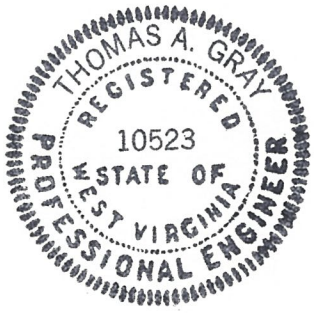
CCR RULE GROUNDWATER STATISTICAL METHOD SELECTION CERTIFICATION
Harrison Power Station Coal Combustion Byproduct Landfill
Harrison County, West Virginia

We the undersigned, being a qualified professional geologist and qualified professional engineers as noted, hereby certify that to the best of our knowledge, information, and belief, the selected statistical methods described herein are currently appropriate for evaluating the groundwater monitoring dataset for the Coal Combustion Byproduct Landfill at the Harrison Power Station and meet the applicable requirements of §257.93(f) of the USEPA's *Disposal of Coal Combustion Residuals from Electric Utilities* Rule (40 CFR, Part 257, Subpart D).

Printed Name:	Jeffrey P. Orient, PG
Certification Date:	October 16, 2017
PA License Number:	PG 000068G
License Expiration:	September 30, 2019
Company/Contact:	Tetra Tech, Inc. 661 Anderson Dr. Foster Plaza 7 Pittsburgh, PA 15220 (412) 921-7090

Printed Name:	Robert C. Baker, PE
Certification Date:	October 16, 2017
PA License Number:	PE 057680E
License Expiration:	September 30, 2019
Company/Contact:	Tetra Tech, Inc. 400 Penn Center Blvd. Suite 200 Pittsburgh, PA 15235 (412) 829-3600

Printed Name:	Thomas A. Gray, PE
Certification Date:	October 16, 2017
WV License Number:	010523
License Expiration:	December 31, 2018
Company/Contact:	Tetra Tech, Inc. 661 Anderson Dr. Foster Plaza 7 Pittsburgh, PA 15220 (412) 921-7090

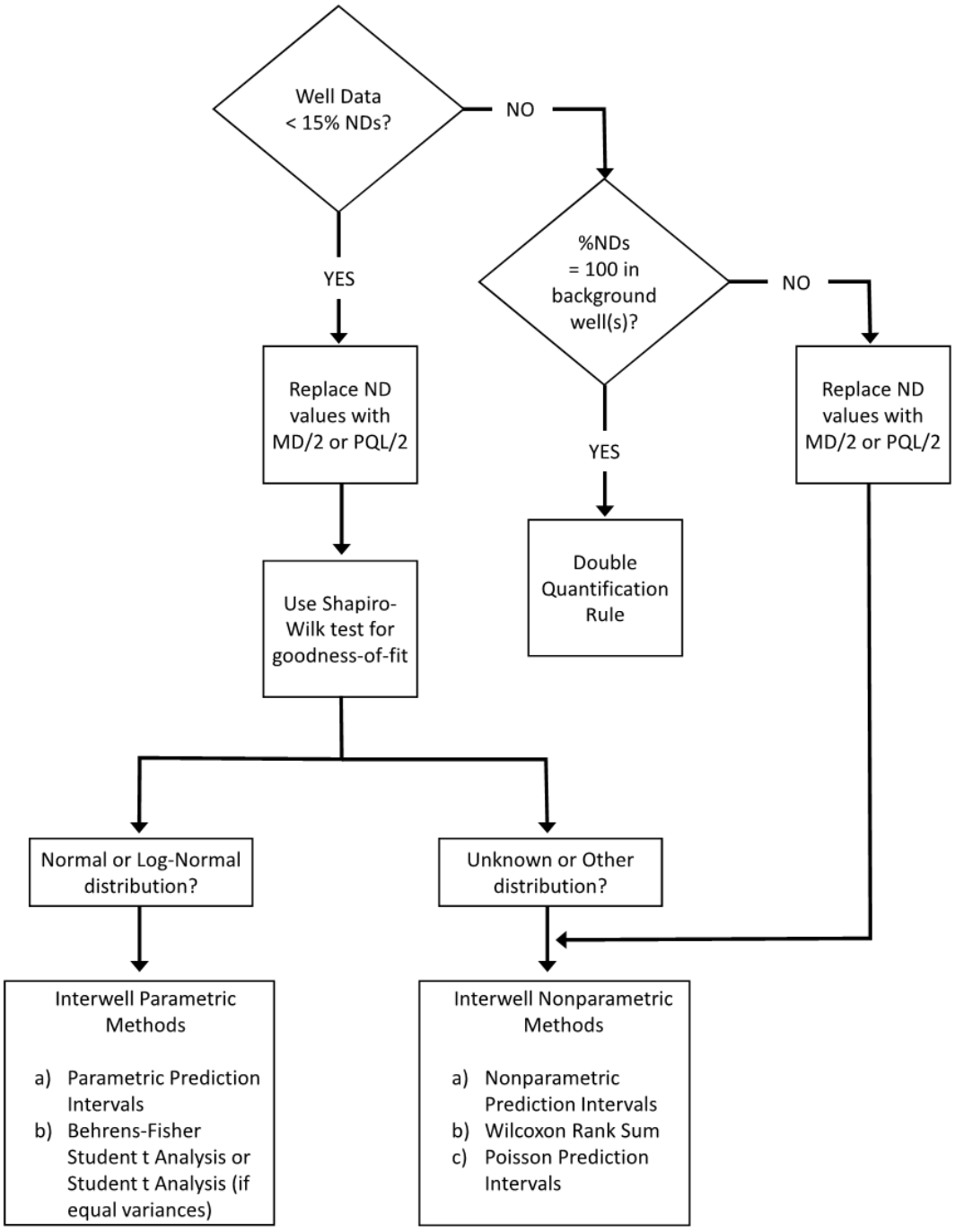


This document summarizes the statistical methods which will be utilized for evaluating groundwater analytical results associated with the Harrison Power Station Coal Combustion Byproduct Landfill's (hereafter referred to as the "Harrison CCR unit") CCR groundwater monitoring program. Based on the attributes of the current data set, the statistical methods were selected among the available methods referenced in 40 CFR 257.93(f) and which meet the performance standards referenced in 40 CFR 257.93(g). Data from the first eight rounds of groundwater analytical results collected at one upgradient/background well and four downgradient wells at the site were evaluated in terms of percent non-detects and data distributions to select the appropriate statistical method for each parameter. This document also summarizes potentially applicable statistical methods that may be used in the future should changes in data distributions and non-detect percentages occur in the Harrison CCR unit dataset.

Overview of Method Selection Approach

Figure 1 presents a decision tree for interwell tests, where separate data are available at background and downgradient wells.

Figure 1. Decision Tree Schematic. Flow Chart for interwell tests, where separate data are available at background and downgradient wells.



The following summarizes the Shapiro-Wilk W test and the statistical methods referenced in Figure 1 that are pertinent to the Harrison CCR unit.

Shapiro-Wilk W Test for Normality and Lognormality

The type of data distribution is required to be determined in order to select an appropriate statistical method [per CCR Rule 40 CFR 257.93(g)(1)]. The Shapiro-Wilk W test is a goodness-of-fit test (two-sided and parametric) on whether the data have been drawn from an underlying normal distribution (Gilbert, 1987). The null hypothesis H_o is that the population has a normal distribution. The alternative hypothesis H_a is that the population does not have a normal distribution. A goodness-of-fit test for lognormal distributions is performed by first taking the natural logarithm of all the data values and then applying the W test on the transformed data. The Shapiro-Wilk W test is valid for sample sizes less than or equal to 50 values, which currently applies to the Harrison CCR unit. For data sets larger than 50, the Shapiro-Francia test is used (Gibbons, 1994).

PARAMETRIC METHODS

Parametric Prediction Intervals for Future Compliance Values

The prediction interval method is one of the statistical methods cited in the CCR Rule [40 CFR 257.93(f)(3)]. Both parametric and non-parametric versions of this statistical test are available (as explained in the Unified Guidance USEPA, 2009), which is cited in the discussion section of the CCR Rule [p. 21401 K(3) and other places]. The parametric prediction interval method calculates upper and lower values, based on background data, against which future values from compliance locations will be compared (USEPA, 1989). This method calculates a parametric prediction interval from all pooled background data for a specified base period from one or more locations. The data are then used to compute a prediction interval for an initial period. The parameter value for each of the compliance location intervals is then compared to the upper bound of the prediction interval. A statistically significant exceedance time period is indicated when the value of an individual measurement for a compliance location exceeds the upper bound of the prediction interval, or the lower bound for pH.

The data or transformed data should be normally distributed. A minimum of four observations per period are recommended for the compliance location data. A minimum of a one year base period of background observations is recommended for construction of the prediction interval. The data should be free of outliers. Based on the current Harrison CCR Unit dataset distribution, this method is appropriate for use at the site for some parameters.

Behrens-Fisher Student t-Analysis

The Student t-test is a one-sided, parametric test that compares the means from two data sets. If confidence ranges for the means overlap, then the two means are not significantly different. This test assumes normally-distributed data.

Satterthwaite's t-test (Iman and Conover, 1983) is a modified form of the standard t-test that is appropriate when the background and site distributions have unequal variances. Testing data sets with unequal variances are called Behrens-Fisher problems. The Student t-test makes three key assumptions: (1) that the two location data sets are independent; (2) not serially correlated; and (3) that both location data sets have normal distributions (Guttman et al.,1971; Gilbert, 1987). If these assumptions are not met, the non-parametric Wilcoxon Rank-Sum test should be used for determining whether the means of

two locations are different (Loftis, et al., 1987). Based on the current Harrison CCR unit dataset distribution, this method is not appropriate for use at the site but has been retained in this document as the method may become applicable for some constituents as more data are collected in the future.

NON-PARAMETRIC METHODS

Non-Parametric Prediction Intervals for Future Compliance Values

The prediction interval method is one of the statistical methods cited in the CCR Rule [40 CFR 257.93(f)(3)]. The non-parametric prediction interval calculates the prediction interval using pooled background data over a specified base period. The background data are pooled from one or more locations. The pooled background data are ranked and the minimum value is identified as the one-sided, lower prediction limit for pH only, P_l, and an appropriate value is identified as the one-sided, upper prediction limit, P_u, depending on the number of background samples (as described in Section 18.3.1 USEPA, 2009). Lower and upper, non-parametric, one-sided confidence limits are computed for the compliance locations. No assumption is made concerning the underlying distribution of the data. However, the assumption is made that the unknown distribution in the background and compliance data is continuous and is the same in both background and compliance datasets in the absence of contamination.

At least four background values and at least one compliance location are needed for this analysis. However, there need not be any actual data in the selected compliance locations if the user only wishes to determine the prediction intervals. If an individual measurement from a compliance location exceeds the upper limit, then a statistically significant exceedance is declared. Normally, only an exceedance of the upper limit is of concern, except for pH. A general discussion of estimating non-parametric prediction limits and alternative verification procedures is given in Gibbons (1994) and in Section 18.3.1 (USEPA, 2009). Based on the current Harrison CCR unit dataset attributes, this method is appropriate for use at the site for some parameters.

Non-Parametric Prediction Intervals for Future Compliance Median

The non-parametric prediction interval method is one of the statistical methods cited in the CCR Rule [40 CFR 257.93(f)(3)]. The USEPA (2009) describes in Section 18.3.1 of the Unified Guidance the various strategies available for setting the upper prediction limit when the background data are non-parametrically distributed and sufficient compliance data are available. In particular, the option of using the median of three future compliance measurements to test against the upper prediction limit is described on page 18-21. For that approach, the user is given the option of setting the upper prediction limit to either the largest, the 2nd-to-largest, or the 3rd-to-largest background measurement. The corresponding confidence limit for each of these choices is affected by the background sample size n_{bg}. The confidence level increases as n_{bg} increases. In addition, for the same sample size n_{bg}, the confidence level decreases as one selects values smaller than the maximum when the prediction limit is chosen to be the jth largest background measurement. A complete statistical table is given on page D-31 of the Unified Guidance (USEPA, 2009). A subset of choices is summarized in Table 1 below, in which only the 95th and 99th percent confidence level values are listed. Note that for the 95% confidence level, only 9 background data values are needed when selecting the maximum background measurement as the upper prediction limit, as compared to needing 24 background values when selecting the third-to-largest background measurement for the upper prediction limit. Based on the current Harrison CCR unit dataset attributes, this method is not appropriate for use at the site but has been retained in this document as the method may become applicable for some constituents as more data are collected in the future.

Table 1. Minimum background (BG) sample size n_{bg} for non-parametric data sets as a function of desired confidence level and the decision to use the j^{th} largest background measurement as the upper prediction limit when proposing to test against the future median of order 3 (i.e., three compliance measurements).

	Upper prediction limit based on using largest BG value	Upper prediction limit based on using 2nd–to-largest BG value	Upper prediction limit based on using 3rd–to-largest BG value
Minimum n_{bg} needed to achieve 95% confidence level	**9**	**16**	**24**
Minimum n_{bg} needed to achieve 99% confidence level	**22**	**39**	**56**

Wilcoxon Rank-Sum Analysis

The Wilcoxon Rank-Sum test is a one-sided, non-parametric test that compares the means from two data series. This method is an alternative statistical test method allowed under the CCR Rule [40 CFR 257.93(f)(5)]. If confidence ranges for the means overlap, then the two means are not significantly different. If multiple background locations are specified, they are pooled. The evaluation is conducted for each compliance location/parameter combination, and determines whether the mean concentration of the specified parameter at the compliance location is statistically higher than the mean concentration of that parameter at the pooled background locations. The Wilcoxon Rank-Sum test assumes that: (1) both data sets contain random values from their respective populations, and (2) in addition to independence within each data set, there is mutual independence between the two sample sets. No assumptions are made about data distribution. The null hypothesis is that the two location means are equal, and the alternative hypothesis is that the two location means are different. Based on the current Harrison CCR unit dataset attributes, this method is not appropriate for use at the site but has been retained in this document as the method may become applicable for some constituents as more data are collected in the future.

Poisson Prediction Interval

The non-parametric prediction interval method is one of the statistical methods cited in the CCR Rule [40 CFR 257.93(f)(3)]. The Poisson prediction interval method calculates upper and lower, one-or two-sided, non-parametric prediction limits, based on background data, against which future data from compliance wells will be compared. The Poisson distribution in statistics is used to model rare events. The Poisson model describes the behavior of a series of independent events that occur while taking a large number of observations. For the purposes of this document, an *event* occurs when the chemical concentration of a sample is above the level of detection. The probability of detection is low but it remains constant from observation to observation.

One of the key distinctions between the Poisson model and other non-parametric models is that the Poisson model utilizes the magnitude of the measured concentrations in its algorithm. Upon selecting a *scaling* parameter, all sampled concentrations for a particular chemical at a location are then converted into an equivalent number of *chemical units* or *counts*. The model then computes the average rate of occurrence of these counts for the chemical from a specified sample set. Finally, it predicts the lower and upper limit for an interval that will contain all of the future measurements of this chemical at the location.

The Poisson model can only be used if there is available at least one background measurement that is detected. As discussed in the Unified Guidance (USEPA, 2009, pp. 6-11 and 6-37), the Double Quantification Rule must be used when none of the background measurements are detected. Based on the current Harrison CCR unit dataset attributes, this method is appropriate for use at the site for some parameters.

Double Quantification Rule

The Double Quantification Rule applies when all data from the background wells have no detected values for a particular constituent. If, during a sampling event, that particular constituent is detected in a downgradient well, a subsequent sample (resampling) would be collected from that well and analyzed. If the downgradient concentration for that constituent in that given well is higher than the reporting limit in both the original sample *and* in the verification resample, then a statistically significant increase determination would be made. Based on the current Harrison CCR unit dataset attributes, this method may be appropriate for use at the site for some parameters.

SELECTED STATISTICAL METHODS

Selected methods for upgradient well MW-5 to downgradient well MW-17, upgradient well MW-5 to downgradient well MW-19, and upgradient well MW-5 to downgradient well MW-20 comparisons by parameter are listed in Tables 2, 3, and 4. Downgradient well MW-18 will not be used until more data are collected, since only three samples were collected due to insufficient water. As indicated in Tables 2, 3, and 4 most of the method selection is driven by the percent non-detects and dissimilar data distributions. The statistical methods considered to be appropriate for the first detection monitoring event to compare up and downgradient wells are either upper parametric or non-parametric prediction limit future values for the Appendix III parameters, except for pH where the lower limit would be used. For Appendix IV parameters, the upper non-parametric prediction limit future values or the Poisson prediction limits will be used for most of the parameters. The Double Quantification Rule will be used when all data from the upgradient/background well has no detected values for a given parameter. The same procedure will be followed for each upgradient to downgradient well pair. As more data are added to the upgradient and downgradient data sets, the percent non-detects and data distributions may change such that other methods could become appropriate for some parameters.

Table 2 - Selected Statistical Methods for Appendix III and IV Parameters for (U)pgradient well MW-5 and (D)owngradient well MW-17.

Appendix III Parameters			
Parameter	**% Non-Detect U/D**	**Data Distributions U/D[a]**	**Planned Methods**
B	0% U, 0% D	UO/Normal	Non-parametric Prediction Limits
Ca	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Cl	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Fl	0 % U, 0% D	Normal/Normal	Parametric Prediction Limits
pH	0% U, 0% D	Normal/ Normal	Parametric Prediction Limits
SO_4	0 % U, 0% D	UO/Normal	Non-parametric Prediction Limits
TDS	0 % U, 0% D	Normal/UO	Non-parametric Prediction Limits

[a]Data distributions are Log N for Log normal and UO for unknown/other.

Appendix IV Parameters			
Parameter	**% Non-Detect U/D**	**Data Distributions U/D[a]**	**Planned Methods**
Sb	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
As	50 % U, 12.5 % D	UO/Normal	Non-parametric Prediction Limits
Ba	0 % U, 0% D	Normal/Normal	Parametric Prediction Limits
Be	100 % U, 100 % D	UO/UO	DQ or Poisson Prediction Limit[b]
Cd	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
T Cr	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Co	100 % U, 100 % D	UO/UO	DQ or Poisson Prediction Limit
Pb	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
Li	0 % U, 0 % D	Normal/ Normal	Parametric Prediction Limits
Hg	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Mo	12.5% U, 0% D	UO/Normal	Non-parametric Prediction Limits
Se	100 % U, >50% D	UO/UO	DQ or Poisson Prediction Limit
Tl	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
Ra 226+228	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit

[a]Data distributions are Log N for Log normal and UO for unknown/other.

[b]DQ is Double Quantitation Rule; If there is a detected value in upgradient well for detection monitoring event, can use Poisson Prediction Limit.

Table 3 - Selected Statistical Methods for Appendix III and IV Parameters for (U)pgradient well MW-5 and (D)owngradient well MW-19.

Appendix III Parameters			
Parameter	**% Non-Detect U,D**	**Data Distributions U/D[a]**	**Planned Methods**
B	0% U, 0% D	UO/Normal	Non-parametric Prediction Limits
Ca	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Cl	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Fl	0 % U, 0% D	Normal/Normal	Parametric Prediction Limits
pH	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
SO_4	0 % U, 0% D	UO/Normal	Non-parametric Prediction Limits
TDS	0 % U, 0% D	Normal/UO	Non-parametric Prediction Limits

[a]Data distributions are Log N for Log normal and UO for unknown/other.

Appendix IV Parameters			
Parameter	**% Non-Detect U,D**	**Data Distributions U/D[a]**	**Planned Methods**
Sb	>50 % U, >50 % D	UO/Normal	Poisson Prediction Limit
As	50 % U, 0 % D	UO/Normal	Non-parametric Prediction Limits
Ba	0 % U, 0% D	Normal/Log Normal	Non-parametric Prediction Limits
Be	100 % U, 100 % D	UO/UO	DQ or Poisson Prediction Limit[b]
Cd	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
T Cr	>50 % U, 42.9 % D	UO/Normal	Poisson Prediction Limit
Co	100 % U, 57.1 % D	UO/Normal	DQ or Poisson Prediction Limit
Pb	>50 % U, 28.6 % D	UO/Normal	Poisson Prediction Limit
Li	0 % U, >50 %	Normal/ UO	Poisson Prediction Limit
Hg	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Mo	12.5% U, 0% D	UO/UO	Non-parametric Prediction Limit
Se	100 % U, 100% D	UO/UO	DQ or Poisson Prediction Limit
Tl	>50 % U, 100 % D	UO/UO	Poisson Prediction Limit
Ra 226+228	>50 % U, 14.3 % D	UO/Log Normal	Poisson Prediction Limit

[a]Data distributions are N for Normal, Log N for Log normal and UO for unknown/other.

[b]DQ is Double Quantitation Rule; If there is a detected value in upgradient well for detection monitoring event, can use Poisson Prediction Limit.

Table 4 - Selected Statistical Methods for Appendix III and IV Parameters for (U)pgradient well MW-5 and (D)owngradient well MW-20.

Appendix III Parameters			
Parameter	**% Non-Detect U,D**	**Data Distributions U/D[a]**	**Planned Methods**
B	0% U, 0% D	UO/Normal	Non-parametric Prediction Limits
Ca	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Cl	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
Fl	0 % U, 12.5% D	Normal/Normal	Parametric Prediction Limits
pH	0% U, 0% D	Normal/Normal	Parametric Prediction Limits
SO_4	0 % U, 0% D	UO/Normal	Non-parametric Prediction Limits
TDS	0 % U, 0% D	Normal/Normal	Parametric Prediction Limits

[a]Data distributions are Log N for Log normal and UO for unknown/other.

Appendix IV Parameters			
Parameter	**% Non-Detect U,D**	**Data Distributions U/D[a]**	**Planned Methods**
Sb	>50 % U, 50 % D	UO/UO	Poisson Prediction Limit
As	50 % U, 0 % D	UO/Normal	Non-parametric Prediction Limits
Ba	0 % U, 0% D	Normal/Normal	Parametric Prediction Limits
Be	100 % U, >50 % D	UO/Normal	DQ or Poisson Prediction Limit[b]
Cd	>50 % U, 50 % D	UO/UO	Poisson Prediction Limit
T Cr	>50 % U, 12.5 % D	UO/Log Normal	Poisson Prediction Limit
Co	100 % U, 100 % D	UO/UO	DQ or Poisson Prediction Limit
Pb	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Li	0 % U, >50 %	Normal/Log Normal	Poisson Prediction Limit
Hg	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Mo	12.5% U, 0% D	UO/UO	Non-parametric Prediction Limit
Se	100 % U, 25% D	UO/Normal	DQ or Poisson Prediction Limit
Tl	>50 % U, >50 % D	UO/UO	Poisson Prediction Limit
Ra 226+228	>50 % U, 25 % D	UO/UO	Poisson Prediction Limit

[a]Data distributions are N for Normal, Log N for Log normal and UO for unknown/other.

[b]DQ is Double Quantitation Rule; If there is a detected value in upgradient well for detection monitoring event, can use Poisson Prediction Limit.

References for Statistical Methods

Gibbons, R. 1994. Statistical Methods for Groundwater Monitoring. John Wiley & Sons, New York, NY, 286 pp.

Gilbert, R. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold Co., NY. 320 pp.

Guttman, I., S. Wilks, and J. S. Hunter. 1971. Introductory Engineering Statistics, John Wiley & Sons, Inc., NY. 249 pp.

Iman, R. and W. Conover. 1983. A Modern Approach to Statistics, John Wiley & Sons, Inc., NY. 497 pp.

Loftis, J., J. Harris, and R. Montgomery. 1987. Detecting changes in ground water quality at regulated facilities, Ground Water Monitoring Review, Winter, pp. 72-76.

USEPA. 1989. Statistical analysis of ground-water monitoring data at RCRA facilities - Interim final guidance. Environmental Protection Agency report EPA/530/SW-89/026, NTIS PB89-151047, Office of Solid Waste, Washington, D.C.

USEPA. 2009. "Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities." Unified Guidance EPA 530/R-09-007. Washington DC: United States Environmental Protection Agency.