STATISTICAL DATA ANALYSIS WORK PLAN
Coal Combustion Residuals Rule Groundwater Monitoring System Compliance
Four Corners Power Plant
Fruitland, New Mexico

Prepared for:

Arizona Public Service Company
400 North 5th Street
Phoenix, Arizona 85004

Prepared by:

Wood Environment & Infrastructure Solutions, Inc.
Phoenix, Arizona

June 10, 2020

Project No. 14-2020-2068

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

%	percent
§	Section
amsl	above mean sea level
ANOVA	analysis of variance
APS	Arizona Public Service
BTV	Background Threshold Value
CCR	coal combustion residuals
CFR	Code of Federal Regulations
CSM	Conceptual Site Model
CWTP	Combined Waste Treatment Pond
DFADA	Dry Fly Ash Disposal Area
DQR	Double Quantification Rule
EDA	exploratory data analysis
FCPP	Four Corners Power Plant
ft	foot, feet
GWPS	groundwater protection standard
KM	Kaplan-Meier
LAI	Lined Ash Impoundment
LDWP	Lined Decant Water Pond
LPL(s)	lower prediction limit(s)
MCL	Maximum Contaminant Level
Multiunit 1	CCR multiunit comprised of LAI and LDWP
RWP	Return Water Pond
RL	reporting limit
ROS	regression order on statistics
SDAWP	Statistical Data Analysis Work Plan
SSI	statistically significant increase
SSL	statistically significant level
SWFPR	site-wide false positive rate
UPL(s)	upper prediction limit(s)
UTL(s)	upper tolerance limit(s)
URS	Upper Retention Sump
TDS	total dissolved solids
VSP	Visual Sampling Plan
USEPA	United States Environmental Protection Agency

1.0 INTRODUCTION

This Statistical Data Analysis Work Plan (SDAWP) was prepared by Wood Environment & Infrastructure Solutions, Inc. (Wood) on behalf of Arizona Public Service (APS) for the Four Corners Power Plant (FCPP) located in Fruitland, New Mexico. The SDAWP details the scope and implementation of statistical criteria and procedures to evaluate site data in accordance with Coal Combustion Residuals (CCR) groundwater monitoring and corrective action requirements detailed in 40 Code of Federal Regulations (CFR) Sections (§) 257.90 through 257.98 (herein referred to as the CCR Rule) (Federal Register, 2018).

This SDAWP was updated in October 2018 to incorporate the evaluation of assessment monitoring data. Minor organizational and editorial changes to existing sections of the report were also made for clarity and readability. The SDAWP was updated again in May 2020 to reflect the addition of the Return Water Pond (RWP) CCR Unit and associated monitoring well network to the FCPP CCR groundwater monitoring program.

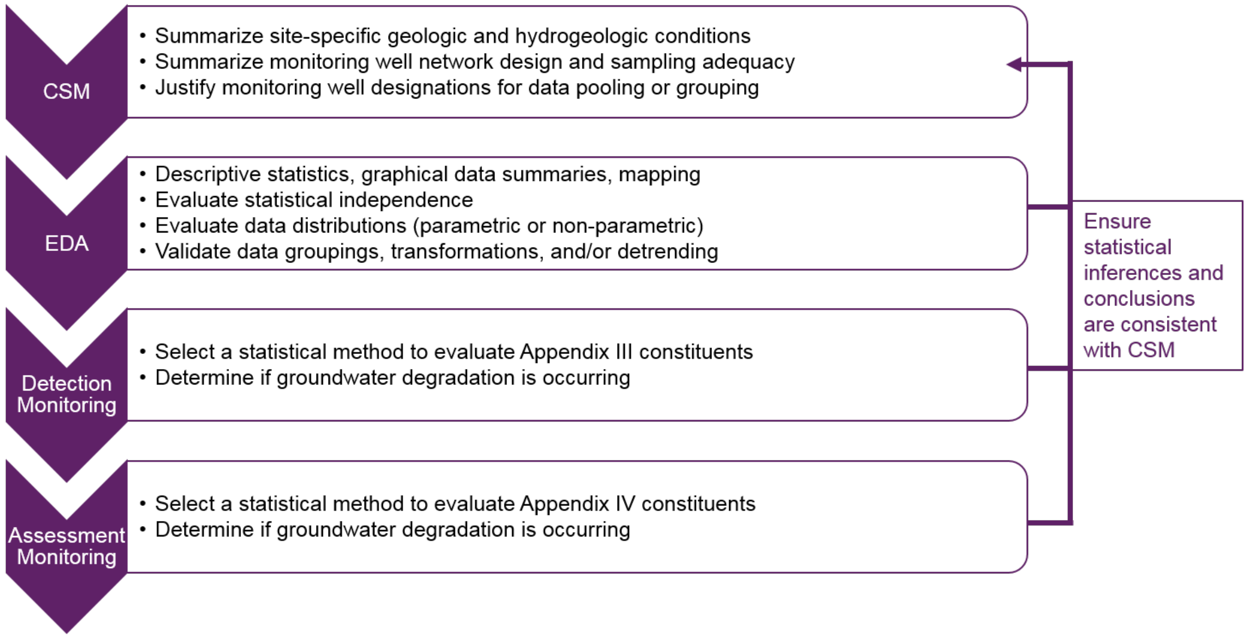
1.1 Objectives

The SDAWP will serve as a reference document throughout the FCPP CCR groundwater monitoring program to:

- Assess the adequacy of sampled data to service statistical procedures (Sections 1.0 and 2.0);
- Select appropriate statistical methods for each constituent in each monitoring well (Sections 2.0 through 5.0);
- Develop background constituent concentration levels, otherwise known as Background Threshold Values (BTVs) (Section 3.0);
- Develop groundwater protection standards (GWPSs) (Section 4.0);
- Identify statistically significant increases (SSIs) in constituent concentrations over BTVs and GWPSs (Sections 3.0 through 5.0); and
- Make recommendations for future sampling and data evaluations (Section 6.0).

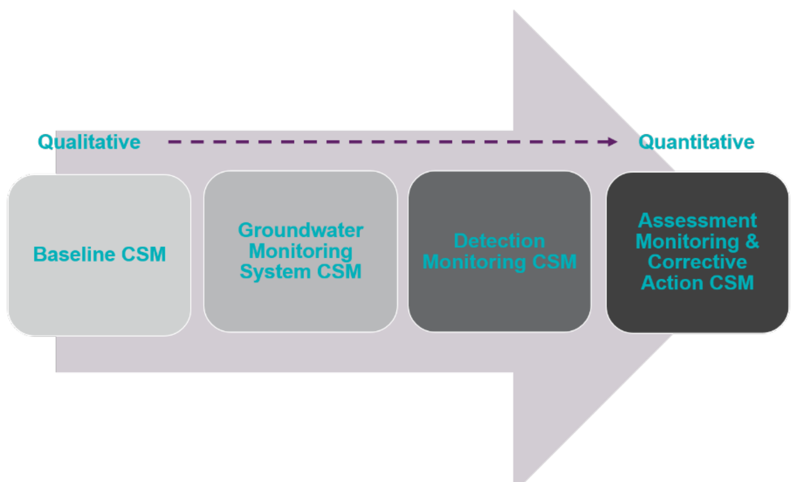
1.2 Purpose

The purpose of this SDAWP is to prescribe a comprehensive workflow that allows practitioners to defensibly evaluate groundwater data and assess if groundwater quality at the FCPP meets the criteria set forth in the CCR Rule. The general workflow for this SDAWP is outlined on the following page:



1.3 Conceptual Site Model

CCR groundwater monitoring systems must collect the right type, quantity, and quality of data to adequately and defensibly assess groundwater quality as set forth in the CCR Rule. Although certification of the FCPP CCR groundwater monitoring systems is being conducted independent of this SDAWP, a baseline conceptual understanding of the site's industrial activities, geology, and hydrogeology is necessary to assess the adequacy of the groundwater monitoring system to sample representative data and statistically evaluate whether groundwater has been adversely impacted by leakage from one or more site CCR units.



The Conceptual Site Model (CSM) constitutes a 'living representation' of a site that helps project members hypothesize, visualize, interpret, and understand site-specific information (USEPA, 2011). This information is utilized throughout different stages of the project lifecycle to make informed decisions regarding monitoring system design, data evaluation, corrective actions, and/or site closure. A baseline CSM establishes a reconnaissance understanding of the site using a framework of preexisting site-specific information that portrays both known and hypothesized information about the site. Development of a baseline CSM for the site is necessary for developing the groundwater monitoring systems. The baseline CSM is used to help determine if the groundwater monitoring system(s) meets the criteria set forth in 40 CFR §257.91 (b)(1) and §257.91(b)(2) and is updated as needed throughout the life of the project.

Wood has relied upon the CCR monitoring well network certification reports (AECOM, 2017 and Wood, 2020) for details of both the previously documented baseline and updated hydrogeological CSMs used to design the CCR groundwater monitoring systems for CCR Units at the FCPP and evaluate collected data. Salient information regarding the hydrogeologic CSM is extracted from these reports (unless noted otherwise) and summarized in the following subsections to document:

- Preexisting site-specific information;
- The adequacy of groundwater monitoring networks to assess groundwater quality; and
- The appropriateness of background and downgradient well classifications for statistically evaluating whether groundwater has been affected by leakage from one or more site CCR units.

This CSM may be refined based on the results of the statistical evaluation of water quality data.

1.3.1 Site Description

The site setting is as follows:

- FCPP is an operating power plant owned by APS and four other utilities:
  - FCPP burns low sulfur coal in two electrical generating units (Units 4 and 5) and has a net generating capacity of 1,540 megawatts.
  - Coal burned at the plant is generally sourced from the nearby Navajo Mine (Navajo Transitional Energy Company, 2016).
  - The plant and associated infrastructure are located on land leased from the Navajo Nation, approximately 20 miles southwest of Farmington, New Mexico (Figure 1).
  - FCPP is situated on the southern bank of Morgan Lake, an approximately 1,300-acre man-made lake that has a maximum storage capacity of 39,000 acre-feet (ft) of water and supplies cooling water to the plant. Morgan Lake was formed by damming a westerly flowing stream (now known as 'No Name Wash') and is replenished by an underground pipeline that routes flow from the San Juan River located approximately 3 miles north of the FCPP. The typical water surface elevation of the lake is 5,330 ft above mean sea level (amsl).
  - Plant infrastructure includes one CCR multiunit (referred to as Multiunit 1) and four single CCR units (Table 1) which are located in the main plant area and to the west of the plant within the FCPP lease boundary (Figure 2).
- The plant is located in a semi-arid climate on the western flank of the San Juan Basin:
  - FCPP is located at an elevation of approximately 5,340 to 5,360 ft amsl in the Colorado Plateau physiographic province of northwestern New Mexico.
  - The area receives an average of 8.6 inches of rain and 12.6 inches of snow per year.
  - The San Juan Basin is a structural depression that lies at the eastern edge of the Colorado Plateau (Dames & Moore, 1988).
  - The dominant geographic feature in the vicinity of FCPP is the Hogback Monocline located to the west of the plant; this monocline is a steep (38 degree) eastward-dipping flank composed of Cretaceous sedimentary rock (Dames & Moore, 1988).
  - The topography of the FCPP area is characterized by rolling terrain, steep escarpments, and incised drainages/arroyos. In the vicinity of the plant, the ground surface is relatively flat,

sloping to the west at approximately 20 ft per mile; however, surface drainage immediately near Morgan Lake flows toward the lake. About one mile west of the plant, the level ground surface drops rapidly to 5,200 ft amsl. Chaco Wash (a.k.a. Chaco River) is located west of this abrupt change in elevation and ephemerally flows north to the San Juan River. Morgan Dam discharges to 'No Name Wash' which flows west of the lake to Chaco Wash.

- The relatively higher-elevation area where the plant operations are located is referred to as the "plant area"; the relatively lower-elevation area west of the plant where the DFADA and Multiunit 1 CCR Units are located is often referred to as the "disposal area".

1.3.2 Site Geology

There are two 'uppermost geologic units' that underlie the FCPP site and immediate vicinity. These units are expected to influence groundwater flow and variations in naturally occurring constituent concentrations across the site. The units are as follows:

- **Pictured Cliffs Sandstone**: The Pictured Cliffs Sandstone is the uppermost geologic unit beneath the plant and the CCR units located in this vicinity (i.e., the Upper Retention Sump [URS], Return Water Pond [RWP], and the Combined Waste Treatment Pond [CWTP] as depicted in Figure 2). This geologic unit is a fine- to medium-grained marine sandstone. The lower portions of the Pictured Cliffs Sandstone represent a transitional sequence between this formation and the underlying Lewis Shale as indicated by alternating thin beds of very fine-grained sandstone and silty shale. The Pictured Cliffs Sandstone forms a capstone on an exposed cliff face located between the plant site and the CCR units located to the west (i.e., the Lined Ash Impoundment [LAI], Lined Decant Water Pond [LDWP], and the Dry Fly Ash Disposal Area [DFADA]).
- **Lewis Shale**: The Lewis Shale is a marine shale that contains evaporite deposits resulting in naturally occurring saline groundwater conditions. The Lewis Shale is the uppermost geologic unit that underlies the LAI, LDWP, and DFADA and spans west of the Pictured Cliffs Sandstone cliff face approximately 1.5 miles westward to the base of the Hogback Monocline. The regional thickness of the Lewis Shale is approximately 500 ft and is underlain by Cliff House Sandstone. The Lewis Shale consists of a weathered shale subunit overlying a hard, unweathered shale subunit. The thickness of the weathered shale varies between 11 and 47 ft with an average thickness of 30 ft within the vicinity of the site (Dames & Moore, 1988). The weathered shale is not as thick when overlain by Pictured Cliffs Sandstone in the vicinity of the plant site and can be difficult to differentiate within the fine-grained rocks that comprise the gradational contact between the Pictured Cliffs Sandstone and underlying Lewis Shale. The weathered shale contains thin sandstone lenses that vary in thickness from one to seven ft; the sandstone is fine to very fine-grained and cemented by calcium carbonate (Dames & Moore, 1988). The unweathered shale is significantly less permeable than the weathered shale. The unweathered shale is very fine-grained to silty, and contains periodic siltstone and sandstone lenses (Dames & Moore, 1988). The surface of the unweathered shale slopes towards the Chaco Wash at approximately the same slope as land surface (Dames & Moore, 1988) but displays some irregularity resulting in varying levels of saturated thickness in the weathered shale. The Lewis Shale is variably saturated and hydraulically interconnected with alluvial deposits of Chaco Wash. The low-permeability unweathered shale underlying the Pictured Cliffs Sandstone results in a perched saturated zone beneath the plant.

1.3.3 Site Hydrogeology

Three general hydrostratigraphic units are conceptualized beneath the FCPP and associated CCR units that have the potential to interact with releases from CCR units. These hydrostratigraphic units form the basis for the hydrogeologic CSM developed by AECOM (2017) and Wood (2020) for the purpose of designing the site CCR monitoring systems and establish the working basis for statistically evaluating groundwater conditions underlying the site.

The first hydrogeologic unit (Pictured Cliffs Sandstone) is dominant only under the plant area, which is located in an elevated area south of Morgan Lake (Figure 2). Three CCR units (i.e., the former URS, RWP, and CWTP) reside within this area. The Pictured Cliffs Sandstone is the uppermost water bearing unit for the plant area and extends from ground surface (between approximately 5,340 to 5,360 ft amsl) to approximately 5,300 ft amsl in the plant area. Groundwater in this area generally flows northward towards Morgan Lake, which has a surface elevation of approximately 5,330 ft amsl. Construction and operations of the plant have resulted in disturbed surface conditions and associated impacts to groundwater are not well understood. This uncertainty will be considered when interpreting constituent concentrations and any potential impact on adequacy of background well locations.

The second hydrogeologic unit (Weathered Lewis Shale/Alluvium) underlies the Pictured Cliffs Sandstone in the plant area and the Multiunit 1 and the DFADA CCR units (Figure 2) in the disposal area, approximately one mile west of the plant. The weathered Lewis Shale and the hydraulically-connected alluvial deposits along Chaco Wash are designated as the uppermost water bearing unit in the disposal area. Although the Lewis Shale is geologically continuous in this area, it is unsaturated in the vicinity of the DFADA. The water table in the weathered Lewis Shale can exhibit local seasonal fluctuations that are attributed to interactions between rates of groundwater recharge and discharge (Dames & Moore, 1988) from/to Morgan Lake, historical unlined ponds, and Chaco Wash. Groundwater flow generally follows the surface topography and descends to the west-southwest in the disposal area, mainly in the weathered shale and in local alluvial channels that drain toward the Chaco Wash (APS, 2013).

The third hydrogeologic unit (Unweathered Lewis Shale) consists of the unweathered Lewis Shale and is a regionally-extensive confining unit that forms the base of the uppermost aquifers in the plant and disposal areas. Although minor amounts of water may be present in the Unweathered Lewis Shale, this unit is thick (hundreds of feet) and acts as an aquitard between the Weathered Lewis Shale/Alluvium and the underlying Cliff House Sandstone.

1.4 Monitoring System Sampling Adequacy

Multiple monitoring well systems are in place at the FCPP to monitor groundwater conditions beneath the five site CCR units. The installation of these networks is summarized in two reports, both of which identify the systems as compliant with 40 CFR §257.91(a) through (e) (AECOM, 2017 and Wood, 2020). AECOM also prepared a *Sampling and Analysis Plan, Coal Combustion Residual (CCR) Groundwater Monitoring* (2015) to document the methods and procedures used to conduct groundwater sampling and evaluate potential impacts of site CCR units. This Sampling and Analysis Plan was updated by APS in January 2018 and included in the *Annual Groundwater Monitoring and Corrective Action Report for 2017* (Amec Foster Wheeler, 2018).

Sampling coverage and adequacy of the CCR monitoring well networks to facilitate the statistical evaluations detailed in this SDAWP are discussed in the following subsections.

1.4.1 Downgradient Groundwater Monitoring Well Networks

A total of 27 downgradient CCR system monitoring wells are in place at the site to monitor the downgradient groundwater conditions of each CCR unit (Table 2). Eleven of these monitoring wells are installed in the Lewis Shale. The remaining 16 wells are completed in the Pictured Cliffs Sandstone. These wells are grouped by respective CCR unit, as described below:

- **URS Downgradient Wells (Pictured Cliffs Sandstone)**: At the time the downgradient wells at the URS were installed, the groundwater flow direction underlying the URS was radially outward from the CCR unit. On this basis, five wells, MW-66, MW-67, MW-68, MW-69, and MW-70 were installed around the perimeter of the URS. Following removal of the URS and replacement of the unit with a concrete tank, mounding in the vicinity of the former URS subsided; the subsequent direction of groundwater flow was determined to be to the northwest (Wood, 2019), towards Morgan Lake. In 2018, to characterize the extent of impacts from the former unit, monitoring wells MW-83, MW-84, MW-85, and MW-86 were installed downgradient of the unit. Each of these wells are screened within the Pictured Cliffs Sandstone. The grouping of monitoring wells, spatial density, and coverage of the monitoring well network appear representative and adequate.
- **RWP Downgradient Wells (Pictured Cliffs Sandstone)**: The RWP is underlain by the Pictured Cliffs Sandstone hydrostratigraphic unit, which is unsaturated beneath the RWP. The next underlying aquifer (in the Cliff House Sandstone) is separated from the CCR unit by several hundred feet of Unweathered Lewis Shale, a regional aquitard. Thus, the groundwater monitoring system is designed to detect potential releases to the Pictured Cliffs Sandstone. Hydrogeologic conditions suggest that a release from the RWP would migrate vertically downward through the permeable and weathered rocks of the Pictured Cliffs Sandstone to the aquitard created by the Unweathered Lewis Shale and laterally along the surface of the Unweathered Lewis Shale in the northeast dip direction of the unit (Wood, 2020). Therefore, three monitoring wells, MW-88, MW-89, and MW-90 were installed around the downgradient (northeastern) edge of the RWP in the Pictured Cliffs Sandstone and screened directly above the aquitard formed by the Unweathered Lewis Shale. The grouping of monitoring wells, spatial density, and coverage of the monitoring well network appear representative and adequate, given the current understanding of the site.
- **CWTP Downgradient Wells (Pictured Cliffs Sandstone):** Similar to the URS, the groundwater flow direction underlying the CWTP was observed to be radially outward from the CCR unit at the time the monitoring system was installed. Four monitoring wells, MW-62, MW-63, MW-64, and MW-65 were installed around the perimeter of the CWTP. Each of these wells are screened within the Pictured Cliffs Sandstone. The grouping of monitoring wells, spatial density, and coverage of the monitoring well network appear representative and adequate, pending further review.
- **Multiunit 1 Downgradient Wells (Weathered Lewis Shale/Alluvium):** Six downgradient monitoring wells are in place below the toe of the western to southwestern edge of Multiunit 1: MW-7, MW-8, MW-40R, MW-61, MW-75 and MW-76 (Figure 2). Two wells, MW-40R and MW-76, are routinely either dry or have a limited saturated thickness which precludes sampling; the wells are included in the program in case conditions change in the future. In 2018, to characterize the extent of impacts from the unit, monitoring well MW-87 was installed downgradient of the unit at the western lease boundary near the Chaco Wash. The grouping of monitoring wells, spatial density, and coverage of the monitoring well network appear representative and adequate, pending further review. The screened interval for each well resides within the Weathered Lewis Shale/Alluvium.
- **DFADA Downgradient Wells (Weathered Lewis Shale/Alluvium):** Four existing wells are identified downgradient of the DFADA: MW-10, MW-13, MW-44, and MW-48. Each well, except

MW-48, is screened within the Weathered Lewis Shale/Alluvium. The screened interval for MW-48 resides within the Unweathered Lewis Shale. The downgradient DFADA wells are known to be dry; this groundwater monitoring system was designed to detect releases since the next underlying aquifer (in the Cliff House Sandstone) is separated from the CCR unit by several hundred feet of Lewis Shale, a regional aquitard. The grouping of monitoring wells, spatial density, and coverage of the monitoring well network appear representative and adequate, pending further review.

1.4.2 Background Groundwater Monitoring Wells

The purpose of comparison statistical tests is to assess if groundwater conditions downgradient of the CCR unit indicates a potential impact from the CCR unit. Therefore, it is important to adequately establish background conditions that accurately represent the quality of groundwater that has not been affected by a CCR unit (40 CFR §257.91).

Per the CCR monitoring well network certification reports (AECOM, 2017 and Wood, 2020), the following monitoring wells are designated as "background monitoring wells" for the respective geologic and hydrogeologic conditions underlying the FCPP:

- **Background Wells for the Pictured Cliffs Sandstone:** Three wells (MW-71, MW-72, and MW-73) are designated to assess background groundwater quality for the Pictured Cliffs Sandstone. MW-71 and MW-72 are upgradient and MW-73 is cross- to down-gradient of the URS, RWP, and CWTP (40 CFR §257.91[a][1] allows the inclusion of wells that are not hydraulically upgradient of the CCR unit when specific conditions are met). The grouping and adequacy of the three background wells to assess background water quality in the Pictured Cliffs Sandstone appear representative and adequate, pending further review.
- **Background Wells for the Weathered Lewis Shale/Alluvium:** Seven existing wells upgradient of Multiunit 1 and the DFADA, including MW-12R1 (which replaced the now-abandoned MW-12R), MW-43, MW-49A, MW-50A, MW-51, MW-55R and MW-74 are designated to assess background groundwater quality for Weathered Lewis Shale/Alluvium. One of these wells could be potentially affected by water from Multiunit 1 (MW-49A) based on its spatial proximity to the unit (AECOM, 2017). Five wells, MW-12R1, MW-43, MW-50A, MW-51, and MW-55R, are routinely either dry or have a limited saturated thickness which precludes sampling; the wells are included in the program in case conditions change in the future. The grouping and adequacy of these background wells to assess background water quality in the weathered Lewis Shale appear representative and adequate, pending further review.

Background can be established by a single monitoring well or a group of monitoring wells. If a group of monitoring wells is used, these wells should be screened within the same lithologic unit, exhibit similar groundwater chemistry, illustrate similar statistical characteristics, and be consistent with the CSM.

Due to the natural heterogeneity of the geologic and hydrogeologic conditions underlying the FCPP, background constituent concentrations are expected to be spatially heterogeneous (varying) across the site. The site is also expected to exhibit both spatial and temporal heterogeneity attributable to local climatic regimes, potential leakage from Morgan Lake, and potential operational activity at the site. The groundwater monitoring well networks, respective to sampling coverage and frequency, appear to adequately evaluate this spatial and temporal heterogeneity, pending further review.

The adequacy of designated background monitoring wells will be assessed using groundwater elevation data, boron data, total dissolved solids (TDS) data, a working understanding of the spatial heterogeneity of geochemistry underlying the FCPP, and statistical characteristics of constituents of concern. Historical groundwater chemistry data will be consulted during this evaluation but data preceding December 2011 will not be relied upon due to noted "matrix interference issues associated with saline waters" in samples analyzed prior to this date (APS, 2013).

2.0 EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis (EDA) is a diagnostic data evaluation step to assess the groundwater monitoring system's ability to collect the right quantity, quality, and type of data to adequately perform the statistical analyses set forth in 40 CFR §257.93. EDA occurs iteratively throughout the various sample acquisition stages and subsequent data evaluations. Two phases of EDA will occur within the scope of this SDAWP. Phase I EDA will occur after completing the first four rounds of sampling and will serve as a data screening step used to inform the data collection process. The second EDA phase will occur after the first eight rounds of sampling are complete and will continue throughout the monitoring program. Phase II EDA will service two objectives: 1) ensure the correct statistical method will be selected for determining background concentrations and performing statistical comparisons and 2) evaluate if the data meet the statistical inferences and criteria required to establish background threshold levels and perform statistical comparisons.

In general, the statistical inferences and criteria to complete groundwater monitoring under the CCR Rule include:

- the sampled data have no trend (i.e. are statistically stationary);
- the sampled data are statistically independent of each other;
- the sampled data are representative of a single statistical population; and
- the sampled data follow a discernable distribution.

The following subsections detail methods to determine if the data meet these assumptions. If these assumptions are not met, then data transformations will be explored, including detrending, data domaining, and data normalization. In cases where data transformations are ineffective, nonparametric statistics will be considered.

2.1 Data Evaluation Objectives

Diagnostic data evaluations allow practitioners to become familiar with sampled data to service three primary objectives. The first objective is to identify and resolve any anomalous data quality issues in a timely manner. The second objective is to identify data distributions and patterns that allow practitioners to make informed decisions when selecting a defensible statistical method to assess groundwater quality per 40 CFR §257.93 (f)(1) through (5). The third objective is to update the CSM with relevant information to make informed and defensible project decisions.

This SDAWP will implement the following methods as part of EDA. All Phase I EDA methods are applicable in Phase II and denoted by an asterisk in the following summary graphic.

Phase I EDA (after 4 sample rounds)

- Quick spatial interpolation (Section 2.5.1)*
- Time series plots (Section 2.5.3)*
- Histograms and box plots (Section 2.7)*

Phase II EDA (after 8 sample rounds)

- Autocorrelation (Section 2.5.2)
- Trend analysis (Section 2.5.3)
- Data transformations, if necessary (Section 2.6)
- QQ plots and fitting data distributions (Section 2.7)
- Outlier testing (Section 2.8)

2.2 Constituents of Concern

Within the scope of this SDAWP, the CCR Rule Appendix III constituents will be evaluated as part of the EDA process, including:

- Boron
- Calcium
- Chloride
- Fluoride
- pH
- Sulfate
- TDS

If there is an SSI declared at a site CCR unit for one or more of the Appendix III constituent concentrations, then the EDA process will ensue for the following Appendix IV constituents except fluoride (since fluoride has already been subjected to the EDA process as part of the Appendix III constituents):

- Antimony
- Arsenic
- Barium
- Beryllium
- Cadmium
- Chromium
- Cobalt

- Lead
- Lithium
- Mercury
- Molybdenum
- Selenium
- Thallium
- Radium 226 and 228 combined

Groundwater elevation, TDS, and boron will hold particular emphasis throughout the EDA process to assess the adequacy of background well classifications.

2.3 Non-Detects

Non-detects, also known as left-censored measurements, are values that cannot be quantified according to the laboratory method. There are several approaches for numerically representing non-detect data to complete the data evaluations listed within this SDAWP. For the purpose of this SDAWP, simple substitution and censor estimation techniques will be used to numerically represent non-detects. These methods will be selected according to sample size, frequency of detection, and method of data evaluation. Simple substitution and censor estimation techniques are described below.

Imputation for geospatial, geostatistical, and time series analyses (Section 2.4) will conform to the simple substitution criteria detailed in Section 2.3.1. Imputation for establishing background constituent concentrations (Section 3.0 and Section 4.0) and performing statistical comparisons will favor censor estimation techniques, where appropriate, and conform to the criteria set forth in this section.

2.3.1 Simple Substitution

Simple substitution is imputation using a qualitatively-derived value, usually equal to the reporting limit (RL), half the RL, zero, or method detection limit, for a non-detect measurement. The RL represents the lowest level that can be reported by a laboratory. For simple substitution, half the RL will be used if the concentration is undetected ("U" qualifier flag) or if samples are reported as detected but not quantified. Half the RL is assumed to be between zero and the RL, which reflects the maximum likelihood estimate of the mean or median of values uniformly distributed along the interval (i.e. 0 to the RL) (USEPA, 2009). Non-detects that are estimated ("J" qualifier flag) will respect the estimated value as a valid measurement (USEPA, 2009) for statistical purposes. For traditional statistical methods, simple substitution will be considered when the frequency of detection is greater than 85 percent (%) (USEPA, 2009) and/or the sample number is fewer than eight.

2.3.2 Censor Estimation

Censor estimation techniques rely on modeling the underlying data distribution to quantitatively model or estimate values for non-detect measurements. These techniques attempt to fit a sample to a known distribution using a censored estimation method, such as the Kaplan-Meier (KM) estimator or the robust regression on order statistics (ROS) (USEPA, 2009) and generate a model-based estimate of statistical moments or imputed number, respectively. Parametric statistical calculations are then performed using these model-based estimates or imputations. Parametric and nonparametric statistical methods are discussed in more detail in Section 2.6. For traditional statistical methods, censor estimation techniques will be implemented when the sample number is sufficient to discern the underlying data distribution (e.g. normal, lognormal, gamma), the frequency of non-detects are between approximately 10% and 50%, and the sample number is eight or more.

In cases where more than one RL is used, the ROS will be preferred method. Instances where the data do not conform to a discernable data distribution nor fit the criteria set forth in this section, nonparametric statistical methods will be used.

2.3.3 Double Quantification Rule

In cases where the background data are 100% non-detects, the Double Quantification Rule (DQR) is appropriate. The DQR states that if two consecutive samples exceed the RL, then there is enough evidence to declare an SSI (USEPA, 2009).

2.4 Spatio-Temporal Data Dependence

Environmental parameters and processes inherently influence the distribution, fate, and residence of constituents. These parameters and processes are oftentimes correlated in space and/or time, meaning sample data are not completely independent and exhibit some degree of spatial and/or temporal dependence, or correlation. Spatial and temporal EDA methods allow practitioners to evaluate spatial and/or temporal relationships, such as spatial distributions and temporal trends in constituents, over space and time. These methods are critical for: 1) visualizing data and further developing the CSM in terms of screening relationships between groundwater quality, geology, groundwater gradients, and seasonal trends; and 2) ensuring the sample data meet the statistical method assumptions listed under the CCR Rule.

This section discusses spatial and temporal EDA approaches for detecting and assessing data dependence. Section 2.5 discusses methods for managing data dependence to ensure the sample data meet the statistical assumptions.

2.4.1 Quick Spatial Interpolation

Application: Quick spatial interpolation screens for:

- spatial anomalies, dependence, and extents of constituent concentrations in groundwater;
- spatial associations between constituent concentrations and groundwater elevation; and
- changes in spatial groundwater gradients and CCR Rule constituent distributions over time and any potential anomalous data that may warrant further investigation or sampling.

EDA scope: Phase I and Phase II EDA

Selected methods: Selected methods include interpolation by natural neighbor, inverse distance weighted, splines (or other higher order polynomials), and/or nearest neighbor methods.

Interpolation is a generic term representing various methods used to generate maps, or spatial estimates of sampled data in unsampled locations. The quick interpolation methods listed are interpolators that do not make any assumptions regarding the distribution of the sampled data and require limited parameter input(s). More than one quick interpolation method may be selected to test the sensitivity of another quick interpolation method.
An adequate number and spacing of monitoring wells are necessary to map groundwater constituent concentrations. To facilitate meaningful mapping of groundwater constituents, monitoring wells assigned to each CCR monitoring system, in addition to geologically and hydrogeologically relevant FCPP monitoring wells not identified within the CCR monitoring systems, will be considered for quick spatial interpolation.

Quick interpolation maps of constituent concentrations and groundwater gradients will be integrated into the project CSM.

2.4.2 Autocorrelation

Application: Autocorrelation is used to:

- model and quantify the degree of spatial and/or temporal correlation between sampled data;
- identify sampling redundancies in the monitoring well network (in space and time); and
- optimize sampling frequency and monitoring network performance to reduce sampling redundancies.

EDA scope: Phase II EDA

Selected methods: Selected methods include the variogram model and lag plot.

Data dependence will be screened using quick interpolation methods. Data dependence will be quantified and tested using autocorrelation methods.

Autocorrelation quantifies the ability for a measured property, or constituent, to relate to itself in space or time. This notion follows Tolber's First Law of Geography, which states that "everything is related to everything else, but near things are more related than distant things." If values for nearby samples (either in space or time) are similar, then there is some autocorrelation among them, and therefore, the values contain varying degrees of redundant information. Autocorrelation is a valuable data evaluation tool for quantifying the presence of spatial and temporal dependence in sampled data.

Within the scope of this SDAWP, standard methods to quantify autocorrelation include the variogram or lag plot (USEPA, 2009). A lag plot is a useful EDA tool to screen for non-random (e.g. autocorrelated) variation in a sampled data set. If a data set exhibits spatial or temporal autocorrelation a pattern will appear in the lag plot.

The variogram model is useful for assessing sampling adequacy and autocorrelation. The variogram quantifies the ratio of dependent versus independent variation in the sampled data. This ratio is known as the nugget:sill ratio. If the nugget:sill ratio is less than 0.50, the data will be considered spatially or temporally dependent. A variogram model fits a range value to the sample data that represents the extent a sample parameter, or constituent, exhibits autocorrelation. The range can represent a distance value when modelling spatial data or a temporal frequency when modeling temporal data. The range value quantifies the distance or frequency over which a sampled property, or constituent, is considered autocorrelated. The range of autocorrelation can be useful for making informed data-driven decisions including how to best transform a spatial or temporal data set, and optimize sampling frequencies within the groundwater monitoring system(s) to ensure sample independence (Section 2.5). Therefore, optimizing sampling frequencies will minimize sampling redundancies (e.g. autocorrelation) and cost without jeopardizing sampling adequacy (40 CFR §257.94(d)(2)).

The variogram requires that the data meet the assumption of intrinsic stationary, which satisfies the following criteria: the data are stationary (no systematic change in the mean) and the variance depends only on sample separation increment, or separation distance between samples in space or time. Ideally, shorter separation increments will have higher autocorrelation whereas larger separation increments will have lower autocorrelation, which follows the principle of Tobler's First Law of Geography. The variogram also requires a sufficient number of sample data to adequately characterize autocorrelation.

Recommendations for reducing data dependence will include decreasing the sampling frequency for future samples and detrending or domaining the data prior to performing statistical comparison tests (Section 2.5).

2.4.3 Time Series Analysis

Application: Time series analysis is used to:

- screen for potential anomalous data that warrant further investigation;
- screen for temporal trends in constituent concentrations in each monitoring well; and
- test for significance of temporal trends, where identified.

EDA scope: Phase I and Phase II EDA

Selected methods: Selected methods include time series plots and parametric and non-parametric trend analysis.

A time series is a sample data set ordered consecutively by sample date. Plotting constituent concentrations as a time series provides a very quick visual approach to screen monitoring well data for potential outliers and/or temporal trends. In this case, outliers will consist of visually identifying constituent concentrations that do not conform to the historical temporal variations characteristic to a given well, such as extremely high or low concentration values.

Long-term temporal trends exist when a constituent time series shows a discernable pattern of increase or decrease in constituent concentrations over time, thereby indicating that the sample mean is non-stationary over time. The significance and slope of these trends will be evaluated using the Mann-Kendall and the Theil-Sen tests to determine if the increase or decrease in constituent concentrations are significant ($p < 0.05$). The Mann-Kendall and the Theil-Sen tests make no assumptions regarding the data distribution. The Mann-Kendall test does not indicate the slope of the trend. The Theil-Sen test can be used in conjunction with the Mann-Kendall test to assess the magnitude of the slope of the trend.

Temporal trends in groundwater samples can be indicative of natural fluctuations in groundwater conditions and/or impact from anthropogenic activity independent of the CCR Unit operation. All temporal trends will be interpreted through the conceptual site model to explore their origin. If trends are statistically and hydrogeologically justified, the data should be detrended (Section 2.5.1), or otherwise accounted for, when implementing a statistical method pursuant to 40 CFR §257.93. Historical data should be reviewed to determine if they are representative of current site-specific groundwater conditions. The presence of inconsistent trends among wells within the groundwater monitoring network might also suggest spatial heterogeneity in groundwater conditions; in such case(s) the adequacy of an interwell statistical comparison might need reconsideration (Section 2.6).

2.5 Statistical Independence and Data Domaining

The statistical methods in 40 CFR §257.93 assume sampled data are stationary (statistical properties are constant in space and time), independent (exhibit no spatial or temporal relationships between individual samples) and consist of a single sample population.

For the purpose of this SDAWP, a data set that exhibits a statistical mean that changes systematically in space or time is considered non-stationary (meaning the data exhibit a trend). This change can take the form of a linear or non-linear increase or decrease in a constituent concentration in space and/or over time. El Kadi (1995) provides a good overview of stationarity and non-stationarity in the context of groundwater statistics.

The presence of a trend will automatically infer two things:

1. The sample data are statistically dependent (Section 2.4) because the trend itself demonstrates that samples exhibit a distinct relationship in space or time; and
2. The sample data set possibly exhibits more than one statistical population.

In such cases, data detrending (Section 2.5.1) and/or data domaining (Section 2.5.2) methods will be considered.

Quick interpolation (Section 2.4.1), autocorrelation (Section 2.4.2), and/or time series analysis (Section 2.4.3) can assess data dependence and methods in Section 2.5.2 can assess the appropriateness for domaining data.

2.5.1 Data Detrending

Application: Data detrending is used to:

- transform a statistically dependent sample data set into statistically independent sample data set.

EDA scope: Phase II

Selected methods: regression modeling and adjusting the sampling frequency

Data will be considered statistically dependent if:

- there are statistically significant ($p < 0.05$) trends in constituent concentrations sampled over time in individual wells; and/or
- the variogram model exhibits a nugget:sill ratio less than 0.5.

If the data are considered statistically dependent, regression modeling is one option for generating a statistically independent data set. Regression modeling applicability is dependent on data evaluation objectives, data adequacy, and working knowledge of the hydrogeological environment the sample data represent (e.g. the CSM).

Regression modeling can include linear, non-linear (e.g. seasonal), and spatial or temporal regression methods. In general, regression modeling requires identifying the type of trend present, fitting an adequate model to the trend, then performing statistical evaluations on the modeled trend residuals. The trend residuals will be tested for independence using correlation analysis. Goodness of fit criteria will be used to determine if the regression model adequately describes the trend.

If regression methods prove inadequate, alternative methods will be considered, such as data domaining (Section 2.5.2).

If the CSM suggests that temporal trends are intrinsic to the groundwater system and not attributed to a release from the CCR Unit, then it is arguable to decrease the sampling frequency to ensure the sample data are independent. If enough data are available, the variogram model can provide a data-driven minimum time lag necessary to sample independent data. The variogram model should be interpreted within context of groundwater velocity for reasonableness. The groundwater velocity can provide an estimate of the residence time of groundwater at a given location and the variogram time lag should be larger than this residence time. If too few sample data are available to generate a variogram model, the groundwater residence time can help infer an adequate sampling frequency.

2.5.2 Heterogeneity and Data Domaining

Application: Data domaining is used to:

- decompose a multi-population sample data set into respective single population sample data sets; and/or
- transform a non-stationary data set into a stationary data set.

EDA scope: Phase II

Selected methods: box and whisker plots, Levene's test, ANOVA (and non-parametric equivalents), cluster analysis and principal component analysis.

Spatial and temporal heterogeneity are common within groundwater systems and indicates that the groundwater monitoring network is sampling more than one statistical population. Spatial and temporal heterogeneity means there are measurable differences in statistical characteristics among one sample population and the next, whether these populations derive from different locations within the monitoring well network or from different sample periods over time. These differences can be discrete or continuous, where the latter takes form through gradual trends (i.e. the data are non-stationary). It is important to recognize and test these differences to ensure the sample data are grouped properly to perform statistical comparisons in Sections 3.0 and 4.0.

When prepared by region or individual well, box and whisker plots (Section 2.7) can provide a quick visual assessment of spatial heterogeneity within the groundwater monitoring network. Time series plots (Section 2.4.3) can provide a quick visual assessment of temporal heterogeneity in groundwater quality data.

More advanced statistical comparisons are necessary to make defensible conclusions with regard to the presence or absence of spatial and temporal heterogeneity. Several statistical methods are available to test for statistically significant differences between sample populations in space (e.g. sampling different monitoring well locations) and time (e.g. sampling different periods over time). The Levene's test can determine the equality of variance between two sample data sets ($\alpha = 0.01$). If the equality of variance test holds, then a one-way ANOVA test can subsequently determine if there are statistically significant ($\alpha = 0.05$) differences in constituent concentrations between two sample data sets. The Kruskal-Wallis test is the non-parametric alternative to the parametric ANOVA test. More than one statistical comparison test may be considered. Spatial and temporal heterogeneity is constituent dependent, meaning the results of these tests can vary from one constituent to another.

If these statistical evaluations suggest that the monitoring well network is sampling equivalent statistical populations then it is possible to pool the sample data. This is recommended if more than one background well is present for a CCR Unit. Pooling sample data is advantageous because it increases the sample number, which in turn, increases the statistical power of the statistical test.

If the above statistical evaluations suggest more than one sample population is present then the data must be domained, or decomposed, into individual sample populations. More advanced statistical methods, including cluster analysis and/or principal component analysis, can provide data-driven groupings based on information redundancies, or underlying correlations observed in the sampled data. The notion behind cluster analysis and principal component analysis is to group data according to within-group similarities and between-group dissimilarities, where each group represents a unique statistical population. The sample data are then segregated and pooled into individual homoscedastic statistical populations. Statistical analyses are then performed using data groupings, or pooled data. In theory, underlying correlations derive from the environmental properties and processes from which the sample data originate, making these methods ideal for identifying different groundwater types, lithologies and or site geochemistry, for example. The CSM will help interpret data-driven groupings.

If more than one sample population is present in the groundwater monitoring network, it is necessary to determine if there is a representative background population to perform interwell statistical comparisons (Section 2.6) with the corresponding downgradient sample population. If a representative background population is not present then intrawell statistical comparisons (Section 2.6) might be appropriate in downgradient wells.

As alluded to in Section 2.4.3, a non-stationary timeseries data set might contain more than one statistical population. This is oftentimes the case when historical sample data are grouped with more recent sample data. In these cases, the historical sample data might not be representative of current site-specific groundwater conditions. Therefore, excluding the historical temporal samples might produce a stationary sample population representative of current groundwater conditions. The statistical comparison tests in this section can help determine if there are statistically significant differences between historical and recent sample data.

2.6 Interwell versus Intrawell Comparisons

The FCPP groundwater monitoring systems are designed to perform interwell statistical comparisons, with the exception of intrawell statistical comparisons for select monitoring wells and constituents at the CWTP (Wood, 2019). Interwell comparisons are oftentimes referred to as "upgradient-to-downgradient comparisons" (USEPA, 2009) because they compare measurements sampled in background monitoring wells to measurements sampled in monitoring wells that reflect groundwater conditions downgradient of the CCR unit. Interwell comparisons perform poorly in cases where a constituent exhibits spatial and/or temporal heterogeneity such that the statistical mean and variance are not considered representative or constant across the groundwater monitoring network. In such cases, intrawell comparisons are an industry accepted and recommended alternative to interwell comparisons (USEPA, 2009).

An intrawell comparison compares constituent concentrations over time within a single well. For a given monitoring well, sample data that reflect baseline constituent concentrations are compared to sample data (sampled from the same well) to determine if there is a statistically significant increase in constituent concentrations relative to baseline concentrations. In this context, intrawell comparisons help determine if groundwater quality is deteriorating or holding constant at a given location over time. Intrawell comparisons are less useful when baseline constituent concentrations are constructed from only a few sample points and/or groundwater samples are collected post CCR installation, which means that the data are potentially impacted by CCR activity. When faced with these disadvantages, groundwater deterioration is evaluated by testing for statically significant positive trends in constituent concentrations sampled within the well over time.

Intrawell and interwell selections are data-driven and will be constituent dependent, meaning an intrawell comparison might be appropriate for one constituent but interwell comparisons are appropriate for remaining constituents, for example.

2.7 Data Distribution Assessment

Pursuant to 40 CFR §257.93(g)(1), the statistical method used to evaluate groundwater data will be appropriate for the distribution of the constituent (e.g. sample population). Two hierarchies of statistical methods are present in 40 CFR §257.93, including parametric or nonparametric statistical methods. Parametric methods make specific assumptions regarding data distributions. If the sampled data do not fit a theoretical distribution (e.g., normal, lognormal, gamma, etc.) then nonparametric tests are appropriate. Nonparametric tests make no assumptions about the distribution of the sample data and, as such, are oftentimes referred to as distribution-free tests. In general, parametric tests are more powerful than nonparametric tests and will therefore be emphasized for establishing background constituent concentrations and performing statistical comparisons.

Application: Data distribution assessment is used to:

- visualize potential outliers;
- determine if the data follow an identifiable data distribution; and/or
- screen for potential heteroscedasticity.

EDA scope: Phase I and Phase II EDA

Selected methods: Q-Q plots, box and whisker plots, summary statistics, histograms, in addition to the Shapiro-Wilk, Lilliefors, and gamma distribution tests.

The following EDA methods are considered "qualitative" and interpreted through visual assessment of graphic outputs:

- The Q-Q plot compares the sampled data set distribution against a defined distribution. For normal Q-Q plots the theoretical normal distribution is linear in the Q-Q plot. If the sampled data distribution is normal then it will conform to a linear shape comparable to that of the theoretical distribution. The linear correlation coefficient represents the degree of linear correlation between the two distributions. Non-normal or bimodal distributions are apparent when inflection points are observed in the sampled data distribution. Inflections can be indicative of outliers (Section 2.8) or bimodal distributions (more than one sample population present in the data set). In some cases, the correlation coefficient may still be robust even though inflections are present. For this reason, more than one line of statistical evidence is necessary to determine if the sample data set exhibit normality and it is suggested to use at least one formal statistical test described below. Other distributions (lognormal, gamma) can be tested by constructing Q-Q plots based on the appropriate theoretical distribution and interpreted in the same way as above.
- Box plots are a quick tool to screen the location, spread, and shape of the data and underlying sample distribution. A box plot illustrates the 25^{th}, 50^{th} and 75^{th} percentiles of the data in addition to potential outliers (Section 2.8). It is particularly useful to group data to plot multiple box plots to screen for potential heteroscedasticity.
- Histograms also provide a graphical summary of the distribution of a sample data set. The histogram shows equally sized data classes (or bins) on the x-axis and the number of samples (also known as counts) falling within each bin on the y-axis. The histogram is useful for visualizing the center, spread, skewness, and modality of the data. The histogram is also useful for screening outliers (Section 2.8) in the sampled data.

Summary statistics, goodness of fit tests, including the Shapiro-Wilk, Lilliefors, and gamma distribution tests, are numeric statistical tests that evaluate if the sample data distribution fits a pre-defined theoretical data distribution (e.g. normal, lognormal, or gamma). These tests will be performed at a 0.05 level of significance.

- Summary statistics will include calculating the statistical measures (e.g. mean, median, variance, skewness, and kurtosis), minimum and maximum values, and coefficient of variation. If the data exhibit a similar mean and median and little to no skewness then the data likely fit a normal distribution.
- The Shapiro-Wilk test will evaluate if the sampled data fit a normal or lognormal data distribution (ProUCL, 2013). This test is useful for data sets with less than or equal to 50 sample observations. The Shapiro-Wilk test can be applied to raw data to determine if data transformations might be necessary. In such cases, the Shapiro-Wilk test should subsequently be applied to transformed sampled data to test the effectiveness of the data transformation.

- The Lilliefors test is appropriate for larger data sets consisting of 50 or more samples and assesses if the data fit a normal or lognormal data distribution.
- The gamma distribution tests constitute the K-D and A-D tests (ProUCL, 2013). Most positively skewed data follow a lognormal as well as a gamma distribution (ProUCL, 2013). In these cases, the use of a gamma distribution tends to yield more reliable and stable results and will therefore hold preference (ProUCL, 2013).

It is advisable that more than one line of statistical evidence, both graphic and numeric, be provided to defensibly discern the distribution of a sampled data set.

2.8 Outlier Tests

Application: Test for statistically significant ($p < 0.05$) outliers.

EDA scope: Phase II EDA

Selected methods: Dixon or Rosner tests.

Outliers will be tested for significance ($p < 0.05$) using the Dixon's and Rosner's tests. These outlier tests assume the data are normally distributed in the absence of the potential outliers. Therefore, these tests will be performed on transformed data if the data do not exhibit a normal distribution in the presence of the potential outlier(s). More than one line of statistical evidence, such as Q-Q plots and histograms, will be necessary to confirm if a potential outlier should be discarded. The CSM will be incorporated into this decision making to provide reasoning for the abnormal value and to justify its exclusion from the statistical analysis, if appropriate.

3.0 DETECTION MONITORING

The CCR Rule states that by October 17, 2017, a minimum of eight independent samples must be collected to initiate detection groundwater monitoring as required by §257.94(b). This section discusses ensuing statistical tests to assess if there is an SSI over background levels.

The CSM will be iteratively updated with results from the data evaluations detailed in this SDAWP.

3.1 Data Evaluation Objectives

The objective of a detection monitoring program is to evaluate each of the CCR Rule Appendix III constituents (Section 3.2) and to determine, pursuant to 40 CFR §257.93(h), if there is an SSI in constituent concentrations downgradient of the unit compared to background levels.

3.2 Constituents of Concern

The CCR Rule Appendix III constituents are listed in Section 2.2. Statistical evaluations will be performed independently for each constituent.

3.3 Background Comparison Tests

The selection of a defensible statistical test is a data-driven process. Therefore, the selection and declaration of any statistical method(s) herein, pursuant to §257.93, is subject to change as: 1) data characteristics become known through the EDA process and 2) future data present modification(s) to precursor EDA results and/or other known information contained within the CSM.

Background wells will be used to evaluate the quality of water not affected by a CCR unit. Background wells are adequate if they sample groundwater conditions representative of those beneath and downgradient of the CCR unit (assuming it is not leaking). Section 2.6 discusses alternative industry standard procedures for performing groundwater statistical evaluations when background proves inadequate.

3.3.1 Prediction Limits

Within the scope of this SDAWP, upper prediction limits (UPLs) will be calculated to establish background concentration threshold values except for pH, which will also require the calculation of a lower prediction limit (LPL). The UPL belongs to a statistical class of methods called statistical intervals (USEPA, 2009). Intervals are a statistical measure that represent a finite probable range (upper and lower limit) in which a future sample statistic or population parameter is expected to occur (USEPA, 2009). A future sample statistic can constitute a single sample value or a statistical parameter (e.g. mean). For most constituents, the upper interval limit is of interest. A level of confidence is declared based on an error rate (α), which represents the likelihood that the interval does not contain the future sample statistic or population parameter (USEPA, 2009). Measurements falling outside of the interval limit are considered to be significantly different than background at a prescribed level of statistical confidence.

The prediction limit method assumes the background and downgradient sample populations are identical, meaning there is a high probability (1-α) that the prediction limit will contain the future sample value(s) or statistical parameter(s) if the CCR unit is not impacting groundwater. The project CSM (Section 1.0) and EDA (Section 2.0) will provide lines of evidence and guidance as to whether or not designated background and downgradient compliance wells are sampling the same statistical population. Future samples or statistical parameters are collected from downgradient monitoring wells and compared to the constituent UPL established using samples collected from background monitoring wells.

The probability of a future sample to exceed a prediction limit is based on background concentration values but also the design of the monitoring well network (Section 1.0), number of future samples that will be compared to the background prediction limit, and how these comparisons are performed. The Unified Guidance recommends a retesting strategy when using the UPL method to maintain a low false positive occurrence (falsely identifying an SSI) while providing acceptable statistical power (Section 3.4) (USEPA, 2009).

Resampling strategies are in place to ensure an SSI is not falsely declared on account of cumulative random statistical error in future samples. Resampling strategies are applicable for parametric, non-parametric, intrawell, and interwell UPL comparisons. Resampling strategies typically follow a "1 of m" sampling design, where m is the number of resamples necessary to verify a potential SSI. Resampling strategies depend on several criteria, such as the size of the background data set, sampling frequency, interwell versus intrawell comparisons, and the number of active monitoring wells, among other considerations. Only when the analytical data indicate a sample is in exceedance of its UPL is resampling initiated. For a 1 of 2 resampling strategy, as an example, the initial exceedance and a second statistically independent sample must be in exceedance of the UPL to declare an SSI. If the second sample is not in exceedance, then there is insufficient evidence to declare an SSI at that time and the 1 of 2 count is reset to 0. If there is no exceedance in the analytical results then resampling is not necessary.

Resampling strategies are established prior to performing statistical compliance testing and will reflect in the parametric calculation of the UPL through an κ-multiplier (USEPA, 2009). The most appropriate resampling strategy will be selected in consideration of the expected statistical power and sitewide false positive rate (SWFPR) (Section 3.4). The overall defensibility of a resampling strategy decreases when the sample data are statistically dependent (i.e., sampled so close in time that they are correlated), which is usually the case when sampling at a frequency higher than quarterly. The resampling strategy is generally unnecessary when the observed concentrations in downgradient wells are distinctly higher than concentrations observed in background wells (e.g., all samples are order(s) of magnitude higher); in this case, background might be inadequate or a release from the evaluated unit has occurred.

For parametric background data sets exhibiting a linear temporal trend, it is possible to calculate the parametric UPL and LPL around the trend (Section 2.5.1). This SDAWP adapts Equation 10-13 in the ProUCL Technical Guide to calculate the UPL and LPL around a trend (USEPA, 2013). The data must meet the statistical assumptions for the ordinary least squares regression method and exhibit a non-detect frequency of less than 15%.

Non-parametric UPLs will be appropriate for constituents with at least a 50%, but less than 100%, non-detect frequency and/or the data do not exhibit an identifiable distribution. For non-parametric UPLs the upper limit generally reflects the highest or second-highest constituent concentration. It is beneficial that the background data set has a sufficient number of samples to achieve an acceptable SWFPR. A minimum of 8 samples is likely insufficient and, therefore, pooling background might be appropriate and should be explored using the Kruskal-Wallis method (Section 2.5.2). Choosing a 1 of m retesting strategy follows the same logic as presented above for the parametric UPL.

The DQR will be appropriate for constituents exhibiting 100% non-detection frequencies (Section 2.3.3).

3.3.2 Alternatives to Prediction Limits

The declaration of the UPL is pending review of available data. If available data do not lend itself to using the UPL, an appropriate statistical test from the remaining tests listed in 257.93(f) will be chosen.

3.4 Performance Standards

There are performance standards to help ensure that the statistical tests perform adequately to identify the occurrence of a legitimate CCR unit leakage. These performance standards can provide measures of sampling adequacy but also sensitivity of the statistical tests to detect changes in groundwater quality. Within the scope of this SDAWP, these standards consider statistical power, site-wide false positive errors, and retesting strategies.

- **Statistical Power**: The statistical power is the ability for a statistical comparison test to identify a legitimate leakage from a CCR unit. The statistical power will improve as the sample number increases and varies based on the type of statistical method used. For each statistical method, multiple types of statistical comparisons are possible as part of the SDAWP (e.g. parametric, non-parametric, intrawell, or interwell). Therefore, statistical power will reflect the type of statistical tests and will follow method-specific recommendations put forth in the USEPA Unified Guidance (2009).
- **Site-Wide False Positive Rate**: The SWFPR should be considered in balance with statistical power. The SWFPR reflects the risk that a test will falsely indicate there is leakage from a CCR unit (USEPA, 2009). This risk is encountered in each comparison test that is performed as part of the detection monitoring statistical program. Because the number of comparison tests may be large over the

lifespan of a detection monitoring program (e.g., due to repeated sampling) the likelihood of at least one statistical test indicating a false positive is realistic. This is known as the multiple comparisons problem (USEPA, 2009). The multiple comparison problem can be addressed using retesting.

- **Retesting**: Retesting is proposed to achieve a realistic balance between a low SWFPR and maintaining adequate statistical power to detect leakage from a CCR unit. Resampling is a check for transient, marginal increases over a background threshold level that are not really significant but are to be expected as a result of multiple comparisons. In general, retesting overcomes the multiple comparison problem by constructing a set of decision rules that are applied to UPL strategies. The Unified Guidance provides several approaches for establishing decision rules (USEPA, 2009). Retesting schemes for medians and means provide more robust statistical properties (e.g., power and SWFPR) in comparison to other retesting methods and are ideal for detection monitoring programs where multiple sample rounds are anticipated. The chosen approach will affect the choice of κ-multiplier; therefore, the retesting approach needs to be selected prior to calculating the UPL.

3.5 SSI Declaration – Detection Monitoring

If the detection monitoring statistical evaluation indicates there is an SSI over background for one or more constituents, an investigation should ensue to determine if a release from the CCR unit is the cause of the SSI(s). If the data and information within the CSM demonstrate: 1) a release from an alternative source; 2) natural spatial or temporal heterogeneity, and/or 3) sampling or analytical error is the source to the declared SSI, then this demonstration must be made in writing and certified by a qualified professional engineer within 90 days of completing the statistical evaluation. Alternative source demonstrations will rely on available data in addition to information contained within the CSM.

If this demonstration cannot be made within 90 days of the SSI declaration, then the site moves into assessment monitoring (Section 4.0).

4.0 ASSESSMENT MONITORING

The CCR Rule states that a CCR unit must begin assessment monitoring 90 days following a declaration of an SSI over background if during this time there is no supporting evidence presented demonstrating that the SSI results from an alternative source.

Within 90 days of the SSI declaration, and **annually** thereafter, the owner or operator must sample the unit's groundwater monitoring network for Appendix IV constituents, pursuant to §257.95(b). Within 90 days of obtaining sample results, and **semiannually** thereafter, the owner or operator must sample the unit's groundwater monitoring network for all Appendix III constituents and the Appendix IV constituents whose concentrations were detectable during the initial assessment monitoring sample event.

Pursuant to §257.95(h), groundwater protection standards (GWPS) must be established for detectable Appendix IV constituents. For each constituent, the selected GWPS is the higher of the site-specific background level (i.e., BTV), the USEPA's promulgated Drinking Water Maximum Contaminant Level (MCL), or a risk-based alternative GWPS identified in the CCR Rule for constituents without an MCL (i.e., cobalt, lead, lithium, and molybdenum).

This section discusses statistical tests that will be used to assess if one or more Appendix IV constituents are detected at statistically significant levels (SSLs) above GWPSs.

4.1 Data Evaluation Objectives

The objective of an assessment monitoring program is to evaluate each of the CCR Rule Appendix IV constituents and to determine, pursuant to 40 CFR §257.95(g), if a constituent is detected at a SSL above the GWPS in groundwater downgradient of the evaluated unit.

4.2 Constituents of Concern

The CCR Rule Appendix IV constituents are listed in Section 2.2. Statistical evaluations will be performed independently for each constituent.

4.3 GWPS Comparison Tests

The selection of a defensible statistical test is a data-driven process. Therefore, the selection and declaration of any statistical method(s) herein is subject to change as: 1) data characteristics become known through the EDA process and 2) new data present modification(s) to precursor EDA results and/or other known information contained within the CSM.

There are two approaches to performing assessment monitoring. The first is a single-sample statistical comparison, where downgradient sample data are compared to a predefined and fixed value (e.g., the MCL). The second is a two-sample statistical comparison test, where statistical properties of the downgradient sample population are compared to statistical properties of the site-specific background sample population. The two-sample statistical comparison test is only applicable when background is higher than the MCL or alternative risk-based GWPS.

The statistical methods listed in 40 CFR §257.93(f) are adequate for two-sample statistical comparisons, however, are inadequate to compare downgradient sample data to a fixed threshold value (single-sample comparisons). The following sections recommend statistical methods that are appropriate for both single-sample and two-sample statistical comparisons and are hereby incorporated based on 40 CFR §257.93(f)(5) which states that "Another statistical test method that meets the performance standards of paragraph (g) of this section" may be used.

4.3.1 Single-Sample Comparison Tests

The single-sample approach compares the downgradient well constituent concentrations to a fixed value. In this case, the fixed value will be the MCL, an alternative risk-based GWPS or, if higher, a site-specific background level. The statistical hypothesis structure for a single-sample comparison is reversible, unlike a two-sample statistical comparison, such that the same fixed background level can be used for assessment monitoring and later for corrective action testing, if necessary. When the MCL or alternative risk-based GWPS serves as the constituent GWPS, the Unified Guidance (USEPA, 2009) recommends calculating confidence intervals around the downgradient data set's mean or median (pending definition of the data distribution) and comparing the lower confidence limit of this interval to the constituent GWPS; if the lower confidence limit exceeds the constituent GWPS there is enough statistical evidence to declare an SSI. The confidence interval calculation requires a t value at a specified confidence (e.g., $1-\alpha$, where $\alpha = 0.05$); this value should be chosen to achieve a low false positive rate while achieving adequate statistical power. The confidence interval calculation can account for a temporal trend, similar to the UPL calculation with a trend.

The confidence interval will be calculated in accordance with EDA procedures (Section 2.0) and recommendations put forth in the USEPA's Unified Guidance.

The Unified Guidance recommends calculating the upper tolerance limit (UTL) to represent the constituent GWPS when the site-specific background level is higher than the MCL or alternative risk-based GWPS. To determine if the UTL is higher than the MCL or alternative risk-based GWPS, the former will be calculated for each constituent. The UTL is designed to be "a reasonable maximum on the likely range of background concentrations" (USEPA, 2009) and, similar to the MCL or alternative risk-based GWPS, the UTL can accommodate a statistical hypothesis structure that is reversible (i.e., is appropriate for both compliance and corrective action testing, if necessary). In general, the UTL represents a sample concentration range, or coverage, that contains a pre-defined proportion of the underlying statistical population. Most often this pre-defined coverage is equal to 95% (e.g., the 95% upper tolerance limit). The tolerance limit calculation is very similar to the prediction limit calculation but the tolerance limit multiplier ($\boldsymbol{\tau}$) is based on the selected coverage (recommended coverage γ = 95%) and selected confidence (recommended confidence is 95%, or α = 0.05). Since the UTL is treated interchangeably to the MCL or alternative risk-based GWPS in this case, the statistical comparison is performed similarly using the lower confidence limits of the downgradient sample data; if the lower confidence limit of the downgradient sample data exceeds the site-specific UTL there is enough statistical evidence to declare an SSI and possibly justify corrective action. The tolerance limits will be calculated in accordance with the EDA procedures (Section 2.0) and recommendations put forth in the USEPA's Unified Guidance.

4.3.2 Two-Sample Comparison Tests

The two-sample statistical comparison uses site-specific background levels to establish a constituent's GWPS, which may be higher than either the MCL or alternative risk-based GWPS. For this approach, the prediction limit method remains adequate as it did for detection monitoring. This is, in part, because the UPL follows a single statistical hypothesis structure common to detection monitoring (USEPA, 2009). In this specific case, however, the Unified Guidance recommends constructing the upper prediction limit of a mean or median (pending definition of the data distribution) then comparing the mean or median of the downgradient data set to this upper limit. The UPL calculation can account for a temporal trend, if necessary (Section 2.5.1). If the mean or median is in exceedance then there is enough statistical evidence to declare an SSI and possibly justify corrective action. The prediction limits around the mean or median will include a retesting strategy (Section 3.3.1 and Section 3.4) and be calculated in accordance with the EDA procedures (Section 2.0) and recommendations put forth in the USEPA's Unified Guidance.

4.4 Performance Standards

The performance standards in Section 3.4 are applicable to assessment monitoring and will follow method-specific recommendations put forth in the USEPA's Unified Guidance.

4.5 SSL Declaration – Assessment Monitoring

If assessment monitoring demonstrates that all Appendix III and Appendix IV constituents are equal to or below their respective background levels for two consecutive sampling events, then the monitoring program can return to detection monitoring.

Pursuant to 40 CFR §257.95(f), assessment monitoring will continue if there are Appendix III or Appendix IV constituent concentrations above background levels but there is insufficient evidence to declare a constituent is present at an SSL above the GWPS established under 40 CFR §257.95(h).

If assessment monitoring demonstrates an exceedance at an SSL above an Appendix IV constituent's GWPS then the owner or operator must follow criteria set forth under 40 CFR §257.95(g).

5.0 STATISTICAL CONSIDERATIONS FOR UNSATURATED ZONES

For CCR units where the uppermost hydrostratigraphic unit is dry, such as the RWP and DFADA, the groundwater monitoring system is designed to detect a release from the CCR unit, however, the monitoring well screened intervals reside within an unsaturated zone. In this situation, the groundwater monitoring program expects the monitoring wells to either be dry or to contain a minimal amount of water (insufficient for sampling). CCR units that produce inadequate or insufficient groundwater analytical data to perform the statistical evaluations put forth in this SDAWP lend themselves to alternative evaluations to assess if the CCR unit is leaking.

Criterion for indicating a potential leak at the RWP or the DFADA include the measurement of saturated thickness in a given downgradient well that is distinguishable from condensate buildup in the monitoring well(s). If this criterion is met, then the following steps will ensue to help determine if the CCR unit is leaking:

1) Monitor the saturated thickness within the downgradient well on a monthly basis to evaluate if the saturated thickness stays the same or increases over time.
2) If sufficient groundwater is present in a given well:
    a. Collect a groundwater sample and compare the Appendix III analytical concentrations to their respective and established BTVs.
    b. If there is a sample exceedance, execute the resampling strategy (Section 3.3.1) set forth for each respective Appendix III constituent using a quarterly sampling frequency.
3) Reference the CSM to:
    a. Investigate if local climatic regimes could have caused the previously-dry hydrostratigraphic unit beneath the CCR unit to become saturated;
    b. Assess the observed groundwater elevations beneath the CCR unit to evaluate if groundwater mounding indicative of a release is occurring; and
    c. Review the Site's operation and maintenance history to identify reasonable cause for a leak or alternative source.

If the resampling strategy in step 2(b) above confirms there is an SSI over background, then the monitoring program will proceed according to the steps described in Section 3.5. If there is insufficient evidence to declare an SSI over background but there are lines of evidence from the above preliminary evaluation(s) to indicate that the CCR unit might be leaking, then professional judgement will determine if the monitoring program will continue in detection monitoring or proceed to either an ASD or assessment monitoring program, per 40 CFR §257.94(e).

6.0 RECOMMENDATIONS FOR FUTURE EVALUATIONS

The CCR Rule does not declare criteria for updating background values over time. The minimum sampling criteria put forth by the CCR Rule likely does not capture the range of intrinsic temporal variation in constituent concentrations typical for a dynamic groundwater system. Moreover, larger sample numbers will increase the statistical power of the subsequent statistical tests.

Consequently, background limits should be updated periodically (e.g., every two years) until the sample data set is representative of intrinsic temporal variations in groundwater conditions and the sample number produces an adequate statistical power.

To update background, it is appropriate to compare the new background data to existing background data to ensure the two data sets reflect the same sample population; statistical methods listed in Section 2.5.2 are sufficient for performing this assessment. If the statistical comparison tests do not indicate significant differences then the background data can be pooled. If statistically significant differences are present between two sample sets, the data should be reviewed to determine the source of the difference and the sample set that is most representative of current groundwater conditions should hold precedence.

The exception is for background limits that are calculated around a trend line. In these cases, background limits will need updating after every sampling event.

This version of the SDAWP does not include statistical procedures for completing corrective action monitoring. Updates to this SDAWP will be made if a unit transitions into corrective measures and when a remedy is selected.

7.0 SOFTWARE

EDA and detection monitoring statistical evaluations will be performed using ProUCL. ProUCL is a public domain software platform supported by USEPA.

Visual Sampling Plan (VSP) is public domain software supported by the U.S. Department of Energy and Pacific Northwest National Laboratory. This software is useful for assessing data dependence (Section 2.5) and performing sampling optimization.

Other public domain software packages, including R and Spatial Analysis in Macroecology, are defensible and transparent spatial regression and data detrending (Section 2.6.1) software platforms. These software platforms will supplement ProUCL and VSP, as necessary.

Isatis (Geovariances, France) is a well-established geostatistical software platform. This software will be used to validate variogram models and spatial interpolation methods (Section 2.5), as necessary.

8.0 CERTIFICATION

By means of this certification, I certify that I have reviewed this SDAWP and that the statistical methods described herein are appropriate and meet the requirements of 40 CFR §257.93.

Emily H. LoDolce
Printed Name of Registered Professional Engineer

Signature

25912	New Mexico	10 June 2020
Registration No.	Registration State	Date

9.0 REFERENCES

AECOM, 2015. *Sampling and Analysis Plan*, Coal Combustion Residual (CCR) Groundwater Monitoring, Four Corners Power Plant, Arizona Public Service, Farmington, New Mexico. AECOM Job No. 60437300. December 2015.

AECOM, 2017. *CCR Monitoring Well Network Report and Certification*, Four Corners Power Plant, Fruitland, New Mexico. AECOM Job No. 60531071. September 2017.

AECOM, 2018. Boring Log for Well Number MW-12R1. Well completed in April 2018.

Amec Foster Wheeler Environment & Infrastructure, Inc. (Amec Foster Wheeler), 2018. *Annual Groundwater Monitoring and Corrective Action Report for 2017*. Coal Combustion Residuals Rule Groundwater Monitoring System Compliance, Four Corners Power Plant, Fruitland, New Mexico. Project No. 14-2016-2024. January 2018.

Arizona Public Service (APS), 2013. *Four Corners Power Plant Groundwater Quality Data Submittal*.

Dames & Moore, 1988. *Final Report on Hydrogeology (Volume I) for Arizona Public Service Four Corners Generating Station*. D&M Job No. 02353-083-33. March 1988.

El Kadi, A., 1995. Groundwater Models for Resources Analysis and Management. London, CRC Press, Inc.

Federal Register, 2018. *40 Code of Federal Regulations Part 257 – Hazardous and Solid Waste Management System; Disposal of Coal Combustion Residuals From Electric Utilities; Final Rule promulgated April 17, 2015 with Amendments to the National Minimum Criteria (Phase One, Part One) effective August 29, 2018*.

ProUCL, 2013. *ProUCL (Version 5.0.00) User Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations*. EPA/600/R-07/041. Washington D.C. September.

Sakura Engineering & Surveying, 2018. Survey of MW-12R1. Sealed on May 18, 2018.

United States Environmental Protection Agency (USEPA), 2009. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities Unified Guidance*. EPA 530/R-09-007. Environmental Protection Agency Office of Resource Conservation and Recovery.

USEPA, 2011. *Environmental Cleanup Best Management Practices: Effective Use of the Project Life Cycle Conceptual Site Model*. EPA 542-F-11-011. Office of Superfund Remediation and Technology Innovation.

USEPA, 2013. ProUCL (Version 5.0.00) User Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations. EPA/600/R-07/041. Washington D.C. September.

Wood Environment & Infrastructure Solutions Inc. (Wood), 2020. *Groundwater Monitoring Network Certification Report for the Return Water Pond*. Coal Combustion Residuals Rule Groundwater Monitoring System Compliance, Four Corners Power Plant, Fruitland, New Mexico. Report dated June 5, 2020.

TABLES

Table 1
Description of Coal Combustion Residuals Units

CCR Unit	Location	Function	Operation	Size/Construction	History
Upper Retention Sump (URS)	*Plant Area* NW1/4 of Section 36, T29N, R16W	*Single CCR unit*. Impoundment. Surge pond for FGD system.	FGD system discharge is discharged into the sump via 10 plus controlled/monitored lines. Pond contents are recirculated back into the FGD process via a pump chamber located on the south end of the pond. Solids are periodically removed from the sump.	- 1.07 acres in areal extent - Soil-cement liner on bottom and inside slopes	Placed in service around 1983.
Combined Waste Treatment Pond (CWTP)	*East of Plant, Adjacent to Morgan Lake* SE1/4 of Section 25, T29N, R16W	*Single CCR Unit*. Impoundment. Detention pond used for NPDES treatment; settling and stabilization basin for ash-impacted and other Plant wastewater flows prior to discharge to Morgan Lake in accordance with an NPDES permit.	The primary source of water to the CWTP is from hydrobins which separate transport water from bottom ash generated in plant Units 4 and 5. Seven earthen basins in the western edge of the CWTP promote sediment settling prior to the water decanting into the main portion of the CWTP and then overflowing into the cooling water discharge canal at the northeast corner of the pond.	- 13.4 acres in areal extent	Constructed in 1978.
Lined Ash Impoundment (LAI)	*Disposal Area* E1/2 of Section 34, T29N, R16W	*Part of a CCR multiunit with the LDWP* that receives fly ash, flue gas desulfurization (FGD) waste and associated residuals as a slurry from the plant.	Waste is discharged into the pond in the northeast portion of the pond. Decanted flow discharges via a vertical drop inlet structure and through a toe drain into the LDWP.	- 126.8 acres in areal extent (high water line) - 60 mil HDPE liner - 5,364 acre-ft design capacity - 5,275.2 ft AMSL maximum working level	Constructed on top of closed Ash Ponds 4 and 5 and placed in service in 2004.
Lined Decant Water Pond (LDWP)	*Disposal Area* E1/2 of Section 34, T29N, R16W	*Part of a CCR multiunit with the LAI* that receives decanted water from the LAI. Impoundment.	Decanted water is discharged into the pond from the LAI via gravity; the water is pumped from the LDWP back to the plant for reuse in operations.	- 45 acres in areal extent - Two 60 mil HDPE liners separated by a leak detection layer - 435 acre-ft design capacity - 5,213.2 ft AMSL maximum working level	Constructed on top of closed Ash Pond 3 and placed in service in 2004.
Dry Fly Ash Disposal Area (DFADA)	*Disposal Area* SE1/4 of Section 34, T29N, R16W	*Single CCR unit*. Landfill. Disposal of dry fly ash, bottom ash, and construction debris. In the future, FGD solids will be mixed with fly ash at the plant and landfilled in the DFADA.	The DFADA is filled in general accordance with a stacking plan. Leachate generated from the DFADA cells is pumped into trucks and used for dust control or can be transferred to the LDWP.	- 3 conjoined cells (DFADA 1, 2, and 3) with an areal extent of 94.8 acres total - 3,125 acre-ft design capacity - DFADA 1: compacted clay overlain by 60 mil HDPE liner and drainage layer - DFADA 2 and 3: geosynthetic clay liner overlain by 60 mil HDPE liner and drainage layer - Leachate collection system drains each DFADA cell - DFADA 4 is planned and under construction in 2020	Constructed in 2007 (DFADA 1), 2012 (DFADA 2), and 2014 (DFADA 3).

Table 1
Description of Coal Combustion Residuals Units

CCR Unit	Location	Function	Operation	Size/Construction	History
Return Water Pond (RWP)	*Plant Area* NW1/4 of Section 36, T29N, R16W	*Single CCR unit.* Lined impoundment for the temporary storage of FGD system waste, drain down from the LAI, treated sewage wastewater flow, and water pumped from the site seepage collection system.	The RWP consists of two cells; FGD system waste generated at the plant can be discharged into an FGD cell while all other liquids are discharged into a liquid cell. A spillway between the two cells allows liquid in the FGD system waste to decant into the liquid cell. Liquids from the liquid cell are pumped back to the plant for reuse in plant operations.	- 5.1 acres in areal extent - Composite liner system and associated LCRS comprised of a primary 60 mil HDPE liner, a geosynthetic drainage layer, a secondary 60 mil HDPE liner, and an underlying geosynthetic clay liner - 38.6 acre-ft design capacity - 5379 ft AMSL maximum working level	Constructed in 2019 and placed in service June 2020.

Abbreviations:

AMSL - above mean sea level
CCR - Coal combustion residuals
CWTP - Combined Waste Treatment Pond
DFADA - Dry Fly Ash Disposal Area
FGD - flue gas desulfurization
ft - feet
HDPE - high density polyethylene
LAI - Lined Ash Impoundment
LCRS - leak collection and removal system
LDWP - Lined Decant Water Pond
NPDES - National Pollutant Discharge Elimination System
RWP - Return Water Pond
URS - Upper Retention Sump

Table 2
CCR Groundwater Monitoring System Summary

Well	CCR Unit	Well Designation	Hydrogeologic Unit	Date Installed	Borehole Depth [ft bgs]	Top of Casing Elevation [ft AMSL]	Ground Surface Elevation [ft AMSL]	Top of Screen [ft bgs]	Bottom of Screen [ft bgs]	Screen Length [ft]	Top Screen Elevation [ft AMSL]	Bottom Screen Elevation [ft AMSL]	Bottom Borehole Elevation [ft AMSL]
MW-62	CWTP	Downgradient	Pictured Cliffs Sandstone	9/28/2015	20	5,341.87	5,339.37	10.0	20.0	10	5,329.37	5,319.37	5,319.37
MW-63	CWTP	Downgradient	Pictured Cliffs Sandstone	9/25/2015	20	5,337.02	5,337.02	9.0	19.0	10	5,328.02	5,318.02	5,317.02
MW-64	CWTP	Downgradient	Pictured Cliffs Sandstone	9/26/2015	25	5,337.66	5,337.66	10.0	20.0	10	5,327.66	5,317.66	5,312.66
MW-65	CWTP	Downgradient	Pictured Cliffs Sandstone	9/27/2015	20	5,339.74	5,337.24	8.0	18.0	10	5,329.24	5,319.24	5,317.24
MW-10	DFADA	Downgradient	Lewis Shale	3/12/1987	35	5,150.71	5,149.65	13.0	33.0	20	5,136.65	5,116.65	5,114.65
MW-12R1	DFADA	Background	Lewis Shale	4/10/2018	40	5,270.12	5,268.23	22	32	10	5,246.20	5,236.20	5,228.23
MW-13	DFADA	Downgradient	Lewis Shale	8/31/1987	60	5,150.75	5,149.52	34.9	54.9	20	5,114.62	5,094.62	5,089.52
MW-44	DFADA	Downgradient	Lewis Shale	3/28/2012	40	5,146.89	5,145.15	13.5	23.5	10	5,131.65	5,121.65	5,105.15
MW-48	DFADA	Downgradient	Lewis Shale	5/14/2013	60	5,165.96	5,163.43	35.0	60.0	25	5,128.43	5,103.43	5,103.43
MW-55R	DFADA	Background	Lewis Shale	9/13/2015	95	5,243.96	5,241.36	72.9	92.9	20	5,168.46	5,148.46	5,146.36
MW-07	Multiunit 1	Downgradient	Lewis Shale	3/11/1987[a]	60	5,149.32	5,148.29	14.7	34.7	20	5,133.59	5,113.59	5,088.29
MW-08	Multiunit 1	Downgradient	Lewis Shale	3/11/1987[a]	74	5,122.56	5,120.85	27.7	47.7	20	5,093.15	5,073.15	5,046.85
MW-40R	Multiunit 1	Downgradient	Lewis Shale	9/17/2015	25	5,137.43	5,134.83	14.3	24.3	10	5,120.53	5,110.53	5,109.83
MW-43	Multiunit 1	Background	Lewis Shale	3/24/2012	60	5,271.58	5,269.42	16.0	26.0	10	5,253.42	5,243.42	5,209.42
MW-49A	Multiunit 1	Background	Lewis Shale	5/18/2013	68	5,288.62[b]	5,285.29[b]	50.0	65.0	15	5,231.38	5,216.38	5,213.38
MW-50A	Multiunit 1	Background	Lewis Shale	5/7/2013	63	5,335.67	5,333.20	28.0	43.0	15	5,305.20	5,290.20	5,270.20
MW-51	Multiunit 1	Background	Lewis Shale	4/28/2013	80	5,288.14	5,285.14	20.0	30.0	10	5,265.14	5,255.14	5,205.14
MW-61	Multiunit 1	Downgradient	Lewis Shale	9/16/2015	35	5,129.19	5,126.59	24.2	34.2	10	5,102.39	5,092.39	5,091.59
MW-74	Multiunit 1	Background	Lewis Shale	1/18/2017	40	5,219.09	5,216.70	8.1	18.1	10	5,208.60	5,198.60	5,176.70
MW-75	Multiunit 1	Downgradient	Lewis Shale	3/15/2017	41	5,126.80	5,124.80	29.0	39.0	10	5,095.80	5,085.80	5,083.80
MW-76	Multiunit 1	Downgradient	Lewis Shale	3/16/2017	33	5,116.23	5,114.30	11.8	26.8	15	5,102.50	5,087.50	5,081.30
MW-87	Multiunit 1	Downgradient	Lewis Shale	11/28/2018	50	5,076.53	5,074.29	15.0	45.0	30	5,059.29	5,029.29	5,024.29
MW-66	URS	Downgradient	Pictured Cliffs Sandstone	9/27/2015	33	5,344.69	5,344.70	15.0	25.0	10	5,329.70	5,319.70	5,311.70
MW-67	URS	Downgradient	Pictured Cliffs Sandstone	9/11/2015	31	5,352.76[b]	5,353.8[b]	19.6	29.6	10	5,334.42	5,324.42	5,323.02
MW-68	URS	Downgradient	Pictured Cliffs Sandstone	9/10/2015	30	5,353.58	5,353.95	19.0	29.0	10	5,334.95	5,324.95	5,323.95
MW-69	URS	Downgradient	Pictured Cliffs Sandstone	9/9/2015	35	5,357.66	5,355.26	24.3	34.3	10	5,330.96	5,320.96	5,320.26
MW-70	URS	Downgradient	Pictured Cliffs Sandstone	9/30/2015	53	5,371.12	5,368.62	40.0	50.0	10	5,328.62	5,318.62	5,315.62
MW-83	URS	Downgradient	Pictured Cliffs Sandstone	11/29/2018	35	5,343.15	5,341.51	14.0	29.0	15	5,327.51	5,312.51	5,306.51
MW-84	URS	Downgradient	Pictured Cliffs Sandstone	11/18/2018	35	5,338.23	5,339.34	10.0	30.0	20	5,329.34	5,309.34	5,304.34
MW-85	URS	Downgradient	Pictured Cliffs Sandstone	11/18/2018	35	5,352.78	5,353.69	15.0	30.0	15	5,338.69	5,323.69	5,318.69
MW-86	URS	Downgradient	Pictured Cliffs Sandstone	11/17/2018	35	5,338.76	5,338.74	10.0	30.0	20	5,328.74	5,308.74	5,303.74
MW-88	RWP	Downgradient	Pictured Cliffs Sandstone	12/6/2019	31	5365.25	5362.71	20	30	10	5342.71	5332.71	5331.71
MW-89	RWP	Downgradient	Pictured Cliffs Sandstone	12/6/2019	35	5370.21	5367.51	24	34	10	5343.51	5333.51	5332.51
MW-90	RWP	Downgradient	Pictured Cliffs Sandstone	12/7/2019	40	5374.08	5372.93	29	39	10	5343.93	5333.93	5332.93
MW-71	URS/CWTP	Background	Pictured Cliffs Sandstone	3/1/2016	50	5,362.91	5,363.62	22.5	42.5	20	5,341.12	5,321.12	5,313.62
MW-72	URS/CWTP	Background	Pictured Cliffs Sandstone	3/2/2016	61	5,381.62	5,379.09	50.7	60.7	10	5,328.39	5,318.39	5,318.09
MW-73	URS/CWTP	Background	Pictured Cliffs Sandstone	1/18/2017	45	5,353.95	5,351.90	28.9	43.9	15	5,323.00	5,308.00	5,306.90
DMX-01	N/A	Supplementary	Lewis Shale	4/15/1992	39	5,098.02	5,097.49	19.0	39.0	20	5,078.49	5,058.49	5,058.49
DMX-03	N/A	Supplementary	Lewis Shale	4/14/1992	38	5,085.50	5,084.85	18.0	38.0	20	5,066.85	5,046.85	5,046.85
DMX-04	N/A	Supplementary	Lewis Shale	4/15/1992	51	5,073.00	5,072.11	31.0	51.0	20	5,041.11	5,021.11	5,021.11
DMX-06	N/A	Supplementary	Lewis Shale	4/16/1992	35	5,077.40	5,076.42	15.0	35.0	20	5,061.42	5,041.42	5,041.42
IP-01	N/A	Supplementary	Lewis Shale	December 2013	38.5	5,101.81	5,099.39	13.5	38.5	20	5,085.89	5,060.89	5,060.89
IP-02	N/A	Supplementary	Lewis Shale	December 2013	27.5	5,090.79	5,088.27	17.0	27.0	10	5,071.27	5,061.27	5,060.77
IP-03	N/A	Supplementary	Lewis Shale	December 2013	34.5	5,091.08	5,088.68	24.0	34.0	10	5,064.68	5,054.68	5,054.18
IP-04	N/A	Supplementary	Lewis Shale	December 2013	32.5	5,095.92	5,093.46	22.0	32.0	10	5,071.46	5,061.46	5,060.96

Table 2
CCR Groundwater Monitoring System Summary

Well	CCR Unit	Well Designation	Hydrogeologic Unit	Date Installed	Borehole Depth [ft bgs]	Top of Casing Elevation [ft AMSL]	Ground Surface Elevation [ft AMSL]	Top of Screen [ft bgs]	Bottom of Screen [ft bgs]	Screen Length [ft]	Top Screen Elevation [ft AMSL]	Bottom Screen Elevation [ft AMSL]	Bottom Borehole Elevation [ft AMSL]
IP-05	N/A	Supplementary	Lewis Shale	December 2013	41	5,094.43	5,091.88	20.5	40.5	20	5,071.38	5,051.38	5,050.88
MW-01	N/A	Supplementary	Lewis Shale	9/06/1987	21.6	5,140.43	5,138.48	11.6	21.6	10	5,126.88	5,116.88	5,116.88
MW-03	N/A	Supplementary	Lewis Shale	3/13/1987	44.25	5,126.73	5,125.52	14.3	44.3	30	5,111.27	5,081.27	5,081.27
MW-05	N/A	Supplementary	Lewis Shale	3/12/1987	49.1	5,088.50	5,087.31	29.1	49.1	20	5,058.21	5,038.21	5,038.21
MW-06	N/A	Supplementary	Lewis Shale	3/12/1987	48.8	5082.71	5,080.19	28.8	48.8	20	5,051.39	5,031.39	5,031.39
MW-11	N/A	Supplementary	Lewis Shale	3/13/1987	49.9	5,111.96	5,110.48	29.9	49.9	20	5,080.58	5,060.58	5,060.58
MW-15	N/A	Supplementary	Lewis Shale	9/1/1987	52.7	5,093.93	5,092.28	22.2	52.2	30	5,070.08	5,040.08	5,039.58
MW-16	N/A	Supplementary	Lewis Shale	9/2/1987	54.8	5,101.32	5,100.42	35.5	54.8	19	5,064.92	5,045.62	5,045.62
MW-17R	N/A	Supplementary	Lewis Shale	December 2013	32	5,093.09	5,090.43	16.5	31.5	15	5,073.93	5,058.93	5,058.43
MW-18	N/A	Supplementary	Lewis Shale	9/3/1987	55	5,089.10	5,088.06	25.5	55.0	30	5,062.56	5,033.06	5,033.06
MW-19	N/A	Supplementary	Lewis Shale	9/3/1987	49.7	5,127.40	5,126.34	29.2	49.7	21	5,097.14	5,076.64	5,076.64
MW-21	N/A	Supplementary	Lewis Shale	9/4/1987	30	5,155.04	5,154.47	10.6	30.0	19	5,143.87	5,124.47	5,124.47
MW-22	N/A	Supplementary	Lewis Shale	9/4/1987	30.4	5,156.51	5,156.30	10.4	30.4	20	5,145.90	5,125.90	5,125.90
MW-23R	N/A	Supplementary	Lewis Shale	12/6/2013	41.5	5,101.53	5,099.08	21.0	41.0	20	5,078.08	5,058.08	5,057.58
MW-24	N/A	Supplementary	Lewis Shale	9/05/1987	69.7	5,081.65	5,080.41	59.7	69.7	10	5,020.71	5,010.71	5,010.71
MW-26	N/A	Supplementary	Lewis Shale	9/06/1987	50.5	5,139.26	5,138.36	40.5	50.5	10	5,097.86	5,087.86	5,087.86
MW-30	N/A	Supplementary	Lewis Shale	6/7/2010	23	5,091.67	5,092.06	13.0	23.0	10	5,079.06	5,069.06	5,069.06
MW-31	N/A	Supplementary	Lewis Shale	6/7/2010	24	5,092.59	5,089.96	14.0	24.0	10	5,075.96	5,065.96	5,065.96
MW-32	N/A	Supplementary	Lewis Shale	6/7/2010	20	5,087.65	5,084.94	10.0	20.0	10	5,074.94	5,064.94	5,064.94
MW-36R	N/A	Supplementary	Lewis Shale	December 2013	34	5,093.33	5,090.76	13.5	33.5	20	5,077.26	5,057.26	5,056.76
MW-38R	N/A	Supplementary	Lewis Shale	December 2013	39	5,094.12	5,091.41	13.5	38.5	25	5,077.91	5,052.91	5,052.41
MW-45	N/A	Supplementary	Lewis Shale	May 2013	39	5,089.56	5,087.13	24.0	39.0	15	5,063.13	5,048.13	5,048.13
MW-46	N/A	Supplementary	Lewis Shale	May 2013	26	5,064.30	5,061.91	16.0	26.0	10	5,045.91	5,035.91	5,035.91
MW-52	N/A	Supplementary	Lewis Shale	May 2013	82	5,210.41	5,208.06	67.0	82.0	15	5,141.06	5,126.06	5,126.06
MW-54	N/A	Supplementary	Lewis Shale	5/20/2013	91	5,217.82	5,218.38	76.0	91.0	15	5,142.38	5,127.38	5,127.38
MW-56	N/A	Supplementary	Lewis Shale	December 2013	36.5	5,091.49	5,089.14	26.0	36.0	10	5,063.14	5,053.14	5,052.64
MW-57	N/A	Supplementary	Lewis Shale	December 2013	42.5	5,088.30	5,085.70	22.0	42.0	20	5,063.70	5,043.70	5,043.20
MW-60	N/A	Supplementary	Lewis Shale	9/16/2015	25	5144.1	5141.5	14.3	24.3	10	5,127.16	5,117	5,116.50
MW-77S	N/A	Supplementary	Lewis Shale	11/8/2018	80	5094.94	5092.35	24.0	44.0	20	5,068.35	5,048.35	5,012.35
MW-78S	N/A	Supplementary	Lewis Shale	11/13/2018	80	5,088.79	5,086.51	24.0	44.0	20	5,062.51	5,042.51	5,006.51
MW-79S	N/A	Supplementary	Lewis Shale	11/20/2018	58	5,086.90	5,084.35	16.0	36.0	20	5,068.35	5,048.35	5,026.35
MW-80S	N/A	Supplementary	Lewis Shale	11/16/2018	81	5,086.80	5,084.29	35.0	55.0	20	5,049.29	5,029.29	5,003.29
MW-81	N/A	Supplementary	Lewis Shale	11/26/2018	36	5,086.41	5,084.07	13.0	33.0	20	5,071.07	5,051.07	5,048.07
MW-82S	N/A	Supplementary	Lewis Shale	11/27/2018	65	5,093.37	5,091.02	17.0	37.0	20	5,074.02	5,054.02	5,026.02

Notes:

Source of presented information presented is AECOM, 2017 and Sakura Engineering & Surveying, 2017 and 2019

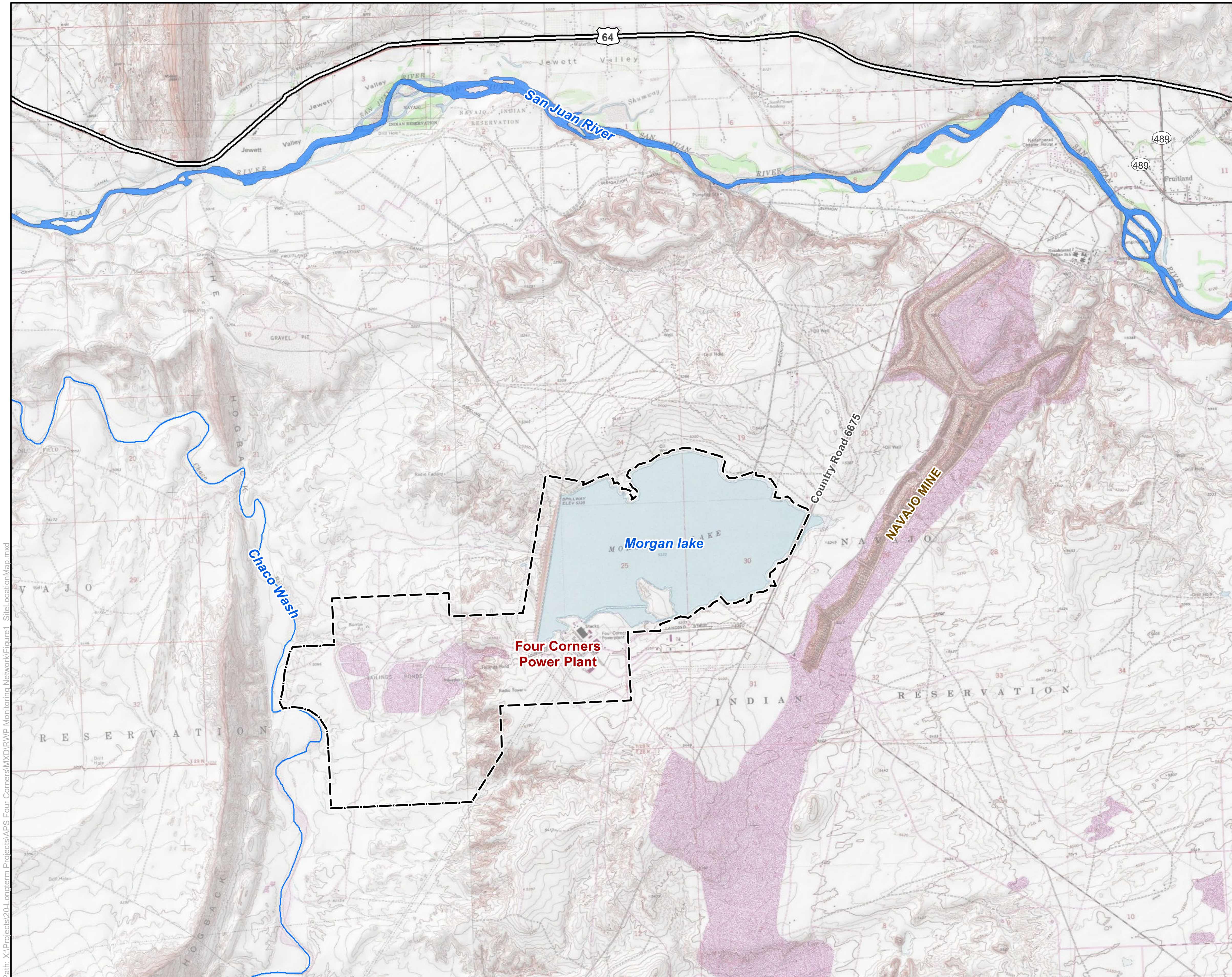
Vertical datum is NAVD 88

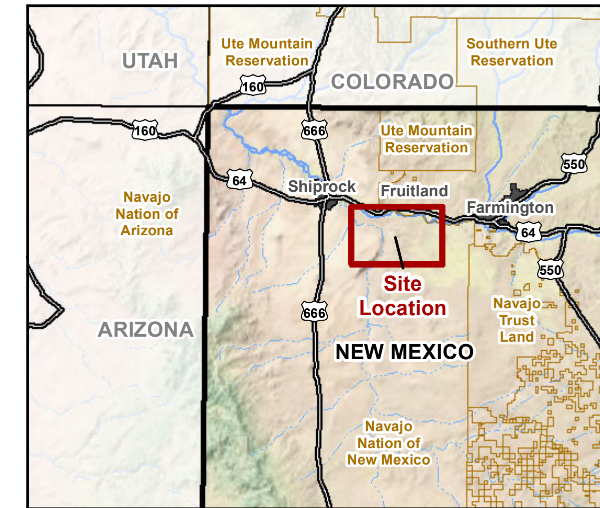
(a) - Estimated

(b) - New surveyed elevation after wellhead modifications

AMSL - Above mean sea level
bgs - below ground surface
CCR - coal combustion residuals
CWTP - Combined Waste Treatment Pond
DFADA - Dry Fly Ash Disposal Area
ft - feet
N/A - Supplementary well not associated with a CCR unit
RWP - Return Water Pond
URS - Upper Retention Sump

FIGURES





Area of Detail

Legend

— — Four Corners Power Plant Lease Boundary

0 0.4 0.8 Miles

N

Arizona Public Service
Four Corners Power Plant
Fruitland, New Mexico

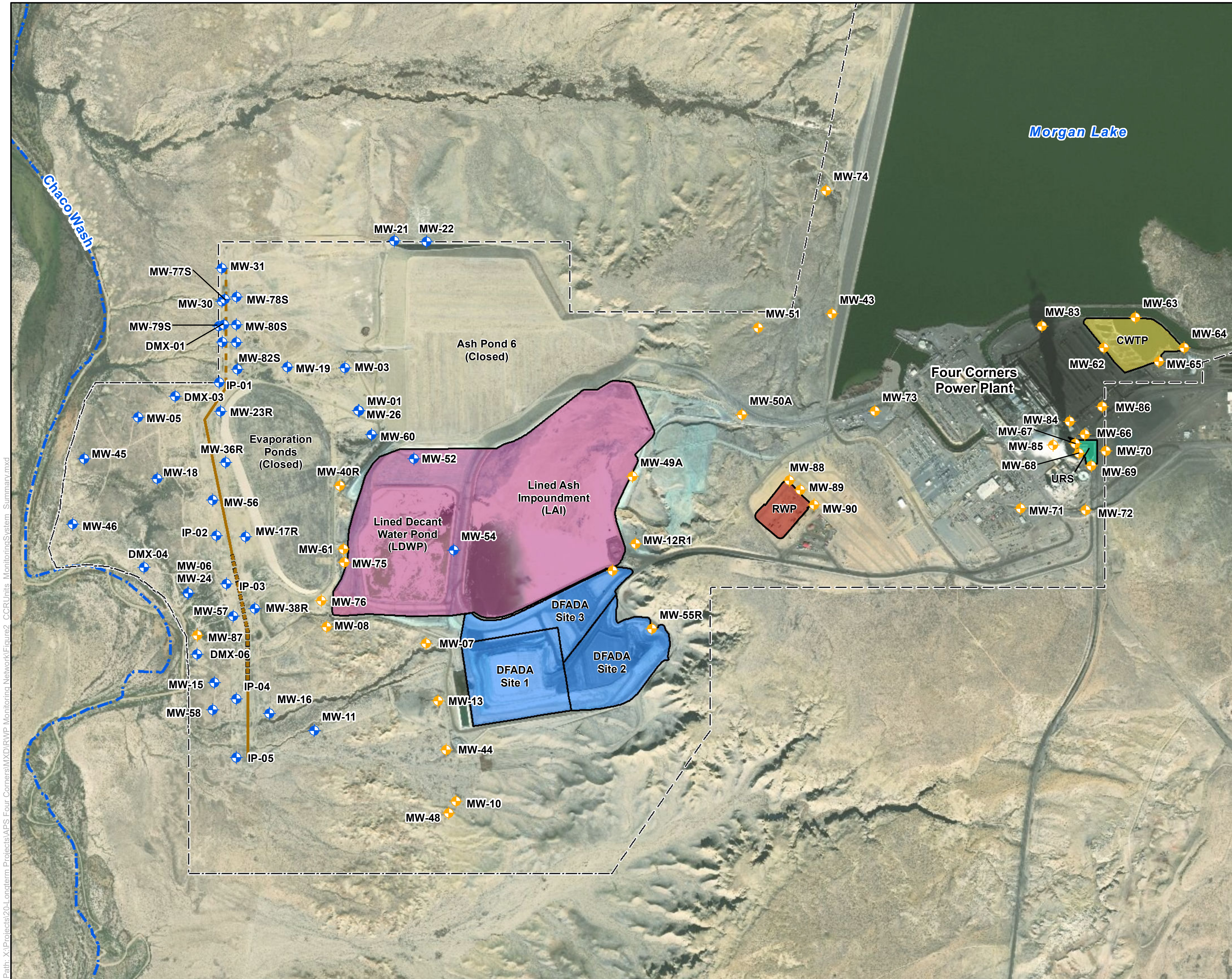
FIGURE 1

Site Location Map

Job No.	14-2018-2068
PM:	EHL
Date:	4/23/2020
Scale:	1"= 0.8 miles

wood.

Path: X:\Projects\20-Longterm Projects\APS Four Corners\MXD\RWP Monitoring Network\Figure1_SiteLocationMap.mxd



Legend

- CCR Monitoring Well Location
- Supplementary Site Monitoring Well Location
- FCPP Lease Boundary
- North Intercept Trench
- South Intercept Trench
- Approximate Extent of High Flow Zone
- Ephemeral Surface Water Feature

CCR Units

- Multiunit 1(LAI and LDWP)
- Dry Fly Ash Disposal Area (DFADA)
- Combined Waste Treatment Pond (CWTP)
- Upper Retention Sump (URS)
- Return Water Pond (RWP)

Notes:

CCR	Coal Combustion Residuals
CWTP	Combined Waste Treatment Pond
DFADA	Dry Fly Ash Disposal Area
FCPP	Four Corners Power Plant
LAI	Lined Ash Impoundment
LDWP	Lined Decant Water Pond
URS	Upper Retention Sump
RWP	Return Water Pond

0 350 700 1,400 Feet

N

Arizona Public Service
Four Corners Power Plant
Fruitland, New Mexico

FIGURE 2

CCR Units and Monitoring System Summary

Job No.	14-2018-2068
PM:	EHL
Date:	4/24/2020
Scale:	1" = 1400'