STATISTICAL ANALYSIS PLAN
AES PUERTO RICO LP, GUAYAMA, PUERTO RICO

PREPARED IN COMPLIANCE WITH USEPA's COAL COMBUSTION RESIDUALS RULE, 40 CFR 257.93

OCTOBER 2023

Prepared for:

AES Puerto Rico, LP
PO Box 1890
Guayama, Puerto Rico 00785

Prepared by:

DNA-Environment, LLC
35 Calle Juan C. Borbon, STE 67-227
Guaynabo, Puerto Rico 00969-5375

DNA-ENVIRONMENT, LLC

STATISTICAL ANALYSIS PLAN
AES PUERTO RICO LP, GUAYAMA, PUERTO RICO

PREPARED IN COMPLIANCE WITH USEPA's COAL COMBUSTION RESIDUALS RULE, 40 CFR 257.93

OCTOBER 2023

Prepared by: _____

Alberto Meléndez

Principal Environmental Consultant

CERTIFICATION

I, Edwin A. Ayala Ramírez, am a qualified professional engineer as defined in 40 CFR §257.53. I have reviewed this Statistical Analysis Plan (AES Puerto Rico LP, Guayama, Puerto Rico) and certify that the selected statistical methods described herein are appropriate and meet the statistical requirements under 40 CFR §257.93.

I am a duly registered Professional Engineer under the laws of the Commonwealth of Puerto Rico.

Edwin A. Ayala Ramírez, P.E. P.E. License No. 16316	10-10-2023 Date

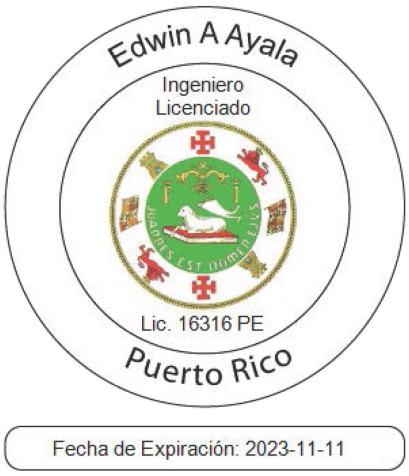

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AES-PR	AES Puerto Rico, LP
CCR	Coal Combustion Residuals
CFR	Code of Federal Regulations
DNA	DNA-Environment, LLC
GWPS	Groundwater Protection Standard
LCL	Lower Confidence Limit
Ln	Natural Logarithm
LPL	Lower Prediction Limit
MCL	Maximum Contaminant Level
mg/L	Milligrams per Liter
PE	Professional Engineer
PQL	Practical Quantitation Limit
RL	Reporting Limit
ROS	Regression of Order Statistics
§	Section
SSI	Statistically Significant Increase
SSL	Statistically Significant Level
SWFPR	Sitewide False Positive Rate
UCL	Upper Confidence Limit
UPL	Upper Prediction Limit
USEPA	United States Environmental Protection Agency
UTL	Upper Tolerance Limit

This Page Intentionally Left Blank

1 INTRODUCTION

1.1 Purpose and Scope

DNA-Environment, LLC (DNA) has prepared this Statistical Analysis Plan (Plan) for the temporary staging area of manufactured aggregate (AGREMAX™) at AES Puerto Rico LP (AES-PR) in Guayama, Puerto Rico. The Plan describes the statistical criteria and procedures that will be employed to evaluate site groundwater data in accordance with the Coal Combustion Residuals (CCR) groundwater monitoring and corrective action requirements under 40 CFR §§257.90 through 257.98. Acceptable statistical methods and performance criteria are prescribed in 40 CFR §257.93.

This Plan updates the procedures presented in the PE-Certified Statistical Analysis Plan prepared in 2017 for the CCR groundwater monitoring program at AES-PR, which was included in the document entitled *Groundwater Monitoring System & Sampling and Analysis Program, AES Puerto Rico LP, Guayama, Puerto Rico* (DNA, August 2017). This updated Plan incorporates additional details to improve clarity regarding the selected statistical procedures and replaces the 2017 Statistical Analysis Plan.

The procedures for collecting, preserving, shipping, and laboratory analysis of the groundwater samples are described in a separate document entitled *Federal CCR Corrective Action Groundwater Monitoring Plan, AES Puerto Rico LP, Guayama, Puerto Rico* (DNA, October 2023).

1.2 Groundwater Monitoring Requirements under the Federal CCR Rule

In April 2015, the United States Environmental Protection Agency (USEPA) issued the final rule that establishes national minimum criteria for existing CCR landfills, surface impoundments, and lateral extensions of those units under 40 CFR 257, Subpart D, which is commonly known as the CCR Rule. Facilities regulated under the CCR Rule are required to install and sample a groundwater monitoring well network to be analyzed for a prescribed list of constituents to evaluate whether its CCR unit has impacted downgradient groundwater quality. The monitored constituents under the CCR Rule are listed in Appendix III (*Constituents for Detection Monitoring*) and Appendix IV (*Constituents for Assessment Monitoring*) to 40 CFR Part 257 and **Table 1** of this Plan. At a minimum, the groundwater monitoring network must include one upgradient and three downgradient monitoring wells in relation to the location of the CCR unit. The groundwater monitoring system requirements are described in 40 CFR §257.91.

The CCR Rule establishes a multi-phase approach for the monitoring of groundwater. Among others, this approach provides for groundwater sampling, analysis, and statistical evaluation of the data and whether further assessment monitoring and corrective action are warranted.
The groundwater monitoring phases listed under the CCR Rule are as follows:

1. Detection Monitoring, which consists of:
    a. Initial eight rounds of monitoring to establish background levels; and
    b. Semiannual Detection Monitoring events (following the initial eight events).
2. Assessment Monitoring, if required.
3. Corrective Action Monitoring (following implementation of corrective measures, if any).

Table 1. Monitored Constituents under the CCR Rule

Appendix III to 40 CFR Part 257 – *Constituents for Detection Monitoring*	**Appendix IV to 40 CFR Part 257 –** *Constituents for Assessment Monitoring*
Boron	Antimony
Calcium	Arsenic
Chloride	Barium
Fluoride	Beryllium
pH	Cadmium
Sulfate	Chromium
Total Dissolved Solids	Cobalt
	Fluoride
	Lead
	Lithium
	Mercury
	Molybdenum
	Selenium
	Thallium
	Radium 226 and 228 combined

The groundwater monitoring program begins by conducting eight independent sampling events where groundwater samples are collected from each upgradient and downgradient well in the groundwater monitoring network. Groundwater samples are analyzed for the constituents listed in Appendices III and IV, and site-specific background levels are calculated from the groundwater dataset obtained from the sampling of the upgradient/background wells. Following the establishment of background levels, detection monitoring for the constituents listed in Appendix III is performed at least semiannually.

The groundwater detection monitoring phase progresses to the next monitoring phase (*i.e.*, assessment monitoring) if statistical evaluation of the constituents listed in Appendix III identifies a statistically significant increase (SSI) above the established background levels for any of the constituents, and it cannot be demonstrated that the increase is attributable to naturally occurring variations in groundwater quality, other sources of contamination, or sampling or analytical error.

If assessment monitoring is warranted, groundwater protection standards (GWPS) must be calculated for each detected constituent listed in Appendix IV. Assessment monitoring consists of an annual sampling event for the analysis of all constituents listed in Appendix IV and semiannual sampling events for all Appendix III constituents and Appendix IV constituents detected in the annual sampling event. If any of the Appendix IV constituents are identified at a statistically significant level (SSL) above the associated GWPS, the nature and extent of groundwater impact must be determined, and corrective action remedy implemented if it cannot be ruled out that the CCR unit has impacted the downgradient groundwater quality.

Following the implementation of corrective measures, a corrective action groundwater monitoring program must be established to document the effectiveness of the corrective action remedy and demonstrate compliance with the GWPS. As in assessment monitoring, corrective action monitoring consists of annual and semiannual sampling events to be analyzed for the constituents listed in Appendix III and Appendix IV to 40 CFR Part 257.

The following section contains a detailed description of the statistical methods to be applied at each CCR groundwater monitoring phase, as applicable.

2 STATISTICAL EVALUATION PROCEDURES

Statistical procedures will be performed in accordance with the USEPA guidance document entitled *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance* (USEPA, 2009), commonly referred to as the Unified Guidance. Graphical and statistical analyses will be conducted using Sanitas™ Statistical Software or similar software (*e.g.*, ProUCL, R statistical software, or others).

2.1 Reviewing and Preparing Data

The following statistical procedures for data screening and preparation will be performed on all upgradient and downgradient groundwater datasets, whether generated during detection, assessment, or corrective action monitoring.

2.1.1 Summary Statistics

Summary statistics (*e.g.*, mean, median, standard deviation) will be calculated for the available datasets. Graphical representations of descriptive statistics may be generated as appropriate.

2.1.2 Identification of Potential Outliers

Time series graphs and side-by-side box plots will be constructed for each well and constituent pair (well/constituent pair) to identify potential outliers visually. The Tukey's Outlier Screening test, Dixon statistical test, or similar procedure will be performed to confirm the presence or absence of outlier values. The Unified Guidance recommends that testing for outliers be performed, but outliers should not be generally removed unless an error or basis for the observed discrepancy can be identified. Potential sources of error may include sampling and analytical errors. Potential discrepancies may include inconsistent sample turbidity and values significantly outside the historical ranges of existing data. Even if excluded from statistical analyses, outlier values should be flagged and maintained in the database to be reevaluated as new data become available.

2.1.3 Temporal Trends

The least-squares linear regression, or the Sen's Slope/Man-Kendall procedure, will be performed to test if a significant temporal trend exists. The least-squares linear regression method will be used when the dataset follows a normal or transformed normal distribution and when the dataset contains less than 15% non-detects. In addition, the regression residuals must be

normally distributed and show equal variance across time. Otherwise, nonparametric methods (*e.g.*, Sen's Slope/Mann-Kendall) will be used to test for significant linear trends.

2.1.4 Testing for Normality

The Shapiro-Wilk or similar test will be performed to test for normality. Whenever possible, non-normally distributed data will be transformed into normally distributed data using the Ladder of Powers procedure. In this method, the data is submitted to the following transformation sequence: x, $x^{1/2}$, x^2, $x^{1/3}$, x^3, ln(x), x^4, x^5, x^6, until a suitable transformation is applied to normalize the data.

2.1.5 Handling of Datasets with Non-Detect Results

Where available, estimated results less than the RL (i.e., "J" flagged data) will be used in the statistical evaluation. Groundwater analytical data with non-detect results will be handled as follows:

- Datasets containing less than 15% non-detects will be replaced with one-half of the reporting limit (RL). The reporting limit to be used for non-detects will be the practical quantitation limit (PQL) as reported by the analytical laboratory (typically identified as "RL" in laboratory analytical reports).
- Datasets containing between 15-50% non-detects will be submitted to the Kaplan-Meier adjustments, regression of order statistics (ROS) adjustments, or similar tests. These methods adjust the mean and standard deviation of the dataset to account for the non-detect values.
- Nonparametric statistics will be used on datasets containing more than 50% non-detects. Non-detects will be set at the RL (*i.e.*, PQL) for statistical testing.
- Note that statistical analyses are not required on well/constituent pairs containing 100% non-detects (refer to the Unified Guidance 2009, Chapter 6).

2.2 Detection Monitoring

During detection monitoring, analytical results will be statistically evaluated using the prediction interval method [40 CFR §257.93(f)93)]. Interwell prediction limits,[1] combined with a 1-of-2 resample plan, have been selected to meet the USEPA's requirement of maintaining a 10% annual sitewide false positive rate (SWFPR) and adequate statistical power.

[1] The method of interwell comparisons (i.e., comparisons of downgradient to upgradient well data) was selected over intrawell comparisons (i.e., comparisons of recent well data to historic background data from the same well) given that groundwater background data did not exist prior to CCR unit placement at AES-PR, and CCR impacts to downgradient wells could not be ruled out based on the downgradient wells concentrations of Appendix III constituents detected following the initial phase of detection monitoring (*See* footnote number 2, below).

2.2.1 Establishing and Updating Background

Upgradient well data will be used to establish background levels for each individual Appendix III constituent.[2] Initially, the dataset from the upgradient wells will be statistically evaluated and handled following the procedures described in **Section 2.1**, *Preparing and Reviewing Data*.

Groundwater constituent concentrations from the pooled upgradient well dataset will be used to compute the upper prediction limit (UPL) for each Appendix III constituent. Parametric prediction limits will be computed when the background data follow a normal or transformed-normal distribution. Nonparametric prediction limits will be calculated when the background data do not follow a normal or transformed-normal distribution or when more than 50% of the data consists of non-detects.

As new background data becomes available, it will be statistically evaluated to verify if the new dataset is representative of existing background values. The Unified Guidance recommends that background values be updated when four to eight new measurements are available to allow for statistical evaluation of the new dataset against the existing dataset. Besides statistically testing for significant trends and outliers, as described in **Section 2.1**, the Welch's *t*-test, or the nonparametric Mann-Whitey test (also known as the Wilcoxon rank-sum test) or similar procedure, should be used to test the new dataset against the existing dataset. If Welch's *t*-test or Mann-Whitney test finds no significant difference between the two groups, then the new data should be combined with the existing background data to calculate an updated UPL. Generally, the level of significance for the Welch's *t*-test is set at an alpha level equal to 0.01 ($\alpha = 0.01$), whereas that for the Mann-Whitney test is set at $\alpha = 0.05$ (if five or more new observations are available, alpha may be set at $\alpha = 0.01$). In case of a significant Welch's *t*-test or Wilcoxon rank-sum test result, a closer investigation of the available data should be performed to determine whether existing or new background datasets are more representative of the current groundwater conditions.

2.2.2 Evaluating Statistically Significant Increases (SSIs)

Once background prediction limits are calculated, upgradient-to-downgradient interwell comparisons will be conducted by comparing the downgradient groundwater sampling results to the prediction limits computed as background concentrations. That is, the concentration of each constituent in individual downgradient wells will be compared to the corresponding background level to determine if a statistically significant increase (SSI) over background exists. An SSL is identified for a given well/constituent pair when the constituent concentration in the downgradient well is higher than the associated background UPL.[3] The detection monitoring program will be based on a 1-of-2 resample plan per the Unified Guidance (*i.e.*, a second independent sample may be collected and analyzed to confirm an initial SSI determination). The

[2] The initial phase of detection monitoring to establish background levels, which consisted of eight rounds of groundwater samples from the monitoring well network at AES-PR, was completed by October 17, 2017 [40 CFR §257.94(b)].

[3] Background pH levels have UPL and lower prediction limit (LPL) values. A statistically significant result is identified when the pH value in a downgradient well is higher than the background UPL or lower than the background LPL.

1-of-2 resample plan will help achieve the USEPA statistical requirements of an annual sitewide false positive rate (SWFPR) of 10% and adequate statistical power.

For any constituent, a confirmed determination of SSI over background may trigger assessment monitoring in the absence of evidence of natural variation, sampling or analytical error, or other sources of contamination.

2.3 Assessment Monitoring

In assessment monitoring, groundwater data is typically compared against a fixed numerical standard, which is established as a groundwater protection standard (GWPS). If assessment monitoring is warranted, the groundwater data will be statistically evaluated using confidence intervals around the mean for parametric or around the median for nonparametric testing. Confidence interval analysis is recommended in the Unified Guidance when comparing compliance well data against a fixed numerical value (*i.e.*, GWPS) to identify the presence or absence of an SSL.

The datasets from upgradient and downgradient wells will be statistically evaluated and handled following the procedures described in **Section 2.1** before computing GWPS and confidence intervals. Individual downgradient well data from each detected Appendix IV constituent will be used to construct confidence intervals and compared against the associated GWPS as described below.

2.3.1 Establishing Groundwater Protection Standards

During assessment monitoring, downgradient well concentrations of detected Appendix IV constituents are statistically compared to the corresponding GWPS. The GWPS for all detected Appendix IV constituents will be calculated in accordance with 40 CFR §257.95(h).

Pursuant to 40 CFR §257.95(h) and the USEPA amendments to 40 CFR §257.95 of July 30, 2018,[4] which promulgated CCR-Rule specified numeric criteria for cobalt (0.006 mg/L), lead (0.015 mg/L), lithium (0.040 mg/L), and molybdenum (0.100 mg/L), the GWPS will be:

- The maximum contaminant level (MCL) established under §§141.62 and 141.66 of 40 CFR Part 257;
- The CCR-Rule specified numeric criteria for constituents for which an MCL has not been established (i.e., cobalt, lead, lithium, and molybdenum); or
- The corresponding background concentration when the background level is higher than the MCL or CCR-Rule specified numeric criteria (see below).

The Upper Tolerance Limit (UTL) will be used to calculate the site background level for each Appendix IV constituent using the pooled upgradient-well data. The parametric UTL, with 95% confidence and 95% coverage, will be calculated for normal or transformed-normal distributions. Nonparametric upper tolerance limits will be calculated when the distribution of the background

[4] *See* Federal Register/Vol. 83, No. 146/Monday, July 30, 2018/Rules and Regulations.

data is not normal or transformed-normal or when the dataset contains more than 50 percent of non-detects. In such cases, the nonparametric UTL will be set at the highest value in the background dataset. When the background dataset contains 100% non-detects, the UTL will be the laboratory reporting limit (*i.e.*, PQL). Appendix IV background values will be updated as described in **Section 2.2.1**, except that the upper tolerance limit will be used to calculate background levels.

2.3.2 Evaluation of Statistically Significant Levels

Under assessment monitoring, the presumption is that the average concentrations of Appendix IV constituents are at or below their respective GWPS unless demonstrated otherwise. Therefore, a statistically significant level (SSL) is detected when the lower confidence limit (LCL) of the mean, or median, exceeds the associated GWPS.

For normal or transformed-normal distributions, a 95% LCL will be constructed from recent data.[5] Once the number of available observations (*i.e.*, results) exceeds 19 data points, the 99% LCL may be computed instead. Nonparametric LCL will be calculated for datasets with greater than 50% non-detects and for datasets that do not follow a normal or transformed-normal distribution. The confidence interval for nonparametric LCL will be set based on the available number of observations.

For downgradient well data exhibiting a statistically significant temporal trend, the confidence interval will be plotted as confidence bands around the predicted trend line. The least-squares linear regression, Sen's Slope/Mann-Kendall, or similar procedures will be performed to test if a significant linear trend exists. **Section 2.1.3** contains a description of the procedures and requirements to test for statistically significant temporal trends. If a statistically significant trend is detected, the LCL (lower bound of the confidence band) will be compared against the GWPS. An SSL is detected when the LCL of the confidence band exceeds the associated GWPS.

If an SSL is detected for one or more Appendix IV constituents, and if it cannot be demonstrated that the increase is attributable to naturally occurring variations in groundwater quality, other sources of contamination, or sampling or analytical error, the nature and extent of the groundwater impact for constituents with SSLs must be undertaken [40 CFR 257.95(g)(1)]. Within 90 days of detecting an SSL for any of the Appendix IV constituents, an assessment of corrective measures must be initiated [40 CFR 257.96(a)], a remedy must be selected [40 CFR 257.97], and corrective action groundwater monitoring program established [40 CFR 257.98(a)(1)] once the selected remedy has been implemented.

[5] Statistical evaluation should be performed on datasets that are representative of existing groundwater quality conditions at the time of evaluation. For example, if a shift (jump) in the mean concentration of a constituent of concern is observed, and the new mean concentration is deemed statistically to be more representative of actual site conditions, then the newer dataset should be used in statistical analysis. Although four data points are the minimum number of observations required to construct a confidence interval, the Unified Guidance recommends at least eight observations.

2.3.3 Comparing Data to Background

Besides performing a statistical evaluation to identify potential SSLs, the downgradient concentrations of the CCR constituents are frequently compared to the background levels. Confidence intervals for each constituent and downgradient well will be constructed from recent data and compared to the respective background upper tolerance limit (UTL) to determine if Appendix III and Appendix IV constituents are at or below background levels. When the upper confidence limit (UCL) is below the background UTL for two consecutive sampling events, it can be concluded that concentrations are at or below background, and the CCR unit may return to detection monitoring [40 CFR 257.95(e)].

2.4 Corrective Action Monitoring

In corrective action monitoring, groundwater data is typically compared against a fixed numerical standard, which is established as a GWPS. The groundwater data will be statistically evaluated using confidence intervals around the mean or median. Confidence interval analysis is the method recommended in the Unified Guidance when comparing compliance well data against a fixed numerical value (*i.e.*, GWPS) to identify the presence or absence of an SSL.

The datasets from upgradient and downgradient wells will be statistically evaluated and handled following the procedures described in **Section 2.1** before computing GWPS and confidence intervals.

2.4.1 Groundwater Protection Standards

During corrective action monitoring, downgradient well concentrations of detected Appendix IV constituents are statistically compared to the GWPS calculated during assessment monitoring pursuant to 40 CFR §257.95(h). As described in **Section 2.3.1**, the GWPS will be as follows: the MCL; the CCR-Rule specified numeric criteria for constituents for which an MCL has not been established (*i.e.*, cobalt, lead, lithium, and molybdenum); or the background concentration when the background level is higher than the MCL or CCR-Rule specified numeric criterion.

The Upper Tolerance Limit (UTL) will be used to calculate the site background levels for each Appendix IV constituent using the pooled upgradient-well data. The parametric UTL, with 95% confidence and 95% coverage, will be calculated for normal or transformed-normal distributions. Nonparametric upper tolerance limits will be calculated when the distribution of the background data is not normal or transformed-normal or when the dataset contains more than 50 percent of non-detects. In such cases, the nonparametric UTL will be set at the highest value in the background dataset. When the background dataset contains 100% non-detects, the UTL will be the laboratory reporting limit (*i.e.*, PQL). Background values for Appendix IV constituents will be updated as described in **Section 2.2.1**, except that the upper tolerance limit will be used to calculate background levels.

2.4.2 Evaluation of Statistically Significant Levels and Effectiveness of Remedy

Under corrective action monitoring, one or more Appendix IV constituents have been demonstrated to exceed their respective GWPS (i.e., an SSL has been identified). Therefore, the

selected remedy is deemed successful when the upper confidence limit (UCL) of the average concentration of the constituent of concern is less than the GWPS.

For normal or transformed-normal distributions, a 95% UCL will be constructed from recent data. Once the number of available observations (*i.e.*, results) exceeds 19 data points, the 99% UCL may be computed instead. Nonparametric UCL will be calculated for datasets with greater than 50% non-detects and for datasets that do not follow a normal or transformed-normal distribution. The confidence interval for nonparametric UCL will be set based on the available number of observations.

For downgradient well data exhibiting a statistically significant temporal trend, the confidence interval will be plotted as 95% confidence bands around the predicted trend line. The least-squares linear regression, Sen's Slope/Mann-Kendall, or similar procedures will be performed to test if a statistically significant linear trend exists. **Section 2.1.3** describes the procedures and requirements to test for statistically significant temporal trends. If a statistically significant trend is detected, the UCL (upper bound of the confidence band) will be compared against the GWPS.

A remedy is considered complete when the upper confidence limits constructed for Appendix IV constituents in wells identified with SSLs have not exceeded the GWPS for three consecutive years [40 CFR 257.98(c)(2)] at all points within the impact plume that lie beyond the monitoring well system as established under 40 CFR 257.91 [40 CFR 257.98(c)(1)]. In that case, the CCR unit may return to assessment monitoring.

3 REFERENCES

DNA (DNA-Environment, LLC). August 2017. *Groundwater Monitoring System & Sampling and Analysis Program, AES Puerto Rico LP, Guayama, Puerto Rico.*

DNA (DNA-Environment, LLC). October 2023. *Federal CCR Corrective Action Groundwater Monitoring Plan, AES Puerto Rico LP, Guayama, Puerto Rico*.

USEPA (United States Environmental Protection Agency). 1998. *Ground Water Sampling Procedure. Low Stress (Low Flow) Purging and Sampling*.

USEPA (United States Environmental Protection Agency). 2009. *Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Unified Guidanc*e. Washington, DC: EPA. EPA 530/R-09-007.