ARIZONA PUBLIC SERVICE COMPANY

STATISTICAL DATA ANALYSIS WORK PLAN

COAL COMBUSTION RESIDUALS RULE GROUNDWATER MONITORING SYSTEM COMPLIANCE

January 30, 2023

Please be advised that, effective September 21, 2022, Wood Environment & Infrastructure Solutions, Inc. Was acquired by WSP. Due to the acquisition, we have changed our name to WSP USA Environment & Infrastructure Inc. No other aspects of our legal entity or capabilities have changed for this report, including our Federal Tax ID which remains 91-1641772. Correspondence for this report should continue to be addressed to the undersigned.

STATISTICAL DATA ANALYSIS WORK PLAN

COAL COMBUSTION RESIDUALS RULE GROUNDWATER MONITORING SYSTEM COMPLIANCE

ARIZONA PUBLIC SERVICE COMPANY

PROJECT NO.: 14-2022-2007
DATE: JANUARY 30, 2023

WSP USA ENVIRONMENT & INFRASTRUCTURE INC.
4600 EAST WASHINGTON STREET, SUITE 600
PHOENIX, AZ 85034-1917

T: 1-602- 733-6000
WSP.COM

CERTIFICATION STATEMENT

I, Rebecca Weaver, as a qualified groundwater scientist and professional engineer have reviewed the *Statistical Data Analysis Work Plan*, Cholla Power Plant, Navajo County Arizona, Project # 14-2022-2007 dated January 30, 2023. I certify that statistical methods described herein are appropriate for the Arizona Public Service Company Cholla Power Plant site as required for compliance with CCR groundwater monitoring, corrective action, closure, and post-closure requirements detailed in 40 CFR §257.90 through §257.104.

PREPARED BY

Rebecca Weaver
Assistant Vice President – Civil Engineer

Registered Professional Engineer (Civil)
CERTIFICATE NO.
54953
REBECCA A.
WEAVER
01/30/2023
DATE SIGNED
ARIZONA U.S.A.

APPROVED[1] BY

I, Maren Henley, a qualified professional, certify that the statistical methods described in this document, as supported by the Statistical Analysis Plan in the facility's Operating Record, are appropriate for evaluating the groundwater monitoring data for the CCR management area.

Maren Henley
Assistant Vice President – Geological Engineer

I, Tim Glover, a qualified professional, certify that the statistical methods described in this document, as supported by the Statistical Analysis Plan in the facility's Operating Record, are appropriate for evaluating the groundwater monitoring data for the CCR management area.

Tim Glover
Assistant Vice President – Environmental Scientist – Geochemistry and Statistics

[1] Approval of this document is an administrative function indicating readiness for release and does not impart legal liability on to the Approver for any technical content contained herein. Technical accuracy and fit-for-purpose of this content is obtained through the review process. The Approver shall ensure the applicable review process has occurred prior to signing the document.

WSP USA Environment & Infrastructure Inc. prepared this report solely for the use of the intended recipient, Arizona Public Safety Company, in accordance with the professional services agreement. The intended recipient is solely responsible for the disclosure of any information contained in this report. The content and opinions contained in the present report are based on the observations and/or information available to WSP USA Environment & Infrastructure Inc. at the time of preparation. If a third party makes use of, relies on, or makes decisions in accordance with this report, said third party is solely responsible for such use, reliance or decisions. WSP USA Environment & Infrastructure Inc. does not accept responsibility for damages, if any, suffered by any third party as a result of decisions made or actions taken by said third party based on this report. This limitations statement is considered an integral part of this report.

The original of this digital file will be conserved by WSP USA Environment & Infrastructure Inc. for a period of not less than 10 years. As the digital file transmitted to the intended recipient is no longer under the control of WSP USA Environment & Infrastructure Inc., its integrity cannot be assured. As such, WSP USA Environment & Infrastructure Inc. does not guarantee any modifications made to this digital file subsequent to its transmission to the intended recipient.

TABLE OF CONTENTS

TABLE OF CONTENTS

FIGURES

TABLE OF CONTENTS

TABLES

ACRONYMS AND ABBREVIATIONS

%	percent
§	Section
AMEC	AMEC Earth & Infrastructure, Inc.
AMSL	above mean sea level
ANOVA	analysis of variance
APS	Arizona Public Service
ASD	Alternative source demonstration
BAM	Bottom Ash Monofill
BAP	Bottom Ash Pond
CCR	coal combustion residuals
CFR	Code of Federal Regulations
CSM	Conceptual Site Model
DQR	Double Quantification Rule
EDA	exploratory data analysis
FAP	Fly Ash Pond
ft	foot, feet
GWPS	groundwater protection standard
KM	Kaplan-Meier
LPL(s)	lower prediction limit(s)
MCL	Maximum Contaminant Level
RL	reporting limit
ROS	regression order on statistics
SDAWP	Statistical Data Analysis Work Plan
SEDI	Sedimentation Pond
SSI	statistically significant increase
SSL	statistically significant level
SWFPR	sitewide false positive rate

TABLE OF CONTENTS

TDS	total dissolved solids
UPL(s)	upper prediction limit(s)
UTL(s)	upper tolerance limit(s)
USEPA	United States Environmental Protection Agency
VSP	Visual Sampling Plan
Wood	Wood Environment & Infrastructure Solutions, Inc.
WSP	WSP USA

1 INTRODUCTION

This Statistical Data Analysis Work Plan (SDAWP) has been prepared in collaboration by WSP USA (WSP) formerly conducting business as Wood Environment & Infrastructure Solutions, Inc. (Wood) and Geoscience Consulting Strategies LLC on on behalf of Arizona Public Service Company (APS) for the Cholla Power Plant (Cholla, or the Site) located near Joseph City, Arizona in Navajo County. This SDAWP details the scope and implementation of statistical criteria and procedures to evaluate site data for CCR surface impoundments and landfills (CCR Units) in accordance with Coal Combustion Residuals (CCR) requirements detailed in 40 Code of Federal Regulations (CFR) Sections (§) 257.90 through 257.98 (herein referred to as the CCR Rule) (Federal Register, 2020). 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.

Previous statistical analysis documentation prepared for the site (Montgomery & Associates, 2017a and Wood, 2018a) present how detection and assessment monitoring data were to be evaluated. This 2022 update to the SDAWP is intended to supplement previous documentation and present a comprehensive approach for performing statistical analyses of site groundwater data from the detection and assessment monitoring programs to corrective action, closure, and post-closure.

1.1 OBJECTIVES

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

- Assess adequacy of sampled data to service statistical procedures (Sections 1.0 and 2.0);
- Select appropriate statistical methods for each constituent and monitoring well pairing (Sections 2.0 through 5.0);
- Develop background constituent concentration levels (Section 4.0);
- Develop groundwater protection standards (GWPSs) (Section 5.0);
- Identify statistically significant increases (SSIs) in constituent concentrations over background levels and statistically significant levels (SSLs) of constituent concentrations above GWPSs (Sections 3.0 and 5.0);
- Determine the appropriate workflow under detection monitoring (Section 3.0), assessment monitoring (Section 4.0), corrective action (Section 6.0) and closure and post-closure (Section 7.0), as necessary; and
- Make recommendations for future sampling and data evaluations (Section 8.0).

1.2 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 Cholla CCR groundwater monitoring system has been 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 effects 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).

The original CCR Groundwater Monitoring System Certification Report was completed in 2017 (Montgomery & Associates, 2017b). An addendum Groundwater Monitoring System Certification Report has since been completed to certify post-2017 updates to the BAP and FAP monitoring systems (Wood, 2022a). The certification reports detail the baseline hydrogeological CSM used to design the CCR groundwater monitoring system for Cholla. Salient information regarding the baseline hydrogeologic CSM is supplemented by the *Well Completion Report for APP Monitoring Wells* prepared by AMEC Environment & Infrastructure, Inc. (AMEC) (AMEC, 2012), and the *Point of Compliance Evaluation Report* prepared by Montgomery & Associates (Montgomery & Associates, 2011). A summary CSM is presented in the following subsections to document:

- Preexisting site-specific information;
- The adequacy of groundwater monitoring network(s) 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.

More detailed information regarding the CSM is present in the *Annual Groundwater Monitoring and Corrective Action Report for 2021,* prepared by Wood (Wood, 2022e). The CSM may be refined based on the results of the statistical evaluation of water quality data.

1.2.1 SITE DESCRIPTION

The Site is located in an arid climate within the Little Colorado River Basin, which receives on average 6 to 12 inches of precipitation annually (Figure 1-1). Elevation of the Site is approximately 5,025 ft amsl within the Colorado Plateau physiographic province of northeastern Arizona. This area is characterized by canyons, high elevations, and narrow, widely-spaced riverbeds. South of the plant (Figure 1-2) is the Little Colorado River, a meandering, perennial stream with intermittent reaches in a large alluvial floodplain. Topography is characterized by rolling terrain, open vistas, and incised drainages and arroyos. At the plant, the ground surface gently slopes south towards the Little Colorado River at approximately 60 ft per mile. Approximately two miles to the north and south of the plant, the ground surface rises out of the alluvial floodplain to elevations of 5,100 to 5,200 ft amsl.

Cholla is a coal-fired power plant with two operating electrical generating Units (Units 1 and 3) as of the date of this SDAWP. Units 1 and 3 have a net generating capacity of 425.9 megawatts. Coal burned at the plant was previously sourced from the McKinley Mine in New Mexico but when the mine was closed in 2009, the source of the coal switched to the Lee Ranch and El Segundo mines near Grants, New Mexico. APS plans to cease coal-fired boiler operations at Cholla no later than April 2025.

The plant and associated infrastructure are located on land owned by APS adjacent to Interstate 40 (I-40) between the cities of Winslow and Holbrook in Navajo County, Arizona (Figure 1-1). The main plant is west-adjacent to the Cholla Reservoir, a cooling pond for Unit 1 and water storage reservoir for the plant. Cholla Reservoir was constructed in the early 1900s for the Joseph City Irrigation Company and configured in its current location and design by APS in 1961 to receive deliveries of groundwater pumped from the nearby Cholla Well Field, which extracts from the C-Aquifer. Typical water surface elevation of the Cholla Reservoir is 5,022 feet (ft) above mean sea level (amsl).

Plant infrastructure includes four single CCR Units, of which further details are presented in Table 1-1:

— The Fly Ash Pond (FAP) is the largest surface impoundment at the site, is 430 acres in surface area, and receives slurried flue gas desulfurization solids and fly ash from the plant. The FAP is constructed within a drainage area by damming an un-named tributary to the Little Colorado River east of Cholla Reservoir and north of I-40 (Figure 1-2). While it is constructed primarily on the relatively impermeable Moqui Member of the Moenkopi Formation, a relatively thin layer of local wash alluvium underlies parts of the FAP (particularly in the vicinity of the dam). The FAP dam has a clay core that extends through the alluvium to bedrock (the Moqui) where the alluvium was less than 20 ft thick at the time of dam construction. In the middle of the dam where the alluvium was greater than 20 ft thick, a cutoff wall was constructed that generally extends approximately 1 to 2 ft into bedrock. Groundwater in the shallow alluvium downgradient of the FAP flows west-southwest, under I-40 and then joins groundwater in the Little Colorado River Alluvium which flows to the northwest. The FAP is in corrective action and currently has a planned closure completion date of late 2028 as part of the impending cessation of coal-fired power generation at the Site. Details of the closure activities are provided in *Cholla Power Plant Closure Plan §257.102(b) Fly Ash Pond, Amendment 2* prepared by AECOM (2020).

— The Sedimentation Pond (SEDI) was a smaller 1.3-acre CCR impoundment located west of the plant (Figure 1-2), hydraulically downgradient from the other three CCR units. Removal of CCR from the SEDI was completed in October 2021. A 60,000-gallon concrete tank was constructed and commissioned to fulfill the function of the SEDI in 2020. The location of the former SEDI is underlain by the Little Colorado River alluvium. Groundwater in the alluvium near the SEDI flows parallel to the direction of Little Colorado River surface flows, approximately northwest. At the time CCR removal was completed in October 2021, the SEDI was in assessment monitoring. Based on the results of groundwater monitoring conducted after completion of CCR removal indicating no GWPS exceedances or exceedances of respective background threshold values of Appendix IV constituents, decontamination and thus closure of the SEDI were determined to be complete as of October 21, 2022.

— The Bottom Ash Pond (BAP) is a larger, 105-acre surface impoundment that receives slurried bottom ash from the plant. The BAP is constructed in the Tanner Wash drainage area by damming a minor tributary to Tanner Wash (Figure 1-2). Tanner Wash is a south-flowing tributary to the Little Colorado River. The northern and western edges of the BAP are constructed on the Moenkopi Formation, whereas the southern edge rests primarily on the local Tanner Wash alluvium. The BAP dam extends to bedrock where bedrock is shallow; in the central portion of the dam, where bedrock is deeper, a slurry cutoff wall extends to an elevation of approximately 4,980 ft amsl. There is an approximately 10 to 20-ft thick layer of alluvium at the base of the cutoff wall above the Moqui member of the Moenkopi Formation. Groundwater in the alluvium near the BAP flows south-southwest along Tanner Wash to its confluence with the Little Colorado River. The BAP is in corrective action and currently has a planned closure completion date of late 2028 as part of impending Site cessation of coal-fired power generation at the Site. Details of the closure activities are provided in *Cholla Power Plant Closure Plan §257.102(b) Bottom Ash Pond, Amendment 2* prepared by AECOM (2020).

- The Bottom Ash Monofill (BAM) is a 41-acre landfill for bottom ash solids from the plant. It is constructed on the Moqui Member of the Moenkopi Formation, and water levels from nearby wells indicate that the Moenkopi is unsaturated beneath the BAM. In the vicinity of the BAM, the groundwater in the C-Aquifer flows to the northwest. The BAM is in detection monitoring. Initiation of closure of the BAM is tentatively planned for late 2025.

1.2.2 SITE GEOLOGY

The physiographic province where the plant is located, the Colorado Plateau, is typified by horizontal layered sequences of sedimentary rock, primarily sandstones, siltstones, and claystones. At the plant and CCR units, there are three geologic units that are expected to influence groundwater flow and variations in naturally occurring constituent concentrations across the site. The units are as follows (in descending order from ground surface):

- **Little Colorado River, Tanner Wash, and other wash alluviums**: These quarternary alluviums overlie the bedrock formations in localized areas at Cholla and surrounding CCR Units. The alluviums are unconsolidated, heterogeneous, and consists of clay, silt, sand, and gravel. In general, the Tanner Wash (near the BAP) and other wash (near the FAP) alluviums are finer-grained with abundant clay layering, while the Little Colorado River alluvium is less fine-grained. The alluvium ranges in thickness from non-existent to approximately 200 feet, and in general is thickest underneath the plant and Cholla Reservoir. Around the CCR Units, the alluvium ranges from approximately 50 feet thick in the vicinity of the FAP dam to 100 feet thick in the vicinity of the BAP dam.
- **Moenkopi Formation**: The Moenkopi Formation is the uppermost geologic unit beneath the plant and the CCR Units and is present at land surface in areas where the alluvium is non-existent. The thickness of the Moenkopi Formation near the plant ranges from non-existent to over 300 feet thick; where it is sufficiently thick, the Moenkopi Formation acts as an aquitard between the shallow alluvium aquifer to the underlying Coconino Sandstone Aquifer due to the Moqui Member. The Moenkopi Formation consists of three members, described below in descending order:
  - **Holbrook Member**: A relatively permeable, well-consolidated sandstone, the Holbrook Member is the uppermost member of the Moenkopi Formation and is not known to be present in the immediate vicinity of the plant. The Holbrook Member has been noted near the BAP and is variably saturated in localized areas where it is contact with overlying saturated alluvium.
  - **Moqui Member**: The Moqui Member is the primary confining unit within the Moenkopi Formation and consists of maroon and greenish mudstone with abundant gypsum. The Moqui Member is approximately 250 to 300 feet thick near the plant and is known to be variably saturated in localized areas where it is contact with overlying saturated alluvium.
  - **Wupatki Member**: The lowest member of the Moenkopi Formation, the Wupatki Member is approximately 30 to 50 feet thick. The Wupatki Member is comprised of relatively permeable sandstone and is in hydraulic connection with the underlying Coconino Sandstone.
- **Coconino Sandstone**: The Permian-age Coconino Sandstone is the principal lithologic unit of the C-aquifer, a regionally important aquifer for water supply. It is composed of very fine- to fine-grained sandstone with variable permeability depending on the degree of fracturing and cementation. The unit is approximately 375 to 400 ft thick in the vicinity of the plant.

- **Schnebly Hill Formation**: The Schnebly Hill Formation is a very fine-grained, reddish sandstone that is about 300 to 350 feet thick near the plant. It is part of the C-aquifer, but its hydraulic conductivity is about 10 to 28 percent that of the Coconino Sandstone.
- **Supai Formation**: This Pennsylvanian to Lower Permian Supai Formation underlies the Coconino Sandstone. It has minimal impact on the surface operations of Cholla, other than containing an approximately 600-ft thick deposit of halite and anhydrite in the Cholla well field area that impacts groundwater quality both regionally and in the vicinity of the plant.

1.2.3 SITE HYDROGEOLOGY

Two hydrostratigraphic units are conceptualized beneath the plant and associated CCR units. These units form the basis for the hydrogeologic CSM developed by Montgomery & Associates (2011 and 2017) and updated by Wood (2022a).

The first hydrogeologic unit, the Quaternary alluvial aquifer, is present under the plant area, Cholla Reservoir, and the Tanner Wash and Little Colorado River drainage channels. The alluvium in these areas receive recharge from the Little Colorado River and any leakage through anthropogenic features such as the reservoir and the nearby Joseph City Canal. The alluvial aquifer is not used as a drinking water supply but does support a riparian habitat. The alluvium is not uniformly saturated. Where present, groundwater flows generally in the downstream direction of the drainages under which it is present, that is, east to west in the Little Colorado River alluvium and north to south in the Tanner Wash alluvium. Groundwater flow in the Little Colorado River alluvium, especially near the FAP, is also influenced by deeper paleochannels that may not coincide with the present river channel.

The second hydrogeologic unit is the C-aquifer, which consists of the Coconino Sandstone and Schnebly Hill Formation in the vicinity of the plant. Groundwater in this aquifer is under confined conditions in areas north of the Little Colorado River where the Moqui Member of the Moenkopi Formation acts as a confining bed. Groundwater movement in the C-aquifer is generally to the north. However, the Cholla well field (southwest of the plant) has created a cone of depression that has made the groundwater flow in a westerly direction in that area. Near the FAP, the inferred flow of the groundwater is to the west or southwest, possibly due to the broad, northwest-trending anticline that extends from the FAP to near Joseph City.

The alluvial aquifer and the C-aquifer are separated by the Moenkopi Formation, a regional aquitard that creates a barrier between the two aquifers in the vicinity of Cholla. In areas where the C-aquifer in the Coconino Sandstone is confined (primarily north of the Little Colorado River), the Wupatki Member of the Moenkopi has been observed to be water-bearing. In locations at the BAP and FAP where the Holbrook or Moqui Member is in direct contact with saturated alluvium (principally at the base of historical drainage channels where the unit is weathered or fractured from penetrations of monitoring well installations) overlying alluvial groundwater infiltrates into the uppermost layers of the unit. However, the Moqui Member, which is 250 to 300 feet thick in the vicinity of the plant, prevents hydraulic connection between the alluvial aquifer and the C-aquifer and is effectively bedrock when considering water quality conditions and groundwater movement in the alluvial aquifer.

1.3 MONITORING SYSTEM SAMPLING ADEQUACY

Multiple monitoring well systems are in place at Cholla to monitor groundwater conditions beneath the four site CCR Units. The installation of these networks is summarized in CCR Groundwater Monitoring System Certification Reports and are identified as compliant with 40 CFR §257.91(a) through (e) (Montgomery &

Associates, 2017b and Wood, 2022a). Wood has also prepared an updated Sampling and Analysis Plan (Wood, 2022b) to document the methods and procedures used to conduct groundwater sampling and evaluate potential impacts of site CCR units.

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.

1.3.1 DOWNGRADIENT GROUNDWATER MONITORING WELL NETWORKS

A total of 25 downgradient wells are in place at the site to monitor the downgradient groundwater conditions of each CCR Unit (Table 1-2). Downgradient boundary wells assess the groundwater conditions at the boundary of each CCR unit and the remaining downgradient wells evaluate the nature and extent of groundwater conditions associated with each CCR unit. Eighteen of these monitoring wells are installed in the alluvium. Four of the monitoring wells are installed either fully or partially screened into Moqui where it was found to be saturated. The remaining three wells are completed in the Coconino Sandstone. These wells are grouped by respective CCR Unit, as described below:

— **FAP Downgradient Wells (Little Colorado River and other wash alluvial aquifer)**: The groundwater flow direction in the alluvium downgradient of the FAP dam (i.e., the waste boundary) is west-southwest. The alluvial thickness in this area is limited; in some places it may be up to 50 feet thick but in others it is non-existent. On this basis, three downgradient boundary wells, W-123, M-50A, and M-51A (each screened in the Little Colorado alluvium), were originally designated for the FAP in 2017 (Montgomery & Associates, 2017b). In 2018 exceedances of groundwater protection standard(s) (GWPSs) for arsenic, cobalt, fluoride, lithium, and molybdenum were declared for downgradient FAP wells and the CCR Unit was ultimately moved into the corrective action phase of the CCR monitoring program (Wood, 2018b). Per the requirements of 40 CFR §257.95(g)(1), three additional downgradient wells, MW-65A, MW-66A, and MW-67A (each screened in the Little Colorado alluvium), were installed in late 2018 to characterize the nature and extent of the constituents for which the FAP was moved into corrective action (Wood, 2020a). In early 2020 an alternative source demonstration (ASD) was completed for cobalt at the FAP citing inadequate laboratory reporting limits, not the CCR Unit, as the source of the false cobalt exceedances (Wood, 2020b). Investigations into natural variations in aquifer chemistry in the vicinity of the FAP were conducted in 2021 and reduced conditions in groundwater, near downgradient nature and extent well MW-67A specifically, were cited as the source of elevated arsenic in the well (Wood, 2021). Ongoing investigations of groundwater reduction/oxidation process have continued for arsenic exceedances at FAP downgradient boundary well M-51A. To address deteriorating well condition and drilling investigations that indicated W-123 was screened across alluvium and the Moqui, the well was abandoned and replaced in 2021 with W-123R (Wood, 2022c). The monitoring well designations, spatial density, and coverage of the monitoring well network are adequate and representative unless future observations prove otherwise.

— **SEDI Downgradient Wells (Little Colorado River alluvial aquifer):** The groundwater flow direction in the alluvium underlying the SEDI is to the west-northwest, approximately parallel to the surface water flow in the Little Colorado River. Three downgradient boundary wells were originally designated for the SEDI: M-56A, M-57A, and M-58A; these are screened within the Little Colorado River alluvium. No changes have since been made to the SEDI downgradient monitoring system. The monitoring well designations, spatial density, and coverage of the monitoring well network are adequate and representative unless future observations prove otherwise.

- **BAP Downgradient Wells (Tanner Wash alluvial aquifer):** The groundwater flow direction in the alluvium underlying the BAP is generally to the southwest along the Tanner Wash; however, there is a radial component of groundwater flow towards the east-southeast due to hydraulic head from the BAP. Five downgradient boundary monitoring wells were originally designated for the BAP: M-52A, M-53A, W-305, W-306, and W-314; each are screened in the Tanner Wash alluvium. In 2018 exceedances of GWPSs for lithium and cobalt were declared for downgradient BAP wells and the Unit was ultimately moved into the corrective action phase of the CCR monitoring program (Wood, 2018c). In 2019, an ASD was completed for lithium at the BAP citing natural variation of this constituent in groundwater in the aquifer, not the CCR Unit, as the source of the lithium exceedances (Wood, 2019). Per the requirements of 40 CFR §257.95(g)(1), eight additional downgradient wells, MW-71A, MW-72MA, MW-73A, MW-74M, MW-76A, MW-77A, MW-78A, and MW-79A, were installed in 2021 to characterize the nature and extent of cobalt (for which the BAP was moved into corrective action) (Wood, 2022d). Based on an assessment of data collected during the well installation and initial groundwater monitoring, MW-71A, MW-72M, MW-73A, and MW-74M were re-designated as a downgradient boundary wells in 2022 (WSP, 2023). The monitoring well designations, spatial density, and coverage of the monitoring well network are adequate and representative unless future observations prove otherwise.
- **BAM Downgradient Wells (C-Aquifer):** The uppermost hydrogeologic unit underlying the BAM is the C-aquifer in the Coconino Sandstone, which flows towards the northwest in this vicinity. Three downgradient boundary monitoring wells were originally installed to monitor the quality of groundwater passing the waste boundary of the BAM. These wells are named M-59, M-60, and M-61, and they are completed in the Coconino Sandstone. No changes have since been made to the BAM downgradient monitoring system. The monitoring well designations, spatial density, and coverage of the monitoring well network are adequate and representative unless future observations prove otherwise.

1.3.2 BACKGROUND GROUNDWATER MONITORING WELLS

The purpose of background 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 the CCR unit (40 CFR §257.91).

Per the original CCR Groundwater Monitoring System Certification Report (Montgomery & Associates, 2017b), the following monitoring wells are designated as "background monitoring wells" for the four CCR Units:

- **Background Well for the FAP and the BAP (Little Colorado River alluvial aquifer):** The upgradient boundary of the FAP rests on a thick section of the Moenkopi Formation; there is no appreciable saturated alluvium present in the area upgradient from the FAP boundary. Therefore, background well M-64A was installed west of the plant in the Little Colorado River floodplain. The BAP, in the Tanner Wash alluvium, discharges to and is hydraulically connected to the Little Colorado River alluvium. Because hydrogeologic conditions at the BAP prevented installation of an upgradient background well (as they did at the FAP), M-64A also serves as the background well for the BAP. Travel time calculations performed for the original CCR Groundwater Monitoring System Certification Report (Montgomery & Associates, 2017b) indicated that M-64A is located far enough downgradient from the FAP and the BAP to represent unimpacted groundwater; however, it is notable that selection of this background well location is not ideal and has the potential to promote spatial heterogeneity issues in statistical data analysis.

- **Background Wells for the SEDI (Little Colorado River alluvial aquifer):** The groundwater flow direction in the vicinity of the SEDI is to the west-northwest. Background well M-62A is installed in the alluvium on the east (upgradient) side of the SEDI.
- **Background Wells for the BAM (Coconino Sandstone):** The groundwater flow direction in the C-aquifer in the vicinity of the BAM is to the northwest. Background well M-54 is installed in the Coconino Sandstone on the southeast (upgradient) side of the BAM.

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. In the case of the Cholla site, each CCR unit has a single background monitoring well.

Due to the natural heterogeneity of the geologic and hydrogeologic conditions underlying Cholla, 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 natural aquifer variation (as was noted in the BAP lithium ASD [Wood, 2019]), local climatic regimes, potential leakage from Cholla Reservoir, and potential operational activity at the site. The groundwater monitoring well networks, respective to sampling coverage and frequency, are expected to adequately represent this spatial and temporal heterogeneity until proven otherwise.

The adequacy of designated background monitoring wells will continue to be assessed using groundwater elevation data, supplementary analytical data, a working understanding of the spatial heterogeneity of hydrogeochemistry underlying Cholla, and statistical characteristics of constituents of concern. Historical groundwater chemistry data may be consulted during this evaluation. With respect to use of MW-64 as a background well for the FAP and BAP, alternate statistical comparison methods are available to assess CCR unit compliance absent of adequate and representative background sample data. If achieving adequate and representative background is not possible, alternate methods (i.e., intrawell comparisons) are discussed in Section 3.0.

2 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. 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 temporal or spatial trends (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.

This Section 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.

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

— Fluoride

— 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 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 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.

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 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.

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 modeling 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 stationarity, 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.

Selected methods: Selected methods include time series plots and parametric and nonparametric 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 CSM 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 3.0).

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 a statistically independent sample data set.

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.

Selected methods: Box and whisker plots, Levene's test, ANOVA (and nonparametric equivalents), cluster analysis and principal component analysis.

Spatial and temporal heterogeneity are common within groundwater systems and indicate 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 4.0 and 5.0.

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 with statistical confidence. If the equality of variance test holds, then a one-way ANOVA test can subsequently determine if there are statistically significant differences in constituent concentrations between two-sample data sets. The Kruskal-Wallis test is the nonparametric 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 3.0) with the corresponding downgradient sample population. If a representative background population is not present, then intrawell statistical comparisons (Section 3.0) 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 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), 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.

Selected methods: Quantile-Quantile (Q-Q) plots, box and whisker plots, summary statistics, and 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 example, in normal Q-Q plots the theoretical normal distribution is linear. If the sampled data distribution is normal, then it will conform to a linear shape comparable to that of the theoretical normal 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 2.7) 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 (e.g., 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 2.7). 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 2.7) in the sampled data.

Summary statistics and 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 predefined 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 (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 fifty 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.

2.7 OUTLIER TESTS

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

Selected methods: Dixon's or Rosner's 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 INTERWELL VERSUS INTRAWELL COMPARISONS

The Cholla groundwater monitoring systems include designated background monitoring wells to perform interwell comparisons. An interwell comparison is one where samples collected from two different geographic locations within the same water-bearing unit are used to perform the statistical evaluation. In this sense, an interwell comparison is one between two or more monitoring wells. The first geographic location is generally upgradient of the CCR unit and represents groundwater conditions native to those directly beneath the CCR unit (e.g., “background”). The second location represents groundwater conditions downgradient of the CCR unit. Multiple wells can exist at each background and downgradient geographic location. Interwell comparisons perform poorly in cases where an adequate and representative background well location cannot be established. Factors leading to inadequate or non-representative background can include, for example, spatial heterogeneity in groundwater conditions beneath the CCR unit; discontinuous lithologies upgradient versus downgradient of the CCR unit; and/or dry conditions upgradient or adjacent to the CCR unit. These inadequacies can cause an interwell statistical comparison to be meaningless and result in false positive or false negative statistical results. Faced with these background inadequacies, intrawell comparisons are an industry accepted and recommended alternative to interwell comparisons (USEPA, 2009).

An intrawell comparison is one where background geographic locations cannot be adequately established for a CCR unit samples derive from the same geographic location (e.g., downgradient monitoring well) but the samples are divided into different time periods to perform the statistical evaluation. The premise for intrawell testing is to establish a baseline constituent concentration within the monitoring well(s) from which future samples from the same monitoring well(s) will be compared to determine if groundwater conditions are changing (i.e., degrading or improving or holding constant). Intrawell comparisons are an alternative when it is impossible to establish an adequate and representative background location for a CCR Unit. 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 unit 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 (increasing) trends in constituent concentrations within the well over time.

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

In instances where the statistical evaluation declares SSIs over background, and background is questionable in terms of overall adequacy and representativeness, alternative source demonstration efforts should investigate if the background well designation is the cause of the SSI. If so, alternative background locations or intrawell statistical methods should be considered.

4 DETECTION MONITORING

As indicated in the Introduction of this report, previous documentation has been prepared for the site that identifies how detection monitoring data are to be evaluated (Montgomery & Associates, 2017a). *Statistical Analysis of Baseline Groundwater Monitoring Data* (Montgomery & Associates, 2018) presents an initial evaluation of Appendix III constituent data collected during the baseline detection monitoring period (November 2015 through September 2017) based on this documentation.

Since detection monitoring will continue for CCR units that did not transition into Assessment Monitoring as a result of the initial evaluation of Appendix III constituent data, this Section supplements the discussion of statistical methods previously prepared for the site (Montgomery & Associates, 2017a) and further identifies statistical tests to assess if there is an SSI over background levels based on collected detection monitoring data going forward.

4.1 DATA EVALUATION OBJECTIVES

The objective of a detection monitoring program is to evaluate each of the CCR Rule Appendix III constituents (Section 4.2) 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.

4.2 CONSTITUENTS OF CONCERN

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

4.3 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) 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 have been identified 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 3.0 discusses alternative industry standard procedures for performing groundwater statistical evaluations when background proves inadequate.

4.3.1 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 ($1-\alpha$) 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 for identifying an exceedance (Section 4.4) (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, nonparametric, 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 statistically significant increase. 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 a 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 $\varkappa$-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 4.4). 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 is less than 15%.

Nonparametric 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 nonparametric UPLs the upper limit reflects generally 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 eight 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).

4.3.2 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.

4.4 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, sitewide 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, nonparametric, 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 (USEPA, 2009).

— **Sitewide False Positive Rate**: The sitewide false positive rate (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.

4.5 SSI DECLARATION – DETECTION MONITORING

If the detection monitoring statistical evaluation indicates there is an SSI 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 5.0).

5 ASSESSMENT MONITORING

The CCR Rule states that a CCR unit must begin assessment monitoring 90 days following a declaration of an SSI 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, 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 Appendix IV groundwater constituents show an SSL over respective GWPSs.

5.1 DATA EVALUATION OBJECTIVES

The objective of an assessment monitoring program is to evaluate if an Appendix IV constituent downgradient of the CCR unit exhibits an SSL over the respective GWPS, pursuant to 40 CFR §257.95(g).

5.2 CONSTITUENTS OF CONCERN

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

5.3 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 conceptual site model.

There are two approaches to evaluating whether groundwater data comply with GWPSs. 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.

5.3.1 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 protection standard 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 SSL. The confidence interval calculation requires a t value at a specified confidence (e.g., 1-α, 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, which contains a predefined proportion of the underlying statistical population. Most often this predefined coverage is equal to 95% (e.g., the 95% UTL). 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 $\gamma = 95\%$) and selected confidence (recommended confidence is 95%, or $\alpha = 0.05$). 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 SSL 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.

5.3.2 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 (USEPA, 2009) 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 SSL and possibly justify corrective action. The prediction limits around the mean or median will include a retesting strategy (Section 4.3.1 and Section 4.4) and be calculated in accordance with the EDA procedures (Section 2.0) and recommendations put forth in the USEPA's Unified Guidance.

5.4 PERFORMANCE STANDARDS

The performance standards in Section 4.4 are applicable to assessment monitoring and will follow method-specific recommendations put forth in the USEPA's Unified Guidance.

5.5 SSL DECLARATION – ASSESSMENT 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 an SSL above the GWPS established under 40 CFR §257.95(h).

If assessment monitoring demonstrates an SSL over the respective GWPS for any Appendix IV constituent(s), an investigation should ensue to determine if a release from the CCR unit is the cause of the SSL(s) (See Section 4.5), and the owner or operator must follow criteria set forth under 40 CFR §257.95(g).

5.6 RETURN TO DETECTION 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.

6 CORRECTIVE ACTION

Within 90 days of declaring a statistically significant increase of an Appendix IV constituent over the respective GWPS because of a leaking CCR Unit, or noticing a leak from the CCR Unit, the owner or operator of the CCR unit must initiate an Assessment of Corrective Measures, per 40 CFR §257.96 (an extension of 60 days is allowable if warranted). The results of the Assessment of Corrective Measures must be discussed in a public meeting at least 30 days prior to selection of a remedy. Selection of a remedy must consider criteria set forth under 40 CFR §257.97, including the attainment of GWPSs (40 CFR §257.95(h)).

Remedy implementation (40 CFR §257. 98) begins within 90 days of remedy selection. The groundwater monitoring program continues during corrective action implementation to help, in part, determine remedy effectiveness. Pursuant to 40 CFR §257.98(a)(1), the corrective action groundwater monitoring program must:

1. meet requirements of an assessment monitoring program (40 CFR §257.95) (Section 5.0),
2. document the effectiveness of the selected remedy, and
3. demonstrate compliance with GWPSs (Section 5.2).

Pursuant to 40 CFR §257.98(c), a remedy is complete when Appendix IV constituent concentrations are complaint with GWPSs within the spatial extent of the contamination plume for a period of three consecutive years using the statistical procedures and performance standards in 40 CFR §257.93(f) and (g).

Section 5.3.1 discusses the singe-sample comparison test for assessing Appendix IV constituents, which is a reversible test; meaning the same fixed protection standard (i.e., GWPS) can be used for assessment monitoring and corrective action. In either event (assessment monitoring or corrective action), the confidence interval is calculated for downgradient compliance wells and compared to the respective fixed protection standard. For assessment monitoring, if the lower confidence limit of the compliance well data (i.e., lower bound of the confidence interval) exceeds the fixed protection standard, an SSL declaration is made. For corrective action, if the upper confidence limit (i.e., upper bound of the confidence interval) of the compliance well data falls below the fixed protection standard, the remedy is considered effective and complete. It is important to remember the confidence interval reflects the central tendency of the compliance well data (e.g., mean or median value) and, therefore, a single-sample event will likely not transition the CCR unit in or out of compliance. As such, the CCR owner or operator should expect several rounds of sampling to decipher the adequacy of remedy implementation and its overall effectiveness.

7 CLOSURE AND POST-CLOSURE CARE

Closure of CCR units is generally separate from the groundwater monitoring program but there are requirements in the CCR Rule that involve statistically evaluating groundwater data to support assessing whether closure is complete. Requirements generally vary by the closure approach the owner/operator selects for a particular unit. Closure of CCR units can be conducted by either: leaving CCR waste in place (i.e., closure in place) or by removing the CCR waste placed in the unit (i.e., closure by removal). Post-closure care applies to units that close by leaving CCR waste in place.

Pursuant to 40 CFR §257.102(c), closure by removal is when the owner or operator removes and decontaminates all areas affected by releases from the CCR unit. Closure is considered complete when CCR has been removed from all affected areas and groundwater monitoring concentrations comply with GWPSs for Appendix IV constituents. The Federal Register (2015) indicates that in evaluating the completeness of closure by removal, all Appendix IV concentrations should fall below their respective GWPSs for two consecutive sampling events using the statistical procedures in 40 CFR §257.95(g). Further, the statistical methods should follow those in Section 5.3 to determine whether groundwater complies with GWPSs. If a CCR unit is in detection monitoring at the time closure by removal occurs, the evaluation of completeness for closure by removal involves the development of GWPSs for the subject unit as well as assessment of whether the groundwater complies with GWPSs. A total of eight (8) sampling events are generally required for the statistical analysis outlined in this report but the number of independent sampling events necessary for evaluation is data dependent and must be assessed with collected data.

Closure by leaving CCR in place is subject to post-closure requirements put forth under 40 CFR §257.104. This includes maintaining the groundwater monitoring system and monitoring groundwater (40 CFR §257.104(b)(3)). The post-closure period spans 30 years (40 CFR §257.104(c)) unless the CCR unit is under assessment monitoring at the 30-year mark, in which case assessment monitoring will continue until the CCR unit returns to detection monitoring pursuant to 40 CFR §257.95 (Section 5.0). If the CCR unit is in corrective action, the CCR unit will need to achieve corrective action completion status (see Section 6.0) and reach detection monitoring status as put forth in this Section. If the CCR unit is in detection monitoring at the end of 30 years, the CCR unit owner or operator can cease groundwater monitoring.

8 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.

9 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 2.4) and performing sampling optimization.

Other public domain software packages, including R (version 3.5.1) and Spatial Analysis in Macroecology (version 4.0), are defensible and transparent spatial regression and data detrending (Section 2.5.1) software platforms. These software platforms will supplement ProUCL and VSP, as necessary.

Isatis (Geovariances, France) (version 2015) is a well-established geostatistical software platform. This software will be used to validate variogram models and spatial interpolation methods (Section 2.4), as necessary.

BIBLIOGRAPHY

- AMEC Earth & Infrastructure, Inc. (AMEC), 2012. Well Completion Report, Installation of Aquifer Protection Permit Monitor Wells, Arizona Public Service Company, Cholla Power Plant, Navajo County, Arizona. AMEC Job No. 17-2011-4054. May 7, 2012.
- El Kadi, A., 1995. Groundwater Models for Resources Analysis and Management. London, CRC Press, Inc.
- Federal Register, 2020. *40 Code of Federal Regulations Part 257 – Hazardous and Solid Waste Management System; Disposal of Coal Combustion Residuals From Electric Utilities*; Final Rule promulgated at 80 FR 21468 on April 17, 2015 with amendments issued through November 12, 2020.
- Montgomery & Associates, 2011. *Arizona Public Service Cholla Power Plant Point of Compliance Evaluation*. Prepared for APS. January 26, 2011.
- Montgomery & Associates, 2017a. *Cholla Power Plant Coal Combustion Residuals Program – Statistical Method Selected for Evaluation of Groundwater Monitoring Data, Navajo County, Arizona. Document # CH_GW_StatCert_020_20170919.* Prepared for APS. September 19, 2017.
- Montgomery & Associates, 2017b. *Cholla Power Plant Coal Combustion Residuals Program – Design, Installation, and Evaluation of Completeness of Groundwater Monitoring Networks, Navajo County, Arizona. Document # CH_GW_SystemCert_020_20170919.* Prepared for APS. September 19, 2017.
- Montgomery & Associates, 2018. *Cholla Power Plant Coal Combustion Residuals Program – Statistical Analysis of Baseline Groundwater Monitoring Data. November 2015 through September 2017. Navajo County, Arizona.* Prepared for APS. January 12, 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.
- 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, 2018a. *Statistical Data Analysis Work Plan.* Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Cholla Power Plant, Navajo County, Arizona. Dated October 15, 2018.
- Wood, 2018b. *CCR Groundwater Assessment Monitoring Statistical Analysis and Results for the Fly Ash Pond.* Arizona Public Service Cholla Power Plant. Navajo County, Arizona. Dated October 15, 2018.
- Wood, 2018c. *CCR Groundwater Assessment Monitoring Statistical Analysis and Results for the Bottom Ash Pond.* Arizona Public Service Cholla Power Plant. Navajo County, Arizona. Dated October 15, 2018.
- Wood, 2019. *Alternative Source Demonstration for Lithium at the BAP*. Arizona Public Service Cholla Power Plant. Navajo County, Arizona. Dated June 6, 2019.
- Wood, 2021. *Evaluation of Groundwater Redox Conditions at the FAP*. Arizona Public Service Cholla Power Plant. Navajo County, Arizona. Dated January 22, 2021.

- Wood, 2020a. *Hydrogeologic Investigation of the Fly Ash Pond and Bottom Ash Pond.* Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Arizona Public Service Company Cholla Power Plant. Navajo County, Arizona. Dated January 31, 2020.
- Wood, 2020b. *Alternative Source Demonstration for Arsenic and Cobalt at the FAP.* Public Service Cholla Power Plant. Navajo County, Arizona. Dated January 31, 2020.
- Wood, 2022a. *Groundwater Monitoring System Certification Report*. Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Cholla Power Plant, Navajo County, Arizona. Date pending.
- Wood, 2022b. *Groundwater Sampling and Analysis Program*. Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Cholla Power Plant, Navajo County, Arizona. Dated April 11, 2022.
- Wood, 2022c. *Well Completion Report, Abandonment and Replacement Well Program.* Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Cholla Power Plant, Navajo County, Arizona. Dated January 18, 2022.
- Wood, 2022d. *Well Completion Report, Bottom Ash Pond Pre-Design Studies.* Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Cholla Power Plant. Joseph City, Arizona. Date pending.
- Wood, 2022e. *Annual Groundwater Monitoring and Corrective Action Report for 2021.* Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Arizona Public Service Company Cholla Power Plant. Navajo County, Arizona. Dated January 31, 2021.
- WSP USA (WSP), 2023. Groundwater Monitoring System Certification Report. Coal Combustion Residuals Rule Groundwater Monitoring System Compliance. Arizona Public Service Company Cholla Power Plant. Navajo County, Arizona. January 30, 2023.

TABLES

Table 1-1
Description of Cholla Coal Combustion Residual Units

CCR Unit	Function	Operation	Size/Construction	History
Fly Ash Pond (FAP)	*Single CCR unit* - surface impoundment to store slurried fly ash from the plant.	Receives a slurry from the plant that contains primarily fly ash but may also contain some bottom ash, boiler slag, flue gas emission control residuals, boiler cleaning waste, oil/water separator solids, and storm water. Periodically receives solids from the SEDI.	- 430 acres in aerial extent. - Total storage capacity of about 18,000 acre-feet. - Normal operating pool elevation of 5,114 feet amsl.	- Constructed beginning in 1976 and placed into service in 1978. - Unlined; constructed on Moenkopi bedrock and a thin veneer of alluvial sediments. - The dam is constructed of earth fill with a central clay core that extends to bedrock where bedrock is shallow. In the central portion of the dam, where bedrock is deeper, a slurry cutoff wall extends one foot into bedrock or two feet into stiff clay.
Sedimentation Pond (SEDI)	*Single CCR unit* - collects water from drains around plant site, including storm water, process water, plant water, and slurry from plant leaks.	Collects discharge from on-site secondary wastewater treatment plant, effluent from the oil/water separator, vehicle wash water, plant wash water, and FGD wastes from scrubber or scrubber feed tank upsets. Water collected in the SEDI is pumped to Cholla's general water sump for recycling as process water.	- 1.3 acres in aerial extent. - Total storage capacity of 10.5 acre-feet. - Maximum pond depth of 10 feet. - the top of the pond side slope is at 5,019 feet amsl	- Placed into service in 1976. - Lined with a 2-foot-thick layer of compacted clay. - Constructed below grade. - Discharges to the SEDI ceased as of October 2020. - Demolition and excavation of SEDI complete as of October 2021.
Bottom Ash Pond (BAP)	*Single CCR unit* - surface impoundment to store slurried bottom ash from the plant.	Bottom ash is pumped to the BAP as a slurry. The bottom ash settles in the east and west upstream storage cells and the water is decanted to the reservoir and ultimately siphoned back to the plant for reuse. Slurry may also contain fly ash, boiler slag, flue gas emission control residuals, sedimentation pond effluent, cooling tower blowdown, oil/water separator effluent and solids, boiler cleaning waste, and storm water. Periodically receives solids from the SEDI.	- 105 acres in aerial extent. - Total storage capacity of 2,300 acre-feet. - Normal operating pool elevation of 5,117.8 feet amsl.	- Constructed beginning in 1976 and placed into service in 1978. - Unlined; constructed on Moenkopi bedrock and Tanner Wash alluvium. - Consists of a reservoir directly behind the dam and two storage cells upstream of the reservoir. - The dam is constructed of earth fill with a central clay core that extends to bedrock where bedrock is shallow. In the central portion of the dam, where bedrock is deeper, a slurry cutoff wall extends one foot into bedrock or two feet into stiff clay.
Bottom Ash Monofill (BAM)	*Single CCR unit* - landfill for bottom ash solids excavated from the BAP.	Bottom ash that has been drained of water is excavated from the BAP and permanently stored in the BAM.	- 41 acres in aerial extent.	- Placed into service in 1999.

Source:

GEI Consultants, Inc. 2009. *Final Coal Ash Impoundment Specific Site Assessment Report, Arizona Public Service, Cholla Power Plant.* Submitted to Lockheed-Martin Corporation. December 2009.

Abbreviations:

amsl - above mean sea level
BAP - Bottom Ash Pond
BAM - Bottom Ash Monofill
CCR - Coal combustion residuals
FAP - Fly Ash Pond
FGD - flue gas deulfurization
SEDI - Sedimentation Pond

Table 1-2
CCR Groundwater Monitoring System Summary

Well	ADWR Well Index Number	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]
M-54	55-918646	BAM	Background	Coconino Sandstone	10/2/2015	370	5070.71	5068.21	315	365	50	4,753.21	4,703.21	4,698.21
M-59	55-918647	BAM	Downgradient Boundary	Coconino Sandstone	10/21/2015	425	5136.00	5133.86	373	423	50	4,760.86	4,710.86	4,708.86
M-60	55-918649	BAM	Downgradient Boundary	Coconino Sandstone	11/1/2015	450	5151.18	5148.69	395	445	50	4,753.69	4,703.69	4,698.69
M-61	55-918648	BAM	Downgradient Boundary	Coconino Sandstone	11/13/2015	420	5127.58	5124.95	365	415	50	4,759.95	4,709.95	4,704.95
M-64A	55-920353	FAP/BAP	Background	LCR Alluvium	2/9/2017	69	4991.90	4988.90	30	60	30	4,958.90	4,928.90	4,919.90
M-52A	55-918657	BAP	Downgradient Boundary	Tanner Wash Alluvium/Moenkopi - Moqui Member	9/22/2015	83	5049.36	5047.08	20	70	50	5,027.08	4,977.08	4,964.08
M-53A	55-918651	BAP	Downgradient Boundary	Tanner Wash Alluvium	9/22/2015	38	5044.68	5042.09	10	35	25	5,032.09	5,007.09	5,004.09
MW-71A	55-926812	BAP	Downgradient Boundary	Tanner Wash Alluvium	9/28/2021	30	5050.680	5050.15	15.0	25.0	10	5,035.15	5,025.15	5,020.15
MW-72M	55-926814	BAP	Downgradient Boundary	Moenkopi Formation - Moqui Member	9/19/2021	125	5049.670	5050.54	59.0	69.0	10	4,991.54	4,981.54	4,925.54
MW-73A	55-926813	BAP	Downgradient Boundary	Tanner Wash Alluvium	9/23/2021	26	5049.190	5049.66	11.0	21.0	10	5,038.66	5,028.66	5,023.66
MW-74M	55-926815	BAP	Downgradient Boundary	Moenkopi Formation - Moqui Member	9/23/2021	125	5049.070	5049.74	45.0	65.0	20	5,004.74	4,984.74	4,924.74
MW-76A	55-926103	BAP	Downgradient	Tanner Wash Alluvium	5/1/2021	52	5033.890	5032.24	16.0	26.0	10	5,016.24	5,006.24	4,980.24
MW-77A	55-926104	BAP	Downgradient	Tanner Wash Alluvium/Moenkopi Formation - Moqui Member	4/30/2021	65	5031.020	5029.78	44.0	64.0	20	4,985.78	4,965.78	4,964.78
MW-78A	55-926105	BAP	Downgradient	Tanner Wash Alluvium	5/4/2021	107	5036.950	5035.05	66.0	96.0	30	4,969.05	4,939.05	4,928.05
MW-79A	55-926106	BAP	Downgradient	Tanner Wash Alluvium	5/17/2021	177	5040.890	5038.24	135.0	165.0	30	4,903.24	4,873.24	4,861.24
W-305	55-906364	BAP	Downgradient Boundary	Tanner Wash Alluvium	10/7/1983	102	5046.80	5044.65	80	100	20	4,964.65	4,944.65	4,942.65
W-306	55-506365	BAP	Downgradient Boundary	Tanner Wash Alluvium	10/11/1983	52	5046.74	5044.78	30	50	20	5,014.78	4,994.78	4,992.78
W-314	55-533814	BAP	Downgradient Boundary	Tanner Wash Alluvium	1/27/1992	63	5051.10	5051.32	41	61	20	5,010.32	4,990.32	4,988.32
M-50A	55-918641	FAP	Downgradient Boundary	LCR Alluvium	9/18/2015	32	5038.18	5035.65	9	29	20	5,026.65	5,006.65	5,003.65
M-51A	55-918640	FAP	Downgradient Boundary	LCR Alluvium	9/19/2015	14	5041.77	5039.10	7	12	5	5,032.10	5,027.10	5,025.10
MW-65A	55-922299	FAP	Downgradient	LCR Alluvium	11/15/2018	25	5027.86	5026.21	9	19	10	5,017.31	5,007.31	5,001.21
MW-66A	55-922300	FAP	Downgradient	LCR Alluvium	11/14/2018	60	5033.35	5032.46	24	49	25.1	5,008.86	4,983.76	4,972.46
MW-67A	55-922301	FAP	Downgradient	LCR Alluvium	11/16/2018	50	5025.38	5024.05	15	45	30.1	5,009.45	4,979.35	4,974.05
W-123*	55-506587	FAP	Downgradient Boundary	Moenkopi Formation - Moqui Member	11/4/1983	40	5039.838	5038.136	14	29	15	5,024.14	5,009.14	4,998.14
W-123R	55-926116	FAP	Downgradient Boundary	LCR Alluvium	5/12/2021	67	5041.06	5038.83	35	65	30	5,003.83	4,973.83	4,971.83
M-56A	55-918661	SEDI	Downgradient Boundary	LCR Alluvium	10/7/2015	100	5023.17	5020.63	40	85	45	4,980.63	4,935.63	4,920.63
M-57A	55-918660	SEDI	Downgradient Boundary	LCR Alluvium	10/8/2015	100	5023.82	5021.16	40	85	45	4,981.16	4,936.16	4,921.16
M-58A	55-918659	SEDI	Downgradient Boundary	LCR Alluvium	10/13/2015	100	5023.84	5021.24	39	84	45	4,982.24	4,937.24	4,921.24
M-62A	55-918658	SEDI	Background	LCR Alluvium	11/17/2015	97	5020.87	5021.01	39	84	45	4,982.01	4,937.01	4,924.01

Source:

AMEC Earth & Infrastructure, Inc., 2012. *Well Completion Report, Installation of Aquifer Protection Permit Monitor Wells, Arizona Public Service Company, Cholla Power Plant, Navajo County, Arizona.* AMEC Job No. 17-2011-4054. May 7, 2012.

Montgomery & Associates, 2017. Cholla Power Plant Coal Combustion Residuals Program–Design, Installation, and Evaluation of Completeness of Groundwater Monitoring Networks. Navajo County, Arizona. September 19, 2017.

Wood Environment & Infrastructure Solutions, Inc. Surveying, 2018 and 2019.

Martin Land Surveys, Inc., 2020 and 2021.

Notes and Abbreviations:

ADWR - Arizona Department of Water Resources
AMSL - Above mean sea level (Vertical datum is NAVD 88)
BAM - Bottom Ash Monofill
BAP - Bottom Ash Pond
bgs - below ground surface
CCR - Coal combustion residuals
FAP - Fly Ash Pond
ft - feet
LCR - Little Colorado River
NA - Not Available
SEDI - Sedimentation Pond
* Abandoned well
**Approximate - elevation based on measured stickups

FIGURE

0 0.75 1.5 3
Miles

Job No.	14-2022-2007
PM:	MBH
Date:	3/28/2022
Scale:	1" = 1.5 miles

N

The map shown here has been created with all due and reasonable care and is strictly for use with Wood Environment & Infrastructure Solutions, Inc. Project Number 14-2022-2007. This map has not been certified by a licensed land surveyor, and any third party use of this map comes without warranties of any kind. Wood Environment & Infrastructure Solutions, Inc. assumes no liability, direct or indirect, whatsoever for any such third party or unintended use.

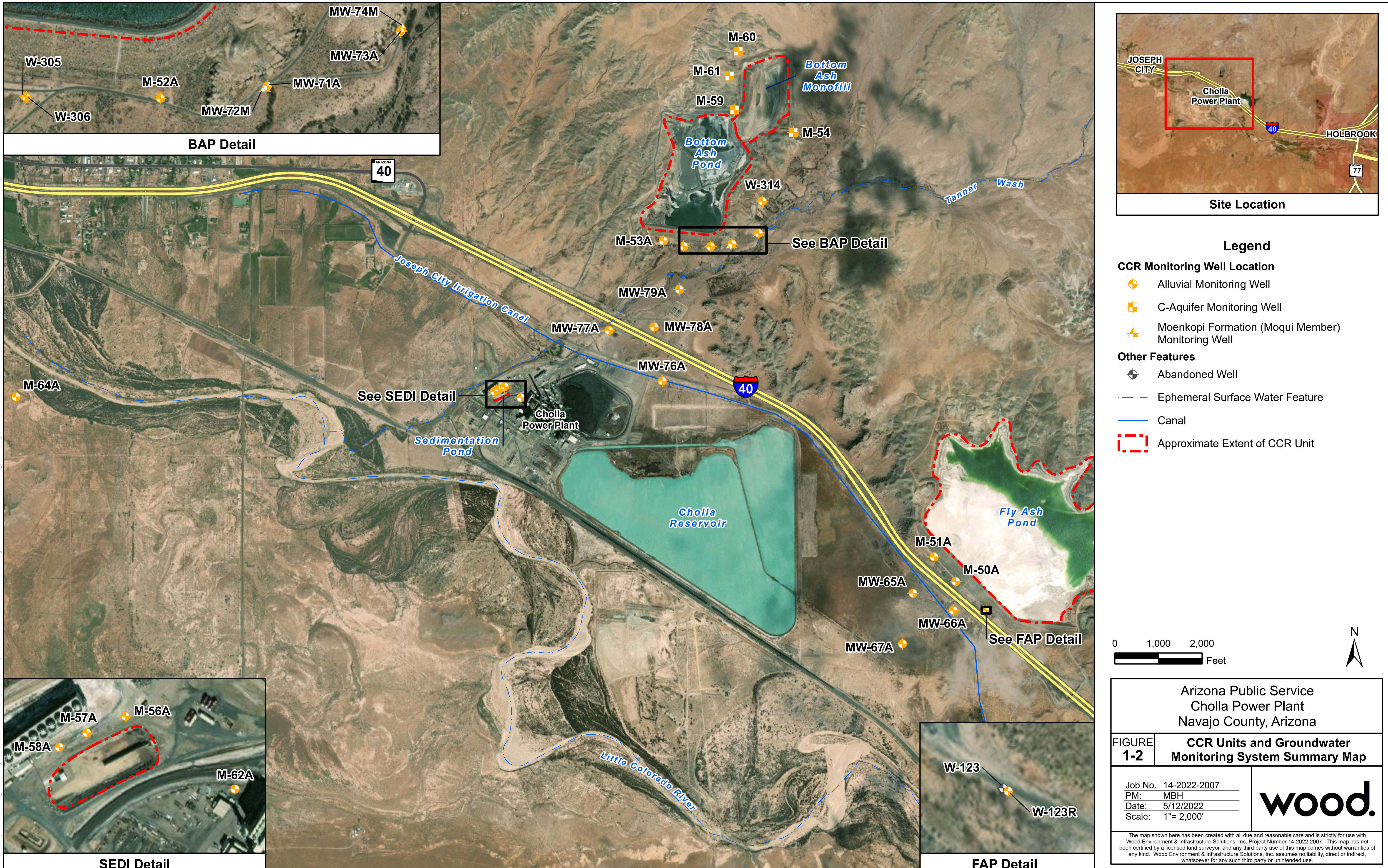
Arizona Public Service
Cholla Power Plant
Navajo County, Arizona

Site Location Map

FIGURE
1-1

wood.

Path: X:\Projects\20-Longterm Projects\APS Cholla Compliance Support\MXD\Sample and Analysis Plan\2022\Sample and Analysis Plan\Sample and Analysis Plan.aprx



Arizona Public Service
Cholla Power Plant
Navajo County, Arizona

FIGURE 1-2 CCR Units and Groundwater Monitoring System Summary Map

Job No. 14-2022-2007
PM: MBH
Date: 5/12/2022
Scale: 1"= 2,000'

wood.

The map shown here has been created with all due and reasonable care and is strictly for use with Wood Environment & Infrastructure Solutions, Inc. Project Number 14-2022-2007. This map has not been certified by a licensed land surveyor, and any third party use of this map comes without warranties of any kind. Wood Environment & Infrastructure Solutions, Inc. assumes no liability, direct or indirect, whatsoever for any such third party or unintended use.

Path: X:\Projects\20-Longterm Projects\APS Cholla Compliance Support\MXD\Statistical Data Analysis Work Plan\SDAWP Figures May 2022.aprx