Statistical Method Certification
Ash Landfill - Hunter Power Plant

Castle Dale, Utah

Prepared For:
Hunter Power Plant
Highway 10, S of Castle Dale
Castle Dale, UT 84513

PacifiCorp
1407 West North Temple, Suite 280
Salt Lake City, Utah 84116

Prepared By:
Water & Environmental Technologies
480 East Park Street
Butte, Montana 59701
406.782.5220

STATISTICAL METHOD CERTIFICATION

The undersigned, hereby certify that the statistical method selected to evaluate groundwater quality for coal combustion residual unit Ash Landfill at the Hunter Power Plant in Castle Dale, Utah satisfies the criteria contained in the Code of Federal Regulations, 40 CFR, Part 257 and 261, *Hazardous and Solid Waste Management System; Disposal of Coal Combustion Residuals from Electric Utilities; Final Rule*. The required statistical method narrative is attached. The discussion of the statistical method is also included in the *Sampling and Analysis Plan, Ash Landfill, Hunter Power Plant, Castle Dale, Utah*, 2017, which is part of the facility operating record.

Dated this 16 day of October, 2017.

By: Steven Matthew Anderson

Its: Senior Engineer

REGISTERED PROFESSIONAL ENGINEER
No. 8049139
STEVEN MATTHEW ANDERSON
10/16/2017
STATE OF UTAH

Dated this 16 day of October, 2017.

By: David J Erickson

Its: Principal Hydrogeologist

LICENSED PROFESSIONAL GEOLOGIST
DAVID J. ERICKSON
PG-2250
10/16/17
UTAH

Statistical Method Narrative Description

The upper tolerance limit approach has been selected to evaluate background and downgradient groundwater quality. The upper tolerance method was selected, because it will support an examination of groundwater quality over time, regardless of the size of the data set. Meaning, a larger dataset and a smaller dataset with similar characteristics, should have a similar tolerance limits over time. In addition, constituents exceeding the groundwater protection standard, will likely result from conditions originating from the CCR unit, not a change in the size of the data set. Using this approach, an upper tolerance limit for each constituent will be established from the background data distribution and compared to the level of each constituent in each compliance well, as specified in the *Final Rule*. The following provides the required narrative discussion of the selected statistical approach. It is taken directly from Section 5.0 of the *Sampling and Analysis Plan & Well Documentation Ash Landfill Hunter Power Plant, Castle Dale, Utah*.

> *Groundwater quality will be assessed using upper tolerance limits (UTLs) by comparing upgradient/background groundwater concentrations for Appendix III and Appendix IV constituents, with individual downgradient groundwater wells. The data measured from the upgradient/background wells, will be used to compute a UTL which will serve as the groundwater protection standards / background values. Data obtained from the downgradient wells will be compared individually to the UTLs to determine if the site complies with the Final Rule. The software package Sanitas© v.2016, will be used to compute the UTLs and perform the comparisons.*
>
> *However, during the period of the groundwater monitoring program, if a comparable statistical software program is or may become available or Sanitas is updated, PacifiCorp reserves the right to change software packages. In addition, during the period of the groundwater monitoring program, if it becomes apparent a change in the statistical analysis method(s) is warranted, PacifiCorp reserves the right to use any other statistical analysis method allowed under the Final Rule. If the statistical analysis method(s) are updated, the certification will be revised to reflect the change in statistical method.*
>
> *As part of this evaluation, groundwater data will be examined for characteristics that impact how the UTL is computed. These characteristics include the:*
>
> - *Number of non-detect results,*
> - *Data distribution, and*
> - *Site-wide false-positive rate (SWFPR)*
> - *Spatial and seasonal variability.*
>
> *Summary statistics and other statistical characteristics of the data will be examined for completeness. Each of these characteristics are described below.*
>
> ***Non-Detects.*** *The majority of datasets contain non-detect values. These values complicate statistical analysis, because non-detect results are reported as being less than a reporting or detection limit (e.g. < 0.010 mg/L) rather than a fixed concentration. Calculations that incorporate these values are non-trivial. The EPA Unified Guidance (USEPA 2009) recommends using one-half of the detection limit in place of the non-detect values if the data have fewer than 15% non-detects. The method detection limit (MDL) is used in this situation. If the data contain more than 15% non-detects, but fewer than 50%, the Kaplan-Meier estimator*

will be used to compute estimated values for non-detects. If more than 50% of the data points are non-detects, a non-parametric UTL will be used. The non-parametric UTL is the maximum observed value from the upgradient / background wells.

Data Distribution. The shape or distribution of the data will be assessed prior to computing unit-specific UTL values. To support this evaluation, histograms and normal-quantile plots will be developed for each of the constituents. They will be used to determine if the data set contains statistical outliers. Outliers will be examined and addressed on a case-by-case basis to ensure that appropriate action is taken. The Shapiro-Wilk test will also be used to assess normality. If the p-value associated with the test is greater than or equal to 0.05, the data will be considered normally distributed and a parametric UTL will be computed using the background measurements. If the p-value is less than 0.05, then a non-parametric UTL will be computed using the background measurements. The parametric UTL is computed using the formula below:

$UTL = \bar{x} + K \times S$

Where:

$\bar{x}$ = the average of the background data
K = multiplier from Tables 13-1 thru 1-18 (Appendix B, EPA 1989)
S = standard deviation of the background data

If the data are not normally distributed, the ladder of powers method will be used to determine if a reasonable transformation can be found that will produce normal data. The ladder of powers tests different monotonic transformations of the data, such as the natural logarithm or square, to see if the transformed data have a normal distribution. If a transformation within the ladder of powers can be found that produces normal data, a parametric UTL will be computed using the transformed data. If a transformation is identified, it will be applied to both upgradient / background and downgradient groundwater data prior to comparison. A non-parametric UTL will be computed for data that are not normally distributed and cannot be transformed to be normal. The non-parametric UTL is equivalent to the largest value measured in the upgradient / background wells.

Statistical Hypothesis Testing. The UTL is a statistical hypothesis test, which means it has both a null and alternative hypothesis associated with it. The null hypothesis for comparison purposes, is the constituent concentration in the downgradient well(s), is less than or equal to the concentration of the same constituent in the upgradient / background well(s). The alternative hypothesis is the constituent concentration is greater in the downgradient well(s), than in the background well(s).

Each hypothesis test has a significance level, α, that must be determined. The significance level is the chance that a Type I decision error is made. In the case of groundwater monitoring, a Type I error occurs when the downgradient wells are found to have a higher constituent concentration when they are, in fact, within the range of background. Because decisions are made using sampling, the chance of this error occurring cannot be eliminated, but it can be controlled. An α of 0.05 is most commonly used with groundwater monitoring hypothesis tests.

However, as with most groundwater monitoring studies, many hypothesis tests are completed. An α of 0.05 indicates that if 20 hypothesis tests are done on downgradient wells, and all are below upgradient / background concentrations, it is expected by random chance, one of the values will erroneously conclude that the downgradient wells have a higher constituent concentration than the background wells. This particular study consists of testing 21 different constituents for several downgradient wells. This means approximately 100 tests will be done for each site. Thus, α must be adjusted to reduce the chance of a Type I error to an acceptable level. The SWFPR is the chance that a Type I error occurs at least once when all of the tests are performed. It can be used to select an appropriate α for each of the individual tests. EPA recommends using a SWFPR of no greater than 0.10 (USEPA 2009). The formula for computing the SWFPR is:

$$\mathbf{SWFPR} = 1 - (1 - \alpha)^{CW}$$

Where:
α = the significance level for each individual hypothesis test
c = the number of constituents with at least one detected value
w = the number of downgradient wells at the site

The SWFPR will be set at a level close to 0.10 and then an appropriate α will be computed. This cannot be determined until all of the data are available, because statistical tests will not be performed for constituents that are not detected in the downgradient wells. Thus, they should not be included in the computation of α. The software package Sanitas© v.2016 will be used to select the α value to control the SWFPR.

Seasonal and Spatial Variability*. Data will be examined for seasonal and spatial variability. If either is found in the data, the variability will be controlled and corrected to ensure that the UTL comparison is appropriate. Sanitas© v.2016 will be used to test for seasonal and spatial variability as well as, provide the necessary corrections when computing the UTLs or performing other statistical computations.*

Data Evaluation. *Once complete data sets have been compiled, and data has undergone a full evaluation, modifications to the data sets may be necessary, to ensure the most representative monitoring wells are being utilized to assess groundwater quality. These changes may include removal of an upgradient / background or downgradient well data from the statistical examination, due to measured conditions that do not represent the uppermost aquifer. If these circumstances arise during evaluation of CCR unit data, the reasoning for modifications will be fully discussed in the annual monitoring report.*