ftn Associates Ltd.
water resources / environmental consultants

STATISTICAL ANALYSIS PLAN
PLUM POINT ENERGY STATION

PREPARED IN COMPLIANCE WITH THE
EPA COAL COMBUSTION RESIDUAL RULE
TITLE 40 CODE OF FEDERAL REGULATIONS PART 257

SEPTEMBER 29, 2017

STATISTICAL ANALYSIS PLAN
PLUM POINT ENERGY STATION

PREPARED IN COMPLIANCE WITH THE EPA COAL COMBUSTION RESIDUAL RULE
TITLE 40 CODE OF FEDERAL REGULATIONS PART 257

Prepared for

Plum Point Services Company, LLC
Plum Point Energy Station
2732 South County Road 623
Osceola, AR 72370

Prepared by

FTN Associates, Ltd.
3 Innwood Circle, Suite 220
Little Rock, AR 72211

FTN No. 14590-1624-001

September 29, 2017

PROFESSIONAL ENGINEER'S CERTIFICATION

With this certification, I certify that I am a qualified professional engineer as defined in §257.53 of Title 40 Code of Federal Regulations Part 257, that I have reviewed this statistical analysis plan, and that the statistical methods decribed herein are appropriate and meet the requirements of §257.93 of 40 CFR Part 257.

Dana L. Derrington, Arkansas PE #16372

09-29-2017
Date

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

1.0 INTRODUCTION

Plum Point Services Company, LLC, operates a landfill for the disposal of coal combustion residuals (CCRs) at the Plum Point Energy Station located near Osceola, Arkansas. The landfill is permitted by the Arkansas Department of Environmental Quality (ADEQ) under Permit No. 0303-S3N-R1 and is operated in accordance with Arkansas Pollution Control and Ecology Commission (APCEC) Regulation No. 22 requirements. In April 2015, the Environmental Protection Agency (EPA) published new regulations for the disposal of CCRs under Title 40 Code of Federal Regulation (40 CFR) Part 257. Facilities regulated under this new rule are required to implement a groundwater monitoring program that includes statistically evaluating groundwater quality data to determine if the CCR unit is impacting groundwater quality at the facility's downgradient (compliance) boundary.

This statistical analysis plan (SAP) has been prepared in accordance with the groundwater monitoring requirements of 40 CFR Part 257[1] and the recommendations provided in EPA's *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance* (Unified Guidance) (EPA 2009). This document describes procedures for the statistical treatment of groundwater quality data for detection and assessment monitoring at the landfill, as applicable, and the associated recordkeeping, notification, and publicly accessible internet site requirements. This SAP does not describe procedures for the collection and analysis of groundwater samples. The facility is required to collect groundwater samples in accordance with the applicable requirements of §257.93.

[1] References herein that begin with "§257" refer to sections or subsections of 40 CFR Part 257.

2.0 REGULATORY BACKGROUND

2.1 Analytical Requirements

Groundwater monitoring at the landfill is required for the list of parameters shown in Table 2.1 during detection and assessment monitoring, as applicable, in accordance with 40 CFR Part 257. To facilitate the collection and analysis of samples that are representative of groundwater quality, the facility has a groundwater sampling and analysis plan (FTN 2017) that identifies procedures for sample collection, preservation, and shipment; analytical procedures; chain-of-custody procedures; and quality control procedures in accordance with the applicable requirements of §257.93.

Table 2.1. Constituents for groundwater monitoring, 40 CFR Part 257.

Appendix III to Part 257 – Constituents for Detection Monitoring	
Boron	Sulfate
Calcium	Total dissolved solids (TDS)
Chloride	pH*
Fluoride	
Appendix IV to Part 257 – Constituents for Assessment Monitoring	
Antimony	Lead
Arsenic	Lithium
Barium	Mercury
Beryllium	Molybdenum
Cadmium	Selenium
Chromium	Thallium
Cobalt	Radium 226 and 228 combined
Fluoride	

*Field-measured parameter.

2.2 Sampling Frequency

In accordance with §257.94(b), a minimum of eight independent samples must be collected and analyzed for the list of appendix III and appendix IV constituents at each monitoring well prior to October 17, 2017, at existing CCR landfills. These first eight values comprise the initial background data set for detection or assessment monitoring, as applicable.

Sample independence is discussed in Section 3.0. Recommended methods and procedures for evaluating and updating background data sets are discussed in Section 4.0.

Following the collection of the initial eight background values, groundwater will be monitored at the landfill throughout the active life and post-closure period, as required by §257.94(b). The facility will collect groundwater samples at the frequencies required by §257.94(b) through (d) for detection monitoring and §257.95(b) through (d) for assessment monitoring, as applicable. In general, routine groundwater sampling and analysis is required on a semiannual basis during detection and assessment monitoring. An alternate sampling frequency is allowed by §257.94(d) and §257.95(c), but in no case can it be less than annual (see Section 7.1.1).

2.3 Statistical Methods Allowed by §257.93(f)

Pursuant to §257.93(f), groundwater data must be statistically tested with one or more of the following tests: analysis of variance (ANOVA), tolerance intervals (tolerance limits), prediction intervals (prediction limits), control charts, or another method that meets the performance criteria of §257.93(g). Each constituent must be separately tested for each monitoring well. Recommended statistical tests for the PPES monitoring program are discussed in Section 5.0 for detection monitoring and in Section 6.0 for assessment monitoring.

2.4 Performance Standards Required by §257.93(g)

The chosen statistical test(s) must comply with the following performance standards, as appropriate, as described in §257.93(g):

> *1. The statistical method used to evaluate groundwater monitoring data shall be appropriate for the distribution of constituents. Normal distributions of data values shall use parametric methods. Non-normal distributions shall use non-parametric methods. If the distribution of the constituents is shown by the owner or operator of the CCR unit to be inappropriate for a normal theory test, then the data must be transformed or a distribution-free (non-parametric) theory test must be used. If the distributions for the constituents differ, more than one statistical method may be needed.*

2. *If an individual well comparison procedure is used to compare an individual compliance well constituent concentration with background constituent concentrations or a groundwater protection standard, the test shall be done at a Type I error level no less than 0.01 for each testing period. If a multiple comparison procedure is used, the Type I experiment wise error rate for each testing period shall be no less than 0.05; however, the Type I error of no less than 0.01 for individual well comparisons must be maintained. This performance standard does not apply to tolerance intervals, prediction intervals, or control charts.*

3. *If a control chart approach is used to evaluate groundwater monitoring data, the specific type of control chart and its associated parameter values shall be such that this approach is at least as effective as any other approach in this section for evaluating groundwater data. The parameter values shall be determined after considering the number of samples in the background data base, the data distribution, and the range of the concentration values for each constituent of concern.*

4. *If a tolerance interval or a predictional interval is used to evaluate groundwater monitoring data, the levels of confidence and, for tolerance intervals, the percentage of the population that the interval must contain, shall be such that this approach is at least as effective as any other approach in this section for evaluating groundwater data. These parameters shall be determined after considering the number of samples in the background data base, the data distribution, and the range of the concentration values for each constituent of concern.*

5. *The statistical method must account for data below the limit of detection with one or more statistical procedures that shall at least as effective as any other approach in this section for evaluating groundwater data. Any practical quantitation limit that is used in the statistical method shall be the lowest concentration level that can be reliably achieved within specified limits of precision and accuracy during routine laboratory operating conditions that are available to the facility.*

6. *If necessary, the statistical method must include procedures to control or correct for seasonal and spatial variability as well as temporal correlation in the data.*

3.0 DATA REVIEW AND DATA REQUIREMENTS

Data should be reviewed prior to application of the statistical methods allowed by §257.93(f) to ensure data meet the statistical assumptions of the selected methods to evaluate groundwater quality. According to the Unified Guidance, most test methods require the following characteristics: statistical independence of the measured values, temporal and spatial stationarity, lack of statistical outliers, and the requisite data distribution for the applied test. This section reviews some of the methods and procedures available to screen groundwater quality data to ensure the data meet these assumptions, as required by the performance standards of §257.93(g). This section also discusses the handling of non-detect data and trace values, required by §257.93(g)(5).

3.1 Exploratory Data Analysis

Exploratory data analysis (EDA) is a procedure that includes using graphs such as time-series plots, box-and-whisker plots, and probability plots to provide a visual summary of the data. These tools provide an overview of the data and are useful for gaining insight to data characteristics such as the presence of trends, outliers, and spatial and/or temporal variation. A description of each of these methods and tools to prescreen data sets can be found in the Unified Guidance (Chapter 9).

3.2 Data Requirements

3.2.1 Sample Independence

3.2.1.1 Sampling Frequency

In accordance with the groundwater monitoring requirements of 40 CFR Part 257 and recommendations found in the Unified Guidance, groundwater samples should be collected at a frequency that ensures that the resulting data are statistically independent. Previous EPA guidance (EPA 1989) encouraged the use of the Darcy equation to determine the minimum time interval between sampling events; however, the Unified Guidance discourages over-reliance on this method. While this approach increases the probability that samples are physically

independent, it is not a guarantee that they will be statistically independent. This is due to the chemical characteristics of the aquifer (e.g., not all parameters travel at the same velocity as groundwater) and the assumptions of the Darcy equation itself (e.g., it assumes laminar flow). ASTM D6312-17 (ASTM International 2017) and Gibbons (1994a) recommend that groundwater be sampled no more frequently than quarterly to avoid autocorrelated data. Further information on evaluating groundwater values for statistical independence can be found in Chapter 14 of the Unified Guidance. If necessary, data correlation should be corrected in accordance with the performance standards of §257.93(g)(6).

3.2.1.2 Field and Laboratory Duplicates

Field or laboratory duplicates should not be treated as independent values or observations. While valuable for assessing the reproducibility of field techniques or the variability of laboratory methods, duplicate results are not statistically independent. Further information regarding the use of these data can be found in the Unified Guidance (Chapter 6).

3.2.2 Stationarity

In the context of groundwater quality statistics, stationarity refers to a data population that contains a mean and variance that generally remain the same over time and/or space. Lack of stationarity can be exhibited by spatial variability in data sets across a well network, temporal changes such as trends and/or seasonal variations, and heteroscedasticity (unequal variances). These data characteristics should be evaluated, controlled, and/or corrected in accordance with §257.93(g)(6). Methods and procedures to assess and treat non-stationarity are described in the Unified Guidance (Chapters 11, 13, and 14).

3.2.3 Outliers

Outliers are values that are extremely different than other values in a given data set. According to Helsel and Hirsch (2002), outliers can be the result of one of the following causes:

1. A measurement or recording error,
2. An observation from a population not similar to that of most of the data, or

3. A rare event from a single population that is skewed.

The consideration of outliers is a necessary component of defining an appropriate data distribution, as required by §257.93(g)(1). Screening data for potential outliers is discussed below and should follow methods described in the Unified Guidance.

EDA tools such as time-series plots, box-and-whisker plots, and probability plots are helpful in identifying possible outliers and can be used prior to the application of formal outlier tests such as Dixon's, Rosner's or Tukey's. Dixon's and Rosner's outlier tests are appropriate for normally distributed data populations, whereas Tukey's can be applied as a pseudo-nonparametric test on populations that are not normally distributed (or cannot be mathematically normalized). The Unified Guidance describes the use of each of these tests and their respective test assumptions (Chapter 12).

If suspect data are evident, results should be reviewed for evidence of error (e.g., calculation and/or transcription errors, field or laboratory instrument mis-calibration, elevated sample turbidity) and verified with the analytical laboratory if possible. If errors are found, they should be corrected in the database. If applicable, verification sampling (retesting) may be considered to confirm or disconfirm irregular data. Retesting is further discussed in Section 5.0.

Data that are identified as outliers with independent evidence of an error should not be used in statistical computations. Data excluded from the analysis should be documented along with the reason(s) for removal. As recommended by the Unified Guidance (Section 6.3.3 of Chapter 6), some low- or high-magnitude outliers may be excluded from statistical analysis even though evidence of error cannot be identified. The decision to include or exclude potential outliers in background data should be based on available technical information and professional judgment.

3.2.4 Data Distribution

Most statistical tests used for groundwater data analysis have explicit assumptions regarding the distribution of the data being tested. For this reason, and pursuant to §257.93(g)(1),

it is recommended that the distributions of the background data sets be characterized prior to the application of formal statistical procedures. As a general rule, it is recommended that summary statistics, including the mean, median, variance, standard deviation, and interquartile range, be calculated and reviewed. Data should further be evaluated to determine if the population is normally distributed, or can be mathematically transformed to be normally distributed, to determine if parametric tests can be appropriately used. If data cannot be transformed, then a nonparametric test should be considered. The Unified Guidance describes methods and procedures to evaluate data distributions in Chapter 10.

3.3 Non-Detect and Trace-Value Data

Methods for handling non-detect and trace-value data are required by §257.93(g)(5). Non-detect data are results that fall below a constituent method detection limit (MDL) and results are typically qualified with a "U." For parametric tests that rely on a normalized data distribution, large numbers of non-detect data are problematic. When the percentage of non-detect data is fairly low (i.e., between 10% and 15%), a simple substitution method can be used and non-detect results can be substituted with one-half the value of the laboratory practical quantification limit (PQL) or reporting limit (RL). When non-detect data comprise greater than 15% but less than 50% of the data set, censored estimation techniques such as Kaplan-Meier or robust regression on order statistics (ROS) are recommended. If the data set is comprised of greater than 50% non-detects, then a nonparametric test is recommended. For the circumstance where 100% of the data are non-detect for a specific well-parameter pair, the Unified Guidance recommends considering the use of the Double Quantification Rule (discussed in Section 5.0).

Trace values are estimated concentrations that fall between the MDL and a laboratory PQL or RL. Trace values are usually qualified with a "J" and should be treated as valid detections for the purpose of statistical analyses.

Further information regarding the appropriate treatment of non-detect data can be found in Chapter 15 of the Unified Guidance. Other methods for handling non-detect and trace values may be used if appropriate and recommended by the Unified Guidance.

4.0 ESTABLISHING AND UPDATING BACKGROUND DATA SETS

This section describes recommended methods and procedures to evaluate the initial eight background values collected in accordance with §257.94(b), to update background data sets over the life of the monitoring program, and to evaluate and use data from replacement monitoring wells should the need occur to decommission and replace a monitoring well over the life of the program. This section also discusses the use of interwell versus intrawell statistical comparisons for groundwater at the landfill, since this concept must be considered prior to evaluating appropriate background. As identified in the Unified Guidance and in Section 8.0, the term "background" refers to the natural or baseline groundwater quality at a site. Background conditions can range from an uncontaminated aquifer to a historically contaminated site with baseline conditions that are unaffected by recent releases that are actionable under the Resource Conservation and Recovery Act (RCRA). The Unified Guidance and this SAP use the terms "background" and "baseline" interchangeably.

4.1 Interwell versus Intrawell Comparisons

An interwell comparison is an "upgradient to downgradient" comparison. Data from the downgradient compliance wells are statistically compared to data from a hydraulically upgradient or background well, or to pooled data from two or more upgradient or background wells. An intrawell comparison is a "within well" comparison. New compliance data from the well are compared to earlier background data from the well itself.

In the case of the interwell comparison, several hydrogeological and hydrogeochemical conditions must be met in order for the comparison and resulting statistical test to be valid. According to the Unified Guidance (Chapter 6), these factors generally include the following:

1. Changes in downgradient wells are only caused by landfill activity. Placement of compliance wells downgradient of other sources of impact will make it difficult to determine the cause of changing water quality.
2. Upgradient and downgradient wells are screened in the same aquifer and at the same approximate hydrostatic position.

3. Groundwater flows in a definable, consistent pathway from upgradient to downgradient.
4. Groundwater flows at a sufficient velocity such that the same groundwater sampled at the upgradient wells is subsequently sampled at downgradient wells over the course of an evaluation period (e.g., semiannual or annual sampling frequency).

Historical information from the landfill shows the direction of groundwater flow in the aquifer to be variable, and shifts on a seasonal basis to the east, west, north, or south (FTN 2017). Consequently, condition 3 listed above is not met and intrawell comparisons for groundwater quality are warranted for the Plum Point Energy Station landfill groundwater monitoring program.

4.2 Establishing Background

The initial background data should be reviewed to ensure that the values reflect current natural or baseline water quality. To do this, data can be screened using the EDA tools listed in Section 3.1 to identify potential trends, outliers, and spatial variability. Background data sets should be further evaluated using a formal trend test, such as the Mann-Kendall test and Theil-Sen trend line, described in Section 5.0, and the outlier tests summarized in Section 3.0. If trends are indicated in background populations, testing strategies that either correct for, or are not sensitive to, temporal variation may be required. Statistical outliers may need to be excluded from a background data set, particularly if values are found to be in error. Documentation and treatment of outliers should be handled as summarized in Section 3.0 and as recommended by the Unified Guidance.

4.3 Updating Background

Because some long-term fluctuation in background levels may be possible even though a given well has not been impacted by the landfill, the Unified Guidance recommends periodically updating background data to include more recent observations. Increasing the number of background samples also increases the statistical power of the tests (see Section 5.0 for a discussion of statistical power). The methods described in this section adhere to methods

described in the Unified Guidance and will help to ensure that background data sets do not include data that may reflect impact from onsite landfilling activities. Other methods may be used to update background data if alternate methods are appropriate.

Background populations for well-parameter pairs are considered for a potential update when at least four compliance data values are available and if no confirmed statistically significant exceedance has been indicated since the last background update. A statistically significant exceedance can be either a statistically significant increase (SSI) or a statistically significant decrease (SSD) in the case of pH and are discussed further in Section 5.0. The terms "statistically significant exceedance" and "SSI/SSD" are used interchangeably in the Unified Guidance and in this SAP. In the event that a well-parameter pair has an unconfirmed SSI/SSD, background data should not be considered for an update until retesting has been completed (retesting is discussed in Section 5.0). If the initial exceedance is disconfirmed by retesting, the exceedance should be treated in accordance with recommendations in Chapter 5, Section 5.3.3, of the Unified Guidance (regarding updating background data).

Prior to updating, existing background and compliance populations for each well-parameter pair will be tested with a matched-pair test, such as the intrawell Mann-Whitney (Wilcoxon Rank-Sum) or a similar test. The intrawell Mann-Whitney (Wilcoxon Rank-Sum) test evaluates whether the measurements from the compliance data are statistically higher or lower than the existing background population. Where testing indicates that the compliance data are not statistically different from the existing background data, data sets are updated.

Updated background populations should be further evaluated for the presence of statistically significant outliers and trends. In some cases, extreme low- or high-magnitude outliers may be excluded from statistical analysis, even if no independent evidence of error is identified for the cause of the anomalous data. As advised in the Unified Guidance, select removal of extreme outliers without knowledge of error may be warranted, but removal of all outliers can mask real and legitimate changes in background data. Subsequent to outlier screening, background should be tested for the presence of trends using the Mann-Kendall/ Theil-Sen trend line or a similar test to determine the presence of statistically significant

increasing or decreasing trends. The presence of linear trends in background data may violate key assumptions of some statistical tests and require an alternate approach to testing the data.

Other methods and procedures may be used to update background data sets as appropriate and if recommended by the Unified Guidance.

4.4 Background Data for Replacement Wells

In the event that the facility has to decommission and replace a monitoring well, data from the two wells should be examined to determine if the data collected from them are statistically similar. In general, a "replacement well" is located sufficiently close to the original or decommissioned well such that the two monitored points are hydrogeochemically similar. The advantage of this is that background data from the original/decommissioned well can be utilized for statistical comparisons with newer data collected from the replacement well. Prior to using data from the original well as background for the replacement well, a matched-pair test, such as the Mann-Whitney (Wilcoxon Rank-Sum) or a similar test, should be applied to evaluate whether the data sets are considered to come from the same population.

5.0 DETECTION MONITORING

This section describes the statistical program design and methods and procedures to determine statistically significant exceedances in groundwater quality during detection monitoring. Recordkeeping and notification requirements for detection monitoring are described in Section 7.0.

5.1 Site-Wide False Positive Rate and Statistical Power

The Unified Guidance recommends that detection monitoring programs have adequate statistical power and a site-wide false positive rate (SWFPR) value of 10% over a one-year period of testing. Using the Unified Guidance standard, the yearly SWFPR is fixed at 10%, and as such, the semiannual SWFPR is fixed at 5%. The magnitude of the per-test alpha will vary depending on how many statistical tests are required per semiannual evaluation.

To gauge statistical power, the Unified Guidance recommends the use of the EPA Reference Power Curve (ERPC) to estimate the ability of any individual test to identify a statistically significant exceedance from background. Any single statistical test should have the ability to detect an exceedance 55% to 60% of the time at three standard deviations (3σ) above background, and 80% to 85% of the time at 4σ above background. The statistical power of the detection monitoring program should be evaluated against the ERPC on an annual basis or when changes are made to background data sets (e.g., when background is updated).

5.2 Evaluating for Statistically Significant Exceedances

As described in Section 2.0, §257.93(f) allows the following tests to evaluate for statistically significant exceedances in groundwater quality: ANOVA, tolerance limits, prediction limits, control charts, or another method that meets the performance criteria of §257.93(g). The Unified Guidance recommends prediction limits combined with retesting to maintain a low SWFPR and to achieve the recommended target rates for identifying statistically significant exceedances (power). Control charts are also acceptable but require the data be normally distributed (or transformed-normal). There is not a nonparametric version of the control

chart, unlike prediction limits, which can be used as both parametric and nonparametric tests. The Unified Guidance strongly discourages the use of ANOVA for the purpose of determining SSIs/SSDs due to problems associated with excessive SWFPRs. ANOVA is also an inherently interwell test, and results would be invalid at this facility for reasons discussed in Section 4.0.

This SAP describes the use of prediction limits with retesting. It also describes the use of formal trend tests in the event that background data contain linear trends that cannot be corrected, as recommended by the Unified Guidance. This section also summarizes the use of the Double Quantification Rule for data sets that contain 100% non-detect values in background. Other methods may be applied if allowed by §257.93(f) and if used in accordance with recommendations in the Unified Guidance.

5.2.1 Prediction Limits with Retesting

A prediction limit tests for the likelihood that a new monitoring value (compliance value) comes from the same population as background data. Prediction limit analysis combined with retesting (verification sampling) is effective at reducing a monitoring program's SWFPR and improving the statistical power of the monitoring program. Parametric or nonparametric prediction limits may be applied, as appropriate, and are described below. Retesting strategies are discussed in Section 5.2.2. Methods and procedures used to calculate prediction limits are detailed in Chapters 18 and 19 of the Unified Guidance.

The prediction limit test assumes that the data being tested are statistically independent and exhibit stationarity. For parametric tests, data should be normally distributed or mathematically transformed-normal prior to analysis. A minimum of 8 to 10 background values are recommended by the Unified Guidance for this test. If data characteristics do not meet the assumptions of the test, alternate testing methods (such as a formal trend test) may be warranted, as recommended by the Unified Guidance.

A parametric prediction limit can be applied to a normally distributed (or transformed-normal) data set where non-detects are less than 50%. The parametric prediction limit (PL) is computed according to the general equation:

$$PL = \overline{X} + Ks$$

Where: $\overline{X}$ = the sample mean of background data at the same sampling location,
K = a multiplier, the value of which will depend on the type of prediction limit under construction and the landfill's retesting strategy, and
s = standard deviation of background.

A nonparametric prediction limit may be constructed in the event that (1) background data do not follow a normal (or transformed-normal) distribution, (2) greater than 50% of the background data are non-detect, or (3) a parametric test would be considered inappropriate for a given sample size. For a nonparametric test, the highest or second-highest value from the background data may be specified and used as the upper limit of the prediction limit.

With the exception of pH, calculated prediction limits should be one-sided to estimate the upper concentration limit associated with the tested data. A two-sided prediction limit should be calculated for pH, since increasing or decreasing values can indicate groundwater quality deterioration.

The confidence level associated with prediction limit testing is discussed in Section 5.1.

5.2.2 Retesting Strategy

In the event of an unexplained statistical exceedance, retesting should be implemented by obtaining one or more verification samples to confirm or disconfirm the initial result. Because of the slow rate of groundwater flow beneath the landfill, this SAP recommends using a "1 of 2" retesting strategy. The "1 of 2" strategy requires obtaining one verification sample within the same monitoring period as the initial exceedance. Sufficient time must elapse between sampling events for the data to be statistically independent. For this reason, a "1 of 3" strategy, where two verification samples are collected within the same monitoring period as the initial exceedance, is

not recommended due to the slow rate of groundwater flow at the site and the need to maintain statistical independence of the samples (see Section 3.2.1.1).

If retesting using the "1 of 2" approach disconfirms the initial exceedance, then a declaration of a statistical exceedance is not required, as recommended by the Unified Guidance. Unless the initial (disconfirmed) exceedance can be documented as an outlier or other anomaly, it should be retained in the database as a valid measurement in addition to the resample result (Unified Guidance Chapters 5 and 19). If retesting confirms the initial exceedance, the facility is required to declare a confirmed statistical exceedance.

5.2.3 Analysis of Trends

A data set can be tested for the presence of statistically significant trends using a combination of the Mann-Kendall test and Theil-Sen trend line tests. Both tests are nonparametric, meaning they do not depend on normal distribution of the data. The Mann-Kendall tests for the presence of an upward or downward slope at a given confidence level. The Theil-Sen trend line provides an estimate of the slope magnitude. Further information regarding how these tests are constructed can be found in the Unified Guidance (Chapter 17). Other formal trend tests may be used if appropriate and if recommended by the Unified Guidance.

If a statistically significant increasing trend (or statistically significant decreasing trend in the case of pH) is indicated for a period-of-record data set, then evidence exists of possible deteriorating groundwater quality and the facility is required to declare a statistical exceedance, as recommended by the Unified Guidance.

5.2.4 Double Quantification Rule

For circumstances where a well-parameter pair is 100% non-detect, the Unified Guidance recommends the use of the Double Quantification Rule, which states that "[a] confirmed exceedance is registered if any well-constituent pair in the '100% non-detect' group exhibits quantified measurements (i.e., at or above the reporting limit [RL]) in two consecutive sample and resample events." The Unified Guidance recommends a minimum of 10 to 15 consecutive

values be non-detect prior to applying this rule. Further considerations for applying this rule can be found in Chapter 6 of the Unified Guidance.

5.3 Determination of a Statistically Significant Exceedance above Background

If a confirmed statistical exceedance is found during detection monitoring, the landfill is required to follow the actions required by §257.94(e)(1) through (3), as applicable, for assessment monitoring. Assessment monitoring and the associated statistical approach is described in Section 6.0 of this SAP. Recordkeeping and notification requirements for the finding of a confirmed statistical exceedance are summarized in Section 7.0.

6.0 ASSESSMENT MONITORING

This section describes the methods and procedures to statistically test groundwater quality data during assessment monitoring. Recordkeeping and notification requirements for assessment monitoring are described in Section 7.0.

6.1 Initiation of Assessment Monitoring

Pursuant to §257.95(a), the landfill is required to initiate assessment monitoring if statistical analysis shows a confirmed statistical exceedance for one or more appendix III parameters and a successful alternate source demonstration (ASD) has not been made in accordance with §257.94(e)(2). As discussed in Section 7.1.2, an ASD is a written demonstration showing that a source other than the landfill caused a statistically significant exceedance (e.g., laboratory error, sampling error, upgradient or offsite influences). Within 90 days of the confirmed SSI/SSD and annually thereafter for the duration of assessment monitoring, the landfill is required to collect samples at all wells in the monitoring network and analyze them for the list of appendix IV constituents [§257.95(b)]. Within 90 days of obtaining the results from the first assessment monitoring sampling event, and on a semiannual basis thereafter during assessment monitoring, the landfill is required to collect samples from all wells for the analysis of the list of appendix III constituents and detected appendix IV constituents from the first assessment monitoring sampling event (and/or from subsequent annual appendix IV monitoring events) [§257.95(d)(1)].

6.2 Establishing Groundwater Protection Standards

In accordance with §257.95(d)(2), groundwater protection standards (GWPSs) will be established for all detected appendix IV constituents if assessment monitoring is required. Per §257.95(h), GWPSs will be established as follows:

1. For constituents for which a maximum contaminant level (MCL) has been established under 40 CFR Part 141, §§141.62 and 141.66, the GWPS will be the MCL for that constituent [§257.95(h)(1)];

2. For constituents for which an MCL has not been established, the GWPS will be the background concentration for the constituent established from wells in accordance with §257.91 [§257.95(h)(2)]; or
3. For constituents for which the background concentration is higher than the MCL, the GWPS will be the background concentration [§257.95(h)(3)].

Chapter 7 of the Unified Guidance describes methods and procedure for establishing GWPSs.

6.3 Statistical Comparison to a GWPS

During assessment monitoring, data must be evaluated to determine if groundwater concentrations of one or more appendix IV parameters are at statistically significant levels (SSLs) above a GWPS [§257.95(g)]. The statistical methods allowed by §257.93(f) are appropriate for comparing compliance data to a background population; however, they are not appropriate for statistically comparing data to a fixed standard (GWPS). For this type of analysis, the Unified Guidance recommends confidence interval testing against a fixed GWPS (Chapters 7 and 21). Both parametric and nonparametric intervals can be constructed, and are discussed below.

The purpose of assessment monitoring is to ascertain whether groundwater quality has deteriorated such that a population statistic (e.g., mean, median, or percentile) estimated for a parameter at a well exceeds a GWPS at an SSL. Because the true boundaries of the statistical population (i.e., the full range of concentrations that can be expected for a parameter at a well) are unknown, a confidence interval test can be used to estimate those boundaries based on the population statistic. The population statistic is used to estimate the upper and lower boundaries of the population, and the alpha level of the test is the degree of confidence that the true population statistic falls within those calculated boundaries. As such, there are times when the GWPS may be exceeded by an individual measurement, but that alone is not necessarily indicative of an SSL with respect to the GWPS. Rather, a range of values represented by the confidence interval is compared to a fixed GWPS to determine if groundwater concentrations have increased above the GWPS at an SSL.

Interpreting the results of the confidence interval test is straightforward. If the upper limit of the confidence interval is below the GWPS, it can be concluded that there is statistically significant evidence that groundwater concentrations (with respect to the population statistic used) lie below the GWPS. If the upper and lower limits of the confidence interval straddle the GWPS, then it is unknown whether groundwater concentrations exceed or lie below the GWPS, and the facility is required to continue assessment monitoring. If the lower limit of the interval is above the GWPS, then groundwater concentrations are considered to be greater than the GWPS at an SSL.

Similar to the statistical tests used for detection monitoring, a confidence interval test assumes the data tested are statistically independent, exhibit stationarity, and follow a normal or transformed-normal distribution in the case of parametric evaluations. As such, data should be examined using EDA, outlier tests, and trend tests. Data shifts (up- or downward departures from a previously stable population) and the presence of linear trends should be taken into consideration when determining the appropriate window of data tested. The Unified Guidance recommends preferably eight or more values per compliance well to construct a confidence interval. Fewer values may be used; however, a minimum of four data points is required to construct a parametric confidence interval and a minimum of seven data points is required to construct a nonparametric confidence interval. Treatment of trace values and non-detect data should be handled as summarized in Section 3.0 of this SAP and as recommended by the Unified Guidance (Chapter 15).

The methods for determining the degree of confidence and associated statistical power for confidence interval testing are described in Chapter 22 of the Unified Guidance. Further information on how to calculate and apply confidence intervals is described in Chapters 7 and 21 of the Unified Guidance. Other statistical tests may be used to evaluate groundwater values against a GWPS if appropriate and recommended by the Unified Guidance.

6.4 Return to Detection Monitoring

Pursuant to §257.95(e), the landfill may return to detection monitoring if the concentrations of all constituents listed in appendix III and appendix IV are shown to be at or

below background values, using the statistical procedures in §257.93(f). For this reason, assessment monitoring data should be compared to background values on a periodic basis.

If compliance groundwater quality data appear to be near or at background values, the landfill may opt to increase the frequency of sampling and analysis (e.g., from annual to semiannual) for the full list of appendix IV parameters. Increasing the frequency of analysis for both lists from the required annual analysis to semiannual or quarterly may provide an opportunity to return to detection monitoring sooner, so long as the data show that water quality has returned to background levels.

6.5 Determination of Statistically Significant Levels above a GWPS

If statistical analysis indicates that measured values of one or more appendix IV parameters are at an SSL above a GWPS, the landfill is required to follow the actions required by §257.95(g)(1) through (4), as applicable. Reporting and notification requirements for the finding of statistically significant levels above a GWPS are summarized in Section 7.0.

7.0 RECORDKEEPING, NOTIFICATION, AND PUBLIC INTERNET SITE REQUIREMENTS

This section presents the recordkeeping, notification, and publicly accessible internet site (CCR website) requirements for detection and assessment monitoring. Recordkeeping requirements for groundwater monitoring are included in §257.105(h), which identifies documents that must be placed in the landfill's operating record. Notification requirements for groundwater monitoring are included in §257.106(h), which identifies the notifications that must be sent to the relevant state director and/or appropriate tribal authority. For the Plum Point Energy Station landfill, the appropriate state director is ADEQ. The CCR website requirements for groundwater monitoring are included in §257.107(h), which identifies documents that must be maintained on the website for public access.

This SAP does not identify any of the recordkeeping, notification, or CCR website requirements related to assessment of corrective measures or corrective action.

7.1 Annual Groundwater Monitoring and Corrective Action Report

In accordance with §257.90(e) and §257.105(h)(1), the facility is required to place the first annual groundwater monitoring and corrective action report in the facility's operating record by January 31, 2018, and by January 31 annually thereafter. The annual report is required to describe the monitoring and corrective action program, summarize key actions completed, document problems encountered and actions taken to resolve the problems, and project key activities for the upcoming year. Within 30 days of placing the annual report in the facility's operating record, the landfill is required to place the annual report on the facility's website in accordance with §257.107(h)(1) and provide notification to the director of ADEQ that the report is available on the website and in the operating record in accordance with §257.106(h)(1).

The report will contain the following information in accordance with §257.90(e)(1) through (5):

1. A map, aerial image, or diagram showing the CCR facility and the groundwater monitoring well network, including the well identification numbers, for the CCR program [§257.90(e)(1)];

2. Identification of any monitoring wells that were installed or decommissioned during the preceding year, along with a narrative description of why those actions were taken [§257.90(e)(2)];
3. A summary of sampling activities and data, including the number of groundwater samples that were collected at each well, sampling dates, and whether each sample was required by the detection monitoring or assessment monitoring programs [§257.90(e)(3)];
4. A narrative discussion of any transition between monitoring programs (e.g., the date and circumstances for transitioning from detection monitoring to assessment monitoring in addition to identifying any constituent(s) with statistically significant exceedances with respect to background levels) [§257.90(e)(4)]; and
5. Other information required by §§257.90 through 257.98 [§257.90(e)(5)].

If the landfill is in assessment monitoring, the annual report is required to also include the following information pursuant to §257.95(d)(3):

1. The recorded concentrations of appendix III and appendix IV parameters,
2. The background concentrations identified for all constituents listed in appendix III and appendix IV, and
3. The GWPSs established for appendix IV parameters.

Other documentation that must be included in the annual groundwater monitoring and corrective action report, should the need arise, are an alternate sampling frequency demonstration and/or an alternate source demonstration. This documentation and associated recordkeeping and notification requirements are described in the following sections (Sections 7.1.1 and 7.1.2).

7.1.1 Alternate Sampling Frequency Demonstration

If the landfill anticipates the need for an alternative sampling frequency to that described in §257.94(b) and §257.95(b) due to insufficient groundwater flow for sampling, the landfill may document the need for less-frequent sampling based on consideration of the following factors [§257.94(d)(1) and (2); §257.95(c)(1) and (2)]:

1. Lithology of the aquifer and unsaturated zone;
2. Hydraulic conductivity of the aquifer and unsaturated zone; and
3. Groundwater flow rates.

For an alternate sampling frequency demonstration for detection monitoring, the documentation must show that the alternate frequency will be no less effective in detecting a release within a timeframe that will not materially delay establishment of an assessment monitoring program. For an alternate sampling frequency demonstration for assessment monitoring, the documentation must show that the alternate frequency will be no less effective in detecting a release within a timeframe that will not materially delay initiation of remediation measures. The documentation must be certified by a qualified professional engineer stating that the demonstration for an alternative groundwater sampling and analysis frequency meets the requirements of §257.94 (for detection monitoring) or §257.95 (for assessment monitoring).

The landfill must include the demonstration providing the basis for the alternative monitoring frequency and the certification by a qualified professional engineer in the annual groundwater monitoring and corrective action report.

7.1.2 Alternate Source Demonstration for Statistically Significant Results

If a statistically significant exceedance over background levels is detected, the landfill may demonstrate that the SSI/SSD resulted from an alternate source, such as a sampling error, statistical artifact, or natural variation in groundwater quality. The alternate source demonstration must be completed and certified by a qualified professional engineer within 90 days of the SSI/SSD detection.

The alternate source demonstration and engineer's certification must be included in the annual groundwater monitoring and corrective action report required by §257.90(e). If a successful demonstration is completed within the 90-day period, the landfill is required to continue with detection monitoring.

7.2 Other Required Recordkeeping and Notifications

Outlined below are the recordkeeping and notifications that are required based on the statistical results of detection or assessment monitoring (beyond those identified in Section 7.1). As previously stated, this SAP does not identify any of the recordkeeping, notification, or CCR website requirements related to assessment of corrective measures or corrective action.

7.2.1 Notification of the Establishment of Assessment Monitoring

Within 30 days of entering assessment monitoring, the landfill is required to place a notification in the facility's operating record in accordance with §257.105(h)(5) that the landfill has established an assessment monitoring program. Within 30 days of placing the notice in the operating record, the landfill is required to place the notice on the facility's website [§257.107(h)(4)] and notify the director of ADEQ that the notice has been placed in the facility's operating record and on the website [§257.106(h)(4)].

7.2.2 Analytical Results

The landfill is required to place the sampling results from assessment monitoring in the facility's operating record as those results are available [§257.95(d)(1) and §257.105(h)(6)]. As described in Section 7.1, these results are also required in the annual report.

7.2.3 Return to Detection Monitoring

If the concentrations of all constituents listed in appendix III and appendix IV are shown to be at or below background values for two consecutive sampling events, the landfill can resume detection monitoring [§257.95(e)]. The landfill is required to prepare a notification stating that detection monitoring has been resumed and place it in the facility's operating record within 30 days of returning to detection monitoring [§257.105(h)(7)]. Within 30 days of placing the notice in the operating record, the landfill is required to place the notice on the facility's website [§257.107(h)(5)] and notify the director of ADEQ that the notice has been placed in the facility's operating record and on the website [§257.106(h)(5)].

7.2.4 Detection of Statistically Significant Level above a GWPS

If one or more constituents in appendix IV are detected at a statistically significant level above a GWPS, the landfill is required to place a notice in the facility's operating record within 30 days of identifying the exceedance(s) [§257.95(g) and §257.105(h)(8)]. Within 30 days of placing the notice in the operating record, the landfill is required to place the notice on the facility's website [§257.107(h)(6)], and notify the director of ADEQ that the notice has been placed in the facility's operating record and on the website [§257.106(h)(6)].

7.2.5 Response to Determination of Statistically Significant Levels

As previously stated, if statistical analysis indicates that measured values of one or more appendix IV parameters are at an SSL above a GWPS, the landfill is required to initiate the actions and notifications of §257.95(g)(1) through (4), as applicable. The requirements of §257.95(g)(1) through (4) are summarized below:

- The landfill is required to characterize the nature and extent of the release and any relevant site conditions that may affect selection of a remedy. The characterization must include (1) the installation of additional monitoring wells necessary to define the contaminant plume(s), including at least one additional monitoring well at the facility boundary in the direction(s) of contaminant migration; (2) sampling of all wells for the list of appendix III and appendix IV list of constituents; and (3) an estimation of the quantity of material released [§257.95(g)(1)].
- The landfill must to notify all persons who own or reside on the land that directly overlies any part of a plume that has migrated offsite [§257.95(g)(2)].
- The landfill must initiate assessment of corrective measures in accordance with §257.96 within 90 days of finding that any of the constituents listed in appendix IV have been detected at an SSL above a GWPS or demonstrate that a source other than the CCR unit caused the contamination [§257.95(g)(3)]. The demonstration must be made in accordance with §257.95(g)(3)(ii). If a successful demonstration has not been made at the end of the 90-day period, the landfill is required to initiate the assessment of corrective measures requirements under §257.96 [§257.95(g)(4)].

8.0 GLOSSARY OF SELECT TERMS[2]

alpha (α) level: Decimal level of significance or false positive error of a statistical test

alternate source demonstration (ASD): A written demonstration to show that a source other than the landfill has caused a statistically significant exceedance, either during detection monitoring [§257.94(e)(2)] or assessment monitoring [§257.95(g)(3)(ii)].

autocorrelation: Correlation of values of a single variable data set over successive time intervals

background: Natural or baseline groundwater quality at a site; can be characterized by upgradient, historical, or sometimes sidegradient water. Background data should reflect historical conditions unaffected by the activities the groundwater monitoring program is designed to assess. These conditions could range from an uncontaminated aquifer to a historically contaminated site unaffected by recent RCRA-actionable contaminant releases. In this document, "background" is typically a reference to the background data set to which compliance values are compared.

baseline: see "background"

confidence level: Degree of confidence associated with a statistical estimate or test, denoted as "1–α."

EPA Reference Power Curve (ERPC): Recommended standards for comparing performance of RCRA statistical methods in detection monitoring; based on individual prediction limit using 10 background samples and an alpha level of 0.01.

false negative: Finding no statistically significant difference when there is a physical difference in the underlying populations or between a single population and a fixed compliance standard; also known as Type II error.

false positive: Finding a statistically significant difference when there is no physical difference in the underlying populations or between a single population and a fixed compliance standard; also known as a Type I error. The alpha (α) level is the probability of committing a Type I error.

interwell: Comparisons between distinct monitoring wells; "upgradient to downgradient" comparison.

intrawell: Comparisons over time at a given monitoring well between early and later measurements; "within well" comparison.

method detection limit (MDL): The minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero in a specific matrix

nonparametric test: Statistical test that does not depend on knowledge of the distribution of the sampled population

[2] This glossary contains adapted definitions of terms from the Unified Guidance.

normal distribution: A family of symmetric continuous probability distributions defined by two finite parameters, the mean and variance

outlier: Value unusually discrepant from rest of a series of observations

parametric test: Statistical test that depends upon or assumes observations from a particular probability distribution or distributions

percentile: The specific value of a distribution that divides the distribution such that p percent of the distribution is equal to or below that value. If the 95^{th} percentile is X, it means that 95% of the values in the statistical sample are less than or equal to X.

population: All possible measurements/values over a period of time at a given location, series of locations, or over a spatial or volumetric extent

practical quantification limit (PQL): Lowest concentration level for an analytical method that can be reliably achieved within specified limits of precision and accuracy under routine laboratory operating conditions

probability: Quantitative measure of uncertainty about the occurrence of a random or uncertain event

probability distribution: Numerical statistical pattern associated with a population of measurements; many common patterns can be described using mathematical formulas.

regression on order statistics (ROS): either parametric or robust techniques for fitting non-detect data to a single distribution

reporting limit: Lowest concentration level for an analytical method that can be reliably measured by a laboratory

seasonality: The presence of seasonal effects on groundwater quality observations; effects may be natural or man-made.

site-wide false positive rate (SWFPR): design probability of at least one statistically significant finding among a network of statistical test comparisons at a group of uncontaminated wells

statistical power: Strength of a test to identify an actual release of contaminated groundwater or difference from a compliance standard

statistically significant exceedance: Statistical difference exceeding a test limit by a large enough margin to account for data variability and chance. A statistically significant exceedance can be either a statistically significant increase (SSI) or a statistically significant decrease (SSD) in the case of pH.

variance: A measure of spread or dispersion calculated as the average of squared differences from the mean in a set of data or a population

verification resampling or retesting plan: A plan to collect an additional sample or samples to confirm or disconfirm an initial statistically significant exceedance

9.0 SELECTED REFERENCES

ASTM International. 2017. *D6312: Standard Guide for Developing Appropriate Statistical Approaches for Groundwater Detection Monitoring Programs at Waste Disposal Facilities*. West Conshohocken, PA: ASTM International.

EPA. 1989. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Interim Final Guidance* [EPA 530-SW-89-026]. Washington, DC: US Environmental Protection Agency, Office of Solid Waste, Waste Management Division.

EPA. 1993. *Solid Waste Disposal Facility Criteria, Technical Manual* [EPA 530-R-93/017]. Washington, DC: US Environmental Protection Agency, Office of Solid Waste and Emergency Response. November 1993.

EPA. 2009. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Unified Guidance*. US Environmental Protection Agency, Office of Resource Conservation and Recovery, Program Implementation and Information Division. Washington, DC.

EPA. 2015. Hazardous and Solid Waste Management System; Disposal of Coal Combustion Residuals From Electric Utilities; Final Rule. Title 40 *Code of Federal Regulations*, Part 257. 2015 edition.

FTN. 2017. *Conceptual Site Model, Plum Point Energy Station.* Little Rock, AR: FTN Associates, Ltd.

FTN. 2017. *Groundwater Sampling and Analysis Plan, Plum Point Energy Station.* Little Rock, AR: FTN Associates, Ltd.

Gibbons, R.D. 1994a. The folly of Subtitle D statistics: When greenfield sites fail. In *Proceedings of WasteTech '94, National Solid Waste Management Association, Washington, DC, January 13-15, 1994*. Quoted in EPA 2009, p. 14-19.

Gibbons, R.D. 1994b. *Statistical Methods for Groundwater Monitoring*. New York: John Wiley & Sons, Inc.

Helsel, D.R., and R.M. Hirsch. 2002. *Statistical Methods in Water Resources* [USGS Techniques of Water-Resources Investigations Book 4, Chapter A3]. Reston, VA: US Geological Survey. 524 pp. Available online at https://pubs.usgs.gov/twri/twri4a3/html/toc.html

Puls, R.W., and M.J. Barcelona. 1996. *Low-Flow (Minimal Drawdown) Ground-Water Sampling Procedures* [EPA 540/S-95/504]. Technical Support Project, Groundwater Forum, Issue Paper. Washington, DC: US Environmental Protection Agency, Office of Research and Development, Office of Solid Waste and Emergency Response. April 1996.

Sanitas Technologies, Inc. 2014. *Sanitas Statistical Software, User Guide, version 9.5*. Shawnee, KS: Sanitas Technologies.