April 17, 2019

Ms. Shannon Bennett
Tennessee Valley Authority
Generation Construction Projects
1101 Market Street
Chattanooga, Tennessee 37402

Subject: Statistical Methods Certification
TVA CCR Rule Groundwater Quality Monitoring Program
John Sevier Fossil (JSF) Plant
Bottom Ash Pond
Rogersville, Hawkins County, Tennessee
Terracon Project Number 8618P185

Dear Ms. Bennett:

Terracon Consultants, Inc. (Terracon) has reviewed the *Statistical Methods Certification for Compliance with the Final Coal Combustion Residuals Rule (40 CFR §257.93)* prepared by Dr. Kirk Cameron, MacStat Consulting, Ltd, for application at the above referenced facility and monitored unit. Based upon our review of this document it is our opinion that, to the best of our knowledge, information, and belief:

1. The information contained in this certification is prepared in accordance with the accepted practice of engineering;
2. The information contained therein is accurate as of the date of this letter; and,
3. The selected statistical methods are appropriate for evaluating the groundwater monitoring data for the referenced CCR management area at the John Sevier Fossil Plant in Hawkins County, Tennessee and that the referenced methods meet the requirements described in 40 CFR 257.93.

We appreciate the opportunity to be of service to TVA in this matter. If we can be of further assistance please contact us.

Sincerely,
Terracon Consultants, Inc.

George K. Flores, P.E.
Program QA/QC

GEORGE KEVIN FLORES
REGISTERED ENGINEER
AGRICULTURE
COMMERCE
No. 107336
STATE OF TENNESSEE

4/17/19

Dallas E. Whitmill, P.E.
Program Manager

Tennessee Valley Authority 1101 Market Street Chattanooga, Tennessee 37402

Statistical Methods Certification
for Compliance with the Final Coal Combustion Residuals Rule (40 CFR §257.93)

John Sevier Fossil Plant

CCR Groundwater Monitoring Network

1. Introduction

The U.S. Environmental Protection Agency's (USEPA's) final Coal Combustion Residuals (CCR) Rule establishes a comprehensive set of requirements for the management and disposal of coal ash in landfills and surface impoundments by electric utilities. The Tennessee Valley Authority (TVA) John Sevier Fossil Plant (JSF), located in Rogersville in Hawkins County, Tennessee, Tennessee, has a CCR unit that is subject to the CCR Rule: Bottom Ash Pond.

This report includes a summary of the statistical methodology selected for evaluating groundwater monitoring data at the above mentioned CCR unit and supports compliance with requirements outlined in Sections 257.93(f) and 257.93(g) of the CCR Rule. To develop the most appropriate methods to validate assumptions, evaluate groundwater data, and develop background concentrations, the statistical methodology is generally based on USEPA's 'Unified Guidance' (2009).

The statistical methods for evaluating groundwater data in detection monitoring described in this document are consistent with method/paragraph (3) of Section 257.93(f), which includes a prediction interval or control chart procedure. In assessment monitoring or corrective action, the methods described in **Section 4** of this document — confidence intervals and confidence bands — are consistent with Unified Guidance recommendations, and are also justified under method/paragraph (5) of Section 257.93(f), namely "Another statistical method that meets the performance standards of paragraph (g) of this section."

Groundwater monitoring activities commenced in January 2019 and, at the time of this report, TVA contractors obtained the minimally prescribed number of samples (i.e., “eight independent samples for each background and downgradient well”) to comply with the initial baseline requirements included in §257.90(b) of the CCR Rule. Results from the background wells will be pooled when interwell testing is used for the statistical methods; otherwise, baseline data from each compliance well will serve as well-specific background in the event that intrawell methods are appropriate. The groundwater monitoring dataset will continue to be refined as additional data becomes available throughout the life of the CCR Rule groundwater monitoring program.

As the chosen statistical methodology is applied to groundwater quality data for determining background concentrations and evaluating compliance well sampling results, revisions to the statistical methods may be warranted to ensure the groundwater monitoring data is evaluated appropriately. TVA reserves the right to use any other statistical test(s) that, as allowed by the CCR Rule, would meet the performance standards established by §257.93(g) of the CCR Rule. If the statistical analysis method(s) described herein are subsequently revised, TVA will submit a Management of Change form into its Operating Record along with the revised copy of the backup documentation followed by its posting onto the corresponding TVA-managed CCR Rule web site. Pursuant to the CCR Rule requirements, the enclosed Professional Engineer’s certification record will also be revised to conform to the description of the updated statistical methods and data set modifications.

2. Development of Background

2.1 Interwell vs. Intrawell

When data from multiple background wells are available, a determination will be made as to whether the background data appear to come from the same population or whether there is evidence of statistically significant spatial variation at the facility. Data for each constituent will be plotted using box plots to assist in this determination, allowing concentrations within and across wells to be visualized. Analysis of Variance (ANOVA) will be utilized to statistically evaluate whether or not spatial variation is statistically significant.

Conventionally, interwell statistical tests are used to evaluate whether or not compliance wells are consistent with, and in the expected range of, background. These tests are generally appropriate when there is no significant spatial variation at the site, and the natural groundwater gradient flows from the background wells to the compliance locations. In the event of significant spatial variation among the background wells, it may be reasonable to assume similar variation among the compliance wells, independent of any groundwater contamination. Under such conditions, it may be difficult to make valid interwell comparisons between compliance wells and background locations, since apparent differences may reflect natural spatial variability rather than evidence of groundwater contamination.

As an alternative, USEPA's Unified Guidance recommends switching from interwell methods to intrawell methods when it can be reasonably demonstrated that no pre-existing contamination from current practices or waste management at the regulated facility is present. More generally, intrawell methods may also be needed when there is insufficient data from (upgradient) background wells or when interwell methods will not adequately address the question of a change in groundwater quality at compliance locations. The latter can occur, for instance, when the uppermost aquifer underlying a site is discontinuous, or when compliance wells are screened in different hydrostratigraphic zones.

Intrawell tests compare the most recent sample(s) from a given well to historical measurements at the same well, rapidly detecting changes over time at a given location. When appropriate, intrawell methods remove the confounding factors of spatial variation in well-to-well concentration levels. In these cases, EPA recommends intrawell methods, such as intrawell prediction limits with retesting, as an acceptable alternative to interwell testing.

The overarching goals in selecting either interwell or intrawell testing will be to:

- Ensure that statistical comparisons will be adequately sensitive to detecting a facility release;
- Ensure that data used in testing reflect current background conditions; and
- Avoid confusing an impact caused by a release from the facility with a difference between wells caused by heterogeneous subsurface conditions.

2.2 Background Screening

Credible and adequate background data is the most important aspect to developing accurate and sensitive statistical limits. Standard parametric prediction and control chart limits for groundwater assume that the background data (1) are representative of current background conditions; (2) are statistically stable over time (i.e., not trending);

(3) do not include (extreme) outliers; (4) can be normalized, possibly via transformation; and (5) are sufficient in size to accurately estimate the variability in the underlying groundwater population, and thus be sensitive to a persistent change in groundwater concentrations. Nonparametric prediction limits do not make assumption (4), but otherwise have the same requirements as parametric limits.

To test these assumptions, any proposed background data will be screened prior to constructing statistical limits. Time series plots and formal trend tests will be used to check stability; box plots and formal outlier tests to identify and check for outliers. The statistical pattern of the data along with the history and hydrogeology of the site will be used to gauge how well the data mimic current background conditions.

If average background concentration levels are changing over time (i.e., trending), the prospective background data may need to be truncated, removing older data to ensure that the resulting limits continue to represent current natural conditions. Confirmed outliers will be flagged and de-selected from prospective background data prior to establishing statistical limits. Any values flagged as outliers will be summarized in periodic reporting.

Probability plots and normality tests, adjusted for the presence of non-detects (Cameron, 2017), if any, will be used to identify and test best-fitting distributional models for the background data. If the data can be fit to a normal distribution (i.e., 'normalized') — possibly via mathematical transformation — then a parametric prediction limit or control chart will be constructed. If the data cannot be normalized, a nonparametric prediction limit will be constructed instead.

The size of the background dataset impacts both the accuracy (false positive rate) and sensitivity (statistical power) associated with a prediction limit or control chart comparison. The CCR rule requires at least 8 baseline observations prior to the start of monitoring, but often more background data is needed to meet EPA performance requirements for groundwater tests, especially at larger well networks. These requirements are discussed below (**Section 3.1**).

2.3 Screening Intrawell Background

For intrawell statistical limits, background data for each constituent-well pair will be designated from the historical/past data for that pair. Because the background data in this case derive from a compliance location, it is necessary to first demonstrate that there have been no suspected prior impacts at the compliance well that might distort the prospective 'background' data or be unrepresentative of natural background conditions at that location. If intrawell analyses are not feasible due to elevated concentrations in compliance wells relative to concentrations upgradient of the facility, interwell analyses will be initially utilized until or unless further evidence supports the use of intrawell testing.

The data screening procedure is designed to check for stable and representative background conditions, and to account for existing groundwater quality as a result of previous and/or present activities in the area. Once waste is placed, ongoing groundwater monitoring is designed to quickly identify a change in groundwater quality. While pre-waste data is ideal for characterization of groundwater quality prior to waste placement, many facilities do not have such information. The screening process is used, then, to characterize data either upgradient of the facility and/or collected during the baseline sampling period, as well as to identify whether possible contamination is already present at compliance points.

2.4 Periodic Updating of Background

Background data will be updated for interwell statistical limits by consolidating more recent sampling observations with historical background data after every two to four new sampling events. Any new outliers in the combined background data will be flagged and, if confirmed, removed prior to construction of statistical limits. This updating process will not only increase the background sample size, but will also reduce the incidence of false positives when using nonparametric prediction limits and increase the statistical power of parametric prediction or control chart limits.

When using intrawell methods, periodic updating of background data will be considered only after at least four new observations are available. Data will be tested for suitability of updating by statistically comparing the new sampling observations with the screened background data. To help verify that the most recent observations represent non-impacted groundwater, a two-sample test such as the Wilcoxon Rank Sum will be conducted, comparing the current background to the more recent observations. If the most recent data are not found to be statistically different from the older data, background may be updated by consolidating the newer observations with the current background.

3. Detection Monitoring Tests

Prediction limits and control charts are recommended by USEPA as primary techniques for detection monitoring. Prediction limits and control chart limits are statistical thresholds estimated from background. If any new compliance observations exceed the upper prediction or control chart limit, a potential statistical exceedance will be flagged. Retesting will then be conducted by collecting one or more independent resamples of the same well-constituent pair to confirm or disconfirm the initial exceedance. Any confirmed exceedance will be recorded as an SSI.

To conduct retesting, the pass one-of-m method, as described in the Unified Guidance (Chapter 19), allows for an efficient plan to confirm or disconfirm a potential SSI over background identified during detection monitoring. Depending on the background sample size, the target site-wide false positive rate, and the available time period in which to collect independent resamples, either a 1-of-2 or 1-of-3 method will be used when retesting is needed.

Under the CCR rule, prediction limit or control chart tests will be implemented for all detected Appendix III parameters. Note that one parameter, pH, will require both upper and lower prediction or control chart limits. In that case, a potential SSI will be flagged whenever a new compliance measurement is either less than the lower statistical limit or higher than the upper statistical limit.

Parameters with all non-detects do not require formal testing, but will be evaluated using USEPA's Double Quantification Rule (DQR). The DQR assumes that a significant change in groundwater quality has occurred whenever two consecutive detections of a parameter are observed after no previous detections. It is similar in nature to a nonparametric prediction limit with a single retest (1-of-2).

3.1 Statistical Performance Requirements

The Unified Guidance recommends two general criteria when designing a statistical detection monitoring program in order to meet RCRA (and, by reference, CCR) statistical performance requirements: (1) an annual site-wide false positive rate (SWFPR) of no more than 10%, and (2) statistical power of a site's 'weakest' test greater than or equal to the minimum benchmark power represented by the EPA reference power curves.

The first criterion informs the accuracy of statistical testing, limiting the occurrence of spurious (false) SSIs. The second criterion guides the sensitivity of testing, ensuring an adequate chance of identifying real changes in groundwater quality. In practical terms, the annual SWFPR is distributed evenly among the total number of well-constituent pairs and among the total number of statistical evaluations per year. Statistical limits will be constructed with sufficient background size and retesting in order not to exceed the per-pair portion of the overall false positive risk. Similarly, site-specific power curves associated with each distinct type of test will be constructed and compared to the EPA reference power curves to ensure adequate statistical power.

The CCR Rule indicates that if an SSI over background is confirmed at the waste boundary for one or more Appendix III constituents during detection monitoring, then the owner or operator of the CCR unit must, within 90 days: 1) establish an assessment monitoring program, 2) demonstrate that a source other than the CCR unit caused the SSI over background, or 3) demonstrate that the SSI over background resulted from error in sampling, analysis, statistical evaluation, or natural variation in groundwater quality. Written documentation must also be completed and certified by a qualified professional engineer within the 90 day timeframe.

4. Assessment Monitoring and Corrective Action

To implement assessment monitoring, the CCR rule requires that all Appendix IV constituents be sampled, with any detected parameters added to the list of parameters sampled semi-annually. To statistically evaluate these parameters, concentration data will be compared to Groundwater Protection Standards (GWPS) through the use of confidence intervals or their variant, confidence bands. The GWPS for each constituent will be established as either the Maximum Contaminant Level (MCL) or a statistical limit based on background if either no MCL is available or background concentrations are higher in concentration than the established MCL. On an annual basis, all Appendix IV parameters must be sampled and newly detected parameters added to the list of parameters sampled semi-annually.

4.1 Confidence Intervals

For each well-constituent pair, a trend test will be run to determine whether there is evidence of a significant trend. If not, a parametric confidence interval around the population mean will be constructed at the 99% confidence level when the compliance data follow a normal distribution, and around the geometric mean (or population median) when the data follow a transformed-normal distribution.

In the event that the data do not pass a normality test and cannot be normalized via a transformation, non-parametric confidence intervals will be constructed. The confidence level associated with a non-parametric interval depends on the number of observations used to construct the interval. When a well-constituent pair does not have sufficient sample size to achieve 99% confidence, a confidence interval with less confidence will be constructed, but updated after each new sampling event until the desired confidence level is reached. The pair will also continue to be reported and tracked using time series plots and/or trend tests until enough data are available.

In assessment monitoring, a well is determined to be out of compliance (SSI) when the lower confidence limit (LCL), and thus the entire interval, exceeds the GWPS, as discussed in USEPA's Unified Guidance. Assessment of corrective action is initiated at that time, with remediation efforts evaluated through the continuing use of confidence intervals and confidence bands to determine remedial effectiveness.

4.2 Confidence Bands

If the compliance data at a given well-constituent pair show evidence of a mildly or highly significant trend, a linear regression line will be fit to the data and a confidence band with 99% confidence will be constructed around the trend line. Confidence bands will only be constructed on pairs with at least 8 independent observations.

To evaluate compliance with regulatory standards, the lower edge of the confidence band at the most recent sampling event will be compared to the GWPS. If the lower edge exceeds the GWPS at that point in time (thus guaranteeing the entire vertical cross-section of the band also exceeds the GWPS at that point), an SSI will be recorded. If the lower edge of the band does not exceed the GWPS, no SSI will have occurred. As new sampling events are collected, the trend estimate will be updated along with the confidence band.

4.3 Corrective Action

If and when corrective action is initiated, this information will be placed in the operating record and, if possible, an alternate source demonstration will be made. If there is evidence of a release or if an alternate demonstration is not made regarding any exceedances of GWPS, efforts will be made to characterize the nature and extent of the release and initiate the assessment of corrective action measures.

Once remediation activities begin, semi-annual sampling will continue and confidence intervals and/or confidence bands will monitor the progress of remediation efforts. Confidence intervals and bands are compared to GWPS or other risk-based criteria to determine when clean-up levels are achieved.

Although in corrective action the same statistical techniques are used, the manner of the comparison is different from that in assessment monitoring. In corrective action a well-constituent pair is declared ‘clean’ when the entire confidence interval or cross-section of the confidence band at the most recent sampling event falls below a specified clean-up limit or GWPS (i.e., the upper confidence limit [UCL] or upper confidence band [UCB] falls below the regulatory limit). Alternatively, compliance is achieved when the lower confidence limit (LCL) or lower confidence band (LCB) for every Appendix IV parameter does not exceed the GWPS for a period of three consecutive years.

5. Certification

I, George K. Flores, being a Professional Engineer in good standing in the State of Tennessee, do hereby certify, to the best of my knowledge, information and belief:

1. that the information contained in this certification is prepared in accordance with the accepted practice of engineering;
2. that the information contained herein is accurate as of the date of my signature below; and
3. that the selected statistical methods in this certification report are appropriate for evaluating the groundwater monitoring data for the CCR management area (Bottom Ash Pond) at the John Sevier Fossil Plant in Hawkins County, Tennessee and that these methods meet the requirements described in 40 CFR 257.93.

6. Bibliography

Cameron, K (2017) ‘On-the-fly’ goodness of fit and outlier testing for left-censored data. In JSM Proceedings, Section on Statistics and the Environment, Alexandria, VA, American Statistical Association, 3445-53.

Gilbert, RD (1987) Statistical Methods for Environmental Pollution Monitoring. Professional Book Series, Van Nos Reinhold.

Helsel, DR and Hirsch, RM (1992) Statistical Methods in Water Resources. Elsevier.

U.S. Environmental Protection Agency (2009) Statistical Analysis of Groundwater Monitoring Data at RCRA facilities: Unified Guidance. USEPA Office of Solid Waste, EPA 530-R-09-007.