USGS
science for a changing world

DOE/ID-22252

Prepared in cooperation with the U.S. Department of Energy

Optimization of the Idaho National Laboratory Water-Quality Aquifer Monitoring Network, Southeastern Idaho

Scientific Investigations Report 2021–5031

U.S. Department of the Interior
U.S. Geological Survey

Cover. Big Lost River near the Idaho National Laboratory diversion dam, eastern Idaho. Photograph by U.S. Geological Survey, June 2019.

Optimization of the Idaho National Laboratory Water-Quality Aquifer Monitoring Network, Southeastern Idaho

By Jason C. Fisher, Roy C. Bartholomay, Gordon W. Rattray, and Neil V. Maimer

DOE/ID-22252

Prepared in cooperation with the U.S. Department of Energy

Scientific Investigations Report 2021–5031

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2021

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit https://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit https://store.usgs.gov/.

Suggested citation:
Fisher, J.C., Bartholomay, R.C., Rattray, G.W., and Maimer, N.V., 2021, Optimization of the Idaho National Laboratory water-quality aquifer monitoring network, southeastern Idaho: U.S. Geological Survey Scientific Investigations Report 2021–5031 (DOE/ID-22252), 63 p., https://doi.org/10.3133/sir20215031.

Software releases associated with this publication:
Fisher, J.C., 2020, inldata—Collection of datasets for the U.S. Geological Survey-Idaho National Laboratory aquifer monitoring networks: U.S. Geological Survey software release, R package, https://doi.org/10.5066/P9PP9UXZ.

Fisher, J.C., 2021, ObsNetQW—Assessment of a water-quality aquifer monitoring network: U.S. Geological Survey software release, R package, https://doi.org/10.5066/P9X71CSU.

ISSN 2328-0328 (online)

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
foot per day (ft/d)	0.3048	meter per day (m/d)
	Radioactivity	
picocurie per liter (pCi/L)	0.037	becquerel per liter (Bq/L)
	Hydraulic gradient	
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)
	Transmissivity	
foot squared per day (ft^2/d)	0.09290	meter squared per day (m^2/d)

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Volume	
liter (L)	1.057	quart (qt)
liter (L)	0.2642	gallon (gal)
	Mass	
microgram (µg)	0.000000035274	ounce, avoirdupois (oz)
milligram (mg)	0.000035274	ounce, avoirdupois (oz)

Datums

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Supplemental Information

Concentrations of chemical constituents in water are given in either milligrams per liter (mg/L) or micrograms per liter (µg/L).

Activities for radioactive constituents in water are given in picocuries per liter (pCi/L).

Abbreviations

AFT	accelerated failure time
AIC	Akaike information criterion
AIC_c	AIC with correction for small sample sizes
CFA	Central Facilities Area
CFCs	chlorofluorocarbons
CI	confidence interval
CSU	combined standard uncertainty
DL	detection limit
DOE	U.S. Department of Energy
EDF	empirical distribution function
ESRP	eastern Snake River Plain
GA	genetic algorithm
INL	Idaho National Laboratory
INTEC	Idaho Nuclear Technology and Engineering Center
ISLPGA	island parallel genetic algorithm
LD-MDL	less than a long-term method detection level
LOOCV	leave-one-out cross validation
LRL	laboratory reporting level
MCL	maximum contaminant level
MDL	method detection level
MLMS	multilevel monitoring systems
MRL	minimum reporting level
NRF	Naval Reactors Facility
NWQL	National Water Quality Laboratory
OK	ordinary kriging
QA	quality assurance
QC	quality control
QK	quartile kriging
R^2	coefficient of determination
R^2_{wls}	R^2 statistic for a model fitted by weighted least squares regression
RMSD	root-mean-square deviation
RWMC	Radioactive Waste Management Complex
SE	standard error
SEM	standard error of the mean statistic
SSE	weighted sum of squared errors
TAN	Test Area North
USGS	U.S. Geological Survey

Optimization of the Idaho National Laboratory Water-Quality Aquifer Monitoring Network, Southeastern Idaho

By Jason C. Fisher, Roy C. Bartholomay, Gordon W. Rattray, and Neil V. Maimer

Abstract

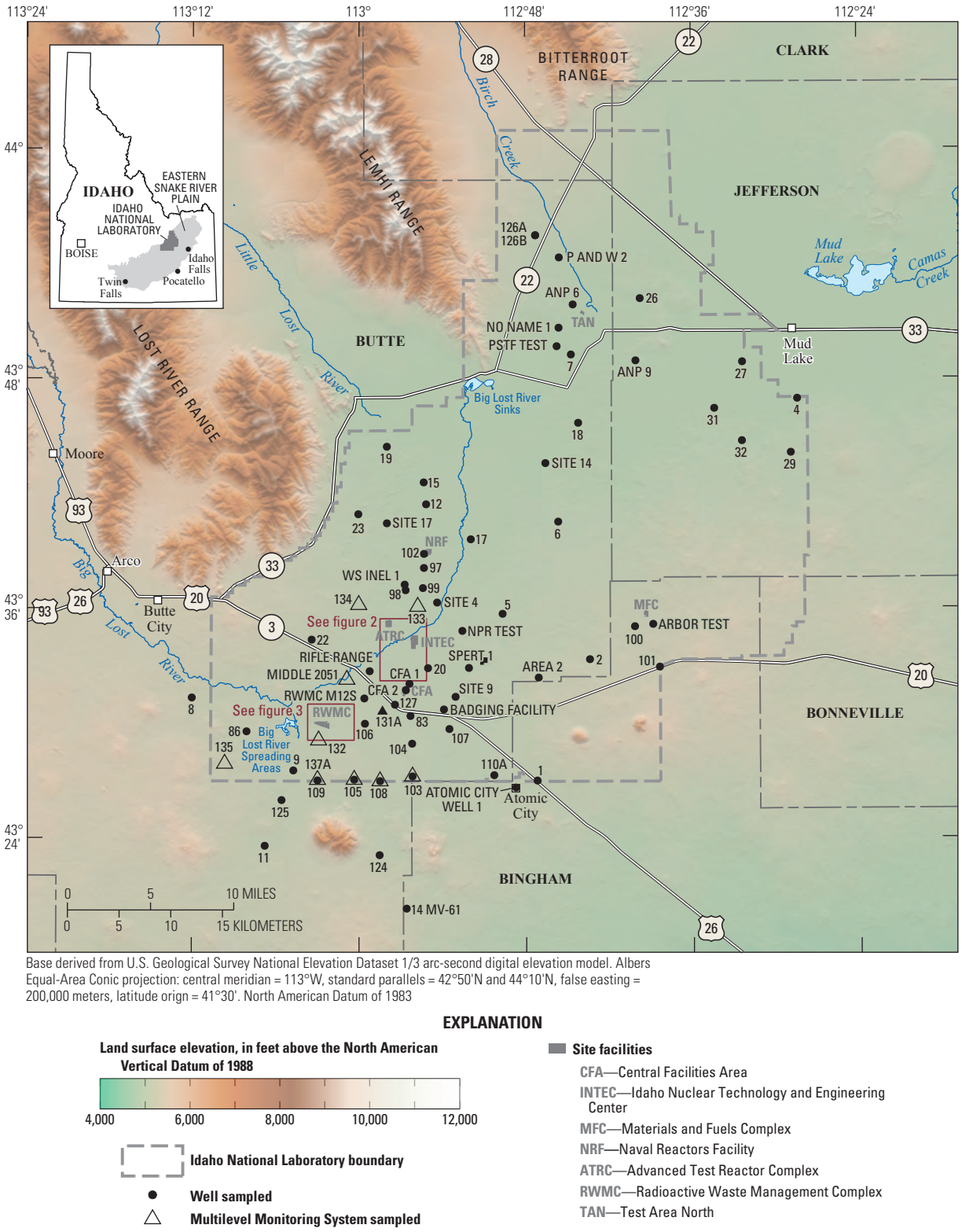
Long-term monitoring of water-quality data collected from wells at the Idaho National Laboratory (INL) has provided essential information for delineating the movement of radiochemical and chemical wastes in the eastern Snake River Plain aquifer, southeastern Idaho. Since 1949, the U.S. Geological Survey, in cooperation with the U.S. Department of Energy, has maintained as many as 200 wells in the INL water-quality monitoring network. A network design tool, distributed as an R package, was developed to evaluate and optimize groundwater monitoring in the existing network based on water-quality data collected at 153 sampling sites since January 1, 1989. The objective of the optimization design tool is to reduce well monitoring redundancy while retaining sufficient data to reliably characterize water-quality conditions in the aquifer. A spatial optimization was used to identify a set of wells whose removal leads to the smallest increase in the deviation between interpolated concentration maps using the existing and reduced monitoring networks while preserving significant long-term trends and seasonal components in the data. Additionally, a temporal optimization was used to identify reductions in sampling frequencies by minimizing the redundancy in sampling events.

Spatial optimization uses an islands genetic algorithm to identify near-optimal network designs removing 10, 20, 30, 40, and 50 wells from the existing monitoring network. With this method, choosing a greater number of wells to remove results in greater cost savings and decreased accuracy of the average relative difference between interpolated maps of the reduced-dataset and the full-dataset. The genetic search algorithm identified reduced networks that best capture the spatial patterns of the average concentration plume while preserving long-term temporal trends at individual wells. Concentration data for 10 analyte types are integrated in a single optimization so that all datasets may be evaluated simultaneously. A constituent was selected for inclusion in the spatial optimization problem when the observations were sufficient to (1) establish a two-range variability model, (2) classify at least one concentration time series as a continuous record block, and (3) make a prediction using the quantile-kriging interpolation method. The selected constituents include sodium, chloride, sulfate, nitrate, carbon tetrachloride, 1,1-dichloroethylene, 1,1,1-trichloroethane, trichloroethylene, tritium, strontium-90, and plutonium-238.

In temporal optimization, an iterative-thinning method was used to find an optimal sampling frequency for each analyte-well pair. Optimal frequencies indicate that for many of the wells, samples may be collected less frequently and still be able to characterize the concentration over time. The optimization results indicated that the sample-collection interval may be increased by an of average of 273 days owing to temporal redundancy.

Introduction

The Idaho National Laboratory (INL), operated by the U.S. Department of Energy (DOE), encompasses about 890 square miles of the eastern Snake River Plain (ESRP) in southeastern Idaho (fig.1; app. 1). The INL was established in 1949 to develop atomic energy, nuclear safety, defense programs, environmental research, and advanced energy concepts. Wastewater disposal sites at the Test Area North (TAN), the Naval Reactors Facility (NRF), the Advanced Test Reactor Complex, and the Idaho Nuclear Technology and Engineering Center (INTEC) (figs. 1–3) have contributed radioactive- and chemical-waste contaminants to the ESRP aquifer. These sites incorporated various wastewater disposal methods, including lined evaporation ponds, unlined percolation (infiltration) ponds and ditches, drain fields, and injection wells. Waste materials buried in shallow pits and trenches within the Subsurface Disposal Area at the Radioactive Waste Management Complex (RWMC) also have contributed contaminants to groundwater.



Base derived from U.S. Geological Survey National Elevation Dataset 1/3 arc-second digital elevation model. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

EXPLANATION

Land surface elevation, in feet above the North American Vertical Datum of 1988

4,000 6,000 8,000 10,000 12,000

Idaho National Laboratory boundary

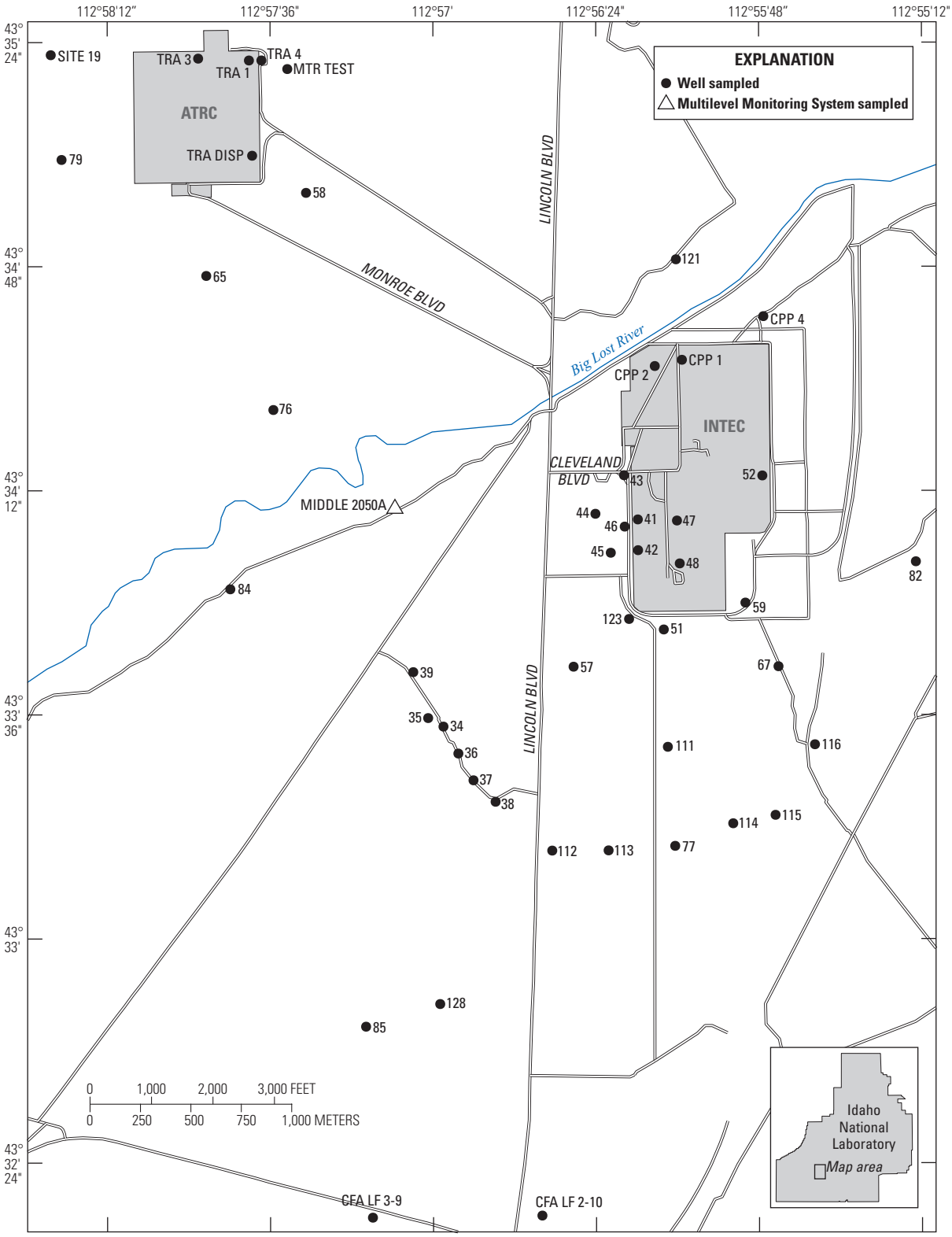
Well sampled

Multilevel Monitoring System sampled

Site facilities

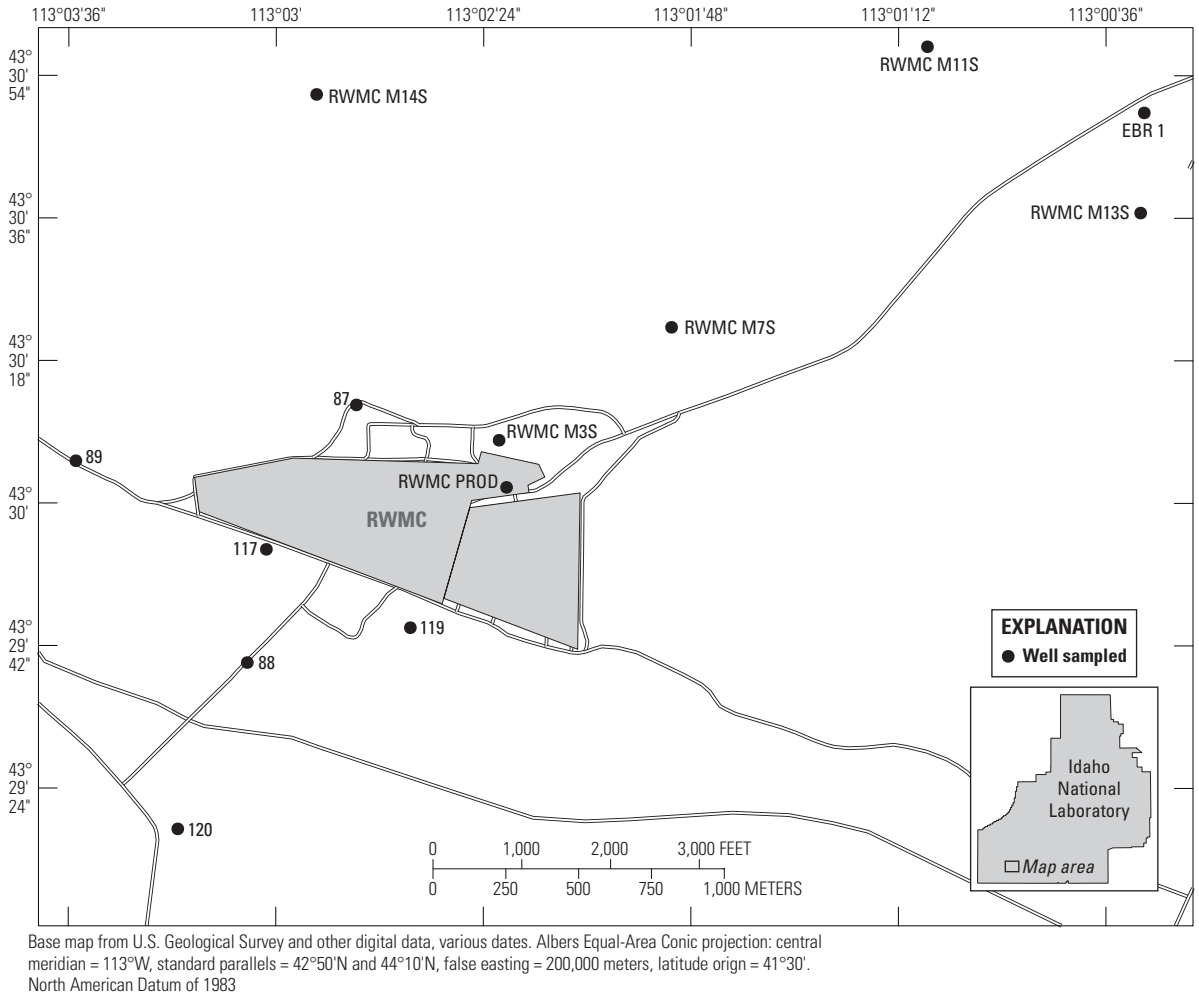
- **CFA**—Central Facilities Area
- **INTEC**—Idaho Nuclear Technology and Engineering Center
- **MFC**—Materials and Fuels Complex
- **NRF**—Naval Reactors Facility
- **ATRC**—Advanced Test Reactor Complex
- **RWMC**—Radioactive Waste Management Complex
- **TAN**—Test Area North

Figure 1. Location of monitor wells selected as part of this study and selected facilities at and near the Idaho National Laboratory, Idaho, 1989–2018.



Base map from U.S. Geological Survey and other digital data, various dates. Albers Equal-Area Conic projection central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

Figure 2. Location of wells in the U.S. Geological Survey aquifer water-quality monitoring network at the Advanced Test Reactor Complex (ATRC) and Idaho Nuclear Technology and Engineering Center (INTEC), Idaho National Laboratory, Idaho, 1989–2018.



Base map from U.S. Geological Survey and other digital data, various dates. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

Figure 3. Location of wells in the U.S. Geological Survey aquifer water-quality monitoring network at the Radioactive Waste Management Complex (RWMC), Idaho National Laboratory, Idaho, 1989–2018.

Since 1949, the U.S. Geological Survey (USGS) has worked in cooperation with the DOE at the INL to define the following:

1. The quality and availability of water for human consumption;
2. The usability of the water for supporting construction and cooling of facilities, and for diluting concentrated waste streams;
3. The location and movement of contaminants in the ESRP aquifer and perched groundwater zones;
4. The sources of recharge to the aquifer;
5. An early detection network for contaminants moving past the INL boundaries; and
6. The processes controlling the origin and distribution of contaminants and naturally occurring constituents in the aquifer (Ackerman and others, 2010).

Since its inception, this water-quality monitoring program at the INL has included a network that once numbered as many as 200 wells. The network of wells has been sampled over the years for tritium, strontium-90, iodine-129, cesium-137, plutonium-238, plutonium-239 and -240 (undivided), americium-241, gross alpha- and gross beta-radioactivity, sodium, bromide, chloride, fluoride, sulfate, nitrate, chromium and other trace elements, volatile organic compounds (VOCs), and total organic carbon (TOC) (Bartholomay, 2013; Davis and others 2013). Most of the wells in this network were constructed as open boreholes, and many are open to the aquifer throughout their entire depth below the water table.

The INL water-quality monitoring network has provided vital information for waste management in the ESRP aquifer. Data from this network have been used to identify contaminant concentrations and to define patterns of waste migration in the aquifer and perched groundwater zones (Bartholomay and others, 2017, p. 7). Improving the efficiency of the monitoring network is desirable because of high network-maintenance costs and funding constraints. The design of a

long-term monitoring (LTM) network is dependent on the spatial and temporal distribution of constituents in the aquifer. These distributions are extremely complicated in the ESRP given its diverse geology, perched alluvial conditions that overlie the regional aquifer, variable fluxes between groundwater and surface water, rapid preferential flow in certain geologic layers, and the long and uncertain history of wastewater disposal at INL facilities.

A more efficient LTM network may be established by evaluating the value of observations measured at each sampling site and the degree to which observations are statistically redundant. Spatially redundant wells were identified for removal from the network and the frequency of sampling was reduced where temporal redundancy was identified in the sampling record. In this report, a *heuristic* optimization procedure was used to redesign the existing INL aquifer water-quality monitoring network. Heuristic is a technique for efficiently guiding the process of optimization; it does not guarantee that the best solution will be determined. This study was conducted by the USGS in cooperation with the DOE.

Purpose and Scope

This report presents an optimization analysis of water-quality data collected from selected wells completed in the ESRP aquifer at and near the INL to identify and remove redundancy in the existing monitoring network, for the purpose of reducing LTM costs while incurring a minimal loss of statistical information. Redundancy is defined by Cameron (2004) as the ability of a reduced-dataset to accurately reconstruct features or characteristics that were estimated from the full-dataset. The cost savings derived from the removal of sampling sites (or locations) from the existing network, or reduction in sampling frequency, is realized by not collecting (and analyzing) the additional water samples. Spatial and temporal redundancy were examined using two different approaches; that is, the spatial and temporal components of the optimization were performed separately.

Spatial optimization was performed by removing redundant sampling sites from the existing monitoring network. The redundancy of a removed site was evaluated by assessing the ability of the reduced-dataset to accurately represent plume maps interpolated for selected analytes using the full-dataset, where a full-dataset map is assumed to provide a realistic estimate of the concentration plume in the aquifer. Maps of the spatial distribution of analyte concentrations in groundwater were predicted by kriging, a geostatistical method that interpolates concentration values for locations between sampling sites. The quality of the network design also entails consideration of other (sometimes competing) objectives including (1) minimizing the interpolation error to ensure that the best spatial coverage is retained in the reduced-monitoring network; (2) safeguarding against the removal of sites with significant long-term trends that may be useful for evaluating the effectiveness of remediation efforts at the INL; and (3) safeguarding against the removal of sites with repeated sampling at regular intervals over multiple years, and showing large variability in analyte concentrations, so as to preserve the long-term history of the sampling program. The presence and slope of long-term trends were estimated using survival analysis, a regression method that accounts for censored concentration data. A local regression analysis (also known as scatterplot smoothing) was used to estimate the presence and variability of sampling at a well. The final decision on the number of sampling sites to remove from the existing monitoring network was left as a management decision and entails a trade-off between cost savings and information loss.

Temporal optimization was performed on a well-by-well basis using an iterative-thinning method proposed by Cameron (2003). This method examines whether the historical sampling frequency for a given well location and constituent may be reduced because of temporal redundancy in the sampling events. Sampling events are removed from the sampling record to characterize the level of redundancy in trends (Cameron, 2004, p. 91). A local regression analysis was used to estimate trend (that is, the long-term trend with seasonality) over the sampled data range. The procedure used to evaluate the ability of a reduced-dataset to accurately represent the existing trend was as follows:

1. Estimate the uncertainty around the trend fitted to the full-dataset,
2. Estimate the trend fitted to the reduced-dataset, and
3. Calculate the proportion of this trend that is within the uncertainty interval of the existing trend.

Repeated random sampling was used to safeguard against irregular trends that may arise from the selection of a single set of sampling events to remove from the existing sampling record. The number of sampling events that may be removed was constrained by an upper limit placed on the proportion of the trend that may lie outside the uncertainty interval of the existing trend, thus ensuring an acceptable level of accuracy in trend estimates. By maximizing the fraction of sampling events to remove from the historical record, an instance of the reduced-dataset was found and its average sampling interval was estimated. An optimal sampling frequency was recommended for each well-analyte combination.

Geohydrologic Setting

The INL is located on the west-central part of the ESRP. The ESRP is a northeast-trending structural basin about 200 mi long and 50–70 mi wide (fig. 1). The basin, bounded by faults on the northwest and by downwarping and faulting on the southeast, has been filled with basaltic lava flows interbedded with terrestrial sediments. The basaltic rocks and sedimentary deposits combine to form the ESRP aquifer, which is the primary source of groundwater for the ESRP.

The ESRP aquifer is one of the most productive aquifers in the United States (U.S. Geological Survey, 1985, p. 193). Groundwater generally moves from northeast to southwest, and eventually discharges to springs along the Snake River downstream from Twin Falls, Idaho, about 100 mi southwest of the INL (fig. 1). Groundwater moves horizontally through basalt interflow zones and vertically through joints and interfingering edges of basalt flows. Infiltration of surface water, heavy pumpage, geohydrologic conditions, and seasonal fluxes of recharge and discharge locally affect the movement of groundwater (Garabedian, 1986). The ESRP aquifer is recharged primarily from infiltration of applied irrigation water, infiltration of streamflow, groundwater inflow from adjoining mountain drainage basins, and infiltration of precipitation (Ackerman and others, 2006).

At the INL, depth to water in wells completed in the ESRP aquifer ranges from about 225 ft below land surface in the northern part of the INL to more than 900 ft below land surface in the southeastern part of the INL (Bartholomay and others, 2017, p. 21). A significant proportion of the groundwater moves through the upper 200–800 ft of basaltic rock (Mann, 1986, p. 21). Ackerman (1991, p. 30) and Bartholomay and others (1997, table 3) reported transmissivity values for basalt in the upper part of the aquifer ranging from 1.1 to 760,000 ft^2/d. The lateral hydraulic gradient at the INL ranges from 2 to 10 ft/mi, with an average of 4 ft/mi (Davis and others, 2013, fig. 9). Horizontal flow velocities of 2 to 26 ft/d have been calculated based on the movement of various constituents in different areas of the aquifer at and near the INL (Robertson and others, 1974; Mann and Beasley, 1994; Cecil and others, 2000; Plummer and others, 2000; and Busenberg and others, 2001). These flow velocities equate to a travel time of about 50–700 years for water beneath the INL to travel to springs that discharge at the terminus of the ESRP groundwater-flow system near Twin Falls, Idaho (fig. 1). Localized tracer tests at the INL have shown that vertical- and horizontal-solute transport rates are as high as 60–150 ft/d (Nimmo and others, 2002; Duke and others, 2007).

Olmsted (1962), Robertson and others (1974), and Busenberg and others (2001) classified groundwater at the INL based on chemical types derived from dissolution of the rocks and minerals within the recharge source areas. Olmsted's Type A water consisted of calcium and magnesium concentrations that constituted at least 85 percent of the cations, and bicarbonate constituted at least 70 percent of the anions. Type A water is present in the central and western part of the INL. Type A water is attributed to seepage loss from the Big Lost River and from groundwater underflow from the Big Lost River, Little Lost River, and Birch Creek drainage basins to the west and northwest of the INL (fig. 1) that contain alluvium derived from Paleozoic carbonate rocks from the surrounding mountains.

Olmsted's Type B water, which is characterized by higher equivalent fractions of sodium, potassium, fluoride, and silica than Type A water, underlies much of the eastern part of the INL and is often referred to as regional water. The groundwater originates from the area northeast of the INL that is composed of a much higher fraction of rhyolitic and andesitic volcanic rocks than mountains west and northwest of the INL that contribute to Type A water. Busenberg and others (2001) used age dating techniques of chlorofluorocarbons (CFCs), sulfur hexafluoride, and tritium/helium to further classify the regional water at the INL into two types based on the recharge type of the young fraction of groundwater. Water in the southeastern part of the INL represented a binary mixture of old (water greater than 40 and 55 years old that did not contain tritium or CFCs, respectively) regional groundwater underflow with young water derived from rapid, focused recharge, probably from precipitation infiltration. Water in the northeastern part of the INL is (1) old, regional groundwater underflow that is mixed with local rapid, focused recharge; (2) slow, diffuse areal recharge through the unsaturated zone; and (3) agricultural return flow from the Mud Lake area (fig. 1).

Previous Investigations

The USGS INL Project Office has examined the hydrologic conditions and distribution of selected wastewater constituents in groundwater and perched groundwater at the INL since operations began in 1949. Numerous previous investigations on the hydrology, water quality, and geology have been conducted by INL contractors, State agencies, and the USGS. The USGS provides a list of references to published reports from its previous INL studies at the USGS INL Project Office web page (U.S. Geological Survey, 2020).

In 2010, the USGS INL Project Office began studies to optimize their LTM networks. Fisher (2013) performed an optimization of the water-level monitoring network using a kriging-based genetic algorithm method. He identified 40 wells that could be removed from the monitoring network without significant loss of accuracy.

For the optimization of the aquifer water-quality monitoring network, three reports initially were planned, of which this report was the third. Bartholomay and others (2012) used statistical methods to determine trends for selected constituents for 67 wells and 7 surface-water sites sampled at the INL that were considered unaffected by wastewater disposal. They determined that chloride trends in wells influenced by recharge from the Big Lost River either decreased or had variable increases and decreases because of wet and dry cycles of precipitation and runoff. Wells influenced by regional recharge showed increasing trends for chloride, sodium, sulfate, and nitrate, and increases were attributed to agricultural or other anthropogenic influences upgradient from the INL. Some wells near the NRF and Power Burst Facility showed increasing trends, possibly owing to wastewater disposal at those facilities.

Davis and others (2015) used statistical methods to determine trends for selected constituents from 64 aquifer wells and 35 perched groundwater wells at the INL that were believed to be influenced by wastewater disposal. Trend test results for

tritium and strontium-90 concentrations in aquifer wells indicated that nearly all wells had decreasing or no trends. Trend test results for chloride, sodium, sulfate, nitrate, chromium, trace elements, and TOC concentrations in aquifer wells also indicated that most wells had decreasing or no trends. Decreasing trends were attributed to discontinued disposal practices and dilution and dispersion in the aquifer. Sodium concentrations showed increasing trends in wells at Central Facilities Area (CFA) and downgradient. Carbon tetrachloride had increasing trends in a couple of wells near RWMC, but all other wells showed decreasing or no trend.

Optimizing LTM networks with the aim of removing data from the system because they add little to no beneficial information has received increasing attention in the recent past. An in-depth description of the optimization problem was provided by Cameron (2004). Previous efforts to eliminate redundancy in existing groundwater monitoring networks have separately examined the temporal and spatial components of redundancy. Johnson and others (1996) performed a temporal optimization by reducing sampling frequency. Reed and others (2000) performed a spatial optimization by reducing the number of sampling locations. And Cameron and Hunter (2002) reduced redundancy both spatially and temporally by performing separate optimizations for each domain—an approach that was also used in this report.

Use of prediction uncertainty is another promising approach for network design. Fienen and others (2010) use a PEST framework to evaluate the uncertainty of a model prediction to determine observations that may be excluded from an existing hydrologic monitoring network.

Computer Software

Using a software development methodology, we took a highly reproducible approach for optimizing the USGS INL water-quality aquifer monitoring network. Reproducibility requires archiving and documenting all datasets and computer source code used to optimize/analyze the monitoring network—an undertaking made easier by the advances in open source software, open file formats, and cloud computing. The collection of source code and processing instructions used to optimize the monitoring network was placed in a software package referred to as **ObsNetQW** (Fisher, 2021). The collection of datasets available for the USGS INL water-quality and water-level monitoring networks was placed in a software package named **inldata** (Fisher, 2020). These packages are an extension of the R-programming language (R Core Team, 2019) and allow for easy, transparent, and cross-platform distribution of their content by enforcing a set of formal format standards. A manual describing package datasets and processing programs (also known as functions in R) is provided in appendix 2.

Sources and Descriptions of Data

Water-Quality Data Collection

The USGS monitoring of groundwater quality in the ESRP aquifer beneath the INL and vicinity is an ongoing long-term program that began in 1949. Water samples collected from wells in the monitoring network were analyzed for some combination of concentrations of tritium, strontium-90, cesium-137, plutonium-238, plutonium-239 and -240 (undivided), americium-241, gross alpha and beta-particle radioactivity; chromium, sodium, chloride, and sulfate; nutrients including nitrite plus nitrate (as nitrogen [N]), nitrite (as N), orthophosphate (as phosphorus [P]), and ammonia (as N); and VOCs. This report presents an analysis that uses a subset of this water-quality dataset that was collected during 1989–2018 (U.S. Geological Survey, 2019). Data collected prior to 1989 were omitted from analysis to eliminate bias introduced through (1) samples being tainted by cable drilling, and (2) changes in the analytical testing laboratory for chemical constituents and organic compounds.

Starting in 1989, water samples were collected from 133 ESRP aquifer wells (figs. 1–3; app. 1), or 153 sampling sites when accounting for multilevel monitoring (table 1). During 2005–12, 10 wells were equipped with multilevel monitoring systems (MLMS) that allow water-quality measurements to be acquired at isolated depths (figs. 1–2). Since 1989, water samples have been analyzed for chemical constituents and organic compounds at the USGS National Water Quality Laboratory (NWQL) in Lakewood, Colorado. Prior to 1989, water samples were analyzed by various laboratories for chloride, chromium, sodium, and nitrate (Wegner, 1989). Water samples have been analyzed for radionuclides at the DOE Radiological and Environmental Sciences Laboratory at the INL since samples were first collected. Many of the samples collected in the 1950s, 1960s, and 1970s were collected during or immediately after cable drilling, and some of the samples had a high probability of containing impurities introduced by the drilling (Robertson and others, 1974, app. B).

Methods used to sample and analyze for selected constituents generally follow the guidelines established by the USGS (Goerlitz and Brown, 1972; Stevens and others, 1975; Wood, 1976; Thatcher and others, 1977; Claassen, 1982; Wershaw and others, 1987; Fishman and Friedman, 1989; Faires, 1993; Fishman, 1993; and U.S. Geological Survey, variously dated). Water samples were collected according to a quality-assurance plan for water-quality activities conducted by personnel at the USGS INL Project Office. The plan was finalized in June 1989 and revised in March 1992, in 1996 (Mann, 1996), in 2003 (Bartholomay and others, 2003), in 2008 (Knobel and others, 2008), and in 2014 (Bartholomay and others, 2014). The plan is available for inspection at the USGS INL Project Office.

Table 1. Wells in the U.S. Geological Survey aquifer water-quality monitoring network, eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 1989–2018.

[**Well name:** Local well name with site locations shown in figs. 1, 2, and 3. **Site No.:** Site identifier assigned by the U.S. Geological Survey (https://waterdata.usgs.gov/nwis). **Well depth:** feet bls, feet below land surface. For wells instrumented with Multilevel Monitoring Systems, well depth represents a sampling-port depth. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of replicate samples:** Number of replicate samples used to estimate variability (random measurement error) of analytical results. Replicate samples are composed of a pair of environmental samples collected sequentially, over a period of 12 hours or less, from a monitoring well]

Well name	Site No.	Well depth (feet bls)	Period of record (mm-dd-yy)	Number of environmental samples	Number of replicate samples
ANP 6	435152112443101	295	07-02-92 to 10-18-17	51	5
ANP 9	434856112400001	322	04-13-94 to 10-18-11	55	7
ARBOR TEST	433509112384801	790	02-15-89 to 04-21-11	72	9
AREA 2	433223112470201	876	09-20-90 to 10-15-18	52	5
ATOMIC CITY WELL 1	432638112484101	639	04-05-89 to 04-09-18	82	9
BADGING FACILITY	433042112535101	644	04-28-89 to 04-11-18	66	6
CFA 1	433204112562001	639	01-25-89 to 04-11-18	134	8
CFA 2	433144112563501	681	01-25-89 to 10-10-17	122	8
CFA LF 2-10	433216112563301	716	10-20-93 to 04-23-18	73	7
CFA LF 3-9	433216112571001	500	10-22-93 to 04-02-07	32	6
CPP 1	433433112560201	586	04-26-89 to 04-19-18	101	6
CPP 2	433432112560801	605	01-31-89 to 10-04-18	90	7
CPP 4	433440112554401	700	04-26-89 to 10-18-16	83	5
EBR 1	433051113002601	1,075	04-28-89 to 04-20-11	76	9
HIGHWAY 3	433256113002501	750	04-05-89 to 10-16-18	79	7
MIDDLE 2050A	433409112570515	539	09-30-05 to 06-12-18	21	2
MIDDLE 2051	433217113004901	1,177	09-29-05 to 06-28-17	16	0
MIDDLE 2051	433217113004903	1,128	09-28-05 to 06-13-18	21	2
MIDDLE 2051	433217113004906	876	09-28-05 to 06-28-17	19	2
MIDDLE 2051	433217113004909	771	09-27-05 to 06-13-18	23	5
MTR TEST	433520112572601	588	03-27-89 to 04-02-18	91	11
NO NAME 1	435038112453401	552	05-22-91 to 04-11-18	69	6
NPR TEST	433449112523101	600	06-20-91 to 10-11-18	67	7
P AND W 2	435419112453101	378	04-18-89 to 04-02-18	83	7
PSTF TEST	434941112454201	319	07-13-89 to 10-18-11	61	6
RIFLE RANGE	433243112591101	620	04-10-02 to 10-17-18	32	4
RWMC M11S	433058113010401	624	03-22-00 to 10-11-18	35	3
RWMC M12S	433118112593401	572	03-22-00 to 10-11-18	33	3
RWMC M13S	433037113002701	643	03-22-00 to 10-11-18	36	4
RWMC M14S	433052113025001	635	03-22-00 to 10-11-18	37	5
RWMC M3S	433008113021801	633	02-26-93 to 10-23-18	51	6
RWMC M7S	433023113014801	628	03-04-93 to 10-22-18	53	7
RWMC PROD	433002113021701	685	01-18-89 to 12-12-18	424	12
SITE 14	434334112463101	717	04-18-89 to 10-10-18	86	8
SITE 17	434027112575701	600	06-18-91 to 03-26-18	55	5
SITE 19	433522112582101	860	04-12-89 to 10-16-18	75	9
SITE 4	433617112542001	495	04-18-91 to 04-11-18	63	7

Table 1. Wells in the U.S. Geological Survey aquifer water-quality monitoring network, eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 1989–2018.—Continued

[Well name: Local well name with site locations shown in figs. 1, 2, and 3. **Site No.:** Site identifier assigned by the U.S. Geological Survey (https://waterdata.usgs.gov/nwis). **Well depth:** feet bls, feet below land surface. For wells instrumented with Multilevel Monitoring Systems, well depth represents a sampling-port depth. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of replicate samples:** Number of replicate samples used to estimate variability (random measurement error) of analytical results. Replicate samples are composed of a pair of environmental samples collected sequentially, over a period of 12 hours or less, from a monitoring well]

Well name	Site No.	Well depth (feet bls)	Period of record (mm-dd-yy)	Number of environmental samples	Number of replicate samples
SITE 9	433123112530101	1,057	04-05-89 to 03-27-18	75	7
SPERT 1	433252112520301	653	04-28-89 to 04-11-18	66	7
TRA 1	433521112573801	600	04-27-89 to 11-06-13	54	5
TRA 3	433522112573501	602	04-27-89 to 10-16-18	62	9
TRA 4	433521112574201	965	11-13-89 to 04-04-17	67	11
TRA DISP	433506112572301	1,267	02-01-89 to 10-16-18	103	9
USGS 1	432700112470801	630	05-30-91 to 10-18-18	66	6
USGS 100	433503112400701	750	04-19-89 to 04-03-18	94	10
USGS 101	433255112381801	842	04-19-89 to 10-18-18	81	7
USGS 102	433853112551601	445	06-08-90 to 05-09-18	129	9
USGS 103	432714112560701	1,297	04-14-89 to 04-18-05	108	6
USGS 103	432714112560702	1,279	09-25-07 to 06-26-18	16	1
USGS 103	432714112560704	1,240	09-25-07 to 06-26-18	17	2
USGS 103	432714112560708	1,098	10-01-07 to 06-26-18	17	1
USGS 103	432714112560712	1,014	10-01-07 to 06-26-18	16	1
USGS 104	432856112560801	700	04-14-89 to 10-16-18	141	16
USGS 105	432703113001801	1,300	04-21-89 to 10-16-07	74	6
USGS 105	432703113001807	1102	09-17-09 to 06-27-18	15	2
USGS 105	432703113001811	982	09-17-09 to 06-27-18	13	1
USGS 105	432703113001815	862	09-18-09 to 06-27-18	14	1
USGS 106	432959112593101	760	04-14-89 to 10-17-18	99	11
USGS 107	432942112532801	690	04-11-89 to 03-27-18	87	9
USGS 108	432659112582601	1,196	04-21-89 to 04-29-08	75	6
USGS 108	432659112582602	1,194	09-20-10 to 06-25-18	11	1
USGS 108	432659112582606	1,060	09-22-10 to 06-26-13	7	2
USGS 108	432659112582610	904	09-20-10 to 06-26-13	7	2
USGS 109	432701113025601	800	04-21-89 to 04-23-18	87	8
USGS 11	432336113064201	704	05-01-89 to 04-09-18	94	10
USGS 110A	432717112501502	644	10-25-95 to 10-18-18	53	5
USGS 111	433331112560501	560	01-05-89 to 04-17-18	94	9
USGS 112	433314112563001	509	03-30-89 to 10-01-18	134	7
USGS 113	433314112561801	556	03-31-89 to 04-17-18	129	7
USGS 114	433318112555001	560	03-31-89 to 10-01-18	134	11
USGS 115	433320112554101	581	04-06-89 to 10-01-18	131	9
USGS 116	433331112553201	572	04-06-89 to 04-17-18	130	7
USGS 117	432955113025901	655	01-12-89 to 10-16-18	148	9
USGS 119	432945113023401	705	01-12-89 to 04-23-18	144	11

Table 1. Wells in the U.S. Geological Survey aquifer water-quality monitoring network, eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 1989–2018.—Continued

[Well name: Local well name with site locations shown in figs. 1, 2, and 3. **Site No.:** Site identifier assigned by the U.S. Geological Survey (https://waterdata.usgs.gov/nwis). **Well depth:** feet bls, feet below land surface. For wells instrumented with Multilevel Monitoring Systems, well depth represents a sampling-port depth. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of replicate samples:** Number of replicate samples used to estimate variability (random measurement error) of analytical results. Replicate samples are composed of a pair of environmental samples collected sequentially, over a period of 12 hours or less, from a monitoring well]

Well name	Site No.	Well depth (feet bls)	Period of record (mm-dd-yy)	Number of environmental samples	Number of replicate samples
USGS 12	434126112550701	563	06-15-90 to 03-26-18	128	9
USGS 120	432919113031501	705	01-12-89 to 10-09-18	174	8
USGS 121	433450112560301	475	03-25-91 to 04-04-18	81	8
USGS 123	433352112561401	515	03-25-91 to 10-02-18	74	7
USGS 124	432307112583101	800	04-21-94 to 04-09-18	71	7
USGS 125	432602113052801	774	04-27-95 to 10-09-18	65	7
USGS 126A	435529112471301	648	11-08-00 to 04-19-11	29	4
USGS 126B	435529112471401	472	11-08-00 to 10-10-18	36	4
USGS 127	433058112572201	596	09-27-00 to 04-03-18	42	4
USGS 128	433250112565601	615	10-31-01 to 10-16-18	28	4
USGS 131A	433036112581803	1,157	10-24-12 to 06-15-15	5	0
USGS 131A	433036112581806	1,058	10-29-12 to 06-15-15	5	0
USGS 131A	433036112581810	842	10-29-12 to 06-19-18	8	0
USGS 131A	433036112581815	632	10-29-12 to 06-19-18	10	2
USGS 132	432906113025018	787	09-05-06 to 06-20-18	17	0
USGS 133	433605112554312	480	09-24-07 to 06-12-18	19	3
USGS 134	433611112595815	652	09-28-06 to 06-29-11	12	2
USGS 134	433611112595819	590	09-27-06 to 06-11-18	16	1
USGS 135	432753113093609	861	09-15-09 to 06-14-18	14	1
USGS 137A	432701113025801	895	10-23-12 to 06-19-17	6	0
USGS 137A	432701113025803	862	10-23-12 to 06-19-17	6	0
USGS 137A	432701113025805	784	10-23-12 to 06-18-18	8	0
USGS 137A	432701113025807	718	10-24-12 to 06-18-18	8	0
USGS 14 MV-61	432019112563201	752	04-01-89 to 10-09-18	95	8
USGS 15	434234112551701	610	06-06-90 to 10-24-11	70	4
USGS 17	433937112515401	498	12-14-89 to 04-09-18	87	4
USGS 18	434540112440901	329	10-12-90 to 04-12-18	53	7
USGS 19	434426112575701	399	04-03-89 to 04-02-18	84	8
USGS 2	433320112432301	699	05-28-91 to 04-12-18	51	7
USGS 20	433253112545901	658	04-01-89 to 04-17-18	89	8
USGS 22	433422113031701	657	04-05-89 to 04-25-11	60	8
USGS 23	434055112595901	458	05-21-91 to 10-09-18	68	8
USGS 26	435212112394001	267	05-23-91 to 04-10-18	67	6
USGS 27	434851112321801	312	03-24-89 to 04-02-18	83	7
USGS 29	434407112285101	426	06-12-91 to 10-11-18	55	7
USGS 31	434625112342101	428	06-12-91 to 04-10-18	55	7
USGS 32	434444112322101	392	06-12-91 to 04-10-18	56	7

Table 1. Wells in the U.S. Geological Survey aquifer water-quality monitoring network, eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 1989–2018.—Continued

[Well name: Local well name with site locations shown in figs. 1, 2, and 3. **Site No.:** Site identifier assigned by the U.S. Geological Survey (https://waterdata.usgs.gov/nwis). **Well depth:** feet bls, feet below land surface. For wells instrumented with Multilevel Monitoring Systems, well depth represents a sampling-port depth. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of replicate samples:** Number of replicate samples used to estimate variability (random measurement error) of analytical results. Replicate samples are composed of a pair of environmental samples collected sequentially, over a period of 12 hours or less, from a monitoring well]

Well name	Site No.	Well depth (feet bls)	Period of record (mm-dd-yy)	Number of environmental samples	Number of replicate samples
USGS 34	433334112565501	700	03-31-89 to 04-16-18	85	8
USGS 35	433339112565801	579	03-31-89 to 10-03-18	87	8
USGS 36	433330112565201	567	03-31-89 to 04-16-18	123	9
USGS 37	433326112564801	572	03-31-89 to 10-03-18	95	11
USGS 38	433322112564301	724	03-31-89 to 04-16-18	90	9
USGS 39	433343112570001	492	03-31-89 to 10-21-13	119	9
USGS 4	434657112282201	553	06-04-91 to 10-19-11	60	5
USGS 41	433409112561301	666	04-07-89 to 10-04-18	84	7
USGS 42	433404112561301	678	04-07-89 to 04-19-18	85	7
USGS 43	433415112561501	564	04-20-89 to 10-02-18	84	7
USGS 44	433409112562101	650	04-07-89 to 04-18-18	104	10
USGS 45	433402112561801	651	04-07-89 to 10-02-18	95	8
USGS 46	433407112561501	651	04-07-89 to 04-18-18	98	9
USGS 47	433407112560301	651	01-31-89 to 10-04-18	102	9
USGS 48	433401112560301	750	04-06-89 to 04-19-18	92	9
USGS 5	433543112493801	494	09-26-90 to 04-12-18	67	6
USGS 51	433350112560601	647	04-17-89 to 04-18-18	84	9
USGS 52	433414112554201	602	04-07-89 to 10-03-18	86	7
USGS 57	433344112562601	582	01-05-89 to 10-02-18	127	7
USGS 58	433500112572502	503	04-13-89 to 04-02-18	90	14
USGS 59	433354112554701	590	04-17-89 to 04-18-18	90	11
USGS 6	434031112453701	620	09-26-90 to 10-20-11	45	6
USGS 65	433447112574501	498	02-17-89 to 04-04-18	138	9
USGS 67	433344112554101	694	04-11-89 to 10-01-18	74	9
USGS 7	434915112443901	903	05-20-91 to 04-11-18	66	7
USGS 76	433425112573201	718	04-07-89 to 04-05-18	99	11
USGS 77	433315112560301	586	03-31-89 to 10-01-18	87	8
USGS 79	433505112581901	702	04-01-89 to 04-05-18	88	8
USGS 8	433121113115801	812	05-01-89 to 04-10-18	82	7
USGS 82	433401112551001	693	04-14-89 to 04-17-18	125	9
USGS 83	433023112561501	752	04-13-89 to 04-25-11	79	10
USGS 84	433356112574201	505	04-01-89 to 10-18-18	85	8
USGS 85	433246112571201	614	04-01-89 to 04-16-18	88	9
USGS 86	432935113080001	691	04-21-89 to 06-06-18	81	6
USGS 87	433013113024201	673	01-04-89 to 04-10-18	141	9
USGS 88	432940113030201	663	01-04-89 to 10-22-18	196	9
USGS 89	433005113032801	714	01-04-89 to 05-09-18	140	9

Table 1. Wells in the U.S. Geological Survey aquifer water-quality monitoring network, eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 1989–2018.—Continued

[Well name: Local well name with site locations shown in figs. 1, 2, and 3. **Site No.:** Site identifier assigned by the U.S. Geological Survey (https://waterdata.usgs.gov/nwis). **Well depth:** feet bls, feet below land surface. For wells instrumented with Multilevel Monitoring Systems, well depth represents a sampling-port depth. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of replicate samples:** Number of replicate samples used to estimate variability (random measurement error) of analytical results. Replicate samples are composed of a pair of environmental samples collected sequentially, over a period of 12 hours or less, from a monitoring well]

Well name	Site No.	Well depth (feet bls)	Period of record (mm-dd-yy)	Number of environmental samples	Number of replicate samples
USGS 9	432740113044501	654	04-21-89 to 10-17-18	86	5
USGS 97	433807112551501	510	04-19-89 to 03-26-18	153	9
USGS 98	433657112563601	508	04-19-89 to 10-15-18	149	7
USGS 99	433705112552101	440	04-19-89 to 10-11-18	152	7
WS INEL 1	433716112563601	490	04-19-89 to 04-05-18	102	5

Sample collection methods varied for several of the wells during the history of sampling. Permanent pumps were installed at various dates with most installations occurring from 1985 to the early 1990s. Prior to installation of pumps, wells were sampled using a portable thief sampler. Some of the samples collected with thief samplers were collected at different depths in the aquifer during the same sampling event. When the depths at which thief samples were collected were known, the data from the depth similar to the depth to which the current pump was set were used in the analyses. After pumps were installed, wells were purged for at least three well volumes prior to sample collection until October 2003, when procedures were changed to allow sample collection after one well volume was purged. Studies by Bartholomay (1993) and Knobel (2006) indicated that different purge rates used at the INL did not affect the analytical results for the wells analyzed in the respective studies.

Beginning in 1980, about 10 percent of water samples were collected for quality assurance (QA) purposes. Quality control (QC) water samples collected by the USGS INL Project Office generally include equipment blanks, splits, and blind replicates; however, other types of QC samples also have been collected throughout the history of the program. Comparative studies to determine agreement between analytical results for water-sample pairs by laboratories used by the INL Project Office QA program were summarized by Wegner (1989); Williams (1996, 1997); Rattray (2012); Davis and others, (2013); and Rattray (2014). Wegner (1989) also statistically compared analytical results among different laboratories used from 1980 to 1988. Analyses of water-sample pairs were in statistical agreement for more than 95 percent of the samples compared.

The period of record, sample collection frequency, and list of analytes tested for, varied for all sites in the monitoring network. Since 2003, all sites have been sampled annually, but prior to that time frame, wells were sampled annually, semi-annually, quarterly, or even more frequently depending on the purpose of the sampling program. Some gaps in data occurred when pumps were out for repair, samples were lost, or program changes did not necessitate sampling of the analyte in question.

Nondetect Data

Measurements whose values are known only to be below a threshold (also known as a censoring level) are referred to as nondetects (or "left-censored data" in the statistical literature). Historically, the threshold used to censor analytical results was the minimum reporting level (MRL). The MRL is defined by the NWQL as the smallest measured concentration of a substance that can be measured reliably by using a given analytical method (Timme, 1995). Methods for choosing MRLs are subjective and no single procedure is universally used. In 1996, the NWQL began censoring data at the laboratory reporting level (LRL) for most inorganic constituents. The LRL generally is twice the method detection level (MDL), which is described as the minimum concentration of a substance that can be measured and reported with 99-percent confidence that the analyte concentration is greater than zero (Childress and others, 1999). Analytical results that are less than a long-term method detection level (LT-MDL) are reported as less than the LRL. The LT-MDL differs from the MDL in that it incorporates laboratory variability and is calculated over an extended period of time. Analytical results that are greater than the LT-MDL and less than the LRL are reported as "estimated" values. For this report, the laboratory-estimated values were assumed to be the actual values.

Among the 66 percent of samples that were nondetects, a disproportionally large percentage were recorded for organic compounds. The average percentage of nondetects in each analyte group is as follows: 97 percent for organics,

5.1 percent for inorganics, 0.3 percent for radionuclides, and 0.1 percent for nutrients. For analytes measured for in water samples, the number of values that were recorded as nondetections are shown in table 2.

Interpreting Radionuclides on the Basis of Analytical Results

The randomness of the measurement process for radionuclides requires that the laboratory report the uncertainty associated with a single measurement. This measurement uncertainty is referred to as the combined standard uncertainty (CSU) and may be viewed as the statistical standard deviation of an individual radiological result (McCurdy and others, 2008, p. 3). Laboratories report the CSU at the 68 percent or 1-sigma confidence level (1σ CSU), which is obtained by propagating sources of analytical uncertainty in measurements. McCurdy and others (2008) provide details on interpreting radiological data used by the USGS. The guidelines for interpreting analytical results are based on an extension of a method proposed by Currie (1984).

The measured concentration and associated CSU are used to calculate a confidence interval (CI) about the measured concentration. The CI defines a range of concentration (the upper and lower limits) within which the "true value" lies with a certain degree of probability. For this report, a 95-percent confidence level is used. Assuming that the uncertainty of a radiological result is normally distributed, the 95-percent CI is defined by the measured concentration (C) plus or minus 1.96 times the associated σ CSU for that measurement; that is,

$$\left[C_{\ell}, C_{u}\right] = C \pm 1.96\,(1\sigma\mathrm{CSU}), \tag{1}$$

where

C_{ℓ} and C_{u} are the lower and upper limits of the CI, respectively.

The randomness of the measurement process for radionuclides makes negative values possible. A negative measured value, or negative lower limit on the CI, does not indicate that there is negative radioactivity.

Analysis and Interpretation of Replicate Data

The process of collecting and analyzing water samples from wells includes a number of steps that can affect how accurately samples represent the environment from which they were collected (Mueller and others, 2015, p. 1). Quality-control data collected from *replicate samples* are used to estimate the magnitude of errors in the reported analyte concentrations and inform the selection of target analytes to be included in the network optimization analysis.

Replicate samples are defined as two or more environmental samples collected sequentially in the same location. Replicate samples are used to measure the *variability* (as standard deviation) of constituent concentrations, which is defined as the random error in independent measurements as the result of repeated application of the measurement process under identical conditions (Mueller and others, 2015, p. 31). For many analytes, sampling variability is correlated with the concentration of that analyte, where variability increases with increasing concentration. A two-range model, as described by Mueller and Titus (2005) and Mueller and others (2015, p. 32–34), was used to evaluate the variability by estimating the standard deviation as a function of analyte concentration.

The two-range model is formulated using the mean concentration ($\bar{C}$) and standard deviation ($\bar{s}$) of replicate-paired data. For each of the n-replicate samples collected and analyzed for a constituent in the water-quality monitoring network, the mean replicate concentration is defined as:

$$\bar{C}_i = \frac{C_{a,i} - C_{b,i}}{2}, \quad \text{for } i = 1, \ldots, n, \tag{2}$$

where

$C_{a,i}$ and $C_{b,i}$ are the constituent concentrations measured for in the two environmental samples that compose replicate pair i.

The computation of $\bar{C}_i$ (eq. 2) is not possible if the concentration in either of the replicates is reported as a negative (only applies to radionuclides) or censored value. In this case, replicate pairs that include a negative or censored value were excluded from the analysis. For example, of the 106 sample-replicate pairs analyzed for carbon tetrachloride (table 2), only 21 were included in variability analysis (table 3). About 60 percent of carbon tetrachloride measurements were reported as censored values (table 2). The mean replicate concentrations are sorted from smallest to largest and expressed as $\bar{C}_{[i]}$, where the square brackets around the subscript indicate sorted values. The standard deviation of the i^{th} replicate pair is defined by:

$$s_i = \sqrt{\left(C_{a,\,[i]} - \bar{C}_{[i]}\right)^2 + \left(C_{b,\,[i]} - \bar{C}_{[i]}\right)^2}. \tag{3}$$

Replicate data are split into two subsets, a low concentration range and high concentration range. For concentrations in the low range, variability is estimated as the mean standard deviation of replicates within that range ($\bar{s}$), and expressed as:

$$\bar{s} = \frac{1}{n_1}\sum_{i=1}^{n_1} s_i. \tag{4}$$

where

n_1 is the number of replicates in the low range.

Table 2. Chemical constituents, organic compounds, and radionuclides measured for in water samples collected from wells in the U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.

[**Analyte name:** U.S. Environmental Protection Agency Substance Registry Services (SRS) systematic name. **Code:** Unique identifier assigned by the U.S. Geological Survey. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of sampling sites:** Number of locations where groundwater samples were collected and analyzed for the constituent. **Number of observations:** Number of laboratory measurements of a constituent concentration. **Number of nondetects:** Number of observations with left-censored values, where the real concentration is known to be less than a detection threshold. **Number of replicate pairs:** Number of sample-replicate pairs analyzed for a constituent]

Analyte name	Code	Period of record (mm-dd-yy)	Number of sampling sites	Number of observations	Number of nondetects	Number of replicate pairs
Analyte group—Inorganics, Major, Metals						
Sodium	00930	01-25-89 to 10-22-18	152	4,440	1	395
Analyte group—Inorganics, Major, Non-metals						
Chloride	00940	01-04-89 to 10-23-18	153	6,372	0	548
Sulfate	00945	02-15-89 to 10-22-18	147	2,973	1	237
Fluoride	00950	02-15-89 to 10-01-18	136	609	45	36
Analyte group—Inorganics, Minor, Metals						
Chromium	01030	01-25-89 to 10-18-18	139	2,866	485	256
Analyte group—Nutrient						
Nitrate	00618	02-15-89 to 10-23-18	153	4,172	6	353
Analyte group—Organics, other						
Dibromomethane	30217	02-19-92 to 12-12-18	70	1,352	1,351	90
Dichlorobromomethane	32101	01-04-89 to 12-12-18	119	1,663	1,646	99
Carbon tetrachloride	32102	01-04-89 to 12-12-18	119	1,723	1,029	106
Tribromomethane	32104	01-04-89 to 12-12-18	119	1,663	1,635	99
Chlorodibromomethane	32105	01-04-89 to 12-12-18	119	1,663	1,648	99
Chloroform	32106	01-04-89 to 12-12-18	119	1,663	1,096	99
Toluene	34010	01-04-89 to 12-12-18	119	1,723	1,672	106
Benzene	34030	01-04-89 to 12-12-18	119	1,723	1,722	106
Chlorobenzene	34301	01-04-89 to 12-12-18	119	1,699	1,699	103
Chloroethane	34311	01-04-89 to 12-12-18	109	1,608	1,608	96
Ethylbenzene	34371	01-04-89 to 12-12-18	119	1,723	1,720	106
Chloromethane	34418	01-04-89 to 12-12-18	119	1,642	1,642	99
Methylene chloride	34423	01-04-89 to 12-12-18	119	1,723	1,721	106
Tetrachloroethylene	34475	01-04-89 to 12-12-18	119	1,723	1,375	106
CFC-11	34488	01-04-89 to 12-12-18	109	1,689	1,689	103
1,1-Dichloroethane	34496	01-04-89 to 12-12-18	119	1,723	1,721	106
1,1-Dichloroethylene	34501	01-04-89 to 12-12-18	119	1,723	1,669	106
1,1,1-Trichloroethane	34506	01-04-89 to 12-12-18	119	1,723	981	106
1,1,2-Trichloroethane	34511	01-04-89 to 12-12-18	119	1,678	1,678	103
1,1,2,2-Tetrachloroethane	34516	01-04-89 to 12-12-18	109	1,668	1,668	103
o-Dichlorobenzene	34536	01-04-89 to 12-12-18	119	1,703	1,703	99
trans-1,2-Dichloroethylene	34546	01-04-89 to 12-12-18	119	1,723	1,723	106
1,2,4-Trichlorobenzene	34551	03-23-89 to 12-12-18	111	1,496	1,495	93
m-Dichlorobenzene	34566	01-04-89 to 12-12-18	109	1,669	1,669	96
CFC-12	34668	01-04-89 to 12-12-18	109	1,629	1,529	96
Naphthalene	34696	03-23-89 to 12-12-18	111	1,558	1,554	100

Table 2. Chemical constituents, organic compounds, and radionuclides measured for in water samples collected from wells in the U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.—Continued

[**Analyte name:** U.S. Environmental Protection Agency Substance Registry Services (SRS) systematic name. **Code:** Unique identifier assigned by the U.S. Geological Survey. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of sampling sites:** Number of locations where groundwater samples were collected and analyzed for the constituent. **Number of observations:** Number of laboratory measurements of a constituent concentration. **Number of nondetects:** Number of observations with left-censored values, where the real concentration is known to be less than a detection threshold. **Number of replicate pairs:** Number of sample-replicate pairs analyzed for a constituent]

Analyte name	Code	Period of record (mm-dd-yy)	Number of sampling sites	Number of observations	Number of nondetects	Number of replicate pairs
Analyte group—Organics, other—Continued						
Vinyl chloride	39175	01-04-89 to 12-12-18	119	1,723	1,723	106
Trichloroethylene	39180	01-04-89 to 12-12-18	119	1,723	1,082	106
Hexachlorobutadiene	39702	03-23-89 to 12-12-18	95	1,462	1,461	90
cis-1,2-Dichloroethylene	77093	02-19-92 to 12-12-18	90	1,467	1,467	100
Styrene	77128	01-04-89 to 12-12-18	119	1,663	1,657	99
1,1-Dichloropropene	77168	02-19-92 to 12-12-18	69	1,352	1,352	90
2,2-Dichloropropane	77170	02-19-92 to 12-12-18	69	1,352	1,352	90
1,2,4-Trimethylbenzene	77222	04-09-92 to 12-12-18	89	1,384	1,381	93
Cumene	77223	04-09-92 to 12-12-18	69	1,350	1,349	90
n-Propylbenzene	77224	04-09-92 to 12-12-18	89	1,384	1,384	93
1,3,5-Trimethylbenzene	77226	04-09-92 to 12-12-18	69	1,350	1,350	90
o-Chlorotoluene	77275	02-19-92 to 12-12-18	69	1,352	1,352	90
p-Chlorotoluene	77277	02-19-92 to 12-12-18	69	1,352	1,352	90
Halon 1011	77297	05-19-93 to 12-12-18	87	1,332	1,332	92
n-Butylbenzene	77342	04-09-92 to 12-12-18	69	1,350	1,350	90
sec-Butylbenzene	77350	04-09-92 to 12-12-18	89	1,384	1,384	93
tert-Butylbenzene	77353	04-09-92 to 12-12-18	69	1,350	1,350	90
p-Cymene	77356	04-09-92 to 12-12-18	69	1,350	1,350	90
1,1,1,2-Tetrachloroethane	77562	02-19-92 to 12-12-18	90	1,387	1,387	93
1,2,3-Trichlorobenzene	77613	04-09-92 to 12-12-18	69	1,350	1,349	90
CFC-113	77652	05-19-93 to 12-12-18	64	1,306	1,306	87
Methyl tert-butyl ether	78032	05-19-93 to 12-12-18	87	1,352	1,352	92
Xylene	81551	01-04-89 to 12-12-18	109	1,687	1,679	102
Bromobenzene	81555	02-19-92 to 12-12-18	69	1,352	1,352	90
TTHM4	90867	01-04-89 to 12-12-18	119	1,663	1,649	99
Analyte group—Organics, pesticide						
1,2-Dichloroethane	32103	01-04-89 to 12-12-18	119	1,723	1,723	106
Acrylonitrile	34215	02-19-92 to 12-12-18	67	1,072	1,072	70
Methyl bromide	34413	01-04-89 to 12-12-18	119	1,642	1,642	99
1,2-Dichloropropane	34541	01-04-89 to 12-12-18	109	1,629	1,629	96
p-Dichlorobenzene	34571	01-04-89 to 12-12-18	119	1,703	1,702	99
trans-1,3-Dichloropropene	34699	01-04-89 to 12-12-18	119	1,642	1,642	99
cis-1,3-Dichloropropene	34704	01-04-89 to 12-12-18	119	1,642	1,642	99
1,3-Dichloropropane	77173	02-19-92 to 12-12-18	69	1,352	1,352	90
1,2,3-Trichloropropane	77443	02-19-92 to 12-12-18	70	1,353	1,353	90
Ethylene dibromide	77651	01-04-89 to 12-12-18	109	1,608	1,608	96

Table 2. Chemical constituents, organic compounds, and radionuclides measured for in water samples collected from wells in the U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.—Continued

[**Analyte name:** U.S. Environmental Protection Agency Substance Registry Services (SRS) systematic name. **Code:** Unique identifier assigned by the U.S. Geological Survey. **Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of sampling sites:** Number of locations where groundwater samples were collected and analyzed for the constituent. **Number of observations:** Number of laboratory measurements of a constituent concentration. **Number of nondetects:** Number of observations with left-censored values, where the real concentration is known to be less than a detection threshold. **Number of replicate pairs:** Number of sample-replicate pairs analyzed for a constituent]

Analyte name	Code	Period of record (mm-dd-yy)	Number of sampling sites	Number of observations	Number of nondetects	Number of replicate pairs
Analyte group—Organics, pesticide—Continued						
1,2-Dibromo-3-chloropropane	82625	02-19-92 to 12-12-18	70	1,353	1,353	90
Analyte group—Radiochemical						
Tritium	07000	01-04-89 to 10-23-18	153	6,384	43	541
Strontium-90	13501	01-04-89 to 10-23-18	138	3,647	22	303
Plutonium-238	22012	01-04-89 to 10-22-18	118	962	0	70
Cesium-137	28401	01-04-89 to 10-22-18	137	2,762	14	241
Alpha particle	63018	04-07-08 to 10-18-18	77	749	0	53
Beta particle	80049	04-07-08 to 10-18-18	77	749	0	53

Table 3. Variability models for selected analytes estimated from replicate-paired sample data, U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.

[**Period of record:** For water samples collected during 1989–2018, in mm-dd-yy (month-day-year). **Number of sampling sites:** Number of sites where one or more replicate samples were collected. **Number of replicate pairs:** Number of sample-replicate pairs analyzed for the constituent. For replicate pairs in which one or both values of the pair were censored, the pair was excluded from the variability analysis. **Mean SD:** Mean standard deviation of replicates. **Mean RSD:** Percent mean relative standard deviation of replicates. **Abbreviations and symbols:** mg/L, milligrams per liter; µg/L, micrograms per liter; pCi/L, picocuries per liter; N, Nitrogen; <, less than; ≥, greater than or equal to; ∞, infinity]

Analyte and units	Period of record (mm-dd-yy)	Number of sampling sites	Concentration range	Number of replicate pairs	Mean SD	Mean RSD (percent)
Sodium, mg/L	07-23-90 to 10-11-18	140	<60	366	0.204	1.5
			≥60	6	1.010	1.5
Chloride, mg/L	01-02-90 to 10-11-18	144	<42	420	0.222	1.5
			≥42	87	1.650	2.0
Sulfate, mg/L	10-01-90 to 10-11-18	99	<47	227	0.149	0.5
			≥47	6	1.640	1.3
Fluoride, mg/L	10-01-90 to 10-03-17	21	-∞ to ∞	31	0.008	4.1
Chromium, µg/L	10-01-90 to 10-11-18	90	<24	188	0.443	6.3
			≥24	6	3.090	2.2
Nitrate, mg/L as N	09-24-90 to 10-11-18	132	<4.9	328	0.016	1.3
			≥4.9	6	1.960	16.4
Carbon tetrachloride, µg/L	10-30-90 to 10-12-16	6	-∞ to ∞	21	0.043	1.7
Chloroform, µg/L	07-16-91 to 10-12-16	5	<0.7	9	0.003	0.9
			≥0.7	6	0.012	1.0
Toluene, µg/L	01-13-99 to 04-17-13	4	-∞ to ∞	5	0.040	13.7
Tetrachloroethylene, µg/L	01-13-99 to 06-28-17	6	-∞ to ∞	10	0.021	6.5
1,1-Dichloroethylene, µg/L	10-17-96 to 10-03-17	4	-∞ to ∞	5	0.001	2.1
1,1,1-Trichloroethane, µg/L	10-26-90 to 10-12-16	14	-∞ to ∞	31	0.007	3.6
CFC-12, µg/L	04-28-98 to 04-17-13	4	-∞ to ∞	6	0.008	5.5
Trichloroethylene, µg/L	10-30-90 to 10-12-16	7	<2.2	14	0.004	1.1
			≥2.2	7	0.038	1.4
Tritium, pCi/L	01-02-90 to 10-04-18	117	<10,200	273	83.300	22.6
			≥10,200	42	392.000	2.4
Strontium-90, pCi/L	01-02-90 to 04-17-18	74	<38	164	0.792	36.8
			≥38	6	9.890	18.7
Plutonium-238, pCi/L	10-30-90 to 10-08-15	16	<0.01	12	0.002	48.1
			≥0.01	8	0.008	65.4
Cesium-137, pCi/L	01-05-90 to 04-11-18	70	<25	83	7.840	58.7
			≥25	34	17.400	55.1
Alpha particle, pCi/L	04-07-08 to 04-11-18	39	<4.3	32	1.220	65.5
			≥4.3	7	2.980	55.4
Beta particle, pCi/L	04-07-08 to 04-11-18	48	<6.5	45	0.711	29.1
			≥6.5	6	1.730	18.4

In the high concentration range, the two-range model assumes that the standard deviation linearly increases with concentration; the line of estimated standard deviation is described by a slope equal to the mean relative standard deviation ($\bar{s}_r$) and intercept equal to zero. The $\bar{s}_r$ is calculated from replicated data in the high concentration range and expressed as:

$$\bar{s}_r = \frac{1}{n - n_1} \sum_{i=n_1+1}^{n} \frac{s_i}{\bar{C}_{[i]}}, \tag{5}$$

where

n is the total number of replicates.

The two-range model describing the standard deviation as a function of constituent concentration is defined by:

$$s(C) = \begin{cases} \bar{s}, & \text{for } C < \bar{C}_{[k]} \\ \bar{s}_r C, & \text{for } C \geq \bar{C}_{[k]} \end{cases}, \tag{6}$$

where

$\bar{C}_{[k]}$ is the mean concentration of replicate pair k and represents the boundary between the low-range and high-range concentrations.

Separating replicate data into low and high ranges of concentration was approached by identifying an abrupt structural change (breakpoint) in the variance (s^2, the square of the standard deviation defined by eq. 3) of replicates sorted by their average concentration; the individual variances within each concentration range should be about equal. An optimal breakpoint was calculated using the algorithm described by Bai and Perron (2003). Given n data points, the algorithm generates a piecewise constant sequence of line segments for each data subset. End points of the replicate-data subsets are defined in terms of their index position within the sorted mean replicate concentrations, and are expressed as $\{k_0, k_1, k_2\}=\{1, k, n\}$. Each line segment covers a concentration range and represents those data points by a constant value of variance. Consider the following two-segment piece wise constant function:

$$s_i^2 = \alpha_j + e_i \quad \text{for } i = k_{j-1}, \ldots, k_j \text{ and } j = 1, 2, \tag{7}$$

where

α_j is the constant value of variance within concentration range j, and

e_i is the error at replicate pair i.

The constant values of variance and breakpoint k_1 are estimated by minimizing the total fit error, which is stated in the following optimization formulation:

$$k_1 = \underset{\substack{k_1 \in \mathbb{Z} \\ \alpha_1, \alpha_2 \in \mathbb{R}}}{\arg\min} \sum_{j=1}^{2} \sum_{i=k_{j-1}}^{k_j} \left(\alpha_j - s_i^2\right)^2, \quad \text{subject to: } k_{j-1} - k_j \geq n_{\min} \tag{8}$$

where

$\mathbb{Z}$ is the set integers,

$\alpha_1, \alpha_2 \in \mathbb{R}$ indicates that variance values are in the set of real numbers, and

$n_{\min}$ is the minimum number of replicate pairs in a concentration range.

For this study, a concentration range was required to have at least five replicate pairs. Once a breakpoint k_1 has been determined, it is used to identify the boundary between low-range and high-range concentrations; that is, $\bar{C}_{[k]}$ in equation 6. For some constituents, no breakpoint could be determined because the number of replicate pairs was too small to separate into low and high ranges of concentration, so $\bar{C}_{[k]}$ was set equal to zero; that is, variability was estimated as the measured concentration scaled by the mean relative standard deviation.

Two examples of the development of the two-range model using tritium and chromium concentrations are shown in figure 4. Replicate data were plotted as points and the standard deviations for all ranges of concentrations were plotted as lines. For tritium (fig. 4*A*), the variability of concentrations was estimated using 315 replicates collected from 117 wells. A boundary concentration of 10,200 pCi/L separates the concentration ranges, with replicate variances about equal within each range of concentration. The mean standard deviation in the low range of concentrations (83 pCi/L) is smaller than in the high range of concentrations (392 pCi/L). Standard deviation is defined with more accuracy in the low range of concentrations because of the numerous replicates in this range (n=273). In comparison, the number of replicates in the high range of concentrations is relatively small (n=42). Variability of tritium was estimated at 83 pCi/L in the low range of concentrations, and 2.4 percent of the measured concentration in the high range (table 3).

Variability of chromium concentrations (fig. 4*B*) was estimated using 194 replicates collected from 85 wells. A boundary concentration of 24 µg/L separates the concentration ranges. The mean standard deviation in the low range of concentrations (0.443 µg/L) is smaller than in the high range of concentrations (3.090 µg/L) (table 3). Standard deviation is defined with more accuracy in the low range of concentrations because of the large number of replicates in this range (n=188). In comparison, the number of replicates in the high range of concentrations is sparse (n=6) and indicates a poorly defined model in this range. Variability of chromium was estimated at 0.443 µg/L in the low range of concentrations, and 2.2 percent of the measured concentration in the high range (table 3).

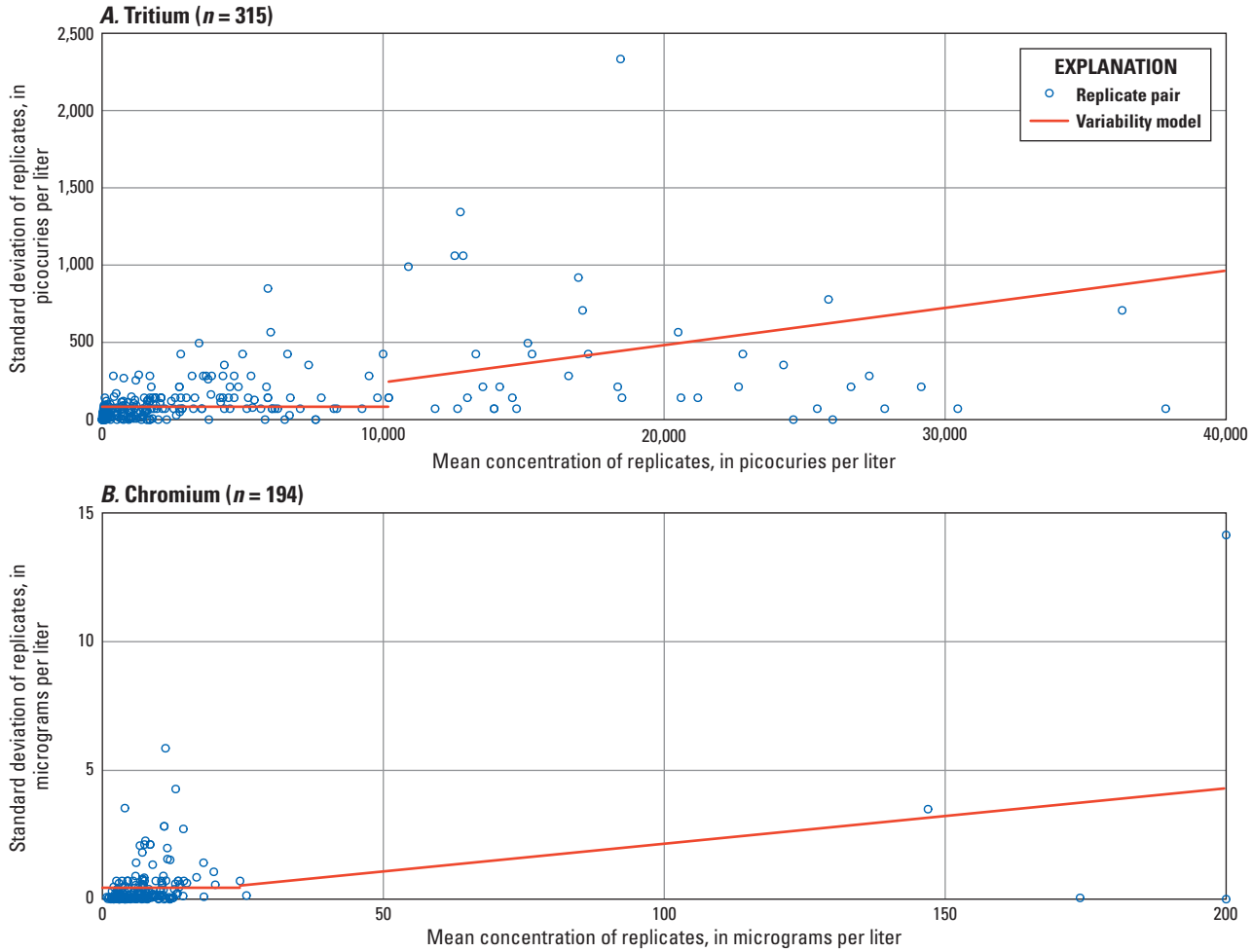


Figure 4. Two-range model fitted to measurements of (*A*) tritium and (*B*) chromium in replicate-paired water samples collected from wells in the Idaho National Laboratory water-quality aquifer monitoring network, 1989–2018. *n*, number of replicates.

Two-range models for each of the selected constituents measured for in replicate samples collected from wells in the INL water-quality network, 1989–2018, are described in table 3 and shown in appendix 3. For sodium, chloride, sulfate, chromium, nitrate, chloroform, trichloroethylene, tritium, strontium-90, plutonium-238, cesium-137, alpha particle, and beta particle, the two-range model is defined by a piece wise constant function with two segments. These two-range models produce the smallest estimates of standard deviation within the low range of concentrations and the largest estimates within the high range of concentrations. For fluoride, carbon tetrachloride, toluene, tetrachloroethylene, 1,1-dichloroethylene, and 1,1,1-trichloroethane, the number of replicate pairs typically was too small to separate into low and high ranges of concentration. The variability of these pairs was modeled as a linear function. For all other analytes, insufficient uncensored-replicate data (less than five replicate pairs) were available to establish a variability model. In this report, the absence of a variability model was the criterion used to exclude an analyte from further analysis. That is, analytes were excluded from analysis when there were insufficient quality-control data to evaluate the variability in analyte concentration.

Background Levels, Summary Statistics, and Maximum Contaminant Levels

Understanding the statistical behavior of pooled measurements of concentration in groundwater is necessary for evaluating the monitoring network. Summary statistics computed from the sample data were used to characterize the behavior of constituent measurements pooled together over space and time. Fixed regularity standards are given for each constituent to provide a point of reference for evaluating the value of an observation. For example, water-quality constituents detected well above background levels are among the most important to monitor because they may indicate groundwater contamination. Sampling sites that exceeded a drinking water maximum contaminant level for a constituent require long-term monitoring to inform the public of potential health risks.

Background concentrations are defined as groundwater influenced by western tributary recharge in the western INL and by eastern regional recharge in the eastern INL. Background concentrations are either naturally occurring or anthropogenic (substances present in the environment as a result of human activities) and are not influenced by waste and wastewater disposal at the INL. Background concentrations in groundwater at or near the INL were compiled from multiple USGS studies going back to 1989 and shown in table 4 for selected analytes.

Summary statistics of selected analytes during 1989–2018 are shown in table 4, with radionuclides represented as a range of potential concentrations defined by the 95-percent confidence interval of the measurement (eq. 1). The product limit method (Kaplan-Meier estimator) from survival analysis (Kaplan and Meier, 1958; She, 1997) was applied to the censored and uncensored concentrations to estimate the 25th percentile (lower quartile), median, mean, and 75th percentile (upper quartile). The uncertainty in the mean value was quantified using the standard error of the mean statistic (SEM) and expressed in table 4 as the measured concentration plus and minus the SEM. The minimum and maximum values were considered as the minimum and maximum concentration, with radionuclide concentrations expressed in table 4 as the measured value plus or minus the 1σ CSU.

The summary statistics for most analytes describe a concentration distribution that is positively skewed with the lower bound near zero and a span of two or more orders of magnitude. Great variation in concentrations and high skewness is typical of locally extreme values surrounded by much smaller values, such as in a groundwater contaminant plume. Waste-disposal practices at the INL have created contaminant plumes of radiochemical and chemical constituents in the ESRP aquifer. Agriculture and other anthropogenic influences upgradient from the INL also have created contaminant plumes in the aquifer (Bartholomay and others, 2017).

The maximum contaminant levels (MCLs), standards set by the U.S. Environmental Protection Agency for drinking water quality, are shown for selected analytes in table 4. A missing MCL value indicates that there is no enforceable standard established for the analyte. The MCL was exceeded one or more times during 1989–2018 for the following analytes: chromium, nitrate, carbon tetrachloride, tritium, strontium-90, and alpha particle. Upper quartile concentrations however, never exceeded the MCL.

Table 5 shows the number of sampling sites, and percentage of the total number of sites, where an analyte was measured one or more times at a concentration greater than the laboratory reporting level (LRL), the upper limit of estimated background concentrations, and the MCL. Exceedance of the LRL indicates that one or more of the measured values were detected. Recall that nondetects (left-censored values) are reported as below the LRL, and the LRL may vary from sample to sample for the same analyte and the same analytical method. Carbon tetrachloride, chloroform, toluene, tetrachloroethylene, and trichloroethylene were recorded at detectable concentrations in less than 10 percent of the 153 sampling sites, thus indicating limited informative data on the variability of these analytes in most of the sites in the monitoring network. The large number of nondetections for these analytes (table 2) and sparsity of sampling sites with detectable concentrations (table 5) can hinder their analysis.

The number of sampling sites where analytes were detected one or more times at concentrations above the upper limit of background ranged from as few as 5 of the 153 sampling sites for 1,1-Dichloroethylene, to as many as 131 sites for tritium (table 5), whereby the number of sites implicitly indicates the historical extent of analyte contamination in the aquifer resulting from waste and wastewater disposal at the INL. For chloride, sulfate, chromium, tritium, strontium-90, and cesium-137, the groundwater affected by contamination is widespread, with concentrations exceeding background at more than 100 sampling sites (or more than two-thirds of the total number of sites in the network). Analytes with very few instances of background exceedance (such as fluoride, carbon tetrachloride, chloroform, tetrachloroethylene, 1,1-Dichloroethylene, trichloroethylene, and beta particle, with less than 20 occurrences each) can bias interpolation estimates because contaminant concentration fields are highly heterogeneous, with locally extreme values near the waste-disposal sites surrounded by much smaller background values.

Measured concentrations for most analytes did not exceed their MCL for drinking water quality (tables 4 and 5).The number of sampling sites where analytes were detected at concentrations greater than the regulatory limit was relatively small for chromium (1 site), alpha particle (1 site), nitrate (2 sites), and carbon tetrachloride (4 sites); in comparison, tritium (19 sites) and strontium-90 (21 sites) each exceeded the MCL in more than 10 percent of the total sampling sites. Sites with unsafe levels of contaminants at any point in time are important because they may be used to trace a pollutant to its source and evaluate the effectiveness of cleanup efforts at the INL.

Table 4. Background concentrations, summary statistics, and maximum contaminant level for selected analytes measured for in water samples collected from wells in the U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.

[**Period of record background range:** Range of background concentrations in the eastern Snake River Plain aquifer during the period of record, where background concentrations are either naturally occurring or anthropogenic, and are not influenced by waste and wastewater disposal at the Idaho National Laboratory. **Summary statistics:** Radionuclides are interpreted as a range of potential concentrations defined by the 95-percent confidence interval of the measurements. Minimum and maximum radionuclide concentrations are plus or minus 1-sigma combined standard uncertainty. Mean concentrations are plus or minus the standard error. **MCL:** Maximum contaminant level (MCL), a standard set by the U.S. Environmental Protection Agency for drinking water quality. **Abbreviations and symbols:** mg/L, milligrams per liter; N, Nitrogen; µg/L, micrograms per liter; pCi/L, picocuries per liter; 25 percent, lower quartile; 75 percent, upper quartile; <, less than; ±, plus and minus; –, MCL has not been established for the analyte]

Analyte and units	Period of record background range	Summary statistics						MCL
		Minimum	25 percent	Median	Mean	75 percent	Maximum	
Sodium, mg/L	[1]8.3–14.8	1	8.85	11.3	15.1±0.316	16.6	98.7	–
Chloride, mg/L	[1]11.8–14.2	1	12.1	16.7	31.9±0.925	30	240	–
Sulfate, mg/L	[1]20.2–21.4	9.5	22.3	26	29.6±0.639	32	176	–
Fluoride, mg/L	[4]0.4–0.5	<0.1	0.194	0.2	0.222±0.011	0.228	1.4	4
Chromium, µg/L	[1]3–4	0.112	3.83	6.35	10.7±0.9	9.43	230	100
Nitrate, mg/L as N	[1]0.655–1	0.035	0.699	0.99	1.39±0.035	1.64	20.6	10
Carbon tetrachloride, µg/L	[4]<0.2	0	0	0	1.49±0.115	2.1	11.6	5
Chloroform, µg/L	[4]<0.1	0.014	0.014	0.018	0.354±0.031	0.4	6.2	80
Toluene, µg/L	[4]<0.1	<0.02	0.1	0.1	0.157±0.072	0.1	63.6	1,000
Tetrachloroethylene, µg/L	[4]<0.1	0.013	0.013	0.013	0.088±0.007	0.141	3.87	5
1,1-Dichloroethylene, µg/L	[4]<0.1	0.011	0.014	0.019	0.032±0.002	0.032	1	7
1,1,1-Trichloroethane, µg/L	[4]<0.1	0.013	0.041	0.114	0.195±0.009	0.314	1	200
CFC-12, µg/L	[4]<0.2	0.1	0.121	0.155	0.163±0.005	0.187	2.6	–
Trichloroethylene, µg/L	[4]<0.1	0.011	0.013	0.022	0.704±0.056	0.8	4.88	5
Tritium, pCi/L	[3]<150	-660±216	15.1	67	2,803±157	1,975	59,500±2,352	20,000
Strontium-90, pCi/L	[4]0	-8±5.88	0.234	0.534	4.25±0.286	3.73	100±196	8
Plutonium-238, pCi/L	[2]0	-0.07±0.059	-0.002	-0.002	-0.001	-0.002	0.1±0.078	–
Cesium-137, pCi/L	[2]0	-138±196	5.48	8.58	7.78±0.26	11.7	90±196	–
Alpha particle, pCi/L	[2]<3	-8±5.88	1.49	1.87	1.71±0.094	2.29	18±7.84	15
Beta particle, pCi/L	[2]<7	-1.8±1.76	2.25	3.05	3.27±0.164	3.45	21.6±3.53	50

[1]From Bartholomay and Hall (2016).

[2]From Knobel and others (1992).

[3]From Michel (1989).

[4]From Orr and others (1991).

Table 5. Number of sampling sites where an analyte was measured one or more times at a concentration greater than the detection threshold, upper limit of background concentrations, and maximum contaminant level, U.S. Geological Survey aquifer water-quality monitoring network, Idaho National Laboratory and vicinity, Idaho, 1989–2018.

[Percentage of the total number of sampling sites in the monitoring network is italicized in parentheses. **Abbreviations and symbol:** LRL, laboratory reporting level; MCL, maximum contaminant level; –, MCL has not been established for the analyte]

Analyte name	Exceeded LRL	Exceeded background	Exceeded MCL
Sodium	151 (*99*)	64 (*42*)	–
Chloride	153 (*100*)	120 (*78*)	–
Sulfate	146 (*95*)	130 (*85*)	–
Fluoride	99 (*65*)	11 (*7*)	0 (*0*)
Chromium	74 (*48*)	127 (*83*)	1 (*1*)
Nitrate	148 (*97*)	94 (*61*)	2 (*1*)
Carbon tetrachloride	10 (*7*)	11 (*7*)	4 (*3*)
Chloroform	4 (*3*)	15 (*10*)	0 (*0*)
Toluene	2 (*1*)	21 (*14*)	0 (*0*)
Tetrachloroethylene	5 (*3*)	10 (*7*)	0 (*0*)
1,1-Dichloroethylene	9 (*6*)	5 (*3*)	0 (*0*)
1,1,1-Trichloroethane	24 (*16*)	30 (*20*)	0 (*0*)
CFC-12	1 (*1*)	6 (*4*)	–
Trichloroethylene	10 (*7*)	9 (*6*)	0 (*0*)
Tritium	127 (*83*)	131 (*86*)	19 (*12*)
Strontium-90	131 (*86*)	123 (*80*)	21 (*14*)
Plutonium-238	118 (*77*)	71 (*46*)	–
Cesium-137	133 (*87*)	127 (*83*)	–
Alpha particle	77 (*50*)	67 (*44*)	1 (*1*)
Beta particle	77 (*50*)	17 (*11*)	0 (*0*)

Classification of Time-Series Data

The time-series analysis described in this report (such as temporal regression and iterative thinning) was constrained by the scarcity of time-series data on analyte concentrations. A sufficient amount of continuous reliable measurements is needed to capture both the natural and human-induced changes in water quality over time. To this end, strict inclusion-exclusion criteria were defined to identify relevant data for time-series analysis. Water-quality data were classified by whether an observation was part of a *continuous record block* (Type-1 data), or not (Type-2 data). A continuous record block is assumed to be suitable for analysis and is defined as the longest period in the time series with (1) at least 15 observations, (2) a 5-year maximum time span between consecutive observations, and (3) a record period that is as least 15 years in duration. Three examples showing the classification of time-series data are presented in figure 5. For tritium in well USGS 20 (fig. 5*A*) and chromium in well USGS 38 (fig. 5*B*), all data reside in a continuous record block and, therefore, are classified as Type-1 data. By contrast, fluoride data in well USGS 15 (fig. 5*C*) are classified as Type-2 data because observations there span a 6-year period (1990–1995), well below the 15-year minimum duration for a continuous record block.

Time-series graphs with Type-1 and Type-2 data classification for selected constituents are shown in appendix 4. Of the 19 selected constituents analyzed for in water samples collected from as many as 153 different sites during 1989–2018, there were a total of 2,413 time series, with 940 (39 percent) composed entirely of Type-1 data, 1,450 (60 percent) composed entirely of Type-2 data, and 23 (1 percent) composed of both Type-1 and Type-2 data. Recall that Type-2 data were excluded from time-series analysis.

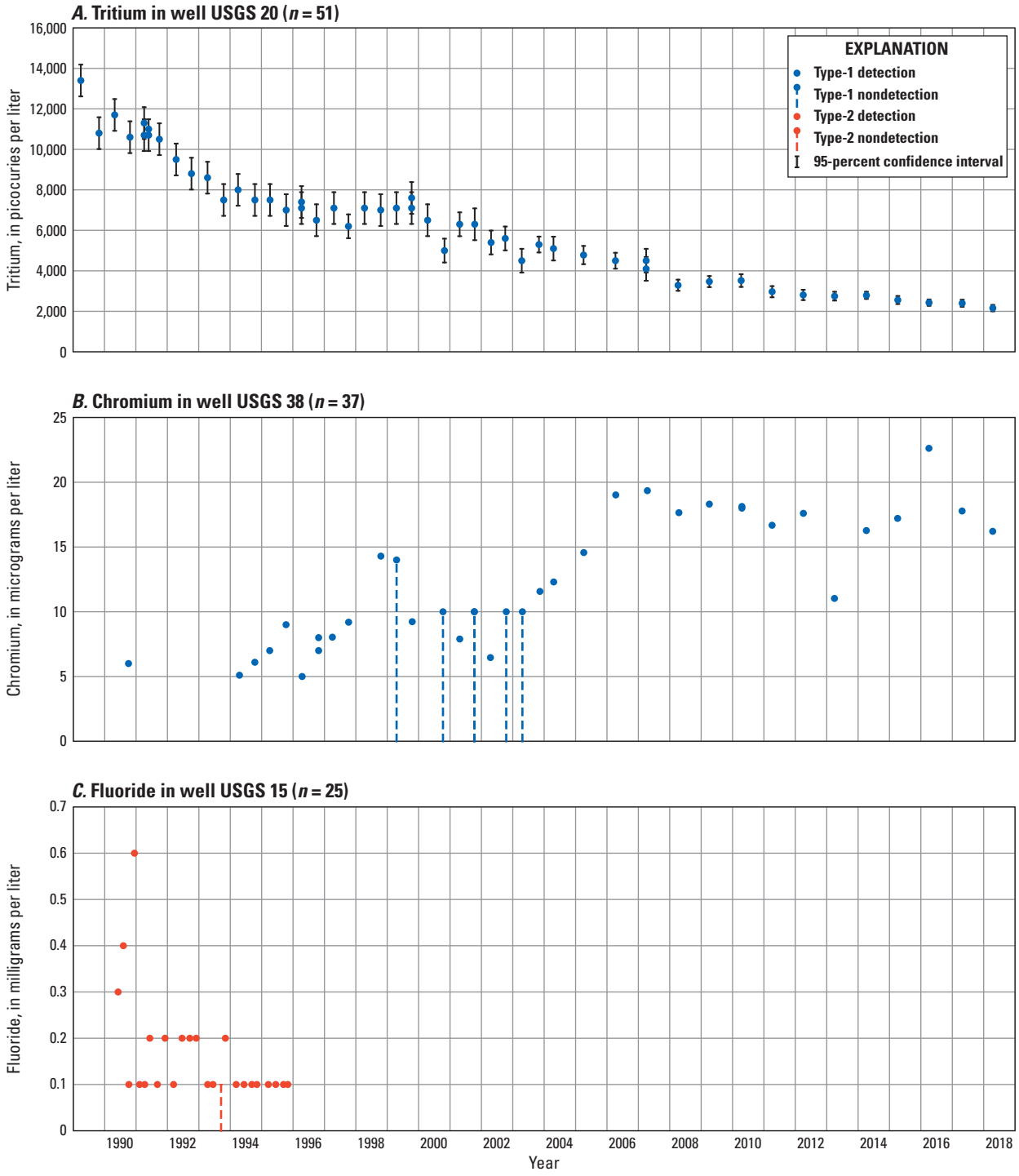


Figure 5. Concentrations of (*A*) tritium in well USGS 20, (*B*) chromium in well USGS 38, and (*C*) fluoride in well USGS 15, Idaho National Laboratory, Idaho, 1989–2018. *n*, number of observations.

Temporal Regression

Inherent in the collection of water-quality data over time is some form of random variation. There are many sources of variation that can affect changes in concentration, such as seasonal variability of groundwater chemistry resulting from infiltrating water reaching the water table, or variability introduced during sampling, preservation, and handling. Regression analysis is used to exclude (or reduce) this random variation and show any underlying trend in the data. For this report, two types of regression techniques were used to model concentration as a function of time for each constituent-well pair: (1) nonparametric local regression, and (2) parametric survival regression. Nonparametric regression differs from parametric regression in that it does not make any underlying assumptions about the distribution of the data.

The two regression techniques complement each other because they each have their own distinctive strengths and weaknesses, whereby the weakness of one technique typically is offset by the strength of the other. Strengths of nonparametric local regression include (1) fitting procedures that are very flexible and well suited for modeling complex structures within the data, and (2) a relative insensitivity to discrepant observations. Weaknesses include (1) results in a fitted model that cannot be expressed as a simple equation, (2) an inability to account for censored values (such as nondetection values), and (3) the use of a partially arbitrary *hyperparameter* to specify the desired degree of smoothness. A hyperparameter is a configuration variable that is external to the model and cannot be directly estimated from the data (Brownlee, 2017).

Strengths of the survival regression technique include (1) results in a fitted model that describes the structure of the underlying data in an easily understood equation (long-term trends may be calculated from this equation), and (2) the ability to account for censored data. Weaknesses include (1) the required specification of a baseline distribution for the response variable (such as a log-normal distribution for constituent concentrations); and (2) relative sensitivity to outliers, discrepant observations, and patterns that may only exist within a limited subset of the data (Jacoby, 2000, p. 609–608). Survival regression also is not flexible enough to account for non-monotonic representation of change. Long-term water-quality trends typically occur in two ways: (1) as a gradual change in concentration over time that is consistently in one direction (monotonic), or (2) as an abrupt change in concentration at a specific point in time (non-monotonic).

Local Regression Analysis

Local regression is an exploratory data-analysis technique that is useful for discovering various characteristics of a time series, such as long-term trend and seasonal components. The general shape of a time series is made apparent by reducing the background variability or "noise," such as the variability introduced to sample measurements because of field procedures and laboratory analysis. The local regression method *LOESS* was used to smooth the concentration time-series data of a constituent measured for in water samples collected from a monitoring well. In the context of the network optimization problem, LOESS also was used to formulate the objective functions for the spatial and temporal components of optimization.

LOESS, originally proposed by Cleveland (1979) and further developed by Cleveland and Devlin (1988), Cleveland and Grosse (1991), and Cleveland and others (1992a), is a nonparametric regression method (that is, it does not make any underlying assumptions about the distribution of the data) and implements a robust locally weighted regression procedure for fitting smooth functions to empirical data points. Low-degree polynomial curves are fitted to localized subsets of the empirical data to develop a smooth function (also known as a "loess curve") that describes the deterministic part (no random elements) of the variation in the observed data, point by point (Freeman and others, 2008, p. 50). A loess curve is formulated here to describe the relation between time and constituent concentration in single well.

Assume that for the i^{th} observation in an unevenly spaced time series, the measured constituent concentration (C_i) and corresponding sample time (t_i) are related by

$$C_i = g(t_i) + e_i, \tag{9}$$

where

g is a deterministic smooth function, and
e_i is the random error of observation i.

Letting $\hat{C}_i$ be an estimate of g at sample time t_i, equation 9 is expressed for all observations as:

$$C_i = \hat{C}_i + e_i \quad \text{for } i = 1,\ldots,n, \tag{10}$$

where

$\hat{C}$ is the set of data points that describes the loess curve, and
n is the number of observations in the time series.

The first step in the LOESS procedure is to define m equally spaced times across the period of record. Denote these times as $\hat{t}_j$, where the subscript j ranges from 1 to m. The loess curve will be predicted at each $\hat{t}_j$; therefore, the temporal resolution of the loess curve is dependent on m. In this study, prediction points were equally spaced at monthly time intervals over the period of record (1989–2018).

For each $\hat{t}_j$, LOESS performs a robust locally weighted linear regression analysis. These regressions are "local" in the sense that each one only uses the subset of observations that fall closest to the prediction point (these are observations in the neighborhood of $\hat{t}_j$) (Jacoby, 2000, p. 583). The proximity of observations to a predication point j is quantified using their time difference, defined as:

$$\Delta_i(\hat{t}_j) = \left| t_i - \hat{t}_j \right| \quad \text{for } i = 1, \ldots, n, \tag{11}$$

where

$\Delta_i(\hat{t}_j)$ is the absolute time difference between observation i and prediction point j;

t_i is the sampling time for observation i; and

$\hat{t}_j$ is the time corresponding to prediction point j.

These time differences are then sorted from smallest to largest and expressed as $\Delta_{[i]}(\hat{t}_j)$, where the square brackets around the subscript indicate sorted values.

The proportion of observations to use in each local regression (p, also known as the smoothing parameter) controls the degree of smoothing in the loess curve, with larger values of p tending to increase the smoothness of the loess curve. If p is too large, the curve is over-smoothed and the regression in unable to capture the underlying pattern of the data. If p is too small, an insufficient number of observations will occur near $\hat{t}_j$, resulting in a regression that captures the noise and the outliers in the data along with the underlying pattern. A version of the Akaike information criterion (AIC) that has a correction for small sample sizes (AIC_c) (Hurvich and others, 1998) was used to choose the smoothing parameter p that best fit the time-series data. The AIC_c is defined as

$$\mathrm{AIC_c} = \ln\left[\frac{1}{n}\sum_{i=1}^{n}\left(C_i - \hat{C}_i\right)^2\right] + 1 + \frac{2\left[\mathrm{tr}(H)+1\right]}{n - \mathrm{tr}(H) - 2}, \tag{12}$$

where

H is the hat matrix (also known as the projection matrix) that describes the influence each C_i value has on each fitted value $\hat{C}_i$, and

$\mathrm{tr}(H)$ is the trace of the hat matrix.

The best fitting value of p was determined by minimizing the value of AIC_c with values of p restricted to lie within the range of values from 0.2 to 0.9. Values of p less than 0.2 tended to overfit the data in each subset and produce numerically unstable estimates. The minimization was solved using a combination of golden section search and successive parabolic interpolation (Brent, 1973).

The number of observations in a local regression (q) is defined as $q = \lfloor pn \rfloor$, where q is the rounded-down number to the nearest integer. This q value is used to identify the maximum time difference for observations in the local regression—the value located in the q position of the sorted time differences, $\Delta_{[q]}(\hat{t}_j)$. Thus, only observations whose time difference is less than or equal to this value, $\Delta_i(\hat{t}_j) \le \Delta_{[q]}(\hat{t}_j)$, are included in the neighborhood around the focal $\hat{t}_j$. The LOESS method implements this neighborhood inclusion criterion by expressing the time difference as a proportion of the maximum time difference (d) and is expressed as

$$d_i = \frac{\Delta_i(\hat{t}_j)}{\Delta_{[q]}(\hat{t}_j)} \quad \text{for } i = 1, \ldots, n. \tag{13}$$

Then, $d_i < 1$ is included in the local regression, and $d_i \ge 1$ is not.

Observations are inversely weighted according to their time difference. Weights are assigned to an observation based on the d_i value (eq. 13); this ensures that observations closer to the focal $\hat{t}_j$ have greater influence over parameter estimates in the local regression model. Weights are defined using the tricube weight function, and defined as:

$$w_i(\hat{t}_j) = \begin{cases} \left[1 - d_i^{\,3}\right]^3, & \text{for } d_i < 1 \\ 0, & \text{for } d_i \ge 1 \end{cases} \quad \text{and } i = 1, \ldots, n, \tag{14}$$

where

$w_i(\hat{t}_j)$ is the weight for observation i in the neighborhood of $\hat{t}_j$.

The weights are set to zero for observations located outside the local neighborhood.

Within a neighborhood of observations, a simple linear model is used to describe the statistical relationship between time (t) and constituent concentration (C). The linear regression model has the form:

$$\hat{C}(t) = \alpha_j + \beta_j t \quad \text{for } \hat{t}_j - \Delta_{[q]}(\hat{t}_j) < t < \hat{t}_j + \Delta_{[q]}(\hat{t}_j), \tag{15}$$

where

$\hat{C}(t)$ is the estimated value of constituent concentration at time t; and

α_j and β_j are the intercept and slope of a line, respectively.

The fitted values of α_j and β_j are determined using weighted least squares regression, which is stated in the following optimization formulation:

$$\alpha_j, \beta_j = \underset{\alpha_j, \beta_j \in \mathbb{R}}{\arg\min} \sum_{i=1}^{n} w_i(\hat{t}_j)\left[C_i - \left(\alpha_j + \beta_j t_i\right)\right]^2, \tag{16}$$

where

$\mathbb{R}$ is the set of real numbers.

Once the regression coefficients have been determined, they are used to calculate the vertical offsets (also known as "local residuals" and denoted as e) between the observation points and fitted line. The local residual is obtained by combing equations 10 and 15:

$$e_i = C_i - \hat{C}_i = C_i - \left(\alpha_j + \beta_j t_i\right) \quad \text{for } i = 1, \ldots, n. \tag{17}$$

To reduce the impact of outlier data (points having large local residual values) on the regression equation (eq. 15), the residual of observation i is scaled by 6 times the median of the absolute value of the local residuals:

$$z_i = \frac{e_i}{6 \ \mathrm{median}\left\{|e_1|, \ldots, |e_n|\right\}}, \tag{18}$$

and used to calculate a robustness weight δ_i for each observation i as follows:

$$\delta_i = \begin{cases} \left[1 - z_i^{\,2}\right]^2, & \text{for } |z_i| < 1 \\ 0, & \text{for } |z_i| \geq 1 \end{cases} \quad \text{for } i = 1,\ldots,n. \quad (19)$$

The robustness weights are set to zero for observations with a residual that is six times the median value (or about 4 standard deviations). Local regression (eq. 16) is then repeated but with new weights equal to $\delta_i w_i(\hat{t}_j)$.

The calculation of new weights (eqs. 16–19) is repeated four times to ensure that outlier data are excluded from the local regression. Once the final regression coefficients have been determined, they are substituted back into equation 15 to obtain a single predicted value on the loess curve evaluated at time $\hat{t}_j$. The steps for local regression (eqs. 12–19) are performed for each of the m prediction points:

$$\hat{C}(\hat{t}_j) = \alpha_j + \beta_j \hat{t}_j, \quad \text{for } j = 1,\ldots,m, \quad (20)$$

and the difference between the highest and lowest predicated concentration $\Delta\hat{C}$ is defined by:

$$\Delta\hat{C} = \max\left(\hat{C}_1,\ldots,\hat{C}_m\right) - \min\left(\hat{C}_1,\ldots,\hat{C}_m\right). \quad (21)$$

Once the prediction points are calculated (eq. 20), the 90-percent confidence intervals around the loess curve are determined. At each predicted point j, confidence limits are defined as:

$$\left[\hat{C}_{\ell,j}, \hat{C}_{u,j}\right] = \hat{C}(\hat{t}_j) \pm t^*_{(1-\alpha/2),\nu}\, S_{E,j}, \quad (22)$$

where

$\hat{C}_{u,j}$, $\hat{C}_{\ell,j}$ are the upper and lower confidence limits for prediction point j, respectively;

α is the specified significance level (0 through 1);

$t^*_{(1-\alpha/2),\nu}$ is the 100$(1-\alpha/2)$ percentage point of the Student's t distribution with ν degrees of freedom; and

$S_{E,j}$ is the estimated standard error for prediction point j.

The significance level for a 90-percent confidence interval is 10 percent, or α=0.1. The procedures for calculating ν and $S_{E,j}$ values in equation 22 are beyond the scope of this investigation; however, a description of these procedures is available in Cleveland and others (1992b, p. 45–46).

Because LOESS relies on the local data structure when performing local fitting (eq. 15), a high density of observations is required to adequately smooth the data since large data gaps in an unevenly spaced time series can lead to highly misleading predictions and inferences. Distorting effects to the loess curve caused by infrequent sampling and a variable sampling rate were mitigated by omitting sample data lying outside a continuous record block (that is, the regression analysis was performed on Type 1 data, as described in section, "Sources and Descriptions of Data"). Because LOESS does not account for censored values, a substitution of one-half the reporting limit was made for nondetection values. This substitution can create an artificial trend, especially for those time-series datasets containing many censored values. For example, toluene concentrations in well EBR 1 prior to January 2004 are reported as less than (<) 0.2 µg/L, and after as <0.1 µg/L. This change indicates enhanced detection of low concentrations of toluene and results in an artificial decreasing trend as shown in appendix 5, fig. 5.13, p. 37. Water-quality data from groundwater samples collected on the same day were averaged to better facilitate iterative thinning described in section, "Iterative Thinning."

Two examples of the application of LOESS to smooth unevenly spaced time-series data are shown in figure 6. These include smoothing of tritium in well USGS 20 (fig. 6*A*) and chromium in well USGS 38 (fig. 6*B*). For concentration measurements of tritium in well USGS 20 (fig 5*A*) (n=51), replicate-paired data were collected on five unique sampling dates. After averaging measurements collected on the same day, 46 observations remained for local regression analysis (n=46). The observation points, calculated loess curve, and 90-percent confidence intervals for tritium in well USGS 20 are shown in figure 6*A*. The loess curve and its confidence intervals are represented as continuous functions—line segments are used to connect adjacent prediction points. The *confidence band* is defined as the area between the upper and lower confidence intervals. An analysis of the loess curve indicates a long-term downtrend of tritium concentrations in the well. The narrow confidence band indicates low uncertainty in the local regression model.

For the 34 measurements of chromium concentration in well USGS 38 (fig. 5*B*) (n=37), replicate-paired data collected on three unique sampling dates were separately averaged, leaving 34 observations for local regression analysis (n=34). The observations, loess curve, and 90-percent confidence band for chromium in well USGS 38 are shown in figure 6*B*. The loess curve indicates a step increase in chromium concentration starting in about 2002 and ending in 2008, with concentrations remaining about constant before and after this step increase. The start of the step increase may be artificially delayed because of the substitution of one-half the reporting limit for the five censored values measured from 1999 through 2003.

Time series graphs showing local temporal trends for selected constituents measured for in water samples from wells in the INL water-quality network, 1989–2018, are presented in appendix 5. The absence of a local-regression model for a constituent at a given well site may indicate that (1) the constituent was not measured for in water samples collected from this well, (2) a local-regression model could not be fit to the measured data, or (3) none of the concentration measurements satisfy the conditions for a continuous record block

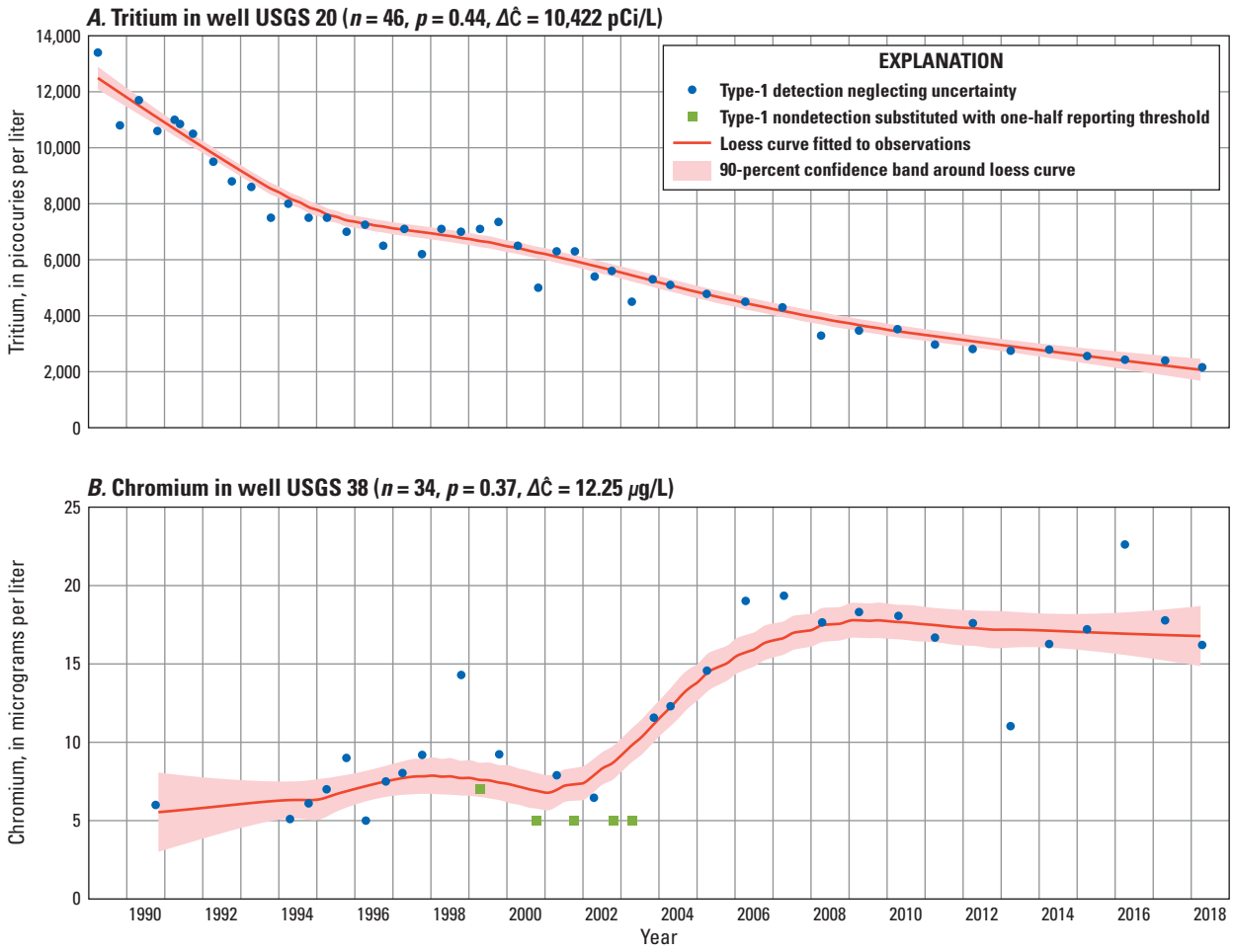


Figure 6. Local-regression model graphs fitted to observations of (*A*) tritium in well USGS 20 and (*B*) chromium in well USGS 38, Idaho National Laboratory, Idaho, 1989–2018. Water-quality data from groundwater samples collected on the same day were averaged. *n*, number of observations; *p*, smoothing parameter; $\Delta\hat{C}$, difference between the highest and lowest predicted concentrations; pCi/L, picocuries per liter; µg/L, micrograms per liter.

(that is, the time series is composed entirely of Type 2 data). For example, there is no loess curve for fluoride concentrations in well USGS 15 because all measurements were identified as Type 2 data (fig. 5*C*).

Survival Regression Analysis

The survival regression model referred to as accelerated failure time (AFT) was used to perform a regression analysis on temporal observations of constituent concentrations in a well. In the context of the network optimization problem, AFT was used to formulate the objective function of the spatial optimization. Unlike ordinary linear regression models, survival methods correctly represent information from censored and uncensored observations in estimating regression parameters. With the presence of censored data, the AFT model (also known as the log-linear model) takes the form of a linear regression model with log-transformation of the response variable (constituent concentration) (Kalbfleisch and Prentice, 2002, chapter 6). As a parametric method, AFT requires that concentration data approximately follow an assumed distribution. A log-normal distribution was chosen for data in this report because most constituent concentrations have positive skewness, a lower bound of zero, and span two or more orders of magnitude. The AFT model is formulated here to describe the relation between time and constituent concentration in single well.

Assuming that for the i^{th} observation in an unevenly spaced time series, the measured constituent concentration (C_i) and corresponding sample time (t_i) are related by

$$\ln(C_i) = \ln[f(t_i)] + \sigma e_i, \tag{23}$$

where

$f(t_i)$ is the regression function evaluated at sample time t_i;

σ is a regression coefficient called the scale parameter;

e_i is the random error in the log-transformed concentration of observation i; and
t is the sample time given as the number of days since January 1, 1970 (the Unix epoch).

A log-linear model is used to represent the regression function, expressed as:

$$f(t) = \exp(\alpha + \beta t), \tag{24}$$

where

exp is the exponential function; and
α and β are regression coefficients, respectively.

Combining equations 23 and 24, the AFT model may be specified as:

$$\begin{aligned} \ln(C_i) &= \ln\left[\exp(\alpha + \beta t_i)\right] + \sigma e_i \\ &= \alpha + \beta t_i + \sigma e_i \end{aligned}. \tag{25}$$

By rearranging equation 25, the random error is expressed for n observations as:

$$e_i = \frac{\ln(C_i) - \alpha - \beta t_i}{\sigma} \quad \text{for } i = 1, \ldots, n. \tag{26}$$

For the i^{th} censored observation, the random error (e_i in eq. 23) may be expressed in terms of the measurement's confidence limits. The upper limit of random error is defined as:

$$e_{u,i} = \frac{\ln(C_{u,i}) - \alpha - \beta t_i}{\sigma}, \tag{27}$$

and lower limit as:

$$e_{\ell,i} = \begin{cases} \dfrac{\ln(C_{\ell,i}) - \alpha - \beta t_i}{\sigma}, & \text{for } \delta_i = 1 \\ -\infty, & \text{for } \delta_i = 0 \end{cases}. \tag{28}$$

where

δ_i is the event indicator for the i^{th} observation with δ_i=1 if the observation is interval censored, and δ_i=0 if the observation is left censored.

For the infinite lower bound of left-censored observations, logarithms of concentration map into a lower bound of zero for the retransformed, original concentration units (Helsel, 2005, p. 202).

The likelihood function (L) defines the likelihood of matching the observed distribution of censored data, and is expressed for n observations as:

$$L = \prod_{i=1}^{n} \left\{\Phi[e_{u,i}]\right\}^{1-\delta_i} \left\{\Phi[e_{u,i}] - \Phi[e_{\ell,i}]\right\}^{\delta_i}, \tag{29}$$

where

$\Phi(e)$ is the standard normal cumulative distribution function of the random error.

A larger value of L indicates an improved fit between the estimated distribution and the observed data. The best fit is determined by choosing values α, β, and σ that maximize L, which is stated in the following optimization formulation:

$$\alpha, \beta = \underset{\alpha,\beta,\sigma \in \mathbb{R}}{\arg\max} \prod_{i=1}^{n} \left\{\Phi\left[\frac{\ln(C_{u,i}) - \alpha - \beta t_i}{\sigma}\right]\right\}^{1-\delta_i} \left\{\Phi\left[\frac{\ln(C_{u,i}) - \alpha - \beta t_i}{\sigma}\right] - \Phi\left[\frac{\ln(C_{\ell,i}) - \alpha - \beta t_i}{\sigma}\right]\right\}^{\delta_i}. \tag{30}$$

The optimization (eq. 30) is solved by setting the partial derivative of the logarithm of L (eq. 29) with respect to each of the regression coefficients equal to zero and using a Newton-Raphson method to iteratively approximate a solution. Once the regression coefficients have been determined, they are substituted back into equation 24 to describe a best-fit smooth curve for the data and a long-term monotonic trend (Δ) estimated for the smooth curve as:

$$\Delta = 100\left[\exp(\beta) - 1\right]365.24, \tag{31}$$

where

Δ is expressed as a percent change per year, with negative values indicating a monotonic decrease in concentration over time.

The predictive strength of the survival regression model was assessed using the McKelvey and Zavoina (1975) pseudo-coefficient of determination (R^2) statistic, which attempts to describe the proportion of variance explained by the model fit and tries to capture the square of the correlation between the fitted and actual values. Pseudo-R^2 ranges from 0 to 1, with higher values indicating better model fit. Unlike the R^2 generated for ordinary least squares regression, the pseudo-R^2 statistic cannot be interpreted independently or compared among time-series datasets. The statistic's usefulness is limited to comparing competing models for the same data.

The p-value for each regression coefficient (α, β, and σ in eq. 25) was used to evaluate the strength of evidence against the null hypothesis, which states that there is no correlation between the coefficient and response variable (Moore and McCabe, 2003), C. If the p-value is less than or equal to a p-value tolerance (called a significance level), the null hypothesis is rejected; that is, the coefficient likely is significant in the model because changes in the coefficient likely are related to changes in the response variable. Conversely, when the

p-value is greater than the significance level, the changes in the coefficient likely are unrelated to changes in the response variable. A significance level of 0.05 was selected for this report because of its traditional usage (Nuzzo, 2014), and a regression model with a coefficient *p*-value greater than 0.05 was omitted. The *p*-value for the overall significance of the survival regression model was used to determine whether the entire model being tested was an improvement over no model at all (Helsel, 2005, p. 203).

Infrequent sampling and a variable sampling rate may result in biased regression coefficients. To mitigate this bias, sample data lying outside a continuous record block were omitted from the regression analysis (that is, the regression analysis was performed on Type 1 data, as described in section, "Sources and Descriptions of Data").

Because concentrations may be censored in survival analysis, a nondetect concentration is represented as left-censored data, and a 95-percent confidence interval about a measured radiological concentration is represented as interval-censored data. For example, consider a measured concentration of 50 pCi/L for tritium. The combined standard uncertainty associated with this measurement was reported by the laboratory as 100 pCi/L. The 95-percent confidence interval is then calculated from equation 1 as:

$$\begin{aligned}\left[C_{\ell}, C_{u}\right] &= 50\ \text{pCi/L} \pm 1.96\,(100\ \text{pCi/L}) \\ &= [-146\ \text{pCi/L},\ 246\ \text{pCi/L}]\end{aligned}. \quad (32)$$

Thus, the concentration measurement of tritium is represented as interval-censored data, with the true concentration estimated with 95-percent confidence to be in the range of -146–246 pCi/L.

The log-transformation of the response variable in the AFT model [$\ln(C_i)$ in eq. 23] requires concentration values to be greater than zero. This is because logarithms are not defined for zero and negative numbers. To ensure that all values are greater than zero, the radiological data were censored using the detection limit (DL); that is, values reported as negative values were changed to less than the DL. The DL is based on instrument sensitivity, sample volumes, analytical procedures, and counting times used in the laboratory. DLs are available from the analyzing laboratory and were reported for selected types of radionuclides by Bodnar and Percival (1982), Bartholomay and others (2003, table 9), and Bartholomay and others (2014, table D1). The DLs shown in table 6 represent typical values—on rare occasions, special arrangements were made to achieve smaller detection limits. Censoring the negative values may obscure some of the underlying trends in the observed data.

For cases where the upper limit of the 95-percent CI was less than the DL for a particular radionuclide, the concentration range was censored to be less than the DL. When the DL was between the upper and lower limits of the 95-percent CI, the concentration range was censored to be less than the upper limit. For example, given a 200 pCi/L DL for tritium (table 6), the 95-percent CI as calculated from equation 32, [-146 pCi/L, 246 pCi/L], is censored to be less than 246 pCi/L. The concentration range remains unchanged for cases where the lower limit was greater than or equal to the DL, or the radiological concentration is reported as a nondetection. The censoring criteria for radionuclides is mathematically expressed as follows:

$$C = \begin{cases} < \text{DL}, & \text{for } C + 1.96\ \text{CSU} < \text{DL} \\ < C + 1.96\ \text{CSU}, & \text{for } C - 1.96\ \text{CSU} < \text{DL} < C + 1.96\ \text{CSU} \\ C \pm 1.96\ \text{CSU}, & \text{for } C - 1.96\ \text{CSU} \geq \text{DL} \\ < C, & \text{for nondetection} \end{cases} \quad (33)$$

For most constituents in this report, concentrations in the survival-regression analysis are represented as an interval of uncertainty in the reported measurement. The effect of including this uncertainty in the trend analysis was a reduction in the trend-detection rate, thereby, helping to avoid invalid conclusions on the presence or absence of trends.

Two examples of the application of survival regression analysis to smooth unevenly spaced time-series data are shown in figure 7. These include a regression analysis

Table 6. Analytical method detection limits of selected radionuclides.

[**Effective period:** Period during which the detection limit was in effect (1989–2018), in mm-dd-yy (month-day-year). **Detection limit:** Based on instrument sensitivity, sample volumes, analytical procedures, and counting times used in the laboratory. pCi/L, picocuries per liter]

Analyte name	Effective period (mm-dd-yy)	Detection limit (pCi/L)	Reference
Tritium	01-01-89 to 03-31-03	500	Bartholomay and others (2003, table 9)
	04-01-03 to 12-31-18	200	Bartholomay and others (2014, table D1)
Strontium-90	01-01-89 to 07-30-98	5	Bodnar and Percival (1982)
	07-31-98 to 12-31-18	2	Bartholomay and others (2014, table D1)
Plutonium-238	01-01-89 to 12-31-18	0.2	Bartholomay and others (2014, table D1)
Cesium-137	01-01-89 to 12-31-18	60	Bartholomay and others (2014, table D1)
Alpha particle	01-01-89 to 12-31-18	3	Bartholomay and others (2014, table D1)
Beta particle	01-01-89 to 12-31-18	2	Bartholomay and others (2014, table D1)

of tritium in well USGS 20 (fig. 7*A*) and chromium in well USGS 38 (fig. 7*B*). The 51 observations (n=51) of measured concentrations represented with their 95-percent uncertainty as interval-censored data are shown in figure 7*A*. The survival regression model is shown as a smooth curve (fig. 7*A*) and is based on estimated regression coefficients of α=10.49 and $\beta = -1.61\times10^{-4}$. Substituting these coefficient estimates into the survival regression function (eq. 24) gives:

$$f(t) = \exp\left(10.49 - 1.61 \times 10^{-4}\, t\right). \quad (34)$$

Thus, on January 1, 2000 (a calendar date represented by the number of days since January 1, 1970, or t=10,957), the predicted tritium concentration based on regression analysis was 6,160 pCi/L. Substituting -1.61×10^{-4} for β in equation 31 gives the long-term monotonic trend (Δ) for tritium in well USGS 20 as:

$$\Delta = 100\left[\exp\left(-1.61 \times 10^{-4}\right) - 1\right]365.24 = -5.9, \quad (35)$$

which indicates a 5.9 percent reduction in tritium concentration per year.

For the 37 measurements of chromium concentration in well USGS 38 (fig. 7*B*), 32 observations were reported without uncertainty and were represented as exact detections, and 5 observations were reported as nondetections and were represented as left-censored data. The censored and uncensored observations of chromium in well USGS 38 are shown in figure 7*B*, along with the fitted survival regression model and the 95-percent prediction band for the fitted survival regression model. The smooth curve indicates an increase in chromium concentration over time, with a long-term monotonic trend of 5.0 percent change per year.

Fitted survival regression models for selected constituents measured for in water samples collected from wells in the INL water-quality network, 1989–2018, are described in appendix 6 and shown in appendix 7. The absence of a survival regression model for a constituent at a given well site may indicate that (1) the constituent was not measured for in water samples collected from this well, or (2) a survival

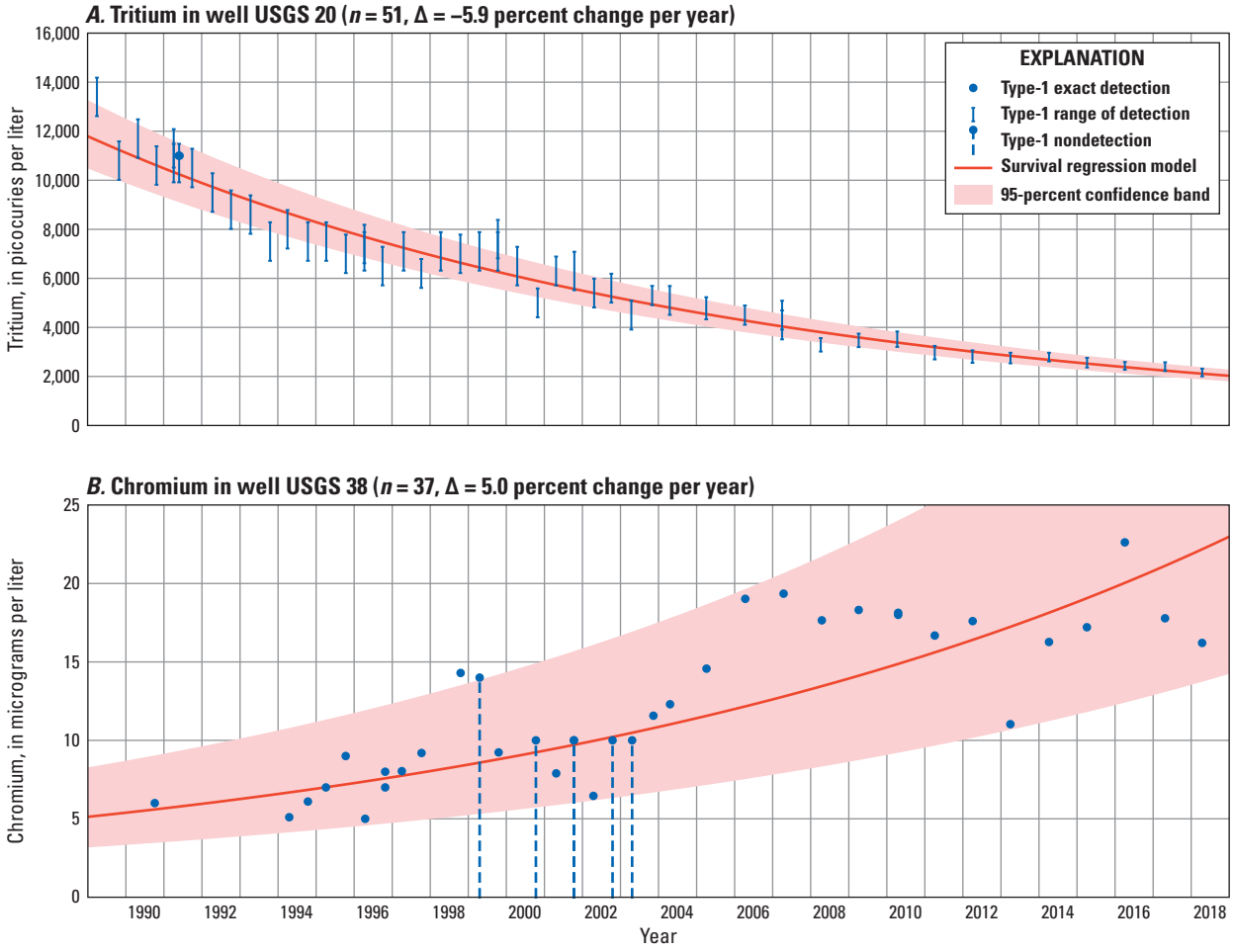


Figure 7. Survival regression function model with measurements of (*A*) tritium in well USGS 20 and (*B*) chromium in well USGS 38, Idaho National Laboratory, Idaho, 1989–2018. *n*, number of observations; Δ, long-term monotonic trend in percent changes in concentration per year.

regression model could not be fit to the measured data or was omitted because of a large *p*-value, or (3) none of the concentration measurements satisfy the conditions for a continuous record block (that is, the time series is composed of Type 2 data).

A visual inspection of regression models indicates an adequate model fit to their observed data values for most constituent-well pairs. The few exceptions of poor model fit occurred when there was an abrupt change in concentration at a specific point in time (non-monotonic characteristics). For example, a relatively late arrival time for groundwater contamination in well USGS 120—in comparison to other wells in the monitoring network—resulted in peak concentrations occurring during the 1999 calendar year for most constituents measured for in water samples collected from that well. This abrupt change in concentration is not represented by the model and indicates the inability of the regression model to account for non-monotonic representations of change. The sensitivity of the spatial optimization results to these rare occurrences of poor survival-regression model fit was not explored in this report.

Spatial Interpolation

The geostatistical technique known as kriging is used to interpolate the concentration for a particular constituent at unmeasured locations in the ESRP aquifer beneath the INL and vicinity. Snyder (2008, p. 19) describes kriging as a type of spatial moving average, where the value at an unmeasured location is estimated as a weighted average of the measured values. The weights assigned to the measured values depend on spatial trends and possible correlations in the data (Bossong and others, 1999, p. 4). Correlation between measurements at two sampling sites is assumed to depend on the separation distance between the two sites. Any depth dependencies among measurements acquired at isolated depths in the same MLMS-equipped well are ignored. For this report, concentrations were depth-averaged to provide the data necessary for two-dimensional kriging. Any bias introduced by depth averaging was assumed negligible because of the small number of wells (about 7 percent) equipped with a MLMS in the existing monitoring network.

Sampling sites that are close together typically have a smaller difference in measured values than those farther apart. The degree of spatial correlation is quantified with the sample variogram, which measures correlation between measurements as a function of distance between the sampling points. Kriging computes an estimate best representing the spatial distribution of the observed values based on a theoretical variogram model that is fitted to the sample variogram and a minimization of the estimation variance (a measure of uncertainty) at measured locations.

Kriging was used to estimate the areal distribution of the time-averaged concentrations for selected constituents. Predicted concentrations were calculated on a regular grid of 0.31-by-0.31 mi (or 0.5-by-0.5 km) resolution across the land surface area defined by a generalization of the convex hull of the monitoring wells. The average separation distance between pairs of monitoring wells was 1.7 km, with only 86 of the 8,778 possible combinations of well pairs separated by less than 0.5 km. Temporal variability was eliminated by averaging the measured concentrations of a constituent in a well during 1989–2018. That is, for the i^{th} monitoring well in the network, the time-averaged concentration for a constituent ($\overline{C}$) is defined as:

$$\overline{C}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} C(\boldsymbol{s}_i, t_j) \quad \text{for } i = 1, \ldots, n, \tag{36}$$

where

$\boldsymbol{s}_i$ is a pair of Cartesian coordinates describing the geographic location (point) of well i;
t_j is the sampling time for observation j;
$C(\boldsymbol{s}_i, t_j)$ is the measured concentration at point $\boldsymbol{s}_i$ and sample time t_j;
m_i is the number of observations in well i; and
n is the number of wells in the monitoring network.

A boldface algebraic symbol (such as $\boldsymbol{s}$ in eq. 36) is used to denote a vector quantity.

The method of time averaging was selected solely for addressing the problem of estimating the spatial distribution of constituent concentrations at a single snapshot in time with observations that are (1) limited in number and sparsely distributed, both spatially and temporally in the monitoring domain; and (2) often reported as a nondetection or background concentration value. Kriging models developed for predicting spatially smoothed and temporally averaged constituent concentration likely will not be successful in predicting observations that are spatially localized and time-varying. However, these models can capture the aggregate plume behavior and can be useful for estimating the occurrence and spatial extent of a constituent in the ESRP aquifer at and near the INL.

For many of the constituents, the areal distribution of time-averaged concentrations can be characterized by locally extreme values near the waste-disposal sites, surrounded by much smaller background or nondetection values. Because kriging estimators are sensitive to a small number of large data values, the highly asymmetric empirical distribution of data may significantly bias kriging estimates.

Quantile kriging (QK) is a transformation-based kriging approach that preserves the relative ranks and spatial structure of data while addressing estimation inaccuracies caused by a skewed distribution (Juang and others, 2001). The approach originally was proposed by Journel and Deutsch (1997) for integrating diverse data types, and Reed and others (2004)

reported that it was well-suited for plume interpolation. QK is formulated here to describe the estimated concentration distribution of a constituent in the study area, and the uncertainty (or error) associated with these kriging predictions.

Because the kriging method does not account for censored values, a substitution of zero was made for nondetection concentration values. This substitution can result in a distorted model of spatial distribution, especially for analytes containing many censored values (such as carbon tetrachloride, 1,1-dichlorothylene, 1,1,1-trichloroethane, and trichloroethylene). However, in such cases, censored values typically are representative of background concentrations, so the substitution of zero has little effect on the predicted concentration distribution of a contaminant plume.

Transformation

The averaged concentrations ($\overline{C}$ in eq. 36) were transformed into standardized ranks (or quantiles) (z) to produce data that are uniformly distributed on the interval from 0 to 1. That is, the non-linear data transformation changes the shape of the original data distribution into a uniform distribution. The transformation was performed using the empirical distribution function (EDF) associated with the averaged concentration sorted from smallest to largest and was denoted as $\overline{C}_{[i]}$, where square brackets around the subscript indicated sorted values. The EDF is mathematically expressed as:

$$z_i = \frac{1}{n}\sum_{j=1}^{n}\begin{cases}1, & \text{for } \overline{C}_{[j]} \le \overline{C}_{[i]}\\ 0, & \text{for } \overline{C}_{[j]} > \overline{C}_{[i]}\end{cases} \quad \text{for } i = 1,\ldots,n, \qquad (37)$$

where

n is the number of time-averaged concentrations.

Spatial interpolation is then performed on the standardized ranks using ordinary kriging (OK) and the predicted values at unmeasured locations are back-transformed into concentration space.

Variograms

Kriging predictors require estimates of the degree of spatial correlation between values of standardized rank (z) separated by different distances ($\boldsymbol{h}$). The separation distance (or Euclidian distance) between any two points ($\boldsymbol{s}_i$ and $\boldsymbol{s}_j$) is defined as:

$$h_{ij} = \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \qquad (38)$$

where

x and y are the easting and northing coordinates, respectively.

The variogram may be used to estimate the degree of spatial correlation present in the data. Because the true variogram can never be known, a nonparametric estimate of the variogram is made using the sample variogram $\tilde{\gamma}$ and computed by averaging variance values $\gamma(h)$ that are in a given h interval (or lag bin). That is, the squared difference in transformed concentrations (z values calculated in eq. 37) is averaged for well pairs separated by a distance that is contained with the same lag bin. Assuming the variance is isotropic with respect to different directions in the areal plain, the sample variogram is defined as:

$$\tilde{\gamma}_k = \frac{1}{2\left|N\left(\tilde{h}_k \pm \delta\right)\right|}\sum_{(i,j)\in N\left(\tilde{h}_k \pm \delta\right)}\left(z_i - z_j\right)^2$$

$$\text{let } N\left(\tilde{h}_k \pm \delta\right) \equiv \left\{\begin{array}{c}\left(\boldsymbol{s}_i, \boldsymbol{s}_j\right) \left| \tilde{h}_k - \delta \le h_{ij} < \tilde{h}_k + \delta \right. \\ \text{for } i,j = 1,\ldots,n \text{ and } i \ne j\end{array}\right\}, \qquad (39)$$

$$\text{and } \tilde{h}_k = \frac{d(k-0.5)}{n_b}, \quad \text{for } k = 1,\ldots,n_b$$

where

$\tilde{h}_k$ is the lag distance coinciding with the midpoint of lag bin k,

δ is half of the lag bin width,

$|N(\tilde{h}_k \pm \delta)|$ is the number of data pairs in lag bin k,

d is the separation distance to which point pairs are included in the variance estimates, and

n_b is the total number bins.

As a rule of thumb, separation distance (d in eq. 39) typically is limited to no more than one-half the maximum separation distance between any two points (Rossi and others, 1992) and each lag bin has at least 30 data pairs (Cressie, 1993). In this report, d was specified at one-half the separation distance between any two monitoring wells, or about 10 mi, a distance subdivided into 20 equal-width bins. Each lag bin is 0.5 mi wide with the number of data pairs within each bin ranging from 107 to 338.

The sample variogram is modeled with a covariance function that represents a theoretical variogram $\gamma(h)$ and provides variance values at any given lag distance. The mathematical function used in this study to describe spatial variability is the Matérn model (Matérn, 1960; Stein 1999), which provides more flexibility in modeling the smoothness of the covariance function, as compared with the other possibilities for the function (such as the spherical model), and it can model many local spatial processes (Minasny and McBratney, 2005). The Matérn variogram model is defined as:

$$\gamma\left(h\right) = \begin{cases} 0, & \text{for } h = 0 \\ g + \left(s - g\right)\left[1 - \frac{2^{1-v}}{\Gamma\left(v\right)}\left(\frac{h}{\rho}\right)^{v} K_v\left(\frac{h}{\rho}\right)\right], & \text{for } h > 0 \end{cases} \qquad (40)$$

where

g, s, ρ, and v represent the nugget effect, sill, range, and shape of the variogram model, respectively.

The nugget effect g is the jump in the variance at the origin (the variance is always zero at h=0), which may be attributed to measurement errors and (or) sources of variation at a scale smaller than the separation distance between wells (Clark, 2010). The sill s is the variance when the model either reaches or becomes asymptotic to a constant value as lag distance increases. The range ρ controls the rate of increase with distance, and the shape parameter v controls the smoothness (or continuity) of the random field that models the spatial variability of the observed data. A v value of 0.5 results in the exponential variogram, whereas a value of infinity is exactly the Gaussian variogram.

Functions $\Gamma(v)$ and $K_v(h / \rho)$ in equation 40 are known as the gamma function and modified Bessel function, respectively. The gamma function is defined for v greater than zero as:

$$\Gamma(v) = \int_0^\infty x^{v-1} \exp(-x)\, dx. \quad (41)$$

The modified Bessel function is of the third kind, of order v, and exponentially scaled is expressed as:

$$K_v(x) = \frac{\pi}{2}\left[\frac{I_{-v}(x) - I_v(x)}{\sin(\pi x)}\right]\exp(x), \quad (42)$$

where

x is h / ρ.

The function $I_v(x)$ in equation 42 is the modified Bessel function of the first kind; that is, defined by:

$$I_v(x) = \left(\frac{x}{v}\right)^v \sum_{k=0}^{\infty} \frac{1}{k!\ \Gamma(v+k+1)}\left(\frac{x}{2}\right)^{2k}. \quad (43)$$

Fitted parameter values of g, s, and ρ in equation 40 were noted for a number of realizations of the shape parameter v, specifically v={0.1, 0.2, …, 9.9, 10}, to explore a wide range of possible variograms. For each realization of v, parameters g, s, and ρ were calculated by minimizing the weighted sum of squared errors (SSE) between the sample variogram $\tilde{\gamma}(\tilde{h})$ and theoretical variogram $\gamma(h)$ evaluated at lag distances $\tilde{h}$; that is:

$$g,s,\rho = \underset{g,s,\rho \in \mathbb{R}}{\arg\min} \sum_{k=1}^{n_h} w_k \left[\tilde{\gamma}\left(\tilde{h}_k\right) - \gamma\left(\tilde{h}_k; g,s,\rho,v\right)\right]^2, \quad (44)$$

subject to: $0 < g < s;\ s > 0;\ \rho > 0$

where

w_k is the weight of the k^{th} point in the sample variogram expressed as the number of data pairs in the bin divided by the square of its lag distance, or:

$$w_k = \frac{\left|N\left(\tilde{h}_k \pm \delta\right)\right|}{\tilde{h}_k^2}. \quad (45)$$

The optimization (eq. 44) is solved by iteratively reweighted least squares (Cressie, 1985). Once the fitted parameters have been determined for each realization of v, they are substituted back into equation 40 to describe a set of fitted variogram models. A leave-one-out cross validation (LOOCV) method was used to choose the best-fitted variogram model from the set of candidate models. The LOOCV method ascertains the predictive performance of the variogram models using the minimum root-mean-square error (RMSE; described in section, “Kriging”).

Because the physical meaning of the range parameter ρ is difficult to interpret in the Matérn model, an estimate is made of the separation distance after which pairs of points are no longer spatially correlated. The estimate is made using the effective range (r), the distance at which the variance value achieves 95 percent of the sill. The Matérn model only asymptotically reaches the sill. The effective range was determined by minimizing the absolute difference between the variogram model evaluated at a distance r and 95 percent of the sill s, and is expressed as:

$$r = \underset{r \in \mathbb{R}}{\arg\min} \left|\gamma(r) - 0.95s\right| \quad \text{subject to: } 0 < r < 50\rho + 1. \quad (46)$$

The optimization (eq. 44) was solved using a combination of golden section search and successive parabolic interpolation (Brent, 1973).

The SSE statistic calculated in equation 44 cannot be interpreted independently or compared among constituent datasets; that is, its usefulness is limited to comparing competing models for the same data. Therefore, to compare among constituents how well the sample variogram is replicated by the theoretical variogram, R^2 was calculated for each constituent. The R^2 statistic for a model fitted by weighted least squares regression (R^2_{wls}) may be expressed as:

$$R^2_{wls} = 1 - \frac{\text{SSE}}{\sum_{k=1}^{n} w_k \left[\gamma\left(\tilde{h}_k\right) - \frac{\sum_{i=1}^{n} w_i \gamma\left(\tilde{h}_i\right)}{\sum_{i=1}^{n} w_i}\right]^2}. \quad (47)$$

The R^2_{wls} normally ranges from 0 to 1, where a value of 1 indicates that the regression predictions perfectly fit the sample variogram. Models were rejected when their R^2_{wls}

value was less than 0.05, which implies that 95 percent of the variability of the semivariance in the sample variogram is unaccounted for by the model.

The sample variogram and fitted variogram model for tritium based on the time- and depth-averaged measurements in 133 monitoring wells are shown in figure 8. Of the 8,778 combinations of well pairs, 5,036 (57 percent) are separated by a distance of less than 10 mi and are included in the variogram analysis (fig. 8*A*). Estimates of variance between values of transformed tritium separated by different lag distances are allocated into 20 lag bins and averaged within each bin. The averaged values are plotted as point symbols at each of the midpoint lag-bin distances and collectively describe the sample variogram (fig. 8*B*). Numbers next to the point symbols refer to the number of sampled data pairs in a lag bin. The variogram model fitted to the sample variogram is drawn as a smooth curve and shows variance initially increasing with lag but later leveling off for larger lags (fig. 8*B*). The Matérn model [$\gamma(h)$ in eq. 40] is fitted with the following parameter coefficients: (1) nugget effect g (or the variance at the origin) of 0.015, in squared units of standardized rank; (2) sill s (or upper bound of the model) of 0.143, in squared units of standardized rank; (3) range ρ of 2.0 mi; (4) effective range r (or lag distance at which the model reaches 95 percent of the sill) of 5.8 mi; and (5) shape parameter v of 0.5, a unitless value (table 7).

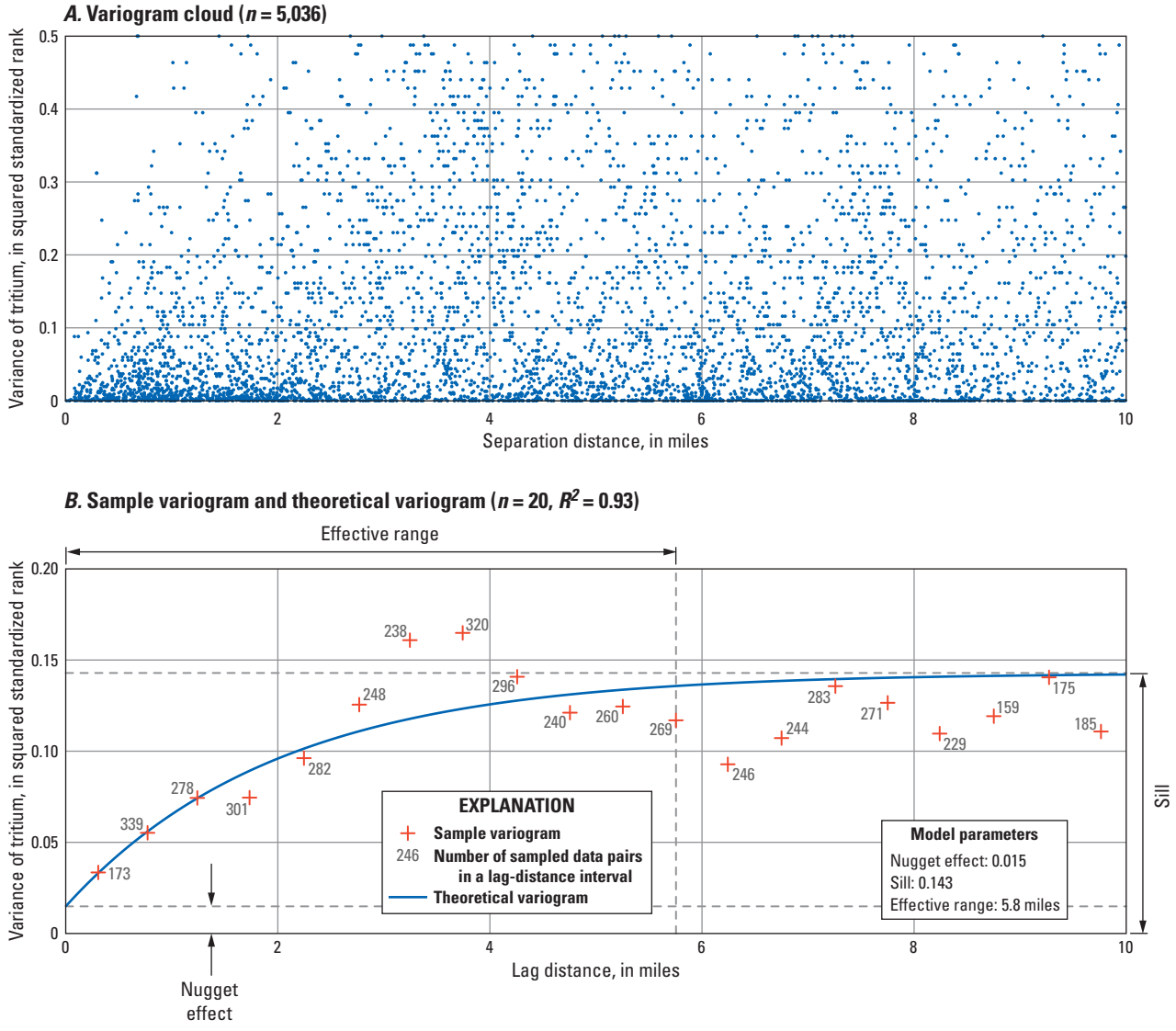


Figure 8. Variogram analysis of tritium measured for in water samples from wells in the Idaho National Laboratory water-quality aquifer monitoring network, averaged during 1989–2018, and transformed into standardized rank space. n, number of points; R^2, coefficient of determination.

Table 7. Fitted theoretical variogram models and kriging performance for selected analytes.

[**Number of observations:** Number of observation points used in the variogram calculation, where concentrations in a well site were averaged over time and represent a single observation. **Variogram:** Describes the theoretical variogram model by its parameter coefficients and fit. A Matérn class model is used for the theoretical variogram. **Nugget effect**: Jump in the variance at the origin, in squared standardized rank. The nugget effect may be attributed to measurement errors and (or) sources of variation at a scale smaller than the separation distance between wells. **Sill:** is the variance, in squared standardized rank, when the model either reaches or becomes asymptotic to a constant value as lag distance increases. **Range:** Controls the rate of increase with distance. **Effective range:** Lag distance, in miles, after which the model remains at 95 percent of the sill. Effective range represents the distance limit beyond which the data are no longer correlated. **Shape:** Controls the smoothness of the random field that models the spatial variability of the observed data. **SSE:** Weighted sum of squared errors between the sample variogram points and the fitted variogram model (ideally 0). R^2_{wls}**:** Coefficient of determination from weighted least squares for the variogram model fitted to the sample variogram values (ideally 1). **Kriging:** Describes the goodness of fit of the kriging model to observed values using leave-one-out cross validation to evaluate model performance. Performance metrics include: **RMSE**, the root-mean-square error (ideally small) in concentration space; and R^2_{cv}, the coefficient of determination from cross validation (ideally 1). **Symbol:** –, model was omitted because its coefficients are inestimable or unreliable]

Analyte name	Number of observations	Variogram							Kriging	
		Nugget effect *g*	Sill *s*	Range *ρ*	Effective range *r*	Shape *v*	SEE	R^2_{wls}	RMSE	R^2_{cv}
Sodium	132	0.005	0.095	0.4	1.1	0.4	6.85×10^{-8}	0.76	7.61	0.29
Chloride	133	0.000	0.101	1.3	3.6	0.4	3.42×10^{-8}	0.93	20.57	0.56
Sulfate	127	0.005	0.084	1.1	2.9	0.4	1.11×10^{-8}	0.96	12.92	0.07
Fluoride	–	–	–	–	–	–	–	–	–	–
Chromium	–	–	–	–	–	–	–	–	–	–
Nitrate	133	0.000	0.158	18.0	36.0	0.2	7.72×10^{-8}	0.81	0.88	0.27
Carbon tetrachloride	111	0.011	2.166	344.5	1,029.8	0.5	3.20×10^{-8}	0.48	0.64	0.30
Chloroform	–	–	–	–	–	–	–	–	–	–
Toluene	–	–	–	–	–	–	–	–	–	–
Tetrachloroethylene	–	–	–	–	–	–	–	–	–	–
1,1-Dichloroethylene	111	0.000	0.108	0.1	0.6	1.8	5.52×10^{-8}	0.83	0.02	0.48
1,1,1-Trichloroethane	111	0.042	0.125	0.2	1.4	5.5	1.67×10^{-8}	0.97	0.09	0.30
CFC-12	–	–	–	–	–	–	–	–	–	–
Trichloroethylene	111	0.022	0.110	15.1	33.2	0.3	2.70×10^{-8}	0.62	0.40	0.08
Tritium	133	0.015	0.143	2.0	5.8	0.5	5.58×10^{-8}	0.93	3,135.83	0.54
Strontium-90	120	0.001	0.119	0.3	1.0	0.7	1.15×10^{-8}	0.98	5.03	0.35
Plutonium-238	–	–	–	–	–	–	–	–	–	–
Cesium-137	–	–	–	–	–	–	–	–	–	–
Alpha particle	–	–	–	–	–	–	–	–	–	–
Beta particle	–	–	–	–	–	–	–	–	–	–

Fitted theoretical variogram models for selected constituents measured for in water samples collected from wells in the INL water-quality network, and temporally averaged during 1989–2018, are described in table 7 and shown in appendix 8. The absence of a variogram model for a constituent may either indicate (1) a failure to converge on a unique set of parameter values—that is, an infinite number of possible combinations of sill and range exist that may be used to fit the model; or (2) a poor correlation between the theoretical and sample variogram points as indicated by an R^2 statistic of less than zero.

A visual inspection of the variograms indicates good model fits to their respective sample variogram for 1,1,1-trichloroethane, chloride, tritium, strontium-90, sodium, nitrate, and sulfate, and adequate model fits for carbon tetrachloride and trichloroethylene (table 7; app. 8). The sample variogram for strontium-90 shows the least amount of local variability in variance values. Cyclic or periodic patterns in the other sample variograms may be attributed to the presence of multiple overlapping plumes and (or) plumes that become discontinuous and move as separate fingers—a common characteristic of groundwater flow through a fractured bedrock aquifer. The variogram model used to approximate the sample variogram does not account for large local variability; rather, it monotonically increases with distance, indicating that the farther two sample sites are apart the more their standardized-rank values tend to differ, on average, from one another (Bossong and others, 1999, p. 13).

Kriging

For each of the selected constituents, an OK method was used to estimate values of standardized rank (z in eq. 37) at unmeasured locations. The OK model represents $z(\mathbf{s})$ as wavering about a constant value (m), and information about the scale and intensity of fluctuations about this constant is provided by the theoretical variogram model (Kitanidis, 1997, p. 120). In mathematical terms, the model is expressed as:

$$z(\boldsymbol{s}) = m + e(\boldsymbol{s}), \tag{48}$$

where

$e(\boldsymbol{s})$ is the stochastic part of z at point $\boldsymbol{s}$ in standardized-rank space, and with a mean of zero.

The deterministic part of z is defined as the expected value (E) of the standardized ranks, denoted by:

$$\mathrm{E}[z(\boldsymbol{s})] = m. \tag{49}$$

To obtain an estimate of z at a point $\boldsymbol{s}_0$ (an estimation point) from transformed measurements of concentration $z(\boldsymbol{s}_1)$, $z(\boldsymbol{s}_2)$, …, $z(\boldsymbol{s}_n)$ requires the following:

(1) The estimate is a linear function of the observed values, that is:

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i, \tag{50}$$

where

$\hat{z}_0$ is the estimate of z at point $\boldsymbol{s}_0$;
z_i is the transformed concentration measurement at point $\boldsymbol{s}_i$;
λ_i is the weighting coefficient corresponding to well site i; and
n is the total number of wells where the constituent was sampled for.

(2) The estimate at sampling points is unbiased, that is:

$$\mathrm{E}[\hat{z}_0 - z_0] = 0. \tag{51}$$

(3) The estimated variance [σ_0^2], or mean square estimation error, at point s_0 should be as small as possible, where the variance is defined as:

$$\sigma_0^2 = \mathrm{E}\left[(\hat{z}_0 - z_0)^2\right]. \tag{52}$$

The unbiased condition (eq. 51), combined with the estimate in equation 50 and the expected value in equation 49, becomes:

$$\begin{aligned} \mathrm{E}\left[\sum_{i=1}^{n}\lambda_i z_i - z_0\right] &= 0 \\ \sum_{i=1}^{n}\lambda_i \mathrm{E}[z_i] - \mathrm{E}[z_0] &= 0 \\ \sum_{i=1}^{n}\lambda_i m - m &= 0 \\ m\left(\sum_{i=1}^{n}\lambda_i - 1\right) &= 0 \end{aligned}. \tag{53}$$

For this condition (eq. 53) to hold for any value of m requires that:

$$\sum_{i=1}^{n}\lambda_i = 1. \tag{54}$$

The estimated variance (eq. 52) in terms of the theoretical variogram model [$\gamma(h)$ in eq. 40] may be computed using the condition in equation 53, and may be expressed as:

$$\sigma_0^2 = -\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(h_{ij}) + 2\sum_{i=1}^{n}\lambda_i\gamma(h_{i0}). \tag{55}$$

Coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n$ are determined by minimizing the estimated variance (eq. 55) subject to the linear constraint of equation 54, or:

$$\lambda_1,\ldots,\lambda_n = \underset{\lambda_1,\ldots,\lambda_n\in\mathbb{R}}{\arg\min} -\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma\left(h_{ij}\right)+2\sum_{i=1}^{n}\lambda_i\gamma\left(h_{i0}\right) \quad \text{subject to: } \sum_{i=1}^{n}\lambda_i = 1 \quad . \quad (56)$$

The minimization is performed algebraically. Once the λ coefficients have been determined, they are substituted back into equation 50 to estimate standardized ranks at a single unmeasured location, and equation 55 to estimate the kriging variance of the predicted value, in standardized rank space. The kriging process is repeated for every active node in the interpolation grid.

Back Transformation

The back transformation from standardized ranks estimated from OK to concentration space is based on the transformation model of equation 37. Recall that the model describes the relation between concentrations ($\bar{C}$; see eq. 36) and standardized ranks (z) using the EDF, a step function that is monotonously increasing. An estimated value of $\hat{z}_0$ (eq. 50) typically will fall between two consecutive standardized ranks z_i and z_{i+1} that correspond to $\bar{C}_{[i]}$ and $\bar{C}_{[i+1]}$ in concentration space. Like in Juang and others (2001, p. 897), a midpoint method is used to interpolate within a standardized-rank interval. The back-transformation model is given by:

$$\hat{C}_0 = \begin{cases} \bar{C}_{[1]}, & \text{for } z_0 \le z_1 \\ \dfrac{\bar{C}_{[i+1]}-\bar{C}_{[i]}}{2}, & \text{for } i\in\mathbb{N}\,\middle|\, z_i < z_o \le z_{i+1} \\ \bar{C}_{[n]}, & \text{for } z_0 > z_n \end{cases} , \quad (57)$$

where

$\hat{C}_0$ is the estimated concentration at an unmeasured location $\boldsymbol{s}_0$, and

n is the total number of point observations.

The square root of the kriging variance (also known as the standard error [SE]) may then be back-transformed into concentration-space by using the same mapping that was discussed for concentration estimates (eq. 55).

The predictive strength of the kriging model was assessed using the LOOCV method (Pebesma, 2004). In this method, the fitted theoretical variogram model is used in a kriging analysis in which an individual observation $\bar{C}_i$ is omitted and a kriging prediction is made at the location of the suppressed observation using the remaining subset of n–1 observations. This process is repeated n times for each observation. The LOOCV procedure is evaluated using the cross-validated RMSE and R^2. The cross-validated RMSE is defined by:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\bar{C}_i-\hat{C}_i\right)^2}{n}}, \quad (58)$$

where

$\bar{C}_i$ is time-averaged concentration in well i, and

$\hat{C}_i$ is predicted value obtained when the kriging model is estimated with the i^{th} observation omitted.

Ideally zero, the RMSE statistic was used to choose the best-fitted theoretical variogram model (provided in eq. 40) from the set of candidate models, and the model with the smallest cross-validated RMSE was selected.

The RMSE statistic calculated in equation 58 cannot be interpreted independently or compared among constituent datasets; that is, its usefulness is limited to comparing competing models for the same data. Therefore, to compare among constituents how well the observed data are replicated by the kriging model, the cross-validated coefficient of determination (R^2_{cv}) statistic was calculated for each constituent, and defined by:

$$R^2_{cv} = 1-\frac{\sum_{i=1}^{n}\left(\bar{C}_i-\hat{C}_i\right)^2}{\sum_{i=1}^{n}\left(\bar{C}_i-\bar{\bar{C}}\right)^2}, \quad (59)$$

where

$\bar{\bar{C}}$ is the arithmetic mean of the time- and depth-averaged concentrations.

The R^2 normally ranges from 0 to 1, where a value of 1 indicates that the kriging predictions perfectly fit the observed data. Models were rejected when their R^2_{cv} value was less than zero.

An example of the application of QK to spatially interpolate tritium observations is shown in figure 9. The kriging analysis is performed using tritium concentrations measured in water samples collected from 133 wells. In each well, recorded tritium concentrations were averaged during 1989–2018, and depth averaged in wells instrumented with multilevel monitoring systems (averaged measurements are shown in app. 9, fig. 9.3*C*). Kriging estimates were made at 8,507 evenly spaced nodes, separated by a distance of 0.31 mi (or 500 m) in the active part of the interpolation grid. The kriging analysis results in two values for each active node location: the predicted concentration (fig. 9*A*; eq. 57) and the kriging SE (fig. 9*B*).

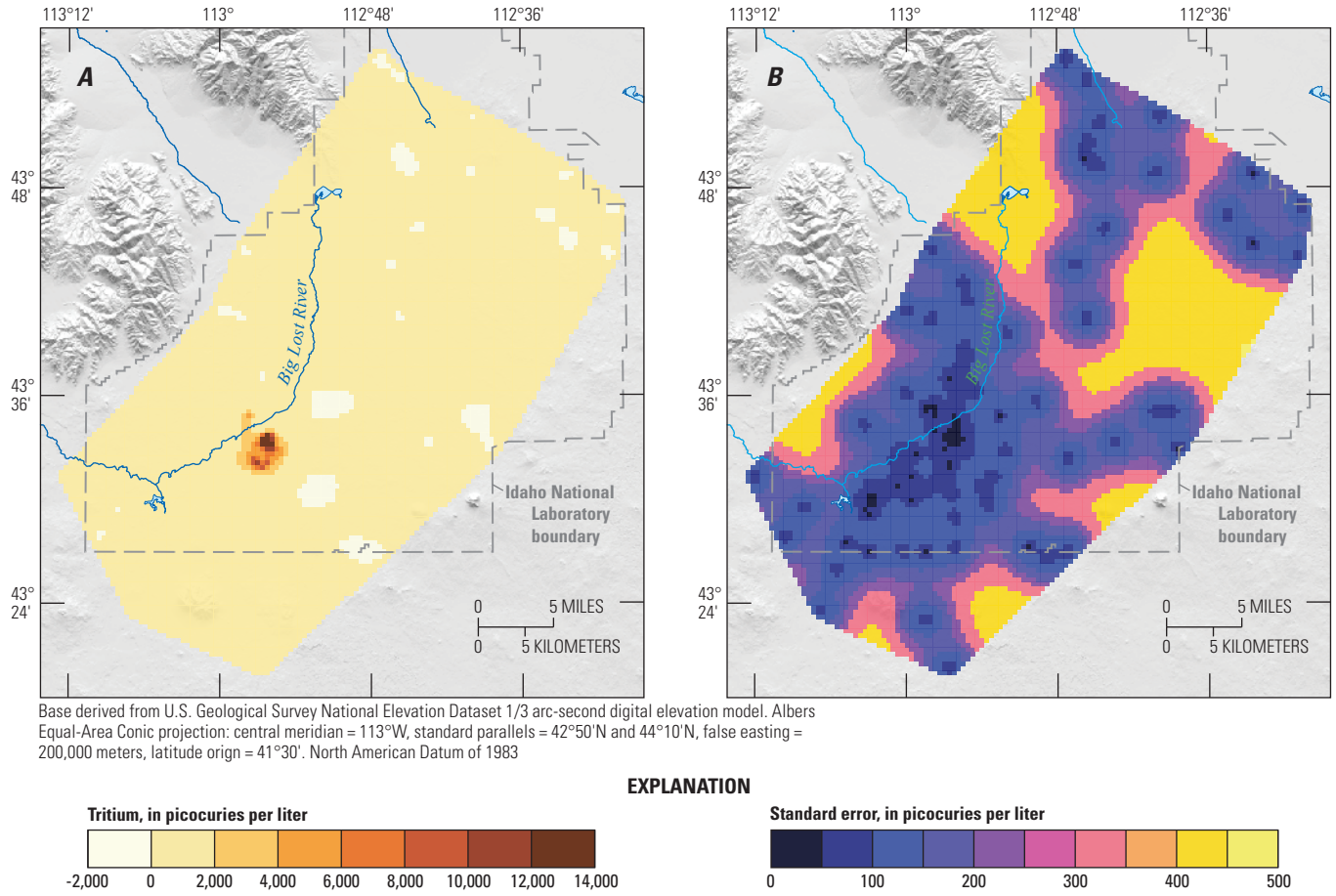


Figure 9. Kriging estimates of the (*A*) prediction surface and (*B*) standard error surface of tritium measured for in water samples collected from wells in the Idaho National Laboratory water-quality aquifer monitoring network and averaged during 1989–2018. Tritium is predicted at points on a regular grid with a spacing of 500 meters and an interpolation domain that is defined by the generalized convex hull of the monitoring sites.

The spatial distribution of predicted tritium concentrations indicates the presence of locally extreme values (greater than 2,000 pCi/L) in groundwater beneath and near the INTEC and Central Facilities Area (CFA) (figs. 1–2; app. 1) with much smaller background concentrations (less than 2,000 pCi/L) in the surrounding area (fig 9*A*). The predicated values agree well with the empirical data, with larger averaged measured concentrations within the boundaries of the predicted contaminant plume. The underestimation of large concentrations by the kriging model is the result of prediction locations not coinciding with measurement locations. For example, the maximum average tritium concentration was 21,642 pCi/L in well USGS 65, whereas the maximum predicted tritium concentration was much smaller at 13,176 pCi/L (app. 9, fig. 9.3).

Every kriging prediction is accompanied by a corresponding measure of the uncertainty associated with the prediction. Figure 9*B* shows the spatial distribution of kriging SE for tritium in units of concentration. Values of SE are basically a scaled version of the distance to the nearest measurement location; that is, SE is small near a sampling site and increases as the density of the monitoring network decreases.

Maps of the back-transformed kriging estimates of concentration and SE are shown in appendix 9 for selected constituents. An examination of the prediction maps indicates the presence of isolated contaminant plumes at and near one or more INL site facilities. Recall that the site facilities are the primary sources of radiochemical and chemical constituents in the ESRP aquifer at the INL (Bartholomay and others, 2000). Elevated concentrations (greater than background levels) of nitrate and sulfate additionally were predicted in the northeastern part of the study area and attributed to the application of fertilizer through irrigation systems in the Mud Lake area northeast of the INL (fig. 1; app. 1). For all constituents, the general spatial characteristics of kriging SE are the same, with small values of SE near the sampling sites and increases in SE as the density of the monitoring network decreases. The rate of increase and maximum estimated SE is based on the spatial correlation of the measured concentrations as depicted by the shape of the underlying theoretical variogram.

Spatial Optimization

Planning Objective

The planning objective for the water-quality monitoring network is to reduce total monitoring costs by removing sampling sites from the existing network because they add little or no information characterizing the concentration plume for selected analytes in the aquifer. In this study, equal monitoring costs are assumed for each sampling site. Although the validity of this assumption is untrue (for example, travel time can account for large variability in monitoring costs), it permits wells to be evaluated exclusively through a geostatistical analysis of the water-quality concentration measurements. An estimate of the true cost savings for a spatially optimized monitoring network is beyond the scope of this study; however, decreases in the total number of sampling sites in a monitoring network will result in a reduction of total monitoring costs.

Design Criteria

To accomplish the established objective of the monitoring network, constituent concentrations should be measured at sampling sites selected to satisfy the following design criteria:

- The total number of sites in the optimized monitoring network is fixed and based on a user-defined number of sampling sites to remove from the existing network (k). Selecting an appropriate value for k is a management decision and typically requires a cost-benefit analysis. To assist decision makers, optimal monitoring networks corresponding to the removal of 10, 20, 30, 40, and 50 sampling sites are included in this report (table 8).
- Plume maps interpolated for selected analytes using the *full-dataset* (data from sampling sites in the existing monitoring network) can be adequately reconstructed using the *reduced-dataset* (data from sampling sites in the reduced monitoring network). The interpolated plume map (prediction surface) estimated from the full-dataset is assumed to provide a realistic estimate of the concentration plume in the aquifer. For a given analyte, spatial accuracy is evaluated using the difference between prediction surfaces estimated using the reduced-dataset and full-dataset. For example, some sites are spatially redundant because their exclusion from the existing monitoring network would have little-to-no effect on predicting the spatial features of the plume.

Table 8. Hyperparameter values that control the optimization of the water-quality monitoring network, eastern Snake River Plain, Idaho.

Hyperparameter	Value
Multi-objective problem	
Number of sampling sites to remove from the existing monitoring network	10, 20, 30, 40, 50
Weighting coefficients on individual objective functions	
Weight on preserving the accuracy of the interpolated plume map (f_1)	10.0
Weight on preserving network coverage (f_2)	1.0
Weight on preserving long-term monotonic trends (f_3)	1.0
Weight on preserving temporal variance (f_4)	0.1
Islands parallel genetic algorithm	
Population size that is distributed evenly among islands	2,000
Number of islands	7
Migration operators	
Proportion of individuals that migrate between islands	10 percent
Number of generations at which exchange of individuals takes place, an epoch event	10
Genetic operators	
Probability of sexual recombination (crossover) between pairs of chromosomes	80 percent
Probability of mutation in a parent chromosome	10 percent
Number of chromosomes to survive to next generation (elitism)	7
Terminating conditions	
Maximum number of consecutive generations without any improvement in the best fitness value	25
Maximum number of generations	500

- Interpolation error (or uncertainty) from the application of kriging should be as small as possible. For example, removal of a sampling site from an area of the monitoring network where few wells exist typically would result in a large increase in the interpolation error (assessed using the normalized sum of kriging variance for all points in the interpolation grid); therefore, this well would not likely be removed.

- The long-term monotonic trend of analyte concentrations at a sampling site should be preserved across time. These trends are important for evaluating the effectiveness of remediation efforts at the INL. That is, retaining sampling sites with significant long-term trends (assessed using the percent change per year estimated from survival regression analysis) is necessary for effectively controlling water pollution in the aquifer. There is also an intrinsic value in preserving data that supports the construction of a survival regression (or trend) model.

- The variability of water-quality measurements should be preserved across time. Sampling sites with concentration time-series showing prominent seasonal fluctuation and long-term trends are important for understanding seasonal water-quality trends and controlling water pollution. Variability is assessed using the concentration range of predicted time-series data estimated from local regression analysis. The removal of sampling sites with small variability preserves the historical variability within the reduced-dataset. There also is an intrinsic value in preserving data that support the construction of a local regression model.

Each of these criteria (with the exception of first design criterion, the number of sites to remove) was converted to a mathematical metric, and the metrics were combined into a single multi-objective function that was used to identify a water-quality monitoring network satisfying the design criteria as a much as possible.

Multi-Objective Problem

The multi-objective problem is formulated as a single-objective optimization where a weighted combination of the design criteria is minimized. In mathematical terms, this is expressed as:

$$\boldsymbol{x} = \underset{x_1,x_2,\ldots,x_k \in \mathbb{Z} \,|\, 1 \le x \le n_e \text{ and } x \notin \boldsymbol{x_0}}{\arg\min} F(\boldsymbol{x}), \qquad (60)$$

where

k is the number of sampling sites to remove from the existing monitoring network,

$\boldsymbol{x}$ is the subset of k-sampling sites selected for removal from the existing monitoring network (a sequential whole number was assigned to each sampling site in the existing monitoring network and used to identify each site),

$\boldsymbol{x}_0$ is the subset of sampling sites that were not considered for removal,

n_e is the number of sampling sites in the existing monitoring network, and

F is the "fitness" function.

The minimum fitness value corresponds to the optimal monitoring network. The fitness function is dependent on the decision variables, a vector of integer values used to identify sampling sites in the existing monitoring network that will not be included in the reduced network ($\boldsymbol{x}$ in eq. 60). The purpose of the optimization solver is to find values of $\boldsymbol{x}$ that minimize the fitness value. The optimization problem was formulated such that a subset of sampling sites ($\boldsymbol{x}_0$ in eq. 60) were not considered for removal from the existing monitoring network. These are sampling sites that are located either in a multilevel completion well or in an open-hole completion well prior to being completed as a multilevel well. The scarcity and intrinsic value of depth-dependent water-quality data makes untenable an argument for the removal of these sites. Of the 153 sampling sites (or 133 monitoring wells) in the existing network, 31 sites (or 11 wells) were excluded from being considered for removal (table 1).

The fitness function is used to evaluate the desirability of a monitoring network design by representing a weighted combination of the chosen design-criteria metrics for multiple constituents. For valid combinations of decision variables, the fitness value is calculated using the function F, given by:

$$F(\boldsymbol{x}) = \sum_{i=1}^{n_c} \left\{ \begin{matrix} w_1 f_{1,i}(\boldsymbol{x}) + w_2 f_{2,i}(\boldsymbol{x}) \\ + w_3 f_{3,i}(\boldsymbol{x}) + w_4 f_{4,i}(\boldsymbol{x}) \end{matrix} \right\}, \qquad (61)$$

where

n_c is the number of selected constituents;

w is the weighting coefficient; and

f is the objective function, a unitless value.

All the design criteria except for the management decision of how many sites to remove from the existing monitoring network (k) are quantified by each of the objective functions: f_1, f_2, f_3, and f_4. The relative influence of each criterion may be established by varying the associated weights: w_1, w_2, w_3, and w_4.

Functions f_1 and f_2 are based on kriging estimates using the existing and reduced monitoring networks, where kriging predictions—and the uncertainty associated with these predictions—are made at nodes within the interpolation grid (app. 9). The function f_1 is a metric selected to minimize the root-mean-square deviation (RMSD) between the predicted concentrations ($\hat{C}_0$ in eq. 57) from kriging of observations in the existing and reduced monitoring networks. The RMSD is normalized by dividing it by the predicated range of concentrations based on kriging of observations in the existing

monitoring network. This normalized value will be close to 1. Individual objective functions were normalized to facilitate the comparison of different objective functions. That is, a transformation is used to ensure that objective functions have similar orders of magnitude.

The function f_1 is mathematically expressed as:

$$f_1(\boldsymbol{x}) = \frac{\sqrt{\frac{1}{n_n}\sum_{i=1}^{n_n}\left[\hat{C}_{e,i} - \hat{C}_{r,i}\right]^2}}{\max\left(\hat{C}_{e,1},\ldots,\hat{C}_{e,n_n}\right) - \min\left(\hat{C}_{e,1},\ldots,\hat{C}_{e,n_n}\right)}, \quad (62)$$

where

n_n is the number of active nodes in the interpolation grid,

$\hat{C}_{e,i}$ is the predicted concentration at node i in the interpolation grid based on kriging of observations at all sites in the existing monitoring network, and

$\hat{C}_{r,i}$ is the predicted concentration at node i based on kriging of observations at sites constituting the reduced monitoring network.

Removal of sites with small differences between measured and estimated values decreases the normalized RMSD more than removing sites with large differences.

The function f_2 is the metric selected to minimize the uncertainty of predicted concentrations at nodes in the interpolation grid (σ_0^2 in eq. 55), and is defined as the ratio between the sum of kriging variance based on kriging of observations at sites constituting the reduced monitoring network (σ^2_r) and the existing monitoring network (σ^2_e), given by:

$$f_2(\boldsymbol{x}) = \frac{\sum_{i=1}^{n_n}\sigma^2_{r,i}}{\sum_{i=1}^{n_n}\sigma^2_{e,i}}. \quad (63)$$

Because kriging variance at each node i depends on the proximity of nearby sampling sites, removal of sites from regions that have sparser data increases kriging variance more than at nodes that are close to other supporting data.

The function f_3 is the metric selected to preserve the long-term monotonic trends estimated using survival regression analysis, and is defined as

$$f_3(\boldsymbol{x}) = \max\left(\left|\left\{\Delta_{i\in\boldsymbol{x}}\right\}\right|\right), \quad (64)$$

where

$\Delta_{i\in\boldsymbol{x}}$ is the fractional change per year of constituent concentration (Δ in eq. 31) at site i of the subset of sites selected for removal from the existing monitoring network.

Removing sites where a trend could not be established, or the slope was small, preserves the significant long-term trends in the reduced monitoring network.

The function f_4 is the metric selected to preserve the temporal variability of predicted concentrations using local regression analysis, and is defined as

$$f_4(\boldsymbol{x}) = \max\left(\tilde{\Delta}\hat{C}_{i\in\boldsymbol{x}}\right), \quad (65)$$

where

$\tilde{\Delta}\hat{C}_i$ is the difference between the highest and lowest predicated concentrations ($\Delta\hat{C}$ in eq. 21) at site i divided by the maximum $\Delta\hat{C}$ for all sites; that is:

$$\tilde{\Delta}\hat{C}_i = \frac{\Delta\hat{C}_i}{\max\left(\Delta\hat{C}_{i\in\{1,\ldots,n_e\}}\right)}. \quad (66)$$

Removing sites where a loess curve could not be established, or the temporal variability was small, preserves the temporal variability more than removing sites with large variability.

For multi-objective problems, identifying a single solution that simultaneously minimizes each objective function (eqs. 62–65) is almost impossible. That is, any single objective value can often only be improved by making at least one of the other objective values worse. Combining the individual objective functions into a single weighted-objective function is subjective, requiring that a decision maker provide the weights. The weighted multi-objective function also is ill-suited for determining tradeoffs among objective functions. Because the objective functions are simply weighted and added to produce a single fitness value, the function with the largest range dominates the solution to the optimization problem. A poor value for the objective function with the larger range makes the overall fitness much worse than a poor value for the function with the smaller range (Bentley and Wakefield, 1997; Fisher 2013).

Objective functions (eqs. 62–65) are formulated such that their function value is close to 1 and dimensionless, thus making it easier to set the weighting coefficients such that they are significant relative to each other and relative to the objective function values (Marler and Arora, 2010, p. 857). For this study, preserving the accuracy of the interpolated plume map (f_1 in eq. 62) by setting w_1 equal to 10 was emphasized, and preserving temporal variability (f_4 as defined in eq. 65) by setting w_4 equal to 0.1 was de-emphasized. Weights associated with preserving network coverage (f_2 in eq. 63) and long-term monotonic trends (f_3 in eq. 64) were set equal to 1 (w_2=w_3=1) (table 8). No other weighting schema were considered to quantify the importance of the weighting choice.

Island Parallel Genetic Algorithm

A genetic algorithm (GA) (Holland, 1975) is used to find the best fitness value (that is, the minimum F value in eq. 60). GAs are adaptive heuristic search algorithms that mimic the mechanics of natural selection and survival of the fittest and are well suited for solving combinatorial optimization problems in which there is a large set of candidate solutions. Koza (1992, p. 18) provides the following definition of a GA:

> "The *genetic algorithm* is a highly parallel mathematical algorithm that transforms a set (*population*) of individual mathematical objects (typically fixed-length character strings patterned after *chromosome* strings), each with an associated *fitness* value, into a new population (i.e., the next *generation*) using operations patterned after the Darwinian principle of reproduction and survival of the fittest and after naturally occurring genetic operations (notably sexual recombination)."

In GA terminology, the array of decision variables in the optimization problem is called a "chromosome," which for the current problem defines the set of sampling sites being considered for removal from the existing monitoring network ($\boldsymbol{x}$ in eq. 60). A chromosome represents a single solution in the *solution space*, the collection of all possible solutions to the optimization problem. In this study, a chromosome describes a single design solution for the reduced water-quality monitoring network (that is, sampling sites to exclude from the existing monitoring network). Each design solution (referred to as a "individual") is assigned a fitness value (F in eq. 60), which summarizes how well the particular set of sites meets the overall design objective (described in section, "Design Criteria"). The GA operates on a collection of individuals referred to as a "population."

The GA problem is well suited for parallel computing because it requires a large number of independent calculations with negligible cost of data communication and synchronization among computer processors. To make use of parallel computing, the type of GA implemented in this report is a coarse-grained parallel GA or island parallel GA (ISLPGA) as described by Scrucca (2017). In an ISLPGA, a population of individuals (that is, candidate solutions) are partitioned into several subpopulations, with each subpopulation assigned a unique and separate "island." A subpopulation is allowed to evolve independently with the occasional exchange of the fittest individuals among islands (referred to as a "migration"). The periods of isolated evolution are called "epochs," with migration occurring at the end of each epoch (except the last) (Martin and others, 1997, p. C6.3:4). The exchange of individuals is used to introduce diversity in a subpopulation, thus avoiding convergence on a *local optimum*, a solution that is optimal within a neighboring set of candidate solutions rather than the best solution among all possible values. Independent GAs are run on the subpopulation of each island, and are assigned to a separate computer processor.

An ISLPGA sensitivity analysis was performed during the preliminary phase of this scientific investigation to determine the best values for parameters that control the spatial optimization of the water-quality monitoring network (table 8); the value of the *hyperparameters* can have a significant effect on the performance of the search. Recall that the hyperparameters are configuration variables that are external to the model and cannot be directly estimated from the data. The process of finding the best hyperparameter values (also known as hyperparameter tuning) for the ISLPGA was based on a trial-and-error approach in which cross validation of the kriging model was used as the primary performance metric to guide hyperparameter tuning, as well as the tradeoff between computational costs and predictive skill.

For this study, a population of 2,000 individuals was distributed evenly between seven separate islands. The length of the epochs (or migration interval) was set equal to 10 generations; that is, migration occurs in 10-generation intervals. The number of individuals that migrate between neighboring subpopulations was set equal to 10 percent of the population. Genetic operators were specified as follows: (1) an 80-percent probability of sexual recombination (crossover) between pairs of chromosomes; (2) a 10-percent probability of mutation in a parent chromosome; and (3) the number of chromosomes to survive to the next generation (elitism) set equal to 25. A general description of these genetic operators is provided by Fisher (2013). The termination of a GA search occurs after 25 generations of no improvement of the best individual within the subpopulation, or for cases when there is no-convergence, after 500 generations.

Optimal Sampling Sites

The water-quality monitoring network was spatially optimized five times, removing 10, 20, 30, 40, and 50 sampling sites, respectively, from the existing network. The number of sites removed from the network (k) has a significant effect on the best fitness value because of the dependence of individual objective functions (f_1,f_2,f_3,f_4) on k. Increasing k results in increased values of the objective function, which, in turn, linearly increases the best fitness value (weighted sum of objectives, eq. 61; table 9). The best fitness value at each epoch of the ISLPGA search (removing 10, 20, 30, 40, and 50 wells, respectively) is shown in figure 10. Recall that an epoch is the number of generations at which exchange of individuals between islands takes place and is equal to 10 generations in this report.

All the convergence curves may be characterized by fitness values that decrease rapidly, and then level off to become asymptotic towards the near-optimal solution (fig. 10). The number of epochs needed to satisfy the convergence (termination) condition increased as the number of removed wells increased (table 9; fig. 10). This should be expected given that the number of possible combinations increases as k is increased. For example, there are 1.4×10^{14} possible network

Table 9. Island parallel genetic algorithm searches summarized for optimal water-quality monitoring networks, eastern Snake River Plain, Idaho.

[**Number removed:** Number of wells removed from the existing monitoring network. Best fitness value: Smallest fitness value determined by the island parallel genetic algorithm (ISLPGA), a unitless value. **Number of epochs:** Number of periods of isolated evolution. An epoch is composed of 10 consecutive generations. **Computational time:** Time required to run the ISLPGA. **Range of objectives:** Range of weighted objective-function values over the ISLPGA search for each design criteria f_1,f_2,f_3, and f_4]

Number removed	Best fitness value	Number of epochs	Computational time (days)	Range of objectives			
				f_1	f_2	f_3	f_4
10	9.46	11	2.6	0.160	0.050	0.100	0.010
20	9.89	28	5.7	0.270	0.080	0.050	0.120
30	10.26	32	5.9	0.540	0.110	0.140	0.120
40	10.67	44	6.6	0.820	0.150	0.110	0.100
50	11.09	50	7.6	0.640	0.220	0.130	0.220

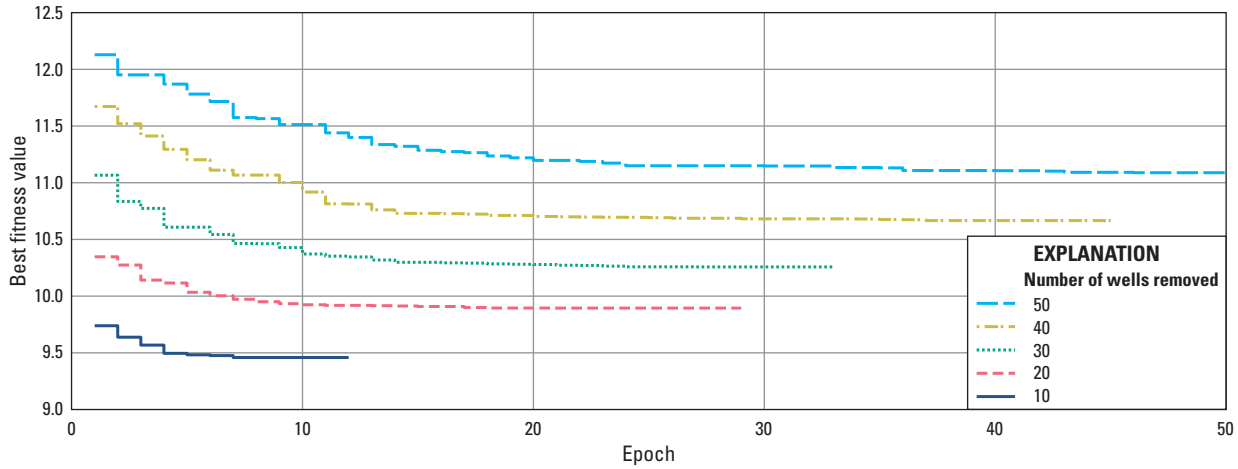


Figure 10. Best fitness value at each epoch of the island parallel genetic algorithm (removing 10, 20, 30, 40, and 50 wells, respectively) from the existing monitoring network, eastern Snake River Plain, Idaho.

configurations when k equals 10 and 5.3×10^{34} configurations when k equals 50. Computational time also increased with increased values of k and was strongly correlated (R^2=0.97) with the number of epochs needed for convergence (table 9). The computational times ranged from 2.6 days (removing 10 sites) to 7.6 days (removing 50 sites) (table 9). The optimization analysis was run in parallel using 7 of the 8 threads available on a 4-core Intel® Xeon® central processing unit E5-1620 v3 running at 3.5 gigahertz and with 32 gigabytes of random-access memory.

The range of weighted objective-function values in solution space indicates the relative influence of each design criterion in determining the optimal solution. For a given ISLPGA run and design criterion, a weighted objective-function value is calculated at each epoch (app. 10); the range of these values is defined as the difference between the largest and smallest value. The range of each weighted objective-function value is given in table 9 for ISLPGA runs based on changing the number of wells to remove from the existing network. As indicated by their ranges, the relative influence of each design criterion on the solution can vary depending on the number of wells to remove.

As intended, the accuracy of the interpolated plume map (f_1) has the greatest control over the evolutionary search. Design criteria f_2,f_3, and f_4 have less control over evolution and only after f_1 has been minimized to its fullest possible extent. The relative influence of criteria f_2,f_3, and f_4 on the solution varies in order of importance. For example, the influence of criterion f_3 (range equal to 0.100) on the solution is greater than the influence of criterion f_4 (range equal to 0.010) when k equals 10. However, when k equals 20, the influence of criterion f_3 (range equal to 0.050) on the solution is less than the influence of criterion f_4 (range equal to 0.120; table 9). For this study, design criteria were divided into primary and

secondary criteria by varying their associated weights. The primary design criterion is f_1, and the secondary criteria are f_2, f_3, and f_4. The order of importance for criteria f_2, f_3, and f_4 was assumed irrelevant to the design of an optimal network.

Sampling sites (where each site represents an individual well) identified for removal from the existing monitoring network are given in table 10. Here, the number of times the site was identified for removal in the five ISLPGA runs (each run corresponding to an optimized monitoring network) is enclosed in parentheses and italicized. For example, 7 of the 10 wells selected for removal were identified in all five ISLPGA runs (wells CPP 2, CPP 4, TRA 4, USGS 109, USGS 121, USGS 125, and USGS 45), 1 of the 10 wells was identified in three of the ISLPGA runs (well USGS 126B), 1 of the 10 wells was identified in two of the ISLPGA runs (well PSTF TEST), and 1 of the 10 wells was identified only once (well TRA 1). Each ISLPGA run provides a unique solution that is entirely dependent on the number of wells to remove from the existing network. That is, the solutions are non-sequential; wells identified for removal in the ISLPGA run removing 10 wells are not required to be part of the solution for the GA run removing 20 wells. The relatively large value of times identified, however, indicates that a consistent group of wells provides little-to-no beneficial added information.

The location of optimal sampling sites that were identified for removal is shown in figure 11; appendix 1, fig. 1.3; and appendix 11. Maps showing the difference between the kriged prediction surface using the existing and reduced monitoring networks are provided in appendix 11. An example of the spatial distribution of the tritium concentration differences predicted from network configurations removing 0 and 50 wells is shown in figure 12 and appendix 11, fig. 11.9. As expected, the largest differences in concentration are located in the solute plume, an area that is difficult to accurately interpolate because of its steep concentration gradients. For all analytes, concentration differences typically are small, thus indicating the effectiveness of the ISLPGA at removing spatially redundant wells.

Throughout the study area, network coverage was adequately preserved in the reduced networks. For cases where two wells are separated by a small distance (less than 200 ft), the ISLPGA frequently identified one of these wells for removal, helping to confirm the ability of the algorithms to identify spatial redundancy. For example, well USGS 109 was selected for removal in all five ISLPGA searches and is located 187 ft from well USGS 137A, and well USGS 126B was selected for removal in three of the searches and is located 61 ft from well USGS 126A.

The efficiency of an optimal network design was quantified for each analyte using the following performance measures:

1. RMSD between the kriged concentrations estimated using the existing and reduced monitoring networks (ideally small).
2. The number of sites that have sufficient data for detecting a long-term monotonic trend using a survival regression analysis (ideally large).
3. The average long-term monotonic trend in percent change per year (ideally large).
4. The number of sites that have sufficient data for local regression analysis (ideally large).
5. The average concentration range of predicted values based on local regression analysis (ideally large).

A quantitative comparison among the existing (0 wells removed) and optimal network designs (removing 10, 20, 30, 40, and 50 wells) using these metrics is shown in table 11. As expected, and for each of the selected analytes, the magnitude of the RMSD increases with increasing numbers of wells removed, with the exception of sodium, where the RMSD for 40 wells removed (0.3175 mg/L) was less than the RMSD for 30 wells removed (0.4492 mg/L), and may indicate that the global optima were not determined for this ISLPGA run.

The number of sites with long-term monotonic trends in the reduced network decreases (or does not change) with increasing numbers of wells removed. For carbon tetrachloride, 1,1-dichloroethylene, 1,1,1-trichloroethane, and trichloroethylene, the trend was estimable in a relatively small number of wells in comparison to the other analytes—a set of wells that were always included in the optimized networks. For sodium, chloride, sulfate, nitrate, tritium, and strontium-90, the average trend (in percent change per year) typically increases with increasing numbers of wells removed. Recall that the maximum trend value for sites removed from the existing monitoring network is minimized in the optimization problem (eq. 64). Therefore, the evolutionary search always will be inclined to remove sites with small trend values.

The number of sites with sufficient data for determining a local regression model decreases with increasing numbers of wells removed, whereas for most analytes, the average concentration range of predicted values increases with increasing numbers of wells removed. Recall that the maximum of the normalized concentration range of predicted values for sites removed from the existing monitoring network is minimized in the optimization problem (eq. 65). Therefore, the evolutionary search always will be inclined to remove sites with a smaller predicted concentration range.

As indicated by the optimization results, no single network exists that simultaneously optimizes each of the objectives for all the selected constituent types. That is, any single objective value often can be improved only by degrading at least one of the objective values. Because analyte components of an individual objective function simply are added to produce a single objective value, improvements associated with one analyte component may come at the cost of one or more of the other analyte components. Analyte components are equally weighted within each objective function (eqs. 62–65) to avoid preferencing one or more analytes above the others.

Table 10. Wells identified for removal based on island parallel genetic algorithm searches, eastern Snake River Plain, Idaho.

[**Number removed:** Number of wells removed from the existing monitoring network. **Well name:** Local well name. The number of times the well site was identified for removal in each of the networks five genetic algorithm searches is enclosed in parentheses and italicized]

Number removed	Well name			
10	CPP 2 (*5*)	CPP 4 (*5*)	PSTF TEST (*2*)	TRA 1 (*1*)
	TRA 4 (5)	USGS 109 (5)	USGS 121 (5)	USGS 125 (5)
	USGS 126B (*3*)	[1]USGS 45 (*5*)		
20	ATOMIC CITY WELL 1 (*3*)	BADGING FACILITY (*4*)	CPP 2 (*5*)	CPP 4 (*5*)
	SITE 19 (*4*)	TRA 3 (*4*)	TRA 4 (*5*)	TRA DISP (*4*)
	USGS 104 (*4*)	USGS 107 (*4*)	USGS 109 (*5*)	USGS 121 (*5*)
	USGS 125 (*5*)	USGS 126B (*3*)	USGS 26 (*4*)	USGS 39 (*4*)
	[1]USGS 42 (*4*)	[1]USGS 45 (*5*)	USGS 79 (*4*)	USGS 97 (*4*)
30	ATOMIC CITY WELL 1 (*3*)	BADGING FACILITY (*4*)	CPP 2 (*5*)	CPP 4 (*5*)
	RWMC M12S (*3*)	SITE 19 (*4*)	SITE 4 (*3*)	SITE 9 (*3*)
	TRA 3 (*4*)	TRA 4 (*5*)	TRA DISP (*4*)	USGS 100 (*3*)
	USGS 104 (*4*)	USGS 107 (*4*)	USGS 109 (*5*)	USGS 116 (*3*)
	USGS 121 (*5*)	USGS 125 (*5*)	USGS 126A (*3*)	USGS 26 (*4*)
	USGS 35 (*2*)	USGS 39 (*4*)	[1]USGS 41 (*3*)	[1]USGS 42 (*4*)
	[1]USGS 45 (*5*)	USGS 58 (*3*)	USGS 7 (*2*)	USGS 76 (*2*)
	USGS 79 (*4*)	USGS 97 (*4*)		
40	ANP 9 (*2*)	ATOMIC CITY WELL 1 (*3*)	BADGING FACILITY (*4*)	CFA 2 (*2*)
	CPP 2 (*5*)	CPP 4 (*5*)	HIGHWAY 3 (*2*)	RIFLE RANGE (*2*)
	RWMC M12S (*3*)	SITE 19 (*4*)	SITE 4 (*3*)	SITE 9 (*3*)
	SPERT 1 (*2*)	TRA 3 (*4*)	TRA 4 (*5*)	TRA DISP (*4*)
	USGS 100 (*3*)	USGS 104 (*4*)	USGS 106 (*2*)	USGS 107 (*4*)
	USGS 109 (*5*)	USGS 116 (*3*)	USGS 121 (*5*)	USGS 125 (*5*)
	USGS 126A (*3*)	USGS 126B (*3*)	USGS 127 (*2*)	USGS 15 (*1*)
	USGS 26 (*4*)	[1]USGS 36 (*2*)	USGS 39 (*4*)	[1]USGS 41 (*3*)
	[1]USGS 42 (*4*)	[1]USGS 45 (*5*)	[1]USGS 46 (*2*)	USGS 58 (*3*)
	[1]USGS 59 (*2*)	USGS 7 (*2*)	USGS 79 (*4*)	USGS 97 (*4*)
50	ANP 9 (*2*)	BADGING FACILITY (*4*)	CFA 2 (*2*)	CPP 2 (*5*)
	CPP 4 (*5*)	HIGHWAY 3 (*2*)	PSTF TEST (*2*)	RIFLE RANGE (*2*)
	RWMC M12S (*3*)	SITE 14 (*1*)	SITE 19 (*4*)	SITE 4 (*3*)
	SITE 9 (*3*)	SPERT 1 (*2*)	TRA 3 (*4*)	TRA 4 (*5*)
	TRA DISP (*4*)	USGS 1 (*1*)	USGS 100 (*3*)	USGS 104 (*4*)
	USGS 106 (*2*)	USGS 107 (*4*)	USGS 109 (*5*)	USGS 110A (*1*)
	USGS 111 (*1*) [1]	[1]USGS 114 (*1*)	USGS 116 (*3*)	USGS 12 (*1*)
	USGS 121 (*5*)	USGS 125 (*5*)	USGS 126A (*3*)	USGS 127 (*2*)
	USGS 19 (*1*)	USGS 23 (*1*)	USGS 26 (*4*)	USGS 35 (*2*)
	[1]USGS 36 (*2*)	USGS 39 (*4*)	[1]USGS 41 (*3*)	[1]USGS 42 (*4*)
	[1]USGS 45 (*5*)	[1]USGS 46 (*2*)	[1]USGS 47 (*1*)	USGS 58 (*3*)
	[1]USGS 59 (*2*)	USGS 76 (*2*)	USGS 79 (*4*)	USGS 82 (*1*)
	[1]USGS 85 (*1*)	USGS 97 (*4*)		

[1]Maximum contaminant level exceeded for at least one of the selected analytes.

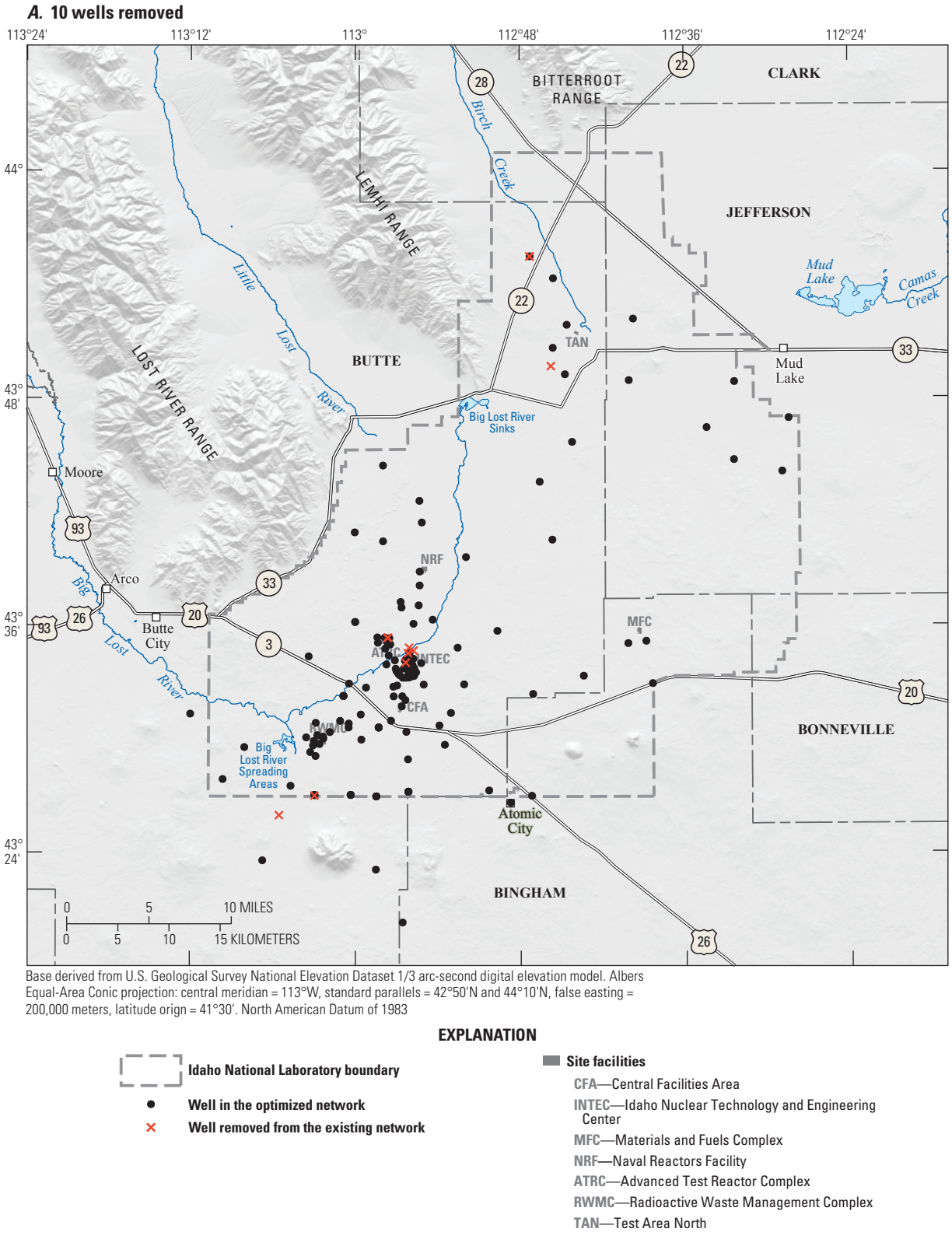
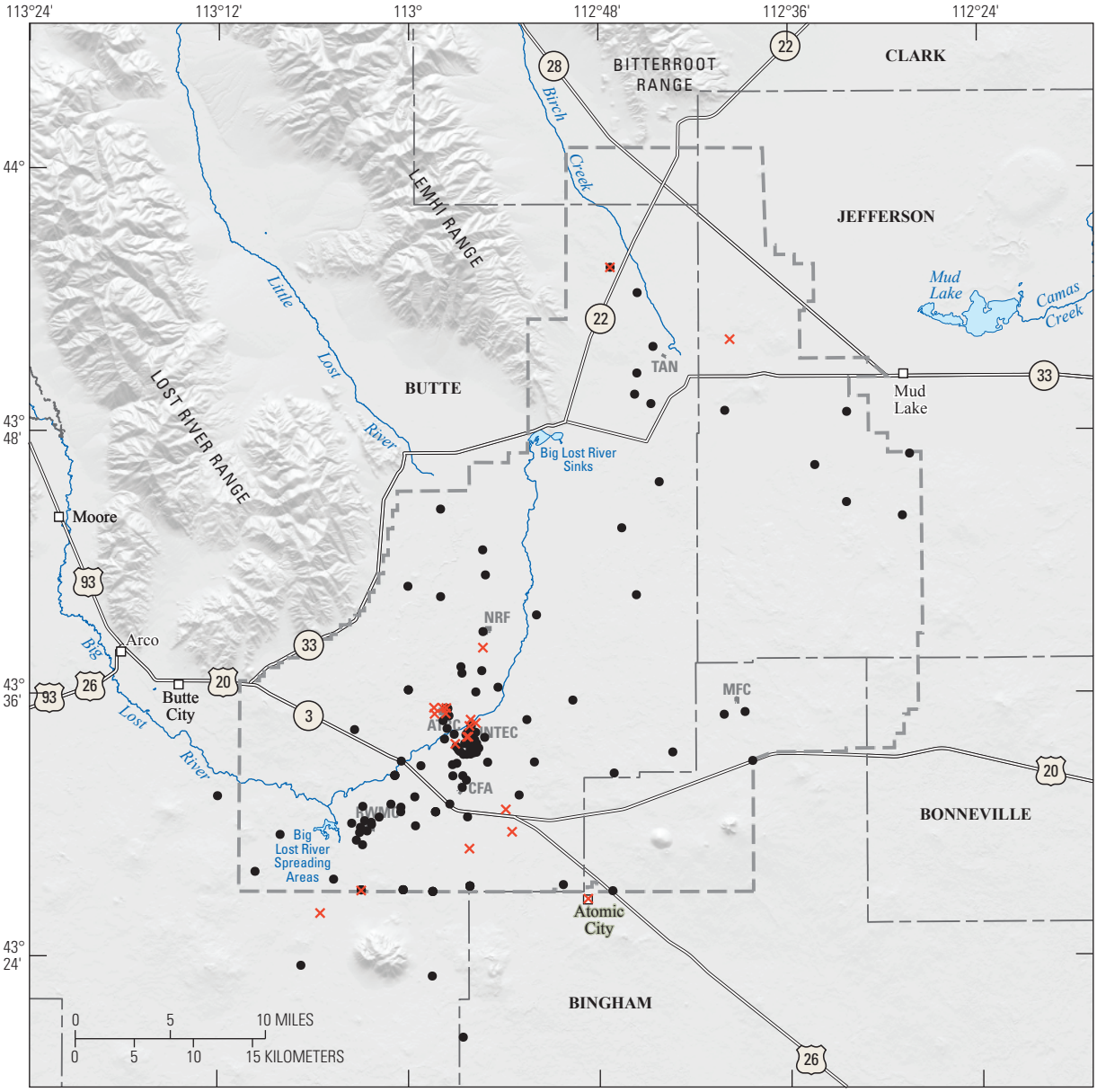


Figure 11. U.S. Geological Survey aquifer water-quality monitoring network after removing (*A*) 10, (*B*) 20, (*C*) 30, (*D*) 40, and (*E*) 50 optimally selected wells, Idaho National Laboratory and vicinity, Idaho.

***B.* 20 wells removed**



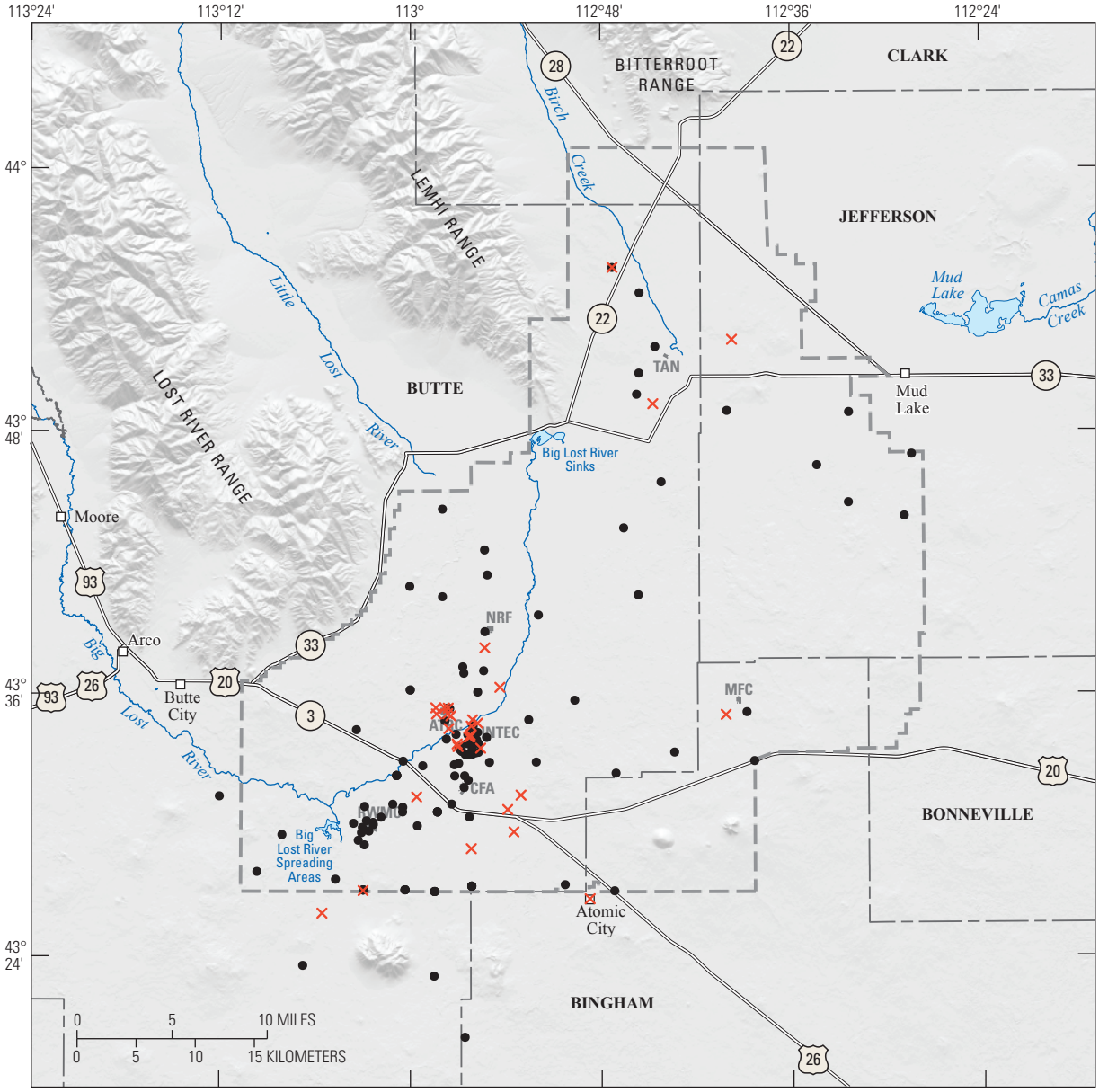
Base derived from U.S. Geological Survey National Elevation Dataset 1/3 arc-second digital elevation model. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

EXPLANATION

- **Idaho National Laboratory boundary**
- **Well in the optimized network**
- **Well removed from the existing network**
- **Site facilities**
  - **CFA**—Central Facilities Area
  - **INTEC**—Idaho Nuclear Technology and Engineering Center
  - **MFC**—Materials and Fuels Complex
  - **NRF**—Naval Reactors Facility
  - **ATRC**—Advanced Test Reactor Complex
  - **RWMC**—Radioactive Waste Management Complex
  - **TAN**—Test Area North

Figure 11.—Continued

C. **30 wells removed**



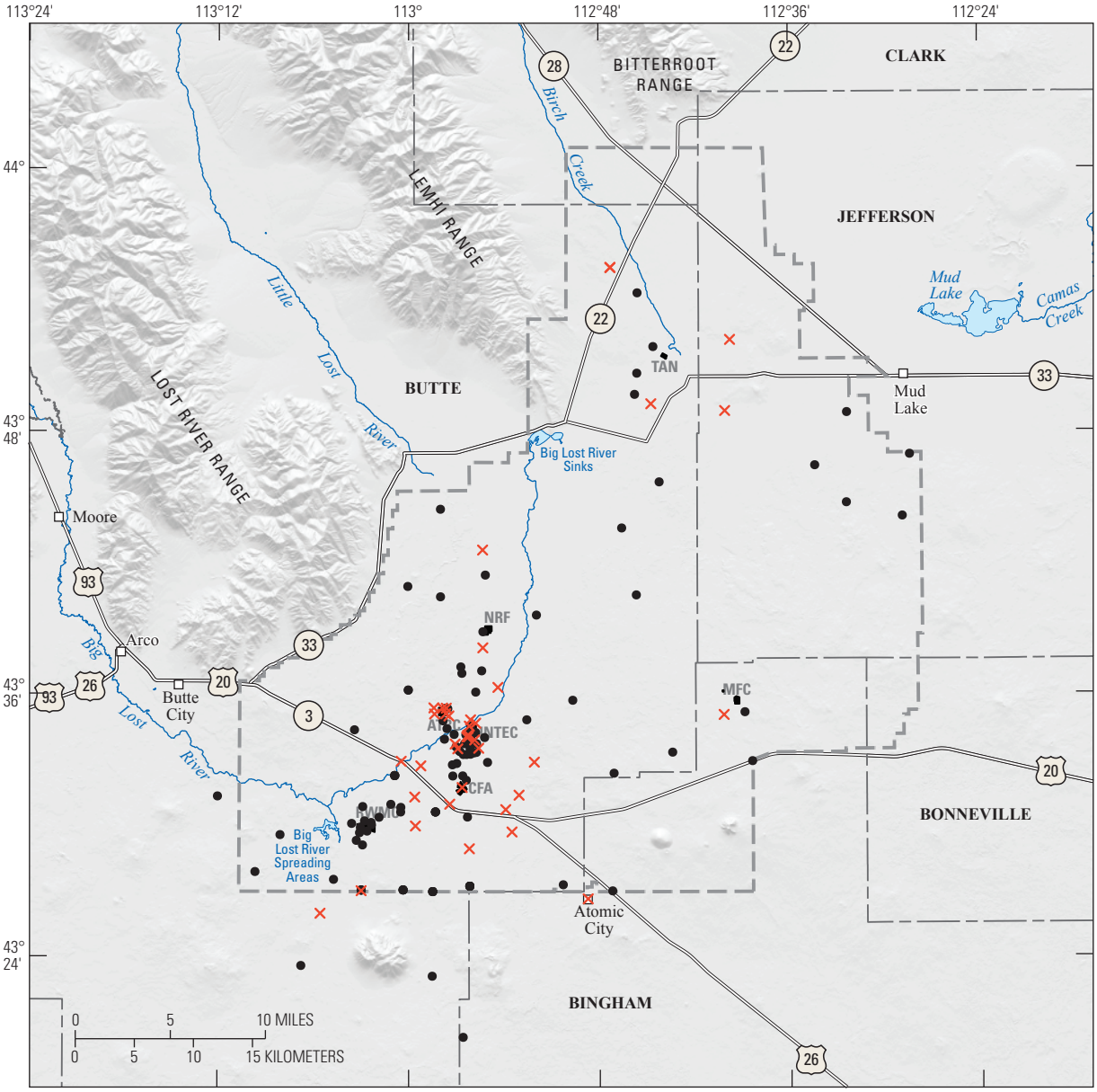
Base derived from U.S. Geological Survey National Elevation Dataset 1/3 arc-second digital elevation model. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

EXPLANATION

- **Idaho National Laboratory boundary**
- **Well in the optimized network**
- **Well removed from the existing network**
- **Site facilities**
  - **CFA**—Central Facilities Area
  - **INTEC**—Idaho Nuclear Technology and Engineering Center
  - **MFC**—Materials and Fuels Complex
  - **NRF**—Naval Reactors Facility
  - **ATRC**—Advanced Test Reactor Complex
  - **RWMC**—Radioactive Waste Management Complex
  - **TAN**—Test Area North

Figure 11.—Continued

D. **40 wells removed**



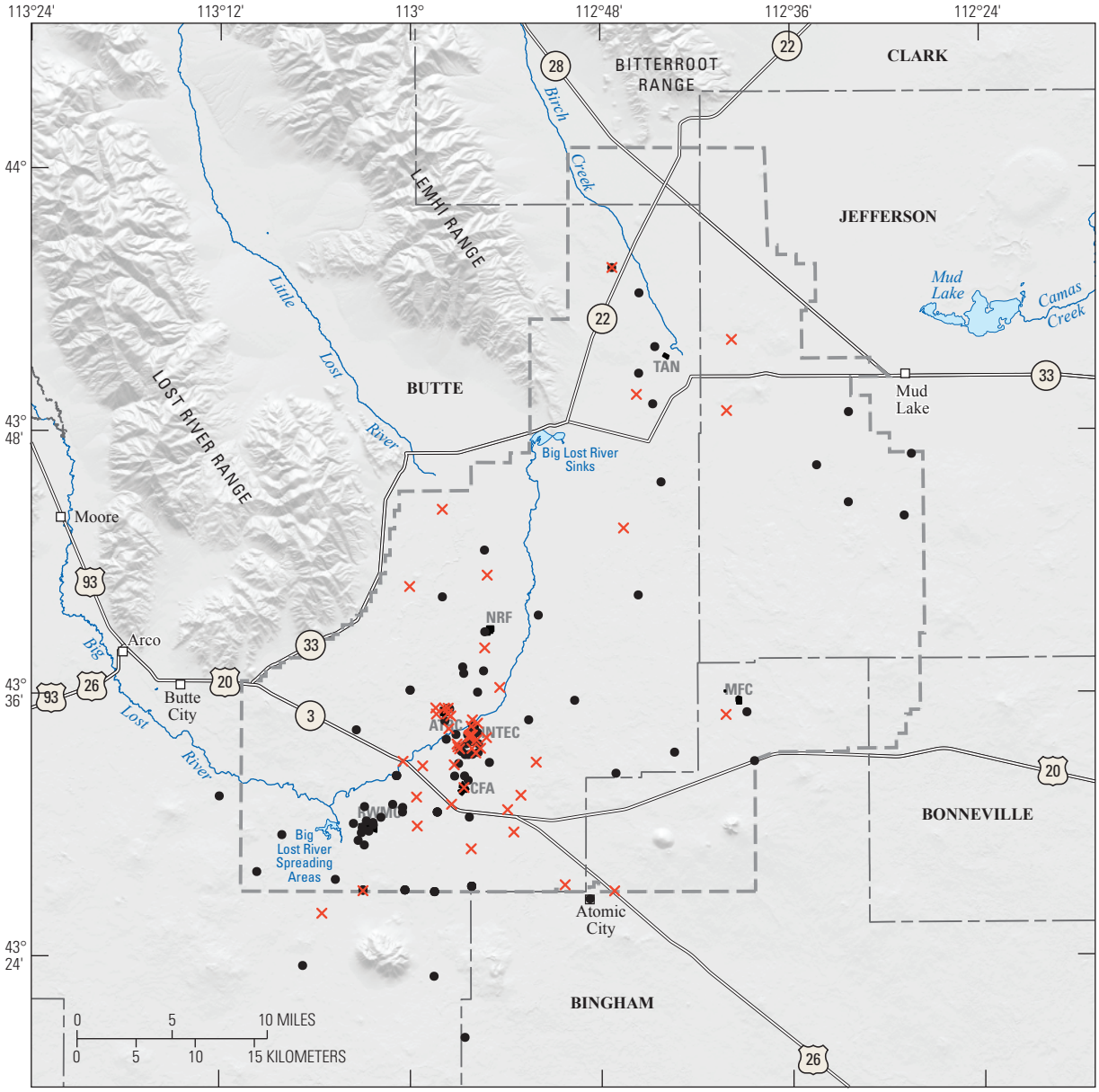
Base derived from U.S. Geological Survey National Elevation Dataset 1/3 arc-second digital elevation model. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

EXPLANATION

- **Idaho National Laboratory boundary**
- **Well in the optimized network**
- **Well removed from the existing network**
- **Site facilities**
  - **CFA**—Central Facilities Area
  - **INTEC**—Idaho Nuclear Technology and Engineering Center
  - **MFC**—Materials and Fuels Complex
  - **NRF**—Naval Reactors Facility
  - **ATRC**—Advanced Test Reactor Complex
  - **RWMC**—Radioactive Waste Management Complex
  - **TAN**—Test Area North

Figure 11.—Continued

E. **50 wells removed**



Base derived from U.S. Geological Survey National Elevation Dataset 1/3 arc-second digital elevation model. Albers Equal-Area Conic projection: central meridian = 113°W, standard parallels = 42°50'N and 44°10'N, false easting = 200,000 meters, latitude orign = 41°30'. North American Datum of 1983

EXPLANATION

- **Idaho National Laboratory boundary**
- **Well in the optimized network**
- **Well removed from the existing network**
- **Site facilities**
  - **CFA**—Central Facilities Area
  - **INTEC**—Idaho Nuclear Technology and Engineering Center
  - **MFC**—Materials and Fuels Complex
  - **NRF**—Naval Reactors Facility
  - **ATRC**—Advanced Test Reactor Complex
  - **RWMC**—Radioactive Waste Management Complex
  - **TAN**—Test Area North

Figure 11.—Continued

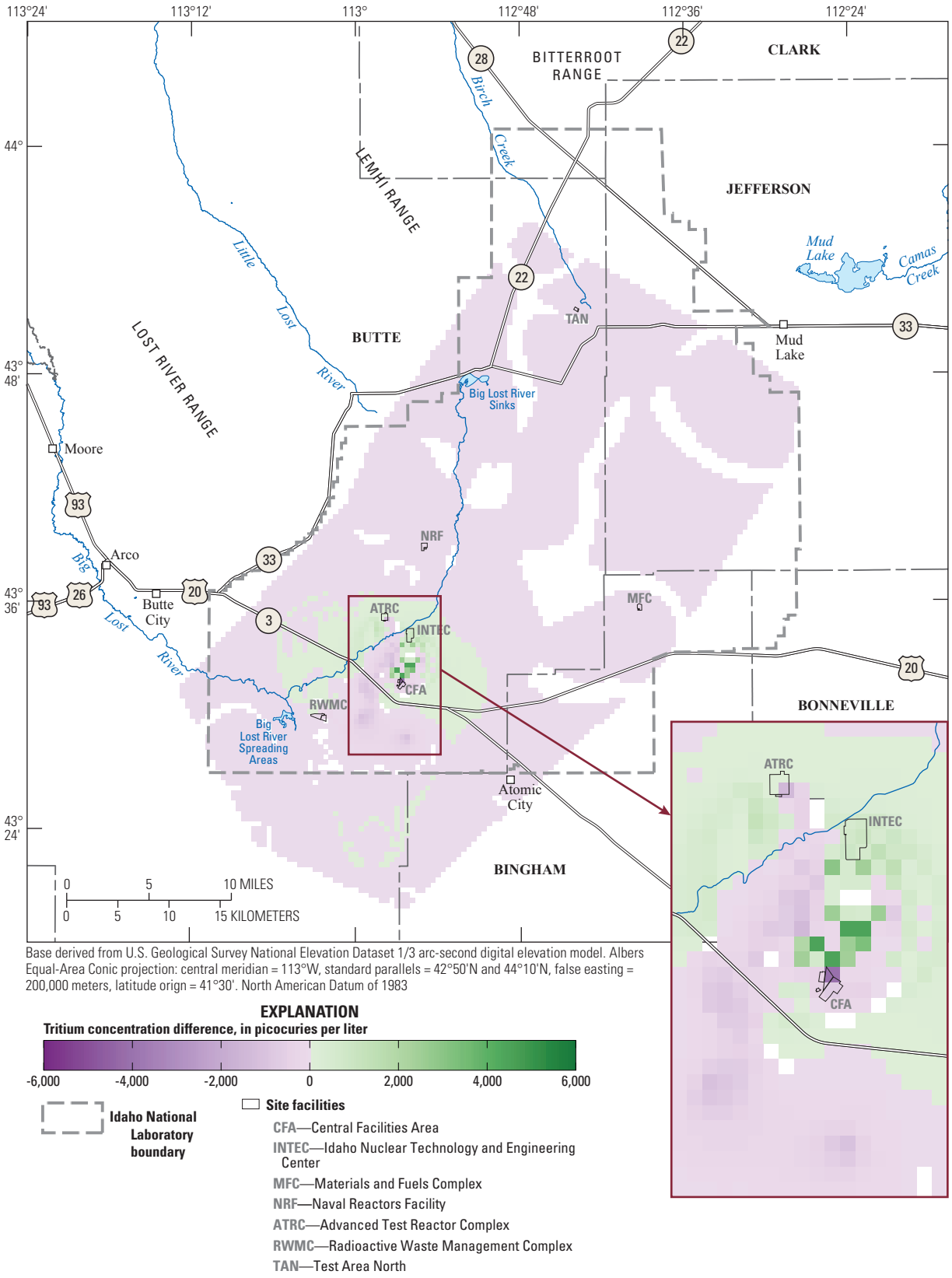


Figure 12. Difference between kriged tritium surfaces using the existing and reduced U.S. Geological Survey aquifer water-quality monitoring network after removing 50 optimally selected wells, Idaho National Laboratory and vicinity, Idaho.

Table 11. Comparison among optimized water-quality monitoring networks for selected constituents, Idaho National Laboratory and vicinity, Idaho.

[**Number removed:** Number of wells removed from the existing monitoring network. **Kriging RMSD:** Quantifies the performance of the quantile kriging analysis using the root-mean-square deviation (ideally small), a measure of the deviation between the kriged concentrations estimated using the existing and reduced monitoring networks. **Survival regression:** Long-term monotonic trend of predicted concentrations in an individual well based on survival-regression analysis. The trend (percent change per year) was estimable in *n* wells and summarized using the **mean** value. **Local regression:** Range of predicted concentrations in an individual well based on local-regression analysis. The range, in concentration space, was estimable in *n* wells and summarized using the **mean** value. **Abbreviations**: mg/L, milligrams per liter; µg/L micrograms per liter; N, nitrogen; pCi/L, picocuries per liter]

Analyte name	Number removed	Kriging RMSD	Survival regression		Local regression	
			n	mean	*n*	mean
Sodium, mg/L	0	0.0000	72	1.545	118	6.156
	10	0.2561	66	1.631	108	6.652
	20	0.2696	60	1.739	98	7.128
	30	0.4492	56	1.823	89	7.609
	40	0.3175	49	1.687	79	7.357
	50	0.4573	42	1.746	69	7.655
Chloride, mg/L	0	0.0000	87	2.072	123	18.880
	10	0.3573	78	2.195	113	20.251
	20	0.4845	72	2.322	103	21.822
	30	0.6982	68	2.346	94	22.945
	40	1.1059	61	2.304	84	23.114
	50	1.2945	52	2.418	74	24.224
Sulfate, mg/L	0	0.0000	54	1.051	73	8.853
	10	0.1373	50	1.094	69	9.230
	20	0.2770	45	1.101	62	9.431
	30	0.3376	39	1.104	54	9.810
	40	0.4146	37	1.128	50	9.995
	50	0.5014	31	1.121	43	10.241
Nitrate, mg/L as N	0	0.0000	59	1.759	105	0.629
	10	0.0203	55	1.800	97	0.667
	20	0.0269	51	1.885	90	0.687
	30	0.0424	49	1.894	84	0.701
	40	0.0469	42	1.778	76	0.697
	50	0.0550	34	1.925	65	0.664
Carbon tetrachloride, µg/L	0	0.0000	3	5.294	28	0.577
	10	0.0044	3	5.294	27	0.599
	20	0.0051	3	5.294	25	0.646
	30	0.0063	3	5.294	24	0.673
	40	0.0067	3	5.294	22	0.734
	50	0.0088	3	5.294	21	0.769
1,1-Dichloroethylene, µg/L	0	0.0000	1	3.011	22	0.082
	10	0.0002	1	3.011	21	0.083
	20	0.0004	1	3.011	20	0.084
	30	0.0005	1	3.011	19	0.086
	40	0.0004	1	3.011	17	0.093
	50	0.0005	1	3.011	16	0.095

Table 11. Comparison among optimized water-quality monitoring networks for selected constituents, Idaho National Laboratory and vicinity, Idaho.—Continued

[**Number removed:** Number of wells removed from the existing monitoring network. **Kriging RMSD:** Quantifies the performance of the quantile kriging analysis using the root-mean-square deviation (ideally small), a measure of the deviation between the kriged concentrations estimated using the existing and reduced monitoring networks. **Survival regression:** Long-term monotonic trend of predicted concentrations in an individual well based on survival-regression analysis. The trend (percent change per year) was estimable in *n* wells and summarized using the **mean** value. **Local regression:** Range of predicted concentrations in an individual well based on local-regression analysis. The range, in concentration space, was estimable in *n* wells and summarized using the **mean** value. **Abbreviations**: mg/L, milligrams per liter; µg/L micrograms per liter; N, nitrogen; pCi/L, picocuries per liter]

Analyte name	Number removed	Kriging RMSD	Survival regression		Local regression	
			n	mean	*n*	mean
1,1,1-Trichloroethane, µg/L	0	0.0000	8	6.055	27	0.165
	10	0.0009	8	6.055	26	0.170
	20	0.0018	8	6.055	25	0.174
	30	0.0032	8	6.055	24	0.179
	40	0.0024	8	6.055	22	0.190
	50	0.0044	8	6.055	21	0.197
Trichloroethylene, µg/L	0	0.0000	5	4.314	23	0.351
	10	0.0001	5	4.314	22	0.365
	20	0.0001	5	4.314	21	0.379
	30	0.0002	5	4.314	20	0.396
	40	0.0015	5	4.314	18	0.437
	50	0.0023	5	4.314	17	0.459
Tritium, pCi/L	0	0.0000	45	8.599	123	5,195.859
	10	37.8545	45	8.599	113	5,638.498
	20	86.2070	41	8.890	103	6,014.921
	30	144.4921	35	9.185	94	6,247.292
	40	147.9324	31	9.091	84	6,627.733
	50	164.3451	24	8.894	74	6,130.093
Strontium-90, pCi/L	0	0.0000	18	5.268	74	5.227
	10	0.0190	18	5.268	69	5.488
	20	0.0756	17	5.465	64	5.675
	30	0.0767	16	5.548	57	5.937
	40	0.1687	13	5.656	51	5.929
	50	0.1748	12	5.785	45	5.626

Temporal Optimization

Iterative Thinning

Temporal optimization was performed using iterative thinning. Originally proposed by Cameron (2004), iterative thinning examines whether the historical sampling frequency for a given well location and constituent may be reduced because of temporal redundancy in the sampling events. The temporal redundancy is assessed using a non-parametric statistical technique of iteratively removing (thinning) observations from time-series data until the temporal patterns in the original (or full) dataset no longer can be identified with confidence.

Iterative thinning is formulated here to identify an optimal *sampling interval* for constituent analysis of future water samples collected from a well. The sampling interval ($\tilde{\Delta}t$) is defined as the median number of days between consecutive and unique sampling dates, or:

$$\widetilde{\Delta t} = \text{median}\left\{\left|t_{i+1} - t_i\right|\right\}, \quad \text{for } i \in \left\{i \in \mathbb{Z} \middle| 1 \le i < n\right\} \tag{67}$$

where

t_i is the sampling time for observation i, and n is the total number of observations.

Given an unevenly spaced time series of constituent concentrations at a well, a subset of observations is selected for removal from the full-dataset in such a way that each observation has an equal probability of being chosen. The number of observations in the subset is expressed as a fraction s of the total number of observations n, and equal to $\lfloor sn \rfloor$. A loess curve is fitted to the reduced-dataset (described in section, "Local Regression Analysis"). Let c_j denote the loess-curve value evaluated at prediction point j (eq. 20). The ability of the reduced-dataset to capture the temporal patterns of trends in the full-dataset is evaluated by calculating the proportion of c_j values that fall inside the 90-percent confidence band (confidence intervals are defined in eq. 22) of the loess curve fitted to the full-dataset. A 90-percent confidence level was chosen, rather than the most commonly used 95-percent confidence level, to ensure that the width of the confidence band did not get so wide as to allow large changes in the loess curve as observations are iteratively removed. This is of particular concern in wells with small sample sizes (Cameron and Hunter, 2002, p. 644–645). The proportion p is defined as:

$$p = \frac{1}{m}\sum_{j=1}^{m}\begin{cases} 1, & \text{for } \hat{C}_{\ell,j} \le c_j \le \hat{C}_{u,j} \\ 0, & \text{for } c_j < \hat{C}_{\ell,j} \text{ or } c_j > \hat{C}_{u,j} \end{cases} \quad (68)$$

where

m is the total number of prediction points.

The iterative-thinning algorithm finds the largest fraction of observations that may be removed from the full-dataset while still resulting in a p value that is greater than 0.8. That is, at least 80 percent of the predicted points on the loess curve fitted to the reduced-dataset must fall within the 90-percent confidence band around the loess curve fitted to the full-dataset. A numerical threshold of 80 percent was chosen to ensure that the characteristics of the original trend were adequately preserved.

To minimize the influence of artifacts that may result from the preferential selection of certain parts of the sampling record, p values (eq. 68) are calculated for many different realizations of the reduced-dataset. For each realization, observations are selected using simple random sampling from the full-dataset. The average of these p values provides a better indication as to how well a reduced sampling interval would be able to preserve the temporal patterns of trends in the full-dataset (Cameron, 2004, p. 95). In mathematical terms, the estimation of s is expressed as:

$$s = \arg\max_{s\in\mathbb{R}} \bar{p} = \frac{1}{N}\sum_{k=1}^{N} p_k, \quad , \quad (69)$$

$$\text{subject to: } \bar{p} > p_{\min}$$

where

$\bar{p}$ is the mean p value calculated from loess curves fitted to different realizations of the reduced-dataset, removing $\lfloor sn \rfloor$ randomly selected observations;

N is the total number of realizations of the reduced-dataset; and

$p_{\min}$ is the minimum permitted $\bar{p}$ value.

In this study, the fraction s (eq. 69) is estimated using an 80-percent confidence band ($\bar{p} = 0.8$), and 500 realization (N=500). Five hundred was chosen as the number of realizations to provide sufficient information to determine the mean of exceeding the 80-percent threshold value. The optimal sampling interval $\widetilde{\Delta t}^*$ is calculated by dividing the historical sampling interval ($\tilde{\Delta} t$ in eq. 67) by the fraction of observations remaining in the optimally reduced-dataset:

$$\widetilde{\Delta t}^* = \frac{\widetilde{\Delta t}}{1-s}. \quad (70)$$

Two examples of the application of iterative thinning to unevenly spaced time-series data are shown in figure 13. These examples include an iterative thinning of tritium in well USGS 20 (fig 13*A*) and chromium in well USGS 38 (fig. 13*B*). Each of these figures shows the loess curves fitted to each of the 500 realizations of the reduced-dataset when the fraction of removed observations is at optimality. In both examples, the bulk of loess curves fitted to realizations of the reduced-dataset are able to preserve the temporal patterns of trends in the full-dataset.

Optimal Sampling Intervals

Iterative thinning was performed on each well-analyte combination with sufficient data for local regression analysis. Recall that the 90-percent confidence band around the loess curve fitted to the full-dataset is shown in appendix 5. For each well-analyte combination, an optimal sampling interval was determined and shown in appendix 12. The sampling reduction is represented as a percent change between median and optimized sampling intervals (Δ%) and was defined as:

$$\Delta\% = \frac{\widetilde{\Delta t}^* - \widetilde{\Delta t}}{\widetilde{\Delta t}} \times 100. \quad (71)$$

Sampling reduction values (app. 12) ranged from a minimum of 0 percent to a maximum of 3,100 percent, with a mean of 113 percent and a standard deviation of 131 percent. Given that multiple constituents are measured for in a water sample collected from a well, the best sampling frequency for a well is a management decision. A conservative approach to selecting the best sampling frequency for a well is to use the minimum of the optimal sampling interval values associated with the well. For example, the most conservative sampling

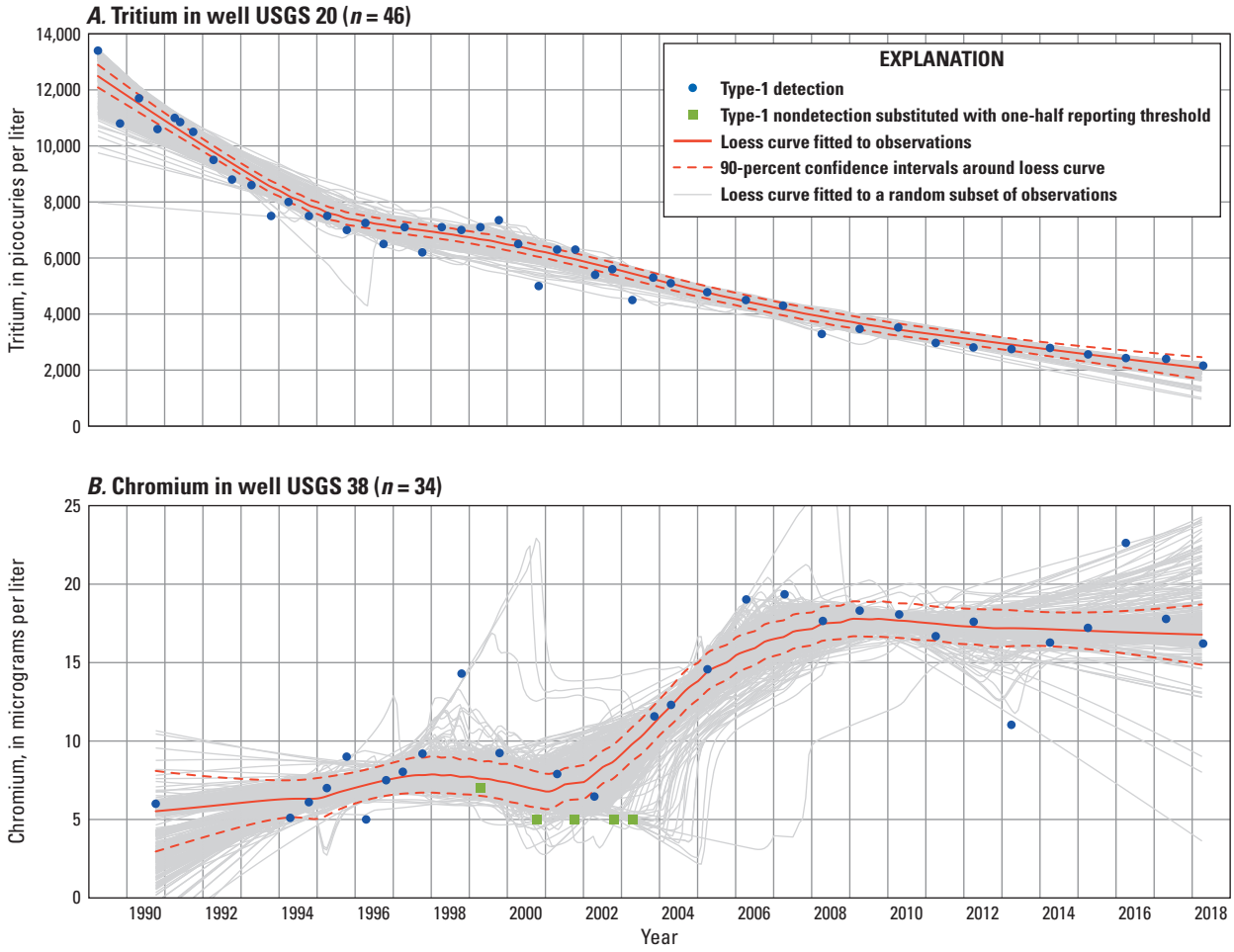


Figure 13. Loess curves computed in the final iteration of the iterative-thinning algorithm, with curves fitted to (*A*) tritium in well USGS 20 and (*B*) chromium in well USGS 38, Idaho National Laboratory, Idaho, 1989–2018. Water-quality data from groundwater samples collected on the same day were averaged. *n*, number of observations.

frequency for well ANP 6 is 1.4 years based on chloride having the smallest optimal sampling interval (1.4 years) relative to the other constituents (1.7–2.4 years).

Summary and Conclusions

Budgetary constraints and the high cost of long-term water-quality monitoring in the eastern Snake River Plain aquifer beneath the Idaho National Laboratory (INL) and vicinity, in southeastern Idaho, have necessitated a reduction in the number of wells (sampling sites) in the existing network. Since 1949, the U.S. Geological Survey, in cooperation with the U.S. Department of Energy, has maintained as many as 200 wells in the INL water-quality aquifer monitoring network. Long-term monitoring of water-quality data collected from these wells has provided essential information for delineating the movement of radiochemical and chemical wastes in the aquifer. The planning objective for the network is to reduce well monitoring redundancy by removing wells and decreasing sampling frequency at locations that add little or no information characterizing the water quality in the aquifer. To accomplish this objective, an evaluation and optimization of groundwater monitoring in the existing network was performed based on water-quality data collected at 153 sampling sites since January 1, 1989. Several options for optimally reduced networks were identified that minimize redundancy while retaining sufficient data to reliably characterize water-quality conditions in the aquifer. Spatial and temporal redundancy were examined using two different approaches; that is, the spatial and temporal components of the optimization were performed separately.

In the spatial optimization, the quality of a water-quality monitoring network design was evaluated using a single weighted-objective function that combines the following individual objective functions: (1) minimizing the interpolation error to ensure that the best spatial coverage is retained in the reduced-monitoring network; (2) safeguarding against the removal of sites with significant long-term trends that may be useful for evaluating the effectiveness of remediation efforts at

the INL; and (3) safeguarding against the removal of sites with repeated sampling at regular intervals over multiple years, and showing large variability in analyte concentrations, so as to preserve the long-term history of the sampling program.

As indicated by the optimization results, no single network exists that simultaneously optimizes each of the individual objective functions for all the selected constituent types. That is, the optimized network represents a tradeoff among sometimes conflicting objectives. Any single objective value may be improved only by degrading at least one of the other objective values. Because analyte components of an individual objective function simply are added to produce a single objective value, improvements associated with one analyte component generally came at the cost of one or more of the other analyte components.

A constituent was selected for inclusion in the spatial optimization problem when the observations were sufficient to (1) establish a two-range variability model, (2) classify at least one concentration time series as a continuous record block, and (3) make a prediction using the quantile-kriging interpolation method. The selected constituents include sodium, chloride, sulfate, nitrate, carbon tetrachloride, 1,1-dichloroethylene, 1,1,1-trichloroethane, trichloroethylene, tritium, strontium-90, and plutonium-238.

Spatial optimization was performed using an island parallel genetic algorithm and statistical analysis to identify near-optimal network designs removing 10, 20, 30, 40, and 50 wells from the existing network. With this method, choosing a greater number of wells to remove results in greater cost savings and decreased accuracy of the average relative difference between interpolated maps of the reduced-dataset and the full-dataset. The reduced-networks were able to reproduce the spatial patterns of the average concentration plumes for the 10 selected constituents while preserving their long-term temporal trends at monitoring sites. The number of sites with sufficient data for determining a local regression model decreases with increasing numbers of wells removed, whereas for most analytes, the average concentration range of predicted values increases with increasing numbers of wells removed. The evolutionary search is inclined to remove sites with a smaller predicted concentration range.

Temporal optimization was used to identify reductions in sampling frequencies by minimizing the redundancy in sampling events. An iterative-thinning method was used to find an optimal sampling frequency for each constituent-well pair. Optimal frequencies indicate that for many of the wells, samples may be collected less frequently and still be able to characterize the concentration over time. The optimization results indicated that the sample-collection interval may be increased by an of average of 273 days owing to temporal redundancy.

Acknowledgments

The authors gratefully acknowledge efforts by the U.S. Geological Survey Idaho National Laboratory hydrologic technicians for sample collection.

References Cited

Ackerman, D.J., 1991, Transmissivity of the Snake River Plain aquifer at the Idaho National Engineering Laboratory, Idaho: U.S. Geological Survey Water-Resources Investigations Report 91–4058 (DOE/ID–22097), 35 p., https://doi.org/10.3133/wri914058.

Ackerman, D.J., Rattray, G.W., Rousseau, J.P., Davis, L.C., and Orr, B.R., 2006, A conceptual model of ground-water flow in the eastern Snake River Plan aquifer at the Idaho National Laboratory and vicinity with implications for contaminant transport: U.S. Geological Survey Scientific Investigations Report 2006–5122, 62 p., https://pubs.usgs.gov/sir/2006/5122/.

Ackerman, D.J., Rousseau, J.P., Rattray, G.W., and Fisher, J.C., 2010, Steady-state and transient models of groundwater flow and advective transport, eastern Snake River Plain aquifer, Idaho National Laboratory and vicinity, Idaho: U.S. Geological Survey Scientific Investigations Report 2010–5123 (DOE/ID-22209), 220 p., accessed March 28, 2019, at https://pubs.usgs.gov/sir/2010/5123/.

Bai, J., and Perron, P., 2003, Computation and analysis of multiple structural change models: Journal of Applied Econometrics, v. 18, no. 1, p. 1–22. https://doi.org/10.1002/jae.659.

Bartholomay, R.C., 1993, Concentrations of tritium and strontium-90 in water from selected wells at the Idaho National Engineering Laboratory after purging one, two, and three borehole volumes: U.S. Geological Survey Water-Resources Investigations Report 93–4201 (DOE/ID-22111), 21 p., https://doi.org/10.3133/wri934201.

Bartholomay, R.C., 2013, Iodine-129 in the Snake River Plain aquifer at and near the Idaho National Laboratory, Idaho, 2010-12: U.S. Geological Survey Scientific Investigations Report 2013–5195, 22 p., https://doi.org/10.3133/sir20135195.

Bartholomay, R.C., Davis, L.C., Fisher, J.C., Tucker, B.J., and Raben, F.A., 2012, Water-quality characteristics and trends for selected sites at and near the Idaho National Laboratory, Idaho, 1949–2009: U.S. Geological Survey Scientific Investigations Report 2012–5169 (DOE/ID 22219), 68 p. plus appendixes, https://doi.org/10.3133/sir20125169.

Bartholomay, R.C., and Hall, L.F., 2016, Evaluation of background concentrations of selected chemical and radiochemical constituents in water from the eastern Snake River Plain aquifer at and near the Idaho National Laboratory, Idaho: U.S. Geological Survey Scientific Investigations Report 2016–5056 (DOE/ID-22237), 19 p., https://doi.org/10.3133/sir20165056.

Bartholomay, R.C., Knobel, L.L., and Rousseau, J.P., 2003, Field methods and quality-assurance plan for quality-of-water activities, U.S. Geological Survey, Idaho National Engineering and Engineering Laboratory, Idaho: U.S. Geological Survey Open-File Report 2003–42 (DOE/ID–22182), 45 p., https://doi.org/10.3133/ofr0342.

Bartholomay, R.C., Maimer, N.V., Rattray, G.W., and Fisher, J.C., 2017, An update of hydrologic conditions and distribution of selected constituents in water, eastern Snake River Plain aquifer and perched groundwater zones, Idaho National Laboratory, Idaho, emphasis 2012–15: U.S. Geological Survey Scientific Investigations Report 2017–5021 (DOE/ID-22242), 87 p., https://doi.org/10.3133/sir20175021.

Bartholomay, R.C., Maimer, N.V., and Wehnke, A.J., 2014, Field methods and quality-assurance plan for water-quality activities and water level measurements, U.S. Geological Survey, Idaho National Laboratory, Idaho: U.S. Geological Survey Open-File Report 2014–1146 (DOE/ID-22230), 66 p., https://doi.org/10.3133/ofr20141146.

Bartholomay, R.C., Tucker, B.J., Ackerman, D.J., and Liszewski, M.J., 1997, Hydrologic conditions and distribution of selected radiochemical and chemical constituents in water, Snake River Plain aquifer, Idaho National Engineering Laboratory, Idaho, 1992 through 1995: U.S. Geological Survey Water-Resources Investigations Report 97–4086 (DOE/ID–22137), 57 p., https://doi.org/10.3133/wri974086.

Bartholomay, R.C., Tucker, B.J., Davis, L.C., and Greene, M.R., 2000, Hydrologic conditions and distribution of selected constituents in water, Snake River Plain aquifer, Idaho National Engineering and Environmental Laboratory, Idaho, 1996 through 1998: U.S. Geological Survey Water-Resources Investigations Report 2000–4192 (DOE/ID–22167), 52 p., https://doi.org/10.3133/wri004192.

Bentley, P.J., and Wakefield, J.P., 1997, Finding acceptable solutions in the Pareto-optimal range using multiobjective genetic algorithms, *in* Chawdhry, P.K., and others, eds., Soft computing in engineering design and manufacturing: London, Springer-Verlag London Limited, p. 231–240.

Bodnar, L.Z., and Percival, D.R., eds., 1982, Analytical Chemistry Branch procedures manual—Radiological and Environmental Sciences Laboratory: U.S. Department of Energy Report IDO-12096, [variously paged].

Bossong, C.R., Karlinger, M.R., Troutman, B.M., and Vecchia, A.V., 1999, Overview and technical and practical aspects for use of geostatistics in hazardous-, toxic-, and radioactive-waste-site investigations: U.S. Geological Survey Water-Resources Investigations Report 98–4145, 70 p., https://doi.org/10.3133/wri984145.

Brent, R.P., 1973, Algorithms for minimization without derivates: Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 195 p.

Brownlee, J., 2017, What is the difference between a parameter and a hyperparameter? *in* Machine learning process: Machine Learning Mastery web page, accessed April 7, 2020, at https://machinelearningmastery.com/difference-between-a-parameter-and-a-hyperparameter/.

Busenberg, E., Plummer, L.N., and Bartholomay, R.C., 2001, Estimated age and source of the young fraction of ground water at the Idaho National Engineering and Environmental Laboratory: U.S. Geological Survey Water-Resources Investigations Report 2001–4265 (DOE/ID-22177), 144 p., https://doi.org/10.3133/wri014265.

Cameron, K., 2003, Optimization at AFP06 using improved GTS: Paper presented at 2003 AFCEE Technology Transfer Workshop, San Antonio, Texas, February 24, 1980.

Cameron, K., 2004, Better optimization of long-term monitoring networks: Bioremediation Journal, v. 8, nos. 3–4, p. 89–107.

Cameron, K., and Hunter, P., 2002, Using spatial models and kriging techniques to optimize long-term ground-water monitoring networks—A case study: Environmetrics, v. 13, nos. 5–6, p. 629–656, https://doi.org/10.1002/env.582.

Cecil, L.D., Welhan, J.A., Green, J.R., Frape, S.K., and Sudicky, E.R., 2000, Use of chlorine-36 to determine regional-scale aquifer dispersivity, eastern Snake River Plain aquifer, Idaho/USA, *in* Nuclear instruments and methods in physics research section B—Beam interactions with materials and atoms: Elsevier, v. 172, nos. 1–4, p. 679–687, https://doi.org/10.1016/S0168-583X(00)00216-0.

Childress, C.J.O., Foreman, W.T., Connor, B.F., and Maloney, T.J., 1999, New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory: U.S. Geological Survey Open-File Report 99–193, 19 p., https://doi.org/10.3133/ofr99193.

Claassen, H.C., 1982, Guidelines and techniques for obtaining water samples that accurately represent the water chemistry of an aquifer: U.S. Geological Survey Open-File Report 82–1024, 49 p., https://doi.org/10.3133/ofr821024.

Clark, I., 2010, Statistics or geostatistics? Sampling error or nugget effect?: Journal of the Southern African Institute of Mining and Metallurgy, v. 110, p. 307–312.

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: Journal of the American Statistical Association, v. 74, no. 368, p. 829–836.

Cleveland, W.S., and Devlin, S.J., 1988, Locally-weighted regression—An approach to regression analysis by local fitting: Journal of the American Statistical Association, v. 83, no. 403, p. 596–610, https://doi.org/10.1080/01621459.1988.10478639.

Cleveland, W.S., and Grosse, E., 1991, Computational methods for local regression: Statistics and Computing, v. 1, no. 1, p. 47–62.

Cleveland, W.S., Grosse, E., and Shyu, M.-J., 1992b, A package of C and Fortran routines for fitting local regression models: Unpublished manuscript, 54 p., accessed February 21, 2018, at https://www.netlib.org/a/.

Cleveland, W.S., Grosse, E., and Shyu, W.M., 1992a, Local regression models, chap. 8 *of* Chambers, J.M., and Hastie, T.J., eds., Statistical models in S: Pacific Grove, California, Wadsworth and Brooks/Cole Advanced Books and Software, p. 309–376.

Cressie, N.A.C., 1985, Fitting variogram models by weighted least squares: Journal of the International Association for Mathematical Geology, v. 17, no. 5, p. 563–586, https://doi.org/10.1007/BF01032109.

Cressie, N.A.C., 1993, Statistics for spatial data (revised ed.): New York, Wiley, 802 p.

Currie, L.A., 1984, Lower limit of detection—Definition and elaboration of a proposed position for radiological effluent and environmental measurements: U.S. Nuclear Regulatory Commission Report NUREG/CR-4007, 139 p.

Davis, L.C., Bartholomay, R.C., Fisher, J.C., and Maimer, N.V., 2015, Water-quality characteristics and trends for selected wells possibly influenced by wastewater disposal at the Idaho National Laboratory, Idaho, 1981–2012: U.S. Geological Survey Scientific Investigations Report 2015-5003 (DOE/ID-22233), 106 p., https://pubs.usgs.gov/sir/2015/5003/.

Davis, L.C., Bartholomay, R.C., and Rattray, G.W., 2013, An update of hydrologic conditions and distribution of selected constituents in water, eastern Snake River Plain aquifer and perched groundwater zones, Idaho National Laboratory, Idaho, emphasis 2009–11: USGS Scientific Investigations Report 2013-5214 (DOE/ID 22226), 89 p., https://doi.org/10.3133/sir20135214.

Duke, C.L., Roback, R.C., Reimus, P.W., Bowman, R.S., McLing, T.L., Baker, K.E., and Hull, L.C., 2007, Elucidation of flow and transport processes in a variably saturated system of interlayered sediment and fractured rock using tracer tests: Vadose Zone Journal, v. 6, no. 4, p. 855–867.

Faires, L.M., 1993, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determinations of metals in water by inductively coupled plasma-mass spectrometry: U.S. Geological Survey Open-File Report 92–634, 28 p., https://doi.org/10.3133/ofr92634.

Fienen, M.N., Doherty, J.E., Hunt, R.J., and Reeves, H.W., 2010, Using prediction uncertainty analysis to design hydrologic monitoring networks—Example applications from the Great Lakes water availability pilot project: U.S. Geological Survey Scientific Investigations Report 2010–5159, 44 p., https://pubs.usgs.gov/sir/2010/5159.

Fisher, J.C., 2013, Optimization of water-level monitoring networks in the eastern Snake River Plain aquifer using a kriging-based genetic algorithm method: U.S. Geological Survey Scientific Investigations Report 2013-5120 (DOE/ID-22224), 74 p., https://pubs.usgs.gov/sir/2013/5120/.

Fisher, J.C., 2020, inldata—Collection of datasets for the U.S. Geological Survey-Idaho National Laboratory aquifer monitoring networks: U.S. Geological Survey software release, R package, https://doi.org/10.5066/P9PP9UXZ.

Fisher, J.C., 2021, ObsNetQW—Assessment of a water-quality aquifer monitoring network: U.S. Geological Survey software release, R package, https://doi.org/10.5066/P9X71CSU.

Fishman, M.J., ed., 1993, v. 93–125. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determination of inorganic and organic constituents in water and fluvial sediments, U.S. Geological Survey Open-File Report, 217 p., https://doi.org/10.3133/ofr93125.

Fishman, M.J., and Friedman, L.C., eds., 1989, Methods for determination of inorganic substances in water and fluvial sediments (3d ed.): U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. A1, 545 p., accessed March 28, 2019, at https://pubs.usgs.gov/twri/twri5-a1/.

Freeman, J.V., Walters, S.J., and Campbell, M.J., 2008, How to display data: New York, Wiley, 109 p.

Garabedian, S.P., 1986, Application of a parameter-estimation technique to modeling the regional aquifer underlying the eastern Snake River Plain, Idaho: U.S. Geological Survey Water Supply Paper 2278, 60 p., https://doi.org/10.3133/wsp2278.

Goerlitz, D.F., and Brown, E., 1972, Methods for analysis of organic substances in water: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. A3, 40 p., https://doi.org/10.3133/twri05A3_1972.

Helsel, D.R., 2005, Nondetects and data analysis—Statistics for censored environmental data: Hoboken, New Jersey, Wiley, 250 p.

Holland, J.H., 1975, Adaptation in natural and artificial systems: Ann Arbor, Michigan, University of Michigan Press, 183 p.

Hurvich, C.M., Simonoff, J.S., and Tsai, C.L., 1998, Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion: Journal of the Royal Statistical Society. Series B. Methodological, v. 60, no. 2, p. 271–293.

Jacoby, W.G., 2000, Loess—A nonparametric, graphical tool for depicting relationships between variables: Electoral Studies, v. 19, no. 4, p. 577–613.

Johnson, V.M., Tuckfield, R.C., Ridley, M., and Anderson, R., 1996, Reducing the sampling frequency of groundwater monitoring wells: Environmental Science & Technology, v. 30, no. 1, p. 355–358. https://doi.org/10.1021/es950335u.

Journel, A.G., and Deutsch, C.V., 1997, Rank order geostatistics—A proposal for a unique coding and common processing of diverse data, *in* Baafi, E.Y., and Schofield, N.A., eds., Geostatistics Wollongton '96, Volume. 1: Dordrecht, The Netherlands, and Boston, Kluwer Academic, p. 174–187.

Juang, K.-W., Lee, D.-Y., and Ellsworth, T.R., 2001, Using rank-order geostatistics for spatial interpolation of highly skewed data in a heavy-metal contaminated site: Journal of Environmental Quality, v. 30, no. 3, p. 894–903.

Kalbfleisch, J.D., and Prentice, R.L., 2002, The statistical analysis of failure time data 2nd ed.: Hoboken, New Jersey, Wiley, 462 p.

Kaplan, E.L., and Meier, P., 1958, Nonparametric estimation from incomplete observations: Journal of the American Statistical Association, v. 53, no. 282, p. 457–481.

Kitanidis, P.K., 1997, Introduction to geostatistics—Applications to hydrogeology: New York, Cambridge University Press, p. 249.

Knobel, L.L., 2006, Evaluation of well-purging effects on water-quality results for samples collected from the eastern Snake River Plain aquifer underlying the Idaho National Laboratory, Idaho: U.S. Geological Survey Scientific Investigations Report 2006–5232 (DOE/ID-22200), 52 p., accessed March 28, 2019, at https://pubs.usgs.gov/sir/2006/5232/.

Knobel, L.L., Orr, B.R., and Cecil, L.D., 1992, Summary of background concentrations of selected radiochemical and chemical constituents in groundwater from the Snake River Plain aquifer, Idaho—Estimated from an analysis of previously published data: Journal of the Idaho Academy of Science, v. 28, no. 1, p. 48–61.

Knobel, L.L., Tucker, B.J., and Rousseau, J.P., 2008, Field methods and quality-assurance plan for quality-of-water activities, U.S. Geological Survey, Idaho National Laboratory, Idaho: U.S. Geological Survey Open-File Report 2008–1165 (DOE/ID–22206), 36 p., accessed March 28, 2019, at https://pubs.usgs.gov/of/2008/1165/.

Koza, J.R., 1992, Genetic programming—On the programming of computers by means of natural selection (complex adaptive systems): London, A Bradford Book, 819 p.

Mann, L.J., 1986, Hydraulic properties of rock units and chemical quality of water for INEL-1—A 10,365-foot deep test hole drilled at the Idaho National Engineering Laboratory, Idaho: U.S. Geological Survey Water-Resources Investigations Report 86–4020 (DOE/ID–22070), 23 p., https://doi.org/10.3133/wri864020.

Mann, L.J., 1996, Quality-assurance plan and field methods for quality-of-water activities, U.S. Geological Survey, Idaho National Engineering Laboratory, Idaho: U.S. Geological Survey Open-File Report 96–615 (DOE/ID-22132), 37 p., https://doi.org/10.3133/ofr96615.

Mann, L.J., and Beasley, T.M., 1994, Iodine-129 in the Snake River Plain aquifer at and near the Idaho National Engineering Laboratory, Idaho, 1990–1991: U.S. Geological Survey Water-Resources Investigations Report 94–4053 (DOE/ID-22115), 27 p., https://doi.org/10.3133/wri944053.

Marler, R.T., and Arora, J.S., 2010, The weighted sum method for multi-objective optimization—New insights: Structural and Multidisciplinary Optimization, v. 41, no. 6, p. 853–862, https://doi.org/10.1007/s00158-009-0460-7.

Martin, W.N., Lienig, J., and Cohoon, J.P., 1997, Island (migration) models—Evolutionary algorithms based on punctuated equilibria, chap. C6.3 *of* Bäck, T., Fogel, D.B., and Michalewicz, Z., eds., Handbook of evolutionary computation: Bristol, United Kingdom, Institute of Physics, p. 1–16.

Matérn, B., 1960, Spatial variation—Stochastic models and their applications to some problems in forest survey sampling investigations: Report of the Forest Research Institute of Sweden, v. 49, no. 5, 144 p.

McCurdy, D.E., Garbarino, J.R., and Mullin, A.H., 2008, Interpreting and reporting radiological water-quality data: U.S. Geological Survey Techniques and Methods, book 5, chap. B6, 33 p., https://doi.org/10.3133/tm5B6.

McKelvey, R.D., and Zavoina, W., 1975, A statistical model for the analysis of ordinal dependent variables: The Journal of Mathematical Sociology, v. 4, no. 1, p. 103–120.

Michel, R.L., 1989, Tritium deposition in the continental United States, 1953–1983: U.S. Geological Survey Water-Resources Investigations Report 89–4072, 46 p., https://doi.org/10.3133/wri894072.

Minasny, B., and McBratney, A.B., 2005, The Matérn function as a general model for soil variograms: Geoderma, v. 128, nos. 3–4, p. 192–207.

Moore, D., and McCabe, G., 2003, Introduction to the practice of statistics (4th ed.): New York, W.H. Freeman and Company, 438 p.

Mueller, D.K., Schertz, T.L., Martin, J.D., and Sandstrom, M.W., 2015, Design, analysis, and interpretation of field quality-control data for water-sampling projects: U.S. Geological Survey Techniques and Methods, book 4, chap. C4, 54 p., https://doi.org/10.3133/tm4C4.

Mueller, D.K., and Titus, C.J., 2005, Quality of nutrient data from streams and ground water sampled during water years 1992–2001: U.S. Geological Survey Scientific Investigations Report 2005–5106, 27 p., https://pubs.usgs.gov/sir/2005/5106/.

Nimmo, J.R., Perkins, K.S., Rose, P.E., Rousseau, J.P., Orr, B.R., Twining, B.V., and Anderson, S.R., 2002, Kilometer-scale rapid transport of naphthalene sulfonate tracer in the unsaturated zone at the Idaho National Engineering and Environmental Laboratory: Vadose Zone Journal, v. 1, no. 1, p. 89–101.

Nuzzo, R., 2014, Scientific method—Statistical errors: Nature, v. 506, no. 7487, p. 150–152.

Olmsted, F.H., 1962, Chemical and physical character of ground water in the National Reactor Testing Station, Idaho: U.S. Atomic Energy Commission, Idaho Operations Office Publication IDO-22043-USGS, 142 p., accessed March 28, 2019, at https://digital.library.unt.edu/ark:/67531/metadc100251/.

Orr, B.R., Cecil, L.D., and Knobel, L.L., 1991, Background concentrations of selected radionuclides, organic compounds, and chemical constituents in ground water in the vicinity of the Idaho National Engineering Laboratory: U.S. Geological Survey Water-Resources Investigations Report 91-4015 (DOE/ID-22094), 52 p., https://doi.org/10.3133/wri914015.

Pebesma, E.J., 2004, Multivariable geostatistics in S—The gstat package: Computers & Geosciences, v. 30, no. 7, p. 683–691.

Plummer, L.N., Rupert, M.G., Busenberg, E., and Schlosser, P., 2000, Age of irrigation water in ground water from the eastern Snake River Plain aquifer, South-central Idaho: Ground Water, v. 38, no. 2, p. 264–283, https://doi.org/10.1111/j.1745-6584.2000.tb00338.x.

R Core Team, 2019, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed on March 28, 2019, at https://www.r-project.org/.

Rattray, G.W., 2012, Evaluation of quality-control data collected by the U.S. Geological Survey for routine water-quality activities at the Idaho National Laboratory, Idaho, 1996–2001: U.S. Geological Survey Scientific Investigations Report 2012–5270 (DOE/ID-22222), 74 p., https://doi.org/10.3133/sir20125270.

Rattray, G.W., 2014, Evaluation of quality-control data collected by the U.S. Geological Survey for routine water-quality activities at the Idaho National Laboratory and vicinity, southeastern Idaho, 2002-08: USGS Scientific Investigations Report 2014–5027 (DOE/ID-22228), 66 p., https://doi.org/10.3133/sir20145027.

Reed, P.M., Ellsworth, T.R., and Minsker, B.S., 2004, Spatial interpolation methods for nonstationary plume data: Ground Water, v. 42, no. 2, p. 190–202.

Reed, P.M., Minsker, B.S., and Valocchi, A.J., 2000, Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation: Water Resources Research, v. 36, no. 12, p. 3731–3741.

Robertson, J.B., Schoen, R., and Barraclough, J.T., 1974, The influence of liquid waste disposal on the geochemistry of water at the National Reactor Testing Station, Idaho, 1952–1970: U.S. Geological Survey Open-File Report 73–238 (IDO–22053), 231 p., https://doi.org/10.3133/ofr73238.

Rossi, R.E., Mulla, D.J., Journel, A.G., and Franz, E.H., 1992, Geostatistical tools for modeling and interpreting ecological spatial dependence: Ecological Monographs, v. 62, no. 2, p. 277–314.

Scrucca, L., 2017, On some extensions to GA package—Hybrid optimisation, parallelisation and islands evolution: The R Journal, v. 9, no. 1, p. 187–206.

She, N., 1997, Analyzing censored water quality data using a non-parametric approach: Journal of the American Water Resources Association, v. 33, no. 3, p. 615–624.

Snyder, D.T., 2008, Estimated depth to ground water and configuration of the water table in the Portland, Oregon area: U.S. Geological Survey Scientific Investigations Report 2008–5059, 40 p., accessed April 23, 2019, at https://pubs.usgs.gov/sir/2008/5059/.

Stein, M.L., 1999, Interpolation of spatial data—Some theory for kriging: New York, Springer, 247 p.

Stevens, H.H., Ficke, J.F., and Smoot, G.F., 1975, Water temperature—Influential factors, field measurement, and data presentation: U.S. Geological Survey Techniques of Water-Resources Investigations, book 1, chap. D1, 65 p., accessed March 28, 2019, at https://pubs.usgs.gov/twri/twri1-d1/.

Thatcher, L.L., Janzer, V.J., and Edwards, K.W., 1977, Methods for determination of radioactive substances in water and fluvial sediments: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. A5, 95 p., accessed March 28, 2019, at https://pubs.usgs.gov/twri/twri5a5/.

Timme, P.J., 1995, National Water Quality Laboratory, 1995 services catalog: U.S. Geological Survey Open-File Report 95–352, p. 92., https://doi.org/10.3133/ofr95352.

U.S. Geological Survey, 1985, National water summary, 1984—Hydrologic events, selected water-quality trends, and ground-water resources: U.S. Geological Survey Water-Supply Paper 2275, 467 p., https://doi.org/10.3133/wsp2275.

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed June 11, 2019, at https://doi.org/10.5066/F7P55KJN.

U.S. Geological Survey, 2020, Idaho National Laboratory project office: U.S. Geological Survey Idaho National Laboratory project office web page, accessed in 2020, at https://www.usgs.gov/centers/id-water/science/idaho-national-laboratory-project-office.

U.S. Geological Survey, [variously dated], National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1–A9., accessed March 28, 2019, at https://water.usgs.gov/owq/FieldManual/.

Wegner, S.J., 1989, Selected quality assurance data for water samples collected by the U.S. Geological Survey, Idaho National Engineering Laboratory Idaho, 1980 to 1988: U.S. Geological Survey Water-Resources Investigations Report 89–4168 (DOE/ID–22085), 91 p., https://doi.org/10.3133/wri894168.

Wershaw, R.L., Fishman, M.J., Grabbe, R.R., and Lowe, L.E., eds., 1987, Methods for the determination of organic substances in water and fluvial sediments (revised ed.): U.S. Geological Survey Techniques of Water-Resource Investigation, book 5, chap. A3, 80 p., https://doi.org/10.3133/twri05A3.

Williams, L.M., 1996, Evaluation of quality assurance/quality control data collected by the U.S. Geological Survey for water-quality activities at the Idaho National Engineering Laboratory, Idaho, 1989 through 1993: U.S. Geological Survey Water-Resources Investigations Report 96–4148 (DOE/ID–22129), 116 p., https://doi.org/10.3133/wri964148.

Williams, L.M., 1997, Evaluation of quality assurance/quality control data collected by the U.S. Geological Survey for water-quality activities at the Idaho National Engineering Laboratory, Idaho, 1994 through 1995: U.S. Geological Survey Water-Resources Investigations Report 97–4058 (DOE/ID–22136), 87 p., https://doi.org/10.3133/wri974058.

Wood, W.W., 1976, Guidelines for collection and field analysis of ground-water samples for selected unstable constituents: U.S. Geological Survey Techniques of Water-Resource Investigation, book 1, chap. D2, 24 p., accessed on March 28, 2019, at https://pubs.usgs.gov/twri/twri1-d2/.

Appendix 1. Interactive Web Maps

Appendix 1 contains interactive maps showing the location of monitoring wells, long-term monotonic trends, optimal sampling sites, historical sampling frequencies, and optimal sampling intervals. Appendix 1 is an HTML file available for download at https://doi.org/10.3133/sir20215031.

Appendix 2. Software User Manual

Appendix 2 is a software user manual describing package datasets and processing programs (also known as functions in R). Appendix 2 is contained in an HTML file available for download at https://doi.org/10.3133/sir20215031.

Appendix 3. Graphs Showing Replicate-Paired Data and Variability Models

Appendix 3 shows two-range models for each of the selected constituents measured for in replicate samples collected from wells in the Idaho National Laboratory water-quality network, Idaho, 1989–2018. Appendix 3 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 4. Time-Series Graphs Showing Water-Quality Measurements

Appendix 4 shows time-series graphs with Type-1 and Type-2 data classification for selected constituents. Appendix 4 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 5. Time-Series Graphs Showing Local Regression Models

Appendix 5 shows time-series graphs with local temporal trends for selected constituents measured for in water samples from wells in the Idaho National Laboratory water-quality network, Idaho, 1989–2018. Appendix 5 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 6. Table Summarizing Survival Regression Models

Appendix 6 is a table that describes fitted survival regression models for selected constituents measured for in water samples collected from wells in the Idaho National Laboratory water-quality network, Idaho, 1989–2018. Appendix 6 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 7. Time-Series Graphs Showing Survival Regression Models

Appendix 7 shows time-series graphs for fitted survival regression models for selected constituents measured for in water samples collected from wells in the Idaho National Laboratory water-quality network, Idaho, 1989–2018. Appendix 7 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 8. Graphs Showing Variogram Models

Appendix 8 shows graphs for fitted theoretical variogram models for selected constituents measured for in water samples collected from wells in the Idaho National Laboratory water-quality network, Idaho, and temporally averaged during 1989–2018. Appendix 8 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 9. Maps Showing Kriging Estimates and Observations

Appendix 9 shows maps of kriging estimates and uncertainties for selected constituents measured for in water samples collected from wells in the Idaho National Laboratory water-quality aquifer monitoring network, Idaho, 1989–2018. Appendix 9 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 10. Graphs Showing Weighted Objective-Function Values

Appendix 10 shows maps of the weighted objective-function values throughout the Island Parallel Genetic Algorithm search—removing 10, 20, 30, 40, and 50 wells from the existing U.S. Geological Survey water-quality monitoring network at the Idaho National Laboratory, Idaho. Appendix 10 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 11. Maps Showing Difference Between Kriged Prediction Surfaces

Appendix 11 shows maps of the difference between the kriged prediction surface using the existing and reduced U.S. Geological Survey water-quality monitoring network at the Idaho National Laboratory and vicinity, Idaho. Appendix 11 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Appendix 12. Table Summarizing Reductions in Sampling Interval

Appendix 12 provides a table showing for each well-analyte combination an optimal sampling interval at the Idaho National Laboratory and vicinity, Idaho. Appendix 12 is an Adobe Acrobat® PDF file available for download at https://doi.org/10.3133/sir20215031.

Publishing support provided by the U.S. Geological Survey
Science Publishing Network, Tacoma Publishing Service Center

For more information concerning the research in this report, contact the
Director, Idaho Water Science Center
U.S. Geological Survey
230 Collins Road
Boise, Idaho 83702-4520
https://www.usgs.gov/centers/id-water

ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20215031