



PNNL-ACT-10120

# 2021 Hydrological/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana

May 2022

CJ Thompson  
E Koehler  
PJ Chamberlain

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY  
*operated by*  
BATTELLE  
*for the*  
UNITED STATES DEPARTMENT OF ENERGY  
*under Contract DE-AC05-76RL01830*

Printed in the United States of America

Available to DOE and DOE contractors from the  
Office of Scientific and Technical Information,  
P.O. Box 62, Oak Ridge, TN 37831-0062;  
ph: (865) 576-8401  
fax: (865) 576-5728  
email: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)

Available to the public from the National Technical Information Service  
5301 Shawnee Rd., Alexandria, VA 22312  
ph: (800) 553-NTIS (6847)  
email: [orders@ntis.gov](mailto:orders@ntis.gov) <<https://www.ntis.gov/about>>  
Online ordering: <http://www.ntis.gov>

# 2021 Hydrological/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana

May 2022

CJ Thompson<sup>1</sup>  
E Koehler<sup>2</sup>  
PJ Chamberlain<sup>3</sup>

Prepared for  
Talen Montana, LLC, Colstrip, Montana, and Genie Land Company,  
Colstrip, Montana, under an Agreement for Commercializing Technology  
with the U.S. Department of Energy, Contract DE-AC06-76RLO1830

Pacific Northwest National Laboratory  
Richland, Washington 99354

---

<sup>1</sup> Pacific Northwest National Laboratory, Richland, Washington

<sup>2</sup> KC Harvey Environmental, LLC, Bozeman, Montana

<sup>3</sup> Enviro-Sci Consulting, Big Lake, Alaska



## Executive Summary

Pacific Northwest National Laboratory/Battelle–Pacific Northwest Division and KC Harvey Environmental, LLC personnel have monitored and collected water-quality data in the drainage basins of Cow Creek and Pony Creek in Rosebud County, Montana, since 1984. This work is performed for the signatories of Stipulation 12(d) under contract with Talen Montana (formerly PPL Montana LLC), which is located in Colstrip, Montana. This report presents the collection, analysis, and interpretation of the water-quality and hydrologic data obtained during 2021.

The objective of the 2021 water monitoring study was to determine whether there were changes in water-quality and hydrological parameters downgradient from the effluent holding pond (EHP) east of Colstrip Power Plant Units 3 and 4 on Genie Land Company property since the initial 1984–1986 baseline study was completed. A further objective was to establish the current characteristics and conditions of water-quality and hydrological parameters so the effects of any ongoing or future events associated with the EHP and related operations could be evaluated. Specific conductance and boron and sulfate concentrations have been monitored since 1984 as indicators of leakage from the EHP. In 2010, bromide was added to the list of monitored constituents as an additional, unique indicator for detecting potential impacts from EHP operations.

There were no reportable releases from the EHP during 2021.

Groundwater levels measured in the alluvial wells in 2021 averaged about 0.93 ft lower than those in the previous year, consistent with the dry weather pattern in 2021. Between the June and October sampling campaigns, water levels declined in all monitored wells. The 2021 groundwater levels do not indicate any major changes in groundwater flow patterns for Cow Creek and South Fork Cow Creek.

Statistical analyses of conductivity, sulfate, and boron data identified seven site-level trends where water quality had changed from previous years; no drainage-level trends were detected. At the site level, Cow Creek had increasing trends for conductivity and boron at PW 735, which taps a deep confined aquifer. Genie spring (GSP) 2 had an increasing trend for boron. Pony Creek had an increasing trend for conductivity and a decreasing trend for boron at Genie old well (GOW) 4, an increasing trend for conductivity at GSP 4, and an increasing trend for boron at GSP 8. These findings do not indicate any apparent impacts on groundwater quality from the operations of the EHP.

Using data from the entire observation period (1984–1987 and 1990–2021) and based on the selected sampling sites, the Pony Creek drainage was determined to have the lowest mean levels of conductivity, sulfate, and boron. Water from the Cow Creek drainage was distinctly highest in conductivity and sulfate, and the South Fork Cow Creek drainage means fell between the values from the other drainages except for boron, for which the mean was identical to that of Cow Creek.

Bromide was formally added as a monitoring constituent in 2010 in response to Talen Montana's recent addition of bromide to the flue gas scrubber solution. Bromide data from the 2021 sampling events provided no evidence of impingement of water from the EHP in the study area.

Continuous-monitoring probe measurements in private monitoring well 736 showed relatively stable conductivity and an early spring increase in water level followed by a gradual decline in groundwater elevation through summer. The probe data do not indicate any releases from the EHP during 2021.



## Acronyms and Abbreviations

DTW	depth to water
EHP	effluent holding pond
ERA	Environmental Resources Associates
ft	foot (feet)
GAS	Genie alternative supply
GEL	General Engineering Laboratories
GKW	quality control sample
GNW	Genie new well
GOW	Genie old well
gpd/ft <sup>2</sup>	gallons per day per square foot
GSP	Genie spring
GSW	Genie surface water
in.	inch(es)
µm	micron(s)
LOWESS	locally weighted scatterplot smoothing
MAD	median absolute deviation
MDL	method detection limit
mg/L	milligram(s) per liter
mL	milliliter(s)
MPV	most probable value
MW	megawatt(s)
NSI	NSI Lab Solutions
OD	outside diameter
PLR	Predictive Likelihoods Ratio
PNNL	Pacific Northwest National Laboratory
PNWD	Battelle–Pacific Northwest Division
ppm	parts per million
PVC	polyvinyl chloride
PW	private monitoring well
QA	quality assurance
QC	quality control
RPD	relative percent difference
µS/cm	microsiemens per centimeter
TDS	total dissolved solids
TOC	top of well casing
W-1	Montana State University-installed well



# Contents

Executive Summary .....	iii
Acronyms and Abbreviations .....	v
1.0 Introduction.....	1.1
1.1 Study Area .....	1.2
1.2 Study Purpose and Scope.....	1.3
1.3 Report Contents and Organization.....	1.4
2.0 Hydrological Evaluation .....	2.1
2.1 Groundwater Flow System .....	2.2
2.2 Cow and Pony Creek Monitoring Network .....	2.9
2.3 Results and Discussion .....	2.10
3.0 Water Quality in the Cow Creek and Pony Creek Drainages .....	3.1
3.1 Water Sampling .....	3.1
3.2 Sampling Protocols.....	3.1
3.2.1 Alluvial Wells.....	3.1
3.2.2 Spring and Surface Water.....	3.3
3.2.3 Sample Handling, Quality Control Samples, and Field Measurements .....	3.4
3.2.4 Sample Custody.....	3.5
3.3 Sample Analysis .....	3.6
3.4 Quality Control for Water Samples .....	3.6
3.5 Statistical Analysis of Chemical Data.....	3.10
3.5.1 Objective 1: Identify Any Time Trends at the Site or Drainage Level in the Parameters Conductivity, Sulfate, and Boron.....	3.11
3.5.2 Objective 2: Estimate and Compare Overall (Time-Averaged, Site-Averaged) Drainage-Mean Values for Cow Creek, South Fork Cow Creek, and Pony Creek .....	3.11
3.5.3 Sites and Data Used in the Statistical Analyses.....	3.11
3.5.4 Testing Methods .....	3.13
3.5.5 Results of Statistical Analysis .....	3.14
3.5.6 Overall Conclusions of the Statistical Analysis.....	3.15
3.6 Results and Discussion .....	3.16
3.6.1 Sampling Data from the Genie Alternative Supply Wells.....	3.17
3.6.2 Sampling Data from Pony Creek Drainage .....	3.17
3.6.3 Sampling Data from South Fork Cow Creek.....	3.23
3.6.4 Sampling Data from Cow Creek .....	3.26
3.6.5 Comparison of Cow Creek, South Fork Cow Creek, and Pony Creek Drainages.....	3.31
3.6.6 Analysis of Bromide.....	3.33
3.6.7 Continuous-Monitoring Probe Results .....	3.34

4.0	Conclusions.....	4.1
5.0	References.....	5.1
	Appendix A Comments Provided by Signatories of Stipulation 12(d).....	A.1
	Appendix B Field Record Sheet and Chain-of-Custody Form Used for Sample Control .....	B.1
	Appendix C Analytical Methods Used for Water Analysis .....	C.1
	Appendix D Photos of Spring and Surface-Water Sites .....	D.1
	Appendix E Site and Drainage-Mean Values for Data Collected in 2021 .....	E.1
	Appendix F Analytical Results for Water Samples Collected in 2021 .....	F.1
	Appendix G Statistical Analysis – Methods and Results.....	G.1

# Figures

Figure 1.1.	Colstrip Units 1–4 (from right to left).....	1.1
Figure 1.2.	Location of study area (modified from Montana Department of State Lands and U.S. Office of Surface Mining 1983).....	1.3
Figure 1.3.	Map showing the location of the Colstrip Units 3 and 4 EHP near the headwaters of Cow Creek.....	1.5
Figure 2.1.	Annual precipitation at Colstrip, Montana, 1984–2021. Data were compiled from the National Oceanic and Atmospheric Administration (see footnote on previous page).....	2.2
Figure 2.2.	Monthly distribution of annual precipitation at Colstrip, Montana, 1984–2021. Data were compiled from the National Oceanic and Atmospheric Administration (see footnote on previous page).....	2.2
Figure 2.3.	Measuring DTW with a water-level meter.....	2.4
Figure 2.4.	Monitoring locations in the study area.....	2.7
Figure 2.5.	Water table in the alluvial aquifers of Pony Creek, Cow Creek, and South Fork Cow Creek.....	2.8
Figure 3.1.	Well sampling: (a) Lowering the submersible pump into well GNW 2. (b) Filling sample bottles. Water from the well is pumped through the sampling manifold, a 3 ft section of ¼-in. OD poly tubing, and an in-line filter before flowing into the sample containers. (c) Lowering a bailer into well W-1.....	3.3
Figure 3.2.	Measuring the flow rate at a spring.....	3.4
Figure 3.3.	Preparing an equipment blank sample. Deionized water was pumped from the container on the left through a sample filter and into the sample bottle. ....	3.5
Figure 3.4.	Plots and data for conductivity, sulfate, and boron in the Pony Creek Drainage during 2021. Values are single measurements from the June sampling event.....	3.18
Figure 3.5.	Trend plot for conductivity at GSP 4. Outliers identified in Appendix G, Section G1.2, were omitted from the graph. ....	3.20
Figure 3.6.	Trend plot for conductivity at GOW 4.....	3.20
Figure 3.7.	Trend plot for boron at GOW 4.....	3.20
Figure 3.8.	Trend plot for boron at GSP 8. Outliers identified in Appendix G, Section G1.2, were omitted from the graph. ....	3.21
Figure 3.9.	Trend plots for conductivity, sulfate, and boron in the Pony Creek Drainage, 1985–2021.....	3.22
Figure 3.10.	Plots and data for conductivity, sulfate, and boron in South Fork Cow Creek Drainage during 2021. Values are averages from the June and October sampling events. ....	3.24
Figure 3.11.	Trend plots for conductivity, sulfate, and boron in South Fork Cow Creek, 1984–2021.....	3.25
Figure 3.12.	Plots and data for conductivity, sulfate, and boron values for alluvial wells and springs in Cow Creek Drainage during 2021. Values shown are averages from the July and October sampling events.....	3.28
Figure 3.13.	Trend plot for conductivity at PW 735.....	3.29
Figure 3.14.	Trend plot for boron at PW 735.....	3.29

Figure 3.15.	Trend plot for boron at GSP 2.....	3.30
Figure 3.16.	Trend plots for conductivity, sulfate, and boron in Cow Creek Drainage, 1984–2021.....	3.31
Figure 3.17.	2021 Bromide concentrations from sampling locations within the Cow Creek, South Cow Creek, and Pony Creek Drainages. The sites are ordered according to their position in each drainage (upgradient to downgradient). ND = non-detected result; Dry = site was dry (no sample was collected).....	3.33
Figure 3.18.	Continuous-monitoring probe.....	3.34
Figure 3.19.	Well PW 736 with companion datalogger enclosure, solar panel, and fenced enclosure.....	3.35
Figure 3.20.	Groundwater temperature trend measured at PW 736 in 2021 and 2010–2021 (inset). ....	3.37
Figure 3.21.	Groundwater elevation trend at PW 736 during 2021 in association with monthly total precipitation (columns and secondary y-axis).....	3.38
Figure 3.22.	Groundwater elevation trend at PW 736 from 2010–2021.....	3.39
Figure 3.23.	In situ PW 736 specific conductance trend measured by the datalogger over the 2021 monitoring year compared to field and laboratory specific conductance measured during sampling events. ....	3.39
Figure 3.24.	Specific conductance for PW 736 over the entire period of record 2010–2021.....	3.40

## Tables

Table 2.1.	2021 groundwater-level data (feet). .....	2.5
Table 2.2.	2021 Status of springs and surface-water sites.....	2.6
Table 3.1.	Sites sampled during 2021 and sample acquisition methods. <sup>(a)</sup> .....	3.2
Table 3.2.	Analytical detection and reporting limits. ....	3.6
Table 3.3.	Analytical results for QC samples from the June 2021 sampling trip.....	3.8
Table 3.4.	Analytical results for QC samples from the October 2021 sampling trip. ....	3.9
Table 3.5.	Wells and springs used in the statistical analysis. ....	3.12
Table 3.6.	Time trends detected in 2021. <sup>(a)</sup> .....	3.15
Table 3.7.	95% confidence limits on drainage-mean values using all data. <sup>(a)</sup> .....	3.15
Table 3.8.	Composition of EHP solution (based on a sample collected in May 2015).....	3.16
Table 3.9.	Values for conductivity, sulfate, and boron in samples collected from GAS wells in 2021.....	3.17
Table 3.10.	Drainage-mean values for conductivity, sulfate, and boron in Pony Creek Drainage, 1985–2021.....	3.23
Table 3.11.	Drainage-mean values for conductivity, sulfate, and boron in South Fork Cow Creek, 1984–2021. ....	3.26
Table 3.12.	2021 analysis results from the PW wells. ....	3.30
Table 3.13.	Drainage-mean values for conductivity, sulfate, and boron in Cow Creek drainage, 1984–2021.....	3.32



## 1.0 Introduction

The Colstrip coal-fired electric power generation complex in southeastern Montana consists of four units that previously had a combined generating capacity of 2276 megawatts (MW) (Figure 1.1). Units 1 and 2 were retired in January 2020. Active Units 3 and 4 are currently capable of producing up to 1480 MW of electricity. Solid wastes from the scrubber system of Units 3 and 4 are slurried and piped to a nearby effluent holding pond (EHP) that was placed in service in 1983. In 2004, a paste plant was placed in service to receive and dewater the scrubber slurry. The paste plant takes the 10–15% solids scrubber slurry and converts it into a paste with a solids content ranging from 65–68%. This material then is pumped to the EHP for impoundment.



Figure 1.1. Colstrip Units 1–4 (from right to left).

This report describes the 2021 monitoring and analysis activities for the groundwater and surface-water downgradient from the EHP on Talen Montana (formerly PPL Montana LLC) and Genie Land Company property. Assessment of the hydrologic conditions and water quality in the study area have been performed since 1984.

During the first quarter of 1984, Stipulation 12(d) was signed by the Montana Board of Natural Resources and Conservation, the Department of Natural Resources and Conservation, the Department of Health and Environmental Sciences, the Northern Plains Resource Council, the Rosebud Protective Association, the Genie Land Company, G.M. Garfield, the Rosebud County Board of Commissioners, and the Montana Power Company (now Talen Montana). This action provides for a water monitoring program downgradient (east) from Colstrip Power Plant Units 3 and 4 EHP in the Cow Creek and Pony Creek drainage basins in Rosebud County, Montana. After reorganization of the Montana State government in 1995, the roles and responsibilities of the Board of Natural Resources and Conservation, the Department of Natural Resources and Conservation, and the Department of Health and Environmental Sciences relative to Stipulation 12(d) were transferred to the Department of Environmental Quality and the Board of Environmental Review. Appendix A of this report provides comments about the report by the signatories of Stipulation 12(d).

Stipulation 12(d) provided for a water monitoring study that began in 1984, was suspended from May 1987 through August 1989, and was resumed in September 1989. At that time, Battelle–Pacific Northwest Division (PNWD), operator of the U.S. Department of Energy’s Pacific Northwest National Laboratory, assumed responsibility for project management, data analysis, and data interpretation. KC Harvey Environmental, LLC, based in Bozeman, Montana, is currently responsible for sample collection and handling and site maintenance. Enviro-Sci Consulting, through a contract with PNWD, performs the statistical estimation and testing of constituent trends and drainage-level comparisons of indicator parameter mean values.

Talen Montana also has an extensive groundwater monitoring program in the area upgradient from this study (i.e., in the vicinity of the Units 3 and 4 EHP). The latest site characterization reports are available from the Montana Department of Environmental Quality website.<sup>1</sup>

## 1.1 Study Area

The study area is located near Colstrip, Montana, which is about 30 mi south of Forsyth and about 140 mi east of Billings (Figure 1.2). This region lies in the Great Plains Physiographic Province (Shimer 1972) in the northwestern part of the Powder River Basin. The study area is immediately east of Colstrip and east of the Western Energy Company Rosebud Mine (Figure 1.3; T 2 N, R 42 E and 43 E). The area includes three creeks with intermittent flows: Pony Creek, Cow Creek, and South Fork Cow Creek. The creeks are tributaries to Rosebud Creek, which drains north to the Yellowstone River near the unincorporated town of Rosebud, Montana. The drainages are divided by well-dissected uplands and buttes. Elevation differences between the creeks and adjacent uplands range from 20 to 30 ft near the headwaters to approximately 300 ft in the central part of the study area. The Colstrip Units 3 and 4 EHP complex is near the headwaters of Cow Creek (Figure 1.3). The pond occupies a small drainage tributary to Cow Creek. The EHP and clear well have a combined design surface area of approximately 324 acres and a usable surface area of approximately 180 acres (Hydrometrics 1990).

Throughout this study, Pony Creek is included in comparisons with the other drainages because it is thought to represent background conditions. Pony Creek is hydrologically and geochemically similar to both the Cow Creek and South Fork Cow Creek drainages, but it is presumed to be unaffected by EHP activities. However, local variations in geology may differentially affect water quality. In addition, mining activity has progressed down the Pony Creek–Spring Creek and Pony Creek–Cow Creek divides to a point near Genie spring (GSP) 4 in Pony Creek, but the alluvial stream channel has not been directly disturbed. This mining activity could influence water-quality and hydrological parameters within the drainage.

---

<sup>1</sup> <http://deq.mt.gov/cleanupandrec/Programs/colstrip>

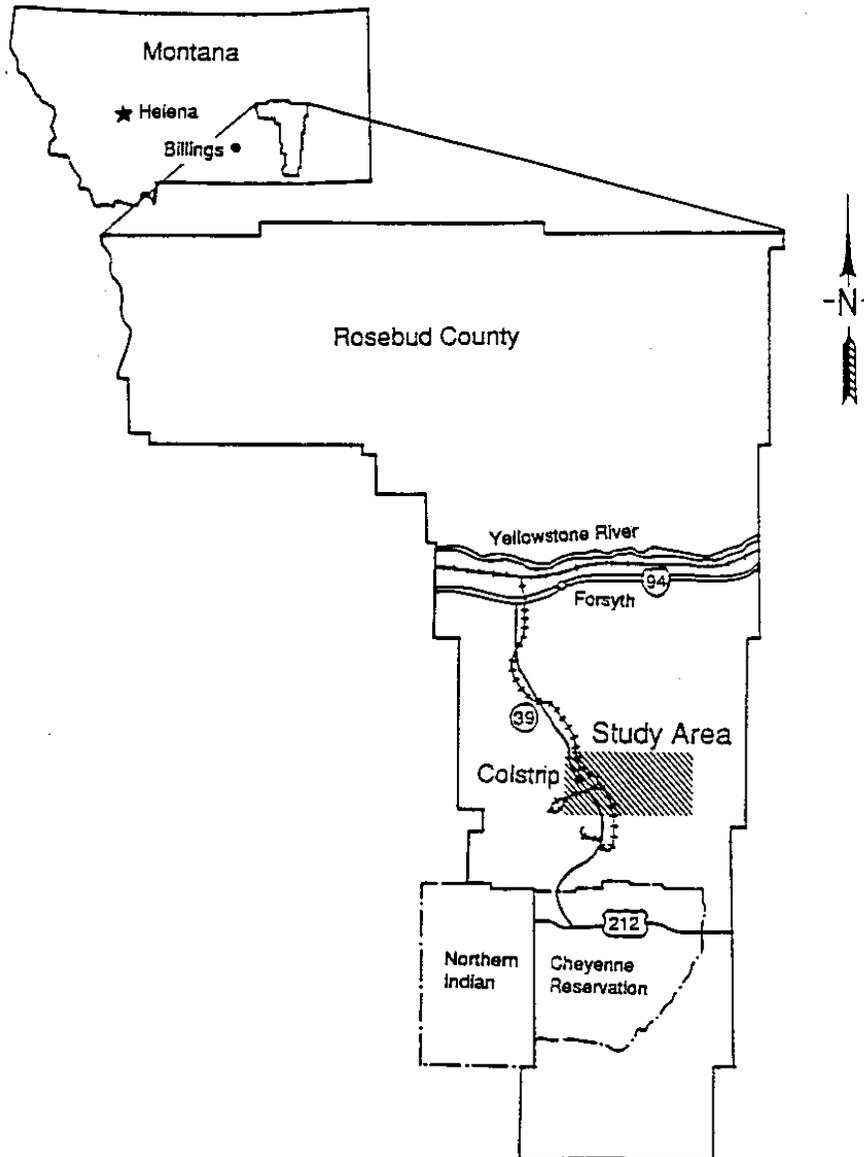


Figure 1.2. Location of study area (modified from Montana Department of State Lands and U.S. Office of Surface Mining 1983).

## 1.2 Study Purpose and Scope

The objective of the 2021 water monitoring study was to include and compare data collected during 2021 with data obtained in previous years to determine whether any changes had occurred in the water-quality constituents or hydrological parameters (water levels and groundwater flow directions) east of the Colstrip Units 3 and 4 EHP. A further objective was to establish the current characteristics and conditions of water-quality and hydrological parameters to evaluate any effects of ongoing or future events associated with the EHP (such as accidental releases) and related operations. For example, in August 2020, Hydrometrics conducted a hydraulic pumping test on the Units 3&4 EHP underdrain to generate the data that can support incorporating the underdrain into the full-scale implementation of the DEQ-approved remedy for the EHP, better estimate volumes within the ash, and evaluate groundwater

conditions directly below the EHP. Details of the pumping test may be obtained from the Montana Department of Environmental Quality website (see reference on previous page).

To meet the study objectives, water samples were collected from surface-water sites, springs, and wells during June and October 2021. Within those water sources are six categories of sample sites: 1) Genie new well (GNW), 2) Genie old well (GOW), 3) Genie alternative supply (GAS) well, 4) GSP, 5) Genie surface water (GSW), and 6) a group of three private monitoring wells (PWs<sup>2</sup>) that were installed in June 2009. The PW wells are located along Cow Creek near the western border of the study area (i.e., just northeast of the EHP). A continuous-monitoring probe that measures specific conductivity, temperature, and depth to water (DTW) was installed in one of the PW wells in December 2010. Data from that system were evaluated to help understand the short-term behavior of the shallow groundwater system in the upper portion of Cow Creek near the EHP.

This report presents the compilation, statistical analysis, and interpretation of the collected data, including water-quality information collected during the June and October 2021 sampling trips. Chemical analysis of the water quality samples was subcontracted to General Engineering Laboratories, LLC, in Charleston, South Carolina.

### **1.3 Report Contents and Organization**

Chapter 2 of this report presents a hydrological evaluation of the study area for 2021. Chapter 3 presents the sampling methods and quality assurance/quality control (QA/QC) results for 2021, statistical detection of water-quality trends at selected sites and for each drainage as a whole, and statistical comparisons of indicator parameter and constituent concentrations at the drainage level. Chapter 3 also describes and provides an interpretation of the water-quality data and site trends in the Cow Creek, South Fork Cow Creek, and Pony Creek drainages. Conclusions are presented in Chapter 4, and cited references are listed in Chapter 5. Appendix A contains comments about the report by the signatories of the Stipulation 12(d). Appendices B through G contain supporting information for the report, including example sampling documentation, a summary of analytical methods, photos of sampling sites, sample data (chemical analysis results), and statistical analysis methods and results.

---

<sup>2</sup> Private monitoring wells were installed by Hydrometrics, Inc.

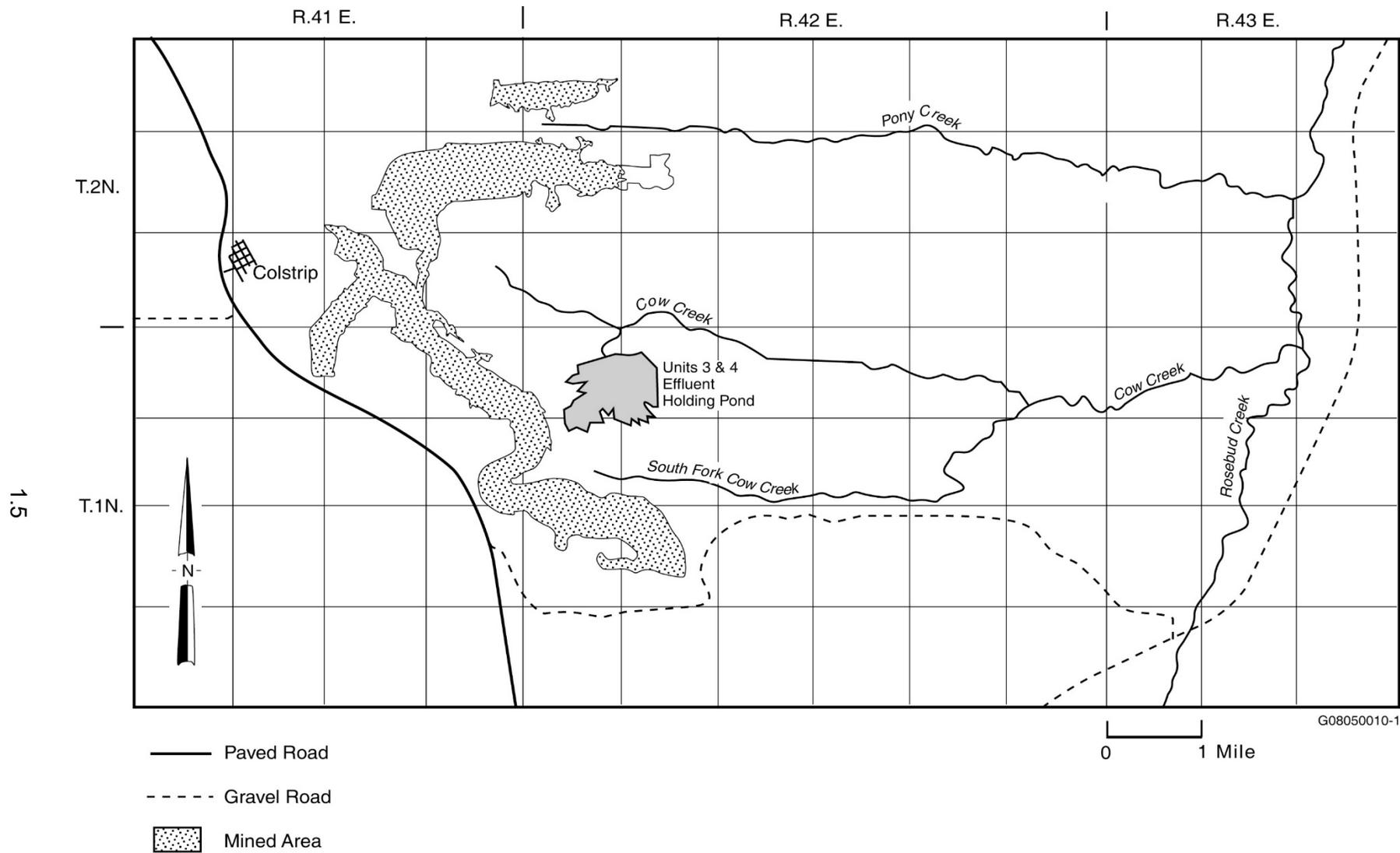


Figure 1.3. Map showing the location of the Colstrip Units 3 and 4 EHP near the headwaters of Cow Creek.



## 2.0 Hydrological Evaluation

This study focuses on monitoring both surface water and groundwater in the Cow Creek and South Fork Cow Creek drainages to evaluate effects from the Colstrip Units 3 and 4 EHP. Pony Creek also is monitored as a control drainage that is presumed to be unaffected by the EHP impoundment. An essential part of this monitoring is the ongoing evaluation of hydrological data, including recent precipitation, because hydrological conditions could influence chemical and contaminant indicator levels in these water sources.

All three streams are characterized as intermittent with little or no flow during most of the year. Most surface-water flow occurs during snowmelt or precipitation events in the spring. Groundwater springs (referred to as springs in this report) result in longer-duration flows of relatively small volume downstream from the source locations. Spring flows vary seasonally, and some are dry during parts of the year.

Temporal patterns of precipitation may influence water chemistry and water level/elevation by increasing or decreasing water flux through the hydrologic system. Therefore, in this study, it is important to monitor annual precipitation amounts and distribution trends throughout the year because the amount of precipitation likely influences indicator parameter levels. Fluctuations in precipitation may be the cause of apparent changes in water quality.

Annual precipitation at Colstrip in 2021 was 8.89 in., which is approximately 43% lower than long-term, historical averages from before 1983 (15.5 in.; Montana Department of State Lands and U.S. Office of Surface Mining 1983) and from 1984–2020 (15.60 in., based on monthly data from the National Oceanic and Atmospheric Administration<sup>1</sup>). The relatively low total for 2021 was less than the amount received in 2020 (11.78 in.) and significantly less than the total for 2019 (22.64 in.). Since this monitoring study began in 1984, the annual precipitation data have exhibited below-average values in the 1980s, near- and below-average values in the 1990s and early 2000s, and large 1–2-year fluctuations since 2004 (Figure 2.1). The five driest years between 1984 and 2021 were 2012 (8.04 in.), 1988 (8.58 in.), 2021 (8.89 in.), 2004 (9.76 in.), and 1984 (11.06 in.), respectively. In contrast, precipitation was above average in 2005, 2007, 2008, 2011, 2013, 2016, 2018, and 2019. The largest fluctuation recorded since 1984 was a decrease of over 16 in. from 2011 to 2012, which was just slightly larger than the increase in precipitation from 2012 to 2013 of approximately 15.5 in.

The monthly distribution of precipitation at Colstrip for 1984–2020 and 2021 is shown in Figure 2.2. In 2021, approximately 38% of the annual precipitation occurred in the spring and summer months, which is 37% less than the normal seasonal distribution of about 75% (Montana Department of State Lands and U.S. Office of Surface Mining, 1983). August, October, and December were the three wettest months in 2021, accounting for 18%, 13%, and 12% of the total annual precipitation, respectively. Overall, the winter months were relatively dry, which is a normal precipitation pattern for the area. November was the driest month with less than 5% of the annual precipitation. Because of limited winter precipitation, frozen (impermeable) surface soil, and insignificant upgradient catchment area, snowmelt and rainfall during winter are negligible sources of water influx for the subsurface drainages. The typically elevated groundwater levels observed during the spring monitoring campaigns are due to antecedent springtime rainfall. Also, when precipitation in late summer is high, measured groundwater levels also tend to be high some weeks later. Therefore, groundwater recharge of the alluvial aquifers, which are described in the next section, is presumably rapid in this hydrogeologic system.

---

<sup>1</sup> Data are available at [https://www.weather.gov/byz/local\\_climate](https://www.weather.gov/byz/local_climate).

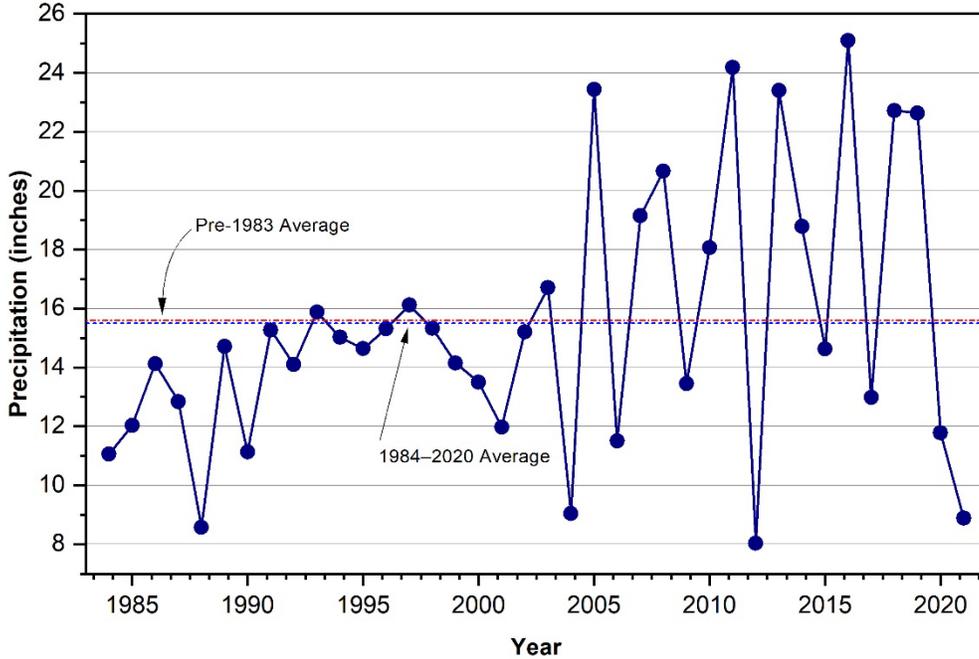


Figure 2.1. Annual precipitation at Colstrip, Montana, 1984–2021. Data were compiled from the National Oceanic and Atmospheric Administration (see footnote on previous page).

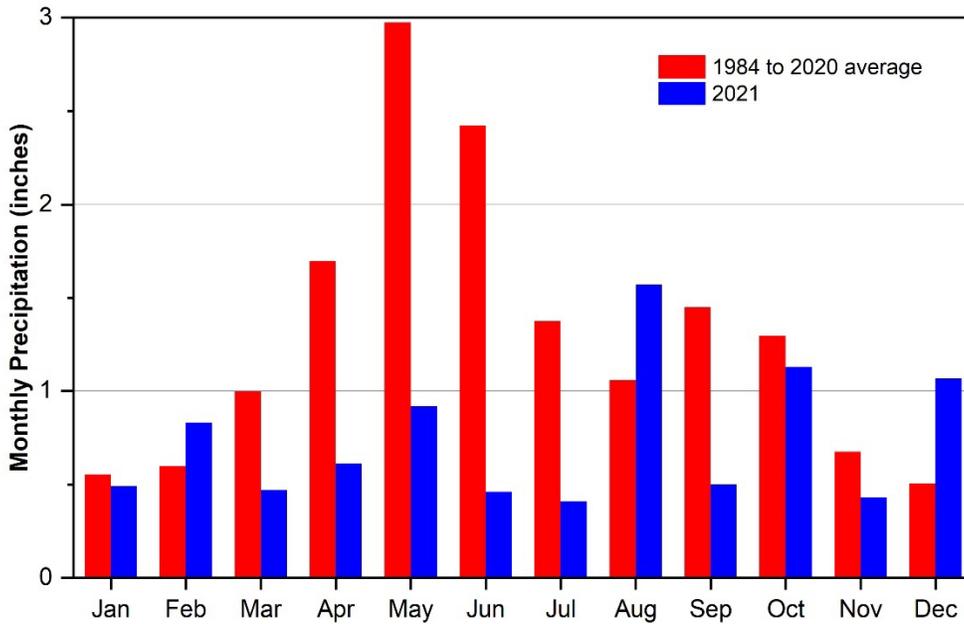


Figure 2.2. Monthly distribution of annual precipitation at Colstrip, Montana, 1984–2021. Data were compiled from the National Oceanic and Atmospheric Administration (see footnote on previous page).

## 2.1 Groundwater Flow System

Groundwater flow in the mine-permitted areas near Colstrip is well documented. The Montana Department of State Lands and U.S. Office of Surface Mining (1983) assessed the hydrology and geology

of the area. Van Voast et al. (1977), Van Voast and Reiten (1988), Erbes (2000), and Metesh (1994) have discussed groundwater hydrology with respect to mining impacts.

The study area is in the northwestern portion of the Powder River Basin. Rock strata near the study area generally slope several degrees downward toward the southeast. The Paleocene Fort Union Formation is exposed at the surface throughout the study area except in the creek valleys, where Quaternary alluvium forms the valley fill. The Rosebud and McKay coal beds are situated in the middle portion of the Tongue River Member of the Fort Union Formation. The Rosebud coal bed is strip-mined in the Colstrip area. Much of the middle Tongue River Member and younger rocks have been eroded in the study area; consequently, sub-McKay Tongue River Member siltstones, sandstones, and coal beds form the surface in the western portion, and the underlying Lebo Shale Member is exposed in the extreme eastern portion. In some places in the past, the coal outcrops have burned, leaving formations of erosion-resistant clinker that caps most of the ridges in the study area.

More detailed stratigraphy information for the immediate study area was obtained from wells drilled in the area. The GNW series of wells was drilled into or through the alluvium in the Pony Creek, Cow Creek, and South Fork Cow Creek drainages. The wells generally penetrate inter-bedded sands, silts, and clays with occasional gravelly zones. Several GAS wells in the upland portions of the study area penetrate the sub-McKay strata, whose stratigraphy consists of alternating layers of sand, silt, siltstone, shale, silty sandstone, sandstone, and thin layers of coal. Lee (1980) noted the formation of these sediments in a deltaic to estuarine environment. Lee further noted the complex depositional processes existing in such systems that result in sedimentary deposits characterized by lenticular beds, heterogeneous lithology, truncated units, and abrupt facies changes (Lee 1980).

The Rosebud and McKay coal seams of the Tongue River Member form the major aquifers in the Colstrip area. The shallow alluvial sediments and sandstone layers between the coal seams also are used occasionally in the area for groundwater supplies, depending on local conditions (Van Voast et al. 1977).

Hydrometrics (1987) and Metesh (1994) described the local groundwater flow in the sub-McKay sandstone, the McKay coal seam, and the overlying alluvium in the upper Cow Creek and South Cow Creek drainages. The reported mean hydraulic conductivity for the coal seams is about 14 gallons per day per square foot (gpd/ft<sup>2</sup>), while the hydraulic conductivity for the overlying surficial alluvial aquifer is reported to average 900 gpd/ft<sup>2</sup> for the Cow Creek drainage (Metesh 1994). Hydraulic conductivity in the Pony Creek alluvium is also expected to be approximately 900 gpd/ft<sup>2</sup>; however, as discussed below, it is likely much higher in the reclaimed backfill in the upper regions where mining has occurred.

Olsen et al. (1987) mapped the water table in the alluvial aquifer in the Cow Creek and South Fork Cow Creek drainages. Both surface water and groundwater generally flow to the east, parallel to the creek drainages and sub-parallel to the regional southeastward dip of the underlying sediments. Hydraulic head contours indicate that groundwater flows down the stream drainages in both the sub-McKay and alluvial sediments. Groundwater appears to discharge from the sub-McKay sediments into the alluvium. The drainage basin boundary near the headwaters of Cow Creek is on Western Energy Company property in Area E of the Rosebud Mine.

Waren and McDannel (2003) evaluated groundwater conditions in mined lands in the Colstrip area; they noted groundwater flow tends to be re-established in backfill materials that replace coal aquifers when mined lands are reclaimed. However, they also noted local perturbations of the potentiometric surface can occur that are persistent in the reclaimed environment. This finding indicates portions of the backfill materials that at least have different hydraulic properties (e.g., higher hydraulic conductivities) from those of the coal bed and adjacent sediments before they were mined. Consequently, these phenomena could affect portions of the Cow Creek and Pony Creek drainages adjacent to mined lands where hydraulic head distributions have changed in the reclaimed flow field. Therefore, as with impacts on flow resulting from

fluctuations in precipitation, any induced changes in the alluvial flow field in these drainages (due to local changes in hydraulic conductivity) may affect concentrations of the sampled parameters there (e.g., increases in flow rate could have a diluting effect on constituent concentrations).

Hydrological measurements for this study were conducted during mid-June and early October 2021 (Figure 2.3). Water-level data are shown in Table 2.1 for all wells. DTW in wells screened in the alluvium ranged from approximately 5 ft to 22 ft in 2021.



Figure 2.3. Measuring DTW with a water-level meter.

Overall, water-level elevations in monitoring wells have fluctuated within a few feet over the course of the study, and changes in the alluvial groundwater levels generally have reflected changes in both annual and seasonal precipitation. During 2021, the overall average groundwater-level measured in all wells was 0.40 ft lower than the corresponding values from the summer and fall sampling trips from the previous year. Individual wells with the greatest year-over-year changes from 2020 to 2021 include GNW 2 (2.6 ft lower in the summer and 2.0 ft lower in the fall), GNW 6 (1.1 ft lower in the summer and 1.5 ft lower in the fall), GOW 11 (1.4 ft lower in the summer and 1.6 ft lower in the fall), and PW 734 (2.1 ft lower in the summer and 2.3 ft lower in the fall).

The GOW well data are listed in Table 2.1. Many of these wells were drilled in the 1950s, and their construction and lithology is not well documented. GOW 5 is deeper and may penetrate through the Lebo Shale Member into the Tullock Member. Geochemical evidence for discerning the source of water to the GOW wells was documented by Olsen et al. (1991) and updated by Thompson et al. (2011). Only two of the GOW well levels are currently measured; GOW 1's average water level was 0.6 ft lower in 2021 than in 2020, while GOW 11's average water level was 1.5 ft lower in 2021.

Table 2.1. 2021 groundwater-level data (feet).

Site	TOC <sup>(a)</sup>	Total Depth	Casing Height (above ground surface)	Measured DTW <sup>(b)</sup> June	June Elev. <sup>(c)</sup> (TOC-DTW)	Measured DTW <sup>(b)</sup> October	Oct. Elev. <sup>(c)</sup> (TOC-DTW)
GNW 1	3047.84	42.5	--	5.30	3042.54	5.34	3042.50
GNW 2	3015.99	33	3.17	15.89	3000.10	18.73	2997.26
GNW 3	2991.86	32.5	2.88	18.60	2973.26	19.23	2972.63
GNW 4	2957.19	30	2.25	8.39	2948.80	9.44	2947.75
GNW 5	2950.64	35.5	3	9.49	2941.15	9.80	2940.84
GNW 6	3059.84	26	--	19.02	3040.82	20.02	3039.82
GNW 7	2989.56	17	2.88	15.93	2973.63	15.77	2973.79
GNW 8	2947.56	32	2.75	9.03	2938.53	9.66	2937.90
GNW 9	3065	--	--	--	--	--	--
GNW 10	--	~20	1.86	Dry	--	--	--
GNW 11	3210	--	1.79	--	--	--	--
GOW 1	2972.38	--	--	21.44	2950.94	22.79	2949.59
GOW 3	3124.29	--	--	--	--	--	--
GOW 4	3024.77	--	--	--	--	--	--
GOW 5	2910.26	--	--	--	--	--	--
GOW 6	3261.66	--	--	--	--	--	--
GOW 11	3065	--	--	22.17	3042.83	23.42	3041.58
GOW 12	2940	105	--	--	--	--	--
PW 734	3083.65	13	1.8	13.79	3069.86	14.48	3069.17
PW 735	3086.21	80	1.7	31.60	3054.61	31.94	3054.27
PW 736	3075.5	37	1.6	21.42	3054.08	21.92	3053.58
GAS 1	3075	100	2.25	40.91	3034.09	41.49	3033.51
GAS 2	2985	200	--	--	--	--	--
GAS 3	3015	195	1.17	--	--	--	--
GAS 4	3050	280	--	--	--	--	--
GAS 6	3080	280	--	--	--	--	--
GAS 7	2925	200	--	--	--	--	--
W-1	2880	--	2.50	11.52	2868.48	12.73	2867.27

(a) TOC = top of well casing. Values are elevations above mean sea level.

(b) DTW = depth to water.

(c) Elevation above mean sea level.

GAS = Genie alternative supply well; GNW = Genie new well; GOW = Genie old well; PW = private monitoring well; W-1 = Montana State University-installed well.

Three PW wells were installed at the upper part of the Cow Creek drainage in June 2009. These closely spaced wells monitor three different vertical intervals. PW 734 is screened in and monitors a tributary to Cow Creek. The screen interval is from 6 to 11 ft; the uppermost aquifer in this region is generally between 8 and 10 ft in depth and is often dry. Well PW 736 also monitors the uppermost aquifer with a screened depth of 16 to 36 ft. Well PW 735 monitors a deeper aquifer with a screened depth interval between 35 and 80 ft. Well PW 734 was sampled in June 2021 but did not have sufficient water for

sampling in October. This well was also dry during several previous sampling trips (2010 spring and fall, 2014 spring, 2015 fall, 2016 spring and fall, and 2017 spring and fall). Water levels in the shallower PW wells may be affected by the timing of monitoring in relation to precipitation events as well as operations associated with the EHP, which include pumping a network of wells installed between the EHP and the PW wells. As discussed in Section 3.6.7, continuous water level and electrical conductivity measurements are recorded in PW-736. Those data showed a water-level increase during March and a corresponding slight overall decrease in electrical conductance, consistent with groundwater recharge from precipitation.

The measuring point elevations (i.e., top of the casing) were resurveyed before the June 2010 measurements for GNW 1, GNW 2, GNW 3, GNW 4, GNW 5, GOW 1, and PW 734. The largest change in measuring point elevation occurred at GNW 5, where the new survey revealed a 33.62 ft lower elevation from the top of the casing. The other resurveyed measuring point elevations increased between 0.01 ft and 2.64 ft. Table 2.1 shows the most recent top-of-casing elevations for all wells. The new measuring point surveys should be accounted for when comparing water-level elevations in this report with earlier reported elevations that were based on inaccurate elevation surveys.

The hydrological conditions of springs and surface-water sites are listed in Table 2.2 (Figure 2.4 shows the site locations). Appendix D contains photos of these sites from the June and October sampling campaigns.

Table 2.2. 2021 Status of springs and surface-water sites.

Site Name	Site Elevation (ft)	June Status	October Status
GSW 1	3047.80	OK(s) <sup>(a)</sup>	OK(s)
GSW 2	3026.51	OK(s)	OK(s)
GSW 3	3005.14	Dry	Dry
GSW 5	2951.85	OK(s)	Dry
GSW 6	2905.43	OK(s)	OK(s)
GSP 1	3036.03	OK(s)	OK(s)
GSP 2 <sup>(b)</sup>	2949.87	OK(s)	OK(s)
GSP 3	3193.42	OK(s)	OK(s)
GSP 4	3275.67	OK(s)	NS <sup>(c)</sup>
GSP 5	3233.67	Dry	NS
GSP 6	3180.00	OK(s)	NS
GSP 7	2952.92	OK(s)	NS
GSP 8	2828.29	OK(s)	NS
GSP 9	3251.31	OK(s)	NS
GSP 10	3120	OK(s)	NS

(a) OK(s) = site okay for sampling (i.e., adequate water was available)

(b) GSP 2 is also known as Stinking Spring.

(c) NS = site not scheduled for sampling.

GSP = Genie spring; GSW = Genie surface water.

Several of the springs provide enough water for livestock watering. GSW 3 and GSP 5 were dry during the June 2021 sampling trip, and GSW 3 and GSW 5 were dry during the October campaign. The remaining spring and surface-water sites were sampled as scheduled. Elevations for springs and surface-water sites were determined at the beginning of the study and have been considered accurate without resurveys. The estimated potentiometric surface is shown in Figure 2.5.

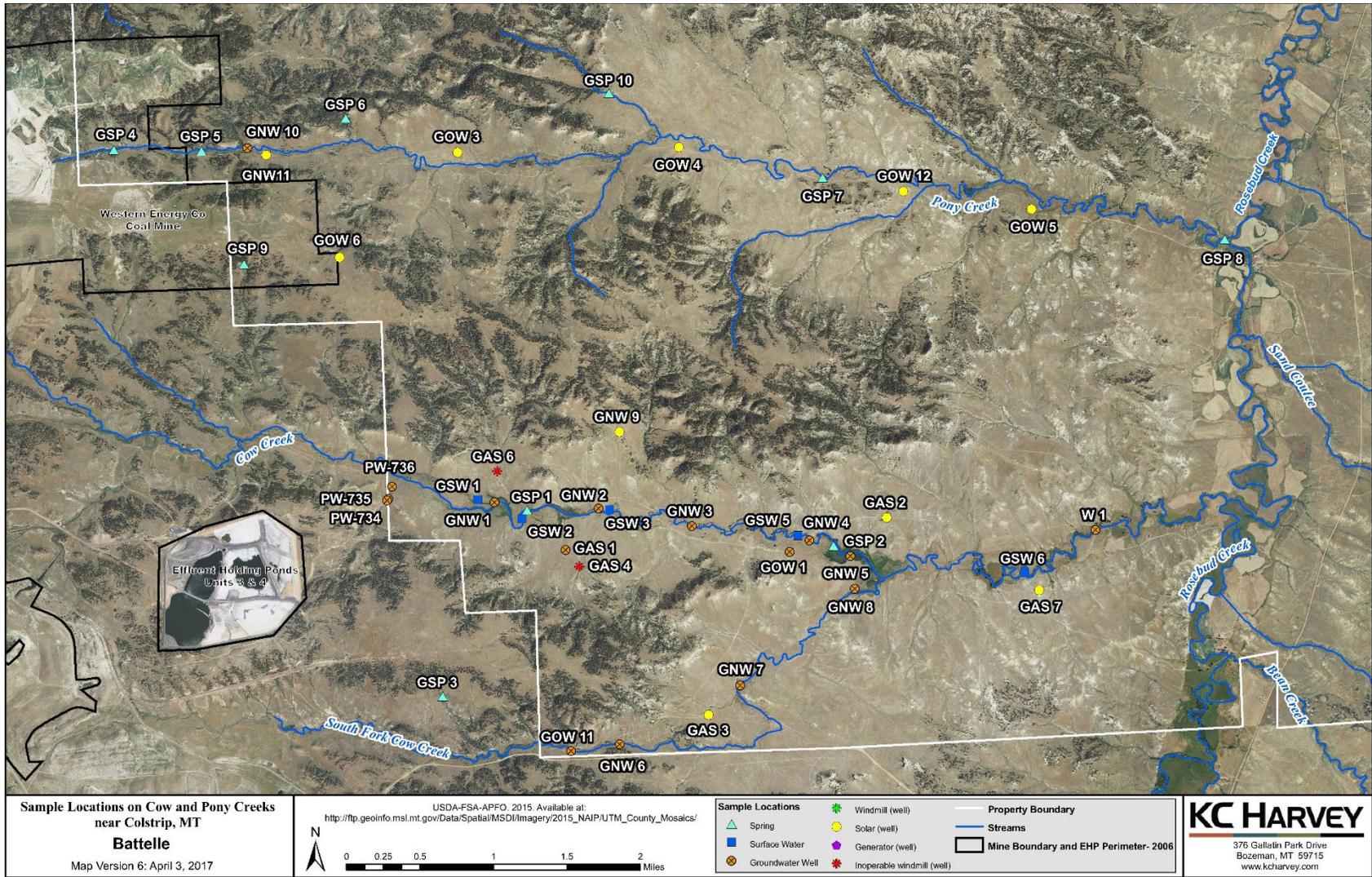


Figure 2.4. Monitoring locations in the study area.

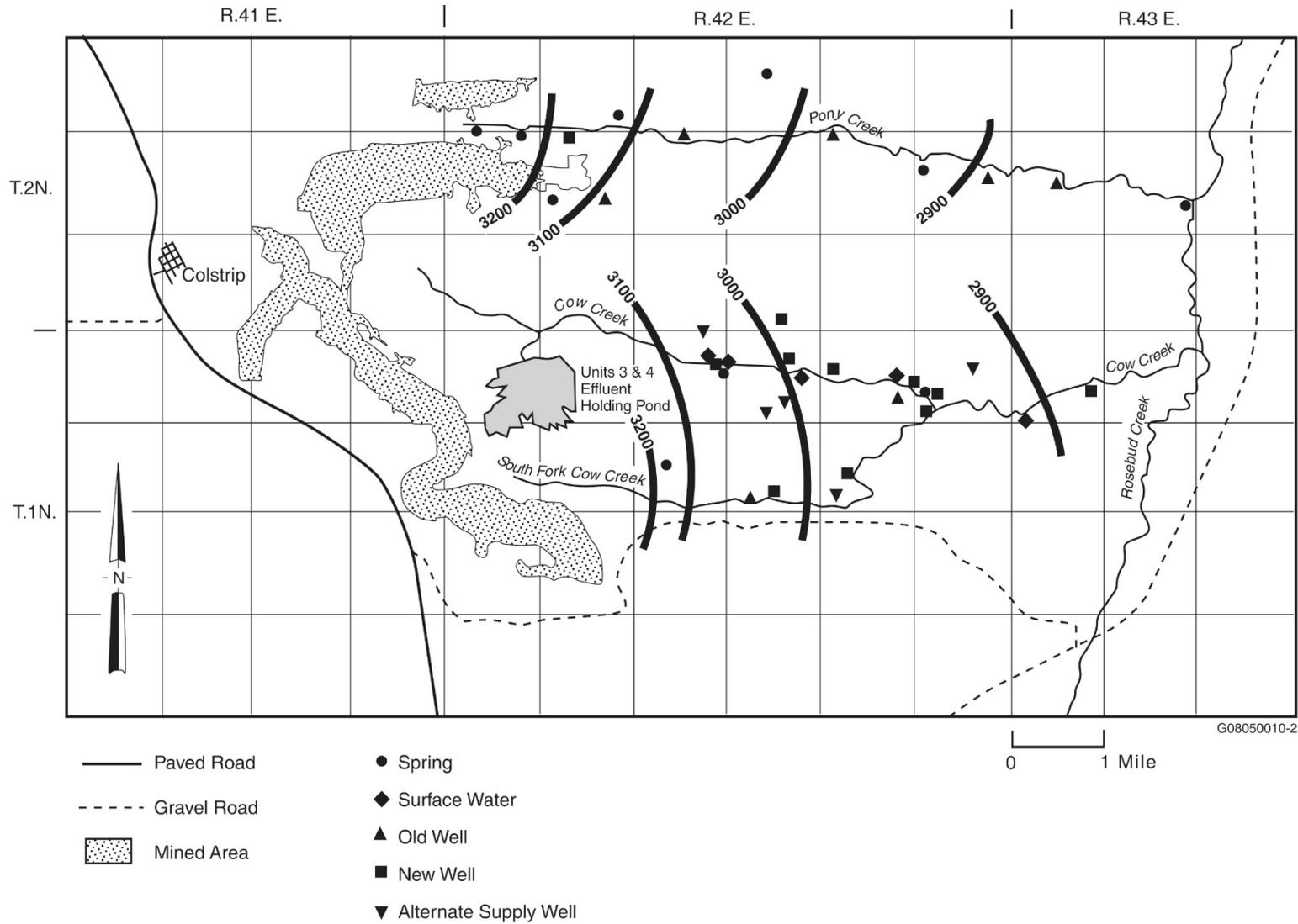


Figure 2.5. Water table in the alluvial aquifers of Pony Creek, Cow Creek, and South Fork Cow Creek.

Relatively minor differences have been observed since the early 1990s in the Cow Creek and South Fork Cow Creek drainages; water levels in October 2021 generally were within several feet of levels observed in the early 1990s. The fall measurement was mapped because it tends to be more representative of base flow conditions; the effects of spring runoff and winter recharge are minimized. Most of the measured water levels are from wells that penetrate the alluvial aquifer. The hydraulic head contours are based on the assumption that the sub-McKay strata are hydraulically connected to the alluvial aquifers in each creek valley (Hydrometrics 1987); thus, a hydraulic connection is assumed for all wells. This assumption is consistent with the hydro-stratigraphic location of these units above the Lebo Shale, a major aquitard in the system. Although the aquifers are hydraulically connected, because of their parallel flow directions and the large distances between monitoring locations in different drainages, it is assumed that any water-quality changes at sites in one drainage do not directly affect the sites in the other drainages.

Groundwater gradients were calculated based on water-level measurements taken in the alluvia for Cow Creek and South Fork Cow Creek in October 1990. The October 2021 measurements do not indicate any significant changes in these calculations. The groundwater gradient in Cow Creek is about 0.008 ft/ft between GSW-1 and GNW 5. The new measuring point survey results do not significantly affect this calculation of the gradient, and the recalculated groundwater elevation at GNW 5 is more consistent with nearby wells GNW 4 and GNW 8. The groundwater gradient in Pony Creek is about 0.01 ft/ft between GSP 4 and GSP 7. Because the hydraulic conductivities in these drainages are essentially unchanged from previous years, the groundwater velocities, which are proportional to the product of gradient and conductivity, are also expected to be unchanged. Therefore, any detected site trends would not appear to be attributable to changes in velocities (i.e., not related to precipitation).

## 2.2 Cow and Pony Creek Monitoring Network

The sampling network includes a combination of wells, springs, and surface-water locations (Figure 2.4 and Plate 1 [inside back cover]). See Section 3.2 for the locations where water samples were collected at least once during 2021. Specific well construction details are included in Table 2.1.

The GNW wells and PW 736 are generally less than 50 ft deep and were drilled to monitor the valley fill alluvium. Wells PW 735, GNW 9, and GNW 11 are deeper and penetrate through the valley fill alluvium into the sub-McKay part of the Tongue River Member. Well PW 734 monitors the uppermost aquifer in a relatively shallow zone at about 8–12 ft below ground surface. All GNW and PW wells are accessible for measuring water levels except for GNW 9 and GNW 11, which have dedicated, solar-powered pump configurations that prohibit measurements.

The GAS wells generally penetrate into the sub-McKay portion of the Tongue River Member and deeper units of the Fort Union Formation. Only GAS 1 is accessible for water-level measurements. The water level has remained fairly constant in this well, typically varying less than 1 ft over the entire long-term study. This behavior is consistent with that of a confined aquifer. Obtaining access to the remainder of the GAS wells for water-level measurements would provide more insight into the regional groundwater flow characteristics of the sub-McKay flow regime. The geochemical characteristics of water from these deeper wells are compared with those of the alluvium in the geochemical discussion (see Section 3.5.1).

The configuration of several of the GOW and GAS wells makes it difficult to obtain water-level measurements at those sites since most of these wells were outfitted with windmills. In 1992, access ports were installed on some of the wells, but because of limited space between the riser pipe and the well casing, obtaining water-level measurements was impractical. Several of the older GOW wells have fallen into disrepair and are no longer used to provide stock water. Solar-powered stock-watering systems have been installed at five locations (Figure 2.4). Water levels cannot be measured in the solar-powered systems, but water-quality samples have been collected at these sites (GAS 7, GNW 9, GNW 11, GOW 4, and GOW 12). As discussed in Section 3 and Appendix G of this report, well GOW 12 is suspected to be

tapping a geochemically different water supply than it accessed before the solar-powered pump installation.

## 2.3 Results and Discussion

Water-level data collected during 2021 confirm previous observations of hydrological conditions and groundwater movement in the Pony Creek, Cow Creek, and South Fork Cow Creek drainages. Water-level measurements for wells withdrawing water from the alluvial materials that fill the valleys of Pony Creek, Cow Creek, and South Fork Cow Creek continue to indicate that groundwater flows in an easterly, down-valley direction toward Rosebud Creek.

Groundwater levels measured in the alluvial wells during 2021 averaged about 0.93 ft lower than those of the previous year. With one exception, all of the summer and fall average levels were lower in 2021. Well W-1 had an identical level in the summers of 2020 and 2021, but the fall level was 0.8 ft lower in 2021. This is consistent with annual changes in precipitation: total precipitation in 2021 was approximately 25% lower than in 2020. Water levels declined in all wells between the June and October 2021 sampling campaigns. The largest decreases occurred at wells GNW 2 (2.84 ft), GOW 1 (1.35 ft), GOW 11 (1.25 ft), and W-1 (1.21 ft).

Groundwater levels measured during 2021 do not indicate any major changes in groundwater flow patterns for Cow Creek and South Fork Cow Creek. The EHP for Colstrip Units 3 and 4 is situated within the drainage near the headwaters of Cow Creek. The headwaters of all three drainages are within areas disturbed by mining.

More work is needed to assess the aquifer properties and better quantify the potential for groundwater movement and potential contaminant migration through groundwater in the study area. Groundwater recharge over much of the alluvial aquifers is suspected to be quite rapid because the majority of recharge occurs during spring. As noted above, this is due to little available wintertime precipitation and seasonally frozen surface soil; therefore, effective groundwater recharge is negligible. Another important aspect of the hydrology in the drainage is the role of creek flow, particularly in combination with recharge potential. Surface-water flow in the creek is likely the fastest path of contaminant migration through the flow system and may explain the simultaneous peaking of boron that occurred at several Cow Creek sites during the mid-1990s (Appendix G). No creek flow was observed within the South Fork Cow Creek, Cow Creek, and Pony Creek drainages during the June and October 2021 sampling trips with the exception of minor flows in GSP 1 (Cow Creek), GSP 6 (Pony Creek), GSP 8 (Pony Creek), and GSP 10 (Pony Creek) in June.

Water levels in the alluvial aquifer generally are 5 ft to 20 ft below the creek bed in the Cow Creek drainage, indicating the creek would tend to lose water to the aquifer. Such losses would depend on the hydrogeology of any given creek segment. Consequently, seasonally flowing surface water is a potential pathway for quick movement and uneven distribution of constituents within a drainage creek. Additional study (e.g., surface-water modeling) would be required to understand the potential for contaminant migration through creek flow and for creek flow to act as recharge to the alluvial aquifer throughout the drainage whenever there is flow.

## 3.0 Water Quality in the Cow Creek and Pony Creek Drainages

Water-quality data gathered from the 2021 sampling trips are presented in Sections 3.1 through 3.3. A statistical analysis of time trends and a comparison of drainages are provided in Section 3.4, followed by interpretation of the results in Section 3.5. Throughout this chapter, concentrations of chemical parameters are listed in units of parts per million (ppm), which is equivalent to milligrams per liter (mg/L). Specific conductivity is reported in units of microsiemens per centimeter ( $\mu\text{S}/\text{cm}$ ). The naming convention for samples is the site name followed by a dash and the sampling trip number. For 2021, the June sampling campaign was trip 74, and the October sampling campaign was trip 75.

### 3.1 Water Sampling

To determine the current water-quality status of the Cow Creek and Pony Creek drainages, three categories of water sources (surface, springs, and wells) were sampled and analyzed. Samples were taken from the six site categories noted in Chapter 1.0: GAS, GNW, GOW, GSP, GSW, and PW. Samples were collected from nine wells and springs in the Pony Creek drainage; a spring and a well on the Cow Creek/Pony Creek divide; 20 wells, springs, and surface-water sites in the Cow Creek drainage; and 6 wells and springs in the South Fork Cow Creek drainage. Table 3.1 lists the sites that were sampled in 2021, and Figure 2.4 shows the sampling locations. Sampling in the Pony Creek drainage and the Cow Creek/Pony Creek divide was conducted during June only. Pony Creek well GNW 10 and spring GSP 5 were dry during the June campaign and were not sampled. Wells in the upper portion of the alluvial aquifer, springs, and surface-water sites in Cow Creek and South Fork Cow Creek were predominantly sampled in both June and October. Cow Creek surface-water site GSW 3 was not sampled in June or October because the site was dry. Additionally, GSW 5, also on Cow Creek, was not sampled in October due to lack of water. Data gathered from the sampling trips are presented in Sections 3.3 and 3.5.

### 3.2 Sampling Protocols

The procedures described in this section are the sample collection techniques used for this hydrological/water-quality study. The protocols were developed to 1) minimize the possible contamination of samples, 2) ensure the samples closely represent the water quality of the sampling site, and 3) ensure accurate identification of samples collected during this study. As each sample site was visited, observations including general weather conditions, sample site conditions, and purging data for wells sampled with a submersible pump or bailer were recorded in a field notebook.

Water samples were collected by the methods most appropriate to the sample source. Table 3.1 summarizes the equipment used at each site sampled during 2021. Springs, wind-driven and solar-powered stock-watering wells, and surface-water sites were sampled with a battery-powered peristaltic pump with an in-line, disposable filter. Twelve wells were sampled with a submersible pump, and two wells were sampled with a bailer due to low yield or high levels of fine sediments.

#### 3.2.1 Alluvial Wells

Water levels were recorded at most GNW- and PW-labeled wells and at GOW 1, GOW 11, GAS 1, and W-1 before they were pumped or bailed. Solar-powered pumps have been installed in wells GNW 9 and GNW 11, and it was impractical to obtain water-level measurements at those sites. The reference point for the water-level measurements was the top of the well casing (marked location or lowest point on the top of the casing). Water levels were recorded to the nearest 0.01 ft using an electric tape water-level indicator.

Table 3.1. Sites sampled during 2021 and sample acquisition methods.<sup>(a)</sup>

South Fork Cow Creek		Cow Creek		Cow/Pony Creek Divide		Pony Creek	
GSP 3	P	PW 734	P	GSP 9	P	GSP 4	P
GOW 11	S	PW 735	S	GOW 6	P	GNW 11 <sup>(b)</sup>	P
GNW 6	S	PW 736	S	--	--	GSP 6	P
GAS 3	P	GSW 1	P	--	--	GOW 3	P
GNW 7	B	GNW 1	S	--	--	GSP 10	P
GNW 8	S	GSW 2	P	--	--	GOW 4 <sup>(b)</sup>	P
--	--	GSP 1	P	--	--	GSP 7	P
--	--	GAS 1	S	--	--	GOW 12 <sup>(b)</sup>	P
--	--	GNW 2	S	--	--	GOW 5 <sup>(b)</sup>	P
--	--	GNW 9 <sup>(b)</sup>	P	--	--	GSP 8	P
--	--	GNW 3	S	--	--	--	--
--	--	GOW 1	S	--	--	--	--
--	--	GSW 5	P	--	--	--	--
--	--	GNW 4	S	--	--	--	--
--	--	GSP 2	P	--	--	--	--
--	--	GNW 5	S	--	--	--	--
--	--	GAS 2	P	--	--	--	--
--	--	GSW 6	P	--	--	--	--
--	--	GAS 7 <sup>(b)</sup>	P	--	--	--	--
--	--	W-1	B	--	--	--	--

- (a) Sample acquisition methods: P = pumped with portable peristaltic pump, S = pumped with 4-in. submersible pump, B = bailed with 3.5-in. outside diameter (OD) PVC bailer, then bailer sampled with peristaltic pump  
(b) Well is pumped by a dedicated, solar-powered pump. Outflow from the pump was sampled in-line (i.e., before flowing into the nearby stock-watering tank).

GAS = Genie spring; GSW = Genie surface water; PVC = polyvinyl chloride; PW = private monitoring well installed by Hydrometrics, Inc.; W-1 = Montana State University-installed well. GSP 2 also is known as Stinking Spring.

Most well samples were acquired with a 4-in. submersible pump (Figure 3.1a and b) or a 3.5-in. OD polyvinyl chloride bailer (Figure 3.1c). A battery-powered peristaltic pump was used to filter water obtained with the bailer. When a well was sampled with the submersible pump, a minimum of 100 gal (i.e., greater than three well volumes) of water were pumped before sampling. This purging process helps ensure that the collected water is representative of the groundwater in the vicinity of the well. The total discharge was recorded on a totalizing flow meter and written in the field notebook to the nearest 0.1 gal. Well GNW 7, a low-yield well, was bailed dry during both the June and October 2021 sampling events and allowed to recharge for 24 hours or more before sampling.

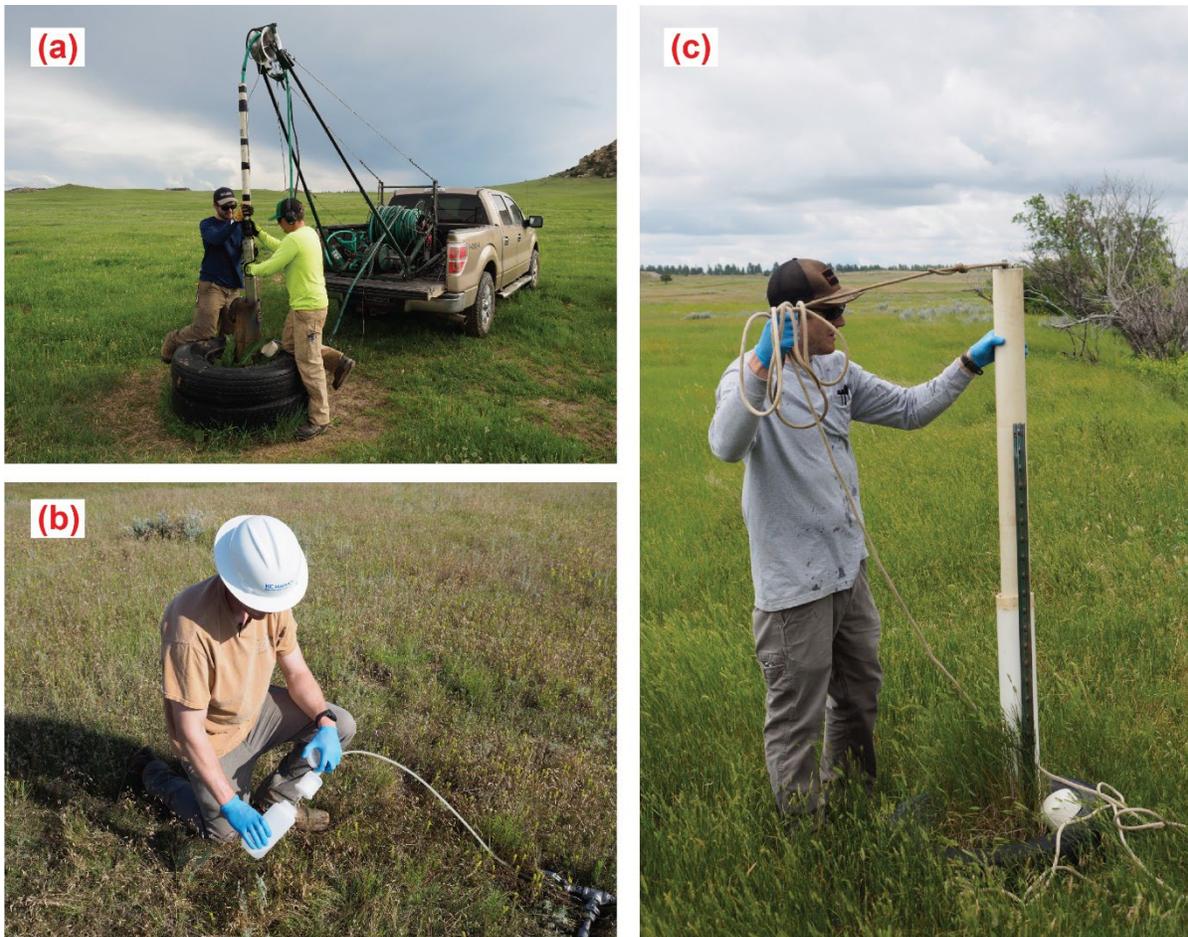


Figure 3.1. Well sampling: (a) Lowering the submersible pump into well GNW 2. (b) Filling sample bottles. Water from the well is pumped through the sampling manifold, a 3 ft section of ¼-in. OD poly tubing, and an in-line filter before flowing into the sample containers. (c) Lowering a bailer into well W-1.

Wells GAS 2, GAS 3, GAS 7, GOW 3, GOW 4, GOW 5, GOW 6, and GOW 12 previously were equipped with windmill-driven pumps but now have solar-powered pumps, which eliminates the need for wind during sampling. Relatively deep wells GNW 9 and GNW 11 are also equipped with solar-powered pumps. All wells with solar-powered pumps were sampled at the pump outfall using a battery-powered peristaltic pump. These wells are generally not purged before sampling because the solar-powered pumps are usually running continuously during daylight hours. In cases where a well pump was not running steadily (e.g., due to overcast conditions), the pump was allowed to operate for several minutes before sampling. The windmill at GAS 4 has not been functional since 2003; therefore, no samples were collected from that location.

### 3.2.2 Spring and Surface Water

Spring and surface-water samples were obtained using a peristaltic pump and an in-line filter assembly to fill the sample bottles directly at the sampling site. Flow measurements or estimates were made at the time of sample collection for all surface-water sites and springs, if possible (Figure 3.2). Appendix D contains photos of the spring and surface-water sites from both sampling campaigns.



Figure 3.2. Measuring the flow rate at a spring.

### 3.2.3 Sample Handling, Quality Control Samples, and Field Measurements

Field samples were preserved based on the recommendations in standard procedures used by the analytical laboratory. Each water sample was divided into two aliquots: 1) 250 mL of a 0.45  $\mu\text{m}$  filtered sample with 1 mL of 1/1 concentrated  $\text{HNO}_3$ /water preservative (for inductively coupled argon plasma analyses for magnesium, calcium, sodium, potassium, and boron) and 2) 1 L of a 0.45  $\mu\text{m}$  filtered sample with no added preservative (for pH, conductivity, alkalinity, total dissolved solids [TDS], bromide, chloride, and sulfate analyses).

Samples were collected from submersible pumped wells using a stainless-steel sampling manifold equipped with a disposable in-line 0.45  $\mu\text{m}$  filter assembly. The filter was flushed with approximately 500 mL of sample water before the sample was collected. Sample water was filtered directly into the appropriate sample container. Spring and surface-water samples were obtained similarly by attaching a 0.45  $\mu\text{m}$  filter to the peristaltic pump tubing, flushing the filter, and delivering a field sample into the appropriate container.

In June 2015, we introduced a sample-mixing protocol at some of the spring and surface-water sites where limited water was available (GSP 1, GSP 2, GSP 3, GSP 4, GSP-5, GSW 1, and GSW 5). Approximately 1.5 L of sample water was pumped through a filter into a 2 L plastic bottle lined with a 1 gal zip-lock bag. The fluid then was mixed thoroughly before pouring it into sample containers. This extra mixing step was designed to improve the consistency of results between metals and anions for low-water sites where the sample concentrations are suspected to vary during the sampling process. We simplified this practice in 2016 by pumping the sample fluid directly into a pre-rinsed 2 L bottle, and fewer charge imbalances were observed at several of the spring and surface-water sites than in previous years. For 2021, this modified pre-mixing protocol was followed at all spring and surface-water sites in both June and October.

At each sampling site, an unfiltered sample aliquot was dispensed into a test tube for field parameter analysis. Measurements of conductivity and pH were made within 8 hours of sampling. These measurements provide initial screening information and help verify laboratory results. Water temperature was determined in situ or in the pump-flow discharge stream with a dial or digital thermometer. These instruments were checked for accuracy with an ice bath and a laboratory mercury thermometer and, in all cases, were found to be accurate within  $\pm 1^{\circ}\text{C}$ . Sample pH was determined using a pH meter calibrated using standard buffer solutions ranging from pH 4 to pH 10. Sample conductivity was measured with a conductivity meter calibrated using various calibration standards ranging from 447 to 12,880  $\mu\text{S}/\text{cm}$ . All calibration data were recorded in the project field notebook.

For both sampling campaigns, field blank and duplicate samples were collected and sent to the analytical laboratory. Equipment blanks were prepared by running distilled water through the peristaltic pump and in-line filter assembly (Figure 3.3). Bottle blanks were prepared by filling sample containers with distilled water. Two sample duplicates were collected during each event: one from a submersible pumped well and the other from a location where the peristaltic pump and in-line filter assembly were used. All samples were placed immediately in coolers with ice.



Figure 3.3. Preparing an equipment blank sample. Deionized water was pumped from the container on the left through a sample filter and into the sample bottle.

### 3.2.4 Sample Custody

Sample custody was managed by Pacific Northwest National Laboratory (PNNL) personnel. After collecting and processing the samples, KC Harvey Environmental, LLC staff transferred the samples to the on-site PNNL representative for the appropriate custodial procedures (i.e., sealing, sample packing, storing, and shipping). An example of a chain-of-custody form is provided in Appendix B.

### 3.3 Sample Analysis

Samples were analyzed by General Engineering Laboratories (GEL) at its main laboratory facility in Charleston, South Carolina. Appendix C lists the analytical methods used by GEL, and Appendix F contains tables of analytical results from both sampling campaigns in 2021. GEL's detection and reporting limits for all measured parameters are listed in Table 3.2.

Table 3.2. Analytical detection and reporting limits.

Parameter	Method <sup>(a)</sup>	Detection Limit <sup>(b)</sup> (ppm)	Reporting Limit <sup>(c)</sup> (ppm)
Alkalinity	2320B	1.45	4.00
Conductivity	120.1	1.00 $\mu$ S/cm	1.00 $\mu$ S/cm
pH	9040	0.01 pH units	0.100 pH units
TDS	2540C	3.40	14.3
Bromide	9056	0.067	0.200
Chloride	9056	0.067	0.200
Sulfate	9056	0.133	0.400
Boron	6010	0.015	0.050
Calcium	6010	0.050	0.200
Magnesium	6010	0.110	0.300
Potassium	6010	0.050	0.150
Sodium	6010	0.100	0.300

(a) Method references are listed in Appendix C.

(b) The detection limit is the lowest concentration that can be measured with 99% confidence. When samples are diluted, the detection limit scales with the dilution factor (e.g., a 10-fold dilution results in a detection limit that is 10 times higher).

(c) The reporting limit is an estimate of the lowest concentration that can be reliably quantified.

### 3.4 Quality Control for Water Samples

Three types of QC samples were submitted to GEL during the two 2021 sampling campaigns to assess sampling and analysis performance. Duplicates, blanks, and blind standards provide measures of reproducibility, contamination, and accuracy, respectively. In all cases, the QC samples were submitted to the analytical laboratory in double-blind fashion (i.e., they were disguised as regular monitoring samples). At the laboratory, an additional level of QC was added by randomly selecting samples to be analyzed in duplicate and by spiking samples to calculate recovery values for measured parameters. Table 3.3 and Table 3.4 list all QC samples sent to the laboratory for analysis and present the analytical results. Appendix C lists the analytical methods used for water analysis.

Analytical results received from GEL were evaluated using Microsoft Excel<sup>®</sup> and Access.<sup>1</sup> Ion balances were calculated on the spreadsheets from the analytical results reported by the laboratory. If an ion balance deviated by more than  $\pm 10\%$ , the analytical laboratory attempted to identify the problem. PNNL scientists reviewed the data for questionable results and typographical errors. GEL was contacted to verify any questionable results. The data were then compared with analytical results from previous sampling trips (Olsen et al. 1987, 1991–2006; McDonald et al. 2007; Thompson et al. 2008–2021). These data were also compared with their means, and a statistical outlier identification procedure was applied (see Appendix G, p. G.2). Any questionable data that could not be reconciled with the analytical

<sup>1</sup> Access and Excel are registered trademarks of Microsoft Corporation.

laboratory were still entered into the database but flagged as questionable and omitted from all statistical tests, as discussed in Section 3.5. Extremely low or high data identified as outliers are listed in Appendix G, Table G.2. Such outliers are errantly extreme data whose causes are unknown but are often observed to occur randomly. Outliers may also occur systematically but are less common. For example, recent data from samples at GOW 12 were omitted from the statistical analysis because they were determined to be systematically low and are not representative of the water in Pony Creek's alluvium. The causes of the extreme values are not considered to be laboratory or sampling related—the data are thought to be representative of the water being sampled—but are instead suspected to be attributable to changes in the water influx source.

The results of the QC samples from the June 2021 sampling trip are summarized in Table 3.3. Two field blanks (GKW 3-74 and GKW 4-74) were collected to evaluate the potential for sample contamination from the bottles and sampling equipment. The bottles used for the blanks were the same type as those used for routine monitoring samples. Results from analysis of all the blanks did not identify any significant contamination for the constituents measured.

Duplicate samples were used to assess the variation in laboratory analysis results (i.e., precision). The values for each constituent or parameter were compared by computing the relative percent difference (RPD), which is the absolute value of the difference between the results divided by their average. RPDs not exceeding 20% are generally considered acceptable if the analytical results are at least five times larger than the laboratory analysis method detection limit (MDL). At levels below five times the detection limit, higher RPDs may occur and be acceptable. The results for duplicate samples (GKW 1-74 and PW 736-74; GKW 2-74 and GSP 2-74) demonstrated good reproducibility for all measured parameters (RPDs were <10%), indicating acceptable sampling and analytical precision.

Laboratory accuracy was evaluated using blind standards, which are samples that contain known levels of constituents. For the June 2021 sampling event, two blind standards were submitted: a sample consisting of a NSI Lab Solutions standard, QC-136 (Minerals QC CRM; sample GKW 5-74), and two Environmental Resources Associates (ERA) standards, WasteWatr™ Trace Metals and Minerals (sample GKW 6-74). The laboratory results were compared with the most probable value (MPV) for each constituent by computing the percent error, which is the measured value minus the MPV expressed as a percentage of the MPV. Results within 20% are generally considered acceptable if the analytical results are at least five times the detection limit. GEL demonstrated strong performance on the blind standards—all results were acceptable, and most of the error percentages were <10%.

Table 3.3. Analytical results for QC samples from the June 2021 sampling trip.

Sample	Date Sampled/ Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (µS/cm)	Br <sup>-</sup> (ppm)	Cl <sup>-</sup> (ppm)	SO <sub>4</sub> <sup>2-</sup> (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GKW-1-74 (Field Duplicate)	16-Jun-21	366	7.48	3060	<0.067	14.9	1550	2710	1.16	221	230	11.8	188	38.43	40.01	-4.04
PW-736-74	16-Jun-21	359	7.47	3060	<0.067	14.9	1550	2710	1.16	225	235	12.0	188	39.04	39.87	-2.10
RPD	—	1.9	NC	0.0	NC	0.0	0.0	0.0	0.0	1.8	2.2	1.7	0.0	—	—	—
GKW-2-74 (Field Duplicate)	16-Jun-21	431	7.65	4610	0.163	27.1	2700	4480	2.04	275	414	22.9	360	64.02	65.60	-2.43
GSP-2-74	16-Jun-21	430	7.59	4660	0.169	26.3	2890	4430	1.99	286	425	22.4	365	65.68	69.51	-5.66
RPD	—	0.2	NC	1.1	NC	3.0	6.8	1.1	2.5	3.9	2.6	2.2	1.4	—	—	—
GKW-3-74	Bottle Blank	1.45	6.08	3.42	<0.067	0.732	0.447	28.6	<0.015	<0.05	<0.11	<0.05	0.578	NC	NC	NC
GKW-4-74	Equip. Blank	1.99	6.06	2.85	<0.067	0.392	0.333	7.14	<0.015	<0.05	<0.11	<0.05	0.552	NC	NC	NC
GKW-5-74	17-Jun-21	58.8	9.23	424	<0.067	41.6	41.0	291	24.9	<0.05	<0.11	33.2	56.2	NC	NC	NC
NSI Lab Solutions Standards	MPV	59.5	NR	391	—	46.3	41.0	259	—	—	—	—	—	NR	NR	NR
Percent Error	—	1.2	NC	8.4	NC	10.2	0.0	12.4	NC	NC	NC	NC	NC	—	—	—
GKW-6-74	17-Jun-21	123	9.24	527	<0.067	44.1	30.9	436	0.998	<0.05	<0.11	<0.05	0.138	NC	NC	NC
ERA Standards	MPV	127	NR	483	NR	48.7	32.8	415	0.938	NR	NR	NR	NR	NR	NR	NR
Percent Error	—	3.1	NC	9.1	NC	9.4	5.8	5.1	6.4	NC	NC	NC	NC	—	—	—

ERA = Environmental Resources Associates; GKW = QC sample; GNW = Genie new well; GSP = Genie spring; GSW = Genie surface water; MPV = most probable value; NC = not calculated; NR = not reported; STD = standard; TDS = total dissolved solids.

Table 3.4. Analytical results for QC samples from the October 2021 sampling trip.

Sample	Date Sampled/ Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (µS/cm)	Br <sup>-</sup> (ppm)	Cl <sup>-</sup> (ppm)	SO <sub>4</sub> <sup>2-</sup> (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GKW-1-75																
(Field Duplicate)	06-Oct-21	426	7.46	4890	0.110	21.2	2640	4620	1.88	269	383	20.0	416	63.54	64.08	-0.86
PW-735-75	06-Oct-21	422	7.44	5000	0.150	21.4	2690	4620	1.88	272	381	19.9	416	63.52	65.05	-2.38
RPD	—	0.9	NC	2.2	30.8	0.9	1.9	0.0	0.0	1.1	0.5	0.5	0.0	—	—	—
GKW-2-75																
(Field Duplicate)	06-Oct-21	400	7.55	3010	0.0677	11.3	1420	2500	1.12	220	213	13.3	177	36.54	37.88	-3.61
GSP-1-75	06-Oct-21	404	7.67	2750	<0.067	11.3	1380	2430	1.09	212	205	13.0	170	35.17	37.13	-5.42
RPD	—	1.0	NC	9.0	—	0.0	2.9	2.8	2.7	3.7	3.8	2.3	4.0	—	—	—
GKW-3-75	Equip. Blank	<1.45	5.92	6.98	<0.067	<0.067	<0.133	8.57	<0.015	<0.05	<0.11	<0.05	0.135	NC	NC	NC
GKW-4-75	Bottle Blank	2.00	5.96	1.10	<0.067	<0.067	<0.133	3.40	<0.015	0.097	<0.11	0.096	0.100	NC	NC	NC
GKW-5-75	07-Oct-21	50.0	9.17	424	<0.067	41.4	40.3	289	22.9	<0.05	<0.11	33.9	54.6	NC	NC	NC
NSI Standards	MPV	59.5	NR	391	NR	46.3	41.0	259	NR	NR	NR	34.0	63.1	NR	NR	NR
Percent Difference	—	16.0	NC	8.4	NC	10.6	1.7	11.6	NC	NC	NC	0.3	13.5	—	—	—
GKW-6-75	07-Oct-21	126	9.24	495	<0.067	46.2	20.9	420	0.860	0.280	0.455	0.163	0.127	NC	NC	NC
ERA Standards	MPV	124	NR	448	NR	47.9	20.7	399	0.893	NR	NR	NR	NR	NR	NR	NR
Percent Difference	—	1.6	NC	10.5	NC	3.5	1.0	5.3	3.7	NC	NC	NC	NC	—	—	—

ERA = Environmental Resources Associates; GKW = QC sample; GNW = Genie new well; GOW = Genie old well; MPV = most probable value; NC = not calculated; NR = not reported; PW = private monitoring well; STD = standard; TDS = total dissolved solids.

A summary of the results of QC samples from the October 2021 sampling trip is presented in Table 3.4. No constituents were detected at significant levels in the equipment or bottle blanks (GKW 3-75 and GKW 4-75, respectively). The duplicate samples collected at PW 735 and GSP 1 demonstrated excellent sampling and analysis reproducibility with all RPDs <10% except for a pair of bromide results from PW 735 that were within a factor of 2 of the MDL. Two blind standards identical to the blind standards that were submitted in June were submitted to the laboratory with the October 2021 samples. The first was the standard from NSI Lab Solutions (GKW-5-75) and the second consisted of two ERA standards (GKW-6-75). All results for the blind standards were acceptable.

A small number of samples from both sampling campaigns were reanalyzed because of potential data concerns. For example, the charge imbalances were unusually high for samples from GNW 6 and GNW 7. Based on historical data from these wells, it appeared that the metals results for the two wells were swapped. The analytical laboratory reanalyzed the samples for metals and obtained similar results. Therefore, a second reanalysis was performed using aliquots from the unpreserved, 1-L sample bottles that are used for other analyses (alkalinity, anions, pH, conductivity, TDS). Results from this second reanalysis agreed much better with historical values and resulted in a reasonable cation-anion charge balance. Consequently, we suspect that the sample bottles for metals were mislabeled or swapped. Another anomaly in the June data set was an extremely low TDS result for sample GKW-2-74 (duplicate of the sample from GSP 2). GEL reanalyzed the sample and obtained a concentration of 4480 mg/L, which agreed well with the duplicate sample result from GSP 2 (4430 mg/L). For the October data set, the sample from GSW 1 was reanalyzed for metals due to a high charge imbalance. The reanalysis results had a significantly lower charge balance error. All of the “improved” results from both sampling campaigns were loaded into the database with comments indicating that the results were from a reanalysis .

Agreement between field and laboratory pH and conductivity readings was reasonable for both sampling campaigns this year. On average, the field pH values were approximately 0.15 pH units lower than the laboratory values. Some of the discrepancies may be due to chemical changes in the samples that occur between the time the measurements are made in the field and at the laboratory (e.g., outgassing of samples from deeper wells and redistribution of inorganic carbon). Most of the conductivity measurements made in the field and at the laboratory agreed within 15%. Notable exceptions in June were pairs of values from GAS 3 (2120  $\mu\text{S}/\text{cm}$  in the field and 2860  $\mu\text{S}/\text{cm}$  at the laboratory), and GSP 3 (1540  $\mu\text{S}/\text{cm}$  in the field and 2030  $\mu\text{S}/\text{cm}$  at the laboratory). In October, there was a large discrepancy in the samples from GSW 2 (430  $\mu\text{S}/\text{cm}$  in the field and 5510  $\mu\text{S}/\text{cm}$  at the laboratory). The reason(s) for the poor agreement at these sites is unknown, but the laboratory values are consistent with historical data. Field data are not used for statistical trend detection.

### **3.5 Statistical Analysis of Chemical Data**

The primary objectives of the 2021 hydrologic/water-quality study were to determine whether there were changes in water quality in the Cow Creek, South Fork Cow Creek, and Pony Creek drainages from the previous year, and to establish the current characteristics and conditions of water quality within the three drainages studied. Stipulation 12(d) identifies two parameters that indicate leakage: conductivity and boron concentration. In addition, high sulfate values are considered to provide a signature of leakage from the EHP. To help meet the study objectives, a statistical analysis was performed on the water-quality data, focusing on these three indicators. The objectives of the statistical analysis are described in the next two sections.

### **3.5.1 Objective 1: Identify Any Time Trends at the Site or Drainage Level in the Parameters Conductivity, Sulfate, and Boron**

The *site-level* analysis is performed to detect any consistently increasing trends at individual sites; this is of primary importance for decision-making. Stipulation 12(d) mandates that if an abrupt increase occurs in conductivity or boron at any well between the Talen Montana/Genie Land Company property line and Stinking Spring (GSP 2) or at any point where an adequate interception system could be constructed, Talen Montana would intercept the reduced-quality water. Therefore, the site-level trend analyses are primarily for detecting and quantifying any increasing trends at the Cow Creek or South Cow Creek sites. The trend analyses for sites at the Pony Creek drainage are useful for comparing and verifying any detected trends at South Cow Creek or Cow Creek. For example, the Pony Creek analysis can help discount that any increasing trends at the other drainages are due to changes in regional hydrogeochemical properties or weather-related events. Pony Creek is hydrologically and geochemically similar to both Cow Creek drainages and is presumed to be unaffected by EHP activities. Therefore, any time trends detected at Pony Creek sites are expected to result from natural changes in drainage flow conditions. However, active mining activity is under way on the north and south sides of Pony Creek approximately 0.5 mi upgradient of GSP 4, and local hydrogeological conditions may differ somewhat between the Cow and Pony Creek drainages because of ongoing activities or natural differences. Alternatively, if transport of EHP effluent by storm-related creek flow events occurred at Cow Creek and South Fork Cow Creek, nearly simultaneous trends at multiple sites could ensue in both drainages. The possibility of contamination by this pathway should therefore be ruled out first before concluding that all increasing trends are hydrogeochemical or weather related.

Our site-level analysis also tests for decreasing trends that may indicate plume movement and reduction of key parameters through the drainage. Under accidental-release conditions that produce a distinct plume, the concentrations of selected parameters at an affected site will eventually decrease as the plume moves further downgradient. Therefore, identification of recent decreases can be useful for characterizing the extent of contamination and for providing insights about the rate of contaminant migration. Decreasing trends may also be caused by naturally changing aquifer characteristics or climatological factors.

In addition to site-level time-trend analyses, *drainage-level* time trends for each of the three indicator parameters are assessed using data from several sites in each drainage. Drainage-trend estimates exhibit considerable uncertainty but are more suitable for qualitatively depicting mass-balance changes within a drainage. Further discussion of drainage trends is provided in Appendix G, Section G.3.

### **3.5.2 Objective 2: Estimate and Compare Overall (Time-Averaged, Site-Averaged) Drainage-Mean Values for Cow Creek, South Fork Cow Creek, and Pony Creek**

Average constituent levels at the drainages are evaluated as part of an ongoing characterization effort. Pony Creek represents a control drainage, although hydrogeologic conditions in Pony Creek are similar but not identical to those in the Cow Creek and South Fork Cow Creek drainages.

### **3.5.3 Sites and Data Used in the Statistical Analyses**

The sites used for the statistical analysis are identified in Table 3.5.

Table 3.5. Wells and springs used in the statistical analysis.

Cow Creek	Pony Creek	South Fork Cow Creek
PW 735 <sup>(a)</sup>	GSP 4	GSP 3
PW 736	GSP 6	GOW 11
GNW 1	GOW 3	GNW 6
GSP 1	GOW 4	GNW 7
GNW 2	GSP 7 <sup>(b)</sup>	GNW 8
GNW 3	GOW 12 <sup>(b)</sup>	--
GOW 1	GSP 8	--
GNW 4	--	--
GSP 2	--	--
GNW 5	--	--

(a) Used for site-level trend analysis only.

(b) Past data used in drainage-level analyses only.

GNW = Genie new well; GOW = Genie old well; GSP = Genie spring; PW = private monitoring well installed by Hydrometrics, Inc.; GSP 2 also is known as Stinking Spring.

In general, these sites are the same as those examined by Olsen et al. (1987, 1991–2006), McDonald et al. (2007), and Thompson et al. (2008–2021). Exceptions include the addition of wells PW 735 and 736, and the exclusion of GOW 12 (after sampling event 60). A new well was installed at GOW 12 in 2014, and the water chemistry of the new well is distinct (e.g., much lower sulfate) from that of the former well. The statistical site analysis will resume inclusion of GOW 12 if these anomalies resolve and after sufficient data for reliable trending have been collected for the new well.

The ordering in Table 3.5 corresponds to the distance of each site from the headwaters (closest first). These wells and springs for sampling groundwater were selected as the ones most likely to yield results representative of the aquifer in question. Site W-1 was not included because it is downstream from the confluence of Cow Creek and South Fork Cow Creek; therefore, any effects at site W-1 are not uniquely attributable to either of these drainages. This precludes its usefulness in identifying differences in drainage characteristics, either temporal or between drainages. Any site trend at W-1 is also of no value for site-based inferences of drainage conditions. However, W-1 data could be useful for other investigations that are currently outside the scope of this report. For example, W-1 could be used to represent a “composite” of Cow and South Fork Cow drainages, in a comparison with selected downgradient sites at Pony Creek. Therefore, W-1 continues to be included in the sampling campaigns.

Data from sampling years 1985–1987 and 1990–2021 were examined. Statistical analyses were performed for conductivity ( $\mu\text{S}/\text{cm}$ ), boron (B, ppm), and sulfate ( $\text{SO}_4^{2-}$ , ppm). Data from samples obtained in 2021 are listed along with other parameters in Appendix F, and the data from previous years are provided in earlier reports (Olsen et al. 1992–2006; McDonald et al. 2007; Thompson et al. 2008–2021).

Outliers in the data were identified and excluded from all statistical analyses as described in Appendix G, Section G.1.2. In this year’s outlier investigation, two sulfate values were declared to be outliers—at GOW 3 (sampling event 10) and at GOW 1 (sampling event 11)—because the concentrations were inordinately low and inconsistent with the adjacent years’ data. Recent results from GSP 7 continue to be retained in the analysis because those results appear to be more representative than data from 2017–2018 based on conditions observed at that site during the sampling trips. The data from GOW 12 after sampling

event 60 continue to be omitted from the drainage time trend and drainage comparisons because their values are unrepresentative of the Pony Creek aquifer.

Outlying data were kept in the database but were omitted from the site- and drainage-level time-trend analyses and drainage comparisons. Outliers excluded as a result of previous outlier analyses are provided in earlier reports (e.g., Thompson et al. 2008–2021), and all cumulatively excluded data are identified in Appendix G, Table G.2.

### **3.5.4 Testing Methods**

Two statistical tests were used to objectively evaluate whether changes in the indicator parameters occurred at individual sites and at the drainages overall during the most recent four years. A cursory description of both methods is provided below; Appendix G includes more detailed information. For Stipulation 12(d) parameters boron and conductivity, the two tests are evaluated as a composite test—a recent increase is concluded if either test determines that an increase occurred.

All conclusions regarding temporal changes and in the comparisons of drainages are based on classical statistical testing, which *assumes a default hypothesis and only abandons that hypothesis if the observed data are highly incompatible with that condition*. The default hypotheses in our analyses are 1) there is no increasing or high-valued trend (and no decreasing trend when testing for decreases), and 2) in the drainage-comparison tests, the drainage constituent means are all equal.

Testing for increases involves assessing how likely it would be to observe our measured data values if the underlying trend curve is actually not increasing. To do this, we generate a large number of simulated data sets under that assumption (e.g., each having a non-increasing trend), evaluate each one using the test criteria described below, and then tally the percentage that show an increase. If this percentage is very low, then we conclude that an underlying increasing trend exists (i.e., the data are inconsistent with the no-trend hypothesis).

#### **3.5.4.1 Conventional Test**

Our “conventional” test has been used for most of the duration of this monitoring program. The test is used to detect both increasing and decreasing trends for all three indicator parameters at individual sites and at each drainage. In brief, the conventional test uses a curve-fitting algorithm to determine whether a steady increase or decrease occurred over the most recent four years. If the recent monitoring data exhibit a change, the curve-fitting algorithm is applied to a large population (100,000) of simulated data sets (each based on the assumption of no recent steady trend with randomized sampling and analysis errors added), and if fewer than 0.07% (0.15% for composite testing) of the data sets exhibit the same type of change (increasing or decreasing) as the original monitoring data, then a trend is concluded.

#### **3.5.4.2 PLR Test**

The Predictive Likelihoods Ratio (PLR) test has been used to assess trends in this program since 2017. This test is applied to detect recent increases in conductivity and boron at the site level only. In the PLR test, a large population (1,000,000) of data sets is generated based on a nonincreasing trend with random measurement errors added. These data sets represent scenarios that could occur if the true trend was not increasing—each data set consists of a randomly selected, non-increasing trend with random measurement errors added. For each data set, a PLR statistic is calculated that is the probability that the data represent an increasing trend divided by the probability that the data belong to the population of data sets with no increasing trend. This statistic tends to be larger when there actually is an increasing trend. Consequently, if fewer than 0.16% of the PLR values exceed the PLR computed from the original data,

we conclude that there is an increasing recent trend (i.e., it is unlikely, under a non-increasing situation, to observe a PLR as large as the one computed from the actual monitoring data).

Performance-wise, the PLR test is far superior to the conventional test because the PLR test has greater detection ability when actual trends exist, and the PLR test's false-detect rate is exactly controllable. The conventional test is nevertheless retained in our analysis as part of the composite test, for all other site trend testing and for drainage-trend tests. The conventional test may perform at its best for longer detected trends and its results often may be visually verified when the data vary only slightly around an increasing estimated trend.

### **3.5.4.3 Evaluation of Drainage Means**

Drainage averages are evaluated by comparing box plots. Medians and 95% confidence intervals are used for estimating and comparing overall drainage means. Appendix G, Section G.3.5 provides more information about the drainage comparisons.

## **3.5.5 Results of Statistical Analysis**

This discussion of results is based on the time-trend and drainage-comparison analyses described in Appendix G of this report. More technical foundations for the methods are provided by Thompson et al. (2014) and Chamberlain (2018).

### **3.5.5.1 Assessment of Drainage Data over Time**

For each constituent, the average of non-outlying data among all sites within each drainage for 1985-1987 and 1990–2021 was examined for strict trends over the most recent four or more years (see Appendix G, Figure G.8). Drainage averages from each campaign were used as individual observations. As indicated in Table 3.6, neither increasing nor decreasing trends were detected at any of the drainages.

### **3.5.5.2 Assessment of Sites over Time**

We examined sampling data for each site for 1985–1987 and 1990–2021 for 1) increasing time trends in conductivity or boron (using the composite test discussed earlier and in Appendix G, Section G.3.3), 2) decreasing trends at any site, and 3) decreasing or increasing trends in sulfate at any site. All trends were defined as continual (steadily increasing or decreasing) over the last four or more years. Test results are given in both Figure G.9 and in the overlaid time plots of the sites in Appendix G, Figure G.10. Summarized results of these trends also are listed below in Table 3.6.

Increasing trends in conductivity were found at two sites along Pony Creek (GOW 4 and GSP 4) and at one Cow Creek site (PW 735). The latter site taps a deeper confined aquifer than that supplying the alluvial wells. Increasing trends for boron were concluded at one Pony Creek site (GSP 8) and two Cow Creek sites (PW 735 and GSP 2). Additionally, a decreasing trend for boron was found at Pony Creek site GOW 4. No sulfate trends, either increasing or decreasing, were found at any site.

Table 3.6. Time trends detected in 2021.<sup>(a)</sup>

Site	Conductivity ( $\mu\text{S}/\text{cm}$ )	B (ppm)	Sulfate (ppm)
Cow Creek	No decreasing site trends	No decreasing site trends	No site trends
PW 735	$+$ <sup>(a,b)</sup>	$+$	
GSP 2		$+$	
Drainage level	No drainage trend	No drainage trend	No drainage trend
South Fork Cow Creek	No site trends	No site trends	No site trends
Drainage level	No drainage trend	No drainage trend	No drainage trend
Pony Creek	No decreasing site trends		No site trends
GOW 4	$+$	$-$ <sup>(c)</sup>	
GSP 4	$+$		
GSP 8		$+$	
Drainage level	No drainage trend	No drainage trend	No drainage trend

- (a) Blue text indicates composite test results (tests for increases in conductivity and in boron; all sites), and black text conveys standard test results (remaining site test combinations and drainage tests).  
 (b) “+” indicates composite test concluded an increasing time trend (5% level of false detection error among all composite tests in the drainage).  
 (c) “-” indicates standard trend test found a decreasing time trend (total 5% level of false-detection error among all non-composite trend tests, including drainage-level tests).  
 GNW = Genie new well; GOW = Genie old well; GSP = Genie spring; PW = Private monitoring well.

### 3.5.5.3 Comparisons between Drainages

Data for all sites within a drainage were pooled across time so the drainage means could be compared (Appendix G, Figure G.11). The estimated drainage means and their 95% confidence intervals are provided in Table 3.7. For the overall observation period (1985–1987 and 1990–2021), the mean values for conductivity and sulfate in the Pony Creek drainage are less than the values for the South Fork Cow Creek drainage, both of which are less than those in the Cow Creek drainage (Table 3.7). For boron, the mean values in both the Cow Creek and South Fork Cow Creek drainages are not statistically different, but each value is greater than the mean in the Pony Creek drainage.

Table 3.7. 95% confidence limits on drainage-mean values using all data.<sup>(a)</sup>

Parameter	Cow Creek	Pony Creek	South Fork Cow Creek
Conductivity ( $\mu\text{S}/\text{cm}$ )	3400 $\pm$ 66	2000 $\pm$ 97	2600 $\pm$ 117
Boron (ppm)	1.40 $\pm$ 0.02	0.60 $\pm$ 0.05	1.37 $\pm$ 0.06
Sulfate (ppm)	2000 $\pm$ 41	770 $\pm$ 63	1240 $\pm$ 89

(a) 1985–1987 and 1990–2021 data

### 3.5.6 Overall Conclusions of the Statistical Analysis

The *drainage time-trend test results* are:

- No recent drainage trends were found at Pony Creek, Cow Creek, or South Fork Cow Creek.

The *site-level time-trend test results* are as follows:

- Increasing trends were found at Cow Creek sites PW 735 (conductivity and boron) and GSP 2 (boron).

- The recent increase in boron at GSP 2 has been gradual (not abrupt), rising by no more than 36% over eight sampling campaigns. This increasing behavior is unique to GSP 2 among Cow Creek well sites in the study area. A separate composite test found no increasing boron trend at upgradient surface-water site GSW 5. Additionally, a separate calculation for the distribution of all possible boron data that could have occurred recently at GSP 2 (given that the underlying trend is increasing) indicates a range of 1.6–2.5 ppm. This range reaches from approximately the midpoint to slightly higher than historic data.
- Increasing trends in conductivity were detected for Pony Creek sites GSP 4 and GOW 4, and decreasing and increasing trends in boron were found at Pony Creek sites GOW 4 and GSP 8, respectively.

Drainage comparisons indicated that the Pony Creek drainage had the lowest values for conductivity, sulfate, and boron. The Cow Creek drainage was distinctly highest in conductivity and sulfate, and the South Fork Cow Creek drainage means fell between the values of the other drainages except for boron, whose mean did not differ from that of Cow Creek.

### 3.6 Results and Discussion

This section summarizes the monitoring data collected during 2021 and is organized into subsections for monitoring locations (alternate supply wells and the three creek drainages), bromide measurements, and continuous-monitoring data at well PW 736. Recent trends identified in the preceding section are interpreted in terms of their significance, based on the magnitude of change relative to previous observations.

Drainage profiles of conductivity, sulfate, and boron are plotted in Figure 3.4, Figure 3.10, and Figure 3.12, respectively, in ensuing sections. These parameters were selected because they are likely indicators of leakage from the Colstrip Units 3 and 4 EHP. For reference, Table 3.8 lists composition data from a sample that was collected from the EHP in 2015. Drainage-mean values of conductivity, sulfate, and boron are plotted for each drainage in Figure 3.9, Figure 3.11, and Figure 3.16 (in ensuing sections) as a function of time from 1985 through 2021. Piper diagrams illustrating the geochemical water types of the GAS wells and sites along the three creek drainages were presented in the 2010 report (Thompson et al. 2011). Because the water chemistry has not changed appreciably in the past 10 years, those plots were not updated for this report.

Table 3.8. Composition of EHP solution (based on a sample collected in May 2015).

Sample Site	Alkalinity (ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br <sup>-</sup> (ppm)	Cl <sup>-</sup> (ppm)	SO <sub>4</sub> <sup>=</sup> (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)
3+4 EHP B-CELL	<4	3.0	32,200	1940	1,160	35,600	34,400	278	560	8,070	132	2,660

The temporal profiles that follow are used qualitatively to support the discussion of the analytical data; for quantitative estimates of drainage time trends and their corresponding tests of statistical significance, refer to Appendix G, Figure G.8. To summarize laboratory results, except where noted, outlying data (as identified in Appendix G, Section G.1.2) were not removed in the calculations of these profiles.

There were no reportable releases from the Colstrip Units 3 and 4 EHP area during 2021.

The site-specific average values were used to generate drainage-mean values (see Appendix E) in each of the tables for Cow Creek, South Fork Cow Creek, and Pony Creek. The individual analytical results for each sample collected during 2021 are listed in Appendix G. All less-than (<) values were ignored when

an average value was calculated for a site. When only a single value was present, no mean value was calculated.

### 3.6.1 Sampling Data from the Genie Alternative Supply Wells

The GAS wells are located in the Cow Creek and South Fork Cow Creek drainages. These wells were installed to provide alternative supplies of water for stock in lieu of surface and upper unconfined groundwater that has elevated sulfate concentrations. These wells are being monitored to ensure that their water quality is not being negatively affected by potentially poor-quality alluvial groundwater within the drainages. Because of their depth, they have the potential to tap deeper aquifers than the GNWs or GOWs in the drainages. Chemical data were used to classify the GAS wells according to their possible water sources. Initial geochemical description of the GAS wells was presented by Olsen et al. (1991) and was updated in the 2010 report (Thompson et al. 2011). Briefly, GAS 1 taps what is suspected to be a shallow alluvial aquifer that contains a magnesium/calcium sulfate and sodium/potassium sulfate mixed-water system very similar to that of GOW 1 and the GNW wells within the Cow Creek drainage. GAS 1 is not being used as an alternate supply well because its water chemistry is similar to that of the GNW wells within the Cow Creek drainage. The geochemistries of water from GAS 2 and GAS 3 are somewhat similar. GAS 2 taps an aquifer containing a sodium sulfate/sodium bicarbonate mixed-water system, and GAS 3 draws water from a sodium sulfate aquifer system.

Wells GAS 1, GAS 2, GAS 3, and GAS 7 were sampled in June and October 2021. The mean analytical values for conductivity, sulfate, and boron from water samples collected in the GAS wells are provided in Table 3.9.

Table 3.9. Values for conductivity, sulfate, and boron in samples collected from GAS wells in 2021.

Site	Comment	Conductivity ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^{2-}$ (ppm)	Boron (ppm)
GAS 1	Mean value	2825	1285	0.972
GAS 2	Mean value	1695	428	0.311
GAS 3	Mean value	2785	1008	0.287
GAS 7	Mean value	2040	594	0.297

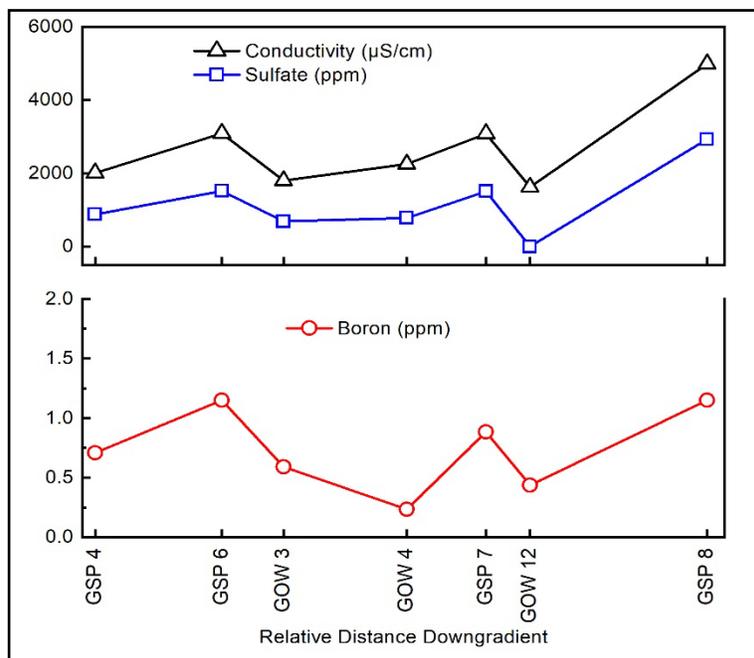
Among the GAS wells, GAS 1 had the highest values for all the indicator parameters. GAS 3 had similar but slightly lower values for conductivity and sulfate (2785  $\mu\text{S}/\text{cm}$  and 1008 ppm, respectively), while the lowest values for both parameters were found at GAS 2 (1695  $\mu\text{S}/\text{cm}$  and 428 ppm, respectively). Boron concentrations ranged from a low of 0.287 ppm at GAS 3 to a high of 0.972 ppm at GAS 1. Boron concentrations at GAS 1 were approximately three times higher than those at the other GAS wells. These relative patterns were essentially unchanged from those of the past 10 years. The overall stability of these parameters and their relatively low values indicate that the GAS wells continue to be unaffected by anthropogenic activities. All chemical data from the GAS wells can be found in Appendices E and F.

### 3.6.2 Sampling Data from Pony Creek Drainage

The Pony Creek drainage contains 12 sampling points: 6 springs, 4 GOW stock-watering wells, and 2 GNW wells. All the sites except GNW 10 and GSP 5 were sampled during June 2021. Both GNW 10 and GSP 5 were dry during the June sampling campaign. Data from GNW 11 and GSP 10 are not representative of the upper alluvial aquifer in the Pony Creek drainage, so they were not included in the statistical analysis (Olsen et al. 1991). Additionally, the values from GOW 12 were omitted for samples after June 2014, because a new well was installed at that location that summer, and subsequent data suggest that the new well taps a deeper source with a different water type that is not representative of the

upper alluvial aquifer. If this anomaly resolves, future site-level trend analyses will include results from GOW 12 after sufficient data have been collected from the new well for reliable trend analyses.

Analytical results for the indicator parameters measured at sample sites in the Pony Creek drainage during 2021 are presented in Figure 3.4. Conductivity values ranged from 1620  $\mu\text{S}/\text{cm}$  at GOW 12 to 4990  $\mu\text{S}/\text{cm}$  at GSP 8, with a drainage-mean value of 2870  $\mu\text{S}/\text{cm}$ . Sulfate concentrations ranged from a non-detected result ( $<0.133$  mg/L) at GOW 12 to 2930 ppm at GSP 8, with a drainage-mean value of 1385 ppm; and boron ranged from 0.233 ppm at GOW 4 to 1.15 ppm at GSP 6 and GSP 8, with a drainage-mean value of 0.786 ppm.



Site	Cond ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^{2-}$ (ppm)	B (ppm)
GSP 4	2010	879	0.707
GSP 5*	—	—	—
GSP 6	3090	1520	1.15
GOW 3	1800	691	0.591
GOW 4	2250	781	0.233
GSP 7	3080	1510	0.885
GOW 12 <sup>†</sup>	1620	<0.133	0.436
GSP 8	4990	2930	1.15
2021 Drainage Mean	2870	1385	0.786

<sup>†</sup>Values for GOW 12 were not included in the drainage mean.

\*GSP 5 was dry during the June sampling campaign.

Figure 3.4. Plots and data for conductivity, sulfate, and boron in the Pony Creek Drainage during 2021. Values are single measurements from the June sampling event.

For 2016 and 2017, data from GSP 7 were not included in the trend analyses and drainage-mean calculations because the values were unusually high compared to historical averages and were formally confirmed to be outliers. This appeared to be caused by dry conditions and apparent evaporation at that

site (Thompson et al. 2018). Since the July 2019 sampling campaign, more water has been present and the concentrations have been more consistent with historical data. Trend testing for GSP 7 resumed in 2020.

Four site-specific time trends were detected in the Pony Creek drainage: an increase in conductivity and a decrease in boron at GOW 4, an increase in conductivity at GSP 4, and an increase in boron at GSP 8. These site trends have been observed in previous years. Figure 3.6, Figure 3.7, Figure 3.5, and Figure 3.8 show the cumulative trend plots for these sites and constituents.

Genie Spring 4, the most upgradient site on Pony Creek, had an increasing trend in conductivity. This trend has been detected at GSP 4 every year since 2008 except 2016, 2018, and 2020. The cumulative trend plot (Figure 3.5) shows strong 1–3 year fluctuations over the past 15 years that make it difficult to assess the significance of short-term behavior. These fluctuations may be related to the highly variable precipitation in the area since about 2001 (Figure 2.1). However, the average conductivity readings since 2006 are generally higher than previous values. We suspect this increase is associated with mining activity near the upper portions of Pony Creek, although specific causes have not been identified.

The increasing trend for conductivity at GOW 4 reflects the slightly higher values since 2018 (Figure 3.6). These changes are relatively minor compared to past fluctuations and do not definitively indicate encroachment of a contaminant plume at this site. Increasing trends for conductivity were also concluded at GOW 4 in 2010 and 2012 with the same interpretation—no obvious plume movement could be concluded based on the relatively small magnitude of the associated 4-year trends and the long-term variability in the entire data set.

A decreasing boron trend was also detected at GOW 4. The boron data plot for GOW 4 (Figure 3.7) shows relatively large swings between 1985 and 1995 and smaller fluctuations (up to  $\pm 0.1$  ppm) during the past 20 years. The estimated trend at this site (Figure G.9) exhibits a gradual decline over the entire study period and was concluded to be correctly decreasing. This is the seventh consecutive year in which a decreasing boron trend was detected at GOW 4. However, the recent decreasing trend is not indicative of plume passage at this site, because similar declines throughout the study period are also exhibited at sites GOW 3 and GOW 12, possibly suggesting insignificant influx of contaminated runoff in the vicinity of these sites prior to 1984 followed by gradual attenuation. Regardless, the boron levels at these sites are all very low. The 2021 boron concentration (0.233 ppm) is very similar to the values from 2016–2020 (0.216–0.237 ppm).

The increasing boron trend identified at GSP 8 is corroborated by the generally increasing concentrations since 2014 (Figure 3.8). An increasing boron trend at this site was also concluded last year. Boron levels ranged from approximately 0.60–0.85 ppm in the mid-1980s and have fluctuated between 0.80 and 1.60 ppm since 1991. The increases over the past four years (1.16–1.51 ppm) are well within the fluctuations observed previously at this location and the 2021 value of 1.15 ppm is approximately 25% lower than the concentration from 2020 (1.46 ppm). Therefore, no conclusions can be made about whether the increasing trend is indicative of plume passage at this site. No increasing or decreasing site trends for boron were detected at the adjacent upgradient location, GOW 5.

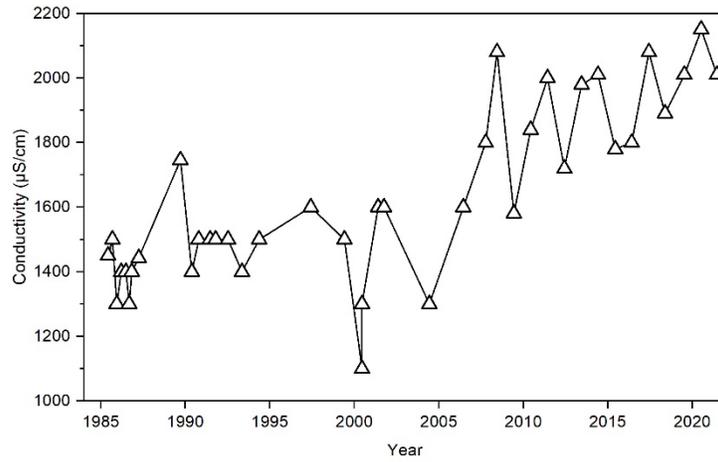


Figure 3.5. Trend plot for conductivity at GSP 4. Outliers identified in Appendix G, Section G1.2, were omitted from the graph.

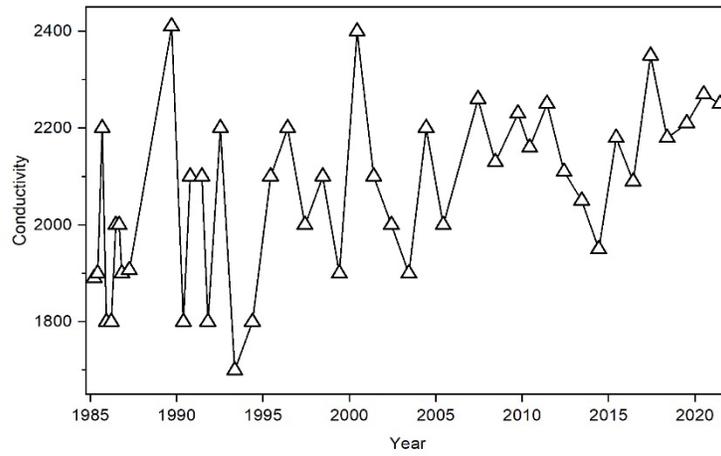


Figure 3.6. Trend plot for conductivity at GOW 4.

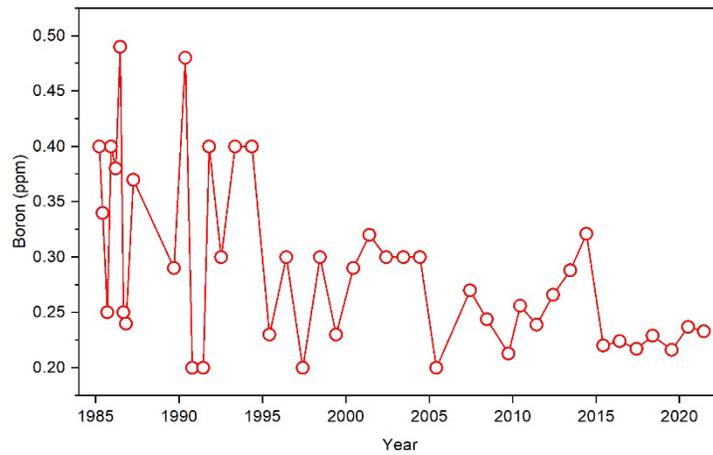


Figure 3.7. Trend plot for boron at GOW 4.

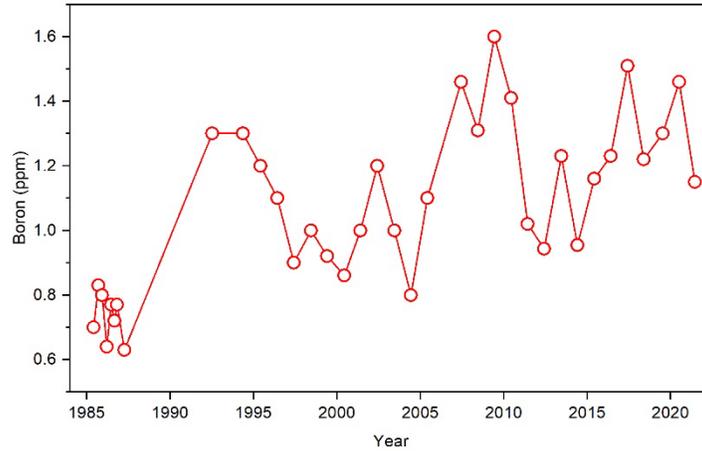


Figure 3.8. Trend plot for boron at GSP 8. Outliers identified in Appendix G, Section G.1.2, were omitted from the graph.

Comparing the drainage means over time (1985–1986 and 1990–2021; Figure 3.9 and Table 3.10) show that conductivity values and sulfate and boron concentrations are strongly correlated. Each parameter shows minor fluctuations superimposed on a gradually increasing tendency since about 1990. No drainage-level trends were detected for any of the parameters in 2021. As seen in the site-overlaid trend estimates in Appendix G, Figure G.10, the overall drainage increases in the parameters are possibly attributable to apparent increases at GSP 4, GSP 6, and GSP 7 (and GSP 8 for boron). In contrast to these spring sites, most of the Pony Creek well sites exhibit non-increasing trends (either flat or tilted downward) with very little variation. These differing features suggest the occurrence of either small influxes of contaminated runoff near the above springs and/or concentrating-diluting influences of evaporation and recharge from runoff.

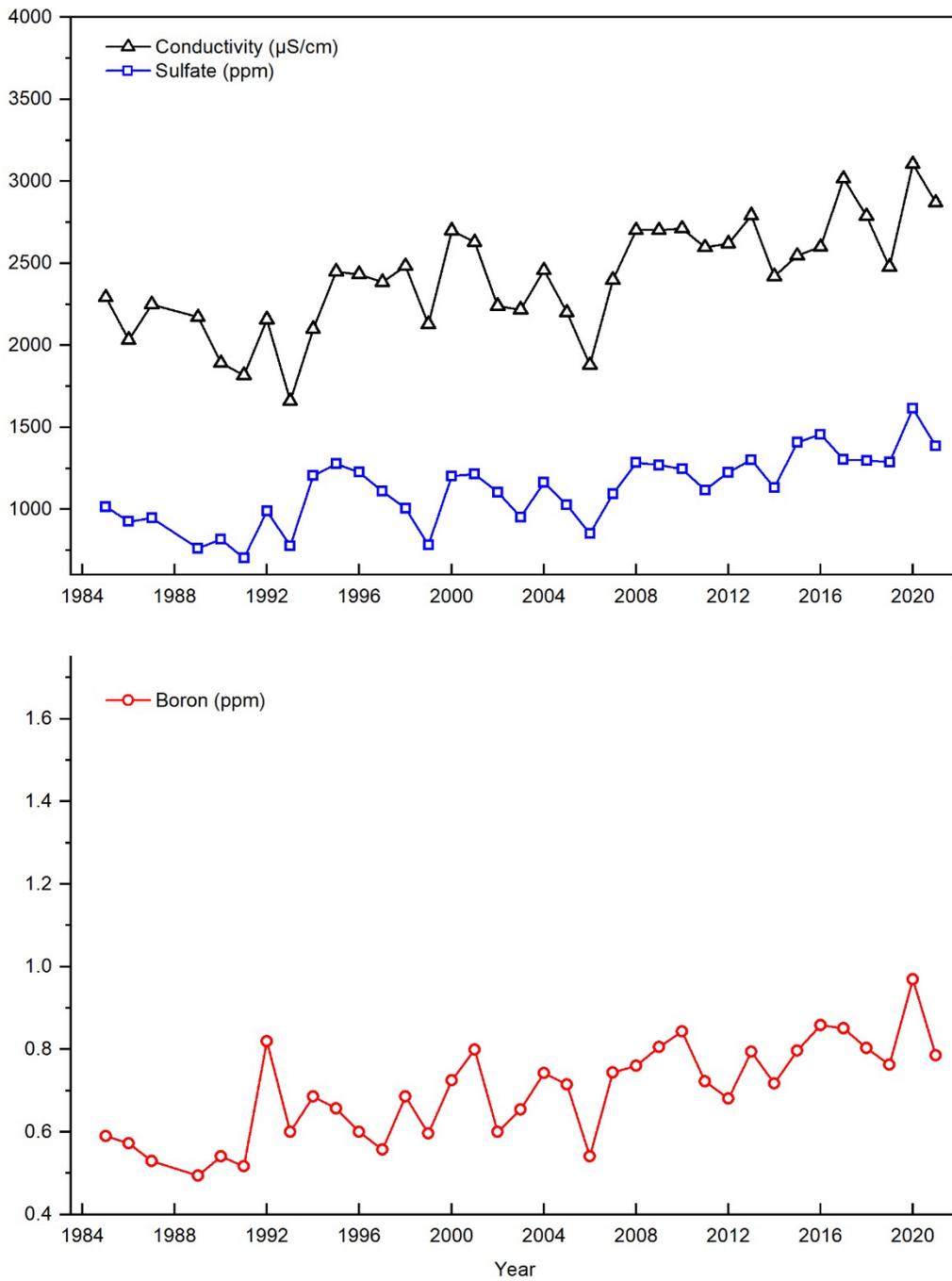


Figure 3.9. Trend plots for conductivity, sulfate, and boron in the Pony Creek Drainage, 1985–2021.

Table 3.10. Drainage-mean values for conductivity, sulfate, and boron in Pony Creek Drainage, 1985–2021.

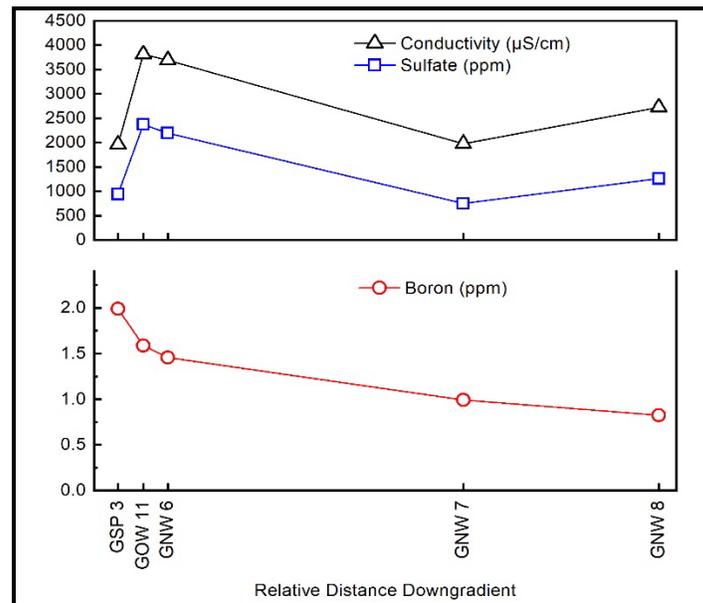
Year	Drainage Mean		
	Conductivity ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^-$ (ppm)	B (ppm)
1985	2292	1015	0.59
1986	2034	926	0.57
1987	2250	948	0.53
1989	2171	762	0.49
1990	1892	818	0.54
1991	1817	703	0.52
1992	2157	988	0.82
1993	1660	778	0.60
1994	2100	1205	0.69
1995	2450	1278	0.66
1996	2433	1227	0.60
1997	2386	1110	0.56
1998	2483	1006	0.69
1999	2129	783	0.60
2000	2700	1203	0.73
2001	2629	1215	0.80
2002	2240	1105	0.60
2003	2217	952	0.65
2004	2457	1163	0.74
2005	2200	1027	0.71
2006	1880	853	0.54
2007	2398	1093	0.74
2008	2701	1284	0.76
2009	2701	1269	0.81
2010	2712	1246	0.84
2011	2597	1117	0.72
2012	2620	1225	0.68
2013	2791	1302	0.79
2014	2421	1132	0.72
2015	2546	1409	0.80
2016	2600	1457	0.86
2017	3016	1302	0.85
2018	2788	1299	0.80
2019	2478	1288	0.76
2020	3104	1614	0.97
2021	2870	1385	0.79

### 3.6.3 Sampling Data from South Fork Cow Creek

There are five primary sampling sites within the South Fork Cow Creek drainage (GSP 3, GOW 11, GNW 6, GNW 7, and GNW 8) and one alternative supply well (GAS 3). GSP 3 is situated on a small tributary that feeds into the main flow of South Fork Cow Creek; the remaining four sites are situated along the main stem of South Fork Cow Creek. All six sites within the South Fork Cow Creek drainage were sampled during both the June and October 2021 sampling trips. Although GAS 3 is included in the South Fork Cow Creek drainage, water from that well is not representative of the upper alluvial aquifer. Therefore, GAS 3 was not included in the statistical analyses.

The 2021 analytical results for indicator parameters at sample sites in the South Fork Cow Creek drainage are presented in Figure 3.10. Values shown are the averages from the June and October sampling events. Conductivity averages ranged from 1960  $\mu\text{S}/\text{cm}$  at GSP 3 to 3810  $\mu\text{S}/\text{cm}$  at GOW 11, with a drainage-mean value of 2831  $\mu\text{S}/\text{cm}$ . Mean sulfate concentrations ranged from 748 ppm at GNW 7 to 2370 ppm at GOW 11, with a drainage-mean value of 1503 ppm. Mean boron concentrations ranged from 0.825 ppm at GNW 8 to 1.99 ppm at GSP 3, with a drainage-mean value of 1.37 ppm.

Figure 3.10 shows spatial trends for conductivity, sulfate, and boron in the South Fork Cow Creek drainage. Conductivity and sulfate were spatially correlated, with maxima near the upper end of the drainage (GOW 11), decreasing levels from GNW 6 to GNW 7, and increasing values from GNW 7 to GNW 8. Boron concentrations were highest at GSP 3 and decreased sequentially down the drainage from GSP 3 to GNW 8. No site-specific trends were detected in the South Fork Cow Creek drainage.



Site	Cond ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^{2-}$ (ppm)	B (ppm)
GSP 3	1960	940	1.99
GOW 11	3810	2370	1.58
GNW 6	3690	2195	1.46
GNW 7	1975	748	0.990
GNW 8	2720	1260	0.825
2021 Drainage Mean	2831	1503	1.37

Figure 3.10. Plots and data for conductivity, sulfate, and boron in South Fork Cow Creek Drainage during 2021. Values are averages from the June and October sampling events.

Figure 3.11 and Table 3.11 show the annual drainage means since monitoring began in 1985. No drainage-level trends were detected for South Fork Cow Creek this year. In 2020, an increasing conductivity trend was concluded, but the change was negligible in magnitude and did not coincide with any site-specific trends. Mean conductivity levels have generally increased since 1999, but the rate of change has been low. The 2021 average, 2831  $\mu\text{S}/\text{cm}$ , is similar to the average from 2010-2020 (2848  $\mu\text{S}/\text{cm}$ ), which indicates little overall change over the past decade.

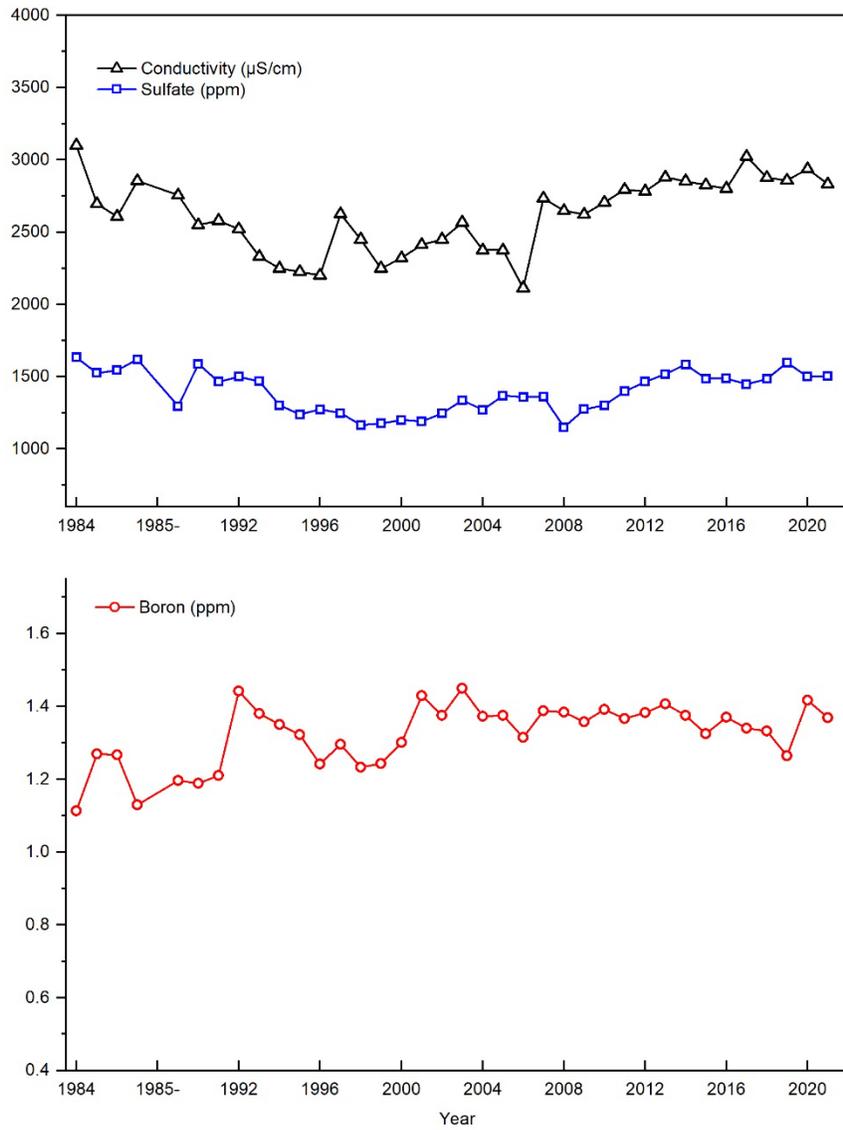


Figure 3.11. Trend plots for conductivity, sulfate, and boron in South Fork Cow Creek, 1984–2021.

Table 3.11. Drainage-mean values for conductivity, sulfate, and boron in South Fork Cow Creek, 1984–2021.

Year	Drainage Mean		
	Conductivity ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^{2-}$ (ppm)	B (ppm)
1984	3100	1633	1.11
1985	2695	1525	1.27
1986	2608	1544	1.27
1987	2853	1620	1.13
1989	2755	1296	1.20
1990	2550	1587	1.19
1991	2580	1465	1.21
1992	2520	1500	1.44
1993	2330	1468	1.38
1994	2250	1301	1.35
1995	2225	1239	1.32
1996	2200	1273	1.24
1997	2625	1248	1.30
1998	2450	1165	1.23
1999	2250	1176	1.24
2000	2321	1200	1.30
2001	2413	1191	1.43
2002	2450	1246	1.38
2003	2567	1337	1.45
2004	2375	1270	1.37
2005	2375	1368	1.38
2006	2113	1359	1.32
2007	2734	1362	1.39
2008	2648	1149	1.38
2009	2623	1275	1.36
2010	2705	1302	1.39
2011	2793	1399	1.37
2012	2782	1466	1.38
2013	2880	1517	1.41
2014	2853	1585	1.38
2015	2826	1486	1.32
2016	2802	1487	1.37
2017	3023	1448	1.34
2018	2876	1484	1.33
2019	2857	1597	1.26
2020	2936	1500	1.42
2021	2831	1503	1.37

### 3.6.4 Sampling Data from Cow Creek

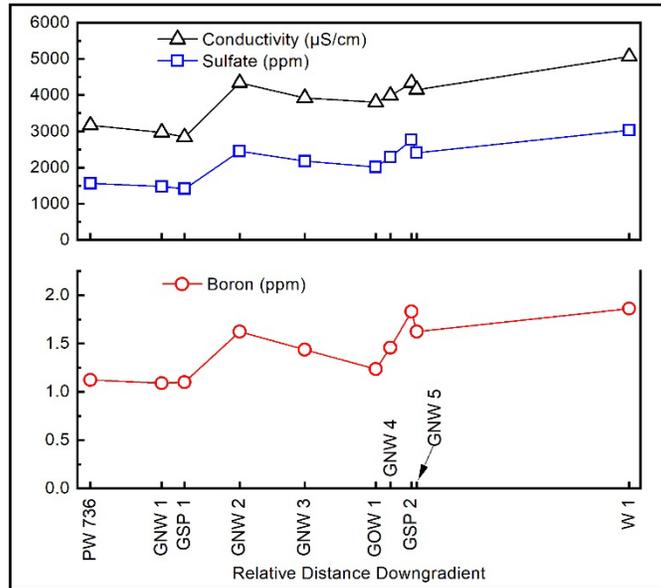
Along Cow Creek, 20 sites were sampled in 2021. This included GAS wells 1, 2, and 7 and the three PW monitoring wells near the western boundary of the study area (PW 734, PW 735, and PW 736; Figure

2.4). However, PW 734 was sampled in June only because the well did not have sufficient water for sampling in October. Surface-water site GSW 3 was dry during both sampling trips, and site GSW 5 was sampled only in June because the site was dry during the October sampling trip. Well GNW 9, the GAS wells, and PW 735 were not included in the drainage-mean calculations because most of them do not monitor the upper alluvial aquifer in Cow Creek. However, PW 735 is nevertheless formally evaluated for time trends in target parameters. Well PW 736 monitors the alluvial aquifer similar to wells GNW 1, 2, 3, 4, and 5, and could be considered a sentinel well that would provide a warning if EHP-contaminated water escaped past the upgradient capture system. Well PW 735 is included in the statistical site-specific trend analyses to determine if there are significant changes in water quality at that deeper location. The 21 monitoring sites in the drainage are divided into three categories: wells, springs, and surface waters. Surface-water sites (GSW) were segregated from the alluvial wells and springs because concentrations of constituents in surface waters are affected by evaporation and dilution, depending on season and weather. Therefore, GSW data also were not included in drainage analysis mean calculations.

Analytical results for the indicator parameters for springs and monitoring wells within the Cow Creek drainage for 2021 are graphed in Figure 3.12. Site mean conductivity values ranged from 2840  $\mu\text{S}/\text{cm}$  at GSP 1 to 5060  $\mu\text{S}/\text{cm}$  at W-1, with a drainage mean of 3723  $\mu\text{S}/\text{cm}$ . Average sulfate concentrations ranged from 1410 ppm at GSP 1 to 3030 ppm at W-1, with a 2061 ppm drainage mean. Mean boron concentrations ranged from 1.09 ppm at GNW 1 to 1.86 ppm at W-1, with a 1.39 ppm drainage mean. The spatial profiles for conductivity, sulfate, and boron in the alluvial wells and springs in Cow Creek drainage (Figure 3.12) show that all three parameters were correlated. Values generally increased down the drainage with localized peaks at GNW 2 and GSP 2. The overall maximum target parameters were found at W-1, which is consistent with tendencies from most previous years.

Three site-specific time trends were identified in the Cow Creek drainage: PW 735 had increasing trends for conductivity and boron, and GSP 2 had an increasing trend for boron. For PW 735, this was the third consecutive year in which an increasing conductivity trend was determined and the second consecutive year for an increasing boron trend. Plots of these constituents vs. time for the respective sites are shown in Figure 3.13, Figure 3.14, and Figure 3.15.

At PW 735, conductivity and boron values exhibit similar tendencies. The data are characterized by a sharp decrease between 2010 and 2012, moderate upward and downward fluctuations from 2012–2017, and generally increasing values from 2016–2021. PW 735 taps a deeper aquifer than most of the alluvial wells along Cow Creek and is monitored for water-quality changes in this deeper zone. It is difficult to conclude whether the increasing trends are related to contaminant plume migration; additional monitoring data and a better understanding of the natural background variation are needed. Moreover, it is unclear whether the aquifer tapped by PW 735 is confined (i.e., hydraulically isolated) in the vicinity of the EHP capture system. If this deeper aquifer is confined, it is still possible that there is some exchange with water from the shallower aquifer at an upgradient location. However, such a scenario is speculative and is not confirmed by our monitoring data. The laboratory-measured and continuous-monitoring probe results for conductivity at nearby well PW 736, which monitors the shallower alluvial aquifer, exhibited a slight, continuous decrease in conductivity over the past four years. Thus, the recent increases at PW 735 do not appear to be impacting the alluvial aquifer currently.



Site	Cond (µS/cm)	SO <sub>4</sub> <sup>2-</sup> (ppm)	B (ppm)
PW 736	3165	1560	1.12
GNW 1	2970	1480	1.09
GSP 1	2840	1410	1.10
GNW 2	4335	2450	1.62
GNW 3	3925	2175	1.44
GOW 1	3800	2015	1.24
GNW 4	3985	2285	1.46
GSP 2	4335	2760	1.83
GNW 5	4150	2410	1.62
W-1 <sup>†</sup>	5060	3030	1.86
2021 Drainage Mean	3723	2061	1.39

<sup>†</sup>Values for W-1 were not included in the drainage mean.

Figure 3.12. Plots and data for conductivity, sulfate, and boron values for alluvial wells and springs in Cow Creek Drainage during 2021. Values shown are averages from the June and October sampling events.

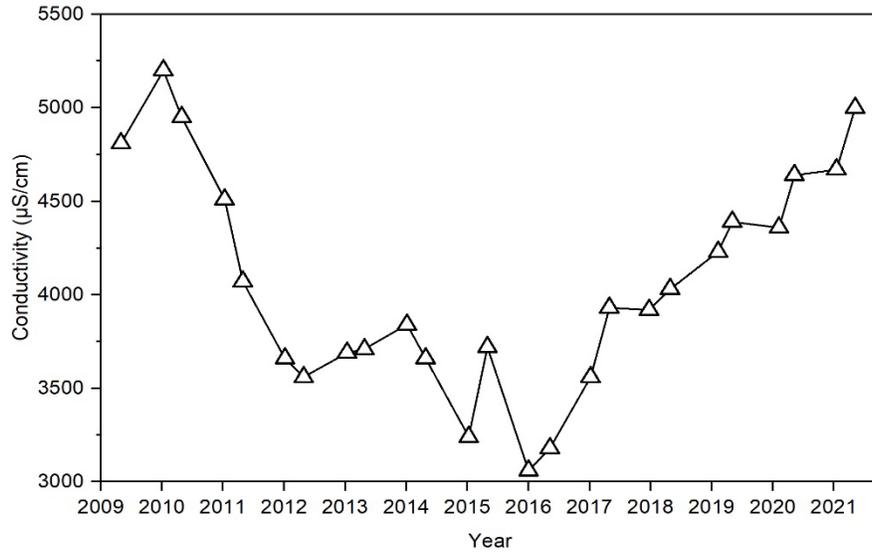


Figure 3.13. Trend plot for conductivity at PW 735.

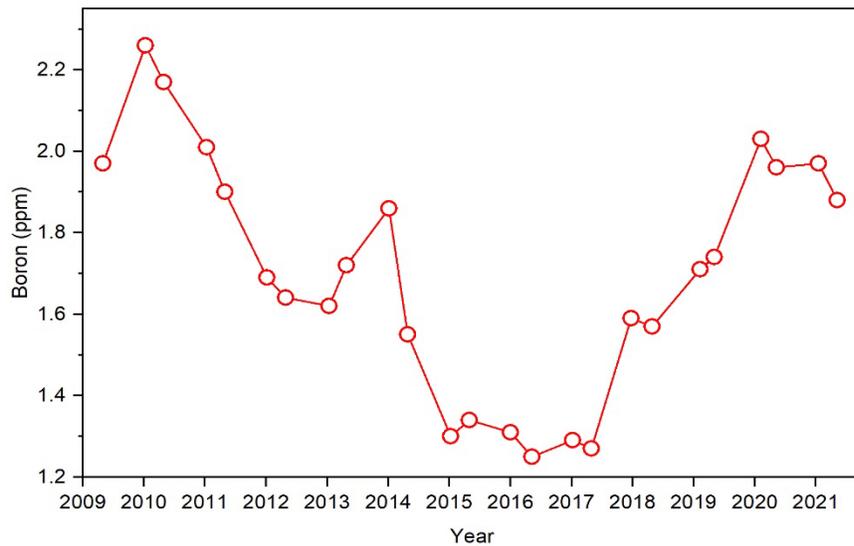


Figure 3.14. Trend plot for boron at PW 735.

The boron data for GSP 2 (Figure 3.15) exhibit frequent high-frequency fluctuations (up to ~0.5 ppm) superimposed on a profile that includes a subtle peak between 1985-2006, a more pronounced peak from 2006-2016, and a gradual increase over the past five years. GSP 2 is specifically mentioned in Stipulation 12(d) as a site where abruptly increased values would trigger interception of the bad quality water. No abrupt increase was found at GSP 2 based on a formal evaluation during the composite trend testing. Based on the conclusion of a recent increasing trend, the distribution of boron data that could have been observed in the October sampling at GSP 2 falls within the range 1.6–2.5 ppm. This corresponds to approximately the midpoint of historic data up to slightly higher concentrations than have been observed previously. There is no confirming evidence that the increasing boron trend at GSP 2 represents an incoming contaminant plume, i.e., no increasing (or decreasing) trends in boron were found at upgradient well GNW 4 or downgradient well GNW 5. Additionally, a special (PLR/conventional composite) test applied to boron at nearby upgradient surface-water site GSW 5 did not detect an increasing boron trend, implying that contaminated stormwater has not impinged on GSP 2. Regarding meteorological causes,

there are too few site increases among the unconfined (directly rechargeable) aquifers in the overall study area and insufficient recharge conditions to conclude that the GSP 2 site's increasing boron trend reflects a weather-induced concentrating effect.

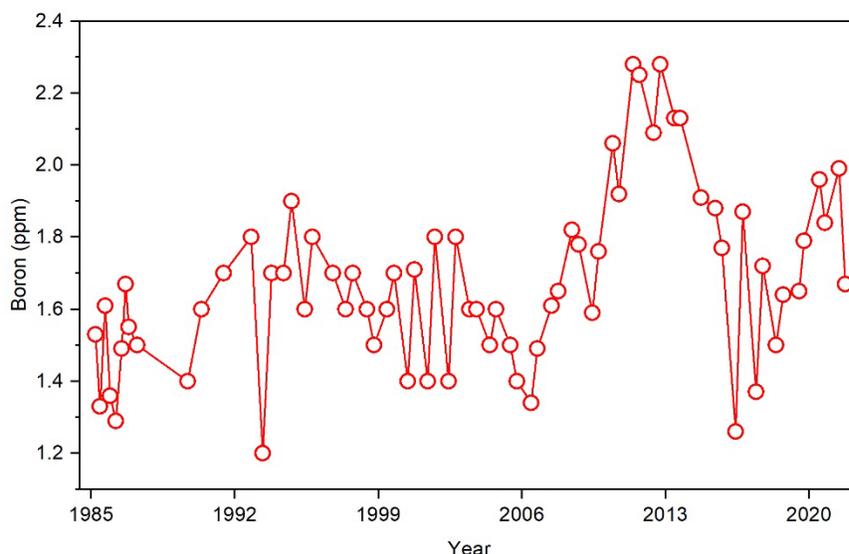


Figure 3.15. Trend plot for boron at GSP 2.

Table 3.12 summarizes the 2021 analytical results from the PW wells located near the western boundary of the study area. Well PW 734 was sampled in June only due to insufficient water in October; the other wells were sampled in June and October. Of the three PWs, PW 736 is believed to be most representative of the alluvial aquifer; therefore, PW 736 is included in the drainage-mean calculations. Well PW 734 taps a region where the uppermost aquifer is relatively shallow (between about 8 and 10 ft), while PW 735 monitors the deeper saturated zone with a screened depth between 35 and 80 ft, and its data are included in the site trend-detection analysis as noted previously. With the exception of pH, all of the parameters in Table 3.12 had their highest average values at PW 735. The highest pH was found at PW 734, followed by PW 736 and PW 735, consistent with the data from last year. Average bromide concentrations were highest at PW 735 followed by PW 736 and PW 734. Bromide results are discussed in more detail in Section 3.6.6. Levels of the indicator parameters conductivity, sulfate, and boron for all three PW wells were within the corresponding ranges of the alluvial wells in the drainage.

Table 3.12. 2021 analysis results from the PW wells.

Site	Comment	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br <sup>-</sup> (ppm)	Cl <sup>-</sup> (ppm)	SO <sub>4</sub> <sup>=</sup> (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)
PW 734	Single Value	361	7.99	4000	<0.06 7	15.1	2160	3770	1.77	208	276	16.3	397	50.77	52.62
PW 735	Mean Value	426	7.41	4835	0.142	21.6	2635	4605	1.92	271	370	18.9	416	62.52	63.99
PW 736	Mean Value	357	7.53	3165	0.076	15.1	1560	2755	1.12	225	237	12.5	186	39.16	40.05

Figure 3.16 and Table 3.13 (in Section 3.6.5) show the drainage-mean values from 1984–2021. The latest conductivity and boron values increased relative to the previous year, while sulfate had a very modest decrease. None of these changes is significant relative to previous fluctuations. No drainage trends were identified for 2021.

### 3.6.5 Comparison of Cow Creek, South Fork Cow Creek, and Pony Creek Drainages

For the overall observation period (1985–1987 and 1990–2021), the mean values for conductivity and sulfate in the Pony Creek drainage are less than those in the South Fork Cow Creek drainage, which are less than those in the Cow Creek drainage. For boron, the means in both the Cow Creek and the South Fork Cow Creek drainages are not statistically different, but both are greater than the mean in the Pony Creek drainage.

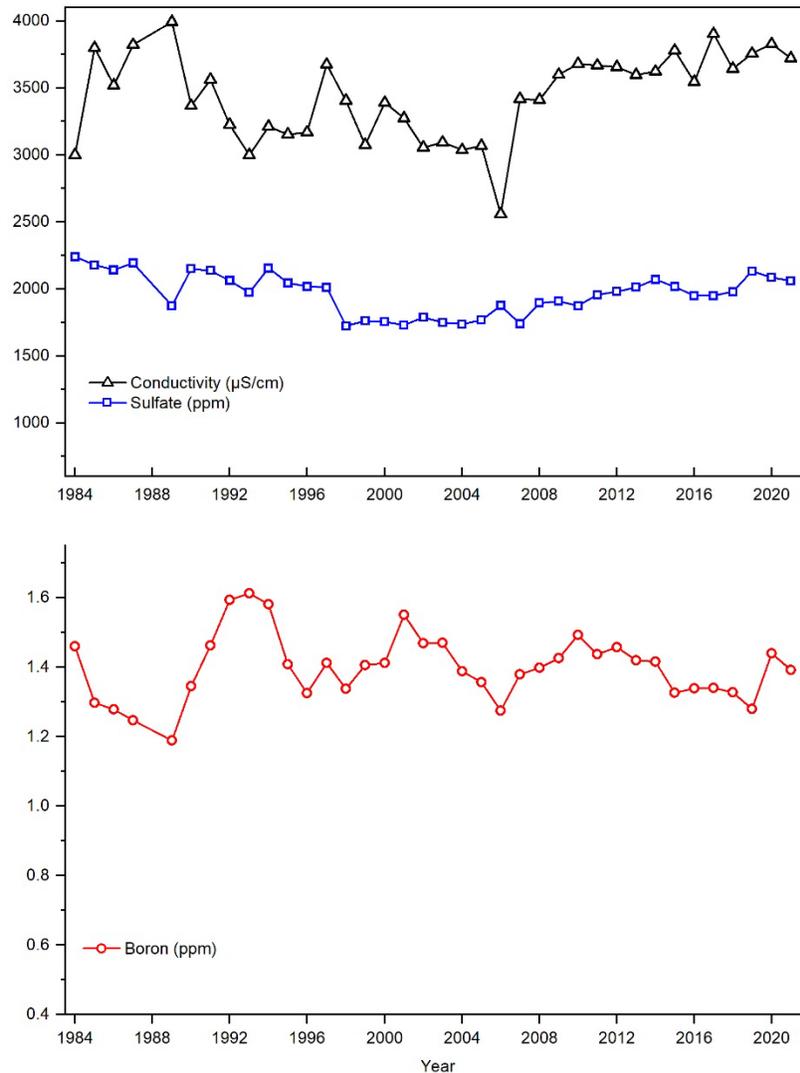


Figure 3.16. Trend plots for conductivity, sulfate, and boron in Cow Creek Drainage, 1984–2021.

Table 3.13. Drainage-mean values for conductivity, sulfate, and boron in Cow Creek drainage, 1984–2021.

Year	Drainage Mean		
	Conductivity ( $\mu\text{S}/\text{cm}$ )	$\text{SO}_4^{2-}$ (ppm)	Boron (ppm)
1984	3000	2240	1.46
1985	3799	2178	1.30
1986	3519	2142	1.28
1987	3823	2193	1.25
1989	3994	1875	1.19
1990	3369	2150	1.35
1991	3563	2138	1.46
1992	3225	2063	1.59
1993	3000	1975	1.61
1994	3213	2156	1.58
1995	3154	2044	1.41
1996	3169	2019	1.33
1997	3675	2013	1.41
1998	3406	1725	1.34
1999	3075	1763	1.41
2000	3391	1755	1.41
2001	3275	1731	1.55
2002	3056	1788	1.47
2003	3094	1750	1.47
2004	3038	1738	1.39
2005	3069	1769	1.36
2006	2559	1876	1.27
2007	3419	1740	1.38
2008	3413	1896	1.40
2009	3599	1909	1.43
2010	3680	1875	1.49
2011	3669	1956	1.44
2012	3657	1981	1.46
2013	3599	2014	1.42
2014	3623	2070	1.42
2015	3779	2016	1.33
2016	3546	1949	1.34
2017	3904	1949	1.34
2018	3644	1979	1.33
2019	3756	2133	1.28
2020	3828	2085	1.44
2021	3723	2061	1.39

### 3.6.6 Analysis of Bromide

In December 2009, driven by a regulatory requirement of the U.S. Environmental Protection Agency, bromide ( $\text{Br}^-$ ) was added to the scrubber solution to enhance the removal of mercury from the flue gas produced by the Colstrip power plant. An EHP sample collected in 2015 contained a bromide concentration of 1940 ppm (Table 3.8). Within groundwater systems, the bromide ion is regarded as a conservative tracer, and its concentration in the EHP is sufficiently high to be useful as a leakage indicator. Therefore, in response to the addition of bromide to the EHP, bromide was added to the formal list of the chemical parameters of interest in 2010.

In previous annual reports, we reported the first 11 years of monitoring results for bromide, along with data from September 2009 (before bromide was added to the EHP) for comparison (Thompson et al. 2011–2021). Several sites have had bromide concentrations that were higher than the 2009 levels, but because the data were highly variable, and almost all of the results were low (i.e., within a factor of three of the MDL), statistical trend testing to address possible EHP contamination has not been performed. Results at these low concentrations are not considered to be reliable for quantitative evaluation.

Monitoring of bromide continued in 2021; Figure 3.17 shows the concentrations for sites along the Cow Creek, South Fork Cow Creek, and Pony Creek drainages. For sites that were sampled in June and October, the average is shown. In cases for which bromide was not detected, the plot is annotated with ND to indicate that an analysis was performed. Also, in situations where the 2021 data for a site consisted of non-detected and detected values, the detected concentration was plotted rather than the average. The MDL was 0.067 ppm as indicated by the horizontal red line on the plot.

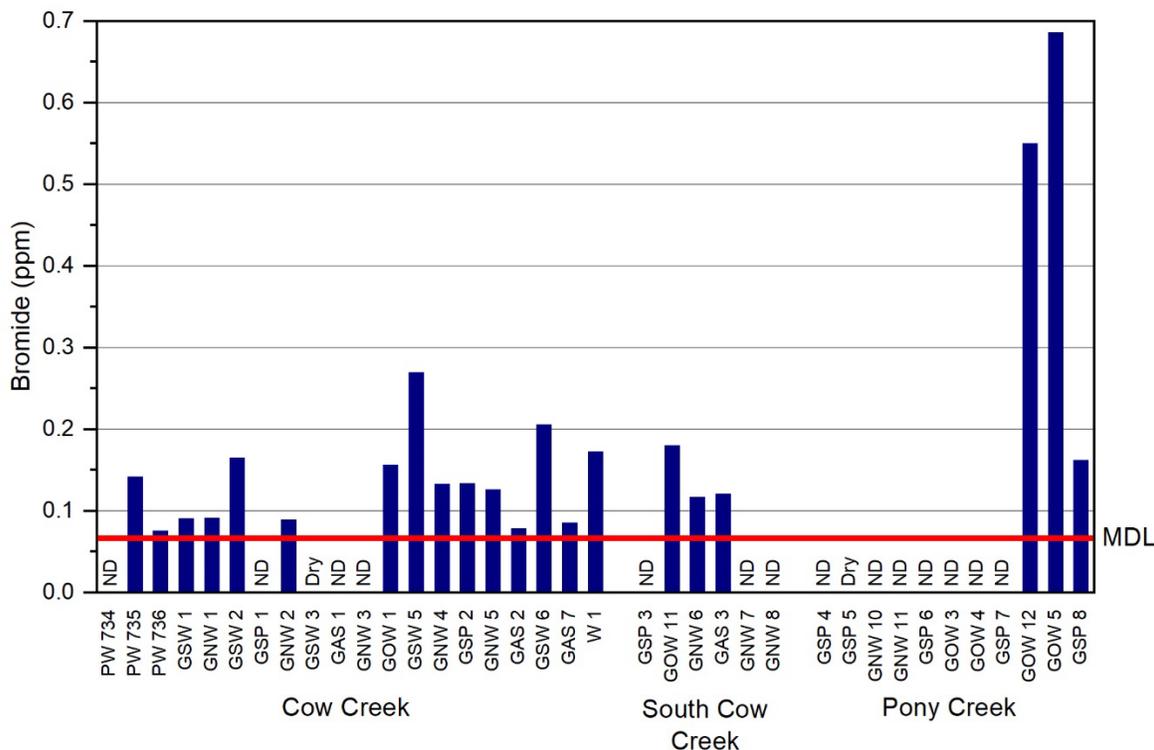


Figure 3.17. 2021 Bromide concentrations from sampling locations within the Cow Creek, South Cow Creek, and Pony Creek Drainages. The sites are ordered according to their position in each drainage (upgradient to downgradient). ND = non-detected result; Dry = site was dry (no sample was collected).

Consistent with previous data, the 2021 bromide results provide no clear evidence of contamination from the EHP in the study area. Of the 25 sites along Cow Creek and South Fork Cow Creek with data in Figure 3.17, only seven had bromide concentrations that exceeded two times the MDL, and only two (GSW 5 and GSW 6) had levels greater than three times the MDL. Evaporation at these surface-water sites may account for their higher concentrations. The highest bromide concentrations, 0.55 and 0.69 ppm, were found at Pony Creek wells GOW 12 and GOW 5, respectively. These wells have had similar concentrations for the past five or more years. As noted earlier in this report, GOW 12 was reconfigured in 2014 and sample data from that well remain unrepresentative of the alluvial aquifer. Limited bromide data are available for GOW 5; the well was not sampled between 2007 and 2015. All of the GOW 5 results have been greater than 0.64 ppm except for a value of 0.33 ppm in 2017 and a non-detected result in 2019 that had an associated MDL of 0.67 ppm due to sample dilution. Both GOW 5 and GOW 12 tap groundwater sources that contain higher levels of bromide than other sites in the study area; thus their recent elevated concentrations are not associated with contamination from the EHP.

The apparent limitation of the current low-concentration bromide data is actually advantageous because it establishes that background levels of bromide at the sites is low, and it is useful for detecting site contamination due to overland runoff (resulting in higher levels). Any groundwater plumes at most sites further downgradient and distant from the PW wells would not yet reflect the recent presence of bromide in the EHP unless there was impingement of contaminated stormwater or snowmelt. Therefore, bromide is currently uniquely valuable as a tracer for contaminated runoff, and its results indicate that no contamination from this pathway has recently occurred. In the future, this analyte will become increasingly useful as a tracer for potential groundwater contamination from the EHP.

### 3.6.7 Continuous-Monitoring Probe Results

An in situ monitoring probe was installed in well PW 736 on December 2, 2010. The probe measures groundwater specific conductance, temperature, and DTW at 4-hour intervals. The data are electronically recorded using an on-site datalogger and periodically are uploaded via satellite to an online database.

Well PW 736 is in the uppermost reaches of Cow Creek, downgradient from the Units 3 and 4 EHP. The monitoring probe (Figure 3.18) was installed in the well casing several feet below the groundwater surface and secured in place using a cable and fixed anchor point. The probe is part of a data station that is located at the ground surface above the well (Figure 3.19). The station includes a datalogger, battery, solar panel, and satellite connection—all of which work to power the probe, record, and relay probe measurements. Probe measurements for the first full year were reported in 2012 (Thompson et al. 2012). The following section presents the data collected in 2021 by the submersible continuous-monitoring probe.



Figure 3.18. Continuous-monitoring probe.



Figure 3.19. Well PW 736 with companion datalogger enclosure, solar panel, and fenced enclosure.

### 3.6.7.1 Data Record

This section summarizes data collected every four hours from January 1, 2021 through December 31, 2021, and includes discussions of groundwater specific conductance, elevation, and temperature. Specific conductance, or electrical conductivity, is the same parameter measured by the analytical laboratory and is an index of dissolved solutes. In addition, the datalogger records data-station battery voltage. The performance of the solar array and data-station battery voltage are recorded to monitor the station's status and maintenance needs. During 2021, the solar array and battery functioned properly. The batteries were replaced in June as a precautionary measure.

Since its installation in December 2010, the monitoring station has recorded approximately 29,350 readings. The 2021 data includes 1880 readings. During the year, several data records failed to be recorded. For example, from August 19 to October 10, data were not recorded due to the telemetry service provider inadvertently canceling the service. Other instances of failed data records were encountered as blank cells in the downloaded file. These were infrequent and constituted approximately 2% of the total data records. False records also were recorded during field sampling in June when the probe was removed from the well for inspection and when the probe re-equilibrated after being redeployed in the well. These data gaps are presented in Figure 3.20, Figure 3.21, and Figure 3.23 in Section 3.6.7.2.

Nevertheless, as will be illustrated below, the data logging system is extremely valuable, not only for verification of field and laboratory conductivity data, but also for obtaining a better understanding of the temporal character of all three parameters that it measures. The system provides a unique capability for evaluating the impacts from precipitation events and dry periods. Installation of continuous-monitoring systems at downgradient sites and in other drainages would be considerably beneficial in this study.

During the June 2021 sampling campaign, well PW 736 was sampled after inspecting the probe. As part of the inspection, the probe calibration was evaluated by immersing the probe in a standard solution of

known conductivity (1413  $\mu\text{S}/\text{cm}$ ) and comparing the readings with the expected value. The probe read 1423  $\mu\text{S}/\text{cm}$  and the field team concurred that recalibration was not necessary. The probe was then replaced in the well after visual confirmation of the probe's condition and sample collection had been completed. After redeployment of the probe, the specific conductance readings decreased by approximately 33  $\mu\text{S}/\text{cm}$  and gradually returned to the pre-sampling level over the next 10 days. This period accounted for 60 false readings that were removed from the data record because they were unrepresentative of the true conductivity conditions.

During the October 2021 sampling event, the probe was again placed in a standard solution of known conductivity (1413  $\mu\text{S}/\text{cm}$ ). The probe read 1427  $\mu\text{S}/\text{cm}$ , and the field team concurred that recalibration was unnecessary. Well PW-736 was sampled following the probe removal. After inspecting the monitoring probe and datalogger, the telemetry service provider was contacted for troubleshooting support due to not receiving data records since August 19, 2021. The service provider concluded that the service was inadvertently canceled and on October 10, 2021, the service was reinstated.

### **3.6.7.2 Results and Discussion**

Groundwater temperatures during 2021 ranged from 10.29 to 10.69  $^{\circ}\text{C}$  with a mean temperature of 10.52  $^{\circ}\text{C}$  (Figure 3.20). The minimum groundwater temperature for 2021, 10.29  $^{\circ}\text{C}$  occurred on multiple occasions in July and August. The 11-year (period of record) groundwater temperature fluctuation ranges from 10.28 to 11.02  $^{\circ}\text{C}$  with a mean temperature of 10.64  $^{\circ}\text{C}$ . Seasonal variation in aquifer temperature is evident; maximum temperatures occur in February/early-March and minimum temperatures occur in August/early-September, indicating, for this site, approximately a 7-month lag in heat balance relative to that of ambient air temperature. Consistent with data records from previous years, the daily groundwater temperature variation was greatest between early spring and mid-summer and is potentially related to aquifer recharge events. Smaller daily temperature variation from August to February may suggest that little recharge occurs during this period.

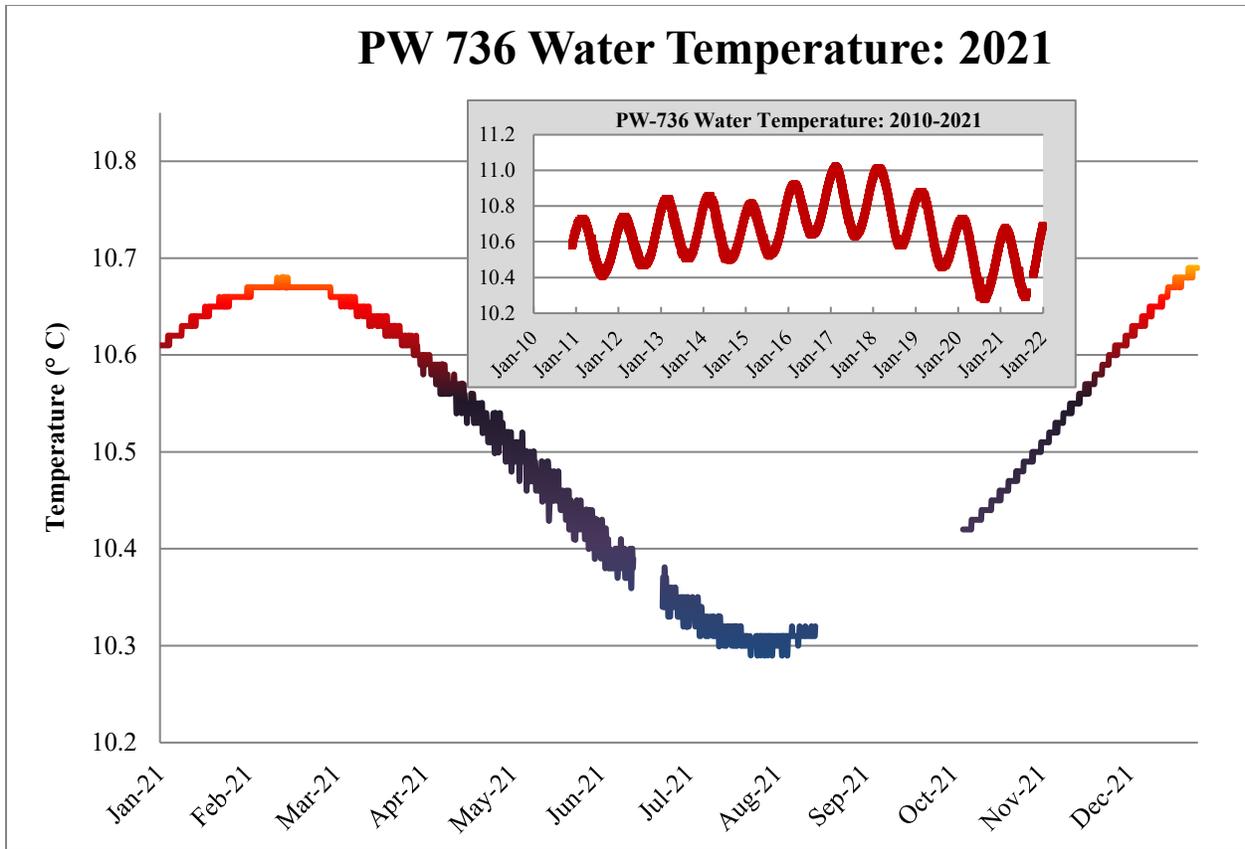


Figure 3.20. Groundwater temperature trend measured at PW 736 in 2021 and 2010–2021 (inset).

The groundwater elevation in PW 736 fluctuated 0.89 ft during 2021, ranging from 3053.52 to 3054.41 ft above mean sea level (Figure 3.21). Groundwater elevation peaked in March and decreased through the remainder of the year. Maximum precipitation occurred during the month of August. No conclusions can be made regarding an influence on the water table during this period since the monitoring probe was not recording data. However, a correlation between precipitation and groundwater recharge has been observed during previous years and is typical of a shallow, unconfined aquifer. The groundwater elevations presented in Figure 3.21 are referenced to the surveyed elevation of the top of the well casing which is 3075.5 ft above sea level.

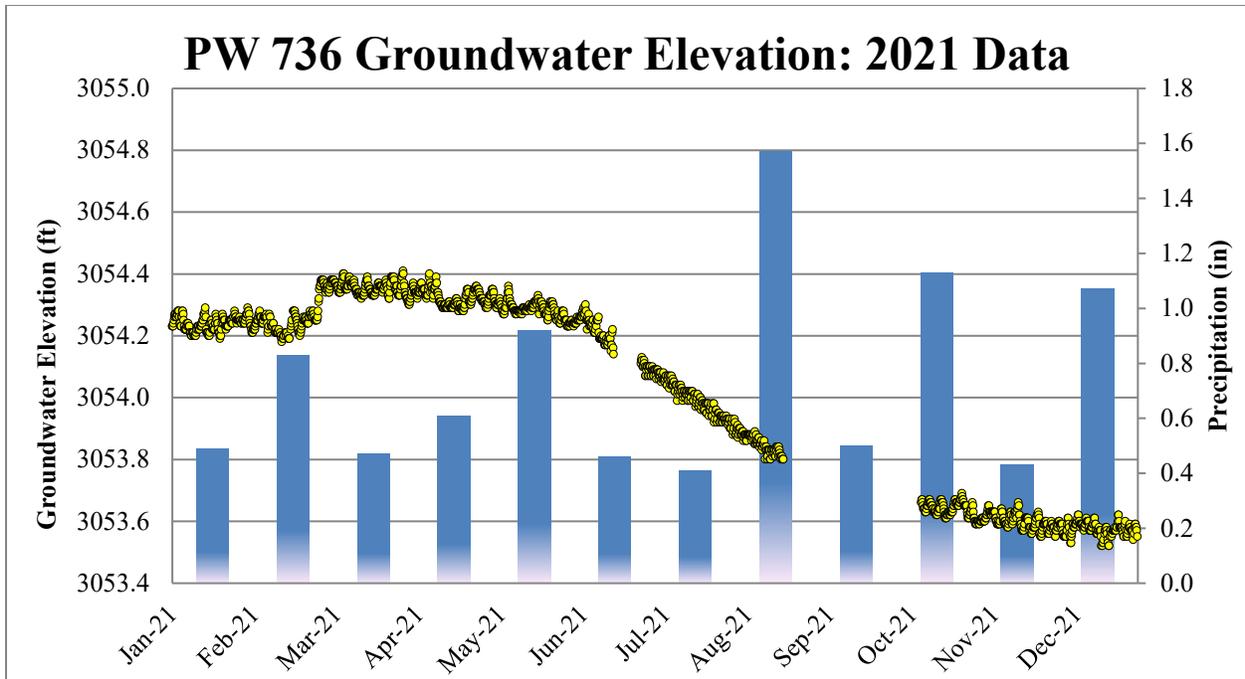


Figure 3.21. Groundwater elevation trend at PW 736 during 2021 in association with monthly total precipitation (columns and secondary y-axis).

The 10-year fluctuation in groundwater elevation is 3.89 ft, ranging from 3052.23 and 3056.12 ft (Figure 3.22). These data suggest that the aquifer is seasonally recharged by spring precipitation and snowmelt during years such as 2011, 2017, 2018, 2019, 2020, and 2021; and to a lesser degree in years such as 2015 and 2016. Rapid increases in groundwater elevation have occurred during May 2011, February 2017, March 2019, and March 2020. In May 2011, 10 in. of precipitation was recorded due to a flood event, resulting in a rapid increase in groundwater elevation. Rapid increases also occurred during February 2017, March 2019, March 2020, and March 2021 due to snowmelt events.

During 2021, the groundwater elevation rapidly increased approximately 0.2 ft and peaked during early-March. This behavior is consistent with a shallow, unconfined alluvial aquifer with a vadose zone capable of appreciable water storage capacity. If precipitation input exceeds the storage capacity of the unsaturated zone, then water table recharge occurs. The daily variation observed in groundwater temperature during the late winter to early summer period suggests some vadose zone input, but a change in water table surface elevation generally is not observed despite daily temperature variations, which suggests these inputs are typically minimal. For 2021, the influence of water inputs from the EHP, if any, on groundwater elevation is unknown. However, nothing in the 2021 data set suggests pulse inputs affecting groundwater elevation.

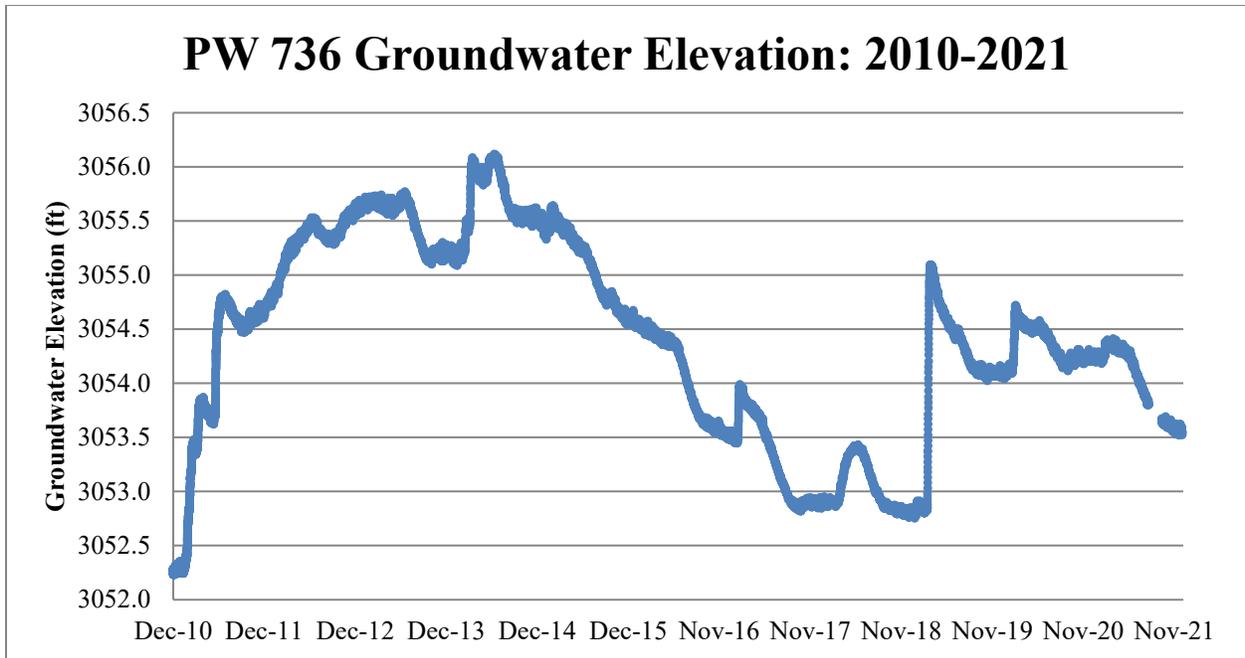


Figure 3.22. Groundwater elevation trend at PW 736 from 2010–2021.

Groundwater specific conductance measured at PW 736 during 2021 ranged from 2884 to 2936  $\mu\text{S}/\text{cm}$  over the course of the monitoring year (Figure 3.23). The 11-year trend in specific conductance is shown in Figure 3.24.

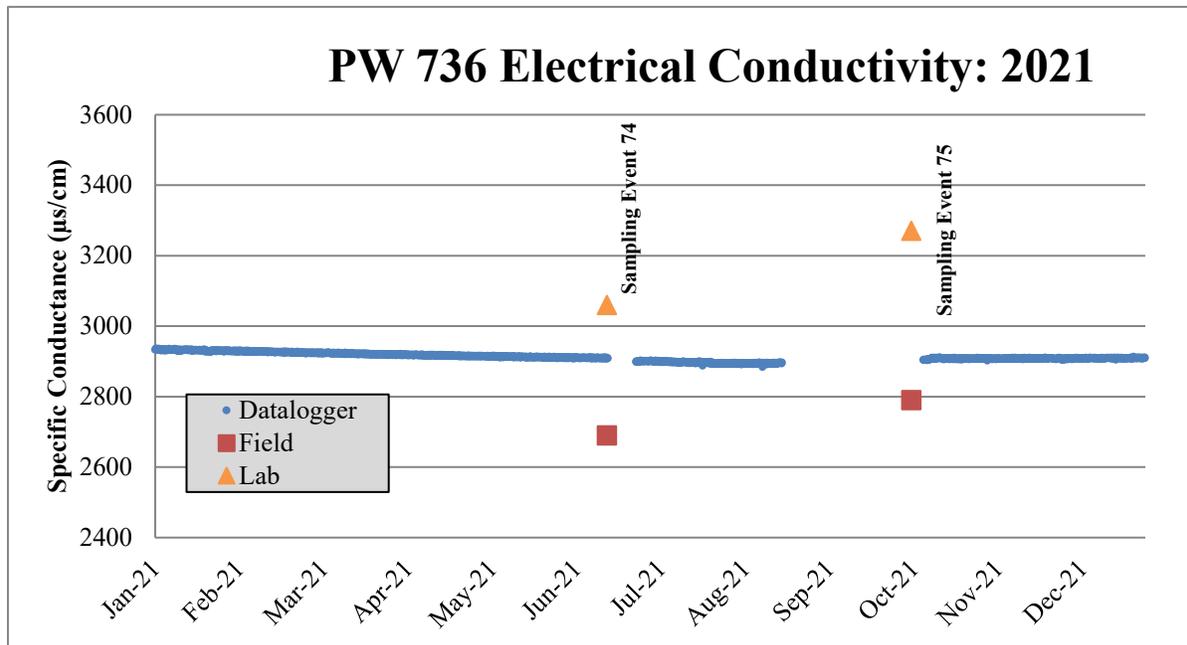


Figure 3.23. In situ PW 736 specific conductance trend measured by the datalogger over the 2021 monitoring year compared to field and laboratory specific conductance measured during sampling events.

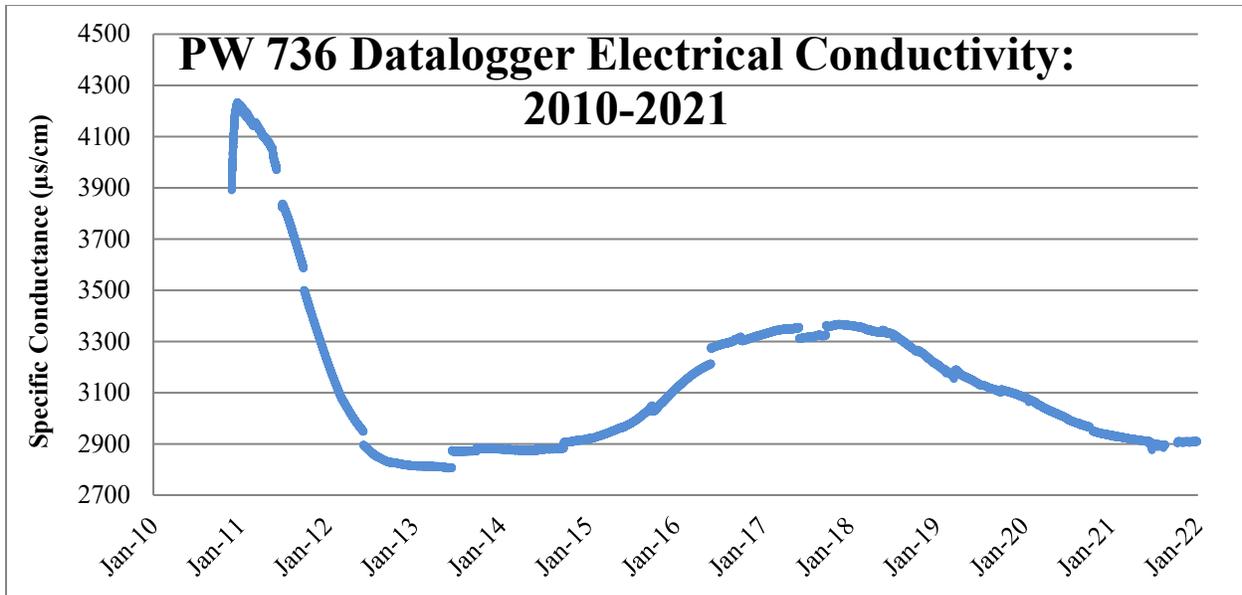


Figure 3.24. Specific conductance for PW 736 over the entire period of record 2010–2021.

Three measurements of specific conductance were collected during each sampling event. The datalogger provides a continuous measurement at 4-hour intervals, and additionally, a field sample was collected and measured within a few hours of collection. The third conductivity result is a laboratory result measured days to weeks after receiving the samples. The resulting variance in the sample results indicates generally good agreement between the measurements by different instruments at different times. For example, in the June 2021 sampling event, the datalogger reported 2910  $\mu\text{S}/\text{cm}$ , the field result was 2690  $\mu\text{S}/\text{cm}$ , and the corresponding lab value was 3060  $\mu\text{S}/\text{cm}$ . The maximum difference between these measurements was 370  $\mu\text{S}/\text{cm}$ . During the October 2021 sampling event, the datalogger reported 2906  $\mu\text{S}/\text{cm}$  while the field result was 2790  $\mu\text{S}/\text{cm}$ , and the corresponding lab value was 3270  $\mu\text{S}/\text{cm}$ . The maximum difference between these measurements was 480  $\mu\text{S}/\text{cm}$ . For the June and October 2021 sampling events, differences between datalogger and lab analysis measurements of specific conductance were 150 and 364  $\mu\text{S}/\text{cm}$ , respectively. The RPD between the datalogger and lab analysis measurements were 5.0% and 11.8% for the June and October 2021 sampling events, respectively. Thus, the differences between datalogger and lab analysis measurements of specific conductance measurements during 2021 are consistent with past years. These results confirm that the continuous-monitoring probe provides acceptably accurate and precise measurements of specific conductance.

The long-term trend in conductivity shown in Figure 3.24 suggests that electrical conductivity values are stable over long periods of time with only slight changes between 2012 and 2021. This trend suggests that pulse inputs from the EHP over short periods of time have not resulted in increased salinity of the shallow aquifer during the past 11 years. The cause(s) of apparent aquifer salinity spikes prior to January 2011 are not shown by this data set, but the decrease in salinity in 2011 was a significant trend. Small annual changes in conductivity were measured during the 2012–2021 period with a relatively stable trend during 2021. A slight decreasing trend was noted from January 2021 until mid-July/early August, then a negligible increasing trend was observed through December. Specific conductance varied approximately 52  $\mu\text{S}/\text{cm}$  over the course of the year.

The degree to which regional aquifer characteristics vary over time with respect to salinity is unknown. However, spring and early summer recharge events and other climatological factors (e.g., amount of

snowmelt) are likely the primary causes of the fluctuations. The profile in Figure 3.24 is similar in shape to the estimated conductivity trend for PW 736 (Figure F.9).

### **3.6.7.3 Data Summary**

The continuous-monitoring probe in well PW 736 continues to record important groundwater data downgradient from the Units 3 and 4 EHP. The specific conductance, groundwater elevation, and temperature data provide unique and useful supplements to biannual analytical sampling. Continued monitoring will allow long-term groundwater table fluctuations and specific conductance trends to be observed. No unusual specific conductance, temperature, or elevation measurements were observed during 2021, suggesting an absence of accidental releases from the EHP. Probe/datalogger installations in other monitoring wells and springs or other surface water in the project area would improve our understanding of downgradient groundwater and surface-water temporal behavior in Cow Creek and the South Fork of Cow Creek. Also, such installations would facilitate rapid detection, characterization, and mitigation response in the event of significant leakage or a contaminated stormwater event.



## 4.0 Conclusions

Since 1984, staff from PNNL/PNWD and KC Harvey Environmental, LLC have collected hydrologic and water-quality data in the drainage basins of Cow Creek and Pony Creek in Rosebud County, Montana. Information from this monitoring program is used to determine whether water-quality and hydrological parameters have changed from previous years and to provide a basis for evaluating the effects of any ongoing or future events associated with Colstrip Power Plant Units 3 and 4 EHP, which is located upgradient from Cow Creek and Pony Creek.

There were no reportable releases from the EHP during 2021.

Total precipitation during 2021 was 8.89 in., which is 43% lower than the average of 15.6 in. measured since 1984 and about 25% less than the amount for 2020. During 2021, August, October, and December were the wettest months, accounting for 18%, 13%, and 12% of the total annual precipitation, respectively. The winter months were relatively dry, which is a normal precipitation pattern for the area.

Groundwater levels measured in the alluvial wells during 2021 averaged about 0.93 ft lower than those of the previous year. With one exception, all wells had lower water levels in summer and fall than in the previous year; well W-1 had an identical level in the summers of 2020 and 2021. Between the June and October sampling campaigns, water levels declined in all monitored wells. The largest decreases occurred at wells GNW 2 (2.84 ft), GOW 1 (1.35 ft), GOW 11 (1.25 ft), and W-1 (1.21 ft). Overall, the lower water levels in the past year are consistent with the lower amount of precipitation in 2021.

Statistical analysis identified five sites (GOW 4, GSP 2, GSP 4, GSP 8, and PW 735) where water quality exhibited continual trends from the past four or more years. Four site-specific trends were identified for Pony Creek: GOW 4 had an increase for conductivity and a decrease for boron, GSP 4 had an increase for conductivity, and GSP 8 had an increase for boron. Among Cow Creek sites, relatively deep well PW 735 had increases in conductivity and boron, and GSP 2 had an increase in boron. No drainage trends were found for Pony Creek, Cow Creek, or South Fork Cow Creek. Overall, these findings do not indicate any definitive impacts on groundwater quality from the operations of the EHP.

The relatively new PW wells on Cow Creek near the western border of the study area have been monitored for water quality since 2009. Similar to last year, the levels of conductivity, sulfate, and boron in well PW 736 were comparable to those in most of the alluvial wells in the drainage, while the nearby deeper well PW 735 had markedly higher values for those parameters. PW 735 had increasing trends for conductivity and boron, but long-term variation in the estimated trends makes it difficult to assess whether those trends are related to contaminant plume migration. Additionally, it is unclear whether the aquifer tapped by PW 735 is confined (hydraulically isolated) in the vicinity of the capture system. Conductivity, sulfate, and boron levels in the neighboring shallower well PW 734 fell between those of PW 735 and PW 736.

Boron levels have generally increased at GSP 2 over the past five years, but the magnitude of the change is within the range of historical variability at this site. Stipulation 12(d) states that abrupt increases at this site would trigger interception of the bad quality water. However, the recent changes have been gradual, and there is no confirming evidence at upgradient well GNW 4 or downgradient well GNW 5 that the increasing trend represents an incoming contaminant plume.

Using data from the entire observation period (1984–1987 and 1990–2021) and based on the selected sampling sites, water from the Pony Creek drainage had the lowest levels of conductivity, sulfate, and boron. Water from the Cow Creek drainage was distinctly highest in conductivity and sulfate, and mean

values for water from the South Fork Cow Creek drainage fell between the values of the other drainages except for boron, for which the mean was the same as that from Cow Creek.

Monitoring bromide as a potential indicator of contamination from the Units 3 and 4 EHP continued in 2021. Two sites in the lower portion of Pony Creek, GOW 12 and GOW 5, had the highest bromide concentrations of approximately 0.6 and 0.7 ppm, respectively, in June. Previous data indicate that these wells tap groundwater sources with higher concentrations of bromide than other sites in the study area. At most other sites, detected bromide concentrations were within three times the MDL. Consequently, the most recent results provide no apparent evidence of encroachment of water from the EHP in the study area either via groundwater transport or by overland flow of rainwater/snowmelt.

The continuous-monitoring probe installed in well PW 736 in December 2010 collects frequent temperature, DTW, and conductivity data to better understand the behavior of the alluvial aquifer near the EHP and provide insight into the local alluvial hydrology in the study area in general. Similar to previous years' data, the 2021 probe measurements show fluctuations in temperature that, along with probe groundwater-level measurements, appear to be related to aquifer recharge. Specific conductance was relatively stable during 2021, varying by approximately 52  $\mu\text{S}/\text{cm}$ . This year's continuous-monitoring data do not indicate any releases from the EHP.

## 5.0 References

- Chamberlain, P. J. 2018. "The Bayes Factor Reinterpreted and Renamed as the Predictive Likelihoods Ratio - It's a UMP Composite vs. Composite Test on Multiple Parameters." Technical Bulletin No. 3, Enviro-Sci Consulting.
- Erbes D. 2000. *Geochemical Transformations and Hydrologic Recovery Observed in a Coal-Strip-Mine Spoils Aquifer*. Billings Land Reclamation Symposium 2000, Striving for Restoration, Fostering Technology and Policy for Reestablishing Ecological Function, Reclamation Research Unit Publication 00-01, Montana State University, Bozeman, Montana.
- Hydrometrics. 1987. Quantity Management of Groundwater in the Upper Cow Creek Drainage, Colstrip, Montana. Hydrometrics, Inc., Helena, Montana.
- Hydrometrics. 1990. Evaluation of 1989 Hydrologic Monitoring Data from the Montana Power Company Process Pond System at Colstrip, Montana. Hydrometrics, Inc., Helena, Montana.
- Lee RW. 1980. *Geochemistry of Water in the Fort Union Formation of the Northern Powder River Basin, Southeastern Montana*. Water Resources Investigations Open-File Report 80-336, U.S. Geological Survey, Helena, Montana.
- McDonald JP, KB Olsen, JL Smoot, PJ Chamberlain, SR Jennings, and JD Goering. 2007. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2006*. Battelle – Pacific Northwest Division, Richland, Washington.
- Metesh J. 1994. *Cumulative Hydrologic Impacts on Rosebud Creek from Development Near Colstrip*. Montana Bureau of Mines and Geology Open-File Report No. 273, Butte, Montana.
- Montana Department of State Lands and U.S. Office of Surface Mining. 1983. *Comprehensive Environmental Impact Study: Western Energy's Rosebud Mine*. Helena, Montana.
- Olsen KB, DA Myers, MR Toland, and FF Munshower. 1987. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. Battelle–Pacific Northwest Division, Richland, Washington.
- Olsen KB, JL Smoot, and JW Johnston. 1991. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1990*. Battelle–Pacific Northwest Division, Richland, Washington.
- Olsen KB, JL Smoot, and PJ Chamberlain. 1992. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1991*. Battelle – Pacific Northwest Division, Richland, Washington.
- Olsen KB, JL Smoot, and PJ Chamberlain. 1993. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1992*. Battelle – Pacific Northwest Division, Richland, Washington.
- Olsen KB, JL Smoot, and PJ Chamberlain. 1994. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1993*. Battelle – Pacific Northwest Division, Richland, Washington.
- Olsen KB, JL Smoot, and PJ Chamberlain. 1995. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1994*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1996. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1995*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1997. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1996*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1998. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1997*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1999. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1998*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2000. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1999*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2001. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2000*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2002. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2001*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2003. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2002*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2004. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2003*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2005. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2004*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2006. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2005*. Battelle – Pacific Northwest Division, Richland, Washington.

Shimer JA. 1972. *Field Guide to Landforms in the United States*. The Macmillan Company, New York.

Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2008. *2007 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. Battelle – Pacific Northwest Division, Richland, Washington.

Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2009. *2008 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4063, Battelle – Pacific Northwest Division, Richland, Washington.

- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2010. *2009 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4182, Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2011. *2010 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4261, Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2012. *2011 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4352, Battelle–Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2013. *2012 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-22498, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2014. *2013 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10010, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2015. *2014 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10027, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2016. *2015 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10043, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2017. *2016 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10056, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, PD Thorne, PJ Chamberlain, SR Jennings, and C Lucy. 2018. *2017 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10066, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, PJ Chamberlain, and C Lucy. 2019. *2018 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10080, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, PJ Chamberlain, and E Koehler. 2020. *2019 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10098, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, E Koehler, M Larson, and PJ Chamberlain. 2021. *2020 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10108, Pacific Northwest National Laboratory, Richland, Washington.
- Van Voast WA and JC Reiten. 1988. *Hydrogeologic Responses: Twenty Years of Surface Coal Mining in Southeastern Montana*. Memoir 62, Montana Bureau of Mines and Geology, Butte, Montana.
- Van Voast WA, RB Hedges, and JJ McDermott. 1977. *Hydrogeologic Conditions and Projections Related to Mining near Colstrip, Southeastern Montana*. Bulletin 102, Montana Bureau of Mines and Geology, Butte, Montana.

Waren K and A McDannel. 2003. "An Evaluation of Ground Water Conditions in the Colstrip Coal Deposit, Montana." *Abstracts of the Twentieth Annual Meeting American Society for Surface Mining and Reclamation*, Billings, Montana, June 3–6, 2003.

## **Appendix A**

### **Comments Provided by Signatories of Stipulation 12(d)**

## **Appendix A**

### **Comments Provided by Signatories of Stipulation 12(d)**

No formal comments were received for this year's report.

## **Appendix B**

### **Field Record Sheet and Chain-of-Custody Form Used for Sample Control**

## Appendix B

### Field Record Sheet and Chain-of-Custody Form Used for Sample Control

COLSTRIP GROUND WATER SAMPLE FIELD RECORD							
BATTELLE Pacific Northwest Division Richland, WA 99354							
Well Name _____				Project: <u>Colstrip</u>			
Recommended Purge Volume (gal) _____				Date _____			
Purge Flow Rate (gal/min) _____				Calculations: _____			
Pump Type _____							
SAMPLES COLLECTED							
SAMPLE NUMBER(S)	BOTTLE ID	PUMP TYPE	COLLECTOR				
Filtered Anions 1L P Filtered Metals 250 ml P w/HNO <sub>3</sub>  TOTAL NUMBER OF BOTTLES: <b>2</b>							
FIELD MEASUREMENTS							
Previous pH	Prv. Depth	Water Levels (ft)		Time			
Steel Tape# _____	Held			pH			
pH/Cond Meter# _____	Cut			Teml (DEG.C)			
Thermometer# _____	Depth below TC			Cond (um/cm)			
Turbidimeter# _____				Turb (NTU)			
FIELD OBSERVATIONS							
Weather: _____							
General Problems/ Unusual Events: _____							
Equipment Irregularities: _____							
Container Irregularities: _____							
Comments: _____							
Well capped and locked? <u>      </u> Yes <u>      </u> No							
Data Recorded by: _____				Data Checked by: _____			
Chain-of-Custody form # _____				FRF Logbook and Page No: _____			





## **Appendix C**

### **Analytical Methods Used for Water Analysis**

## Appendix C

### Analytical Methods Used for Water Analysis

Listed below are the methods for chemical analysis that were used to measure parameters in the Cow Creek and Pony Creek hydrogeochemical study. Analyses were conducted by GEL Laboratories, LLC, in Charleston, South Carolina.

Bicarbonate alkalinity	<i>Standard Methods for the Examination of Water and Wastewater</i> , 19th Edition, 1995, Method 2320 B
Carbonate alkalinity	<i>Standard Methods for the Examination of Water and Wastewater</i> , 19th Edition, 1995, Method 2320 B
Conductivity	<i>Methods for Chemical Analysis of Water and Wastes</i> , EPA-600/4-79-020, March 1983 and subsequent revisions, Method 120.1
pH	<i>Test Methods for Evaluating Solid Waste, Physical/Chemical Methods</i> , EPA SW-846 Third Edition, November 1986, and its updates, Method 9040
Total dissolved solids	<i>Standard Methods for the Examination of Water and Wastewater</i> , 19th Edition, 1995, Method 2540 C
Trace inductively coupled plasma metals	<i>Test Methods for Evaluating Solid Waste, Physical/Chemical Methods</i> , EPA SW-846 Third Edition, November 1986, and its updates, Method 6010B
Anion analysis (Br <sup>-</sup> , Cl <sup>-</sup> , and SO <sub>4</sub> <sup>-</sup> )	<i>Test Methods for Evaluating Solid Waste, Physical/Chemical Methods</i> , EPA SW-846 Third Edition, November 1986, and its updates, Method 9056A
Ion balances	<i>Standard Methods for the Examination of Water and Wastewater</i> , 19th Edition, 1995, Method 1030 F

## **Appendix D**

### **Photos of Spring and Surface-Water Sites**

## Appendix D

### Photos of Spring and Surface-Water Sites

This appendix contains spring and surface-water site photographs that were taken during the June and October 2021 sampling campaigns. The images convey site conditions at the time of sampling and can be useful when interpreting geochemical results (e.g., surface-water sites can be subject to evaporation, which could result in higher concentrations of measured parameters). Refer to Figure 2.4 for the locations of the sampling sites.

#### D.1 Spring Sites

June	October
GSP 1	
	
GSP 2	
	
GSP 3	
	

June	October
<p data-bbox="776 247 847 279">GSP 4</p> 	
<p data-bbox="776 674 847 705">GSP 5</p>  <p data-bbox="370 1066 782 1098">Site was dry; no sample was collected.</p>	
<p data-bbox="776 1146 847 1178">GSP 6</p> 	

June	October
<p data-bbox="776 247 846 281">GSP 7</p> 	
<p data-bbox="776 672 846 705">GSP 8</p> 	
<p data-bbox="776 1096 846 1129">GSP 9</p> 	

June

October

GSP 10



The normal sample collection point for this site (water collection basin in the upper two photos) did not contain sufficient water for sampling. However, water from the spring was flowing into the nearby stock-watering tank, and a sample was collected from the pipe shown in the lower photo.

## D.2 Surface-Water Sites

June	October
GSW 1	
	
GSW 2	
	
GSW 3	
	
Site was dry; no sample was collected.	Site was dry; no sample was collected.

**June**

**October**

GSW 5



Site was dry; no sample was collected.

GSW 6





## **Appendix E**

### **Site and Drainage-Mean Values for Data Collected in 2021**

## Appendix E

### Site and Drainage-Mean Values for Data Collected in 2021

Table E.1. 2021 Reported Values for All Parameters from GAS Wells

Sample	Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)
GAS 1	Mean Value	339	7.60	2825	<0.067	11.6	1285	2155	0.972	120	116	9.08	384
GAS 2	Mean Value	366	8.62	1695	0.0785	12.2	428	1020	0.311	2.98	0.922	2.03	357
GAS 3	Mean Value	334	8.48	2785	0.120	13.8	1008	1860	0.287	9.45	2.78	3.60	598
GAS 7	Mean Value	355	8.72	2040	0.0854	13.6	594	1315	0.297	4.72	1.29	3.08	430

Table E.2. 2021 Reported Values for All Parameters from South Fork Cow Creek Drainage

Sample	Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)
GSP 3	Mean Value	228	7.04	1960	<0.067	7.70	940	1605	1.99	154	141	13.1	84.5
GOW 11	Mean Value	362	7.56	3810	0.180	24.5	2370	3890	1.58	340	370	12.4	205
GNW 6	Mean Value	332	7.60	3690	0.117	23.8	2195	3515	1.46	299	343	11.2	186
GNW 7	Mean Value	432	7.70	1975	<0.067	8.02	478	1510	0.990	161	120	8.13	119
GNW 8	Mean Value	334	7.66	2720	<0.067	11.6	1260	2155	0.825	128	142	8.39	306
Drainage Mean		337	7.51	2831	0.149	15.1	1503	2535	1.37	216	223	10.7	180

Table E.3. 2021 Reported Values for All Parameters from Pony Creek Drainage

Sample	Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)
GSP 4	Single Sample	353	7.39	2010	<0.067	8.33	879	1690	0.707	147	187	16.3	41.8
GSP 5	Not Sampled	—	—	—	—	—	—	—	—	—	—	—	—
GSP 6	Single Sample	636	7.39	3090	<0.067	11.3	1520	2860	1.15	291	271	14.0	112
GOW 3	Single Sample	357	7.61	1800	<0.067	6.04	691	1370	0.591	157	126	8.36	79.2
GOW 4	Single Sample	293	8.53	2250	<0.067	10.6	781	1460	0.233	8.52	4.00	2.61	475
GSP 7	Single Sample	406	8.17	3080	<0.067	22.2	1510	2680	0.885	203	242	14.4	198
GOW 12 <sup>(a)</sup>	Single Sample	707	8.63	1620	0.550	106	<0.133	930	0.436	2.28	0.710	1.82	371
GOW 5 <sup>(a)</sup>	Single Sample	697	8.66	1710	0.686	137	<0.133	971	0.449	2.34	0.815	1.97	389
GSP 8	Single Sample	305	8.87	4990	0.162	44.5	2930	4960	1.15	200	418	9.57	506
Drainage Mean		392	7.99	2870	0.162	17.2	1385	2503	0.786	168	208	10.9	235

(a) Values for GOW 12 and GOW 5 were not included in the drainage mean.

Table E.4. 2021 Reported Values for All Parameters from Cow Creek Drainage

Sample	Comments	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)
PW 736	Mean Value	357	7.53	3165	0.076	15.1	1560	2755	1.12	225	237	12.5	186
GNW 1	Mean Value	342	7.54	2970	0.0918	17.0	1480	2645	1.09	232	219	15.2	161
GSP 1	Mean Value	408	7.58	2840	<0.067	11.8	1410	2435	1.10	222	214	12.2	176
GNW 2	Mean Value	458	7.60	4335	0.0894	24.4	2450	4215	1.62	276	368	20.6	312
GNW 3	Mean Value	410	7.56	3925	<0.067	21.2	2175	3585	1.44	258	323	18.0	272
GOW 1	Mean Value	398	7.62	3800	0.156	20.4	2015	3460	1.24	204	258	11.2	387
GNW 4	Mean Value	416	7.68	3985	0.133	22.2	2285	3765	1.46	270	338	16.8	279
GSP 2	Mean Value	422	7.56	4335	0.134	24.7	2760	4405	1.83	282	412	21.2	340
GNW 5	Mean Value	414	7.54	4150	0.126	19.4	2410	1966	1.62	270	342	18.1	310
W 1 <sup>(a)</sup>	Mean Value	440	7.53	5060	0.173	22.0	3030	4915	1.86	290	400	19.5	488
Drainage Mean		403	7.58	3723	0.115	19.6	2061	3248	1.39	249	301	16.2	269

(a) Values for W 1 were not included in the drainage mean.



## **Appendix F**

### **Analytical Results for Water Samples Collected in 2021**

## Appendix F

### Analytical Results for Water Samples Collected in 2021

Table F.1. Sample Site: GAS 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GAS-1-74	16-Jun-21	343	7.59	2830	<0.067	11.8	1300	2180	0.979	125	118	9.18	391	33.19	34.26	-3.18
GAS-1-75	06-Oct-21	335	7.62	2820	<0.067	11.4	1270	2130	0.966	116	113	8.99	378	31.76	33.46	-5.23
MEAN		339	7.60	2825	<0.067	11.6	1285	2155	0.972	120	116	9.08	384	32.47	33.86	

Table F.2. Sample Site: GAS 2

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GAS-2-74	16-Jun-21	371	8.66	1690	<0.067	12.5	424	1010	0.320	2.99	0.898	2.13	369	16.33	16.60	-1.64
GAS-2-75	06-Oct-21	360	8.59	1700	0.0785	12.0	431	1030	0.302	2.98	0.947	1.93	345	15.28	16.51	-7.73
MEAN		366	8.62	1695	0.0785	12.2	428	1020	0.311	2.98	0.922	2.03	357	15.81	16.56	

Table F.3. Sample Site: GAS 3

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GAS-3-74	15-Jun-21	338	8.50	2860	0.145	14.5	1030	1880	0.294	9.52	2.90	3.60	602	26.99	28.61	-5.83
GAS-3-75	05-Oct-21	330	8.47	2710	0.096	13.2	985	1840	0.280	9.38	2.66	3.60	593	26.57	27.48	-3.35
MEAN		334	8.48	2785	0.120	13.8	1008	1860	0.287	9.45	2.78	3.60	598	26.78	28.05	

Table F.4. Sample Site: GAS 7

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GAS-7-74	16-Jun-21	358	8.65	2040	<0.067	14.0	612	1280	0.303	4.02	1.27	2.38	438	19.42	20.30	-4.42

GAS-7-75	06-Oct-21	352	8.80	2040	0.0854	13.1	575	1350	0.291	5.41	1.31	3.77	421	18.79	19.38	-3.11
MEAN		355	8.72	2040	0.0854	13.6	594	1315	0.297	4.72	1.29	3.08	430	19.10	19.84	

Table F.5. Sample Site: GNW 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-1-74	16-Jun-21	347	7.51	3080	<0.067	17.0	1490	2640	1.14	235	222	15.8	166	37.61	38.44	-2.18
GNW-1-75	06-Oct-21	338	7.58	2860	0.0918	16.9	1470	2650	1.04	229	216	14.7	156	36.36	37.84	-4.00
MEAN		342	7.54	2970	0.0918	16.95	1480	2645	1.09	232	219	15.2	161	36.99	38.14	

Table F.6. Sample Site: GNW 2

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-2-74	16-Jun-21	456	7.58	4430	<0.067	27.4	2390	4160	1.57	270	366	20.1	308	57.49	59.65	-3.69
GNW-2-75	06-Oct-21	460	7.63	4240	0.0894	21.4	2510	4270	1.68	281	370	21.0	316	58.74	62.06	-5.50
MEAN		458	7.60	4335	0.0894	24.4	2450	4215	1.62	276	368	20.6	312	58.12	60.86	

Table F.7. Sample Site: GNW 3

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-3-74	16-Jun-21	415	7.54	4030	<0.067	24.1	2170	3610	1.48	259	329	18.3	276	52.46	54.16	-3.18
GNW-3-75	06-Oct-21	404	7.57	3820	<0.067	18.4	2180	3560	1.39	257	317	17.6	267	50.97	53.99	-5.76
MEAN		410	7.56	3925	<0.067	21.2	2175	3585	1.44	258	323	18.0	272	51.71	54.07	

Table F.8. Sample Site: GNW 4

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-4-74	16-Jun-21	416	7.53	4120	0.181	24.8	2340	3860	1.45	278	349	16.8	284	55.36	57.74	-4.20

GNW-4-75	06-Oct-21	415	7.83	3850	0.0847	19.7	2230	3670	1.46	263	328	16.8	274	52.45	55.28	-5.26
MEAN		416	7.68	3985	0.133	22.2	2285	3765	1.46	270	338	16.8	279	53.91	56.51	

Table F.9. Sample Site: GNW 5

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-5-74	16-Jun-21	424	7.5	4270	0.175	19.5	2430	72.9	1.64	274	347	18.1	314	56.34	59.62	-5.67
GNW-5-75	06-Oct-21	405	7.59	4030	0.0771	19.3	2390	3860	1.61	266	338	18.1	305	54.81	58.40	-6.35
MEAN		414	7.54	4150	0.126	19.4	2410	1966	1.62	270	342	18.1	310	55.57	59.01	

Table F.10. Sample Site: GNW 6

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-6-74	15-Jun-21	329	7.53	3760	0.163	24.9	2220	3570	1.43	308	355	10.9	190	53.12	53.50	-0.73
GNW-6-75	05-Oct-21	335	7.67	3620	0.0715	22.6	2170	3460	1.48	290	331	11.5	181	49.87	52.52	-5.18
MEAN		332	7.60	3690	0.117	23.8	2195	3515	1.46	299	343	11.2	186	51.49	53.01	

Table F.11. Sample Site: GNW 7

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-7-74	17-Jun-21	393	7.85	1970	<0.067	7.91	777	1490	0.920	162	121	7.79	118	23.37	24.26	-3.74
GNW-7-75	06-Oct-21	470	7.54	1980	<0.067	8.13	720	1530	1.06	160	120	8.47	120	23.29	24.62	-5.54
MEAN		432	7.70	1975	<0.067	8.02	748	1510	0.990	161	120	8.13	119	23.33	24.44	

Table F.12. Sample Site: GNW 8

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-8-74	16-Jun-21	337	7.65	2780	<0.067	11.7	1280	2110	0.835	127	141	8.38	310	31.64	33.72	-6.38
GNW-8-75	06-Oct-21	330	7.68	2660	<0.067	11.5	1240	2200	0.815	129	143	8.40	303	31.60	32.74	-3.56
MEAN		334	7.66	2720	<0.067	11.6	1260	2155	0.825	128	142	8.39	306	31.62	33.23	

Table F.13. Sample Site: GNW 9

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GNW-9-74	16-Jun-21	341	8.50	2560	<0.067	12.8	867	1630	0.294	7.93	2.29	3.08	537	24.02	25.23	-4.92
GNW-9-75	06-Oct-21	328	8.43	2460	0.0901	12.6	850	1640	0.284	8.65	2.53	3.41	543	24.35	24.61	-1.08
MEAN		334	8.46	2510	0.0901	12.7	858	1635	0.289	8.29	2.41	3.24	540	24.18	24.92	

Table F.14. Sample Site: GOW 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-1-74	16-Jun-21	400	7.64	3890	0.181	20.4	2040	3430	1.26	209	261	11.3	387	49.02	51.05	-4.05
GOW-1-75	06-Oct-21	395	7.59	3710	0.132	20.3	1990	3490	1.21	199	256	11.1	387	48.11	49.91	-3.67
MEAN		398	7.62	3800	0.156	20.4	2015	3460	1.24	204	258	11.2	387	48.56	50.48	

Table F.15. Sample Site: GOW 3

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-3-74	17-Jun-21	357	7.61	1800	<0.067	6.04	691	1370	0.591	157	126	8.36	79.2	21.86	21.70	0.74

Table F.16. Sample Site: GOW 4

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-4-74	17-Jun-21	293	8.53	2250	<0.067	10.6	781	1460	0.233	8.52	4.00	2.61	475	21.48	22.42	-4.26

Table F.17. Sample Site: GOW 5

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-5-74	17-Jun-21	697	8.66	1710	0.686	137	<0.133	971	0.449	2.34	0.815	1.97	389	17.16	17.80	-3.71

Table F.18. Sample Site: GOW 6

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-6-74	17-Jun-21	281	8.17	1060	<0.067	2.98	304	714	0.480	86.6	68.6	5.32	29.9	11.40	12.03	-5.39

Table F.19. Sample Site: GOW 11

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-11-74	15-Jun-21	368	7.54	3940	0.196	24.7	2430	3810	1.65	355	377	12.1	208	58.08	58.65	-0.97
GOW-11-75	05-Oct-21	355	7.57	3680	0.165	24.3	2310	3970	1.52	325	363	12.8	202	55.19	55.88	-1.24
MEAN		362	7.56	3810	0.180	24.5	2370	3890	1.58	340	370	12.4	205	56.64	57.26	

Table F.20. Sample Site: GOW 12

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GOW-12-74	17-Jun-21	707	8.63	1620	0.550	106	<0.133	930	0.436	2.28	0.710	1.82	371	16.36	17.13	-4.61

Table F.21. Sample Site: GSP 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-1-74	16-Jun-21	413	7.49	2930	<0.067	12.2	1440	2440	1.11	232	222	11.5	183	38.09	38.58	-1.28
GSP-1-75	06-Oct-21	404	7.67	2750	<0.067	11.3	1380	2430	1.09	212	205	13.0	170	35.17	37.13	-5.42
MEAN		408	7.58	2840	<0.067	11.8	1410	2435	1.10	222	214	12.2	176	36.63	37.86	

Table F.22. Sample Site: GSP 2

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-2-74	16-Jun-21	430	7.59	4660	0.169	26.3	2890	4430	1.99	286	425	22.4	365	65.68	69.51	-5.66
GSP-2-75	06-Oct-21	415	7.52	4010	0.0985	23.1	2630	4380	1.67	279	400	20.0	316	61.08	63.71	-4.21
MEAN		422	7.56	4335	0.134	24.7	2760	4405	1.83	282	412	21.2	340	63.38	66.61	

Table F.23. Sample Site: GSP 3

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-3-74	15-Jun-21	224	6.96	2030	<0.067	8.06	985	1590	2.09	164	150	12.8	79.3	24.30	25.21	-3.70
GSP-3-75	05-Oct-21	232	7.13	1890	<0.067	7.34	894	1620	1.89	144	132	13.4	89.7	22.29	23.46	-5.12
MEAN		228	7.04	1960	<0.067	7.70	940	1605	1.99	154	141	13.1	84.5	23.29	24.34	

Table F.24. Sample Site: GSP 4

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-4-74	17-Jun-21	353	7.39	2010	<0.067	8.33	879	1690	0.707	147	187	16.3	41.8	24.95	25.60	-2.54

Table F.25. Sample Site: GSP 6

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-6-74	17-Jun-21	636	7.39	3090	<0.067	11.3	1520	2860	1.15	291	271	14.0	112	42.04	44.68	-6.09

Table F.26. Sample Site: GSP 7

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-7-74	17-Jun-21	406	8.17	3080	<0.067	22.2	1510	2680	0.885	203	242	14.4	198	39.02	40.18	-2.94

Table F.27. Sample Site: GSP 8

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-8-74	17-Jun-21	305	8.87	4990	0.162	44.5	2930	4960	1.15	200	418	9.57	506	66.62	68.36	-2.58

Table F.28. Sample Site: GSP 9

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-9-74	17-Jun-21	274	7.29	1610	<0.067	9.50	682	1230	0.677	120	132	17.6	22.5	18.28	19.95	-8.75

Table F.29. Sample Site: GSP 10

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSP-10-74	17-Jun-21	369	7.54	1250	<0.067	7.65	316	846	0.528	94.8	86.0	5.67	41.1	13.74	14.17	-3.13

Table F.30. Sample Site: GSW 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSW-1-74	16-Jun-21	726	7.9	5650	<0.067	4.15	3430	5710	3.43	408	572	11.1	410	85.53	86.05	-0.60
GSW-1-75	06-Oct-21	632	7.9	7860	0.091	147	5490	9710	3.39	417	783	89.3	840	124.04	131.09	-5.52
MEAN		679	7.9	6755	0.091	75.6	4460	7710	3.41	412	678	50.2	625	104.79	108.57	

Table F.31. Sample Site: GSW 2

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSW-2-74	16-Jun-21	407	8.05	4030	0.143	22.6	2310	3710	1.87	277	358	18.5	299	56.75	56.87	-0.21
GSW-2-75	06-Oct-21	445	8.13	5510	0.188	35.2	3190	5000	2.39	304	488	34.2	431	74.94	76.31	-1.82
MEAN		426	8.09	4770	0.166	28.9	2750	4355	2.13	290	423	26.4	365	65.84	66.59	

Table F.32. Sample Site: GSW 5

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSW-5-74	16-Jun-21	906	7.60	6490	0.270	39.1	3940	6860	2.96	392	681	33.5	557	100.67	101.25	-0.58

Table F.33. Sample Site: GSW 6

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K(ppm)	Na(ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
GSW-6-74	16-Jun-21	837	8.14	7920	0.237	29.8	4950	8510	3.28	376	699	23.8	1040	122.11	120.64	1.21
GSW-6-75	06-Oct-21	617	7.66	16200	0.174	66.7	11300	18100	4.97	399	1610	48.8	2370	256.69	249.49	2.85
MEAN		727	7.90	12060	0.206	48.2	8125	13305	4.12	388	1154	36.3	1705	189.40	185.06	

Table F.34. Sample Site: W 1

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
W-1-74	16-Jun-21	444	7.52	5170	0.209	21.9	3110	4970	1.79	290	405	19.2	495	69.81	74.25	-6.16
W-1-75	06-Oct-21	436	7.54	4950	0.137	22.0	2950	4860	1.93	289	394	19.8	481	68.26	70.76	-3.60
MEAN		440	7.53	5060	0.173	22.0	3030	4915	1.86	290	400	19.5	488	69.04	72.50	

Table F.35. Sample Site: PW 734

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
PW-734-74	16-Jun-21	361	7.99	4000	<0.067	15.1	2160	3770	1.77	208	276	16.3	397	50.77	52.62	-3.57

Table F.36. Sample Site: PW 735

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
PW-735-74	16-Jun-21	430	7.38	4670	0.134	21.7	2580	4590	1.97	270	358	17.9	417	61.52	62.93	-2.26
PW-735-75	06-Oct-21	422	7.44	5000	0.150	21.4	2690	4620	1.88	272	381	19.9	416	63.52	65.05	-2.38
MEAN		426	7.41	4835	0.142	21.6	2635	4605	1.92	271	370	18.9	416	62.52	63.99	

Table F.37. Sample Site: PW-736

Sample	Date Sampled	Alkalinity (as ppm CaCO <sub>3</sub> )	pH	Cond (μS/cm)	Br- (ppm)	Cl- (ppm)	SO <sub>4</sub> = (ppm)	TDS (ppm)	B (ppm)	Ca (ppm)	Mg (ppm)	K (ppm)	Na (ppm)	Cation Sum (meq)	Anion Sum (meq)	% Error
PW-736-74	16-Jun-21	359	7.47	3060	<0.067	14.9	1550	2710	1.16	225	235	12.0	188	39.04	39.87	-2.10
PW-736-75	06-Oct-21	355	7.59	3270	0.076	15.3	1570	2800	1.09	225	239	13.0	185	39.27	40.22	-2.39
MEAN		357	7.53	3165	0.076	15.1	1560	2755	1.12	225	237	12.5	186	39.16	40.04	

## **Appendix G**

### **Statistical Analysis – Methods and Results**

# Appendix G

## Statistical Analysis – Methods and Results

---

*This appendix contains the results of the statistical analyses and an overview of the methods used to conduct them. For more technical documentation of these methods, readers with a deeper statistical background are referred to Appendix F in Colstrip reports of years prior to 2015 (the most complete of these being that in Thompson et al. [2014] in which random field assumptions and expected mean square error of the sampling regime are addressed) and Chamberlain (2018). There is no required statistical background for this appendix, but there are a few terms and concepts whose coverage is outside of the scope of this report—to understand them readers may consult an introductory statistics textbook such as McClave and Sincich (2016).*

---

We performed statistical analysis to 1) determine if data exhibit differences among drainages or manifest recent changes over time and 2) characterize the water quality for each drainage. Specific objectives were as follows:

- Identify any recent time trends in individual site data and for each drainage overall.
- Detect any statistically significant differences between overall parameter means in the Cow Creek, South Fork Cow Creek, and Pony Creek drainages. Pony Creek is used as a “control” drainage, presumed to be unaffected by EHP activities, and is hydrologically and geochemically similar to both Cow Creek drainages. However, water quality in a drainage may change as a result of mining activity taking place therein.
- Estimate overall drainage means for the parameters of interest.

### G.1 Data Used

#### G.1.1 Sample Sites

The parameters conductivity ( $\mu\text{mhos/cm}$ ), boron (ppm), and sulfate (ppm) were examined in the statistical analysis. The sampling sites, listed in Table G.1, were groundwater wells and springs that were selected as being representative of the aquifers of interest.

The sites are ranked in terms of their distance from the headwaters of their respective drainages, with the shortest distance listed first. Wells PW 735 and PW 736, located between the EHP and Cow Creek, were installed in 2009, and the analysis results from sampling there are reported by Thompson et al. (2010–2020) and in Appendix F of this report. An additional site that is downgradient from the confluence of Cow Creek and South Fork Cow Creek drainages (W-1) was not included in the analysis because its location does not support drainage comparisons, and in the presence of time-drainage interactions (Olsen et al. 1992), time comparisons inclusive of this well are not statistically proper. Thus, its inclusion was not appropriate in the drainage comparisons or time-trend analyses. However, because of its location downstream of the confluence of the Cow Creek and South Cow Creek drainages, the cumulative site W-1 data could be useful as a composite representation of selected sites in the lower Cow and South Fork Cow Creek drainages, in a comparison with downstream Pony Creek sites (currently outside the scope of our study).

**Table G.1. Sampling Sites**

Cow Creek	Pony Creek	South Fork Cow Creek
PW 735	GSP 4	GSP 3
PW 736	GSP 6	GOW 11
GNW 1	GOW 3	GNW 6
GSP 1	GOW 4	GNW 7
GNW 2	GSP 7	GNW 8
GNW 3	GOW 12	--
GOW 1	GSP 8	--
GNW 4	--	--
GSP 2	--	--
GNW 5	--	--

GNW = Genie new well; GOW = Genie old well; GSP = Genie spring;  
PW = private monitoring well.

### G.1.2 Identification and Removal of Outliers

This section addresses the removal of inordinately extreme data (high or low) from the time-trend analyses and drainage comparisons and describes the outlier identification method used to identify these values. In the following discussion, we first identify the outlying data that were excluded, then we describe the method used to detect these anomalous data.

Data used in the analysis are presented in Appendix F (2021 sampling data), Olsen et al. (1992–2006), Thompson et al. (2008–2021), and McDonald et al. (2007). Data plots at the site level are given later in Section G.3.4, along with the results of the time-trend analysis. Extremely high or low data, as detected in the analysis of outliers or as identified by PNNL staff, were then examined for validity by scrutinizing sampling records and laboratory QA/QC information. Conclusively anomalous data were then removed from the time-trend analysis as being “outliers.”

Extremely large or small data values were excluded from the analysis according to the methodology and procedure discussed below. Data from GOW 12 since sampling event 60 were again omitted from the drainage time-trend and drainage comparisons, along with the 2017-2018 sampling results for GSP 7 because all of these values are unrepresentative of the Pony Creek aquifer, and therefore their site trend testing was suspended pending the acquisition of additional data and information (Thompson et al., 2018–2021).

The omitted data identified in this and previous years (discussed in earlier reports) are provided in Table G.2, where -36, for example, indicates sampling event 36. All of these data were excluded from the time-trend analyses on the basis of their being classified as spuriously extreme data resulting from unknown causes and not in any way related to the (unknown) actual values.

*In this year’s outlier investigation, two sulfate values were declared to be outliers—sulfate at GOW 3-10, and at GOW 11-11—because they were inordinately low and inconsistent with the adjacent years data.*

The statistical detection of outliers was accomplished by direct inspection of data, comparing data with smoothed values using the LOWESS model fit (discussed below) in the “R” data analysis system (R Development Core Team 2022), examining regression residuals from the smoothed fit (i.e., data values minus fitted values, for each sampling event), and using a formal (objective) outlier rejection rule. The rule stipulates that all data whose residuals from a robust trend fit that are farther than 5.2 median

absolute deviations (MADs) of the residuals from the median residual (approximately zero), should be rejected (X84 outlier rejection rule of Hampel et al. 1986).

**Table G.2.** Data Omitted from the Analysis<sup>(a)</sup>

<b>Cow Creek</b>	<b>Pony Creek</b>	<b>South Fork Cow Creek</b>
GNW 1 -9 (cond) -19 (boron) -36 (boron)	GSP 4 -2 (cond.) -24 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -28 (cond.) -36 (cond.) -38 (cond.) -42 (cond.) -48 (boron)	GSP 3 -4 (cond.) -16 (boron) -36 (boron) -48 (cond., SO <sub>4</sub> <sup>-</sup> ) -54 (cond., SO <sub>4</sub> <sup>-</sup> ) -58 (cond., SO <sub>4</sub> <sup>-</sup> )
GSP 1 -20 (boron) -21 (cond., SO <sub>4</sub> <sup>-</sup> ) -23 (cond., boron, SO <sub>4</sub> <sup>-</sup> )	GSP 6 -13 (boron) -36 (boron)	GNW 6 -15 (boron) -42 (boron) -70 (SO <sub>4</sub> <sup>-</sup> )
	GOW 3 -2 (SO <sub>4</sub> <sup>-</sup> ) -10 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -12 (boron) -18 (boron) -20 (SO <sub>4</sub> <sup>-</sup> )	GNW 7 -1 (cond) -3 (SO <sub>4</sub> <sup>-</sup> )
GNW 2 -44 (SO <sub>4</sub> <sup>-</sup> )	GOW 4 None omitted	GNW 8 -11 (boron) -13 (boron) -36 (boron) -41 (SO <sub>4</sub> <sup>-</sup> )
GNW 3 -13 (boron) -20 (SO <sub>4</sub> <sup>-</sup> ) -36 (boron) -54 (SO <sub>4</sub> <sup>-</sup> )	GSP 7 -2 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -16 (boron) -66 through -69 (cond., boron, SO <sub>4</sub> <sup>-</sup> )	GOW 11 -11 (SO <sub>4</sub> <sup>-</sup> ) -15 (boron) -44 (cond) -60 (cond.)
GOW 1 -13 (boron)		
GNW 4 None omitted	GOW 12 -5 (cond., SO <sub>4</sub> <sup>-</sup> ) -18 (cond.) -10 (cond., SO <sub>4</sub> <sup>-</sup> ) -13 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -18 (SO <sub>4</sub> <sup>-</sup> ) -61 through -69 (cond., boron, SO <sub>4</sub> <sup>-</sup> )	
GSP 2 -13 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -15 (boron) -16 (boron) -24 (boron) -60 (cond., boron, SO <sub>4</sub> <sup>-</sup> )	GSP 8 -12 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -14 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -16 (SO <sub>4</sub> <sup>-</sup> ) -18 (cond., boron, SO <sub>4</sub> <sup>-</sup> ) -30 (SO <sub>4</sub> <sup>-</sup> )	
GNW 5 None omitted		

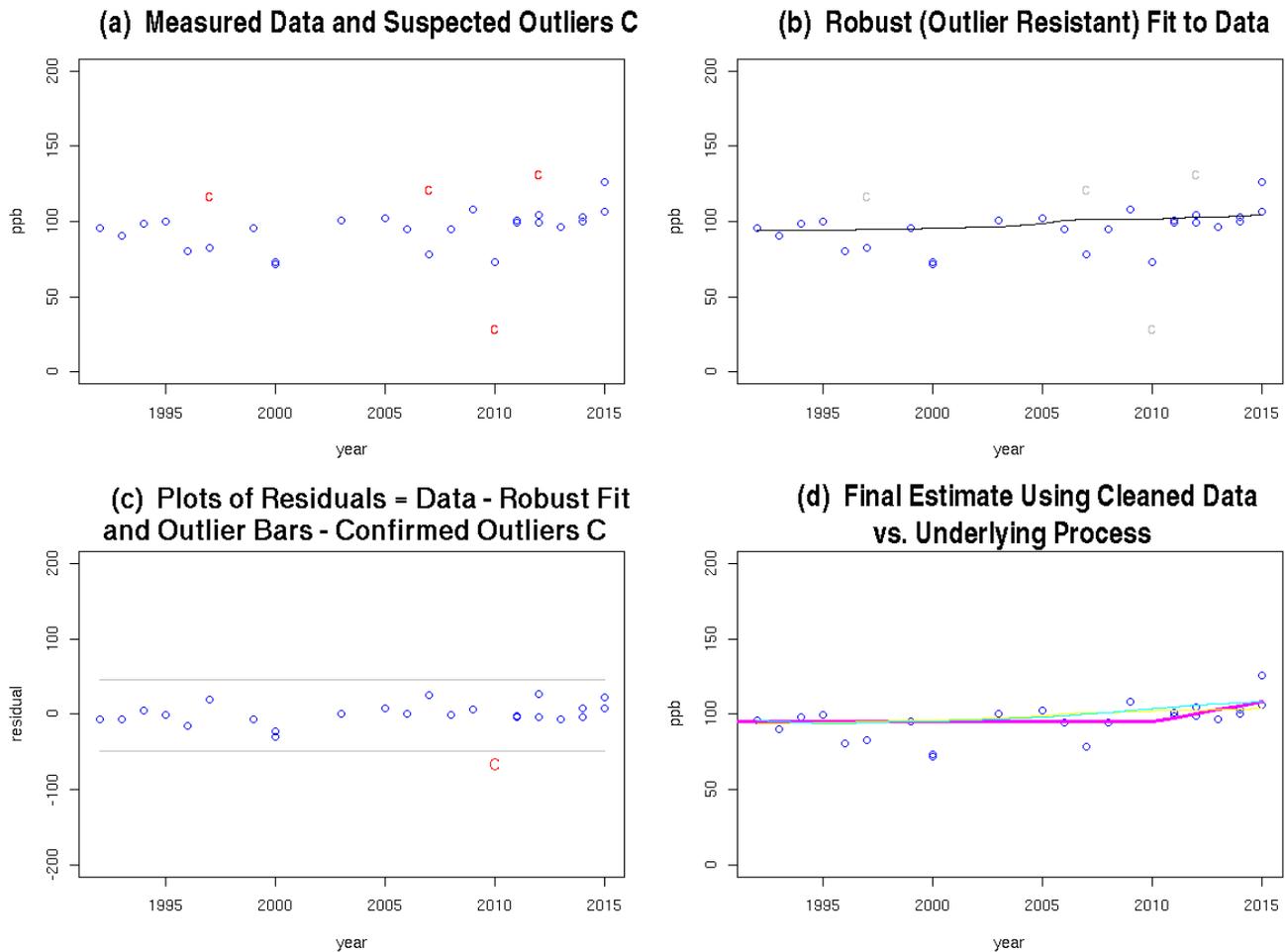
(a) Specific sampling events are indicated by a hyphen and the sampling event number.

A robust estimator of the mean or some other feature of the data's probability distribution is one that is not influenced when a certain proportion of extreme-valued data points is encountered; that estimator is not

sensitive to spuriously large or small values. Therefore, it remains useful as a reference measure of the remaining “clean data,” even when the outlying data are extreme.

In our example, a robust fit over time was applied to the cumulative data from each site (using the locally weighted scatterplot smoothing [LOWESS] model [Cleveland 1981]). The fitted lines and data were plotted together in order to visually inspect for outlying values. Then 1) the regression residuals (i.e., the differences between the data values and the fitted lines evaluated at the data sampling times) were obtained, 2) the median (MED) and MAD of the residuals were computed, and 3) data whose corresponding residuals were outside of the interval defined by  $MED \pm 5.2 \text{ MAD}$  were considered probable outliers.

The procedure, originally described by Thompson et al. (2015), is again depicted in Figure G.1 for a hypothetical example over the years 1990–2015. The data are plotted in Figure G.1(a) and, based on visual observation, suspiciously spurious data—outliers—are labeled “C”. The robust LOWESS fit is overlaid with the data in Figure G.1(b), and its corresponding plot of residuals is given in Figure G.1(c), along with the outlier identification lines. Data whose residuals are farther from zero than  $\pm 5.2 \text{ MAD}$  (flat lines)—in this case only, one data point, labeled “C” in Figure G.1(c)—are outliers according to the above X84 rule.



**Figure G.1.** Procedure for Outlier Analysis – Hypothetical Example (Thompson et al. 2015)

Data identified as outliers were then examined for validity by scrutinizing sampling records and laboratory QA/QC information. In recent years, in this study this rule has been used to determine a definitive acceptance region, rather than a rejection region; i.e., values within the 5.2 MADs from the median are definitely retained in the analysis, and those that are moderately more outlying are usually retained as well, instead of strictly rejecting all data outside the region. Also, data were retained when there was evidence of their validity. For example, there are instances in which the robust LOWESS smoother is too insensitive, not following the actual data, as at several springs (e.g., GSP 7 data from sampling campaigns in the 1990s), where it failed to adequately represent historically confirmed peaks—this resulted in some residuals being (artificially) outside of the  $\pm 5.2$  MAD envelope. In such cases, the data were typically retained in all analyses.

It should be emphasized that the analysis and detection of outliers is focused on finding "false" data that are spuriously high or low, that have been generated from some other probability distribution than that of the actual data. The analysis is not useful for the more difficult task of identifying nonextreme data that derive from a probability distribution that has the same range as the actual data but is skewed or otherwise different as a result of some artificial causal mechanism. There are methods for this type of identification, but this effort is beyond the scope of our study. Had there been notable bias of results due to either field sampling or laboratory analysis, such an inquiry might be justifiable and in such cases it may be possible to transform the skewed data to represent typical values, even in cases of nonzero time trends, rather than remove those data from the analyses. This potential to recover the actual data would also apply to systematic outliers, so in Thompson et al. (2016), a cursory examination of the distributions of outliers (as identified by "o" in Figure G.9) was performed for each applicable site and constituent; there was no conclusive evidence of any systematic causes, largely due to a (fortuitously) small number of outliers. Also the QC results consistently indicate the nonexistence of any bias at the laboratory level. Based on the discussion in Section 3.6.2, the 2017 through 2019 GSP 7 data and GOW 12 data after sampling campaign 60, were qualitatively judged to be systematically outlying; therefore, no application of analysis was necessary to confirm their status as (systematic) outliers.

After the outlying data were removed, another model—SUPSMU—was fitted to the outlier-cleaned data as in Figure G.1(d). The supersmooth (i.e., SUPSMU) estimator reported by Friedman (1984), is a curve-fitting approach in which, for each given point (e.g., time) a straight line is fit using the data from nearby points only, and the final predictions are obtained from (the midpoints on) a moving series of shorter lines that steepen or flatten to accommodate the behavior of the response data (e.g., conductivity). The window widths for the data to be included may change from one time point to the next and are selected so that the local lines are sufficiently sensitive to changes in the data, but not so much as to "chase" them too closely, causing unduly high uncertainty in the predictions. One significant advantage of the SUPSMU approach is that no particular functional form is preselected or imposed on all of the data—the method largely maintains any changes in historical data while flexibly accommodating new changes, which may be either more or less dramatic than historical ones and may also tend in a different direction (increasing or decreasing). For representing "cleaned" data, the SUPSMU model is preferred over the LOWESS model, because SUPSMU has an appropriately higher data sensitivity when the data contain no outliers.

## **G.2 Statistical Testing Posture: Assume No Impact Unless Refuted by Data**

All conclusions regarding temporal changes and in the comparisons of drainages are based on classical statistical testing, which *assumes a default condition or hypothesis, and only abandons the belief that it holds if the observed data are highly incompatible with that condition*. The default hypotheses in this report are 1) there is no increasing or high-valued trend (no decreasing one when testing for decreases),

and 2) in the drainage-comparison tests, the drainage constituent means—based on all historical data pooled together—are all equal.

### G.3 Time-Trend Testing and Drainage Comparisons

*In this report, trend means time-varying, true value.* Our use of the term "trend" here instead means a time-representation of a constituent level, be it cyclic, increasing, decreasing, etc. Also though, our trend is the true, underlying level. The site-level trend is the true level, at a given site, of the constituent at each sampling time: sampling data always contain some "error" component, due to uncontrollable variation in both sampling and analysis, but it's desirable that the data are close to the underlying trend values. This time trend aka "temporal trend" is not necessarily always-increasing or -decreasing over time, or even time-varying, but merely represents the true value at any point in time. Similarly, the term drainage trend does not imply a specific direction of change, only that the true drainage mean over time is being represented.

**Recently Increasing or Decreasing Trends.** (Recently) Increasing or decreasing trends are, in this report, defined as those occurring continuously over at least the most recent 4 years (or eight sampling campaigns). This requirement reflects the expected nature of a contaminant plume while excluding spurious changes over only a few years that are weather-related. This requirement applies to both site- and drainage-level trend direction testing.

#### G.3.1 Testing for Recent Time Trends Using Our Conventional Test

This subsection describes our "conventional test", applied in Colstrip Reports since the mid 1990s. As discussed later the PLR test, first used in Colstrip trend testing in 2018, was additionally applied to Stipulation 12(d) constituents and the results of both tests were examined and a composited conclusion was made. This section however describes the conventional test, in detail, which addresses the following questions: "... are any recent, increasing trends at the site and drainage level valid? Or did we estimate increases 'by chance' while their true level is instead constant over time?" Because the data have some degree of uncertainty, fitted trends are also uncertain; so conclusions, based on visual inspection, that trend directions are truly representative could be highly unreliable (especially when the data vary greatly around a fitted trend line). So we instead examine *the probability of observing an increase in the fitted trend*. We answer the first question above by assuming (in our statistical distribution) that the second one holds true—the true level is actually constant—and if an increasing trend was fitted we will abandon this assumption and conclude that there is indeed an increasing trend if the probability of observing that increasing trend "by chance" is small.

The testing in this section involves 1) computing the above probability, and 2) concluding that a fitted increase, at either the site or drainage level, is valid only if that probability of a "by-chance" occurrence is less than the customary threshold value of 5%. Otherwise the observed increase is concluded to be not valid, having occurred merely by chance—a chance event having a probability greater than 5%. In the same manner we also address the case of decreasing trends, because actual decreases give us insight into a contaminant plume's movement and intensity. Use of this customary value ensures that if our assumption is true, the error rate of incorrectly deciding that an increasing (or decreasing) trend exists is small—less than 5%.

First we must define an increasing trend. An increasing time trend is defined as a continual increase in the fitted time trend (SUPSMU mentioned above) for the most recent 4 or more years. This hypothesis was constructed to provide the earliest detection of recent trends subject to requiring sufficient data for such detection. Ideally, a change in water quality would be detected very early, but with the expectation that a contaminant plume exhibits the gradual trend that commonly occurs at a site under typical conditions, 4

years was selected as the shortest duration that would be considered for the trend to occur. Shorter periods would cause our definition to be overly sensitive to sporadic increases due to sampling/measurement variation or small hydrogeologic changes. See Thompson et al. (2014, Appendix F.2) for further discussion of the groundwater transport mechanisms that would induce varying degrees of increase. Likewise, a decreasing trend requires consecutive decreases in the fitted value for 4 or more recent years.

We compute our test probabilities using a form of Monte Carlo simulation, a computational method in which in general all uncertain factors are randomly generated *repeatedly* from their respective probability distributions, and in each repetition or *iteration*, the equation or mathematical model of interest is tabulated using the generated factors for that iteration. For a large number of iterations, the collection of these tabulated results then represents the *statistical population*—all possible occurrences—of the equation being simulated. In our case the data are the uncertain elements, and for each site or drainage a data series representative of the entire constituent data history is randomly generated repeatedly—100,000 times—under the assumption of no recent trend, and in each iteration a SUPSMU trend curve is fitted to those data. The proportion of the total population (i.e., 100,000) of fitted trends that show recent increases as defined above represents the probability of fitting an increase “by chance.” If this probability is less than 5%, we reject the assumed “by-chance” occurrence and instead conclude that the increasing trend in the originally fitted curve is valid; i.e., the constituent is increasing at that site or drainage.

The details of our simulation and probability testing are illustrated in Figure G.2 and were originally presented by Thompson et al. [2015]) using hypothetical data and are given in the following steps:

1. Examine the fitted trend for consistent changes (see Figure G.2a). Fit a trend to the outlier-cleaned data (using the SUPSMU regression model) and examine this trend for continual increases and continual decreases over the most recent 4 or more years. If no such changes are seen in the trend line, then do not test—the conclusion is that no trend was found. Alternatively, if a continual change is found, such as the increase since 2010 in Figure G.2a, then compute the probability of a chance occurrence and continue with Steps 2 through 5.
2. Compute the trend residuals (see Figure G.2b) by subtracting the fitted trend from the data value at each sampling event date.
3. Construct the assumed trend line (see Figure G.2c). Assume that no recently changing trend exists and construct this hypothesized “true” trend line as 1) the fitted trend for the pre-change period (e.g., before 2010), and 2) the average over the recent trend period, of the fitted trend; use this average for the entire change period as in Figure G.2c.
4. Generate the population of 100,000 fitted trends (see Figure G.2d) using the following procedure. In each Monte Carlo iteration, generate a new data series representing the entire sampling history—another data history that might have occurred—under the assumption that no recent continual trend has occurred, and fit a trend to this series (representing a trend that could have occurred under the same assumption). Instead of generating this data series from a probability distribution as in an ordinary Monte Carlo simulation, randomly resample with replacement, a new series of residuals from the computed residuals; that is, for each time point, select a residual “out of a hat that contains our residuals,” add that residual to the hypothesized trend (from Step 3) to obtain a simulated data value, then return the residual “back into the hat” for possible reselection at another time point.
5. Figure G.2d shows one series of these reconstructed data.
6. Figure G.2, subplots e and i through h and l illustrate four iterations of our approach to Monte Carlo simulation. The resampled residuals for the four iterations are given in Figure G.2, subplots e through h and the corresponding trends fitted to data reconstructed from those residuals are given in Figure G.2, subplots i through l.

7. In each iteration, examine the fitted trend for an increase over the originally identified period (of increase). (If the original trend exhibited a consistent decrease, evaluate the simulated trend fits for decreases instead.)
8. Compute the probability of spuriously occurring trend increases as follows. After the Monte Carlo iterations are completed, compute the proportion of the 100,000 trend fits that exhibited an increase. If that proportion is less than 5%, abandon the position that no increase is occurring and conclude that the originally fitted recent increase is valid (because if the no-increase assumption were true, it is unlikely [less than a 5% chance] that an increasing trend would be fitted). Use of the 5% rule ensures that the “false-detect” error probability is under 5%.
9. As an example of this probability computation, if we tested for an increase using only the four iterations (we actually use 100,000 iterations to obtain a better probability estimate) in Figure G.2, subplots i through l), we note that in one of those iterations we fitted an increase, so our estimated probability of such an increase occurring by chance is one in four, or 25%. So we would say that spurious increases are too likely (chance of 1 in 4, i.e., 25%)—more likely than 5%—for our concluding that our original increase is valid, so we maintain that there is actually no increasing trend. Alternatively, if we again based our conclusion on only four iterations but we instead found that no increases were fitted among them we would conclude that the true trend is an increasing one because our estimated probability of obtaining an increasing fit (if the no-trend hypothesis is true) would be 0 in 4 (i.e., 0%, which is less than 5%). While using just a few iterations has a simplistic appeal, the foregoing estimated probability will vary widely between groups of simulations of four iterations, even from 0% to 100% (i.e., it has a very high uncertainty level). So we instead use many iterations (100,000) to obtain a stable probability estimate which also closely estimates the true probability of fitting an increasing trend (when none exists).

The foregoing example for testing applied to an increasing trend. For sulfate (all sites) and at each drainage, we test for either an increasing trend or a decreasing one, depending upon the direction (of the most recent 4 years) of the originally estimated trend – if this trend is decreasing for 4 or more of the most recent years, we test for a decreasing trend (the number of *decreasing* trends in the above Monte Carlo simulation is instead used in computing the p-value). However, for the two-directional tests, we begin with an allowable by-chance error of 2.5% for each test in the decided direction instead of 5% which we again reduce as explained in Section G.3.3.2 of this report.

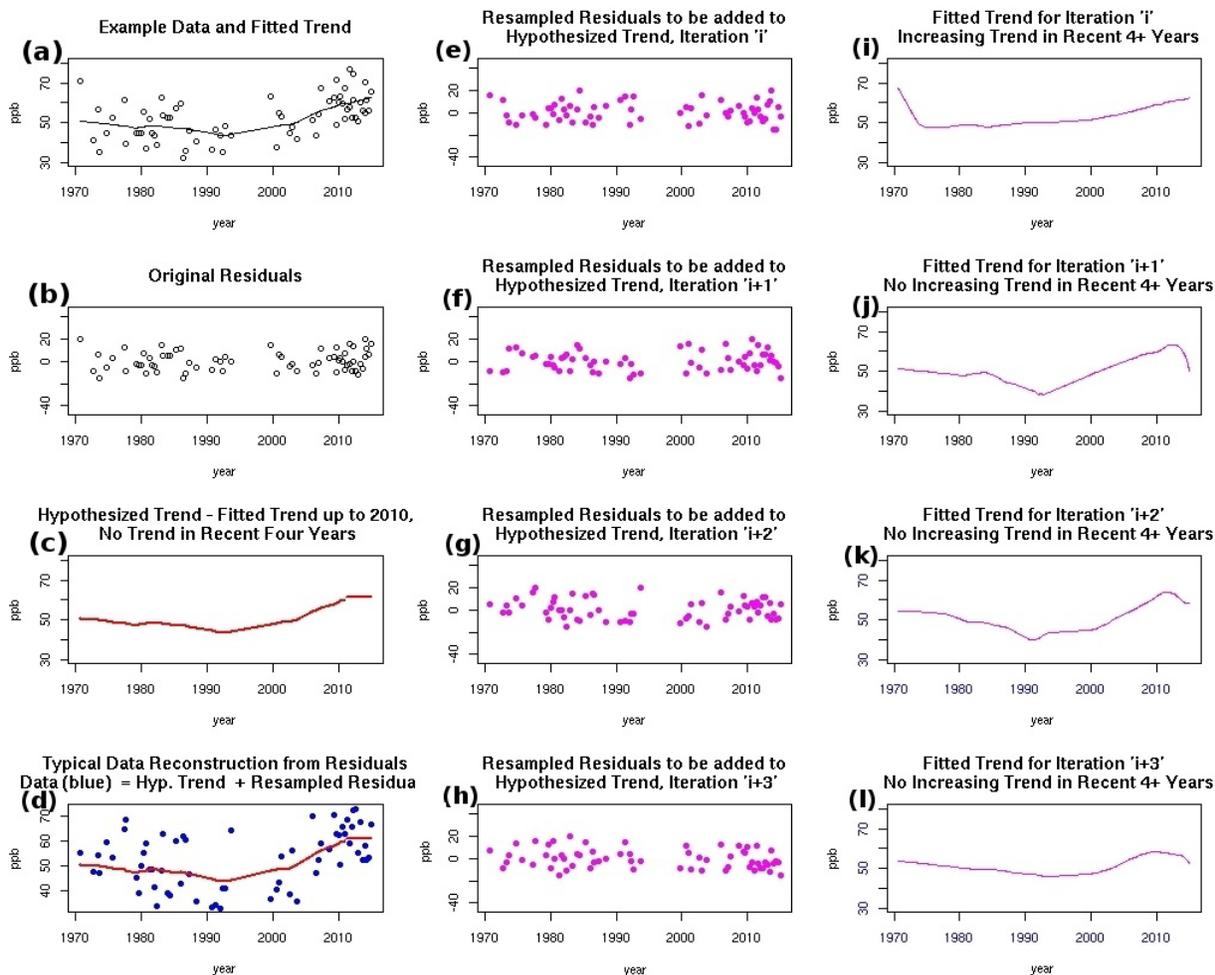
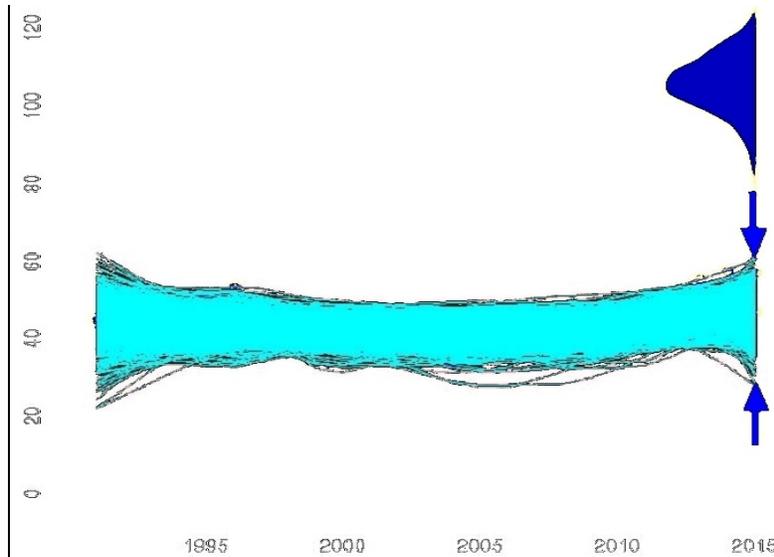


Figure G.2. Procedure for Testing Increasing Trends – Hypothetical Example

### G.3.2 Model Uncertainties – Minimum and Maximum Trend Values at Each Time Point

In addition to the testing described in the previous section, minimum and maximum “possible” estimates were tabulated to describe the uncertainty in our trends (fitted to the original data) using Monte Carlo simulation. But in this simulation, no assumption was made concerning a recent constant mean; the original estimated trend was used as our best estimate of the mean over the entire time history. As with the testing simulation, in each iteration resampled residuals were added to a trend—in this case, the originally estimated trend—to obtain a simulated set of data, and then a trend was fitted to these data. For each time point, the minimum and maximum of the (simulated population of) fitted trends was obtained and plotted along with the data and originally fitted trends. Doing so formed an envelope around the trends that describes, for each time point, the minimum and maximum values that could have occurred from this estimation procedure.

Figure G.3 (Thompson et al. 2015) depicts simulation results for the same generic data in Figure G.2 (i.e., plots of the fitted trend from each iteration). Also, as is shown here for only year 2015, probability distributions of the model could be obtained at some or all of the other time points as well. However, for the site trend fits illustrated in the next sections, we are reporting only the extreme-most values instead of estimated probability distributions.



**Figure G.3.** Uncertainty Assessment on Fitted Model: Overlay of all Simulation Results and Probability Distribution on 2015 Predictions – Hypothetical Data (Thompson et al. 2015)

### G.3.3 Testing for Increasing Trends in Conductivity or Boron: PLR Test Compositing with the Standard Trend Test

#### G.3.3.1 PLR Test for Increasing or Very High Trends

In statistical testing between the default ("null") hypothesis and an alternative hypothesis, the most fundamental test is the Likelihood Ratio test. Most hypothesis tests are based on this one, the actual forms of their *test statistic* being actually derived from an underlying Likelihood Ratio. Originally introduced in the 1930s, this ratio and its test have become the standard basis for statistical hypothesis testing, largely because when applied to two simple competing hypotheses, its ability to detect when the alternative hypothesis is true is maximum possible. This statistic is the ratio of the likelihood of the data's occurring if the alternative is true to the likelihood of their occurrence if the null hypothesis is true: it is also known as the *odds ratio* (unfortunately though this latter reference falsely suggests that we may merely "examine the odds" and make a conclusion, it will presently be shown that the actual testing involves instead determining the *probability distribution* of this ratio and where the observed one scores in that distribution). Also, data are often continuously distributed – they can occur over a continuum rather than taking on discrete values (i.e., whole numbers) so their probability of occurring at any exact value is infinitesimally small: analogous to being represented by a probability their distribution is instead represented as a probability curve (i.e., a "bell curve" or other-shaped one), known as a likelihood function when testing, which is evaluated at the data value. This curve will be different, in shape and/or location, for each hypothesis. Also, if each data point's curve is unaffected by previously measured data, the data are considered to be mutually independent, in which case the likelihood of the joint occurrence of the values all the data have taken is the product of their individual curve values. Therefore, the Likelihood Ratio is usually a ratio of two products of hypothesis-specific curve values.

The Likelihood Ratio is itself a statistic, having a probability curve of its own, and that curve will also be different depending upon which hypothesis is true. If the alternative hypothesis is true, the Likelihood Ratio will tend to be large because the data fit the numerator likelihood better than the denominator one; that is, the curve values in the numerator's product will (on average, geometrically) be higher than those in the denominator's one. But if the null hypothesis is true, the Likelihood Ratio will tend to be small,

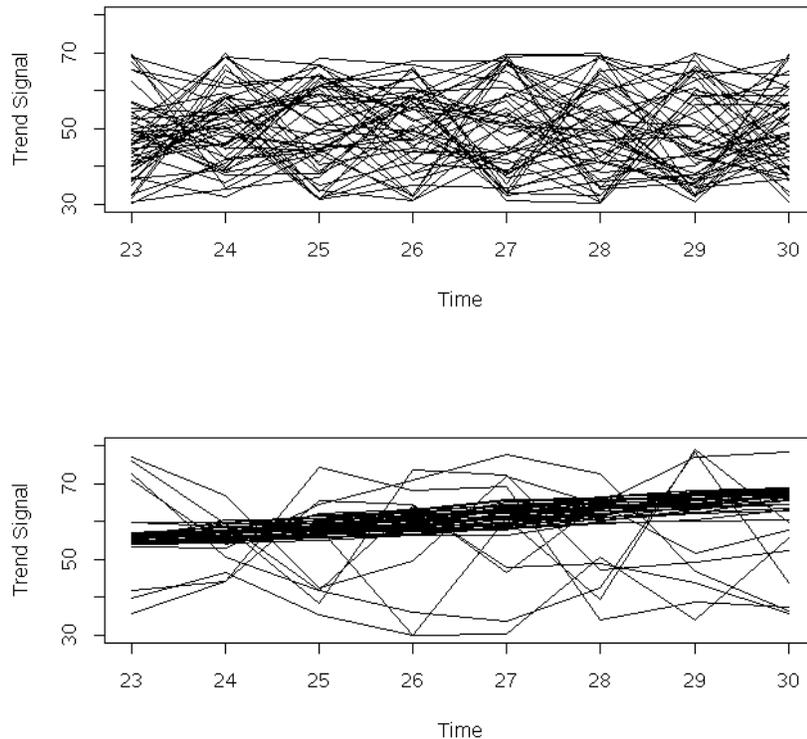
reflecting the better fit of the denominator's Joint Likelihood. However, the two hypothesis-specific ranges that the ratio can take often overlap somewhat, sometimes considerably.

The decision in testing the null hypothesis using the Likelihood Ratio is much the same as the testing decision described in Section G.3.1: the null hypothesis is assumed to be true and if the probability (under that assumption) of the Likelihood Ratio's being as large as (or larger than) the one observed (i.e., the ratio's p-value) is smaller than 5%, we conclude that the ratio is very unlikely to have occurred “by chance” under the null assumption, so we reject the null hypothesis and conclude that the alternative one is true. In this part of the discussion, we again preselect 5% as our acceptable “false-detect” rate: under the null hypothesis, spuriously large Likelihood Ratio values that have p-values less than 5%, and thus trigger (falsely) rejecting the null assumption, occur only 5% of the time. Finally, computing the p-value requires estimating the distribution of the Likelihood Ratio when the null hypothesis is true: in our more complex application, this estimated distribution is obtained by Monte Carlo sampling; that is, by repeatedly generating a series of data under the null hypothesis assumption and evaluating the resulting Likelihood Ratio in each iteration. The resulting “ensemble” of values approximates the population of all possible Likelihood Ratios that could occur if the null hypothesis was true, and if less than 5% of them are larger than the one computed from the original data, we reject the null hypothesis. In testing multiple hypotheses, a false reject rate lower than 5% may be specified for each individual test so that the overall error rate—of one or more false rejects—does not exceed 5%.

The Likelihood Ratio test would be useful if we were testing the null hypothesis of no trend (as in our customary test) versus a specific increasing trend. However in testing the condition of Stipulation 12d where an abrupt increase at Cow Creek is cause for remedial intervention there, in addition to a flat trend's being acceptable, we are also not concerned about small trend oscillations and decreasing trends, whereas large “jumps” in trend and possibly continually increasing ones, trigger intervention. Therefore, the candidate trendforms under each hypothesis are numerous—actually infinite—and the simple Likelihood Ratio framework will not accommodate this realistic complexity. Hence, the PLR (Chamberlain 2018) was employed, as the ratio of Predictive Likelihoods under each hypothesis. If all possible trendforms over the four most recent years are approximated by “dot-to-dot” connections between their values at the eight sampling campaign times, similar to the plots of Figure G.4 (but over the applicable most recent years), they may then be grouped by those that increase relative to a pre-trend reference value (i.e., 50 in these plots) or exhibit large increases in the time period being tested (both shown in the lower plot, having a 4:1 relative frequency of occurrence)—all potentially triggering intervention—and by those that do not (upper plot), being of no concern with respect to Stipulation 12d. (Only 100 trendforms were generated for these plots, though more than  $10^{10}$  are used for our actual testing, which uses a grid of 31 possible values per time point.) Very many trendforms are considered because multiple historic accidental releases of varying durations, coupled with many different precipitation levels and seasons, could potentially result in one of almost an innumerable number of trendforms of various shapes (shapes of overlapping plumes and weather-related “phantom” plumes) at some sites, so this large of a set is expected to include at least one trendform that is close to the actual one.

Each trend in a group represents a possible state that has occurred—a *state of nature*. A *Predictive Likelihood* is a term used in Bayesian statistics when a specific characteristic that determines the shape or location of the probability curve—for example, a parameter, or here a trendform that impacts the Joint Likelihood curve—is considered to have a probability distribution itself: the Predictive Likelihood is the distribution of the joint data curve after probability-weight averaging that curve over all possible values of the characteristic. The probability distribution on the characteristic, known as the *Bayes prior distribution* (or *Bayes posterior distribution* when historic data are combined with a prior distribution) may be limited to discrete values or be defined over a continuum of them, according to the researcher's and/or decisionmaker's belief about the characteristic's value. This probability averaging produces a joint

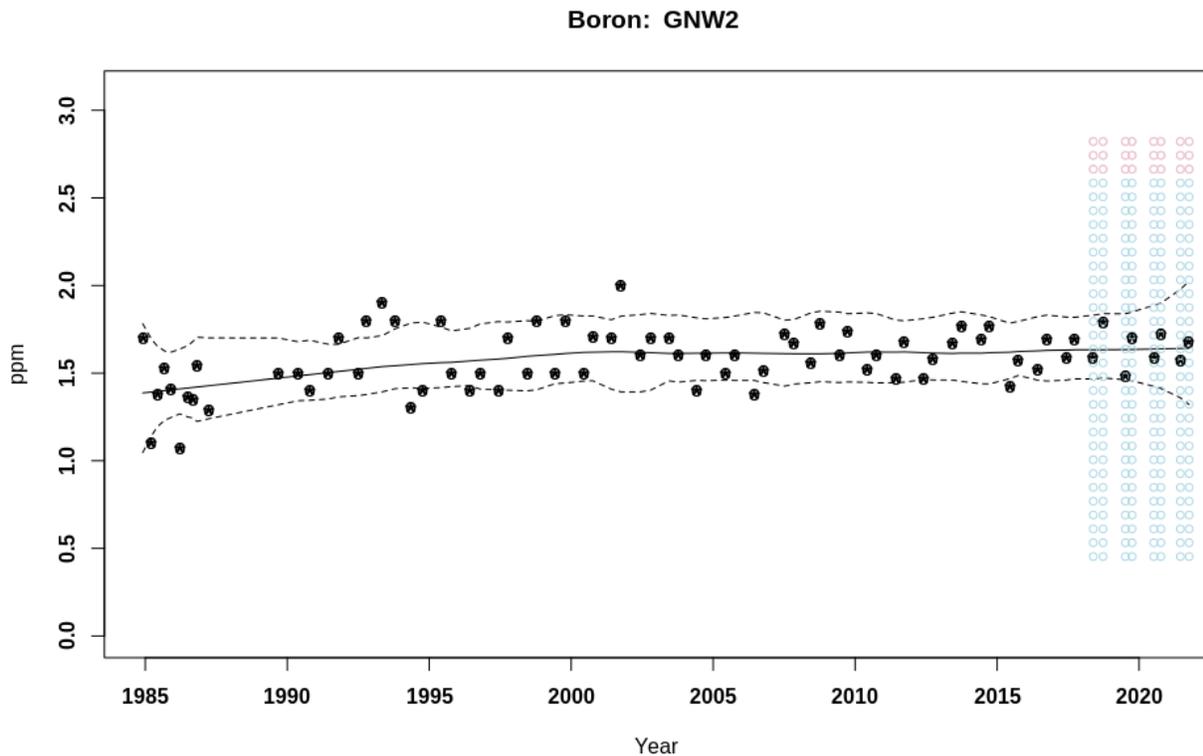
probability curve for the data that is no longer dependent upon only one specific value of that characteristic, but rather is shaped by all of them. The Predictive Likelihood may also be partially evaluated, over only those characteristics of interest; for example, over all trendforms in a hypothesis trendform group, producing a likelihood that is shaped only by the trendforms in that group, representing the accumulation of the combined possible occurrence of each and the plausibility of the data it induces. Therefore in this application, the Predictive Likelihood under the alternative hypothesis that intervention is needed combines, over “all possible” undesirable trends, the data Joint Likelihoods (described earlier) evaluated at each of those trends, whereas under the null hypothesis the Predictive Likelihood combines the Joint Likelihoods applied at each of the “innocuous” trends. The combination used was a simple average of the Joint Likelihoods and represents an equiprobability-weighted average of the occurrences of each trend in a group (innocuous or not), reflecting an “unbiased” viewpoint that each trendform in a group is equally likely. Under the additional unbiased assumption that each trendform group is equally likely to occur, the PLR is the ratio of the two (partial) Predictive Likelihoods just described, and in similarity to the Likelihood Ratio, it will tend to be large if the data suggest more strongly that one or more increasing/high-valued trendforms best fit them, but if instead some innocuous trendforms fit the data better, this ratio will tend to be smaller thus supporting the null hypothesis that there is no concern relative to the Stipulation 12(d). As in our conventional tests, the actual testing involves repeatedly generating a data series under the null hypothesis—here based on a randomly selected innocuous trendform—then evaluating the PLR at each generated series. If less than 5% of these simulated PLR values exceed the PLR computed from the actual data, we say that the observed PLR is uncommonly high if the null hypothesis was true, and we reject that hypothesis (of an innocuous underlying trend, which includes no trend), and we conclude that either there is an increasing recent trend or there are one or more large trend values in the recent years.



**Figure G.4.** Generic Example of 50 Possible Trendforms, Under Non-Increasing Hypothesis (top) and Increasing or High-Valued (exceeding 70) Trendform Hypothesis (bottom) (from Chamberlain, 2018)

The generated data were obtained in a way similar to our conventional test for a trend, as described in Section G.3.1, Steps 1–4, and illustrated in **Figure G.2(a)–(d)** there, but in this case data were reconstructed for only the most recent eight sampling time periods and here, in each iteration a series of random measurement errors (taken from the original error estimates) was added to an innocuous trendform, which was randomly selected as the underlying trend for that series, rather than the errors' being added to an underlying recently flat trendform as was done in the other test. The remaining part of the PLR test procedure is different from the conventional test; whereas in that test a trend was fitted to the simulated data and after all iterations, the p-value was the proportion of those trends that agreed, in direction, with the originally estimated trend's, here in each iteration a PLR was computed from the simulated 8-data series and the p-value was the proportion of all simulated PLR values that exceeded the PLR computed from the data.

The pre-trend reference value used to define an increasing trend (above that value) and, therefore, to identify the triggering group, was the most recent pre-trend value from a local (i.e., moving) polynomial fit similar to supersmooth—30 grid points were constructed using this reference value, usually symmetrically around it. The grid range's half-distance, added to and subtracted from, the reference value to respectively obtain the maximum and minimum grid value, was computed as the range of the data over the entire study period plus 2 error standard deviations, except the lower grid value was set to zero when subtracted half-distances would otherwise result in negative grid values. The extreme, high values that define an “abrupt increase” were taken as the uppermost three grid points. Figure G.5 illustrates the grid for boron at GNW 2.



**Figure G.5.** Example Gridpoints (blue, red, open circles) Used in Approximating All Possible Trendforms

The occurrence (*prior*) probabilities for the subgroups, increasing and high-valued (trends having one or more threshold-exceeding values in eight sampling times), were respectively 0.4 and 0.1, and the

innocuous group's occurrence had the remaining 0.5 prior probability. Returning to the PLR computation, as indicated earlier the PLR was obtained as the ratio of Predictive Likelihoods for the two groups and for each group, the simple average, overall trendforms, of the trendform-specific Joint Likelihoods was tabulated at the generated series. The Joint Likelihoods were as earlier depicted, the products of probability curves evaluated at each of the eight data values (representing four or more most recent years) and the trendform's values at the respective datum times. Specifically, the probability curve used in the Joint Likelihood computations actually characterizes the distribution of original measurement errors, being estimated from them, so the curve was evaluated based on the *difference* between each generated data value and the trendform's value at that time. It should be noted that in each iteration these Joint Likelihood computations were applied to all trendforms in a group—approximately  $2^8$  in the innocuous group and  $30^8$ – $28^8$  in the triggering one—and in all but one case (when evaluating the denominator Predictive Likelihood) they were not the same trendform as the (null-based) one randomly selected and used to construct the data for that iteration. When the difference between datum and trendform was small in magnitude (representing a close match) the curve value was high (peaking around zero [recall that the observed error components—aka residuals—from the original trend fit, are the differences between the data and that fit and can be either positive or negative]). And further, if the product of these curve evaluations was also high, then that trendform closely corresponded to the generated data and its contribution to the average—the Predictive Likelihood for that hypothesis—was also somewhat high. Conversely the trendforms whose values in each time point were distant from the generated data resulted in differences that were too large in magnitude to be in the typical range of the errors, and therefore their resulting curve values were very small or even zero, causing the respective Joint Likelihoods to contribute very little or zero to the Predictive Likelihood for that hypothesis group, and if all contributions were small, the Predictive Likelihood for the competing hypothesis was probably largest. Therefore, because the data are generated under the null assumption, we would expect the PLR's distribution to cover small PLR values (usually fitting the null trendforms best and hence causing the denominator Predictive Likelihood to tend to be larger than the numerator one); i.e., values smaller than the resulting PLR values if the data were generated under the alternative hypothesis. So, we then compare this distribution with the observed PLR and compute the latter's p-value as the percentage of simulated PLR values that exceed the observed one. (It should be emphasized that the observed PLR is evaluated based on the actual data, which have an unknown underlying trendform whose group is also unknown.) If the observed PLR is very large relative to the (null hypothesis based) distribution of the PLR—for example, the observed PLR has a p-value smaller than 5%—we would reject the null hypothesis in favor of the alternative one, concluding that either an increasing trend exists, or one or more underlying means in recent years was very high.

The probability curve, technically referred to as the *probability density function*, was obtained by kernel density estimation (Silverman, 1986) using as its data the original observed errors from a (supersmooth-like) nonparametric trend fit that did not impose any trendform assumptions. Unlike a histogram for example, the kernel density estimate is a continuous function, providing a useful quantitative density value at any error value. This density becomes zero for extremely large or small (i.e., large, negative) errors. For example, within the Joint Likelihood computation, if a generated datum was very distant from the evaluating trendform, their difference was extreme, and the density's resulting near-zero value indicated that the probability of that difference's being a measurement error was extremely low—so, virtually no errors this large or larger in magnitude could occur. This suggests quantitatively that the datum's underlying mean at that time point was different from the trendform value then (and hence dampened the Joint Likelihood for that trendform and its contribution, in the PLR computation, to the overall Predictive Likelihood for the trendform's group).

### **G.3.3.2 The Composite Test for Increasing or High Site Trends: Combining the Best Performance Features of the Trend and PLR Tests**

In 2018, a special study using the data for conductivity and boron at Cow Creek sites was conducted to evaluate the performance of our conventional test and the PLR one. The evaluation involved Monte Carlo simulation repeatedly generating data whose underlying trendform was either flat, or in a separate set of runs, increasing, applying the tests in each iteration, and evaluating both tests' results from all the runs for that trendform. (This simulation was an “outer” one and should not be confused with the “inner” simulations that accomplish the PLR test and that execute the conventional trend test: the tests were applied to each data series generated in the “outer” simulation described here.) Study simulations were applied to the data (through 2017) for each site in Cow Creek and for both conductivity and boron, to assess the “drainage error” rates discussed later, and because for a given constituent, the distributions of the estimated residuals used to generate data in the outer simulation and within each test were site-specific; the residuals each represented the total measurement errors—the sum of (site-specific) sampling error and laboratory error components. Both (testing) error rates were then assessed for each test—the proportion of iterations resulting in false detect and the proportion for which it failed to detect, when the underlying trends were respectively flat, and increasing. At all sites, the PLR test was found to have a much higher rate of detecting increasing underlying trends (i.e., corrected detection rate). As expected, regarding false detection, the conventional test had a false-detect rate lower than 5% (because testing is only done if the originally estimated trend is increasing). From an overall drainage perspective addressing whether one or more Cow Creek sites is increasing in either boron or conductivity, the false-detect results of both test results were approximately equal. They showed very high false-detect rates for the drainage overall (one or more site-level false detects when no increasing trends exist at any site there), while the drainage-level fail-to-detect rate—no detections when at least one site was increasing—were 6% for the current test and approximately zero for the PLR test.

When doing multiple hypothesis tests, the probability of at least one false detect among them is always higher than it is for an individual test, as high as the sum of the individual tests' false detect rates. So to reduce this overall testing error, a common approach is to use Bonferroni's correction (Milliken and Johnson 1984) by setting the individual error rates to the target overall false-detect rate (e.g., 5%), divided by the number of tests; that is, in an individual test, reject the null hypothesis when the observed p-value is lower than this adjusted rate and is rejecting less often. This correction ensures that the total error does not exceed the target rate, but it is often overly conservative and overstates the actual overall rate. For any given test, this correction greatly reduces the false-detection rate, but it increases the fail-to-detect rate, also known as the Type II error. Therefore, in the second study, to achieve the overall drainage false-detect level of 5%, the allowable individual Type I error rates for testing each of the eight Cow Creek sites, for boron and conductivity, were reduced to 0.003 (equaling 5% divided by 16 tests). However, because the conventional test's fail-to-detect rate was very high as a result of this correction, a composite test using both the PLR and conventional test was constructed to exploit the PLR's high correct-detect rate observed for even these more detection-reluctant testing levels. The composite test was as follows: conclude an increasing (or high) trend if either test concludes one. Also, because this test involves two tests, the target error rate for each test was further reduced by one-half to control the event of one or more false detects in that test at each site-constituent combination. The resulting performance of the composite test was remarkable, producing both a low overall false-detect rate in the drainage and a low fail-to-detect rate (when one or more increasing trends in either constituent is occurring at Cow Creek). Therefore, the composite test was applied to the cumulative (boron and conductivity) data in testing for increasing or high site trends in the drainages.

The nonessential recent trend tests - for decreases in conductivity or boron, increases or decreases in sulfate, and drainage-level trend testing in either direction – 98 tests in all – each use the conventional test

alone however, and the individual allowable (type I) error rate for each test was set at 0.05/98, again being based on Bonferroni's correction.

### **Why We Test for Increasing Trends of Any Magnitude Instead of Testing for Only Those That Are Meaningfully Large**

It is important, even essential that we test for any level of increasing trend. Here are the primary reasons:

- **Maintaining Statistical Test Sensitivity in Detecting Actual Increases - AKA the Power of the Test.** If negligibly increasing trends - those below some minimum "cutoff" value - were moved into the innocuous set, because there is considerable overlap in the range of possible data when a trend is below a cutoff and the range when a trend exceeds that cutoff, trends on either side would frequently not be statistically distinguishable. In other words, when there is a significantly increasing trend (i.e., a severe one) just above a cutoff, the probability of a no-increase decision would be high, or at least much higher than it currently is, and so the fail-to-detect rate would be much greater. Moreover, only trends far worse than those that minimally qualify as severe would be reliably detected. *Therefore we choose to maintain a cutoff very low - far away from highly increasing levels, so that we may detect them with high likelihood.*
- Also, when an increasing trend is detected, the supersmooth trend estimate and the posterior probability over all possible increasing trends, respectively provide accurate deterministic and probabilistic trend estimates; *testing for a wide importance range of increasing trends doesn't reduce our ability to correctly estimate any severe trends that may occur.*
- **Characterization of Extent if Contamination Occurs.** Because most Cow Creek sites are hydrogeologically distant from the EHP, any encroaching groundwater plume should have increasingly spread its extent while becoming more dilute (largely due to typical contaminant spreading/dispersing mechanics acting in the flow direction and laterally, arising from varying groundwater velocities within each pore path and soil matrix configuration): therefore constituent levels at these sites should initially manifest a gradual increase rather than a "jump," and so testing for small increases downgradient of a highly contaminated site enables establishing, or at least limiting, the plume's extent – e.g., the plume is just reaching the downgradient site, or, it hasn't arrived there yet.

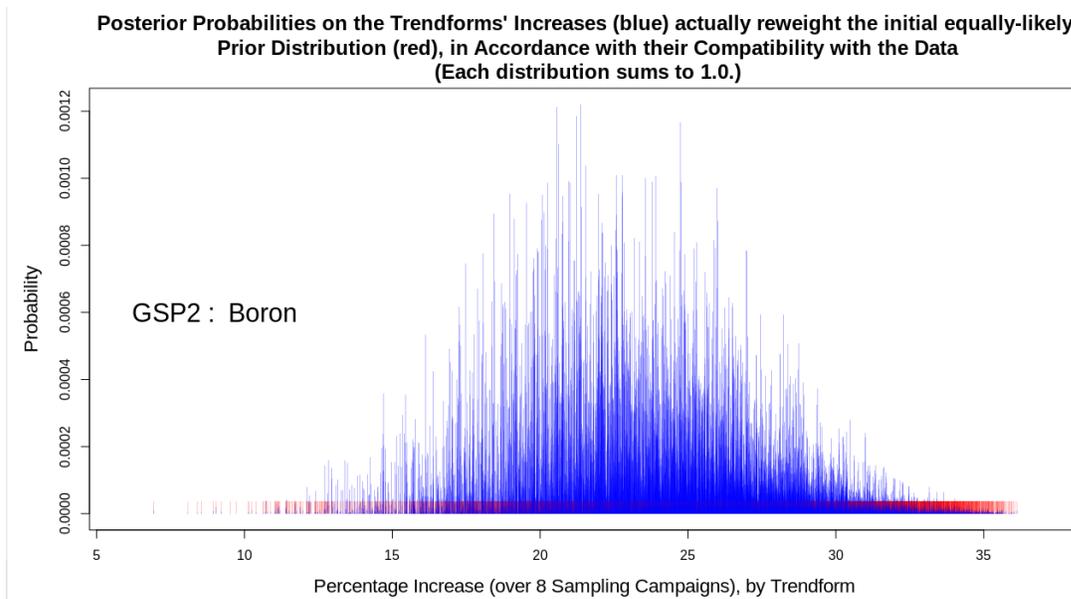
**Characterization of Nature of Contamination.** Further, if no negligible increase was ever detected previously at a highly contaminated site, this would tend to indicate that the source isn't contaminated groundwater at all, but instead is the result of contaminated streamflow/stormwater runoff that is sufficiently close to produce a contaminant breakthrough ("jump") at the site, facilitating our characterization of the nature of the contamination. Finally, if many sites exhibit concurrent, small increases, we also may suspect that the source is contaminated runoff because in the absence of rapid groundwater movement due to karstic or fracture flows, it is impossible for a single plume to affect all sites simultaneously with the onset of a small increase (and it's highly unlikely to observe simultaneous increases if multiple plumes are present in the drainage). Also detection of initially small increasing trends that are widespread in this way enables more timely remedial response.

#### **G.3.3.3 The Posterior Distribution – For Interpreting the Observed PLR, and for Evaluating the Severity of Detected Trends**

In Bayesian analysis, the true "state of nature," here the true trendform, is assumed to have randomly occurred according to a probability distribution on it - a prior distribution – e.g., which reflects our belief about the entire set of trendforms that could possibly occur; to maintain "unbiasedness" our prior distribution assumes that each trendform is equally likely to have occurred. After data are observed, the prior distribution is reweighted – multiplied - by the (trend-specific) data distribution evaluated at each

trendform; the resulting updated distribution on the trendforms - the posterior distribution - is obtained from this reweighting. As described next, *among the initial trendforms considered, those implausible ones, given the data that we've observed, are essentially filtered-out in the posterior distribution and its contribution to the observed PLR.*

The data distribution is similar to a “bell curve” and evaluating it at various trendforms, given the observed data, is similar to evaluating a single data value’s bell curve score when the curve is shifted by varying amounts and directions – when the curve’s peak is close to the datum the score will be high, but the score will decrease as the peak is moved away, in either direction, from the data value – the datum isn’t compatible with distributions that far away. Likewise trendforms that are “distant” from the data will have low scores that will dampen the original prior probabilities at those trendforms, whereas the prior probabilities for trendforms that are close to the data will be augmented by high scores. Because this reweighting refines the prior by augmenting/dampening it at each trendform according as the data are/aren’t compatible with it, the posterior values will be approximately zero at impossible trends given the observed data: in general the posterior gives us an improved, more correct and focused set of probability estimates for the trendforms. The posterior and prior distributions for (approximately) all possible increasing trendforms for boron at GSP 2 are illustrated in Figure G.6; note there the posterior’s redistribution, as compared to the equally likely prior distribution, of probability toward the trendforms with moderately high 8-campaign percentage increases, away from both more extreme ends. The probabilities in this posterior distribution and in the prior one, if we conclude that the actual trendform is increasing, both sum to 1.0. Before making that conclusion, the overall posterior though imputes around 97% probability to this “conditional” posterior and the remaining 3% is imputed to the set of all innocuous trendforms (the posterior’s probability for the third set – of all abruptly increasing – jumping – trendforms was nearly zero for boron at GSP 2).



**Figure G.6.** Equally Likely Prior Distribution, and Posterior (data updated) Distribution on Increasing Trendforms (ordered by percentage increase over 8 sampling campaigns) for Boron at Site GSP 2

Returning to the PLR already described, the observed PLR value (i.e., computed from the actual sample data, not a simulated value within the testing) has a simple interpretation in terms of the posterior. The

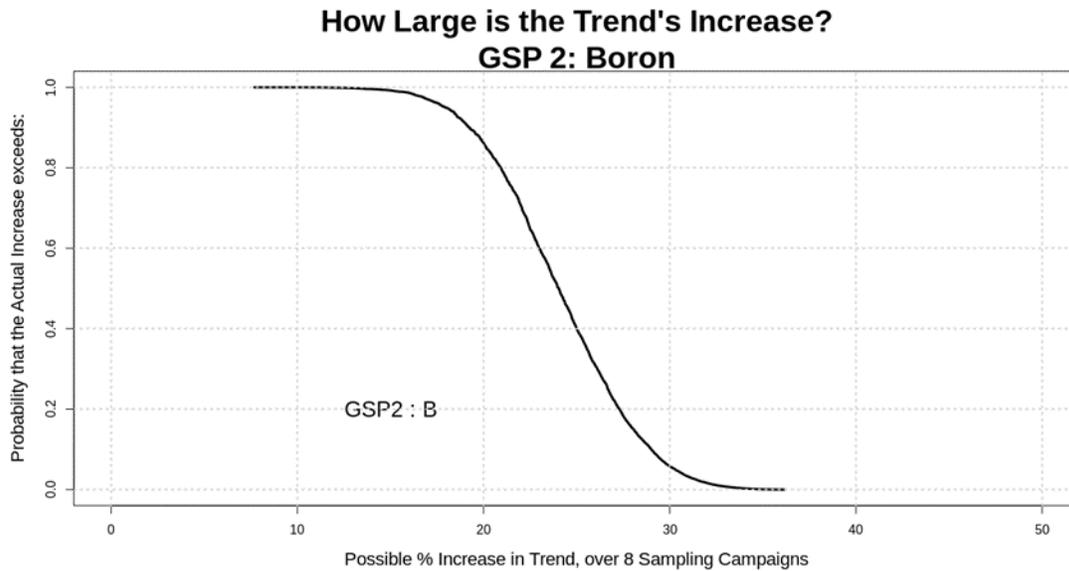
observed PLR is equivalent to the probability, under the posterior distribution, of the true trend's being increasing/high-valued relative to the probability that it's among the innocuous set: the observed PLR accumulates the posterior probabilities over both the increasing/high-valued and innocuous sets, and is computed as their ratio. For boron at GSP 2, noting from the preceding paragraph the 97%-3% probability split, the observed PLR was approximately 0.97/0.03, or 32. As discussed earlier, although the PLR is an odds ratio (one could test based on this value alone) we only compute this value as the statistic for testing: because a large PLR value could possibly occur when the innocuous assumption applies, we don't conclude until we've done the actual (classical) testing described in the previous subsection - the PLR may not be inordinately large after all. So concluding based on merely comparing the observed PLR with an arbitrary, fixed constant would tend to either lead to excessively many false detects or too often failing to detect even when there actually is a contaminant plume at the site. For boron at GSP 2, after testing—repeatedly both generating data under the no-increase assumption and computing their corresponding PLR values—extremely few of the generated PLR values exceeded the observed PLR value of 32; it was too rare to have occurred under that assumption. Therefore, as indicated in Figure G.10, boron at GSP 2 was concluded to be increasing.

So after concluding an increase, we assume that the portion of the posterior relating to increasing trendforms accounts for all the possibilities (its probabilities are rescaled to sum to 1), but instead of then plotting the accumulated probabilities to assess the totaled “less than” probability at every value of percentage increase, we subtract that progressive accumulation from 1.0 and plot the more useful *exceedance probability curve* shown in Figure G.7. ***From this plot, we conclude that the maximum percentage increase in boron at GSP 2 is approximately 36% (i.e., it's impossible – having zero probability – that the true increase is greater than 36%), and it's very likely to be much smaller than 36%.***

Although the PLR test never provides an individual estimated trend, the posterior distribution completely characterizes the probabilities on all the candidate trendforms and so it may be used to obtain a “best estimate” of the true, underlying trendform. One optimal estimate is a probabilistically weighted average of all the trendforms – this is the probability-mean of them. But instead we may focus on only the trendforms' percentage increases just examined, and obtain their mean by the same weight averaging of them, or equally simply, by accumulating the exceedance probability curve as it ranges from 0 to the maximum possible percentage (i.e., the mean of any positive-valued random variable is alternatively obtained by integrating the exceedance probability curve over its entire range). ***The mean percentage increase in boron at GSP 2 is 24%.***

### **G.3.4 Time-Trend Testing Results**

The results for the drainage- and site-level trend analyses are given in Figure G.8 through Figure G.10, where a plus (+) or minus (-) after the title or site, respectively, indicate that a strictly increasing or decreasing recent time trend was concluded. Figure G.9 combines the SUPSMU results with data plots for the sites in which the outlying data points that were removed from the analysis are identified with an “o,” as discussed in Section G.3. Figure G.10 overlays the smoothed site-level results within each drainage.



**Figure G.7.** Probability Curve on Percentage Increase in Boron Trend at Cow Creek Site GSP2, over Recent 4 Years, as Derived from the Posterior Distribution.

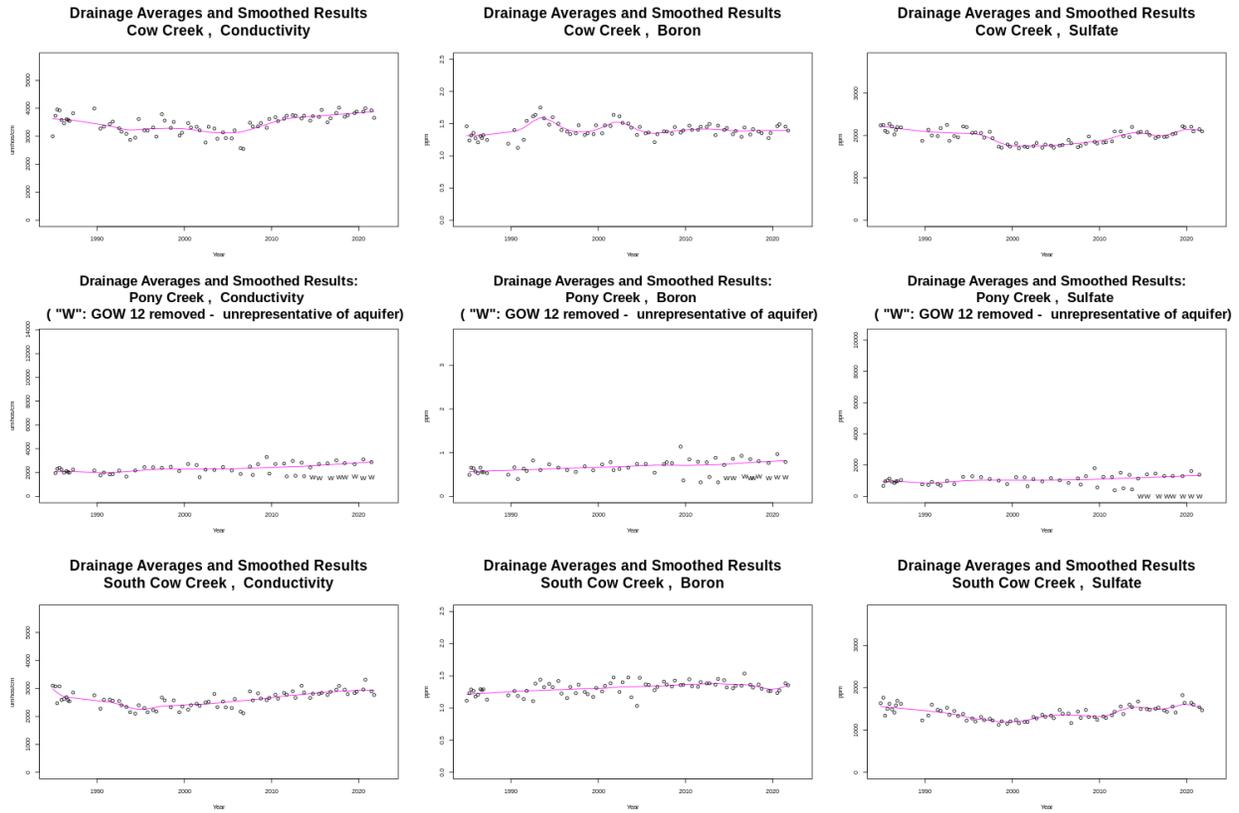
**Site Trend Test Results.** The site trend testing results are exhibited in the individual site plots of Figure G.9 and in the overlaid site plots of Figure G.10. The composite test was used to examine for increasing trends in conductivity and boron, whereas the conventional test was otherwise applied – i.e., for detection of decreases in all constituents and increases in sulfate.

*Increasing trends in conductivity and boron were found at private monitoring well PW 735, which taps a deeper confined aquifer than that supplying the alluvial wells. Boron at Cow Creek site GSP 2 was found to be increasing - gradually not abruptly, estimated to be around 24% but certainly not by not more than 36%, over 8 sampling events. At Pony Creek conductivity was found to be increasing at sites GSP 4 and GOW 4, and for boron increasing and decreasing trends were respectively concluded at GSP 8 and GOW 4. No sulfate trends, either increasing or decreasing, were found at any site.*

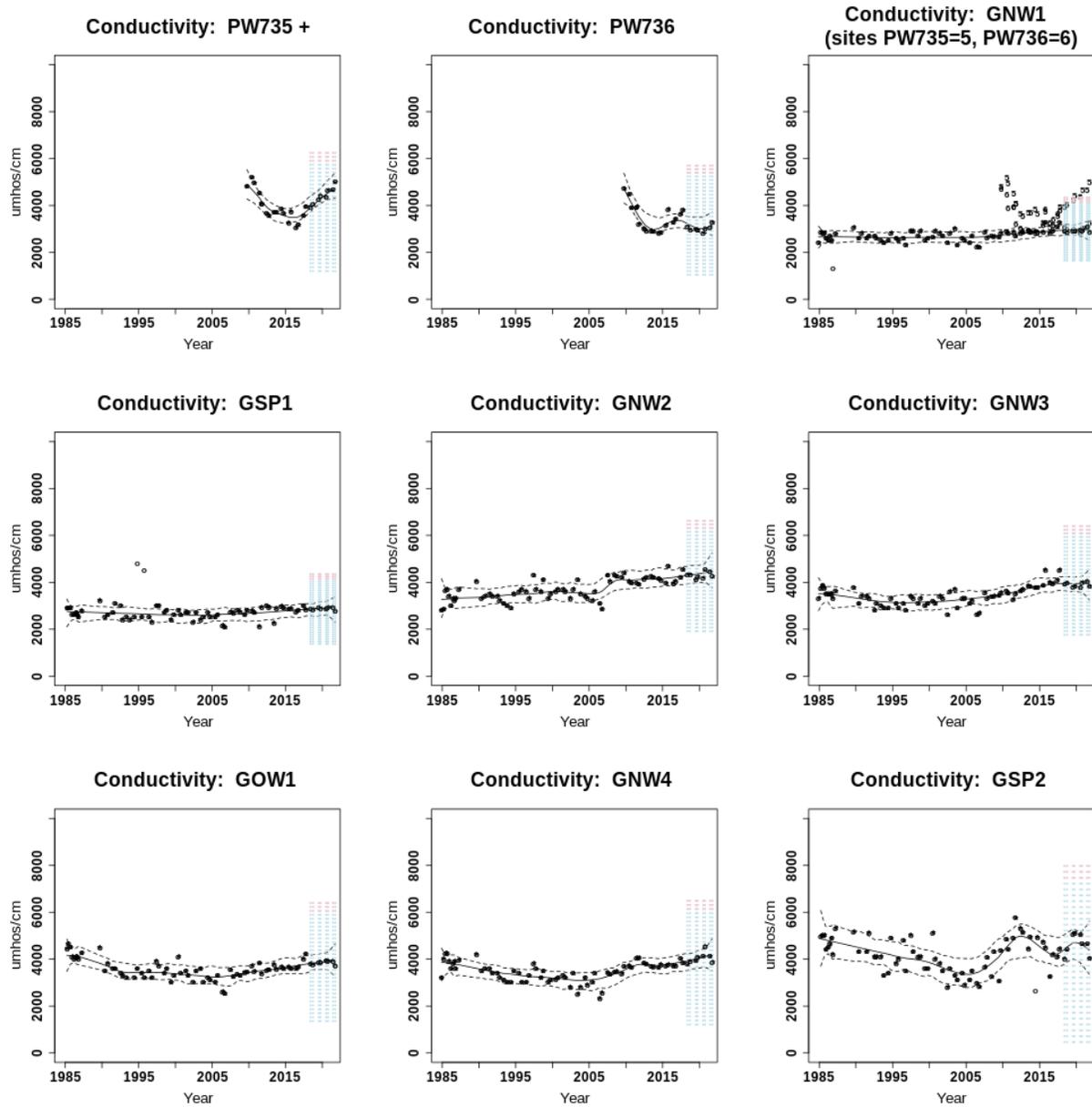
**Drainage-Trend Test Results.**

The results of the drainage-trend tests, for recent increases, and recent decreases are indicated in Figure G.8 and summarized as follows:

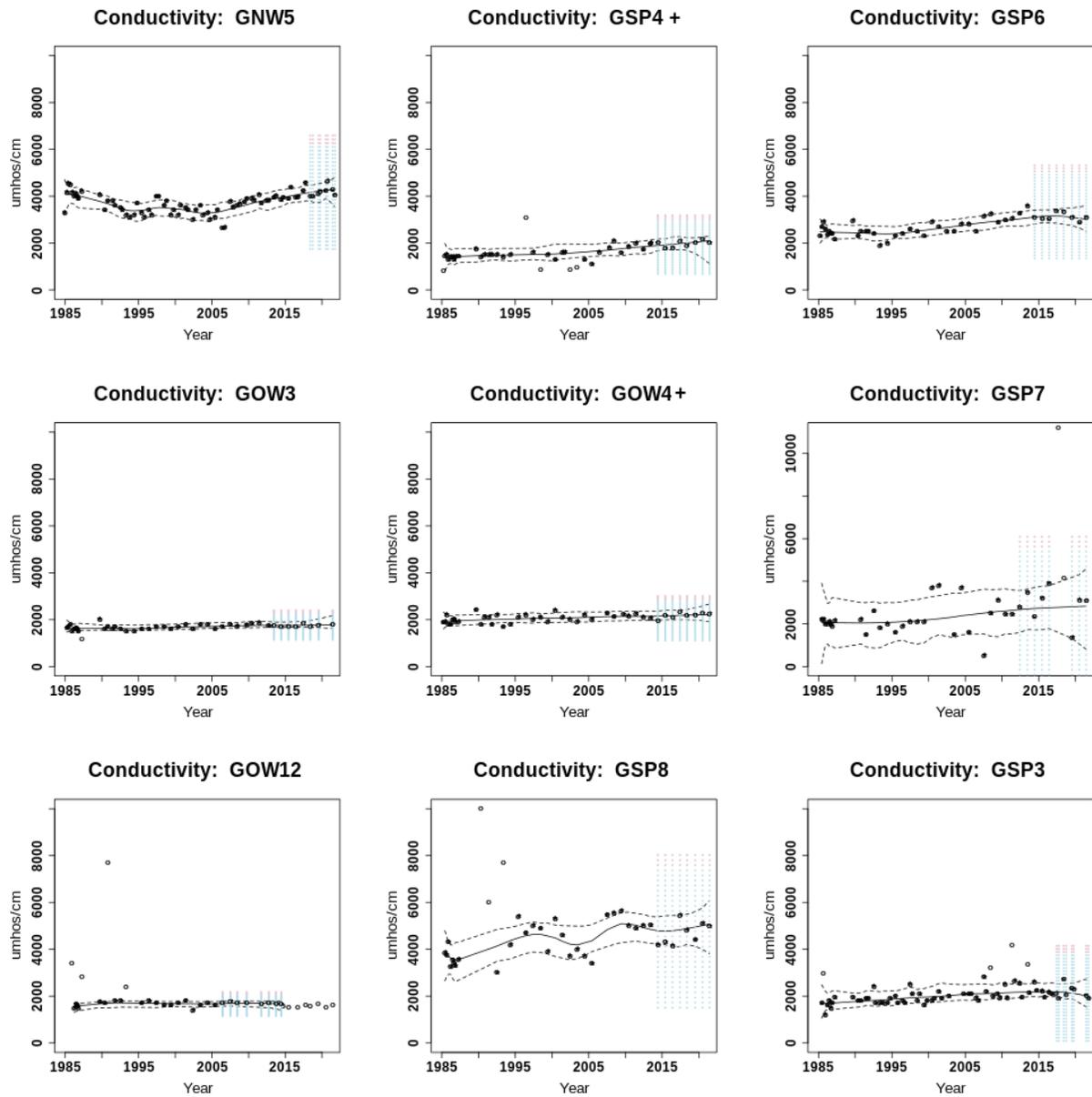
*No recent drainage-level changes were detected. The conventional test was used for all drainage time-trend testing.*



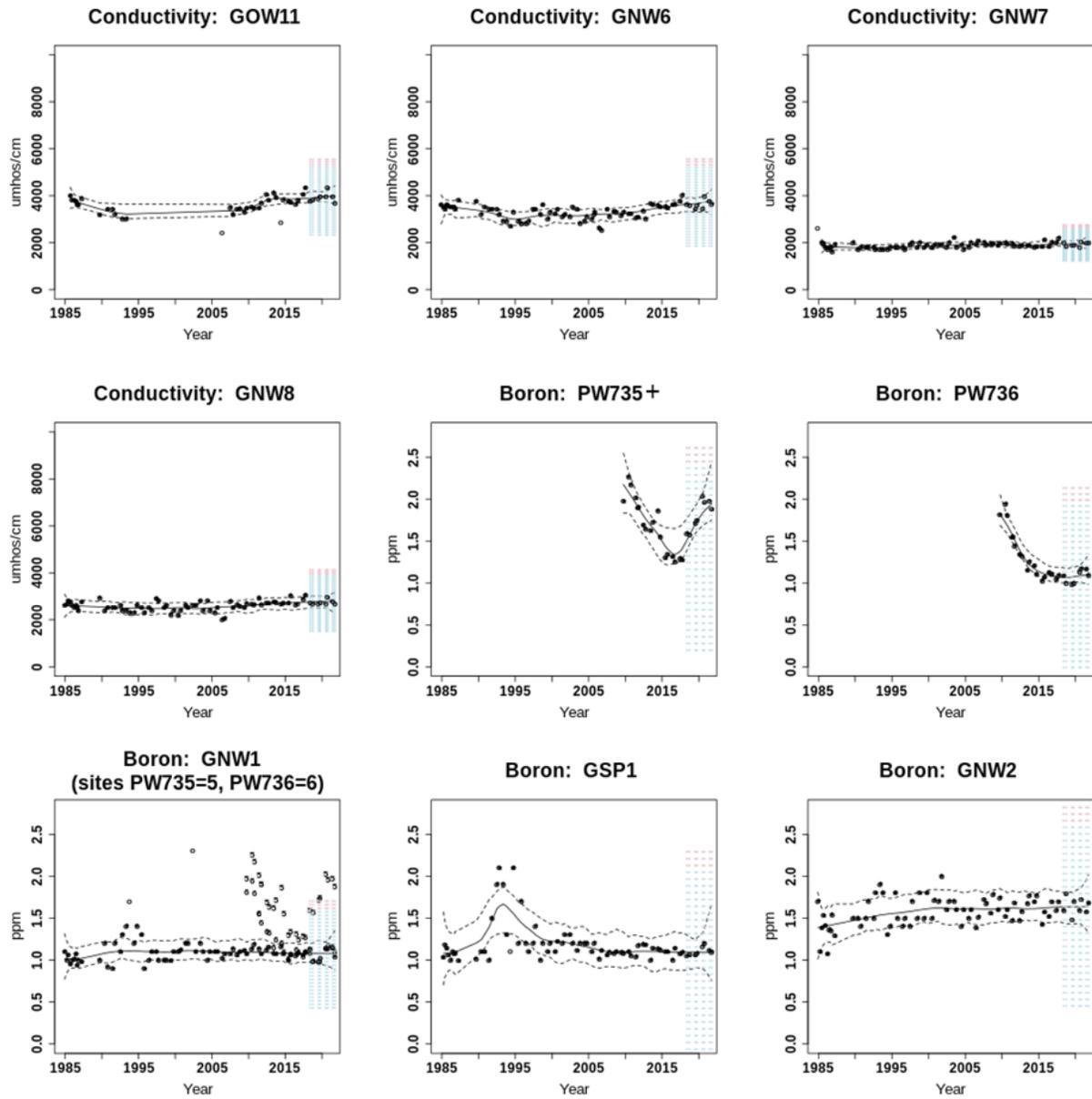
**Figure G.8.** Drainage-Level Trend Analyses. In past reports, a “+” in the title denotes an increasing time trend detected over the most recent 4 or more years at the 5% level of decision error (among all non-Stipulation 12(d) trend tests, site level and drainage level), and conversely a “-” denotes a decrease over that same timeframe. However, neither of these trends was detected in this year’s study.



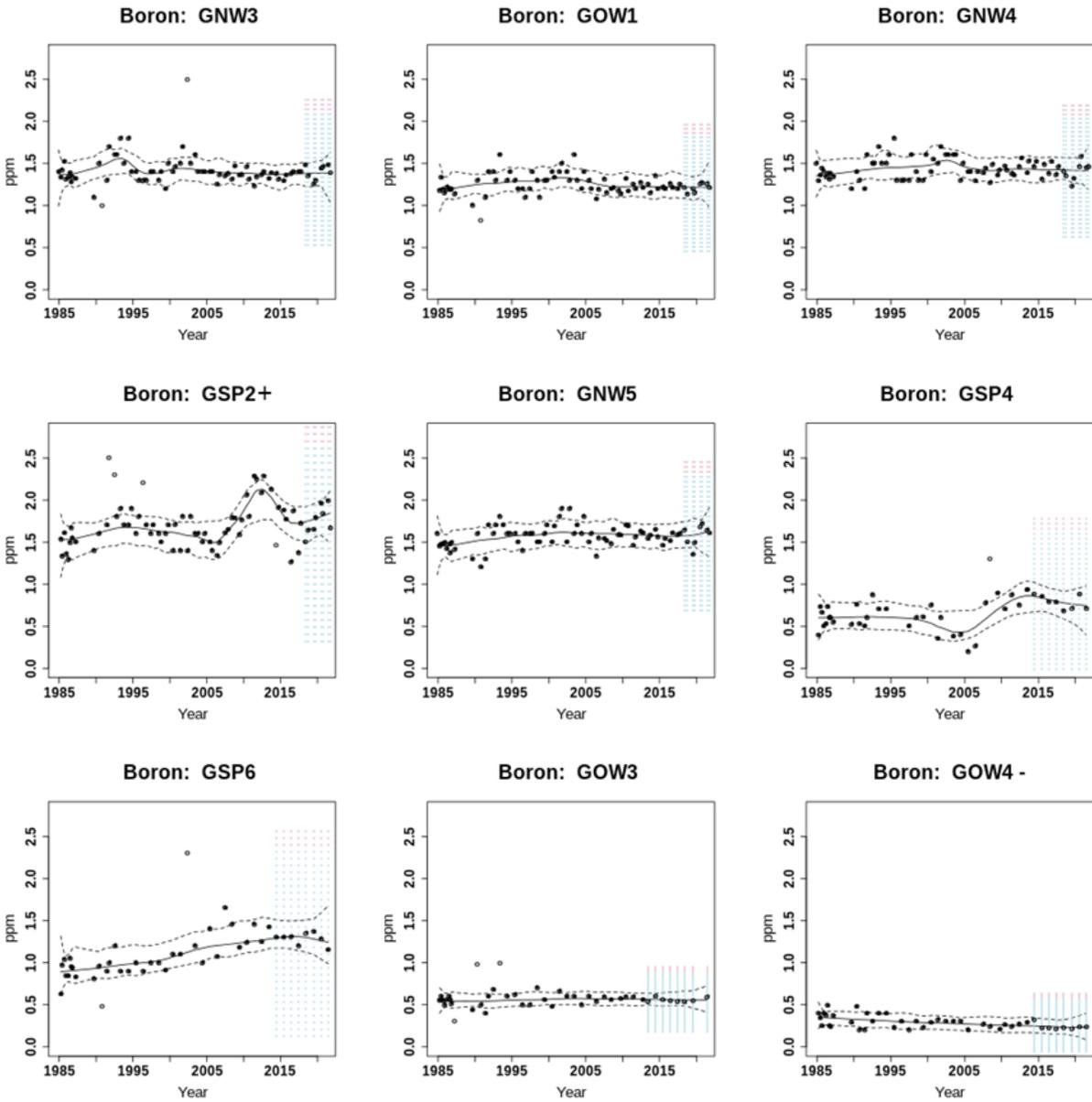
**Figure G.9.** Site-Level Trend Analysis. A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.



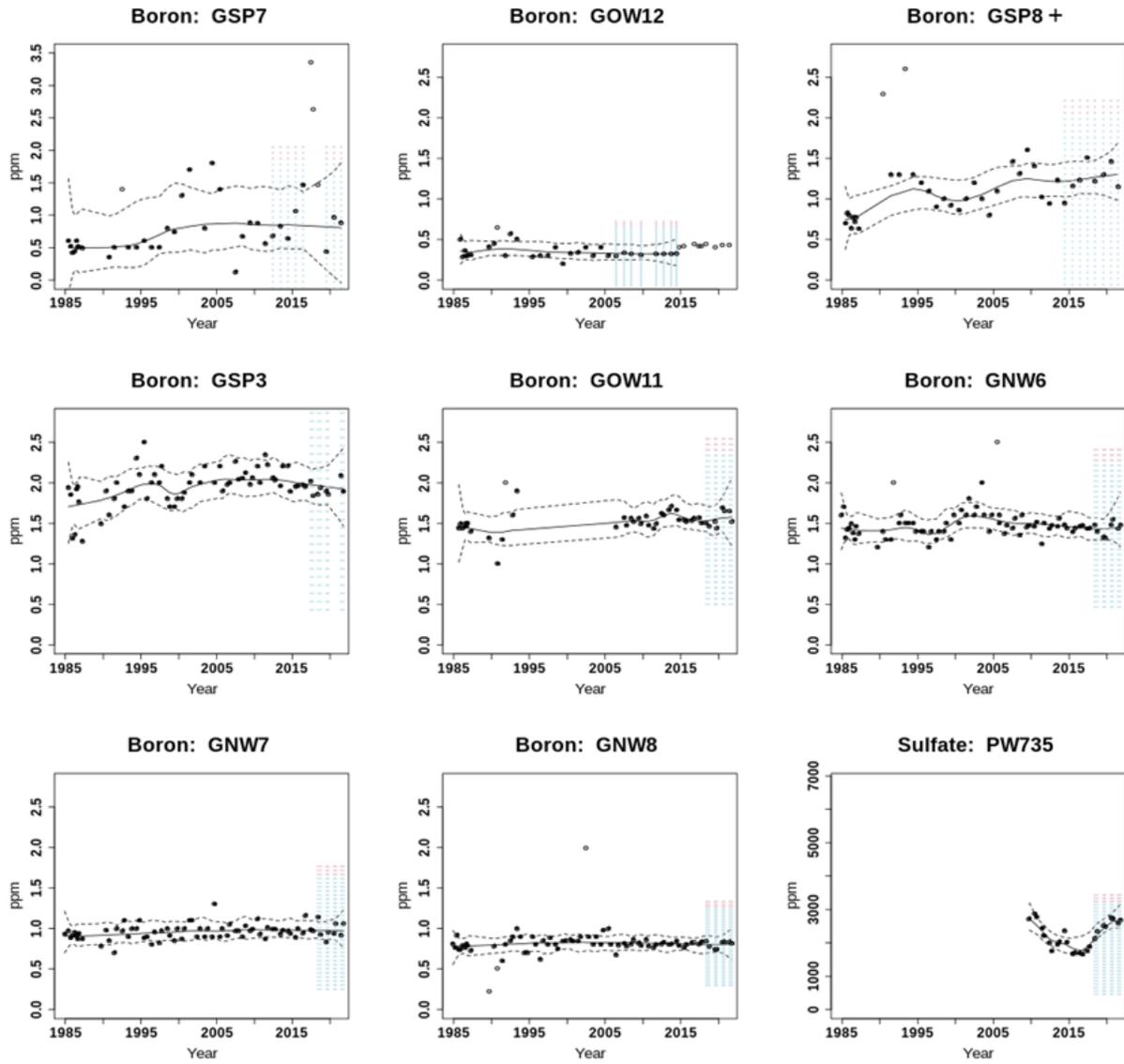
**Figure G.9.** (Continued, 2 of 8) A “+” denotes an increasing trend being found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.



**Figure G.9.** (Continued, 3 of 8). A “+” denotes an increasing trend being found by the composite test at less than the 5% level of *overall decision error for all composite testing in the drainage.*



**Figure G.9.** (Continued, 4 of 8). A “+” denotes an increasing trend found by the composite test at less than the 5% overall decision error for all composite testing in the drainage. A “-” denotes a decreasing trend being found by the conventional test at less than a 5% overall decision error rate among all non-composite site tests and drainage tests (98 tests in all).



**Figure G.9.** (Continued, 5 of 8) A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.

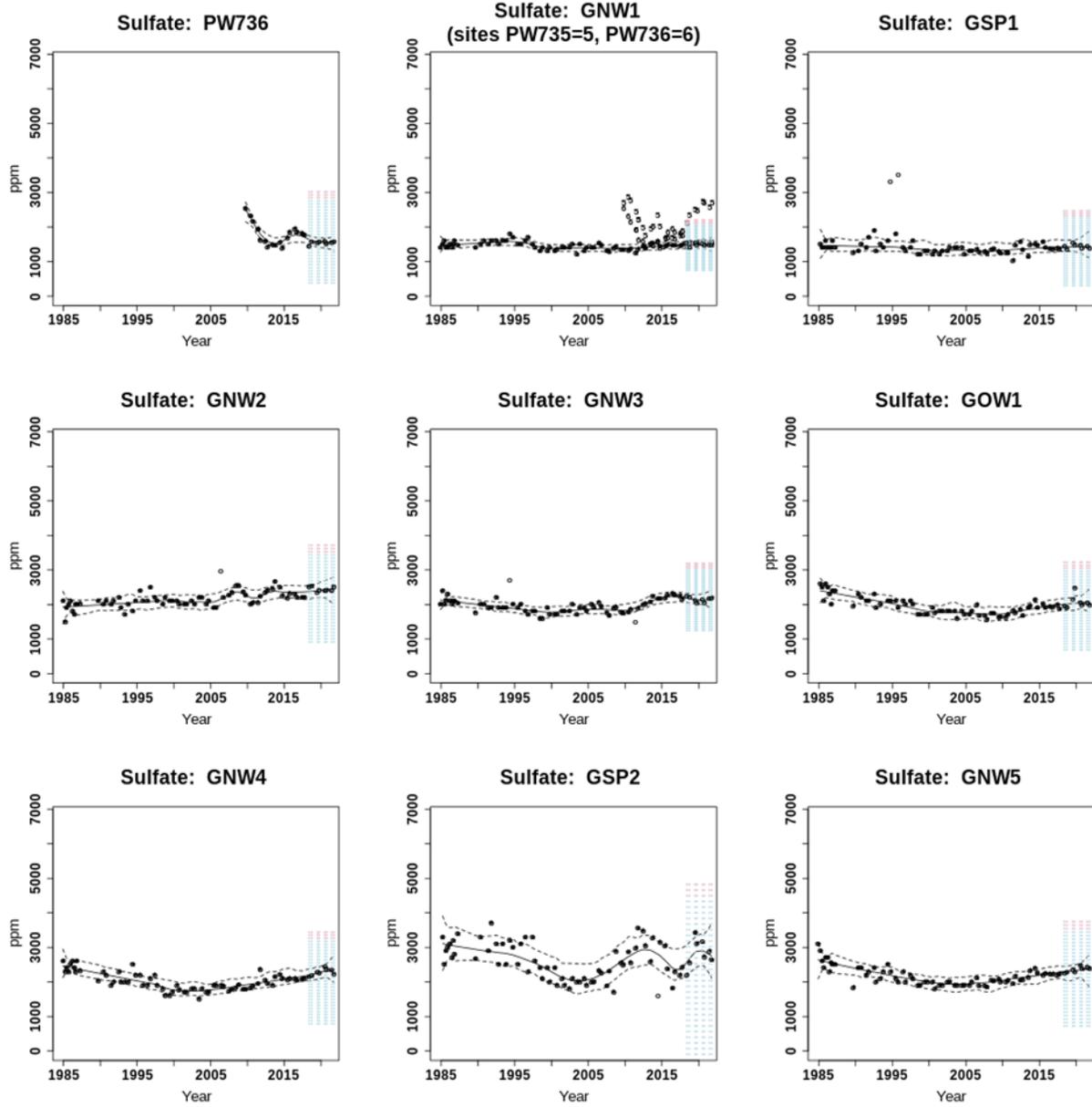


Figure G.9. (Continued, 6 of 8)

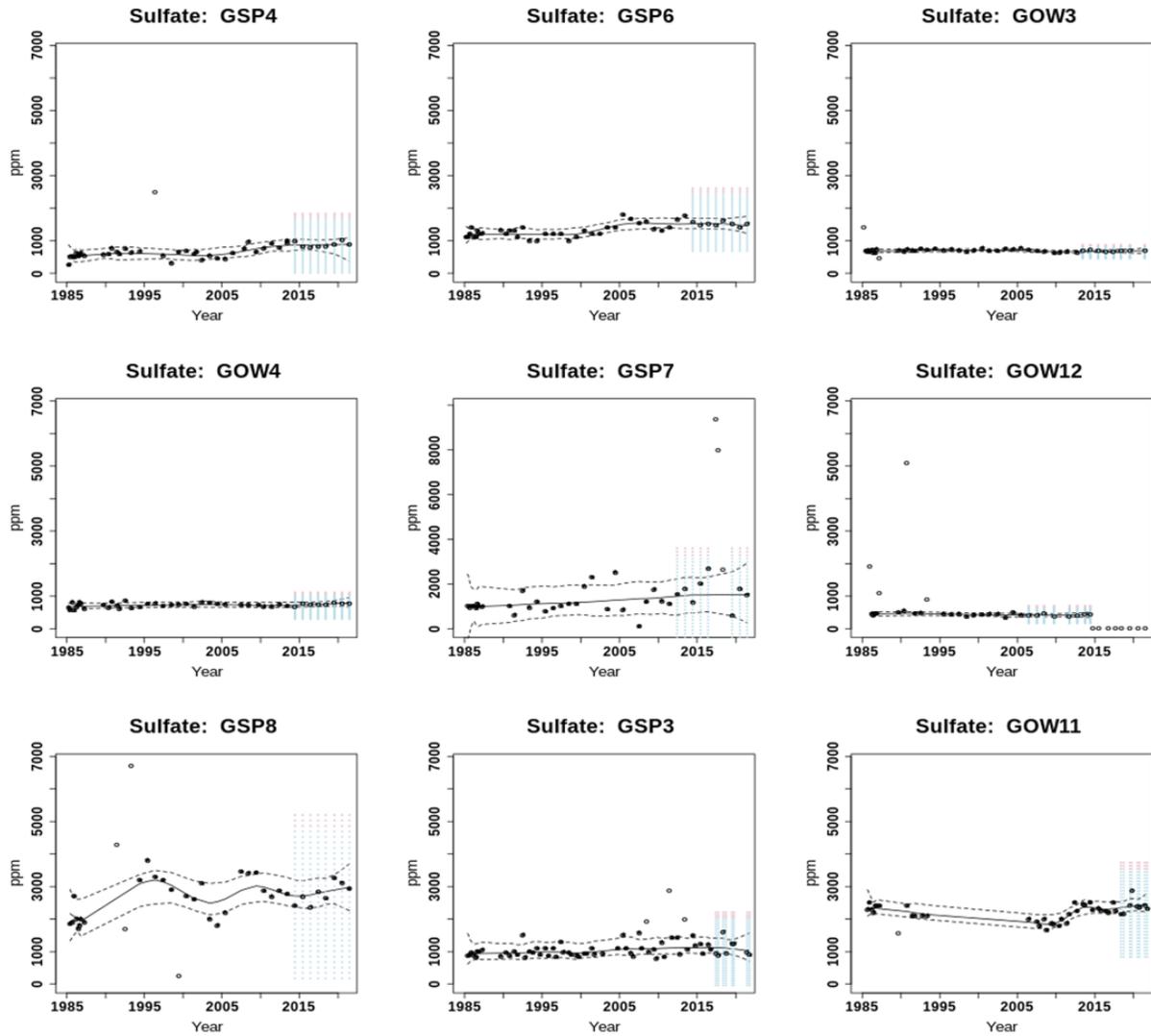


Figure G.9. (Continued, 7 of 8)

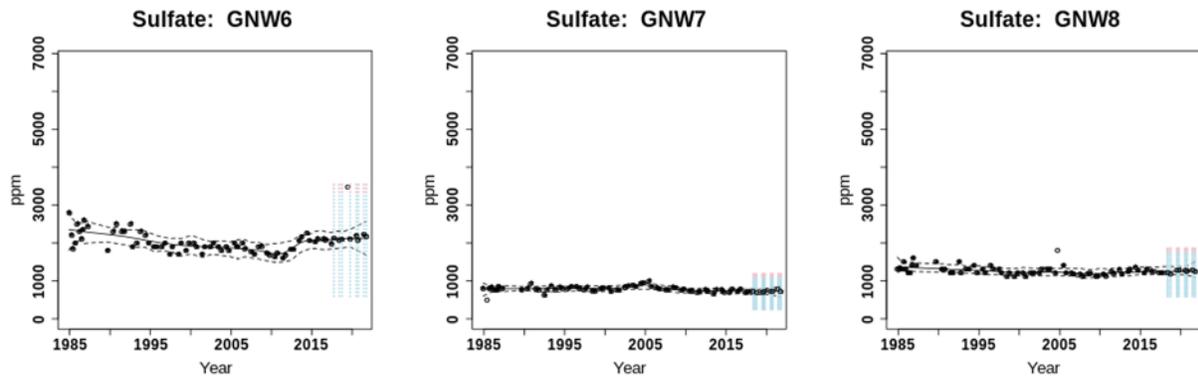
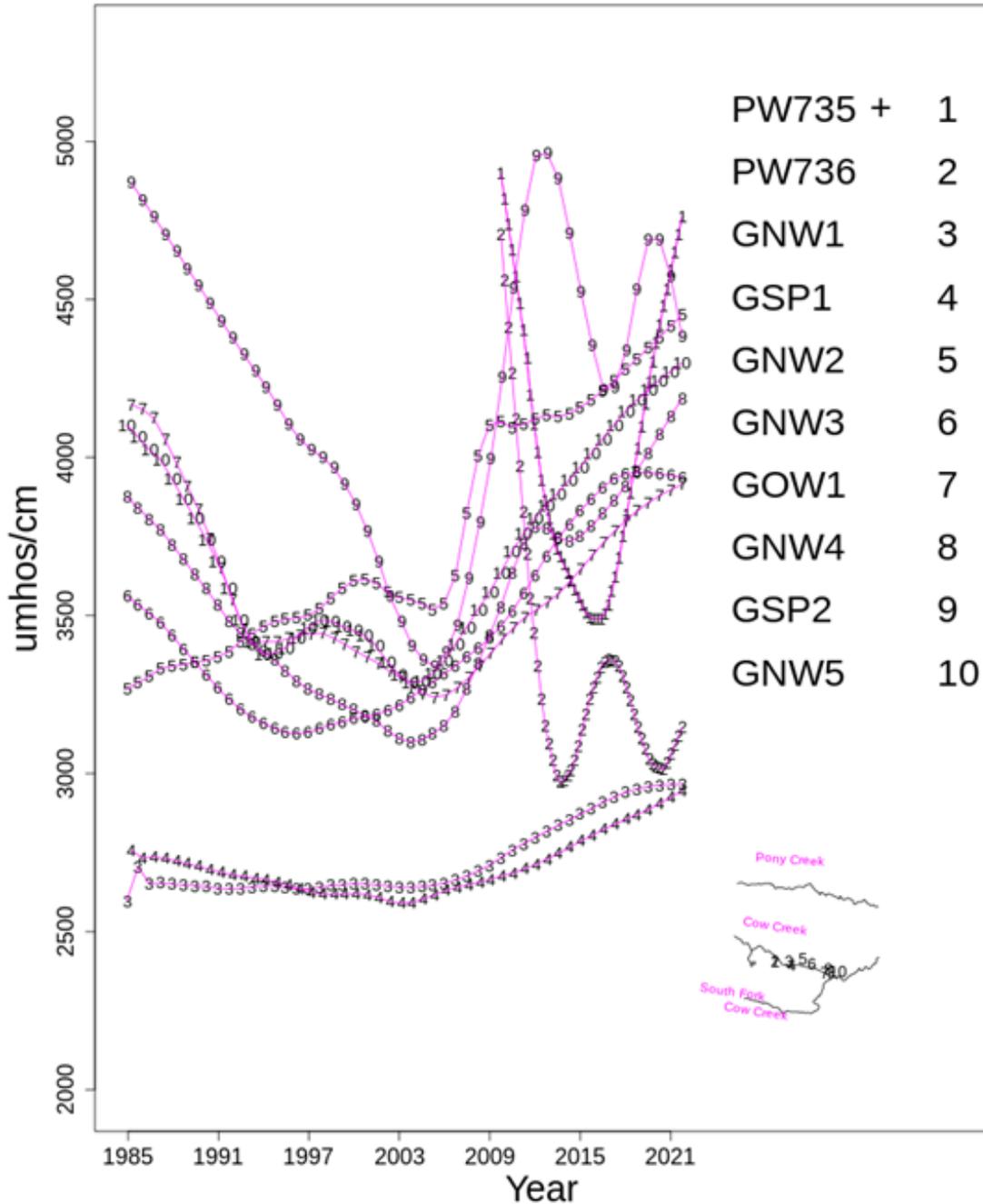


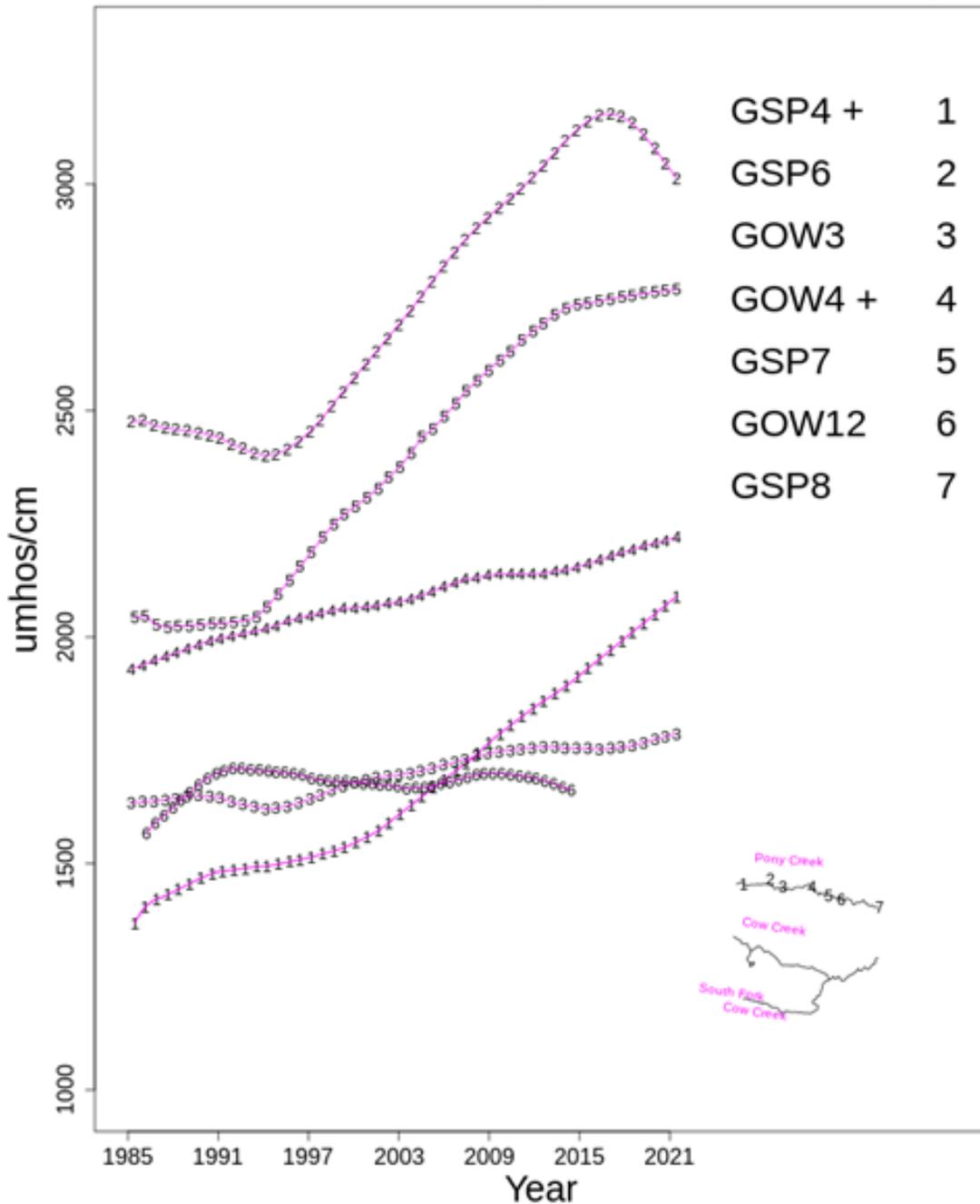
Figure G.9. (Continued, 8 of 8)

### Smoothed Results for Cow Creek Sites: Conductivity



**Figure G.10.** Smoothed Site-Constituent Levels Over Time (1 of 9). A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.

### Smoothed Results for Pony Creek Sites: Conductivity



**Figure G.10.** (Continued, 2 of 9, Trend testing at GOW12 is suspended pending further data.) A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.

### Smoothed Results for South Cow Creek Sites: Conductivity

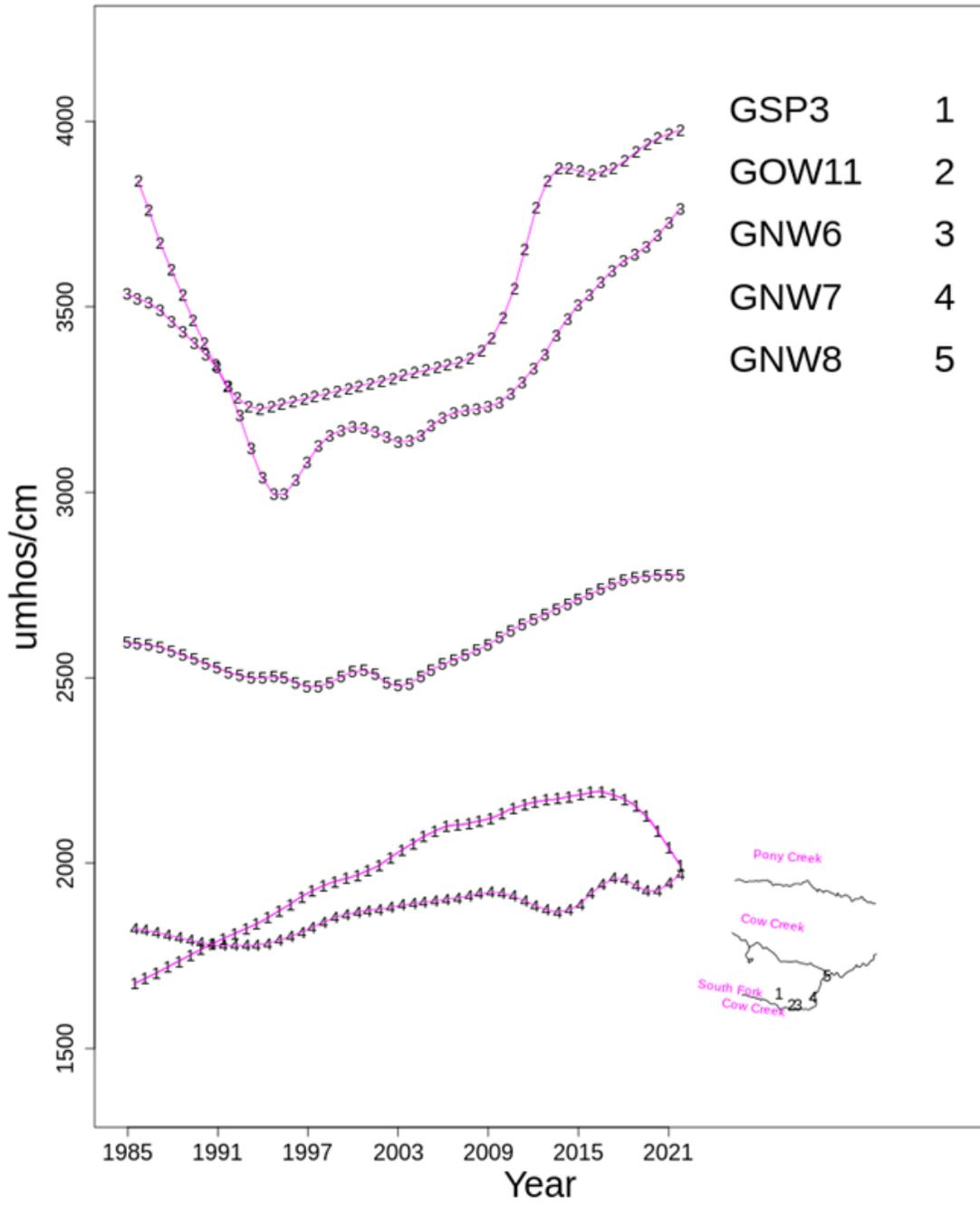
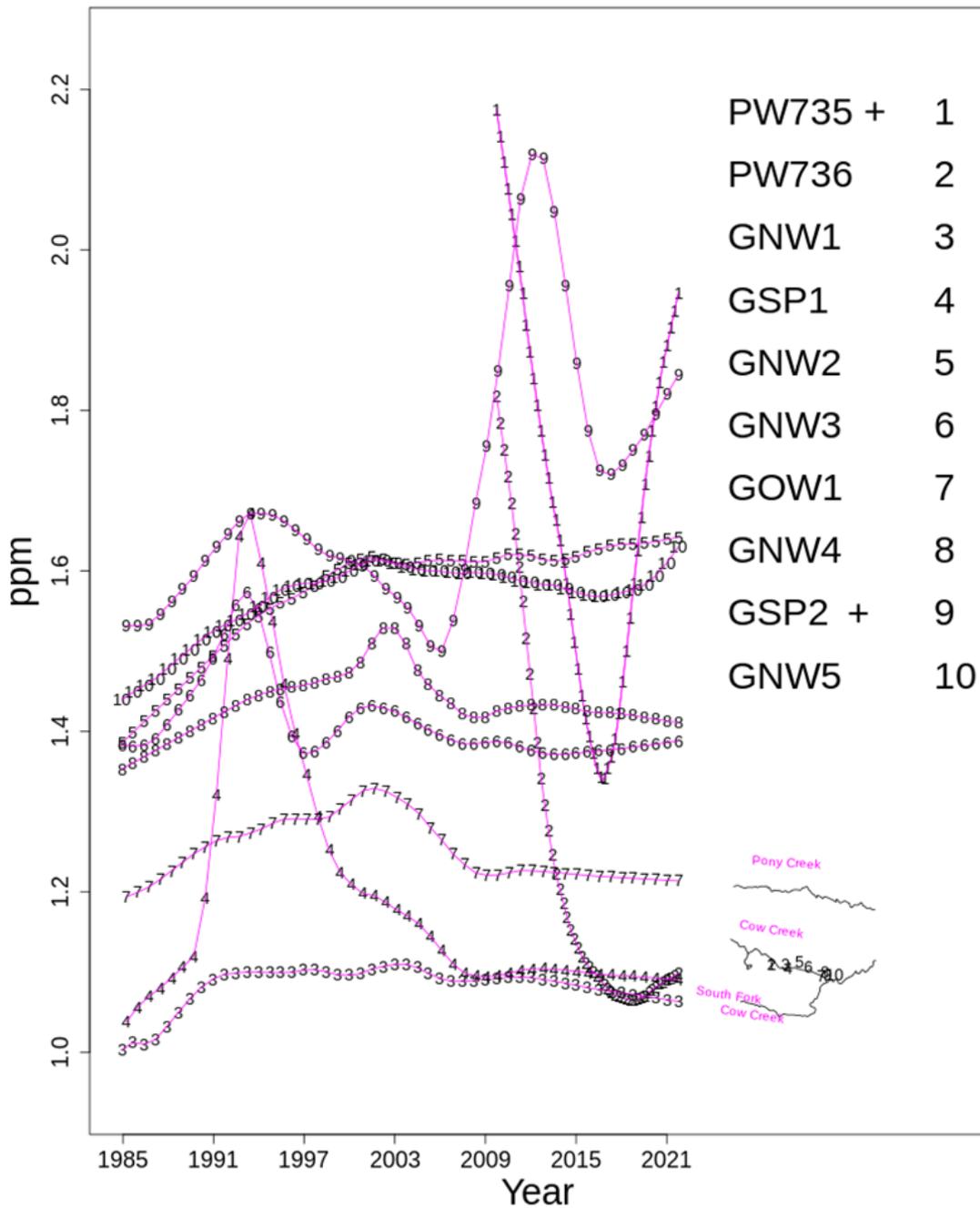


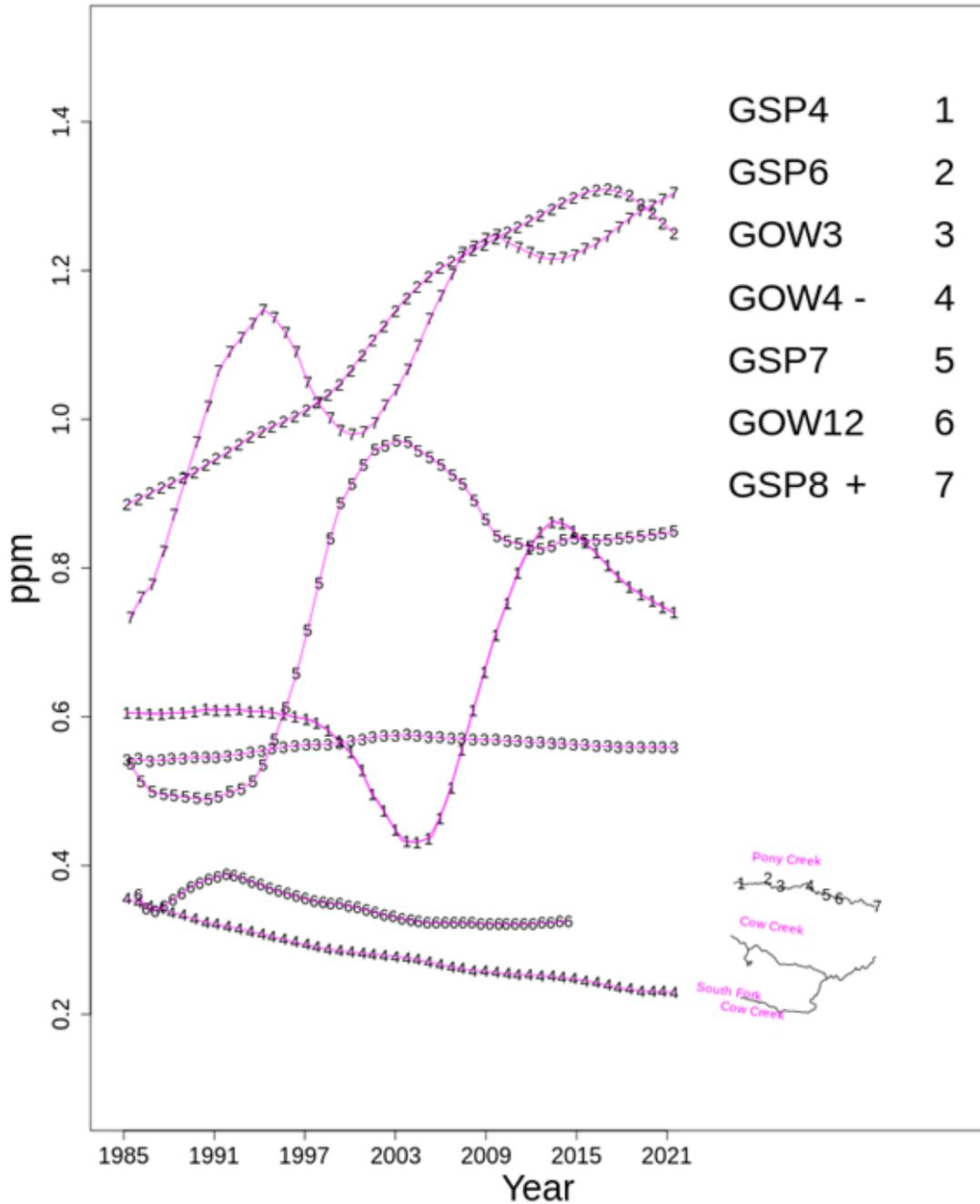
Figure G.10. (Continued, 3 of 9)

### Smoothed Results for Cow Creek Sites: Boron



**Figure G.10.** (Continued, 4 of 9) A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage.

### Smoothed Results for Pony Creek Sites: Boron



**Figure G.10.** (Continued, 5 of 9) A “+” denotes an increasing trend was found by the composite test at less than the 5% level of overall decision error for all composite testing in the drainage. A “-” denotes a decreasing trend was found by the conventional test at less than the 5% level of overall decision error for all non-Stipulation 12(d) tests. Trend testing at GOW12 is suspended pending further data.

Smoothed Results for South Cow Creek Sites: Boron

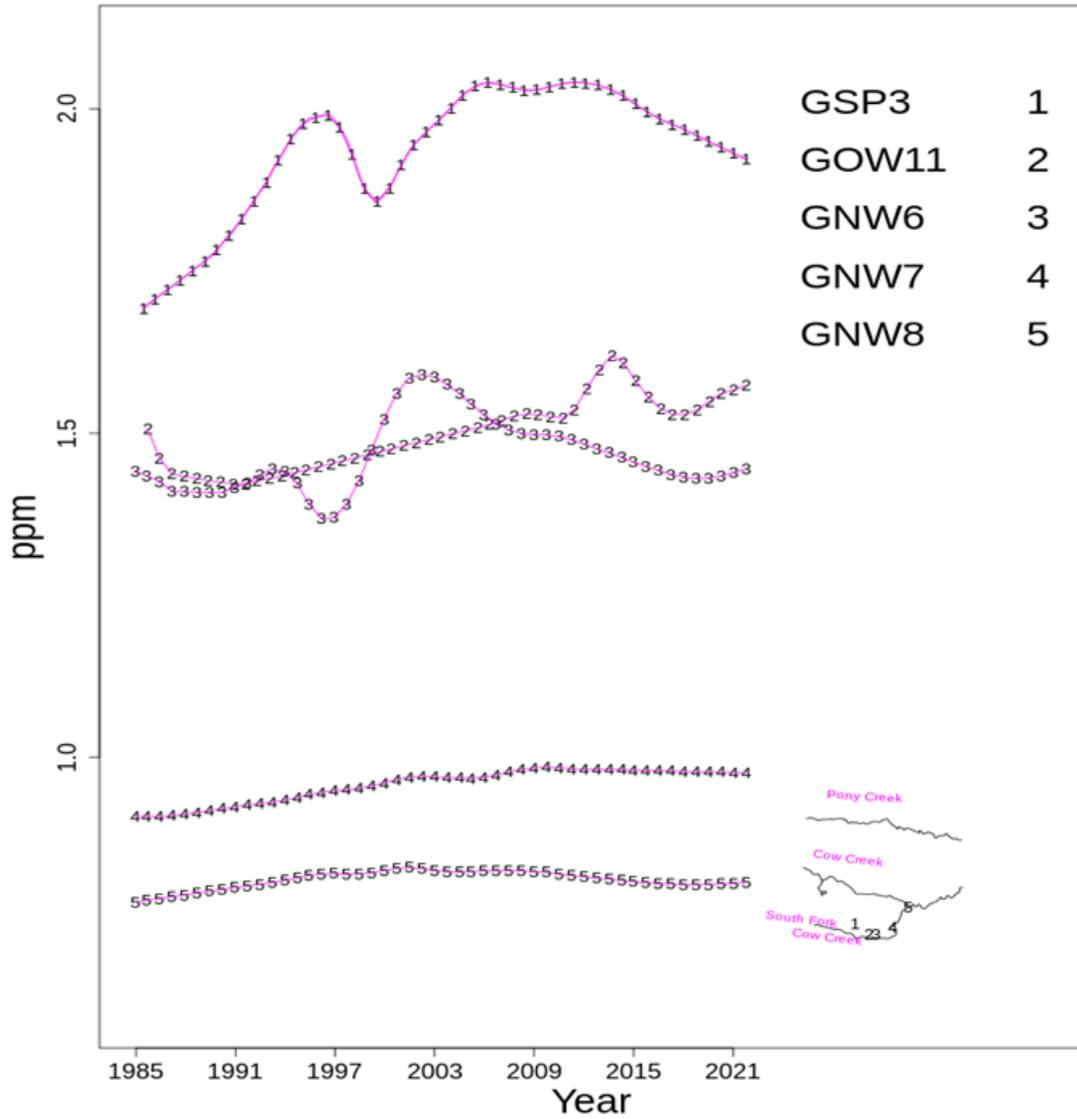


Figure G.10. (Continued, 6 of 9)

### Smoothed Results for Cow Creek Sites: Sulfate

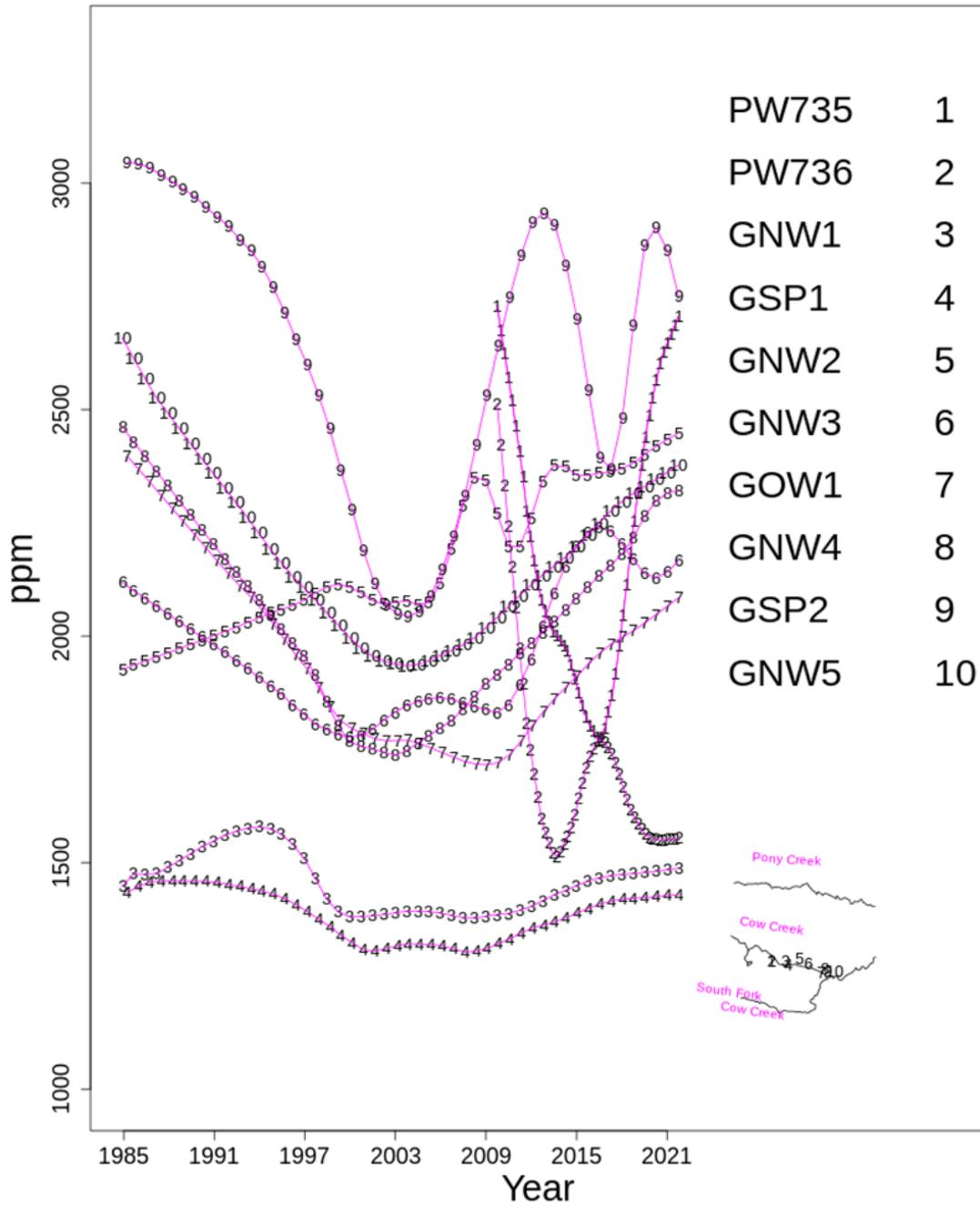


Figure G.10. (Continued, 7 of 9)

### Smoothed Results for Pony Creek Sites: Sulfate

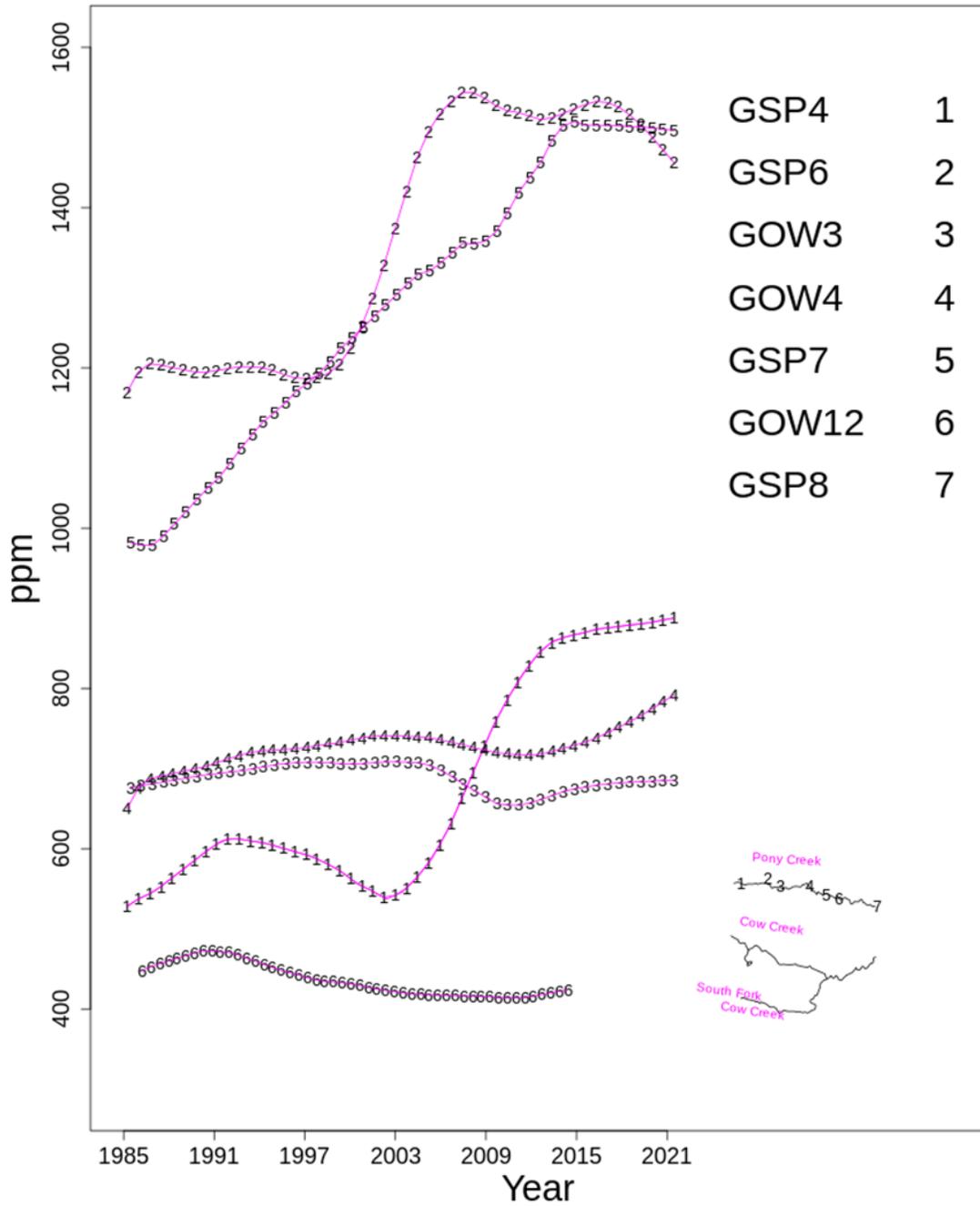


Figure G.10. (Continued, 8 of 9. Trend testing at GOW12 is suspended pending further data.)

### Smoothed Results for South Cow Creek Sites: Sulfate

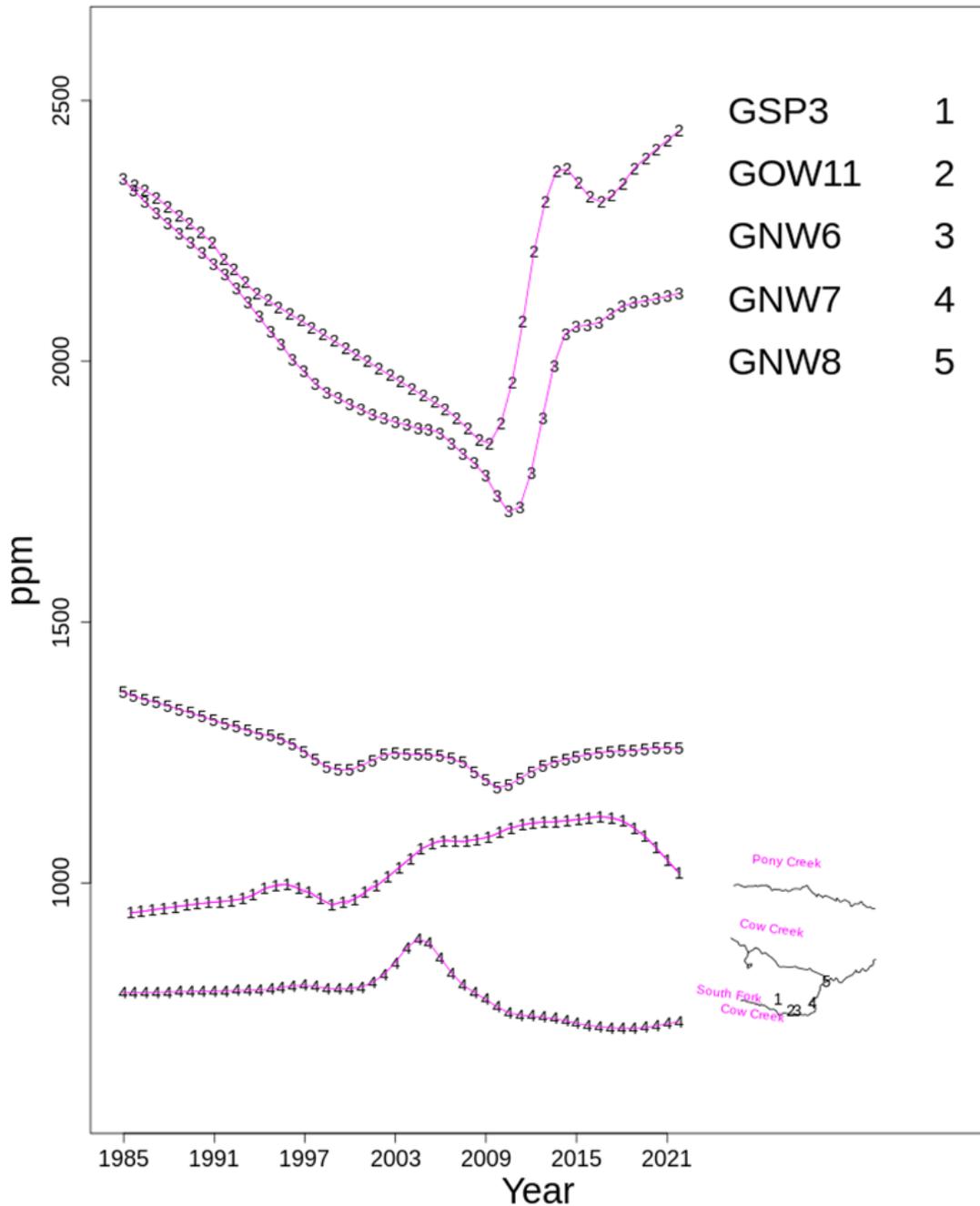


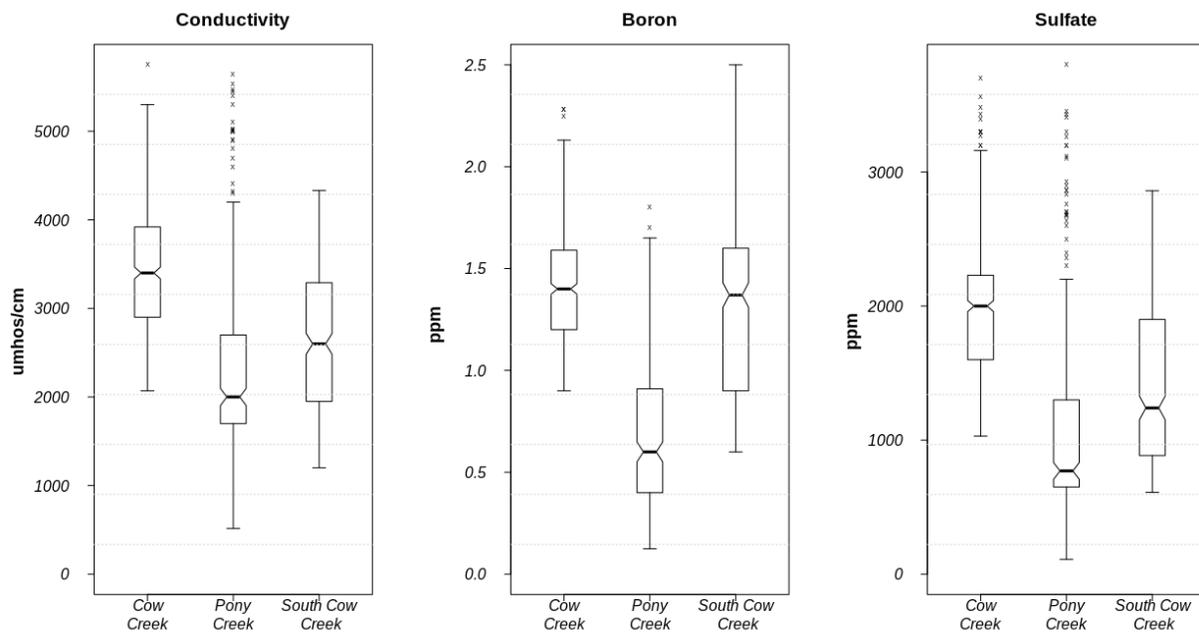
Figure G.10. (Continued, 9 of 9)

### G.3.5 Overall Comparisons between Drainages Using Confidence Interval-Based Tests of Significance

To compare drainages, a graphical test was used based on the box plot (Tukey 1977) shown in Figure G.11. These plots summarize the distributional properties of the data and provide confidence intervals on the median (equaling the mean for symmetric distributions). The centerline within the box represents the sample median of the data, while the lower and upper ends denote the 25<sup>th</sup> and 75<sup>th</sup> percentiles of the data,

respectively. The length of each whisker is either the distance to the next value from the box (lower or higher, related to the lower or higher whisker) or the interquartile range (the difference between the 75% percentile and the 25% percentile) multiplied by 1.5, whichever is smaller. Data points farther away from the end of the whisker are considered to be extreme and are plotted to bring attention to them. The “notch” in the box represents a confidence interval for the median (or mean, under our assumption) that is centered on the sample median. For small data sets, the notch width can be longer than the box; in this case, the box will have “ears.” The confidence interval is a range that is expected to contain the true mean with a specified level of assurance; for example, we are 95% certain that the true mean is contained in this interval. When comparing the underlying means of two sets of data, we initially assume that their means are equal, and we reject this assumption if the two confidence intervals do not overlap.

*The drainage comparisons were done as follows: the means of two or more drainages were judged to be identical (different) if their respective confidence intervals overlap (do not overlap). These confidence intervals are approximately 95%, with a false-detect error rate of about 5% on each comparison (to reduce the overall error rate from all three pairwise comparisons, wider intervals could be used or other approaches may be adopted [Milliken and Johnson 1984]).*



**Figure G.11.** Comparison of Drainages Using All Data.

**Results:** Figure G.11 illustrates the box plot comparisons, based on data from all sampling years. The differences in mean constituent levels between the South Fork Cow Creek and Pony Creek drainages are conclusively nonzero (South Fork Cow Creek has higher means). The conductivity and sulfate means for Cow Creek are higher than the respective means at South Fork Cow Creek, but the mean boron levels are not statistically different in these two drainages.

### G.3.6 Summary of the Statistical Analysis Testing

The *drainage time-trend test results* are:

- No recent drainage trends were found.

The *site-level time-trend test results* are as follows:

- Increasing trends were found at Cow Creek sites PW 735 (conductivity and boron) and GSP 2 (boron).
- At GSP 2, the boron increase is gradual, not abrupt, estimated to be increasing by 24% over the eight recent sampling events, but the increase is certainly not more than 36%.
- Increasing conductivity trends were found at Pony Creek sites GOW 4 and GSP 4.
- Increasing and decreasing trends in boron were found at Pony Creek sites GSP 8 and GOW 4, respectively.
- No trends in sulfate were found at any site, either decreasing or increasing.

Drainage comparisons indicated that the Pony Creek drainage had the lowest values for conductivity, sulfate, and boron. The Cow Creek drainage was distinctly highest in conductivity and sulfate, and the South Fork Cow Creek drainage means fell between the values of the other drainages except for boron, whose mean did not differ from that of Cow Creek.

## G.4 References

Chamberlain, P. J. 2018. "The Bayes Factor Reinterpreted and Renamed as the Predictive Likelihoods Ratio - It's a UMP Composite vs. Composite Test on Multiple Parameters." Technical Bulletin No. 3, Enviro-Sci Consulting.

Cleveland WS. 1981. "LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression." *American Statistician* 35:54.

Friedman JH. 1984. A variable span scatterplot smoother. Laboratory for Computational Statistics, Technical Report No. 5, Stanford University, Palo Alto, California.

Hampel FR, EM Ronchetti, PJ Rousseau, and WA Stahel. 1986. *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley and Sons.

McClave JT and TT Sincich. 2016. *A First Course in Statistics, 12<sup>th</sup> ed.* Pearson, New York, New York.

McDonald JP, KB Olsen, JL Smoot, PJ Chamberlain, SR Jennings, and JD Goering. 2007. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2006*. Battelle – Pacific Northwest Division, Richland, Washington.

Milliken GA and DE Johnson. 1984. *Analysis of Messy Data, Volume 1: Designed Experiments*. Lifetime Learning Publications, Belmont, California.

Olsen KB, JL Smoot, and PJ Chamberlain. 1992. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1991*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1993. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1992*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1994. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1993*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1995. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1994*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1996. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1995*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1997. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1996*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1998. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1997*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 1999. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1998*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2000. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 1999*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2001. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2000*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2002. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2001*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2003. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2002*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2004. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2003*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2005. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2004*. Battelle – Pacific Northwest Division, Richland, Washington.

Olsen KB, JL Smoot, and PJ Chamberlain. 2006. *Hydrogeochemical Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for CY 2005*. Battelle – Pacific Northwest Division, Richland, Washington.

- R Development Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at <http://www.R-project.org>.
- Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York, New York.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2008. *Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana, for 2007*. Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2009. *2008 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4063, Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2010. *2009 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4182, Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2011. *2010 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4261, Battelle – Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and PS Blicher. 2012. *2011 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNWD-4352, Battelle–Pacific Northwest Division, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2013. *2012 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-22498, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2014. *2013 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10010, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2015. *2014 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10027, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2016. *2015 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10043, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, KB Olsen, PD Thorne, PJ Chamberlain, SR Jennings, and LMB Franklin. 2017. *2016 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10056, Pacific Northwest National Laboratory, Richland, Washington.
- Thompson CJ, PJ Chamberlain, and C Lucy. 2019. *2018 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10080, Pacific Northwest National Laboratory, Richland, Washington.

Thompson CJ, PJ Chamberlain, and E Koehler. 2020. *2019 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10098, Pacific Northwest National Laboratory, Richland, Washington.

Thompson CJ, E Koehler, M Larson, PJ Chamberlain, 2021. *2021 Hydrologic/Water Quality Study of Cow and Pony Creek Drainages, Rosebud County, Montana*. PNNL-ACT-10108, Pacific Northwest National Laboratory, Richland, Washington.

Tukey JW. 1977. *Exploratory Data Analysis*. Boston: Addison-Wesley.



## Distribution

### No. of Copies

#### **OFF-SITE (6)**

Mr. David Davenport  
63 Diamond Ranch Lane  
Forsyth, MT 59327

Mr. Gordon Criswell  
Talen Montana, LLC  
P.O. Box 33  
Colstrip, MT 59323

Mr. Karson Kluver  
Genie Land Company  
3034 Rosebud Creek Road  
Forsyth, MT 59327

Ms. Barbara Needham  
Genie Land Company  
3088 Rosebud Creek Road  
Forsyth, MT 59327

Mr. Alex Cunha  
Northern Plains Resource Council  
220 S. 27<sup>th</sup> St. Suite A  
Billings, MT 59101

Mr. Douglas Martens  
Rosebud County Commissioners  
P. O. Box 47  
Forsyth, MT 59327

### No. of Copies

#### **ON-SITE (1)**

Christopher Thompson  
Pacific Northwest National Laboratory  
MS P7-54

#### **ELECTRONIC (4)**

Ms. Terri Mavencamp  
Montana Department of Environmental Quality  
[tmavencamp2@mt.gov](mailto:tmavencamp2@mt.gov)

Ms. Katie Morris  
Montana Department of Environmental Quality  
[katie.morris@mt.gov](mailto:katie.morris@mt.gov)

Ms. Sarah Seitz  
Montana Department of Environmental Quality  
[sarah.seitz@mt.gov](mailto:sarah.seitz@mt.gov)

Ms. Sandy Moisey Scherer  
Montana Board of Environmental Review  
[sandy.moiseyscherer@mt.gov](mailto:sandy.moiseyscherer@mt.gov)

# **Pacific Northwest National Laboratory**

902 Battelle Boulevard  
P.O. Box 999  
Richland, WA 99354  
1-888-375-PNNL (7665)

***[www.pnnl.gov](http://www.pnnl.gov)***