.0	1.16	574.0	1.05
		MOLYBDENUM	7439987	06	5	0	580.0	75.6	581.0	80.3	793.0	1.36	650.0	1.12
		SELENIUM	7782492	06	5	3	26.0	11.1	26.7	13.4	78.3	2.93	39.1	1.46
		TITANIUM	7440326	06	5	2	6.78	6.98	7.38	9.05	44.2	5.99	15.9	2.16
		ZINC	7440666	06	5	2	24.2	13.0	24.3	12.5	53.2	2.19	35.4	1.45

For TSS, the monthly variability factors are estimated assuming 20 days of sampling.
For all other pollutants, monthly variability factors are estimated assuming 4 days of sampling.

APPENDIX D

POLLUTANT-SPECIFIC LONG-TERM AVERAGES

AND VARIABILITY FACTORS

Appendix D. Pollutant-Specific Long-term Averages and Variability Factors

Estimates calculated as Mean of Facility-specific Results. No Imputation Performed
SCC Data Only

Option	Category	Analyte	Cas_NO	Long Term Average (ug/l)	Daily VF	Monthly VF
A	Classicals	COD	C-004	306,000.0	11.9	3.55
		TDS	C-010	37,000,000.0	1.48	1.15
		TSS	C-009	14,300.0	4.10	1.31
A	Metals	ALUMINUM	7429905	198.0	1.70	1.21
		ANTIMONY	7440360	382.0	1.34	1.11
		ARSENIC	7440382	9.52	3.39	1.81
		BORON	7440428	1,710.0	1.23	1.08
		CADMIUM	7440439	62.3	7.76	2.57
		COPPER	7440508	19.6	3.49	1.64
		IRON	7439896	2,030.0	3.40	1.62
		MANGANESE	7439965	518.0	1.17	1.06
		MOLYBDENUM	7439987	579.0	1.31	1.10
		SELENIUM	7782492	53.5	4.95	1.95
		TIN	7440315	33.2	1.73	1.19
		TITANIUM	7440326	4.03	3.25	1.53
		ZINC	7440666	122.0	2.03	1.30
B	Classicals	COD	C-004	351,000.0	11.4	3.44
		TDS	C-010	38,200,000.0	1.26	1.08
		TSS	C-009	5,840.0	4.16	1.28
B	Metals	ALUMINUM	7429905	161.0	1.62	1.19
		ANTIMONY	7440360	347.0	1.48	1.15
		ARSENIC	7440382	8.27	2.01	1.96
		BORON	7440428	1,730.0	1.13	1.04
		CADMIUM	7440439	22.0	6.20	2.24
		COPPER	7440508	10.3	2.18	1.28
		IRON	7439896	130.0	2.08	1.31
		MANGANESE	7439965	545.0	1.16	1.05
		MOLYBDENUM	7439987	581.0	1.36	1.12
		SELENIUM	7782492	26.7	2.93	1.46
		TITANIUM	7440326	7.38	5.99	2.16
		ZINC	7440666	24.3	2.19	1.45

For TSS, the monthly variability factors are estimated assuming 20 days of sampling.
For all other pollutants, monthly variability factors are estimated assuming 4 days of sampling.

APPENDIX E

GROUP-LEVEL VARIABILITY FACTORS

Appendix E. Group-level Variability Factors

SCC Data Only

Option	Category	Number of Pollutants in Each Group	Daily VF	Monthly VF
A	Metals	13	2.03	1.30
B	Metals	12	2.05	1.30

For Metals, monthly variability factors are estimated assuming 4 days of sampling.

Option A Metal Group defined as Al, Sb, As, B, Cd, Cr*, Cu, Fe, Pb*, Mn, Hg*, Mo, Se, Ag*, Sn*, Ti, Zn
Option B Metal Group defined as Al, Sb, As, B, Cd, Cr*, Cu, Fe, Pb*, Mn, Hg*, Mo, Se, Ag*, Sn*, Ti, Zn
*Implies Delta-Lognormal Estimation Criteria Not Met

APPENDIX F

POLLUTANT-SPECIFIC LONG-TERM AVERAGES AND POTENTIAL LIMITATIONS

Appendix F. Pollutant-Specific Long-Term Averages and Limitations

Potential Limitations for Metals are Imputed Based on Group-level Variability Factors within Option
SCC Data Only

Option	Category	Analyte	Cas_NO	Estimated Long Term Average (ug/l)	Potential Daily Maximum Limit (ug/l)	Potential Monthly Average Limit (ug/l)
A	Metals	ALUMINUM	7429905	198.0	401.0	257.0
		ANTIMONY	7440360	382.0	775.0	496.0
		ARSENIC	7440382	9.52	19.3	12.4
		BORON	7440428	1,710.0	3,460.0	2,220.0
		CADMIUM	7440439	62.3	127.0	81.0
		CHROMIUM	7440473	10.0	20.3	13.0
		COPPER	7440508	19.6	39.8	25.5
		IRON	7439896	2,030.0	4,120.0	2,640.0
		LEAD	7439921	47.7	96.8	62.0
		MANGANESE	7439965	518.0	1,050.0	673.0
		MERCURY	7439976	2.64	5.36	3.43
		MOLYBDENUM	7439987	579.0	1,180.0	753.0
		SELENIUM	7782492	53.5	109.0	69.5
		SILVER	7440224	9.49	19.3	12.3
		TIN	7440315	33.2	67.3	43.1
		TITANIUM	7440326	4.03	8.18	5.24
		ZINC	7440666	122.0	248.0	159.0
B	Metals	ALUMINUM	7429905	161.0	330.0	209.0
		ANTIMONY	7440360	347.0	709.0	449.0
		ARSENIC	7440382	8.27	16.9	10.7
		BORON	7440428	1,730.0	3,540.0	2,240.0
		CADMIUM	7440439	22.0	45.1	28.5
		CHROMIUM	7440473	10.0	20.5	13.0
		COPPER	7440508	10.3	21.0	13.3
		IRON	7439896	130.0	267.0	169.0
		LEAD	7439921	46.8	95.7	60.6
		MANGANESE	7439965	545.0	1,120.0	706.0
		MERCURY	7439976	2.00	4.09	2.59
		MOLYBDENUM	7439987	581.0	1,190.0	753.0
		SELENIUM	7782492	26.7	54.6	34.6
		SILVER	7440224	5.00	10.2	6.48
		TIN	7440315	31.5	64.4	40.8
		TITANIUM	7440326	7.38	15.1	9.56
		ZINC	7440666	24.3	49.8	31.5

For Metals, monthly variability factors are estimated assuming 4 days of sampling.

Option A Metal Group defined as Al, Sb, As, B, Cd, Cr*, Cu, Fe, Pb*, Mn, Hg*, Mo, Se, Ag*, Sn, Ti, Zn
Option B Metal Group defined as Al, Sb, As, B, Cd, Cr*, Cu, Fe, Pb*, Mn, Hg*, Mo, Se, Ag*, Sn*, Ti, Zn
*Implies Delta-Lognormal Estimation Criteria Not Met