REPORT

ACTIVE COAL COMBUSTION RESIDUALS LANDFILL GROUNDWATER STATISTICAL METHOD CERTIFICATION

Escalante Generating Station

Prewitt, New Mexico

Submitted To: Tri-State Generation and Transmission Association, Inc.
Post Office Box 33695
Denver, Colorado 80233

Submitted By: Golder Associates Inc.
44 Union Boulevard, Suite 300
Lakewood, Colorado 80228

October 13, 2017 **1783558**

Table of Contents

List of Tables

1.0 INTRODUCTION

Golder Associates Inc. (Golder) has prepared this report on behalf of Tri-State Generation and Transmission Association, Inc. (Tri-State) to certify that the statistical methodology to be employed in the evaluation of groundwater monitoring data for the active coal combustion residuals (CCR) landfill at Tri-State's Escalante Generating Station is appropriate for this purpose, as required under 40 CFR 257.93(f)(6). A narrative description of the statistical method that has been selected for evaluating the groundwater monitoring data is also included.

2.0 BACKGROUND INFORMATION

2.1 Facility

Tri-State owns and operates the Escalante Generating Station, a 273-megawatt coal-fired electric generating plant located near Prewitt, New Mexico. Fly ash, bottom ash, and flue gas desulfurization solids (scrubber solids) generated at Escalante Generation Station are deposited in the site's existing active CCR landfill (the Facility).

2.2 Constituents

Baseline samples (a minimum of eight) were collected between September 2016 and August 2017 and analyzed for the constituents listed in Appendix III and Appendix IV to 40 CFR 257 (Table 1). Additionally, the following field parameters were recorded: pH, specific conductivity, and temperature. Semi-annual Detection Monitoring samples will be analyzed for the constituents on the Appendix III constituent list. If Assessment Monitoring is required, semi-annual samples will also be analyzed for the constituents on the Appendix IV constituent list. Metals will be analyzed as total recoverable metals (i.e., the samples will not be filtered).

The statistical analyses described in this report will be performed for each of the constituents on the Appendix III constituent list during Detection Monitoring and for each of the constituents on the Appendix IV constituent list if Assessment Monitoring is required. With the exception of field pH, statistical analyses on field parameters are not required and will not be conducted.

Table 1. Constituent Lists

Appendix III Constituents
Boron
Calcium
Chloride
Fluoride
pH
Sulfate
Total Dissolved Solids (TDS)

Appendix IV Constituents
Antimony
Arsenic
Barium
Beryllium
Cadmium
Chromium
Cobalt
Fluoride
Lead
Lithium
Mercury
Molybdenum
Radium 226 and 228 combined
Selenium
Thallium

3.0 STATISTICAL METHODOLOGY

The purpose of groundwater monitoring under 40 CFR 257 is to identify whether CCRs contained within the regulated CCR unit are impacting groundwater. To help identify whether an impact has occurred, groundwater data from each monitoring well will undergo statistical analysis. The statistical methodology outlined within this document was selected in accordance with 40 CFR 257.93 and the *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Unified Guidance* (Unified Guidance) (EPA 2009). Additional consideration was given to recommendations provided by the Electric Power Research Institute (EPRI 2015). The statistical methodology is subject to periodic review and revision, with corresponding updates to this document as appropriate.

Baseline diagnostic tests will be conducted using the baseline samples that were collected at each well between September 2016 and August 2017. Statistical analysis will be conducted using MANAGES (EPRI 2017) or similar statistical analysis software.

During Detection Monitoring, comparative statistical analysis will be performed after each sampling event to identify whether concentrations are potentially statistically significant, with the results reported annually. In the event of Assessment Monitoring, comparative statistical analysis will be conducted following the collection of the specified number of confirmation samples discussed for each analysis. Additionally, the baseline data for constituents, as determined by the groundwater monitoring program, will be reviewed periodically (approximately every two to three years based on a semi-annual sampling frequency, or every four to eight samples) to determine if recent results that are not statistically significant can be incorporated into an updated baseline period (specifics of the baseline update procedure are discussed in depth in Section 3.2.6).

3.1 Intra-Well Methodology

40 CFR 257 allows for the use of a series of statistical methods to accurately assess groundwater monitoring data. Intra-well (comparing compliance data to baseline data within a single well) statistics have been identified as an appropriate method for groundwater monitoring at the Facility. The following information provides evidentiary support of this selection in relation to the Unified Guidance (EPA 2009):

- Intra-well methodology allows the establishment of reasonable baseline periods for future tests in wells that may exhibit impacts from other activities at the site and/or historic CCR management. Additionally, measurements occur solely between the baseline period and compliance period of testing within a single well. Through the use of intra-well statistics, changes in water chemistry can be more readily tracked in relation to current CCR handling and containment practices.
- Confounding results stemming from spatial variability can be eliminated.

3.2 Methodology for Baseline Diagnostic Tests

3.2.1 Initial Data Review and Non-detect Handling

Initially, groundwater monitoring data will be plotted on time-series graphs to assess the temporal variability of the data and to visually screen for potential outliers. Temporal variability can be caused by seasonality, changes to the monitored system, changes to the analytical method, recalibration of instruments, and anomalies in the sampling method (EPA 2009).

Non-detect values (ND) are results where the constituent is not detected at a concentration above the Practical Quantitation Limit (PQL) or Reporting Limit (RL) (depending on the nomenclature used by the analytical laboratory). The PQL or RL is the lowest concentration that can be reliably achieved within the specified limits of precision and accuracy during routine laboratory operating conditions. ND values will be managed within the other statistical procedures discussed in this document following this approach:

- For sample populations with < 15% ND, direct substitution of the NDs with ½ the PQL or RL will be employed.
- For sample populations 15% < ND ≤ 50%, the Kaplan-Meier approach will be used to estimate the mean and standard deviation of the population.
- For sample populations with ≥ 50% ND, non-parametric approaches will be used to assess the data, with direct substitution of the NDs with the PQL or RL.

3.2.2 Data Distribution

Parametric statistical tests are based on the assumption that the data are normally distributed or can be transformed to a normal distribution. The distribution of the data will be tested for normality using the Shapiro-Wilk normality test with a 95-percent confidence level (or the Shapiro-Francia test when there are more than 50 results within the dataset). Each constituent from each well will be analyzed separately. Datasets found to be non-normal will be tested for log-normal distribution. When possible, parametric statistical methods will be performed on the normally or log-normally distributed data. Non-parametric statistical methods will be used for datasets that do not exhibit normal or log-normal distribution.

3.2.3 Outlier Analysis

In accordance with the Unified Guidance (EPA 2009), data points will be identified as outliers if the value was an “extreme, unusual-looking measurement” and “inconsistent with the distribution of the remaining measurements.” The Unified Guidance (EPA 2009) recommends testing for outliers within baseline data, but cautions against removal of outliers, unless a likely error or specific discrepancy can be identified, such as recordkeeping errors, unusual sampling and laboratory procedures or conditions, inconsistent sample turbidity, or values significantly outside the range of other results. In accordance with the Unified Guidance (EPA 2009), apparent outliers will be periodically revisited even if initially removed, due to the propensity of groundwater chemistry to change over time.

Outliers will be evaluated and identified through visual inspection and the EPA-recommended Dixon's Test for statistical outliers (or Rosner's Test when there are greater than 25 samples). Dixon's and Rosner's Tests assume that all data values, except the suspected outlier, are normally distributed or log-normally distributed. Consequently, visual inspection of concentrations over time is important in screening for outliers. The effect of removing outliers from the baseline data will usually be to lower the statistical limit (due to a reduction in the standard deviation), resulting in a more conservative statistical limit and improving the odds of detecting increasing concentration levels.

Outliers will be managed as follows:

- Any suspected outlier identified through statistical analysis or visual methods will be reviewed (i.e., through evaluation of the associated analytical report, laboratory narrative, associated laboratory quality assurance/quality control information, and/or field notes) before removal from the dataset to determine if any systematic or systemic errors may have been responsible for the noted anomalous readings. Rejected data points will not be included in the baseline dataset.
- The rationale for the removal of outliers will be documented. The majority of the outliers will likely be isolated values that can be attributed to inconsistent sampling or analytical chemistry methodology resulting in laboratory contamination or other anomalies.
- If an outlier is removed, the normality test (Section 3.2.2) will be re-run to determine if the dataset is normally distributed or log-normally distributed without the outlier.

3.2.4 Trend Analysis

Most statistical tests assume concentrations do not demonstrate temporal correlation. The Sen's Slope methodology is an intra-well statistical analysis of increases or decreases in measured concentrations over time measured by calculating the slope of the linear relationship of concentration and time. The Sen's Slope methodology is paired with the Mann-Kendall test to determine the statistical significance of the calculated Sen's Slope. The methodology involves examining all possible pairs of measurements in the dataset and scoring each pair to determine if a trend exists. The test will be conducted using a target confidence level of 99 percent.

If temporal trends are identified within the dataset, the data will be adjusted to account for the trends, the time period used for the baseline statistics will be reassessed, and/or an alternative statistical method will be used to establish a limit.

3.2.5 Seasonality

Seasonal temporal variability can mask changes in groundwater chemistry. Time-series plots will be observed for visual signs of seasonality, and once enough data has been collected within the program, the data will be evaluated for seasonal variations using the Kruskal-Wallis test. The Kruskal-Wallis test is an intra-well evaluation performed individually for each constituent at each well and requires at least three occurrences of each prospective season with results to calculate the statistic. Datasets that are found to have seasonality will be de-seasonalized for subsequent statistical analysis.

3.2.6 Updating the Baseline Period

The Unified Guidance (EPA 2009) recommends updating the baseline period every two to three years when the sampling frequency is semi-annual, or every four to eight samples. A baseline update will include a review of any revisions to federal and state regulations and EPA statistical guidance documents that may have been promulgated since the previous baseline statistical analysis. The baseline period for a specific constituent will not be updated if verified statistically significant increases (SSIs) have been identified for that constituent, unless a demonstration has been made that the SSI is not related to a release from the Facility.

Prior to inclusion of more recent data in an updated baseline period, a Wilcoxon Rank-Sum test will be conducted. The Wilcoxon Rank-Sum test, also referred to as the Mann-Whitney test, determines if measurements from one population are statistically significantly higher or lower than another population. The test is non-parametric, in that the data being tested is not assumed to fit a specific distribution, such as a normal distribution. When the baseline period is updated in the future, the Wilcoxon Rank-Sum test will be used to compare data from the current baseline period with the more recent data that is intended to be reclassified and included in the updated baseline period. The test will be conducted at a 95-percent confidence level. If the two datasets are drawn from the same population, then the results of the test support updating the prior baseline dataset with the recent data. After the new data is incorporated into the dataset, the baseline diagnostic tests outlined in Section 3.2 will be conducted.

If the Wilcoxon Rank-Sum test detects a significant difference between two sample populations, additional data review will be necessary. The data will be reviewed to determine whether a gradual trend or other change not stemming from a release from the Facility has occurred that was not detected during comparative statistical analysis. At the time of the baseline update, some earlier baseline data might need to be removed from the updated baseline period. Removal of earlier baseline data ensures that future statistical analysis is based on current groundwater conditions, rather than on outdated measurements of groundwater chemistry. Alternatively, outliers identified in the previous baseline period (as described in Section 3.2.3) will be reincorporated into the dataset and reevaluated as potential outliers during the baseline update, unless the outlier(s) were removed due to sampling, laboratory, or other determinant error.

4.0 DETECTION MONITORING

4.1 Statistical Limits

Either a parametric or non-parametric method will be used to generate the baseline statistical limit for each constituent. The statistical method will vary between constituents and will be selected based on the percentage of ND values in the baseline period and the baseline data distribution for each constituent at each well.

For those constituent-well pairs where concentrations of a given analyte are normally or log-normally distributed and have equal to or greater than 50% detections, a parametric prediction limit for future values will be established. Parametric prediction limits will be calculated using a 95% confidence level and the number of future comparisons will be set to eight. Where the concentrations of a given constituent-well pair are not normally or log-normally distributed, or have less than or equal to 50% detections, a non-parametric prediction limit will be used. The non-parametric limit will be assigned as the highest detected value (excluding outliers) or the highest PQL or RL, whichever is greater.

In the case of increasing trends within the baseline period, the data will be adjusted to account for the trends where a source other than the Facility can be identified and/or an alternative statistical method will be used to establish a limit. In the case of downward trends, the trend test will serve as the alternative method and a test for trends will be conducted on the eight most recent data points. When the trend stabilizes, a statistical limit will be established based on the recent data.

4.2 Comparative Statistical Analysis

Comparative statistical analysis will be conducted following each sampling event for Detection Monitoring. For both parametric and non-parametric prediction limits, the comparative statistical analysis will consist of a comparison of Detection Monitoring results (the recent analytical results for each sampling event performed after the baseline data period) to the statistical limit calculated from the baseline data.

The following definitions will be used in discussion of the comparative statistical analysis:

- SSI – is a statistically significant increase (SSI) and is defined as an analytical result that exceeds the parametric or non-parametric statistical limit established by the baseline statistical analysis.
- False-positive SSI – is defined as an analytical result that exceeds the statistical limit but can clearly be attributed to laboratory error or changes in analytical precision, or is invalidated through confirmatory resampling.
- Confirmatory resampling – is designated as the resampling event that occurs within 90 days of detecting an SSI over the statistical limit for determination of a verified SSI.
- Verified SSI – is interpreted as two consecutive SSIs (the original sample and the confirmatory resample for analytical results) for the same constituent at the same well.

The Detection Monitoring program has been developed to identify potential SSIs over baseline values for the Appendix III constituents. This determination will be made within 90 days of receiving the finalized laboratory analytical report(s) for each sampling event and completion of data quality review as necessary to address questions concerning the validity of sampling methods or laboratory analyses. A potential SSI will not be considered a verified SSI until confirmatory resampling is performed.

4.3 Alternative Source Demonstrations and Assessment Monitoring

If a verified SSI is identified in a downgradient well for an Appendix III constituent as part of the Detection Monitoring program, Tri-State will establish an Assessment Monitoring program meeting the requirements of 40 CFR 257.95. Alternatively, Tri-State may demonstrate that a source other than the regulated CCR unit caused the SSI or that the SSI was a result of an error in sampling, analysis, statistical evaluation, or natural variability in the groundwater quality that failed to be captured during baseline data collection. A report documenting the alternative source demonstration (ASD) will be certified by a Professional Engineer registered in the state of New Mexico and placed in the Facility's Operating Record within 90 days of the SSI determination. If a successful ASD is made and documented, Tri-State will continue with the Detection Monitoring program. If, after 90 days, a successful demonstration is not made, Tri-State will initiate an Assessment Monitoring program as described in Section 5.0.

5.0 ASSESSMENT MONITORING

An Assessment Monitoring program will be initiated in the event of a verified SSI of an Appendix III constituent in a downgradient well, unless a successful demonstration is made that an alternative source affected the groundwater chemistry or that the SSI was a result of an error in sampling, analysis, statistical evaluation, or natural variability in the groundwater quality that was not captured during the baseline data collection period.

Many of the statistical comparisons used in Assessment Monitoring require baseline analyses to be completed prior to comparative statistics. Baseline statistics for Assessment Monitoring mirror those for Detection Monitoring and are discussed in further detail in Section 3.2.

5.1 Establishment of Groundwater Protection Standards

The Assessment Monitoring program will comply with 40 CFR 257.95. Within 90 days of triggering the Assessment Monitoring program, Tri-State will sample and analyze each well for the Appendix IV constituents listed in Table 1. A groundwater protection standard (GWPS) will be established for each Appendix IV constituent. The GWPS will fall within one of the following categories:

- For constituents for which a maximum contaminant level (MCL) has been established by EPA (40 CFR 141), the MCL for that constituent will be the GWPS.
- For constituents for which an MCL has not been established by EPA, the statistical baseline concentration for the constituent at the well will be the GWPS.
- For constituents where the statistical baseline level is higher than the MCL, the statistical baseline concentration will be the GWPS.

For those constituents without MCLs or where the statistical baseline level is higher than the MCL, the GWPS will be determined through statistical methods. For those constituents with a normal or log-normal distribution, prediction limits for future means will be developed. This method is similar to prediction limits for future values used in Detection Monitoring; however, the test is designed for a mean of the future values to be compared to the statistical limit instead of individual results. For constituents that have non-normally distributed data or a high percentage of non-detects, prediction limits for future medians will be developed. For both types, the GWPS is considered a one-sided upper prediction limit calculated from the baseline data.

5.2 Comparative Statistics

To determine if Appendix IV constituents have statistically exceeded the associated GWPS, the approaches discussed in this section will be used.

5.2.1 Maximum Contaminant Level (MCL)-Based GWPS

For those constituents with MCLs as the GWPS, a confidence interval approach will be used. Per recommendations provided by the Unified Guidance (EPA 2009) and detailed by EPRI (2015), a confidence interval statistically defines the upper and lower bound (the upper and lower confidence limit) of the true mean of a population. The Unified Guidance (EPA 2009) recommends confidence intervals for Assessment Monitoring. Confidence intervals identify SSIs through comparison against a fixed standard, namely the MCL-based GWPS.

To calculate a confidence interval, at least four samples are required. A confidence interval will only be considered statistically above the associated GWPS if both the upper and lower confidence limits exceed the GWPS.

5.2.2 Baseline-Based GWPS

For those constituents with a GWPS based on baseline concentrations, a prediction interval approach will be used. The Unified Guidance (EPA 2009) recommends prediction intervals for a future mean for data that follows a normal or log-normal distribution and prediction intervals for a future median for data with high percentages of non-detects or non-normally distributed data.

When using prediction intervals for a GWPS, a one-sided prediction interval is calculated using baseline datasets. For parametric comparisons, four measurements are averaged to compare to the GWPS, with statistical significance occurring when the mean is greater than the GWPS. For non-parametric comparisons, the median of three measurements is compared to the GWPS, with statistical significance occurring when the calculated comparative median is greater than the GWPS.

5.3 Returning to Detection Monitoring, Alternative Source Demonstrations, and Assessment of Corrective Measures

If the concentrations of constituents listed in Appendix III and Appendix IV are shown to be at or below baseline values for two consecutive sampling events, Tri-State will return to Detection Monitoring. If the concentrations of any of the Appendix III and Appendix IV constituents are above baseline values, but the Appendix IV constituents are below the established GWPS, Tri-State will continue with Assessment Monitoring. If one or more Appendix IV constituents are detected at statistically significant levels above the established GWPS, Tri-State will place a notification of the exceedance in the Facility's Operating Record and follow the course of action outlined in 40 CFR 257.95(g) and as determined by Tri-State.

If, within 90 days of detecting an Appendix IV constituent at statistically significant levels above the GWPS, a successful demonstration has not been made to indicate an alternative source or that the SSI was a result of an error in sampling, analysis, statistical evaluation, or natural variability in the groundwater quality that was not captured during the baseline data collection period, Tri-State will initiate an assessment of corrective measures in accordance with 40 CFR 257.95(g). The assessment of corrective measures will be completed within 90 days, unless a demonstration is made that additional time is needed to complete the assessment.

6.0 CERTIFICATION

The undersigned Professional Engineer registered in New Mexico certifies that the statistical methodology described in this report is appropriate for evaluation of groundwater monitoring data for the active CCR landfill at Tri-State's Escalante Generating Station, as required under 40 CFR 257.93(f)(6).

THOMAS L. RUTKOWS
NEW MEX
18475
Exp. 12/17
LICENSED PROFESSIONAL ENGINEER

7.0 REFERENCES

Electric Power Research Institute (EPRI), 2015. Groundwater Monitoring Guidance for the Coal Combustion Residuals Rule – 2015 Technical Report. November 2015.

Electric Power Research Institute (EPRI), 2017. MANAGES, Version 4.0. September 2017.

United States Environmental Protection Agency (EPA), 2009. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance*, Office of Resource Conservation and Recovery, EPA-R-09-007, March 2009.

United States Environmental Protection Agency (EPA), 2015. Code of Federal Regulations Title 40 Part 257: Hazardous and Solid Waste Management System; *Disposal of Coal Combustion Residuals from Electric Utilities*. April 17, 2015.

Established in 1960, Golder Associates is a global, employee-owned organization that helps clients find sustainable solutions to the challenges of finite resources, energy and water supply and management, waste management, urbanization, and climate change. We provide a wide range of independent consulting, design, and construction services in our specialist areas of earth, environment, and energy. By building strong relationships and meeting the needs of clients, our people have created one of the most trusted professional services organizations in the world.

Africa	+ 27 11 254 4800
Asia	+ 852 2562 3658
Australasia	+ 61 3 8862 3500
Europe	+ 356 21 42 30 20
North America	+ 1 800 275 3281
South America	+ 56 2 2616 2000

solutions@golder.com
www.golder.com

Golder Associates Inc.
44 Union Boulevard, Suite 300
Lakewood, Colorado 80228
Tel: (303) 980-0540
Fax: (303) 985-2080

Golder Associates

Engineering Earth's Development, Preserving Earth's Integrity