Prepared in cooperation with the Albuquerque Bernalillo County Water Utility Authority

Optimization Assessment of a Groundwater-Level Observation Network in the Middle Rio Grande Basin, New Mexico

Scientific Investigations Report 2020–5007
Version 2, December 2020

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

Optimization Assessment of a Groundwater-Level Observation Network in the Middle Rio Grande Basin, New Mexico

By Andre B. Ritchie and Jeff D. Pepin

Prepared in cooperation with the Albuquerque Bernalillo County Water Utility Authority

Scientific Investigations Report 2020–5007
Version 2, December 2020

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

U.S. Department of the Interior
DAVID BERNHARDT, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia
First release: 2020, online and in print
Revised: December 2020 (ver. 2), online and in print

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://www.usgs.gov/pubprod/.

Suggested citation:
Ritchie, A.B., and Pepin, J.D., 2020, Optimization assessment of a groundwater-level observation network in the Middle Rio Grande Basin, New Mexico (ver. 2, December 2020): U.S. Geological Survey Scientific Investigations Report 2020–5007, 113 p., https://doi.org/10.3133/sir20205007.

ISSN 2328-031X (print)
ISSN 2328-0328 (online)

Contents

Abstract 1
Introduction 1
Purpose and Scope 3
Description of Study Area 3
Background and Previous Investigations 5
ABCWUA Groundwater-Level Observation Network 6
Methods 22
Monitoring Sites Used for Optimization Assessment 22
Sources and Descriptions of Data 22
Data Accuracy 23
Spatial Assessment Methods 23
Background 23
Input Data Description 24
Implementation 24
Spatial Trend Modeling 24
Semivariogram Development 31
Kriged Surface Development and Assessment 31
Multiobjective Function 37
Genetic Algorithm 37
Genetic Algorithm Parameter Selection 42
Number of Sites Removed 43
Kriging Grid Resolution 43
Other Control Parameters 43
Temporal Assessment Methods 47
Background 47
Input Data Description 47
Implementation 47
Parameter Selection 50
Principal Component Analysis Methods 50
Background 50
Input Data Description 51
Implementation and Parameter Selection 51
Optimization Assessment of the Observation Network 62
Spatial Assessment Results 62
Weighted Objective Values 62
Genetic Algorithm Search Performance and Sites Identified for Removal 62
Temporal Assessment Results 79
Principal Component Analysis Results 100
Summary 109
Acknowledgments 110
References Cited 110

Figures

1. Map showing the Middle Rio Grande Basin, New Mexico, and Rio Grande Rift 2
2. Map showing the incorporated places within the study area of this report 4
3. Maps showing the locations of the observation network for *A*, monitoring wells; *B*, piezometers; and *C*, production wells 19
4. Maps showing the locations of sites used for the spatial assessment for the *A*, November 1, 2000, through March 31, 2001 (2001), and *B*, November 1, 2014, through March 31, 2015 (2015), time intervals 25
5. Map showing standard deviations of depth-to-groundwater measurements for the entire period of record used for the spatial assessment 29
6. Map showing uncertainties of groundwater-elevation measurements used for the spatial assessment 30
7. Graphs showing semivariogram analysis of groundwater-elevation residuals for the *A*, 2001 time interval and *B*, 2015 time interval 32
8. Maps showing the interpolated groundwater-elevation kriged surface for the *A*, 2001 time interval and *B*, 2015 time interval 33
9. Maps showing the standard errors from kriging groundwater elevation for the *A*, 2001 time interval and *B*, 2015 time interval 35
10. Maps showing the estimation error from leave-one-out cross validation for the *A*, 2001 time interval and *B*, 2015 time interval 38
11. Scatterplots showing *A*, measured and estimated groundwater elevations and *B*, estimation error and estimated groundwater elevations from leave-one-out cross validation for the 2001 time interval at the omitted sites 40
12. Scatterplots showing *A*, measured and estimated groundwater elevations and *B*, estimation error and estimated groundwater elevations from leave-one-out cross validation for the 2015 time interval at the omitted sites 41
13. Graphs showing best fitness value and computation time response curves of genetic algorithm sensitivity analysis for the control parameters: number of sites removed for the *A*, 2001 time interval and *B*, 2015 time interval; kriging grid resolution for the *C*, 2001 time interval and *D*, 2015 time interval; population size for the *E*, 2001 time interval and *F*, 2015 time interval; elitism rate for the *G*, 2001 time interval and *H*, 2015 time interval; crossover probability for the *I*, 2001 time interval and *J*, 2015 time interval; and mutation probability for the *K*, 2001 time interval and *L*, 2015 time interval 48
14. Graphs showing percentage of monitoring sites lacking available groundwater-level data for the principal component analysis when using *A*, daily mean; *B*, weekly median; and *C*, monthly median values 58
15. Scree plot showing the amount of dataset variance for which each principal component accounted 59
16. Graphs showing the *A*, within-group sum of squares and *B*, mean silhouette width metrics 61
17. Graphs showing response curves of genetic algorithm search performance when removing sites from the observation network: best fitness value for the *A*, 2001 time interval and *B*, 2015 time interval; root-mean-square deviation for the *C*, 2001 time interval and *D*, 2015 time interval; and percent local error for the *E*, 2001 time interval and *F*, 2015 time interval 63

18. Difference maps showing the kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2001 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites 64
19. Difference maps showing the kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2015 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites 68
20. Graphs showing response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites 80
21. Graph showing calculated silhouette width for each clustered monitoring site 101
22. Graphs showing principal component (PC) loadings plots for PCs 1 through 4 showing the final clustered solution: *A*, PC1 loadings versus PC2 loadings; *B*, PC1 loadings versus PC3 loadings; *C*, PC1 loadings versus PC4 loadings; *D*, PC2 loadings versus PC3 loadings; *E*, PC2 loadings versus PC4 loadings; and *F*, PC3 loadings versus PC4 loadings 102
23. Graphs showing clustered monthly median hydrographs plotted using *A*, the same fixed y-axis range and *B*, an independent y-axis range for all plots 104
24. Map showing the spatial distribution of clustered monitoring sites 108

Tables

1. Description of sites in the observation network 7
2. Groundwater elevations used for the spatial assessment 27
3. Control parameter values used for the genetic algorithm sensitivity analysis (base-case value) and spatial assessment (final value) 42
4. Sensitivity of the genetic algorithm to incremental changes in control parameters for the 2001 time interval and 2015 time interval 44
5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014 52
6. Range of weighted objective values for genetic algorithm runs performed by removing sites from the observation network for the 2001 time interval and 2015 time interval 72
7. Summary of genetic algorithm search performance when removing sites from the observation network for the 2001 time interval and 2015 time interval 73
8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval 74
9. Summary of mean interval between monitoring, in days, and median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data at each site 97

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square foot (ft^2)	929.0	square centimeter (cm^2)
square foot (ft^2)	0.09290	square meter (m^2)
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)

Datum

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

Vertical coordinate data collected and stored as National Geodetic Vertical Datum of 1929 (NGVD 29) have been converted to NAVD 88 for this publication. Conversion between NAVD 88 and NGVD 29 varies spatially and is accurate to within about plus or minus 2 centimeters (cm). The reader is directed to the National Geodetic Survey website for VERTCON at http://www.ngs.noaa.gov/TOOLS/Vertcon/vertcon.html for a detailed description of the height conversion methodology.

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

Elevation, as used in this report, refers to distance above the vertical datum.

Abbreviations

ABCWUA	Albuquerque Bernalillo County Water Utility Authority
CRAN	Comprehensive R Archive Network
GA	genetic algorithm
MSW	mean silhouette width
NA	not applicable
NAVD 88	North American Vertical Datum of 1988
PC	principal component
PCA	principal component analysis
PLE	percent local error
PVE	percent of variance explained
R^2	coefficient of determination
RMSD	root-mean-square deviation
RMSE	root-mean-square error
SJCDWP	San Juan-Chama Drinking Water Project
USGS	U.S. Geological Survey
WSS	within-group sum of squares
X	easterly coordinate
Y	northerly coordinate

Optimization Assessment of a Groundwater-Level Observation Network in the Middle Rio Grande Basin, New Mexico

By Andre B. Ritchie and Jeff D. Pepin

Abstract

The U.S. Geological Survey, in cooperation with the Albuquerque Bernalillo County Water Utility Authority (ABCWUA), measures groundwater levels continuously (hourly) and discretely (semiannually and annually) at a network of wells and piezometers (hereafter called the observation network) within the Middle Rio Grande Basin in central New Mexico. Groundwater levels that are measured in this observation network provide a long-term hydrologic dataset that is heavily relied upon to make water management decisions. The desire to upgrade and perform maintenance on this observation network initiated this study, which assesses the spatial and temporal importance of measurements towards optimization of monitoring the observation network to reduce or redirect monitoring costs. This report describes the methods and results of the optimization assessment of this observation network, which included separate spatial and temporal methodologies and an evaluation using principal component analysis (PCA).

Results from the spatial optimization assessment can be used to help identify observation network sites that do not significantly affect the generation of winter groundwater-elevation contour maps of the production zone. Results from the temporal optimization assessment and PCA can also be consulted when deciding which sites to remove from the network, especially for sites that are monitored more frequently than annually. Results from the temporal optimization assessment can be used to inform the minimum monitoring frequency at the observation network required to capture the trends shown in higher frequency monitoring. The PCA results distinguish spatially distributed characteristic water-level trends that can inform the management decisions that are made when using the spatial and temporal optimization assessment results. Reducing the temporal frequency or spatial density of monitoring is ultimately a management decision that depends on the amount of data loss or degradation that is deemed acceptable while still meeting the network objectives of the ABCWUA. This study can also serve as a starting point to a data gap analysis of local aquifer characteristics and help guide enhanced observation network design as needs arise or in advance of future water management decisions.

The results of the spatial optimization assessment indicate that as many as about 20 specified sites can be removed from the observation network with a relatively small loss in the ability to represent the kriged groundwater-elevation surfaces of the production zone that were generated by using median groundwater elevations for two periods: the 2001 time interval and 2015 time interval. This analysis also demonstrated the importance of wells at the margin of the study area and in areas where there are large hydrologic gradients. At many of the 47 hourly monitored sites analyzed in the temporal optimization assessment, temporal trends were well represented for at least one of the reduced monitoring frequencies tested, indicating that a reduced frequency may be sufficient to adequately characterize seasonal and long-term trends. PCA and k-means clustering analysis of the 15 hourly monitored sites that are screened within the production zone indicate that the sites can be categorized into four groups, or clusters, of differing groundwater-level hydrograph characteristics. Except for one cluster, all of the clusters have the potential to be well represented by fewer index monitoring sites.

Introduction

The Albuquerque Bernalillo County Water Utility Authority (ABCWUA) provides water and wastewater services within the Middle Rio Grande Basin (fig. 1) to more than 600,000 water users with a combination of surface water (through the San Juan-Chama Drinking Water Project [SJCDWP]) and groundwater (ABCWUA, 2016). Long-term hydrologic datasets are heavily relied upon to make management decisions involving the limited water resources and competing demands in the Middle Rio Grande Basin. As part of the cooperative Middle Rio Grande Basin monitoring program between the ABCWUA and the U.S. Geological Survey (USGS), groundwater levels are periodically

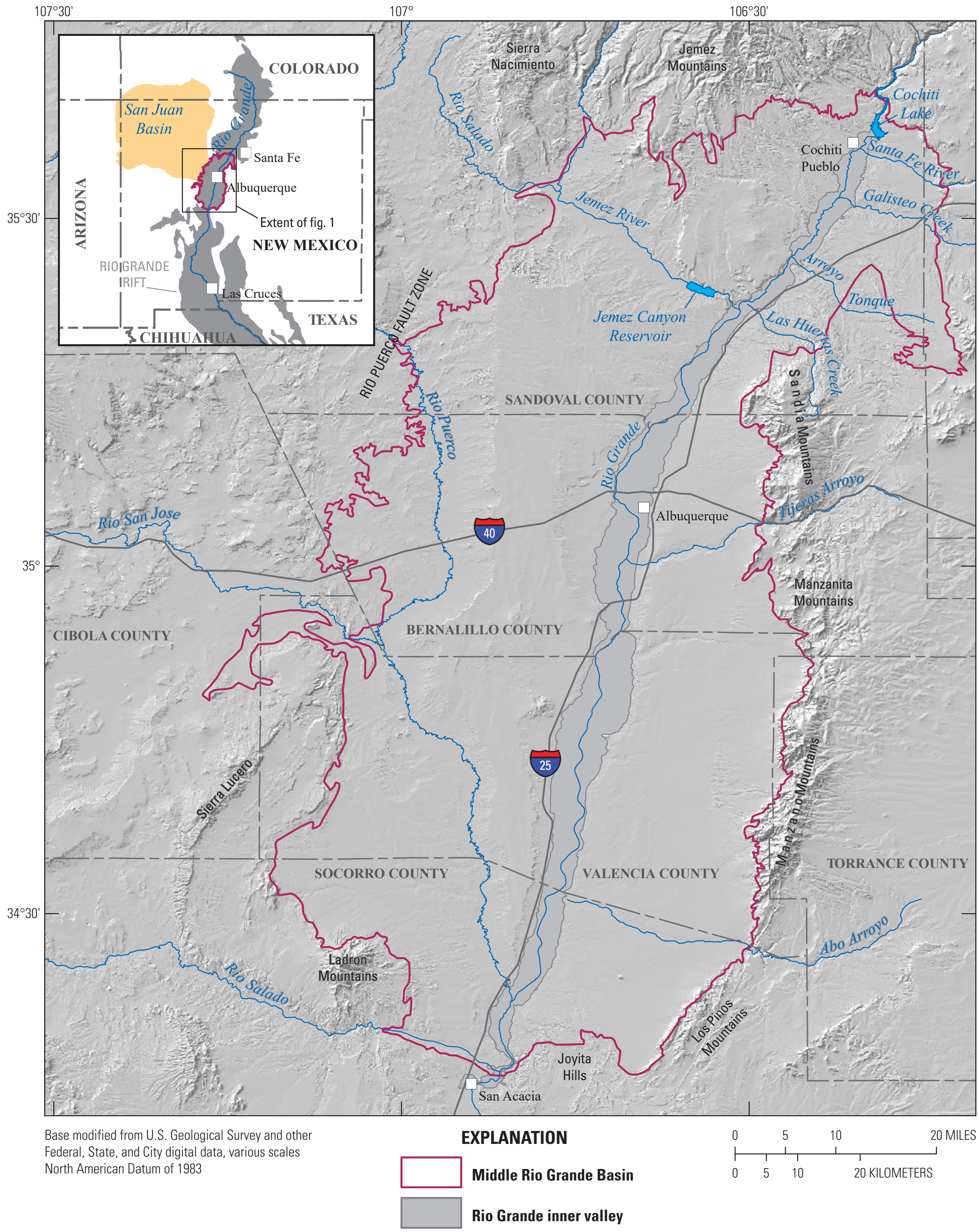


Figure 1. The Middle Rio Grande Basin, New Mexico, and Rio Grande Rift. Extent of Middle Rio Grande Basin from Plummer and others (2012). Extent of San Juan Basin modified from Kernodle (1996). Extent of Rio Grande Rift modified from Hudson and Grauch (2013). Extent of Rio Grande inner valley modified from McAda and Barroll (2002).

measured at a network of wells and piezometers (hereafter called the observation network) (Ritchie and Galanter, 2019). Groundwater levels measured in this observation network have been used to determine magnitudes and rates of groundwater-level change, calibrate groundwater flow models, map horizontal and vertical flow directions, and evaluate groundwater-level response to enhanced aquifer recharge efforts (Bexfield and Anderholm, 2002; Falk and others, 2011; Powell and McKean, 2014; Rice and others, 2014; Galanter and Curry, 2019).

The desire to upgrade and perform maintenance on the observation network initiated an evaluation of options for reducing the data-collection expense of the current observation network while preserving its spatial and temporal accuracy. The optimization assessment presented in this report includes separate spatial and temporal approaches and an evaluation of aquifer and well characteristics using principal component analysis (PCA). The spatial optimization assessment (hereafter called spatial assessment) aimed to determine the spatial importance of monitoring sites. The temporal optimization assessment (hereafter called temporal assessment) provides insight into sites that may be sampled less frequently without a significant loss of information on water-level trends. PCA results can be used to inform the results of the spatial and temporal assessments, as PCA identified characteristic groups of monitoring sites that represent unique properties that may be an additional consideration to be preserved.

Purpose and Scope

The purpose of this report is to present the results of spatial and temporal assessments and PCA of the observation network. This report serves to inform and enhance the ability of water managers to evaluate the consequences of reduced data fidelity for more cost-effective observation network configurations. Two monitoring reduction techniques are considered: (1) removing monitoring sites that have a minimal effect on the overall spatial characterization of the groundwater-elevation surface at a particular time and (2) decreasing the frequency of measurements at monitoring sites that indicate statistically similar temporal trends at lower measurement frequencies. The total number of sites to remove from the observation network or to measure less frequently is a management decision that can be informed by this study.

Description of Study Area

This study is focused on the City of Albuquerque and adjacent areas, New Mexico (called the Albuquerque area; fig. 2), within the Middle Rio Grande Basin of the Rio Grande Rift. The Rio Grande Rift is a north-south trending, tectonically active province of crustal extension that stretches from central Colorado into northern Mexico and consists of several fault-controlled basins (fig. 1) (Bartolino and Cole, 2002; Sanford and others, 2004). The Middle Rio Grande Basin, also known as the Albuquerque Basin, is one of these basins and is about 100 miles long and varies in width between 25 and 40 miles. This basin encompasses about 3,060 square miles within the Rio Grande Rift, with its boundary defined by the extent of consolidated and unconsolidated Cenozoic-age (less than 66 million years ago) deposits that are largely surrounded by uplifts (fig. 1) (Thorn and others, 1993; Plummer and others, 2012). The Middle Rio Grande Basin is also structurally bounded by the Rio Puerco fault zone, which is a transition zone between the San Juan Basin to the northwest and the Middle Rio Grande Basin (fig. 1) (Plummer and others, 2012; Jones and others, 2013).

The Middle Rio Grande Basin (fig. 1) is generally considered to be semiarid with precipitation primarily in the form of summer monsoonal rainstorms (Plummer and others, 2012). Mean annual precipitation recorded in the lower elevations within the basin is typically about 8 inches, whereas that in the surrounding highlands tends to be more than 14 inches. Mean annual snowfall varies greatly with elevation, ranging from 4 inches at Bernardo (about 12 miles north of San Acacia) to 59 inches in the Sandia Mountains (Plummer and others, 2012). The modern Rio Grande is the regional drainage and primary surface-water feature for the Middle Rio Grande Basin; the river originates in the mountains of southern Colorado and flows south through the Rio Grande Rift and the inner valley within the Middle Rio Grande Basin in New Mexico and then towards Texas and Mexico (fig. 1). Tributaries to the Rio Grande within the Middle Rio Grande Basin include the Santa Fe River, Galisteo Creek, Arroyo Tonque, Las Huertas Creek, Jemez River, Tijeras Arroyo, Abo Arroyo, Rio Puerco, and Rio Salado (fig. 1) (Plummer and others, 2012). The Rio Grande and its tributaries are important components of the regional water budget and are heavily relied upon to meet water user needs in portions of the Middle Rio Grande Basin (McAda and Barroll, 2002).

Aquifers are also heavily depended upon by water users throughout the Middle Rio Grande Basin. The aquifer system within the basin consists of unconsolidated to poorly consolidated Quaternary deposits, known as the Rio Grande alluvial aquifer, and older unconsolidated to moderately consolidated middle Tertiary- to Quaternary-age basin-fill deposits of the Santa Fe Group, known as the Santa Fe Group aquifer (Bartolino and Cole, 2002; Plummer and others, 2012; Rankin and others, 2013). The Rio Grande alluvial aquifer consists of channel, flood plain, terrace, and tributary deposits associated with the pre-flood-control-era (1971) flood plain of the Rio Grande (called the inner valley, fig. 1; McAda and Barroll, 2002) that form a thin but extensive aquifer that is as much as 120 feet thick beneath the inner valley (Bartolino and Cole, 2002; Rankin and others, 2013). The Santa Fe Group sediments range in thickness from about 3,000 to more than 14,000 feet and have been divided informally into upper, middle, and lower units on the basis of depositional environment and age (Hawley and Haase, 1992). The Santa Fe Group and Rio Grande alluvial aquifers

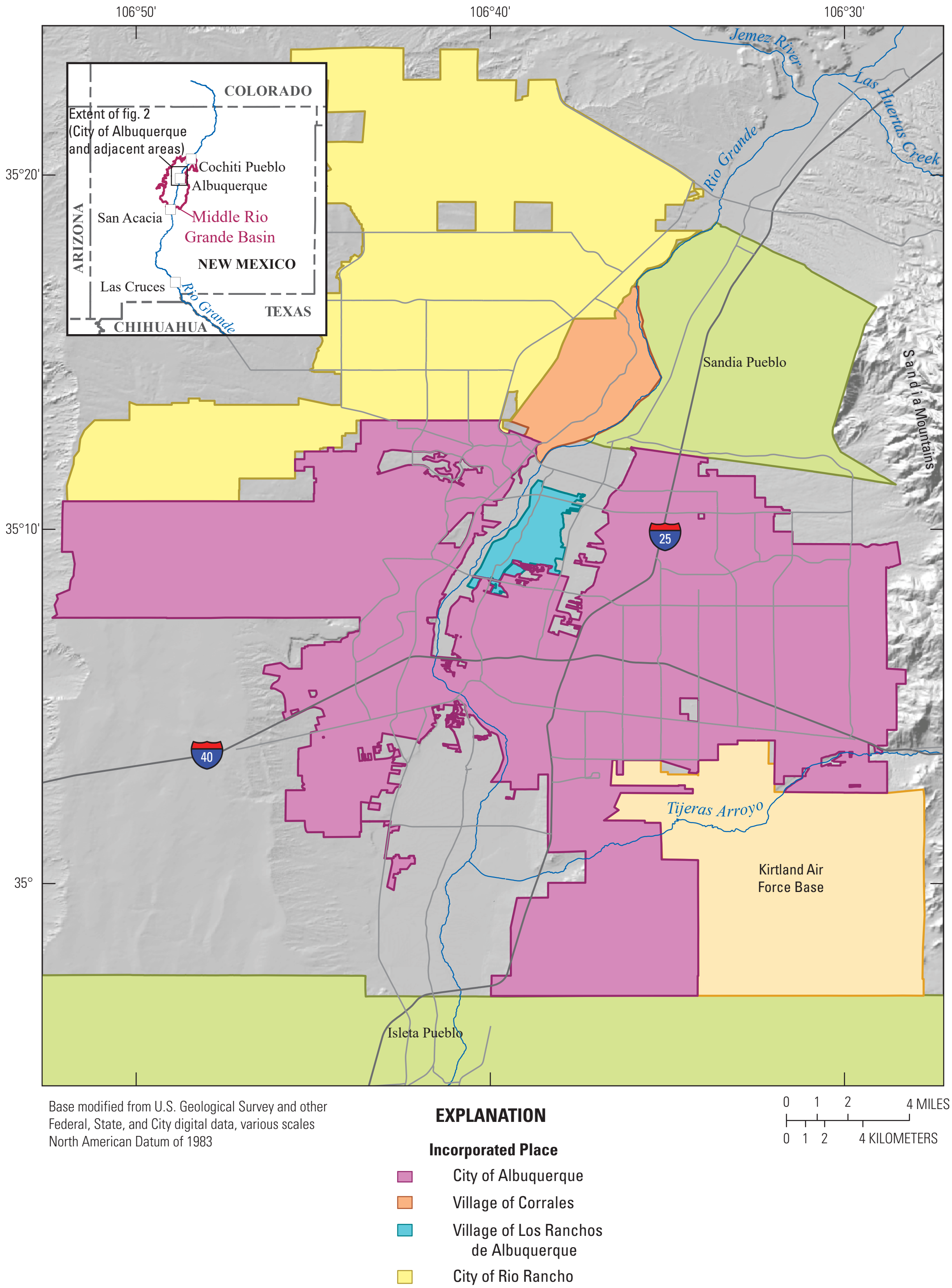


Figure 2. The incorporated places within the study area of this report. Extent of Middle Rio Grande Basin (inset) from Plummer and others (2012).

are hydrologically connected to each other and to the Rio Grande and are collectively called the Santa Fe Group aquifer system (Bartolino and Cole, 2002; Plummer and others, 2012; Rankin and others, 2013). The Santa Fe Group aquifer system is considered an unconfined aquifer, although silt and clay layers sometimes create localized confined to semiconfined conditions (Bartolino and Cole, 2002). Despite the large thickness of the system, only the upper 2,000 feet of the aquifer system is used for groundwater withdrawals because of the high cost of drilling and pumping from deeper and because groundwater in the lower part of the system has been assumed to be of limited quantity and poor quality (Bartolino and Cole, 2002).

Albuquerque and its neighboring communities (fig. 2) rely on both surface water and groundwater to meet their water needs. With a population of 545,852 people, Albuquerque was the most populous city in New Mexico at the time of the 2010 census (U.S. Census Bureau, 2010). The city is bordered on the north by the City of Rio Rancho, the Villages of Corrales and Los Ranchos de Albuquerque, and Sandia Pueblo and is bordered on the south by Kirtland Air Force Base and Isleta Pueblo (fig. 2). Overall, water in the Albuquerque area for municipal, industrial, commercial, and domestic uses is derived from both surface water and groundwater, whereas water for irrigated agriculture is primarily derived from surface water (Rankin and others, 2013).

Historically, the ABCWUA exclusively relied on groundwater from the Santa Fe Group aquifer system for municipal water supply, with groundwater withdrawals peaking at about 127,000 acre-feet per year in 1989 (Bexfield and Anderholm, 2002; ABCWUA, 2016; Galanter and Curry, 2019). Reductions of groundwater withdrawals began in 1990, despite continued population increases, because of conservation programs initiated by the ABCWUA and implemented by the citizens of Albuquerque (ABCWUA, 2016). In December 2008, the ABCWUA began diverting surface water from the Rio Grande as part of the SJCDWP to supplement municipal water supply, allowing for additional reductions in groundwater withdrawals (Powell and McKean, 2014; ABCWUA, 2016; Driscoll and Brandt, 2017; Galanter and Curry, 2019). ABCWUA production wells are typically screened (open to the aquifer) in the Santa Fe Group aquifer system production zone, which consists of the interval from within about 200 feet of the water table to 900 feet or more below the water table (Bexfield and Anderholm, 2002; Powell and McKean, 2014).

Background and Previous Investigations

Groundwater observation network design is defined as the selection of sampling points (spatial), sampling frequency (temporal), and sampling type (constituents) to achieve the observation network objectives (Gangopadhyay and others, 2001; Asefa and others, 2004). These networks typically fall into two categories: (1) groundwater contaminant observation networks and (2) groundwater-level observation networks (Asefa and others, 2004). Network objectives can often compete with one another and may require ambient monitoring, detection monitoring, compliance monitoring, research monitoring, or a combination of these (Heath, 1976; Loaiciga and others, 1992; Gangopadhyay and others, 2001). Objectives of a groundwater contaminant observation network may include initial groundwater contamination detection, plume characterization, and long-term monitoring (Asefa and others, 2004). Objectives of a groundwater-level observation network typically focus on long-term aquifer storage monitoring and aquifer system characterization (Heath, 1976; Asefa and others, 2004). Quantitatively determining observation network sampling frequency and spatial location to achieve network objectives has been the focus of many previous researchers (Olea, 1984; Rouhani, 1985; Massmann and Freeze, 1987a, 1987b; Wagner and Gorelick, 1987, 1989; Meyer and Brill, 1988; Rouhani and Hall, 1988; Scheibe and Lettenmaier, 1989; Loaiciga and others, 1992; Huang and Mayer, 1997; Asefa and others, 2004). Much of their work was aimed at designing groundwater contaminant observation networks and often involved a stochastic determination of network properties by selecting networks that best achieved a mathematical representation of network objectives; this mathematical expression is referred to as an "objective function" or "multiobjective function" (Loaiciga and others, 1992; Nunes and others, 2004b). Multiobjective functions combine several mathematical representations of network objectives into one mathematical expression, and its terms can be weighted to give greater importance to higher priority objectives. Statistical methods to design networks primarily differ in the objective function formulation and the method for exploring various network configurations (Asefa and others, 2004).

Groundwater network optimization (hereafter optimization) sometimes follows network design and involves refining network implementation to better meet network objectives, achieve new objectives, or make monitoring more cost effective. Adjustments may include relocating, adding, or removing sites from the network and (or) modifying the type of monitoring and its sampling frequency at network sites. Optimization commonly involves achieving a fine balance between maintaining consistent sampling sites to promote successful long-term monitoring and accomplishing additional goals of the optimization. Therefore, like groundwater network design, optimization often utilizes advanced statistical techniques and multiobjective functions. While these techniques have been widely applied to develop new networks or expand existing networks, applications to reduce existing networks have been less studied (Nunes and others, 2004a, 2004b).

Reducing the number of sites in an existing observation network is a nonlinear, combinatorial problem (that is, the number of possible combinations of sites is much larger than the number of sites in the network), which can result in a large computational complexity (Nunes and others, 2004b; Fisher, 2013). Heuristic algorithms that can handle the computational complexity of combinatorial problems are often relied upon in optimization efforts that seek to reduce network size (Asefa and others, 2004; Nunes and others, 2004b; Li and Hilton, 2007; Dhar and Datta, 2010; Fisher, 2013). Recently, Fisher (2013) developed a heuristic network design tool, using a genetic search algorithm with a multiobjective function criterion, for optimizing observation networks and applied it separately to two existing groundwater-level observation networks in the eastern Snake River Plain, Idaho. His approach implements universal kriging to create a groundwater-elevation surface for a given study area and then searches for reduced networks that can adequately achieve mathematically represented network objectives. These objectives are combined into a single composite objective function that is optimized for variously sized networks to aid water managers in network size reduction decisions.

Approaches for temporal (that is, sampling frequency) optimization of groundwater observation networks have been presented in several studies (Cameron and Hunter, 2002; Aziz and others, 2003; Cameron, 2004; Herrera and Pinder, 2005; Thakur, 2015). These techniques typically try to reduce a network's temporal redundancy by using trend analysis, Kalman filters, or temporal variogram analysis. Multiple researchers have invoked iterative data-thinning schemes in conjunction with trend analysis to evaluate sampling frequency (Cameron and Hunter, 2002; Cameron, 2004; Thakur, 2015). The iterative thinning approaches often use Sen's (1968) method or local regression for estimating trends and consist of the following components: (1) estimating the trend for the entire time series at a well, (2) iteratively thinning, at random, the time series by subsampling the time series, (3) estimating the trend for the thinned time series, and (4) comparing the trends from the entire and thinned time series. Thinning of the time series at a well continues until the trends from the entire and thinned time series are significantly different (that is, too much of the trend from the thinned time series is outside the bounds of a confidence interval around the trend from the entire time series). The level of thinning in which this occurs is considered the longest temporal frequency that a well can be sampled without significant data loss and thereby informs water managers of minimum sampling frequency. Cameron (2004, p. 95) favored using local regression to estimate trends, indicating that Sen's (1968) method "did not work very well for nonlinear, including seasonal, trends."

PCA is a statistical dimension-reduction technique that can aid in understanding patterns in large datasets (Hastie and others, 2009; James and others, 2013). The method involves recasting the original data variables as a set of linear combinations, which are called principal components (PCs). A large proportion of the original dataset's variance can typically be represented by considering a subset of the PCs. It is in this manner that PCA can reduce the dimensionality of a dataset and thereby ease interpretation (Taylor and Alley, 2001; Hastie and others, 2009; James and others, 2013). This technique has been applied to groundwater observation networks to identify sites that are representative of a larger number of sites (index sites) and distinguish characteristic hydrographs, towards informing a reduction in the number of sites in a network. PCA was used by Winter and others (2000) to identify groups of sites within a groundwater-level observation network that could likely be represented by a few index sites within those groups. These groups were thought to differ by aquifer characteristics and recharge conditions. Gangopadhyay and others (2001) used PCA and a ranking scheme to ascertain a list of network sites that explained the majority of temporal and spatial groundwater-level variations. Khan and others (2008) applied PCA to reduce the size of a groundwater-level observation network in Australia by removing observation sites that accounted for small proportions of dataset variance. In this study, kriged groundwater-elevation surfaces were used to evaluate and visualize how the optimized network compared to the original network. Overall, PCA is a valuable optimization tool that can provide a statistical evaluation of the importance of observation sites and their ability to be represented by index sites.

ABCWUA Groundwater-Level Observation Network

The observation network was established by the ABCWUA in cooperation with the USGS as part of the Middle Rio Grande Basin monitoring program to evaluate the rate and direction of groundwater flow and monitor long-term aquifer storage conditions beneath the Albuquerque area. This observation network consists of monitoring sites that are a combination of monitoring wells, piezometers (a specialized well open to a specific depth in the aquifer, and often of small diameter and constructed in proximity with other piezometers open to different depths [together called a nest]), and production wells (Ritchie and Galanter, 2019). Some of these sites are measured manually using a steel or electric tape (discrete), while others are equipped with electronic pressure transducers and data loggers to obtain measurements at specified sampling intervals (continuous). All sites are measured according to procedures outlined in Cunningham and Schalk (2011). Details regarding the observation network sites considered for this study are provided in table 1, and their spatial distribution is shown in figure 3. As of early 2019, the network included 47 continuous monitoring wells and piezometers recording at an hourly interval, 49 monitoring wells and piezometers monitored semiannually with a steel or electric tape, and a variable subset (approximately 40–50) of production wells monitored annually during a period of decreased seasonal water use (winter) with a steel tape (fig. 3, table 1) (Ritchie and Galanter, 2019). The number of production wells that are monitored annually varies depending on access constraints. Monitoring sites are screened in various parts of the Santa Fe Group aquifer system. A piezometer nest is typically constructed with one piezometer near the water table, one near the middle of the production zone, and one near the bottom of or below the production zone (Bexfield and Anderholm, 2002; Ritchie and Galanter, 2019).

Table 1. Description of sites in the observation network.

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment												
345650106415904	Isleta	piezometer	4,899	1	50	10–40	hourly	No	03/28/1998 to 09/30/2017	0.55	No	No
345650106415903	Isleta	piezometer	4,899	1	185	175–180	hourly	No	02/20/1998 to 09/30/2017	0.97	No	No
345650106415902	Isleta	piezometer	4,899	1	815	805–810	hourly	Yes	02/20/1998 to 09/30/2017	1.2	Yes	Yes
345650106415901	Isleta	piezometer	4,899	1	1,340	1,330–1,335	hourly	Yes	02/20/1998 to 09/30/2017	1.2	No	No
350056106370103	Montessa Park	piezometer	5,094	1	330	260–320	hourly	No	10/01/1998 to 07/25/2017	2.6	No	No
350056106370102	Montessa Park	piezometer	5,094	1	708	698–703	hourly	Yes	10/01/1998 to 09/30/2017	4.5	Yes	Yes
350244106450203	Westgate Heights Park	piezometer	5,241	1	370	320–360	hourly	No	03/31/2001 to 09/30/2017	0.44	No	No
350244106450202	Westgate Heights Park	piezometer	5,241	1	868	858–863	hourly	Yes	07/20/2000 to 09/30/2017	7.2	Yes	Yes
350244106450201	Westgate Heights Park	piezometer	5,241	1	1,290	1,280–1,285	hourly	Yes	07/20/2000 to 09/30/2017	7.5	No	No
350256106390801	San Jose 9	monitoring well	4,943	1	765	188.5–764.5	hourly	Yes	11/02/1988 to 09/30/2017	5.1	Yes	Yes
350307106410602	Armijo	piezometer	4,939	1	1,025	995–1,015	hourly	Yes	06/30/2005 to 09/30/2017	10	Yes	Yes
350307106410601	Armijo	piezometer	4,939	1	1,623	1,593–1,613	hourly	Yes	06/30/2005 to 09/30/2017	9.9	No	No
350534106354703	Del Sol Divider	piezometer	5,210	1	425	315–415	hourly	No	12/07/1996 to 09/30/2017	7.1	No	No
350534106354702	Del Sol Divider	piezometer	5,210	1	842	832–837	hourly	Yes	11/05/1996 to 09/30/2017	11	Yes	Yes
350534106354701	Del Sol Divider	piezometer	5,210	1	1,567	1,557–1,562	hourly	Yes	11/05/1996 to 09/30/2017	11	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350545106335903	Jerry Cline Park	piezometer	5,287	1	510	400–500	hourly	No	04/02/2004 to 09/30/2017	10	No	No
350545106335902	Jerry Cline Park	piezometer	5,287	1	1,050	1,030–1,040	hourly	Yes	04/02/2004 to 09/30/2017	12	Yes	Yes
350545106335901	Jerry Cline Park	piezometer	5,287	1	1,455	1,435–1,445	hourly	Yes	04/02/2004 to 09/30/2017	14	No	No
350638106413703	West Bluff Nest 1	piezometer	5,110	1	433	422–427	hourly	Yes	04/02/1997 to 07/26/2017	1.5	No	No
350638106413702	West Bluff Nest 1	piezometer	5,110	1	689	679–684	hourly	Yes	02/08/1997 to 07/26/2017	6.8	Yes	Yes
350638106413701	West Bluff Nest 1	piezometer	5,110	1	1,095	1,085–1,090	hourly	Yes	02/08/1997 to 07/26/2017	5.2	No	No
350638106413706	West Bluff Nest 2	piezometer	5,110	1	173	143–163	hourly	No	04/05/1997 to 09/30/2017	0.94	No	No
350638106413705	West Bluff Nest 2	piezometer	5,110	1	254	244–249	hourly	No	04/02/1997 to 09/30/2017	1.1	No	No
350638106413704	West Bluff Nest 2	piezometer	5,110	1	328	318–323	hourly	No	02/08/1997 to 06/24/2017	1.1	No	No
350706106390303	Garfield Park	piezometer	4,969	1	93	43–83	hourly	No	12/21/1996 to 09/30/2017	2.7	No	No
350706106390302	Garfield Park	piezometer	4,969	1	582	552–572	hourly	Yes	12/21/1996 to 09/30/2017	4.5	Yes	Yes
350706106390301	Garfield Park	piezometer	4,969	1	1,020	995–1,010	hourly	Yes	12/21/1996 to 09/30/2017	4.9	No	No
350836106395404	Montaño Nest 6	piezometer	4,974	1	182	172–177	hourly	No	09/30/2007 to 09/30/2017	2.3	No	No
350836106395403	Montaño Nest 6	piezometer	4,974	1	568	558–563	hourly	Yes	09/30/2007 to 09/30/2017	4.5	No	No
350836106395402	Montaño Nest 6	piezometer	4,974	1	836	826–831	hourly	Yes	09/30/2007 to 09/30/2017	4.2	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350836106395401	Montaño Nest 6	piezometer	4,974	1	983	972–978	hourly	Yes	09/30/2007 to 09/30/2017	3.9	Yes	Yes
350908106344402	Sister Cities	piezometer	5,242	1	799	789–794	hourly	Yes	01/25/1997 to 09/30/2017	6.7	Yes	Yes
350908106344401	Sister Cities	piezometer	5,242	1	1,308	1,298–1,303	hourly	Yes	01/25/1997 to 09/30/2017	9.5	No	No
350910106414803	Sierra Vista	piezometer	5,109	1	210	140–200	hourly	No	06/04/1998 to 09/30/2017	1.7	No	No
350910106414802	Sierra Vista	piezometer	5,109	1	928	918–923	hourly	Yes	07/31/1998 to 09/30/2017	1.9	Yes	Yes
350910106414801	Sierra Vista	piezometer	5,109	1	1,644	1,634–1,639	hourly	Yes	06/04/1998 to 09/30/2017	3.6	No	No
351114106330603	Nor Este	piezometer	5,453	1	608	538–598	hourly	No	03/20/1998 to 09/30/2017	6.8	No	No
351114106330602	Nor Este	piezometer	5,453	1	1,193	1,183–1,188	hourly	Yes	03/20/1998 to 09/30/2017	6.7	Yes	Yes
351114106330601	Nor Este	piezometer	5,453	1	1,525	1,515–1,520	hourly	Yes	03/20/1998 to 09/30/2017	6.6	No	No
351201106400503	Hunters Ridge Nest 1	piezometer	5,112	1	238	148–228	hourly	No	01/31/1997 to 09/30/2017	1.8	No	No
351201106400502	Hunters Ridge Nest 1	piezometer	5,112	1	855	845–850	hourly	Yes	11/02/1996 to 09/30/2017	3.1	Yes	Yes
351201106400501	Hunters Ridge Nest 1	piezometer	5,112	1	1,518	1,508–1,513	hourly	Yes	11/04/1996 to 09/30/2017	3.1	No	No
351201106400506	Hunters Ridge Nest 2	piezometer	5,112	1	263	238–258	hourly	No	11/02/1996 to 09/30/2017	2.0	No	No
351201106400505	Hunters Ridge Nest 2	piezometer	5,112	1	305	295–300	hourly	No	11/02/1996 to 09/30/2017	2.1	No	No
351201106400504	Hunters Ridge Nest 2	piezometer	5,112	1	359	349–354	hourly	Yes	11/02/1996 to 09/30/2017	2.7	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
351357106323002	Sandia Pueblo	piezometer	5,444	1	1,025	1,015–1,020	hourly	Yes	01/23/1999 to 09/30/2017	2.2	Yes	Yes
351357106323001	Sandia Pueblo	piezometer	5,444	1	1,305	1,295–1,300	hourly	Yes	02/27/1999 to 09/30/2017	2.2	No	No
350138106401103	Rio Bravo Nest 5	piezometer	4,932	1	515	500–510	semiannually	Yes	09/29/1992 to 08/22/2017	0.55	Yes	No
351059106385903	Paseo del Norte Nest 1	piezometer	4,993.9	0.01	600	545–555	semiannually	Yes	12/03/1992 to 08/22/2017	3.5	Yes	No
351035106364703	Paseo del Norte Nest 3	piezometer	5,013	1	543.9	538.9–543.9	semiannually	Yes	08/24/1993 to 08/22/2017	3.6	Yes	No
350552106444601	Arroyo Vista	monitoring well	5,313	1	1,424	520–571; 616–667; 847–899; 1,040–1,091; 1,371–1,424	semiannually	Yes	08/25/2009 to 08/22/2017	0.39	Yes	No
350445106411501	Atrisco #2 (ATR-2)	production well	4,951	1	--	--	annually	Yes	02/22/2000 to 12/14/2016	2.2	Yes	No
350508106411901	Atrisco #3 (ATR-3)	production well	4,954	1	804	180–804	annually	Yes	11/17/1958 to 11/21/2016	2.2	Yes	No
350509106414401	Atrisco #4 (ATR-4)	production well	4,954	1	500	--	annually	Yes	12/04/1996 to 11/19/2015	6.2	Yes	No
350359106362401	Burton #1 (BUR-1)	production well	5,324	1	1,312	--	annually	Yes	12/05/1996 to 11/07/2016	7.2	Yes	No
350421106361001	Burton #2 (BUR-2)	production well	5,287	5	857	425–845	annually	Yes	01/01/1962 to 01/25/2010	30	Yes	No
350440106355801	Burton #3 (BUR-3)	production well	5,218	5	994	358–994	annually	Yes	01/01/1962 to 04/11/2006	32	Yes	No
350355106351501	Burton #5 (BUR-5)	production well	5,279	1	1,170	--	annually	Yes	12/05/1996 to 12/14/2016	7.9	Yes	No
350606106341101	Charles #2 (CHA-2)	production well	5,264	1	1,020	432–996	annually	Yes	01/01/1968 to 01/25/2016	21	Yes	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350615106345901	Charles #5 (CHA-5)	production well	5,223	1	1,400	--	annually	Yes	12/19/1996 to 12/14/2016	8.2	Yes	No
350646106443201	College #1 (COL-1)	production well	5,337	1	1,662	--	annually	Yes	11/15/1996 to 11/04/2015	7.5	Yes	No
350647106440001	College #2 (COL-2)	production well	5,229	1	1,575	550–1,564	annually	Yes	12/01/1978 to 11/04/2015	6.2	Yes	No
351025106341601	Coronado #1 (COR-1)	production well	5,289	1	1,194	479–1,184	annually	Yes	07/13/1974 to 11/08/2016	12	Yes	No
351007106343801	Coronado #2 (COR-2)	production well	5,243	1	1,390	--	annually	Yes	12/18/1996 to 11/21/2016	5.5	Yes	No
351302106424701	Corrales #3 (COA-3)	production well	5,453	1	1,364	650.4–1,351.14	annually	Yes	04/01/1980 to 03/20/2015	16	Yes	No
351215106441301	Corrales #4 (COA-4)	production well	5,470	1	1,357	692–1,339	annually	Yes	09/30/1994 to 03/23/2015	47	Yes	No
350642106401101	Duranes #1 (DUR-1)	production well	4,965	1	924	204–924	annually	Yes	12/08/1959 to 11/21/2016	4.0	Yes	No
350708106405801	Duranes #2 (DUR-2)	production well	4,970	1	804	180–804	annually	Yes	10/01/1958 to 12/08/2015	1.6	Yes	No
350629106405101	Duranes #3 (DUR-3)	production well	4,964	1	950	136–954	annually	Yes	01/01/1954 to 12/14/2016	2.9	Yes	No
350628106411501	Duranes #4 (DUR-4)	production well	4,960	1	950	148–950	annually	Yes	01/01/1953 to 12/27/2016	4.1	Yes	No
350605106411801	Duranes #5 (DUR-5)	production well	4,961	1	950	152–950	annually	Yes	01/01/1953 to 11/07/2016	1.9	Yes	No
350653106403001	Duranes #6 (DUR-6)	production well	4,968	1	--	164–956	annually	Yes	01/01/1953 to 11/07/2016	5.2	Yes	No
350655106395001	Duranes #7 (DUR-7)	production well	4,966	1	814	144–814	annually	Yes	02/01/1953 to 12/27/2016	2.9	Yes	No
350642106422801	Gonzales #1 (GON-1)	production well	5,112	1	970	--	annually	Yes	11/01/1992 to 11/21/2016	9.1	Yes	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350635106415001	Gonzales #2 (GON-2)	production well	5,103	1	1,115	--	annually	Yes	12/18/1996 to 12/14/2016	11	Yes	No
350827106395001	Griegos #1 (GRI-1)	production well	4,975	1	802	232–802	annually	Yes	01/01/1955 to 11/21/2016	4.9	Yes	No
350802106402901	Griegos #3 (GRI-3)	production well	4,969.7	10	916	260–916	annually	Yes	01/01/1955 to 02/16/2010	1.7	Yes	No
350821106390101	Griegos #4 (GRI-4)	production well	4,977	1	804	216–804	annually	Yes	01/01/1955 to 01/29/2014	12	Yes	No
350309106434501	Leavitt #1 (LEA-1)	production well	5,085	1	1,229	--	annually	Yes	12/30/1996 to 11/04/2015	7.7	Yes	No
350223106435401	Leavitt #3 (LEA-3)	production well	5,085	1	1,520	--	annually	Yes	03/06/2000 to 11/04/2015	8.1	Yes	No
350752106342101	Leyendecker #1 (LEY-1)	production well	5,288	1	996	468–996	annually	Yes	01/01/1959 to 01/25/2016	23	Yes	No
350727106340801	Leyendecker #2 (LEY-2)	production well	5,300.74	5	996	468–996	annually	Yes	01/15/1960 to 03/05/2008	33	Yes	No
350819106344001	Leyendecker #3 (LEY-3)	production well	5,270	1	996	204–924	annually	Yes	02/08/1960 to 12/14/2016	22	Yes	No
350815106340601	Leyendecker #4 (LEY-4)	production well	5,328	5	996	480–996	annually	Yes	01/01/1960 to 12/04/2001	34	Yes	No
350430106302401	Lomas #1 (LOM-1)	production well	5,597.66	5	1,300	700–1,300	annually	Yes	09/15/1962 to 12/18/2000	92	Yes	No
350422106312401	Lomas #5 (LOM-5)	production well	5,499	1	1,670	823–1,651	annually	Yes	01/01/1979 to 12/08/2015	16	Yes	No
350408106310101	Lomas #6 (LOM-6)	production well	5,533	1	1,704	876–1,688	annually	Yes	01/01/1979 to 12/27/2016	18	Yes	No
350517106314401	Love #1 (LOV-1)	production well	5,464.7	5	1,096	596–1,096	annually	Yes	01/01/1955 to 02/15/2001	52	Yes	No
350538106333001	Love #8 (LOV-8)	production well	5,317	10	1,455	640–1,440	annually	Yes	06/21/1989 to 01/02/2002	4.5	Yes	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350308106374601	Miles Road #1 (MIL-1)	production well	5,149	5	1,165.5	404.5–1,153.5	annually	Yes	02/14/1997 to 11/22/2010	2.8	Yes	No
350800106315001	Ponderosa #2 (PON-2)	production well	5,602	1	1,581	801–1,569	annually	Yes	01/01/1973 to 12/27/2016	8.4	Yes	No
350851106322001	Ponderosa #6 (PON-6)	production well	5,561	5	1,675	852–1,662	annually	Yes	01/01/1979 to 02/11/2002	15	Yes	No
350401106331401	Ridgecrest #3 (RID-3)	production well	5,387.71	5	1,448	620.5–1,436.5	annually	Yes	11/15/1974 to 01/21/2011	20	Yes	No
350445106334001	Ridgecrest #4 (RID-4)	production well	5,345	1	1,424	572.5–1,412.5	annually	Yes	01/27/1974 to 01/05/2015	22	Yes	No
350420106334401	Ridgecrest #5 (RID-5)	production well	5,355	10	1,470	--	annually	Yes	12/18/1996 to 12/23/2011	25	Yes	No
350316106384801	San Jose #1 (SAJ-1)	production well	4,952	4.3	--	--	annually	Yes	01/23/1997 to 02/14/2005	1.0	Yes	No
350315106390401	San Jose #2 (SAJ-2)	production well	4,947.65	5	471	--	annually	Yes	01/23/1997 to 02/12/2001	3.1	Yes	No
350301106383601	San Jose #3 (SAJ-3)	production well	4,956	1	1,032	--	annually	Yes	01/23/1997 to 11/05/2015	4.9	Yes	No
350648106362501	Santa Barbara #1 (SAB-1)	production well	5,139	1	984	312–672	annually	Yes	01/01/1963 to 11/07/2016	18	Yes	No
350744106333501	Thomas #5 (THO-5)	production well	5,358	1	1,450	--	annually	Yes	02/14/1997 to 12/27/2016	9.8	Yes	No
350720106330401	Thomas #6 (THO-6)	production well	5,411	1	1,536	--	annually	Yes	02/14/1997 to 11/21/2016	11	Yes	No
350719106333401	Thomas #7 (THO-7)	production well	5,346	1	1,475	659–1,460	annually	Yes	04/21/1989 to 11/21/2016	13	Yes	No
350711106323101	Thomas #8 (THO-8)	production well	5,467	1	1,655	--	annually	Yes	02/14/1997 to 11/21/2016	11	Yes	No
350805106354901	Vol Andia #1 (VOA-1)	production well	5,145	1	972	300–972	annually	Yes	01/01/1960 to 12/03/2014	27	Yes	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites used in optimization assessment—Continued												
350732106350101	Vol Andia #2 (VOA-2)	production well	5,210	1	852	360–852	annually	Yes	01/01/1960 to 11/21/2016	21	Yes	No
350803106351101	Vol Andia #4 (VOA-4)	production well	5,204	1	876	372–876	annually	Yes	01/01/1960 to 12/14/2016	21	Yes	No
350809106360901	Vol Andia #5 (VOA-5)	production well	5,112.73	5	900	372–876	annually	Yes	09/01/1960 to 02/27/2002	28	Yes	No
350828106352101	Vol Andia #6 (VOA-6)	production well	5,180.74	5	984	324–984	annually	Yes	10/26/1960 to 01/04/2013	26	Yes	No
350950106434001	Volcano Cliffs #1 (VOC-1)	production well	5,340	1	1,080	528–1,056	annually	Yes	09/24/1968 to 12/14/2016	28	Yes	No
351007106434201	Volcano Cliffs #3 (VOC-3)	production well	5,352.81	5	1,321	666–1,309	annually	Yes	03/06/1980 to 12/11/2006	25	Yes	No
350931106315501	Walker #3 (WAL-3)	production well	5,649	1	1,704	964–1,693	annually	Yes	04/28/1979 to 11/08/2016	14	Yes	No
350918106315401	Walker #4 (WAL-4)	production well	5,631	1	1,626	939–1,613	annually	Yes	01/01/1978 to 11/08/2016	17	Yes	No
351029106332001	Webster #1 (WEB-1)	production well	5,437	1	1,357	--	annually	Yes	02/09/1998 to 11/08/2016	8.1	Yes	No
351013106333501	Webster #2 (WEB-2)	production well	5,387	1	1,346	--	annually	Yes	03/20/2000 to 11/08/2016	7.2	Yes	No
350426106372601	Yale #1 (YAL-1)	production well	5,162	5	960	336–960	annually	Yes	01/01/1963 to 01/25/2012	23	Yes	No
350358106372901	Yale #2 (YAL-2)	production well	5,131	1	1,191	351–1,179	annually	Yes	01/01/1973 to 11/07/2016	11	Yes	No
350435106380101	Yale #3 (YAL-3)	production well	5,084	1	1,004	320–992	annually	Yes	01/01/1973 to 11/07/2016	9.5	Yes	No
350918106425401	Zamora #1 (ZAM-1)	production well	5,170	1	970	--	annually	Yes	09/23/1992 to 11/07/2016	11	Yes	No
350852106425701	Zamora #2 (ZAM-2)	production well	5,160	1	997	440–640	annually	Yes	12/14/1998 to 11/07/2016	7.6	Yes	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites not used in optimization assessment												
341528106533301	--	monitoring well	4,662.4	10	38	--	semiannually	No	02/24/1983 to 08/21/2017	2.5	No	No
342107106530401	Sevilleta Refuge Headquarters	monitoring well	4,862.46	10	223	210–220	semiannually	No	07/08/1975 to 08/21/2017	0.68	No	No
344258106460901	Estes 1	piezometer	4,820	1	135	125–130	semiannually	No	02/23/1983 to 02/27/2017	0.50	No	No
344258106460902	Estes 5	piezometer	4,820	1	300	265–270	semiannually	Yes	02/23/1983 to 02/27/2017	1.0	No	No
350137106410503	Rio Bravo Nest 1	piezometer	4,933	4.3	38.4	28–33	semiannually	No	02/19/1988 to 08/22/2017	0.60	No	No
350137106410502	Rio Bravo Nest 1	piezometer	4,933	4.3	103.8	94–99	semiannually	No	02/19/1988 to 08/22/2017	0.51	No	No
350137106410501	Rio Bravo Nest 1	piezometer	4,933.5	2	148.5	138.5–143.5	semiannually	No	02/19/1988 to 08/22/2017	0.48	No	No
350138106395503	Rio Bravo Nest 2	piezometer	4,934	4.3	48.6	38.6–43.6	semiannually	No	02/19/1988 to 08/22/2017	0.96	No	No
350138106395502	Rio Bravo Nest 2	piezometer	4,934	4.3	91.16	81–86	semiannually	No	02/19/1988 to 08/22/2017	0.92	No	No
350138106395501	Rio Bravo Nest 2	piezometer	4,934	4.3	153.53	143.53–148.53	semiannually	No	02/19/1988 to 08/22/2017	0.82	No	No
350138106393203	Rio Bravo Nest 3	piezometer	4,934	4.3	49.33	39.33–44.33	semiannually	No	02/19/1988 to 08/22/2017	1.5	No	No
350138106393202	Rio Bravo Nest 3	piezometer	4,934	4.3	101	91–96	semiannually	No	02/19/1988 to 08/22/2017	1.4	No	No
350138106393201	Rio Bravo Nest 3	piezometer	4,934	4.3	148	138–143	semiannually	No	02/19/1988 to 08/22/2017	1.3	No	No
350138106401102	Rio Bravo Nest 5	piezometer	4,932	1	22	7–17	semiannually	No	01/09/1992 to 08/22/2017	0.73	No	No
350138106401101	Rio Bravo Nest 5	piezometer	4,932	1	150	135–145	semiannually	No	01/09/1992 to 08/22/2017	0.44	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites not used in optimization assessment—Continued												
350854106403703	Montaño Nest 1	piezometer	4,976	4.3	48.41	40–45	semiannually	No	02/19/1988 to 08/22/2017	1.6	No	No
350854106403702	Montaño Nest 1	piezometer	4,976	4.3	93.4	83.4–88.4	semiannually	No	02/19/1988 to 08/22/2017	1.7	No	No
350854106403701	Montaño Nest 1	piezometer	4,976	4.3	152	140–145	semiannually	No	02/19/1988 to 08/22/2017	1.8	No	No
350836106395603	Montaño Nest 2	piezometer	4,975	4.3	39.7	30–35	semiannually	No	05/23/1988 to 08/22/2017	1.6	No	No
350836106395601	Montaño Nest 2	piezometer	4,975	4.3	147.4	138–143	semiannually	No	02/19/1988 to 08/22/2017	2.3	No	No
350827106391303	Montaño Nest 3	piezometer	4,975.01	1	49.8	40–45	semiannually	No	02/19/1988 to 08/22/2017	2.6	No	No
350827106391302	Montaño Nest 3	piezometer	4,975	4.3	99	90–95	semiannually	No	02/19/1988 to 08/22/2017	2.8	No	No
350827106391301	Montaño Nest 3	piezometer	4,975.02	1	149.8	140–145	semiannually	No	02/19/1988 to 08/22/2017	3.0	No	No
350821106383703	Montaño Nest 4	piezometer	4,977.37	1	50.2	40.2–45.2	semiannually	No	02/19/1988 to 08/21/2017	4.4	No	No
350821106383702	Montaño Nest 4	piezometer	4,977	4.3	93.5	84.5–89.5	semiannually	No	02/19/1988 to 08/21/2017	4.7	No	No
350821106383701	Montaño Nest 4	piezometer	4,977	4.3	131.5	122.5–127.5	semiannually	No	02/19/1988 to 08/21/2017	4.3	No	No
350859106401601	Montaño Nest 5	piezometer	4,979.99	0.1	25	10–20	semiannually	No	12/17/1992 to 08/22/2017	0.63	No	No
350859106401602	Montaño Nest 5	piezometer	4,980.02	0.1	75	60–70	semiannually	No	12/17/1992 to 08/22/2017	0.60	No	No
350859106401603	Montaño Nest 5	piezometer	4,979.85	0.1	150	135–145	semiannually	No	12/17/1992 to 08/22/2017	0.71	No	No
351059106385902	Paseo del Norte Nest 1	piezometer	4,993.52	0.1	25	10–20	semiannually	No	04/06/1992 to 08/22/2017	0.86	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites not used in optimization assessment—Continued												
351059106385901	Paseo del Norte Nest 1	piezometer	4,992.25	0.1	150	135–145	semiannually	No	02/14/1992 to 08/22/2017	2.0	No	No
351057106384203	Paseo del Norte Nest 2	piezometer	4,991.31	0.1	45	30–40	semiannually	No	12/17/1992 to 08/22/2017	1.1	No	No
351057106384202	Paseo del Norte Nest 2	piezometer	4,991.33	0.1	95	80–90	semiannually	No	12/17/1992 to 08/22/2017	1.1	No	No
351057106384201	Paseo del Norte Nest 2	piezometer	4,991.8	0.1	150	135–145	semiannually	No	12/17/1992 to 08/22/2017	2.6	No	No
351035106364701	Paseo del Norte Nest 3	piezometer	5,013	1	68.7	63.7–68.7	semiannually	No	08/24/1993 to 08/22/2017	2.9	No	No
351035106364702	Paseo del Norte Nest 3	piezometer	5,013	1	143.5	138.5–143.5	semiannually	No	08/24/1993 to 08/22/2017	2.9	No	No
350602106210401	Home Oil	monitoring well	6,523	4.3	54.3	--	semiannually	No	04/04/1988 to 08/21/2017	5.8	No	No
350548106383901	City 1	monitoring well	4,960	1	149	138.6–148.6	semiannually	No	08/01/1957 to 08/21/2017	11	No	No
350824106375301	City 2	monitoring well	4,982.71	5	150.3	140.3–150.3	semiannually	No	08/01/1957 to 08/21/2017	17	No	No
350837106393801	City 3	monitoring well	4,975.7	5	152	142–152	semiannually	No	05/09/1957 to 08/21/2017	5.1	No	No
350646106403601	City 4	monitoring well	4,961.69	5	150	140–150	semiannually	No	01/31/1957 to 08/22/2017	5.2	No	No
350454106570401	Cañoncito	monitoring well	5,335	3.2	117	--	semiannually	No	01/21/1958 to 08/23/2017	1.6	No	No
351852106344901	San Miguel	monitoring well	5,155	1	206	--	semiannually	No	05/24/1985 to 02/28/2017	4.2	No	No
350829106420401	La Luz del Sol	monitoring well	5,111	1	250	230–245	semiannually	No	09/28/1978 to 02/27/2017	6.0	No	No
351009106344701	Pino Yards	monitoring well	5,235	10	360	320–360	semiannually	No	06/18/1992 to 08/22/2017	4.3	No	No

Table 1. Description of sites in the observation network.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; DTW, depth to groundwater; PCA, principal component analysis; --, not available; NA, not applicable. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3. Data collected through September 30, 2017, were considered for analysis in this report; therefore, a site with a period of record ending on 09/30/2017 may have additional data collected after this date. A variable subset (approximately 40–50) of the production wells listed are measured annually]

USGS site number	Site name	Site type	Elevation of land surface (feet above NAVD 88)	Accuracy of land surface elevation (± feet)	Well depth (feet below land surface)	Screened interval (feet below land surface)	Measurement frequency	Production zone	Period of record	Standard deviation of DTW (feet)	Used for spatial assessment	Used for PCA
Monitoring sites not used in optimization assessment—Continued												
350602106333201	Charles #4 (CHA-4)	production well	5,328	5	1,056	456–1,032	annually	Yes	01/01/1968 to 12/04/1996	47	No	No
350705106420101	Gonzales #3 (GON-3)	production well	5,102	4.3	1,005	420–940	annually	Yes	01/25/2000 to 01/25/2000	NA	No	No
350249106434201	Leavitt #2 (LEA-2)	production well	5,072	1	1,133	--	annually	Yes	12/30/1996 to 11/04/2015	3.6	No	No
350553106313801	Love #6 (LOV-6)	production well	5,505	1	1,521	756–1,512	annually	Yes	01/01/1973 to 12/08/2015	36	No	No
350834106314901	Ponderosa #4 (PON-4)	production well	5,623	5	1,545	919–1,532	annually	Yes	01/01/1979 to 02/27/2002	22	No	No
350405106322001	Ridgecrest #1 (RID-1)	production well	5,445	5	1,256.5	632.5–1,256.5	annually	Yes	01/01/1973 to 02/13/2006	35	No	No
350427106323401	Ridgecrest #2 (RID-2)	production well	5,419	5	1,512	730–1,500	annually	Yes	01/01/1977 to 01/13/1999	26	No	No
350747106361401	Vol Andia #3 (VOA-3)	production well	5,109.73	5	900	360–852	annually	Yes	09/08/1960 to 01/08/1997	42	No	No
350444106435401	West Mesa #3 (WEM-3)	production well	5,145	4.3	1,365	--	annually	Yes	11/05/1996 to 01/12/2000	2.3	No	No
350442106431801	West Mesa #4 (WEM-4)	production well	5,107.7	5	1,287	386.5–1,274.5	annually	Yes	01/01/1974 to 12/01/2010	11	No	No

A

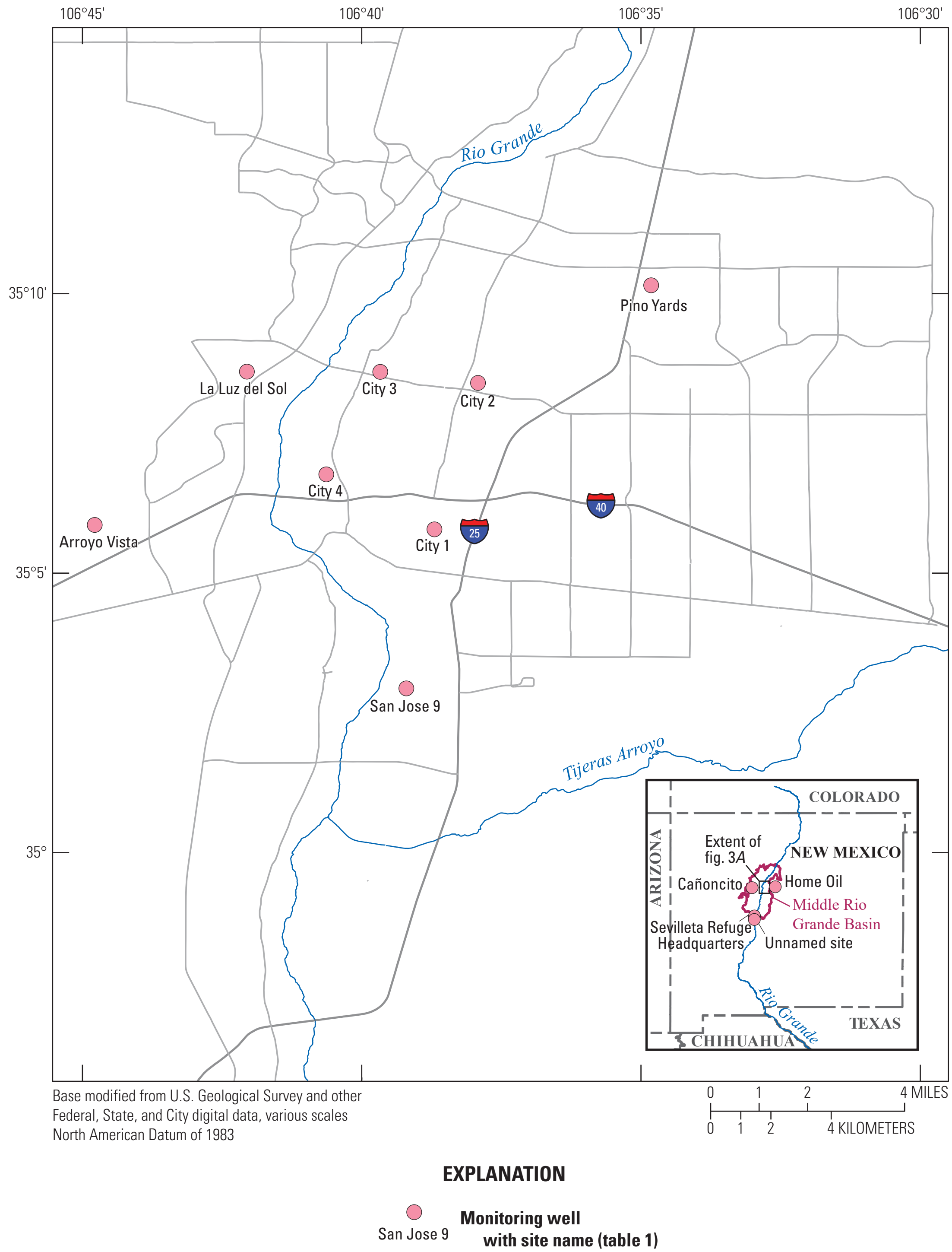
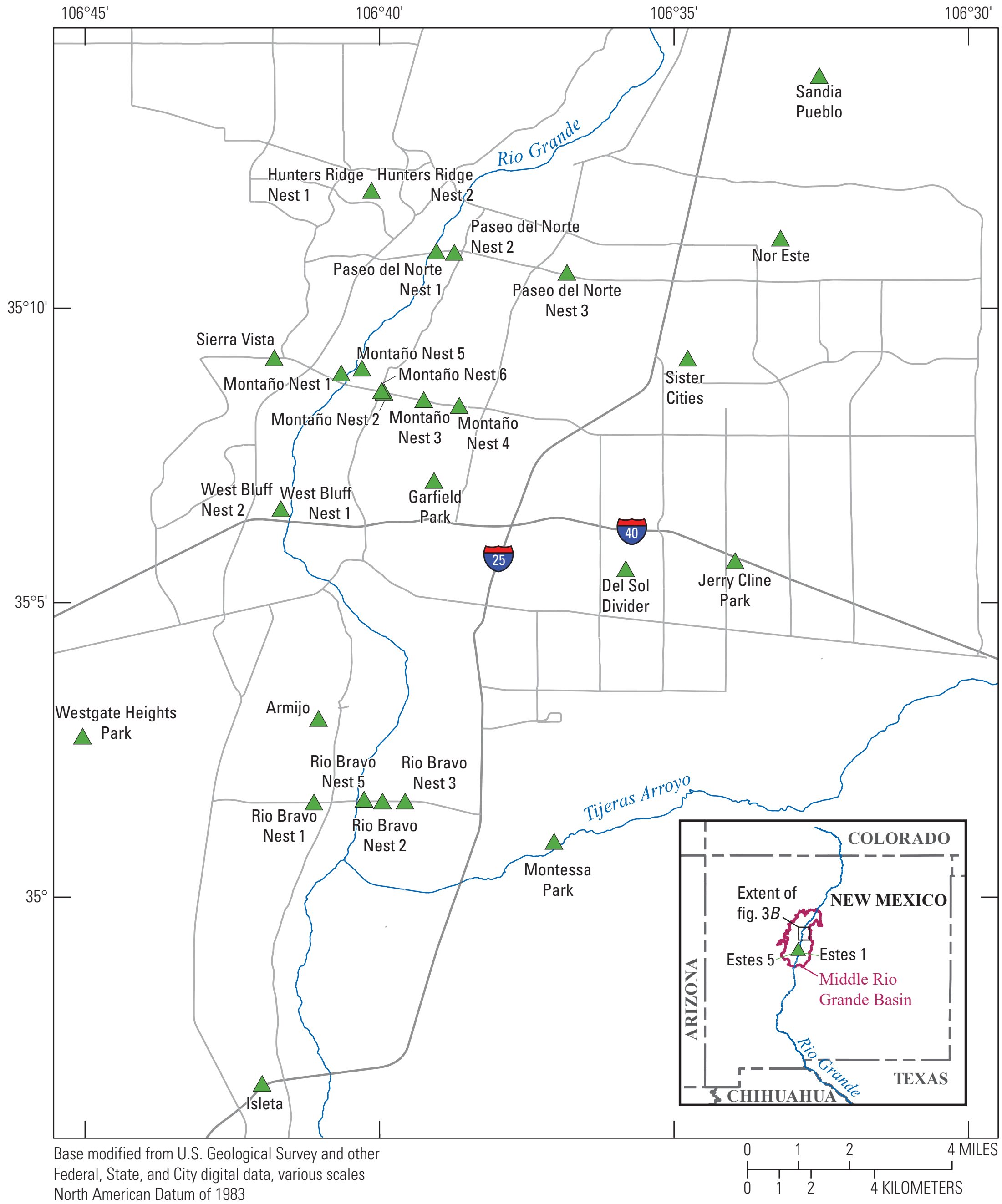


Figure 3. The locations of the observation network for *A*, monitoring wells; *B*, piezometers; and *C*, production wells.

B



EXPLANATION

Isleta — **Piezometer with site name (table 1)**

Figure 3. The locations of the observation network for *A*, monitoring wells; *B*, piezometers; and *C*, production wells.—Continued

C

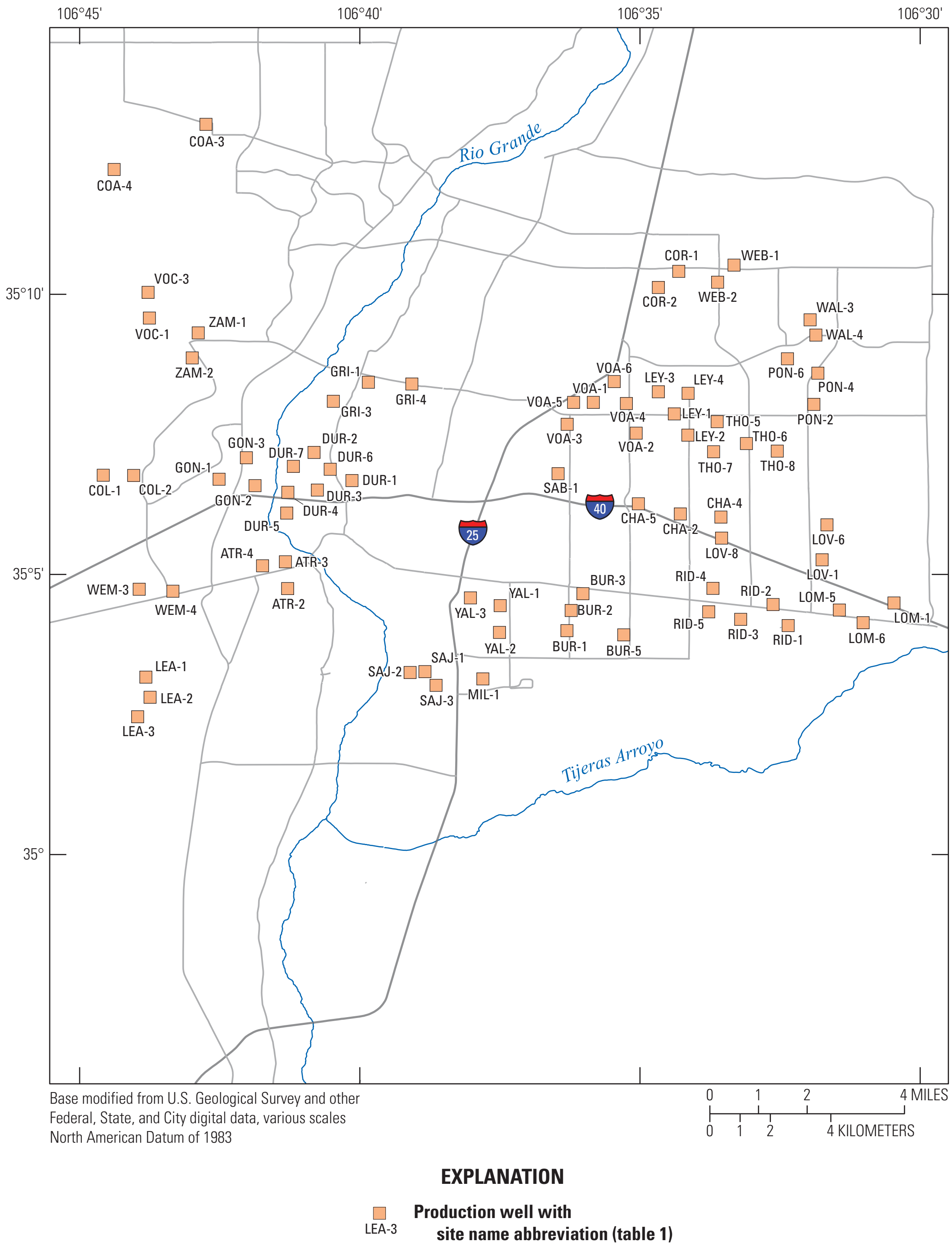


Figure 3. The locations of the observation network for *A*, monitoring wells; *B*, piezometers; and *C*, production wells.—Continued

Methods

The results of the assessment to provide insights for network optimization presented in this report include separate spatial and temporal approaches and an evaluation using PCA. The spatial assessment followed the heuristic procedure presented in Fisher (2013) to identify the spatial importance of monitoring sites. The temporal assessment utilized an iterative thinning technique coupled with statistical trend analyses, similar to the approach presented in Cameron (2004), to identify sites that may be monitored less frequently without significant information loss. In an approach similar to that of Winter and others (2000), PCA was applied to inform the results of the spatial and temporal assessments by identifying groups of monitoring sites that are representative of hydrologic conditions in various spatial regions of the production zone of the Santa Fe Group aquifer system.

Monitoring Sites Used for Optimization Assessment

This optimization assessment focused only on monitoring sites within the observation network that are funded by the ABCWUA and therefore ignored sites that are managed by other entities. This decision ensured that the ABCWUA could still achieve their objectives by solely relying on their network rather than including data from external networks that are outside of their control. The spatial assessment considered a total of 87 monitoring wells, piezometers, and production wells identified as being screened within the production zone (“yes” in “used for spatial assessment” column of table 1). The temporal assessment was performed on the 47 sites with continuous monitoring, where groundwater levels are recorded at an hourly interval (“hourly” in “measurement frequency” column of table 1). The PCA was applied to the 15 sites with continuous monitoring identified as being screened within the production zone (“yes” in “used for PCA” column of table 1). PCA was restricted to these monitoring sites because these sites met the data requirements for accurately applying the PCA technique and because the production zone was of most interest for monitoring purposes.

Monitoring wells and piezometers were classified as being screened in the production zone by computing the difference between the bottom of the screened interval, or the well depth if screen information was not available, and the minimum depth to groundwater measured at the site. The monitoring site was considered to be screened within the production zone if this difference was greater than 200 feet (“yes” in “production zone” column of table 1). This approach likely identified monitoring sites screened below the production zone because, as discussed above, a piezometer nest is typically constructed with one piezometer near the bottom of or below the production zone. Therefore, only those monitoring wells and piezometers identified as being screened in the production zone that were also used to generate the 2016 production zone groundwater-elevation contours within the Albuquerque area (Galanter and Curry, 2019) were included in the spatial assessment and PCA. An exception to this was the semiannually monitored Estes 5 (site no. 344258106460902) piezometer (inset map of fig. 3*B*), which was identified as production zone and was used to generate the 2016 production zone groundwater-elevation contours but was not included in the spatial assessment because it is a remote site (more than 16 miles from the nearest site used in the spatial assessment).

All of the production wells were assumed to be screened in the production zone. The 2016 production zone groundwater-elevation contours (Galanter and Curry, 2019) did not use all production wells within the network because measurements were not available during the winter time interval used to generate the contours (Amy Galanter, USGS, written commun., 2018). In addition, the Leavitt no. 2 (site no. 350249106434201) and Love no. 6 (site no. 350553106313801) production wells were not used to generate the 2016 production zone groundwater-elevation contours because the groundwater-level measurements appeared anomalous relative to surrounding groundwater elevations (Amy Galanter, USGS, written commun., 2018); these wells were therefore not included in the spatial assessment. The groundwater elevation at Leavitt no. 2 (site no. 350249106434201) was about 20 feet higher than the groundwater elevations at nearby Leavitt no. 1 (site no. 350309106434501) and Leavitt no. 3 (site no. 350223106435401) in November 2015 (Amy Galanter, USGS, written commun., 2018). The groundwater elevation at Love no. 6 (site no. 350553106313801) was about 100 feet higher than surrounding groundwater elevations in December 2015 (Amy Galanter, USGS, written commun., 2018). Eight additional production wells (Charles no. 4 [site no. 350602106333201], Gonzales no. 3 [site no. 350705106420101], Ponderosa no. 4 [site no. 350834106314901], Ridgecrest no. 1 [site no. 350405106322001], Ridgecrest no. 2 [site no. 350427106323401], Vol Andia no. 3 [site no. 350747106361401], West Mesa no. 3 [site no. 350444106435401], and West Mesa no. 4 [site no. 350442106431801]) were not included in the spatial assessment because groundwater-level measurements were not available during the time intervals of interest. In total, the spatial assessment was performed on 68 of the 78 observation network production wells and 19 monitoring wells and piezometers identified as being screened in the production zone that were also used to generate the 2016 production zone groundwater-elevation contours (table 1).

Sources and Descriptions of Data

Data used for the optimization assessment were obtained from the USGS National Water Information System (USGS, 2019b) and included geographic coordinates (longitude, latitude, and land surface elevation) and depth-to-groundwater measurements relative to land surface for each of the monitoring sites. Depth-to-groundwater measurements were converted to groundwater elevation by subtracting depth-to-groundwater measurements from land surface elevation. Elevations in this report are expressed as elevation above North American Vertical Datum of 1988 (NAVD 88). The period of record for the monitoring sites included in the optimization assessment is variable, with the earliest groundwater levels recorded in 1953 (table 1). The standard deviation of the depth-to-groundwater measurements at each monitoring site over the entire period of record is a measure of the historical variability, including seasonal fluctuations and long-term trends (table 1).

Data Accuracy

The accuracy of groundwater levels used in this report is related to the methods used to measure land surface elevation and depth to groundwater at each monitoring site. Land surface accuracies for the network are listed in table 1. Land surface elevation accuracy ranged from ±0.01 foot (determined by level or other surveying method) to ±10 feet (interpolated from topographic map), with a mean accuracy of about ±2 feet and a standard deviation of about 3 feet. Daily mean depth-to-groundwater measurements, which are computed by the USGS from the hourly data collected at continuously monitored wells and piezometers, are accurate to ±0.01 foot (USGS, 2019a, 2019b). Depth-to-groundwater measurements recorded with a steel or electric tape at monitoring wells and piezometers that are monitored semiannually and at the production wells that are monitored annually have an accuracy ranging from ±0.01 to ±1 foot (USGS, 2019b).

Spatial Assessment Methods

The spatial assessment followed the kriging-based genetic algorithm (GA) approach of Fisher (2013). This methodology begins by assembling a dataset that includes geographic coordinates, groundwater elevation, depth-to-groundwater-measurement standard deviation, and measurement uncertainty estimates. These data are used to develop a semivariogram model that is fed into a kriging algorithm to create a groundwater-elevation surface (grid or raster) for the study area. Optimized networks are identified by comparing this groundwater-elevation surface to groundwater-elevation surface estimates that are created using spatially reduced networks. The details of this approach are described below and in Fisher (2013).

Background

The spatial optimization algorithm for this assessment was developed using the R programming language (R Core Team, 2019) and was published by Fisher (2013) as the "ObsNetwork" package. This package uses the geostatistical technique known as universal kriging to interpolate the groundwater elevation at unmeasured sites within an observation network (kriged surface). Kriging estimates the values at unmeasured sites as a weighted mean of the measured values; the weights depend on the correlation structure in the data. An empirical semivariogram plot is used to evaluate the correlation between measurements at two monitoring sites. The curvature of this plot is related to the spatial correlation between monitoring sites and provides a data-driven method for kriging to estimate groundwater elevations away from monitoring sites. The empirical semivariogram is fit with the appropriately shaped theoretical semivariogram model (theoretical semivariogram) so that it can be utilized for kriging. Kriging assumes that the mean value of the data being estimated does not change when shifted in space and that the theoretical semivariogram is the same everywhere (stationarity). Groundwater elevations typically have pronounced spatial trends that would violate this assumption. Therefore, a trend model is first subtracted from the groundwater elevations, and the semivariogram and kriged surface are developed for the residuals. The trend model is then added back to the kriged surface to acquire the final kriged groundwater-elevation estimates. This method of kriging with a trend is known as universal kriging. Kriging also requires the estimation variance at monitoring sites (square of estimated value minus measured value) to be zero so that groundwater-elevation data are well represented. Prediction uncertainties associated with kriging are estimated as the standard error (square root of the estimated variance from kriging). The kriging standard error is zero at monitoring sites and increases as the density of monitoring sites decreases. Further information on the kriging formulation used by the "ObsNetwork" package can be found in Fisher (2013).

Once the kriged surface is developed, the "ObsNetwork" package employs a GA to evaluate the effectiveness of various reduced networks of user-specified size at reproducing the kriged surface. GAs explore different network configurations by mimicking the mechanics of natural selection and survival of the fittest (Fisher, 2013). These algorithms have several control parameters that influence how the solution space is explored and are typically selected using sensitivity analysis (Fisher, 2013). Optimal networks are identified as those that minimize a specified multiobjective function. The multiobjective function used in this study and Fisher (2013) contained four terms that represent different optimization objectives as follows: (1) minimize overall interpolation error in the kriged surface when using the reduced network, (2) preserve local anomalies in the kriged surface, (3) keep sites with large groundwater-level variability over time, and (4) maintain sites with higher measurement accuracy. These terms are summed into a single fitness value but can be weighted independently to emphasize certain priorities. Output of the optimization algorithm includes a list of sites identified for removal, a prediction difference raster showing the ability to reproduce the original kriged surface with the reduced network, and several additional reduced-network performance metrics.

Input Data Description

The spatial assessment was performed independently for two winter time intervals: November 1, 2000, through March 31, 2001 (hereafter called the 2001 time interval), and November 1, 2014, through March 31, 2015 (hereafter called the 2015 time interval). These time intervals were chosen to maximize the number of unique monitoring sites with data in at least one of the time intervals. Median groundwater-elevation measurements were included at 78 monitoring sites in the 2001 time interval (fig. 4*A*) and at 65 monitoring sites in the 2015 time interval (fig. 4*B*). In total, 87 unique monitoring sites were included in the spatial assessment by using the 2001 time interval and 2015 time interval (table 1).

The data required for each monitoring site for the spatial optimization algorithm included site number, standard deviation of the depth-to-groundwater measurements at the site over the entire period of record, estimated uncertainty of groundwater-elevation measurements, and median winter groundwater elevation for the time interval of interest (tables 1 and 2). The standard deviations were used by the GA to retain sites with large groundwater-level variability over time (third objective in the multiobjective function), and the groundwater-elevation uncertainties were used to retain monitoring sites with higher measurement accuracy (fourth objective in the multiobjective function). Median depth-to-groundwater measurements for both the 2001 time interval and 2015 time interval were converted to groundwater elevation above NAVD 88 (table 2). The standard deviations of the depth-to-groundwater measurements ranged from 0.39 to 92 feet (fig. 5, table 1), with a mean of about 13 feet and a median of 8.2 feet. Uncertainty of groundwater-elevation measurements was estimated as the accuracy of the land surface elevation (fig. 6, table 1), which varied between ±0.01 and 10 feet with a standard deviation of about 2 feet and a mean of about ±2 feet.

Implementation

The spatial assessment was performed in the R programming language (version 3.5.1; R Core Team, 2019). The implementation of the spatial assessment is described in the following sections.

Spatial Trend Modeling

Groundwater-elevation trend within the study area was estimated separately for the 2001 time interval and 2015 time interval so that trend could appropriately be accounted for during semivariogram development and universal kriging. A linear polynomial function of the coordinate variables (eq. 1) was used in both instances.

$$z(s) = \beta_0 + \beta_1 x(s) + \beta_2 y(s) \tag{1}$$

where

$z(s)$ is median groundwater elevation at point s, in feet above NAVD 88;

β_0 is deterministic unknown trend coefficient, in feet above NAVD 88;

β_i is deterministic unknown trend coefficients (i = 1 through 2);

$x(s)$ is easterly coordinate at point s, in feet; and

$y(s)$ is northerly coordinate at point s, in feet.

Residual groundwater elevations for use in semivariogram development and kriging were calculated as the median groundwater elevations minus the appropriate trend model estimates. Residuals ranged from −62 to 177 feet and −119 to 46 feet for the 2001 time interval and 2015 time interval, respectively. Standard deviations of residuals for the 2001 time interval and 2015 time interval were similar, with 2001 at approximately 32 feet and 2015 at approximately 26 feet. Both the 2001 time interval and 2015 time interval had a residual mean of 0.0 feet and displayed minimal spatial bias with Pearson correlation coefficients of 0.0 relative to X and Y coordinates.

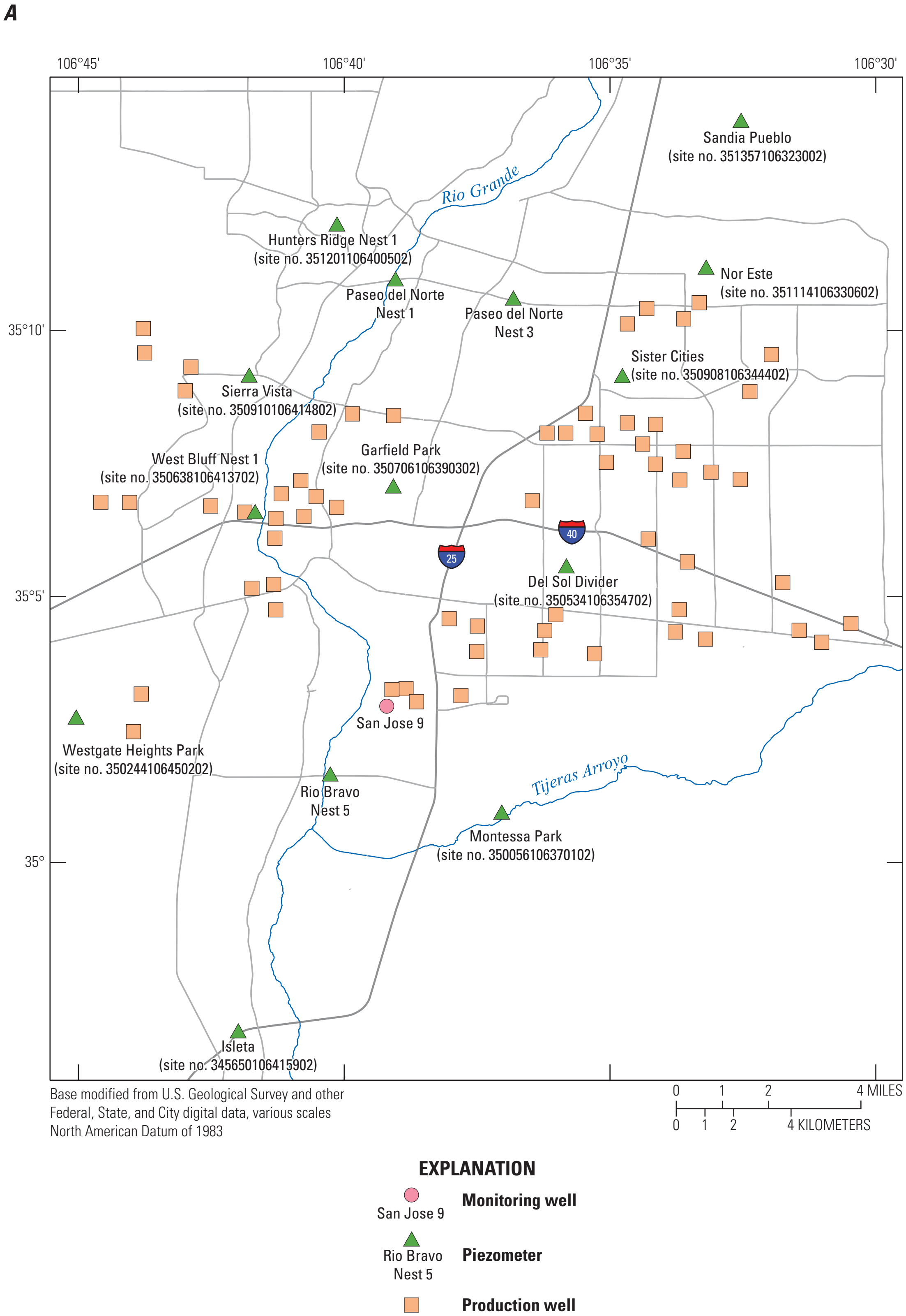


Figure 4. The locations of sites used for the spatial assessment for the *A*, November 1, 2000, through March 31, 2001 (2001), and *B*, November 1, 2014, through March 31, 2015 (2015), time intervals.

B

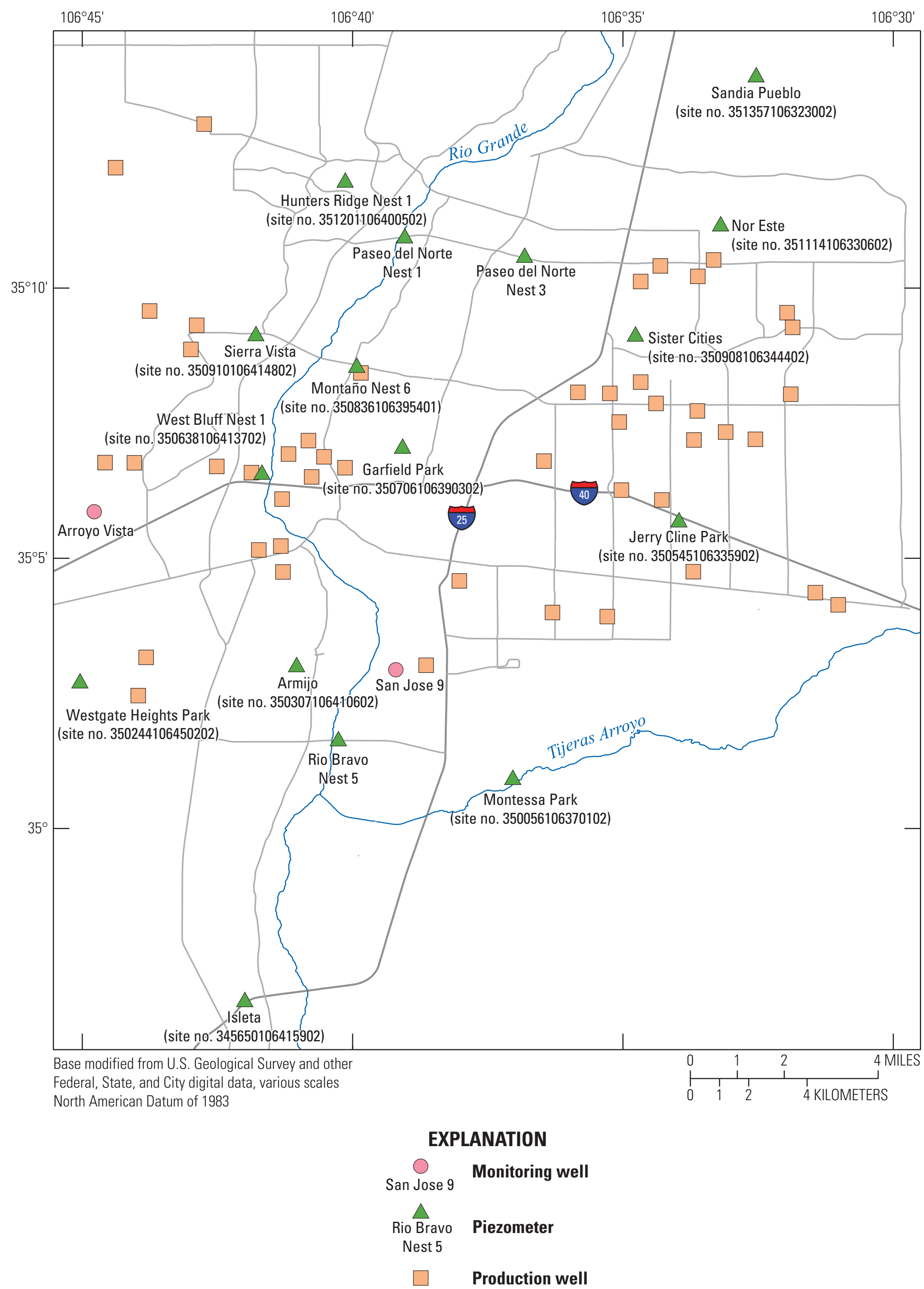


Figure 4. The locations of sites used for the spatial assessment for the *A*, November 1, 2000, through March 31, 2001 (2001), and *B*, November 1, 2014, through March 31, 2015 (2015), time intervals.—Continued

Table 2. Groundwater elevations used for the spatial assessment.

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; NA, not applicable]

USGS site number	Site name	2001 median winter groundwater elevation (feet above NAVD 88)	2015 median winter groundwater elevation (feet above NAVD 88)
345650106415902	Isleta	4,891.94	4,890.31
350056106370102	Montessa Park	4,878.89	4,883.78
350244106450202	Westgate Heights Park	4,882.25	4,898.70
350256106390801	San Jose 9	4,908.64	4,917.92
350307106410602	Armijo	NA	4,911.66
350534106354702	Del Sol Divider	4,865.00	NA
350545106335902	Jerry Cline Park	NA	4,868.05
350638106413702	West Bluff Nest 1	4,948.56	4,936.64
350706106390302	Garfield Park	4,922.08	4,930.11
350836106395401	Montaño Nest 6	NA	4,944.05
350908106344402	Sister Cities	4,890.98	4,902.46
350910106414802	Sierra Vista	4,954.67	4,950.78
351114106330602	Nor Este	4,910.14	4,924.06
351201106400502	Hunters Ridge Nest 1	4,950.92	4,948.68
351357106323002	Sandia Pueblo	4,956.20	4,956.95
350138106401103	Rio Bravo Nest 5	4,919.36	4,920.12
351059106385903	Paseo del Norte Nest 1	4,948.79	4,951.08
351035106364703	Paseo del Norte Nest 3	4,936.24	4,940.31
350552106444601	Arroyo Vista	NA	4,916.06
350445106411501	Atrisco #2	4,937.63	4,940.71
350508106411901	Atrisco #3	4,942.80	4,944.86
350509106414401	Atrisco #4	4,939.49	4,940.89
350359106362401	Burton #1	4,864.46	4,879.22
350421106361001	Burton #2	4,860.52	NA
350440106355801	Burton #3	4,859.50	NA
350355106351501	Burton #5	4,861.84	4,875.23
350606106341101	Charles #2	4,849.39	4,867.62
350615106345901	Charles #5	NA	4,879.93
350646106443201	College #1	4,902.16	4,915.86
350647106440001	College #2	4,895.26	4,906.20
351025106341601	Coronado #1	4,894.24	4,912.43
351007106343801	Coronado #2	4,907.20	4,919.70
351302106424701	Corrales #3	NA	4,929.02
351215106441301	Corrales #4	NA	4,824.00
350642106401101	Duranes #1	4,932.95	4,938.87
350708106405801	Duranes #2	4,952.10	4,951.58
350629106405101	Duranes #3	4,946.52	4,946.09
350628106411501	Duranes #4	4,948.50	NA
350605106411801	Duranes #5	4,948.43	4,944.98
350653106403001	Duranes #6	4,934.88	4,940.96
350655106395001	Duranes #7	4,954.76	4,954.06
350642106422801	Gonzales #1	4,926.39	4,913.60
350635106415001	Gonzales #2	4,943.49	4,939.77
350827106395001	Griegos #1	4,944.17	4,947.43

Table 2. Groundwater elevations used for the spatial assessment.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988; NA, not applicable]

USGS site number	Site name	2001 median winter groundwater elevation (feet above NAVD 88)	2015 median winter groundwater elevation (feet above NAVD 88)
350802106402901	Griegos #3	4,952.27	NA
350821106390101	Griegos #4	4,937.76	NA
350309106434501	Leavitt #1	4,882.05	4,902.62
350223106435401	Leavitt #3	4,877.95	4,896.90
350752106342101	Leyendecker #1	4,863.63	4,879.55
350727106340801	Leyendecker #2	4,860.99	NA
350819106344001	Leyendecker #3	4,870.76	4,888.75
350815106340601	Leyendecker #4	4,870.93	NA
350430106302401	Lomas #1	5,035.48	NA
350422106312401	Lomas #5	4,852.68	4,870.53
350408106310101	Lomas #6	4,854.83	4,871.38
350517106314401	Love #1	4,859.37	NA
350538106333001	Love #8	4,849.85	NA
350308106374601	Miles Road #1	4,886.67	NA
350800106315001	Ponderosa #2	NA	4,853.72
350851106322001	Ponderosa #6	4,888.73	NA
350401106331401	Ridgecrest #3	4,847.85	NA
350445106334001	Ridgecrest #4	4,846.88	4,868.80
350420106334401	Ridgecrest #5	4,847.97	NA
350316106384801	San Jose #1	4,910.48	NA
350315106390401	San Jose #2	4,851.45	NA
350301106383601	San Jose #3	4,914.03	4,923.90
350648106362501	Santa Barbara #1	4,874.16	4,885.80
350744106333501	Thomas #5	4,865.56	4,889.71
350720106330401	Thomas #6	4,863.22	4,884.18
350719106333401	Thomas #7	4,861.70	4,889.62
350711106323101	Thomas #8	4,859.00	4,884.55
350805106354901	Vol Andia #1	4,877.97	4,887.40
350732106350101	Vol Andia #2	4,864.70	4,879.22
350803106351101	Vol Andia #4	4,872.45	4,885.63
350809106360901	Vol Andia #5	4,880.69	NA
350828106352101	Vol Andia #6	4,881.81	NA
350950106434001	Volcano Cliffs #1	4,878.19	4,883.74
351007106434201	Volcano Cliffs #3	4,889.74	NA
350931106315501	Walker #3	4,908.26	4,908.24
350918106315401	Walker #4	NA	4,901.49
351029106332001	Webster #1	4,897.06	4,916.21
351013106333501	Webster #2	4,893.05	4,911.45
350426106372601	Yale #1	4,882.94	NA
350358106372901	Yale #2	4,886.00	NA
350435106380101	Yale #3	4,893.40	4,905.10
350918106425401	Zamora #1	4,924.81	4,917.54
350852106425701	Zamora #2	4,910.45	4,910.28

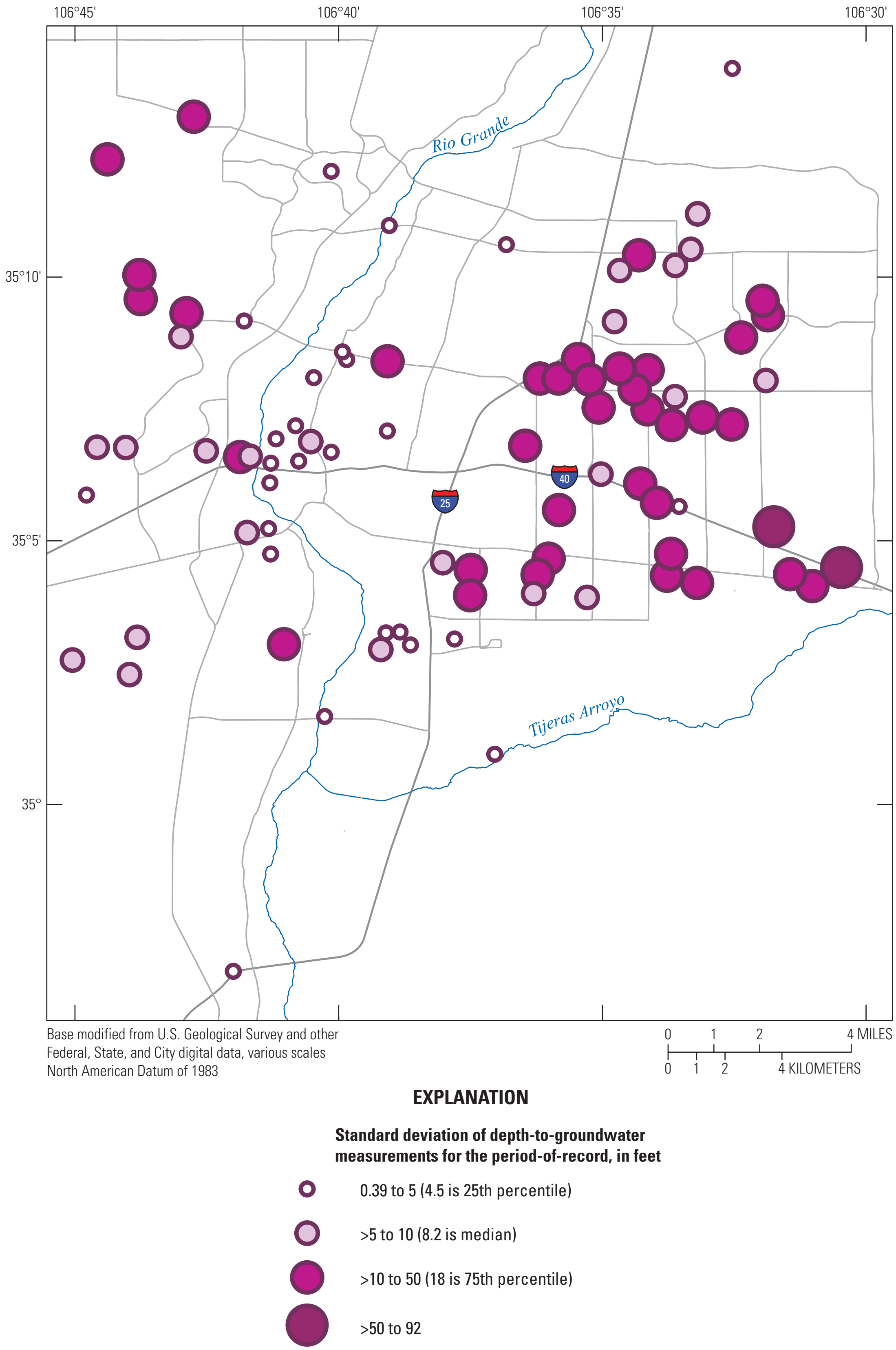


Figure 5. Standard deviations of depth-to-groundwater measurements for the entire period of record used for the spatial assessment.

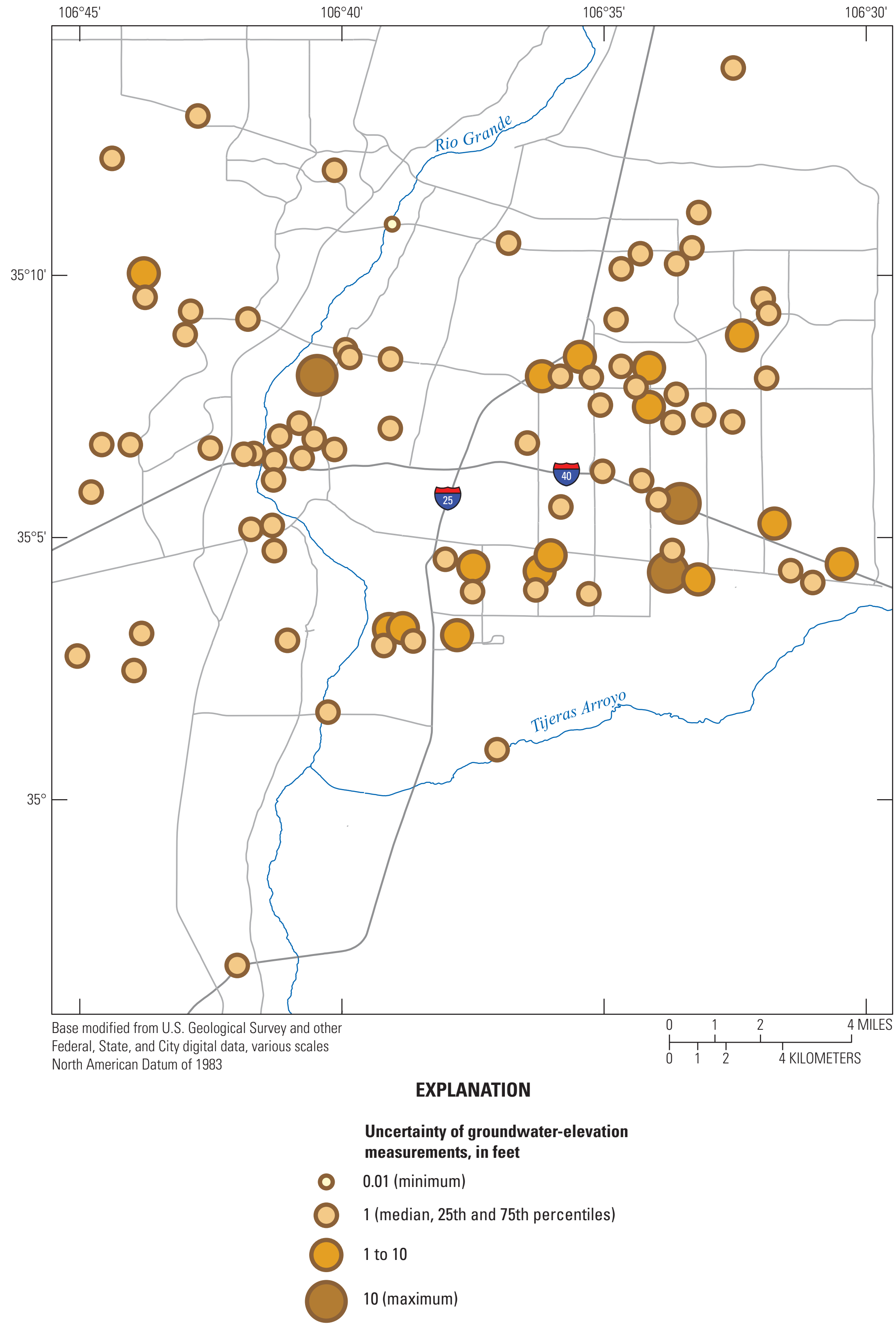


Figure 6. Uncertainties of groundwater-elevation measurements used for the spatial assessment.

Semivariogram Development

Separate empirical and theoretical semivariograms were developed for the 2001 time interval and the 2015 time interval by using the residuals from the trend modeling. Empirical semivariograms were estimated using the “variogram” function within the R package “gstat,” version 2.0-0 (Pebesma and Graeler, 2019). This function calculates the semivariance for all pairs of monitoring sites as one-half the variance of the difference between the residual groundwater elevation at the sites. The empirical semivariogram is then computed by grouping the semivariance values into bins based on the distance (lag) between monitoring site pairs and averaging the semivariance values within each bin. The bin width was set to 2,500 feet, about the minimum distance between sites, and the spatial separation distance to which monitoring site pairs were included in the semivariance estimates was set at 50,000 feet (about half of the maximum separation distance between monitoring site pairs; Fisher, 2013).

Kriging relies on a continuous theoretical semivariogram that is fit to the experimental semivariogram. Fitting semivariograms requires the consideration of semivariogram shape, which is strongly influenced by three components called the nugget, sill, and range. The nugget is the semivariance value at a lag of zero, which typically is related to noise in the data and is user specified by visual inspection. The sill is the semivariance value in which the mean semivariance no longer changes with increasing distance. The range is the distance in which the sill is reached and indicates the distance over which data are correlated. The theoretical semivariograms were modeled using the “fit.variogram” function within the R package “gstat,” version 2.0-0 (Pebesma and Graeler, 2019). Circular, exponential, gaussian, linear, and spherical models were all considered in the fitting procedure. For the 2001 time interval, the theoretical semivariogram was modeled as spherical with no nugget, a sill of 1,140 square feet, and a range of 33,000 feet (fig. 7*A*). For the 2015 time interval, the theoretical semivariogram was modeled as spherical with no nugget, a sill of 770 square feet, and a range of 41,000 feet (fig. 7*B*). The evident differences in the semivariograms from the 2001 time interval and 2015 time interval are likely related to changes in the spatial distribution of groundwater elevations over time and, to a lesser degree, differences in monitoring sites used for each time interval.

Kriged Surface Development and Assessment

Universal kriging was used to interpolate groundwater elevation and quantify uncertainty associated with the interpolated values for the 2001 time interval and 2015 time interval. Interpolation was performed on a uniform grid oriented north-south and east-west, with a square block side length of 500 feet. The spatial resolution of the grid was chosen as less than the minimum spatial separation distance between monitoring sites. The area of analysis was defined by the generalized boundary of drawdown from Galanter and Curry (2019) and the spatial extent of the monitoring sites used in the interpolation. The “krige” function within the R package “gstat,” version 2.0-0 (Pebesma and Graeler, 2019), was used to perform the kriging interpolation.

The interpolated groundwater elevations for the 2001 time interval (fig. 8*A*) ranged from 4,838 to 5,015 feet above NAVD 88, with trends similar to the 2002 groundwater-elevation contours developed for the production zone by Bexfield and Anderholm (2002). The interpolated groundwater elevations for the 2015 time interval (fig. 8*B*) ranged from 4,825 to 4,957 feet above NAVD 88, with trends similar to the 2016 groundwater-elevation contours developed for the production zone by Galanter and Curry (2019). The largest differences between the interpolated groundwater elevations and the published contours are near the north and southeast margins of the grid; these differences are attributed to nearby monitoring sites that are managed by other entities and were not included in the interpolation.

Standard errors from kriging ranged from 2.5 to 42 feet for the 2001 time interval (fig. 9*A*) and from 1.9 to 33 feet for the 2015 time interval (fig. 9*B*). The north and south margins of the grid had the largest prediction uncertainty for the 2001 time interval, and the south margin of the grid had the largest prediction uncertainty for the 2015 time interval. Larger standard error results from the lack of monitoring sites in these areas.

Leave-one-out cross validation was used to (1) test the performance of the theoretical semivariogram model, (2) test the validity of the stationarity assumption for kriging, (3) identify important monitoring sites for interpolating the groundwater elevation, and (4) identify sites where groundwater levels are most dynamic. The leave-one-out cross validation was performed using the “RunCrossValidation” function within the R package “ObsNetwork,” version 1.0.0 (Fisher, 2013). The leave-one-out cross validation removes each measurement point from the dataset one by one and estimates the groundwater elevation at the removed site by using the theoretical semivariogram model developed for the entire dataset to krige with the remaining data. The estimation error from leave-one-out cross validation was calculated as the difference between the measured and estimated groundwater elevation at the omitted site.

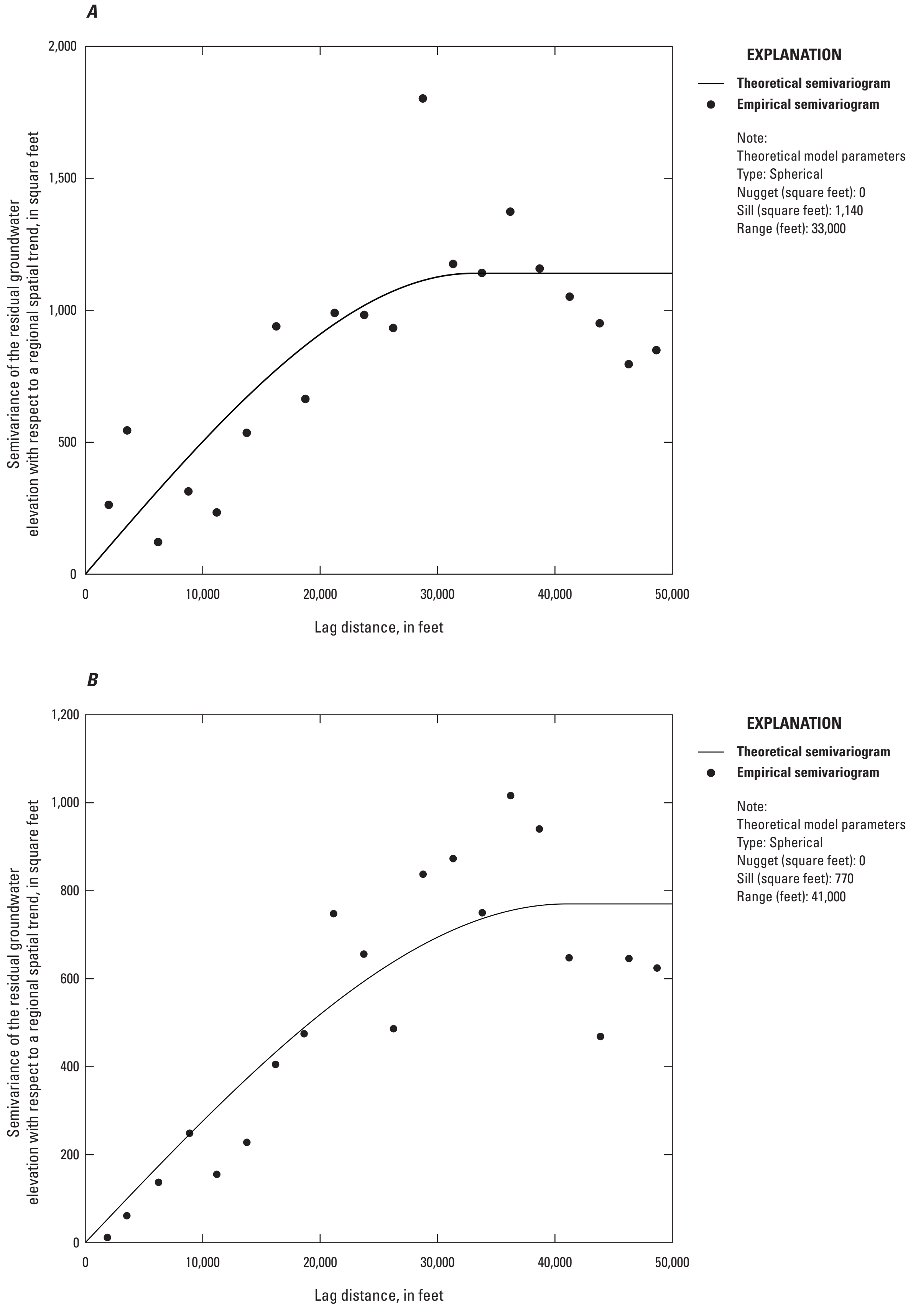
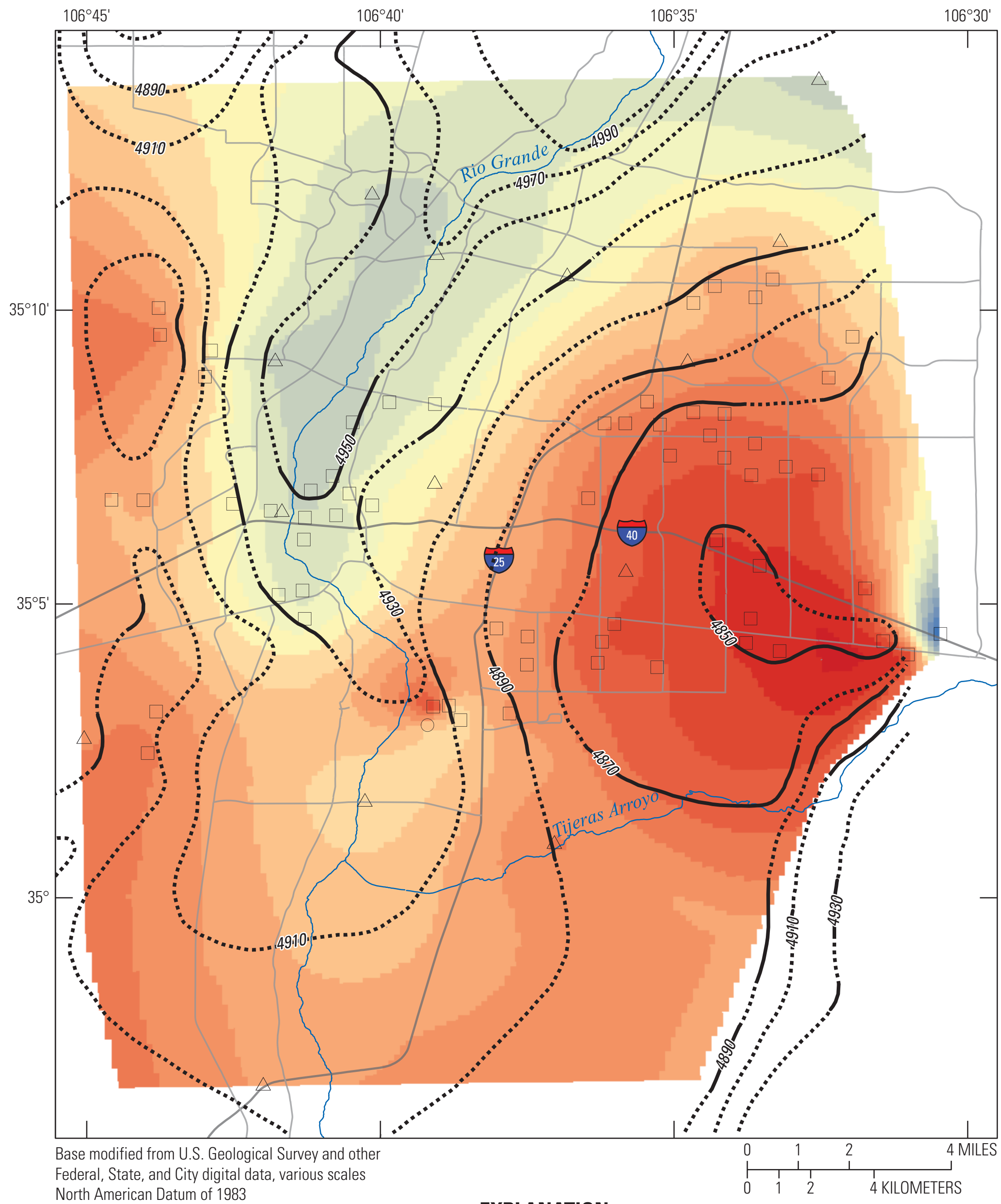


Figure 7. Semivariogram analysis of groundwater-elevation residuals for the *A*, 2001 time interval and *B*, 2015 time interval.

A



EXPLANATION

Interpolated 2001 time interval groundwater elevations from the application of kriging, in feet above North American Vertical Datum of 1988

- ≥ 4,838 to 4,840
- > 4,840 to 4,850
- > 4,850 to 4,860
- > 4,860 to 4,870
- > 4,870 to 4,880
- > 4,880 to 4,890
- > 4,890 to 4,900
- > 4,900 to 4,910
- > 4,910 to 4,920
- > 4,920 to 4,930
- > 4,930 to 4,940
- > 4,940 to 4,950
- > 4,950 to 4,960
- > 4,960 to 4,970
- > 4,970 to 4,980
- > 4,980 to 4,990
- > 4,990 to 5,000
- > 5,000 to 5,010
- > 5,010 to 5,015

Groundwater-level contours from Bexfield and Anderholm (2002)— Water-level contour, 2002. Interval 20 feet. Dashed where inferred. Datum is sea level

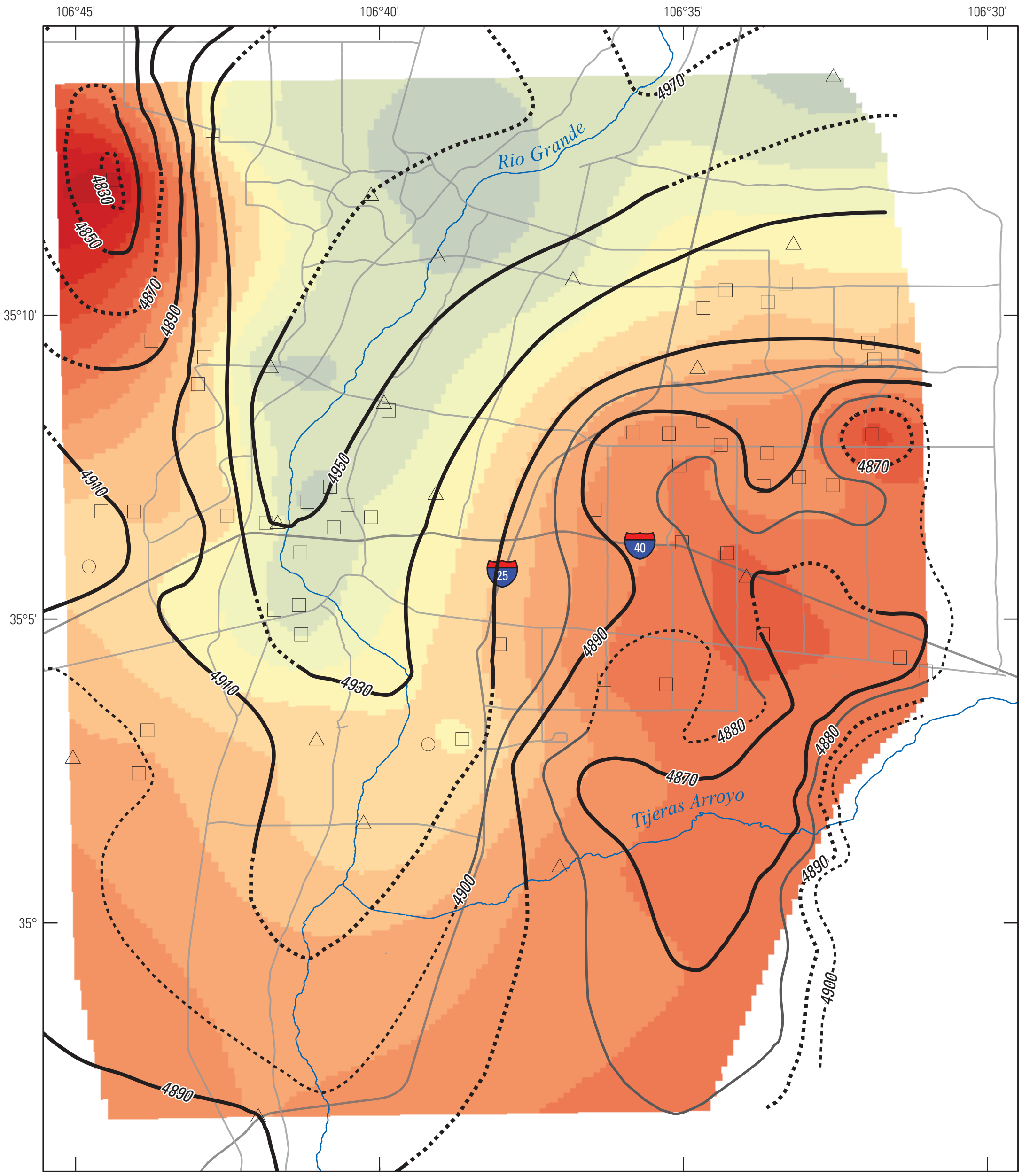
- at 20 foot interval
- at 20 foot interval with greater uncertainty

Monitoring site

- ○ Monitoring well
- △ Piezometer
- □ Production well

Figure 8. The interpolated groundwater-elevation kriged surface for the *A*, 2001 time interval and *B*, 2015 time interval.

B



Base modified from U.S. Geological Survey and other Federal, State, and City digital data, various scales
North American Datum of 1983

EXPLANATION

Interpolated 2015 time interval groundwater elevations from the application of kriging, in feet above North American Vertical Datum of 1988

- ≥ 4,825 to 4,830
- > 4,830 to 4,840
- > 4,840 to 4,850
- > 4,850 to 4,860
- > 4,860 to 4,870
- > 4,870 to 4,880
- > 4,880 to 4,890
- > 4,890 to 4,900
- > 4,900 to 4,910
- > 4,910 to 4,920
- > 4,920 to 4,930
- > 4,930 to 4,940
- > 4,940 to 4,950
- > 4,950 to 4,957

Groundwater-level contours from Galanter and Curry (2019)— Shows elevation at which water levels would have stood in tightly cased wells, 2016. Dashed where approximately located. Contour interval, in feet, is variable. Datum is North American Vertical Datum of 1988

- at 10 foot interval
- at 20 foot interval
- at 10 foot interval with greater uncertainty
- at 20 foot interval with greater uncertainty

Monitoring site

- Monitoring well
- Piezometer
- Production well

Figure 8. The interpolated groundwater-elevation kriged surface for the *A*, 2001 time interval and *B*, 2015 time interval.—Continued

A

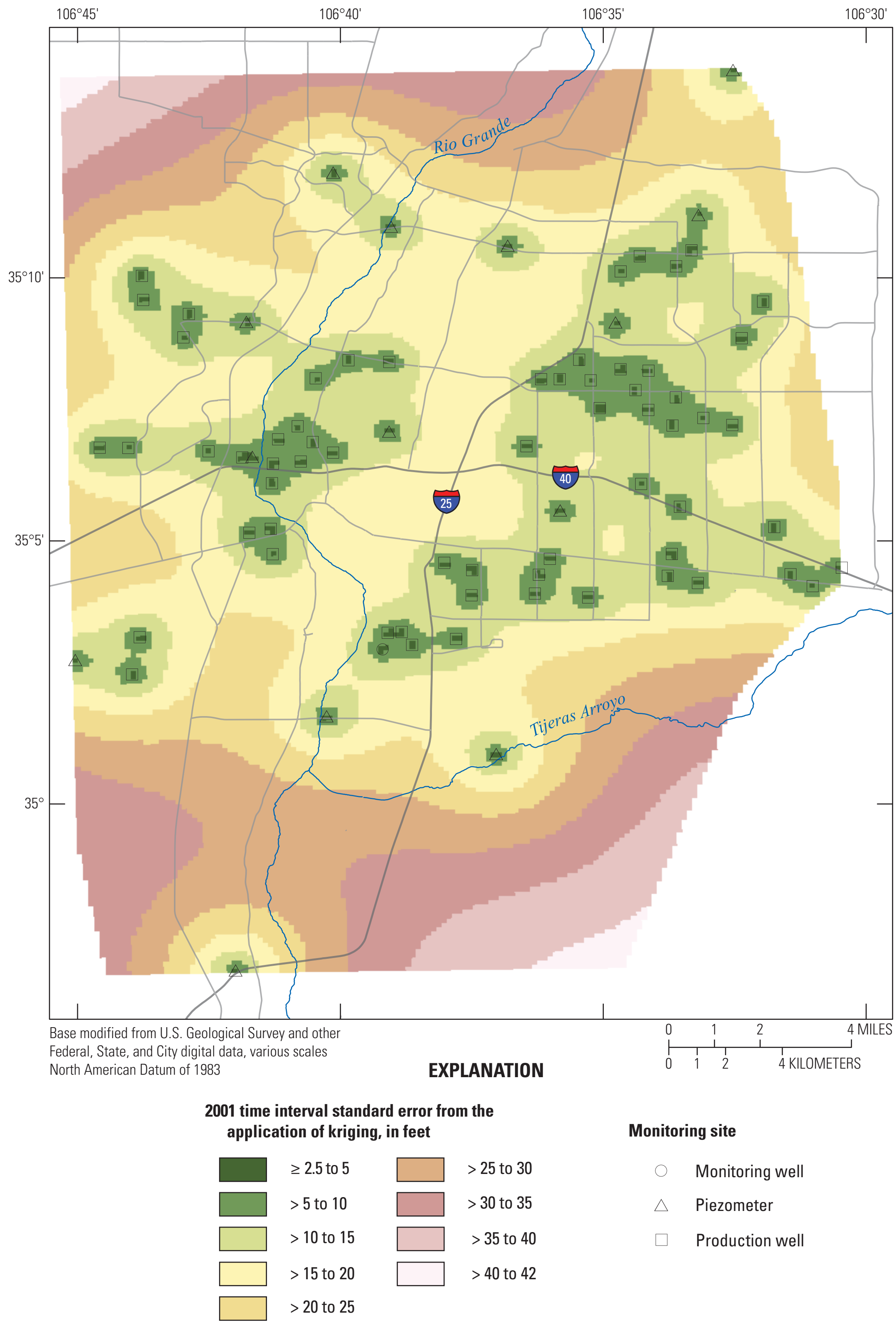


Figure 9. The standard errors from kriging groundwater elevation for the *A*, 2001 time interval and *B*, 2015 time interval.

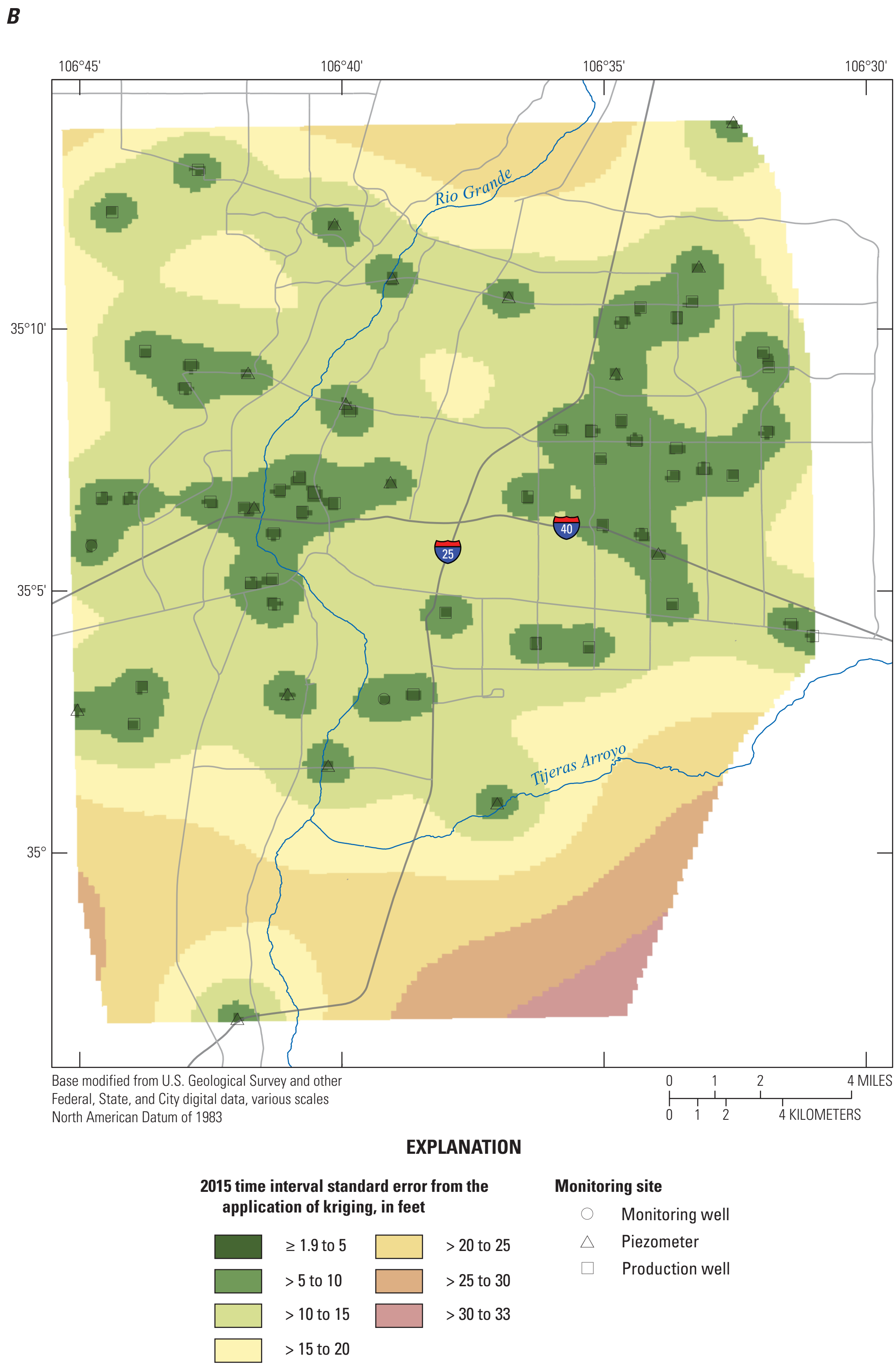


Figure 9. The standard errors from kriging groundwater elevation for the *A*, 2001 time interval and *B*, 2015 time interval.—Continued

Estimation errors from leave-one-out cross validation for the 2001 time interval ranged from −74 to 175 feet (fig. 10*A*), with a mean of 1.8 feet and a standard deviation of 25 feet. Estimation errors for the 2015 time interval ranged from −91 to 54 feet (fig. 10*B*), with a mean of −0.18 feet and a standard deviation of 16 feet. The mean estimation errors, which would ideally be zero, were small (with a magnitude less than 2 feet) for the 2001 time interval and 2015 time interval, indicating an absence of systematic errors from kriging (Fisher, 2013). Sensitive monitoring sites for interpolating groundwater elevation (sites with larger magnitudes of estimation error [Fisher, 2013]) during the 2001 time interval and 2015 time interval are generally located west of the Rio Grande and near the margins of the grid (figs. 10*A* and *B*). Large magnitudes of estimation error at monitoring sites west of the Rio Grande and near the north and south margins of the grid likely result from a low sampling density in these areas. For the 2001 time interval, large magnitudes of estimation error at monitoring sites near the east margin of the grid and in a region within about 1 mile east of the Rio Grande and 4 miles south of Interstate 40 likely result from the substantial change in groundwater elevation between monitoring sites.

Scatterplots of measured groundwater elevation and estimated groundwater elevation at the omitted sites are shown in figure 11*A* for the 2001 time interval and figure 12*A* for the 2015 time interval. The Pearson correlation coefficient between the measured and estimated groundwater elevations was 0.75 for the 2001 time interval and 0.83 for the 2015 time interval, indicating that the theoretical semivariogram models well represent the general structure of the data. The slope of the linear regression through the points (black dashed line) is similar to a 1:1 line (gray solid line) (figs. 11*A* and 12*A*), indicating that the estimates are not conditionally biased (Fisher, 2013). A strong linear correlation between estimation error and the estimated groundwater elevation at the omitted sites was not observed for either time interval (figs. 11*B* and 12*B*; Pearson correlation coefficients of −0.17 and −0.07 for the 2001 time interval and 2015 time interval, respectively), with estimation error equally scattered around a horizontal line, indicating that the estimation error is independent of the magnitude of the estimated groundwater elevation from cross validation and that stationarity may be assumed for the residual groundwater elevations used for kriging (Fisher, 2013). The largest magnitudes of estimation error for the 2001 time interval and 2015 time interval were at Lomas no. 1 (site no. 350430106302401) and Corrales no. 4 (site no. 351215106441301), respectively (figs. 11 and 12, table 1). Lomas no. 1 (site no. 350430106302401) is near the east margin of the grid (fig. 10*A*), and the relatively large estimation error likely results from the substantial change in groundwater elevation between monitoring sites in this area. Corrales no. 4 (site no. 351215106441301) is near the west margin of the grid (fig. 10*B*), and the relatively large estimation error likely results from the low monitoring density in this area.

Multiobjective Function

The objective function terms were quantified following Fisher (2013). The first term represents the overall interpolation error in the kriged surface when using a reduced network and is quantified as the mean of the standard error (square root of the estimated variance from kriging) at all points (or nodes) in the kriging grid. The second term represents the difference between measured and estimated groundwater elevations and is quantified as the root-mean-square error (RMSE) at sites removed from the network. The third term represents the variability of groundwater levels over time and is quantified as the mean standard deviation of depth-to-groundwater measurements over the entire period of record (fig. 5, table 1) for the sites removed from the network. The fourth term represents the accuracy of groundwater-elevation measurements and is quantified as the mean uncertainty of groundwater-elevation measurements (fig. 6, table 1) at sites not removed from the network. These terms can be weighted to emphasize terms of greater importance. The sum of the weighted terms represents the fitness value.

Genetic Algorithm

A GA was used to identify optimal networks that minimized the multiobjective function. In the following discussion, a chromosome represents a single solution for the monitoring sites to remove from the network, with the monitoring sites representing genes. The GA transforms a set (population) of individual objects (chromosomes), each with an associated fitness value, into a new population (for example, the next generation) by using operations mimicking survival of the fittest, sexual recombination, and genetic mutation (Fisher, 2013). The GA operates on the solution space of multiple chromosomes (called the population), where the number of chromosomes is specified as the population size. The following discussion of GA is adapted from Fisher (2013).

A GA run begins (first generation or iteration) with a population of randomly selected chromosomes. The fitness value is computed for each chromosome in the population, and a subset of chromosomes with the smallest fitness values are copied (survive) to a new population (next generation). The specified elitism rate controls the number of chromosomes that are guaranteed to survive to the next iteration. The rest of the chromosomes that compose the new population can be modified from the previous population on the basis of the specified crossover and mutation probabilities.

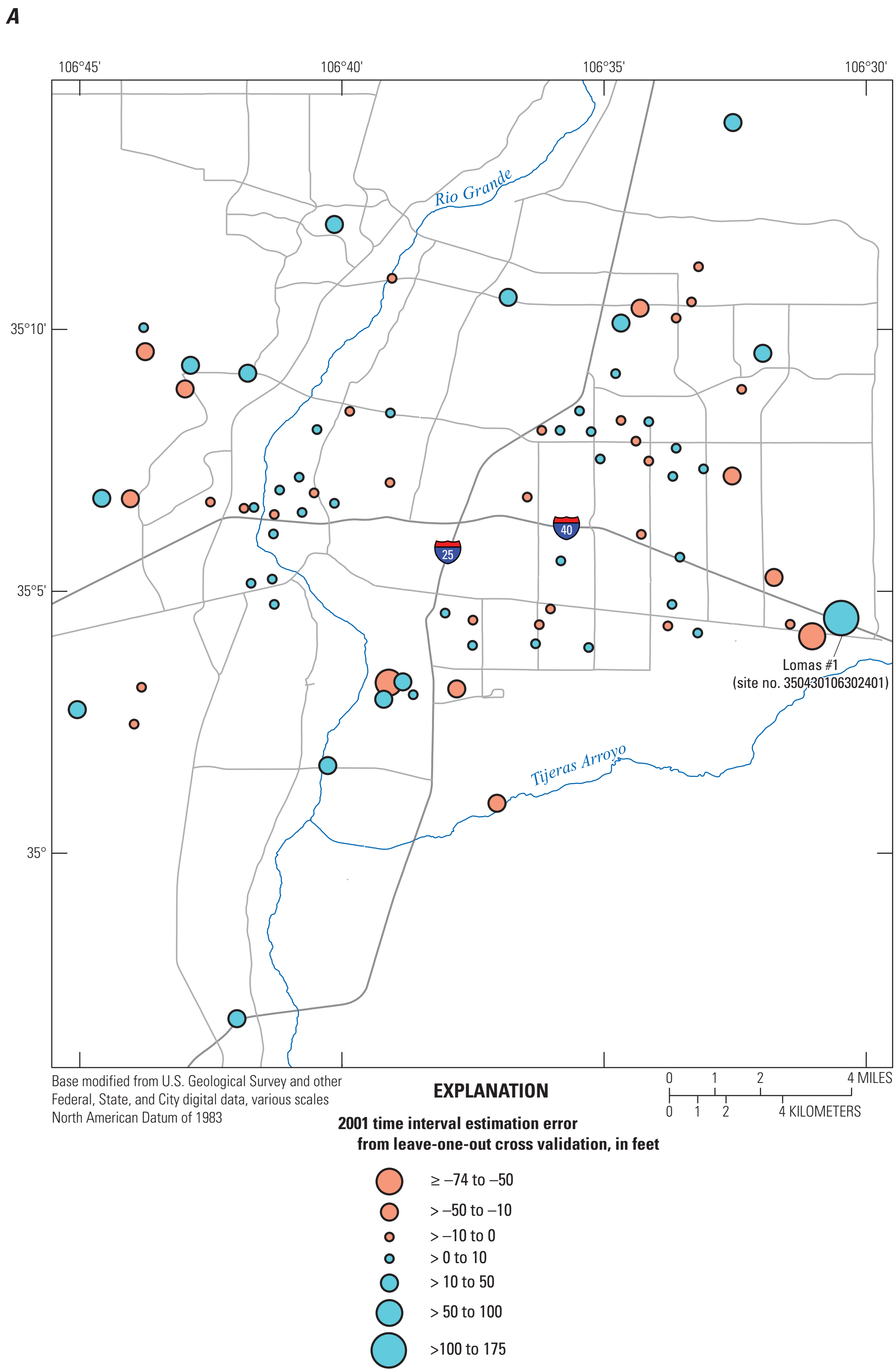


Figure 10. The estimation error from leave-one-out cross validation for the *A*, 2001 time interval and *B*, 2015 time interval.

B

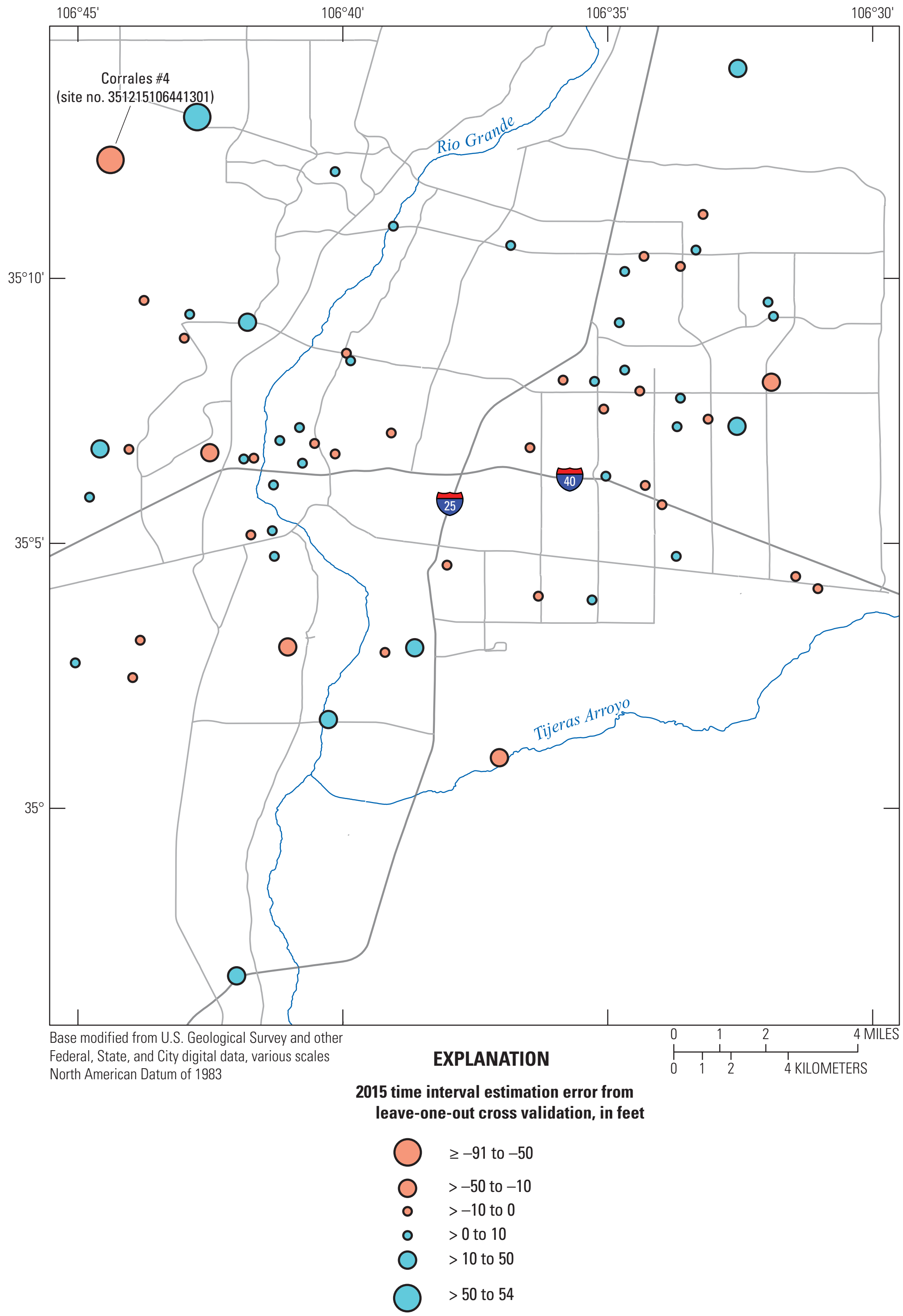


Figure 10. The estimation error from leave-one-out cross validation for the *A*, 2001 time interval and *B*, 2015 time interval.—Continued

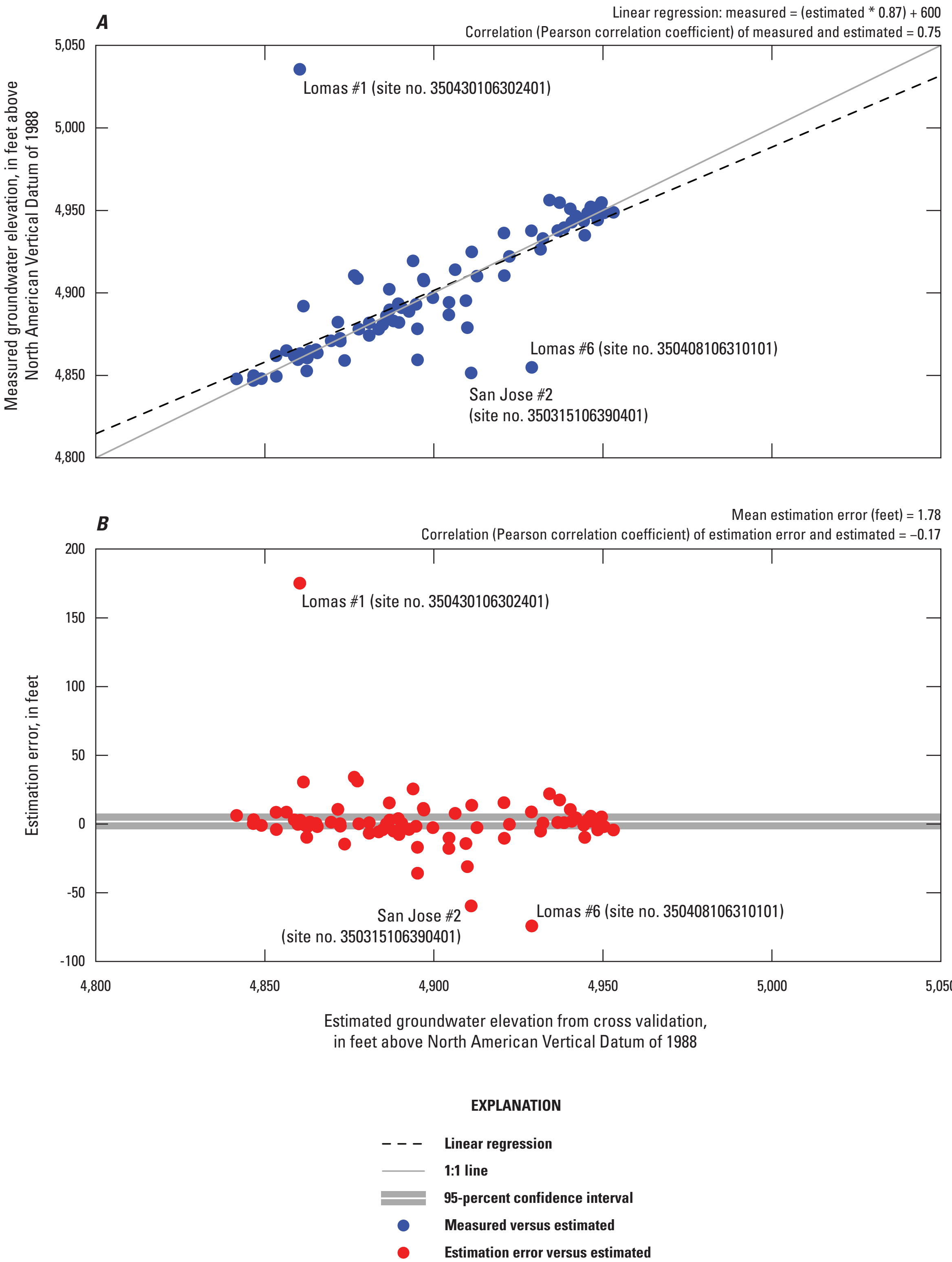


Figure 11. *A*, Measured and estimated groundwater elevations and *B*, estimation error and estimated groundwater elevations from leave-one-out cross validation for the 2001 time interval at the omitted sites.

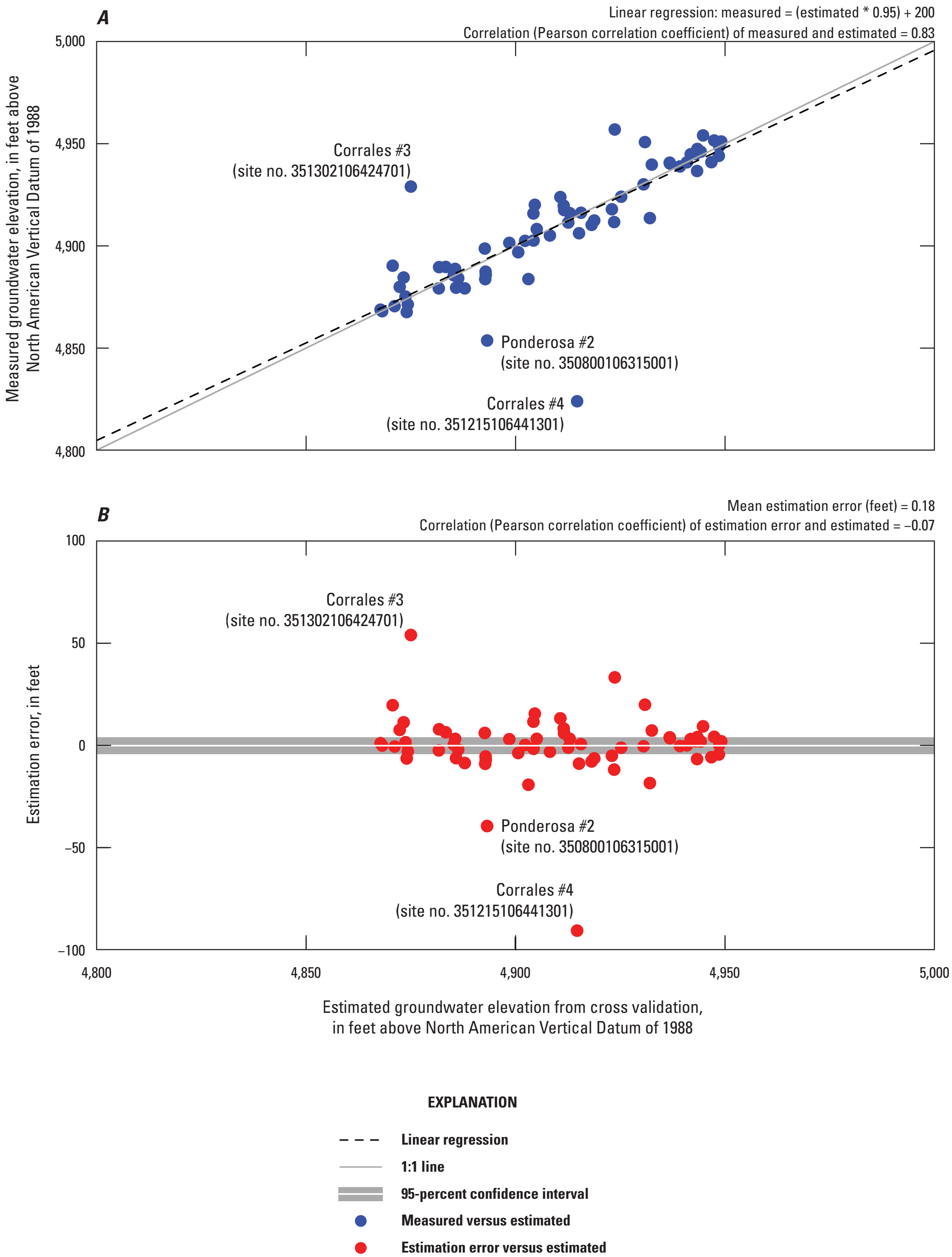


Figure 12. *A*, Measured and estimated groundwater elevations and *B*, estimation error and estimated groundwater elevations from leave-one-out cross validation for the 2015 time interval at the omitted sites.

Crossover (also called sexual recombination) combines parts of two “parent” chromosomes to produce two “child” chromosomes. To discourage the introduction of duplicate monitoring sites into the child chromosomes during crossover, a child with duplicates is aborted up to a specified maximum number of times, and crossover is reattempted. An upper limit is set on the maximum number of times a child chromosome can be aborted to avoid biasing crossover. If the limit is exceeded, the child survives, but its fitness value is set artificially large (penalized) as the product of the specified penalty coefficient and the number of duplicate monitoring sites. The crossover probability is the probability that crossover will occur between two parents.

Mutation follows crossover and involves altering a gene on a child chromosome. The monitoring site that is mutated and the replacement site are randomly selected. The mutation probability is the probability that mutation will occur on a child chromosome.

The GA continues to iterate through new generations until one of the following specified criteria is reached: (1) the maximum number of iterations or (2) the maximum number of consecutive iterations without any improvement in the minimum (best) fitness value. Finding a globally optimal solution is not guaranteed with either of these criteria. However, the probability of finding a globally optimal solution increases as the magnitude of these criteria increases.

Genetic Algorithm Parameter Selection

Version 1.0.0 of the “ObsNetwork” (Fisher, 2013) R package was used to perform the spatial assessment. The “OptimizeNetwork” command of this package couples the aforementioned kriging utilities to the GA to identify optimal network configurations. Input parameters to “OptimizeNetwork” control the resolution of the kriging grid, the number of sites to remove from the observation network, the weighting coefficients applied to the objective function terms, the penalty assigned to GA solutions with duplicate monitoring sites, and the performance of the GA (table 3).

Table 3. Control parameter values used for the genetic algorithm sensitivity analysis (base-case value) and spatial assessment (final value).

[Base-case value, control parameter values specified as base-case conditions in the sensitivity analysis. Final value, control parameter values used in the optimization assessment of the observation network. RMSE, root-mean-square error]

Control parameter	Unit	Base-case value	Final value
Kriging analysis			
Length of square block side in uniform kriging grid (kriging grid resolution)	feet	2,500	1,500
Multi-objective problem formulation			
Number of sites to remove from the observation network (number of sites removed)	unitless	20 (2001 interval) 15 (2015 interval)	2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40
Weighting coefficients on individual objective functions			
Weight on the mean standard error function	unitless	100	100
Weight on the RMSE function	unitless	1	1
Weight on the mean standard deviation function	unitless	1	1
Weight on the mean measurement uncertainty function	unitless	1	1
Penalty function			
Penalty coefficient	feet	1,000,000	1,000,000
Genetic algorithm			
Number of chromosomes in a population (population size)	unitless	2,000	1,750
Genetic operations			
Fraction of the population that is guaranteed to survive to the next iteration (elitism rate)	unitless	0.05 (5 percent)	0.1 (10 percent)
Probability that crossover will occur between two parents (crossover probability)	unitless	0.80 (80 percent)	0.80 (80 percent)
Probability that mutation will occur on a child chromosome (mutation probability)	unitless	0.05 (5 percent)	0.175 (17.5 percent)
Maximum number of times a child chromosome can be aborted during crossover	unitless	10	10
Terminating conditions			
Maximum number of iterations	unitless	100	1,000,000
Maximum number of consecutive iterations without any improvement in the best fitness value	unitless	1,000,000	500

Following the approach of Fisher (2013), the optimization assessment emphasized observation network spatial coverage by applying a larger weighting coefficient to the first term (table 3; weighting coefficient of 100 applied to the mean standard error) and smaller weighting coefficients to the remaining terms (table 3; weighting coefficient of 1 applied to the second term, RMSE; to the third term, mean standard deviation; and to the fourth term, mean measurement uncertainty) in the multiobjective function prior to calculating the fitness value. The weighted objective values were then summed to produce the fitness value, which was minimized (best fitness value) during the GA search of the solution space (Fisher, 2013). The assignment of weights to the objective function terms is subjective, and the multiobjective formulation is poorly suited for determining tradeoffs between and simultaneously minimizing each objective (Fisher, 2013). Therefore, the sensitivity of the GA to changes in the weighting coefficient applied to each objective was not tested. Not testing the sensitivity of the GA to changes in the weighting coefficient applied to each objective was deemed appropriate on the basis of past studies (Fisher, 2013) and the network objectives of the ABCWUA.

Following the approach of Fisher (2013), a series of GA sensitivity runs were performed for six of the control parameters (table 4; number of sites removed, kriging grid resolution, population size, elitism rate, crossover probability, and mutation probability) for the 2001 time interval and 2015 time interval to find reasonable parameter values ("final value" column of table 3) for performing the spatial assessment. The sensitivity analysis involved incrementally changing one of the parameters while holding all other parameters constant ("base-case value" column of table 3, and bolded control parameter values in table 4). The GA iteration stopping criteria were set at 100 for the maximum number of iterations and 1,000,000 for the maximum number of consecutive iterations without any improvement in the best fitness value (table 3). Following Fisher (2013), four metrics were analyzed to assess GA performance in the sensitivity runs: (1) best fitness value, (2) computation time, (3) consecutive iterations without any improvement in the best fitness value (number of times best fitness value repeated), and (4) the percentage of chromosomes that invoke the penalty function (percent penalty) (table 4). Figure 13 presents response curves for best fitness value and computation time for all GA sensitivity runs for the 2001 time interval and 2015 time interval. GA performance is considered best when both the best fitness value and computation time are near minimums (Fisher, 2013). The probability of a GA solution being globally optimal increases with the number of consecutive iterations without any improvement in the best fitness value (Fisher, 2013). Therefore, metric 3 gives some indication of whether a global solution was found. GA performance decreases as the number of calls to the penalty function increases (Fisher, 2013). The sensitivity analysis ignores all interdependencies between the control parameters, an approach that is not strictly valid (Fisher, 2013). However, as documented in Fisher (2013), this approach provides an adequate means of evaluating the GA sensitivity and identifying reasonable parameter values.

Number of Sites Removed

Figures 13*A* and *B* show that increasing the number of sites removed increased the best fitness value (optimal networks are those that minimize the fitness value). This trend is expected on the basis of the results of previous studies and because the first two terms of the multiobjective function (mean standard error and RMSE) have a strong dependence on the number of sites removed (Fisher, 2013). The effect of the number of sites removed on computation time was small (less than 5 hours), peaking at 15 sites removed for the 2001 time interval (fig. 13*A*) and 10 sites removed for the 2015 time interval (fig. 13*B*). The nonlinear relationship between the number of sites removed and computation time was observed by Fisher (2013) and was attributed to the computational cost of kriging. Nonlinearity in the computation time response curves could also be related to percent penalty, which increased as the number of sites removed increased (table 4). Calls to the penalty function do not require kriging; therefore, increasing the percent penalty results in a decrease in computation time (Fisher, 2013).

Kriging Grid Resolution

Decreasing the resolution of the kriging grid (increasing the length of a square block side) generally increased the best fitness value (figs. 13*C* and *D*). Conversely, computation time exponentially decreased with decreasing grid resolution (figs. 13*C* and *D*). Computation time showed a dramatic decrease between lengths of 500 and 1,000 feet (figs. 13*C* and *D*). A length of 1,500 feet ("final value" column of table 3) was selected for the spatial assessment as the tradeoff between the best fitness value and computation time.

Other Control Parameters

The sensitivity analysis indicated near optimal solutions by using the following control parameters: a population size of 1,750 (figs. 13*E* and *F*; "final value" column of table 3), an elitism rate of 0.1 (10 percent) (figs. 13*G* and *H*; "final value" column of table 3), a crossover probability of 0.80 (80 percent) (figs. 13*I* and *J*; "final value" column of table 3), and a mutation probability of 0.175 (17.5 percent) (figs. 13*K* and *L*; "final value" column of table 3). The parameter values selected for the spatial assessment compare favorably with values used by Fisher (2013).

Table 4. Sensitivity of the genetic algorithm to incremental changes in control parameters for the 2001 time interval and 2015 time interval.

[A **bolded** control parameter value indicates base-case values (see table 3)]

Control parameter		2001 time interval				2015 time interval			
Name	Value	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty
Number of sites removed	5	2,117	3.6	20	0.012	1,531	3.5	43	0.026
	10	2,127	4.1	52	0.23	1,541	3.9	35	0.25
	15	2,137	4.3	2	1.0	1,548	3.8	22	2.0
	20	2,154	4.1	8	5.7	1,570	2.9	88	26
	25	2,175	1.9	6	50	1,599	1.4	54	56
	30	2,209	1.7	22	55	1,641	1.2	73	62
Kriging grid resolution (feet)	500	2,142	37	51	8.4	1,536	35	1	2.3
	1,000	2,142	10	2	11	1,544	10	12	2.0
	1,500	2,149	3.4	56	11	1,546	4.1	1	4.2
	2,000	2,159	2.4	33	15	1,554	2.6	54	2.9
	2,500	2,156	2.1	47	12	1,550	2.2	1	1.4
	3,000	2,150	2.0	9	10	1,553	2.0	73	1.5
	3,500	2,161	2.2	5	8.1	1,562	1.8	23	2.1
	4,000	2,194	3.0	38	10	1,580	2.4	15	1.3
	4,500	2,192	3.0	24	6.1	1,586	2.9	1	2.4
	5,000	2,183	2.7	1	12	1,570	2.7	14	2.4
Population size	250	2,149	0.32	9	1.9	1,562	0.30	2	1.1
	500	2,155	0.59	56	3.7	1,555	0.55	7	2.5
	750	2,159	0.83	2	8.6	1,556	0.79	2	2.6
	1,000	2,157	1.1	12	12	1,553	1.1	30	1.6
	1,250	2,152	1.3	5	10	1,554	1.3	1	4.5
	1,500	2,157	1.5	0	15	1,554	1.6	0	2.4
	1,750	2,149	1.8	4	14	1,552	1.8	0	2.0
	2,000	2,154	2.0	15	14	1,554	2.1	7	2.0
	2,250	2,156	2.4	13	10	1,553	2.4	30	1.2
	2,500	2,157	2.6	12	12	1,553	2.6	15	1.9
	2,750	2,159	2.8	2	14	1,554	2.9	44	2.4
	3,000	2,154	3.1	21	12	1,553	3.0	1	3.3
	3,250	2,148	3.4	33	11	1,552	3.4	17	2.6
	3,500	2,154	4.0	11	8.8	1,550	4.0	11	1.7
	3,750	2,149	4.0	57	15	1,550	4.2	17	2.4
	4,000	2,151	5.4	20	13	1,551	5.9	47	2.5
	4,250	2,155	8.1	34	12	1,550	8.1	4	2.6
	4,500	2,156	5.6	39	10	1,551	5.1	31	2.5
	4,750	2,157	5.3	53	13	1,552	5.5	77	2.3
	5,000	2,154	7.9	54	13	1,553	8.9	12	1.8

Table 4. Sensitivity of the genetic algorithm to incremental changes in control parameters for the 2001 time interval and 2015 time interval.—Continued

[A **bolded** control parameter value indicates base-case values (see table 3)]

Control parameter		2001 time interval				2015 time interval			
Name	Value	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty
Elitism rate	0.01	2,157	3.7	25	13	1,552	3.7	18	2.3
	0.025	2,159	3.8	17	10	1,551	3.9	3	1.9
	0.05	2,153	3.6	41	13	1,551	3.9	7	1.9
	0.075	2,154	3.8	10	9.4	1,553	3.8	33	2.6
	0.1	2,152	3.7	13	12	1,550	3.9	9	2.3
	0.125	2,146	4.0	26	7.0	1,552	3.9	0	2.0
	0.15	2,158	3.9	5	8.7	1,551	3.9	29	2.1
	0.175	2,150	4.0	1	7.0	1,554	3.9	61	1.8
	0.2	2,157	3.7	13	10	1,554	3.9	25	2.9
	0.225	2,156	3.5	20	15	1,552	4.0	3	2.1
	0.25	2,152	3.8	3	10	1,550	4.0	22	1.5
	0.275	2,152	3.8	12	10	1,552	3.9	17	2.5
	0.3	2,152	2.6	10	10	1,551	2.5	32	2.6
	0.325	2,155	2.1	9	14	1,551	2.5	13	2.1
	0.35	2,150	2.5	5	11	1,554	2.5	31	1.8
	0.375	2,154	2.2	5	12	1,550	2.5	20	3.1
	0.4	2,158	2.2	20	12	1,550	2.6	1	1.7
	0.425	2,151	2.2	18	11	1,550	2.5	26	2.4
	0.45	2,156	2.3	1	10	1,551	2.5	39	2.3
	0.475	2,154	2.2	3	14	1,553	3.8	54	1.9
	0.5	2,156	2.3	78	10	1,549	4.1	3	2.2
Crossover probability	0.5	2,153	1.6	45	3.7	1,551	1.5	23	1.0
	0.525	2,152	1.6	10	3.9	1,553	1.6	13	1.2
	0.55	2,154	1.7	14	4.3	1,551	1.6	0	1.7
	0.575	2,155	1.8	76	3.5	1,558	1.7	54	1.7
	0.6	2,154	1.7	75	7.7	1,555	1.8	74	0.91
	0.625	2,159	2.1	12	5.9	1,553	1.9	14	1.7
	0.65	2,159	3.2	31	4.1	1,552	3.0	6	1.8
	0.675	2,157	3.2	55	8.2	1,552	3.2	45	1.6
	0.7	2,156	3.3	3	9.6	1,551	3.3	44	1.5
	0.725	2,156	3.4	21	8.1	1,551	3.4	8	1.3
	0.75	2,159	2.3	26	6.1	1,553	2.4	15	2.3
	0.775	2,152	2.2	43	11	1,553	2.2	0	2.2
	0.8	2,154	2.4	17	6.5	1,551	2.3	1	2.0
	0.825	2,157	2.3	23	12	1,554	2.4	4	1.8
	0.85	2,153	2.4	13	12	1,552	2.5	22	1.9
	0.875	2,155	2.4	20	13	1,550	2.5	0	2.8
	0.9	2,156	2.5	4	13	1,554	2.5	3	3.7
	0.925	2,151	2.3	5	23	1,549	2.6	5	1.5
	0.95	2,157	2.6	5	15	1,553	2.7	34	2.8
	0.975	2,153	2.7	5	16	1,552	2.7	25	3.7

Table 4. Sensitivity of the genetic algorithm to incremental changes in control parameters for the 2001 time interval and 2015 time interval.—Continued

[A **bolded** control parameter value indicates base-case values (see table 3)]

Control parameter		2001 time interval				2015 time interval			
Name	Value	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty	Best fitness value (feet)	Computation time (hours)	Number of times best fitness value repeated	Percent penalty
Mutation probability	0.01	2,153	2.4	7	1.8	1,548	2.3	37	0.26
	0.025	2,155	2.2	3	6.8	1,551	2.2	9	1.1
	0.05	2,158	2.0	61	13	1,551	2.1	3	2.1
	0.075	2,154	1.9	23	17	1,552	2.1	27	3.2
	0.1	2,154	1.7	6	24	1,553	2.0	62	6.1
	0.125	2,149	1.6	72	30	1,553	2.0	15	5.6
	0.15	2,152	1.6	4	32	1,555	2.0	3	9.2
	0.175	2,153	1.5	63	34	1,550	2.0	10	9.0
	0.2	2,158	1.5	88	35	1,555	1.9	27	13
	0.225	2,160	1.4	17	41	1,550	1.9	67	14
	0.25	2,147	1.5	28	39	1,550	2.0	46	13
	0.275	2,154	1.4	28	41	1,552	2.1	58	12
	0.3	2,157	1.4	16	43	1,552	2.2	0	14
	0.325	2,154	1.4	23	44	1,551	2.2	7	13
	0.35	2,158	1.5	52	45	1,549	2.2	56	16
	0.375	2,147	1.5	1	45	1,550	2.3	14	14
	0.4	2,152	1.5	15	46	1,550	3.2	23	14
	0.425	2,152	1.6	6	46	1,550	3.9	5	16
	0.45	2,149	1.6	3	46	1,550	4.0	61	15
	0.475	2,154	1.5	60	48	1,550	3.1	18	15
	0.5	2,153	1.5	55	50	1,550	2.4	1	15

Temporal Assessment Methods

The temporal assessment followed an iterative thinning approach coupled with statistical techniques for identifying trend, similar to the approach presented in Cameron (2004). This methodology began by assembling daily mean depth-to-groundwater measurements at each monitoring site. The daily mean data at each monitoring site were then aggregated into a monthly median depth to groundwater. The temporal trend of the monthly median depth-to-groundwater measurements was computed for each monitoring site. Iterative thinning, at random, of the monthly median data at each monitoring site was performed for a range of reduced monitoring frequencies. The trend of the reduced frequency data was computed and compared to the trend of the monthly data.

Background

The trends in groundwater levels through time were estimated using local regression (Cameron, 2004). The trend estimation technique employed in this report can be used for complicated trends, including trends with seasonal components (Cameron, 2004). Seasonal fluctuations in groundwater levels are evident in many hydrographs from the observation network (Ritchie and Galanter, 2019). Local regression was also used to estimate a confidence interval around the trend (Cameron, 2004).

Input Data Description

The temporal assessment was performed for the period of record at continuously monitored sites within the observation network (“hourly” in “measurement frequency” column of table 1). Daily mean depth-to-groundwater measurements computed by the USGS from the hourly data were aggregated into monthly median values for use in the temporal assessment. Initial tests of the temporal assessment method using daily data revealed a narrow confidence interval around the local regression because of the large number of daily measurements (thousands of daily measurements over the period of record at each site). Thus, monthly median depth-to-groundwater measurements were used because the confidence interval estimated around the trend on the daily data was deemed too restrictive to produce meaningful results from this analysis.

Implementation

Computer software utilized to perform the temporal assessment was written in the R programming language (version 3.5.1; R Core Team, 2019). The “loess” function within the R package “stats” (version 3.5.1; base R) was used to perform local regression and estimate the upper and lower confidence intervals of the local regression.

The temporal assessment procedure performed for each monitoring site was as follows:

1. Compute the local regression fit and two-sided (upper and lower) confidence interval for the regression fit on the monthly median data,
2. Randomly remove a fraction of the monthly median depth-to-groundwater data and recompute the local regression fit by using the reduced dataset,
3. Calculate the fraction of estimates from the local regression fit on the reduced dataset that are within (less than or equal to the upper confidence interval and greater than or equal to the lower confidence interval) the confidence interval for the regression fit on the monthly data,
4. Calculate the mean reduced monitoring frequency as the mean of the number of days between subsequent depth-to-groundwater measurements in the reduced dataset (assuming each monthly value falls on the first day of the month),
5. Repeat items 2 through 4 for 10,000 iterations by using the same fraction for data removal,
6. Calculate the median fraction from item 3 over all iterations,
7. Calculate the mean of the mean reduced monitoring frequency from item 4 over all iterations, and
8. Repeat items 2 through 7 by using a different fraction for data removal that increases the amount of data that is removed.

This procedure was performed for a range of data removal fractions that generally corresponded to a range of reduced monitoring frequencies: bimonthly (every 2 months), quarterly (every 3 months), semiannually (every 6 months), annually (every 12 months), every 2 years, and every 5 years.

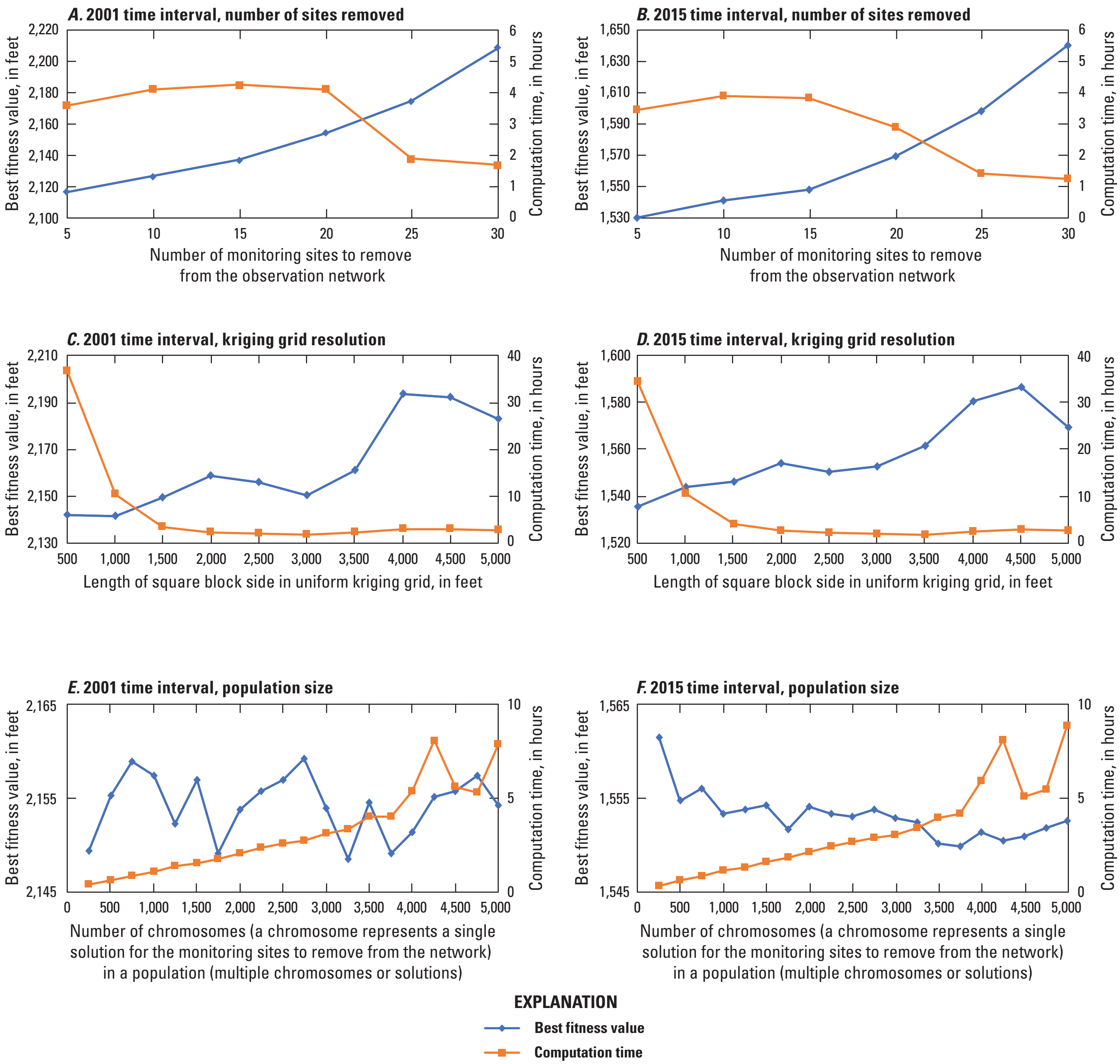


Figure 13. Best fitness value and computation time response curves of genetic algorithm sensitivity analysis for the control parameters: number of sites removed for the *A*, 2001 time interval and *B*, 2015 time interval; kriging grid resolution for the *C*, 2001 time interval and *D*, 2015 time interval; population size for the *E*, 2001 time interval and *F*, 2015 time interval; elitism rate for the *G*, 2001 time interval and *H*, 2015 time interval; crossover probability for the *I*, 2001 time interval and *J*, 2015 time interval; and mutation probability for the *K*, 2001 time interval and *L*, 2015 time interval.

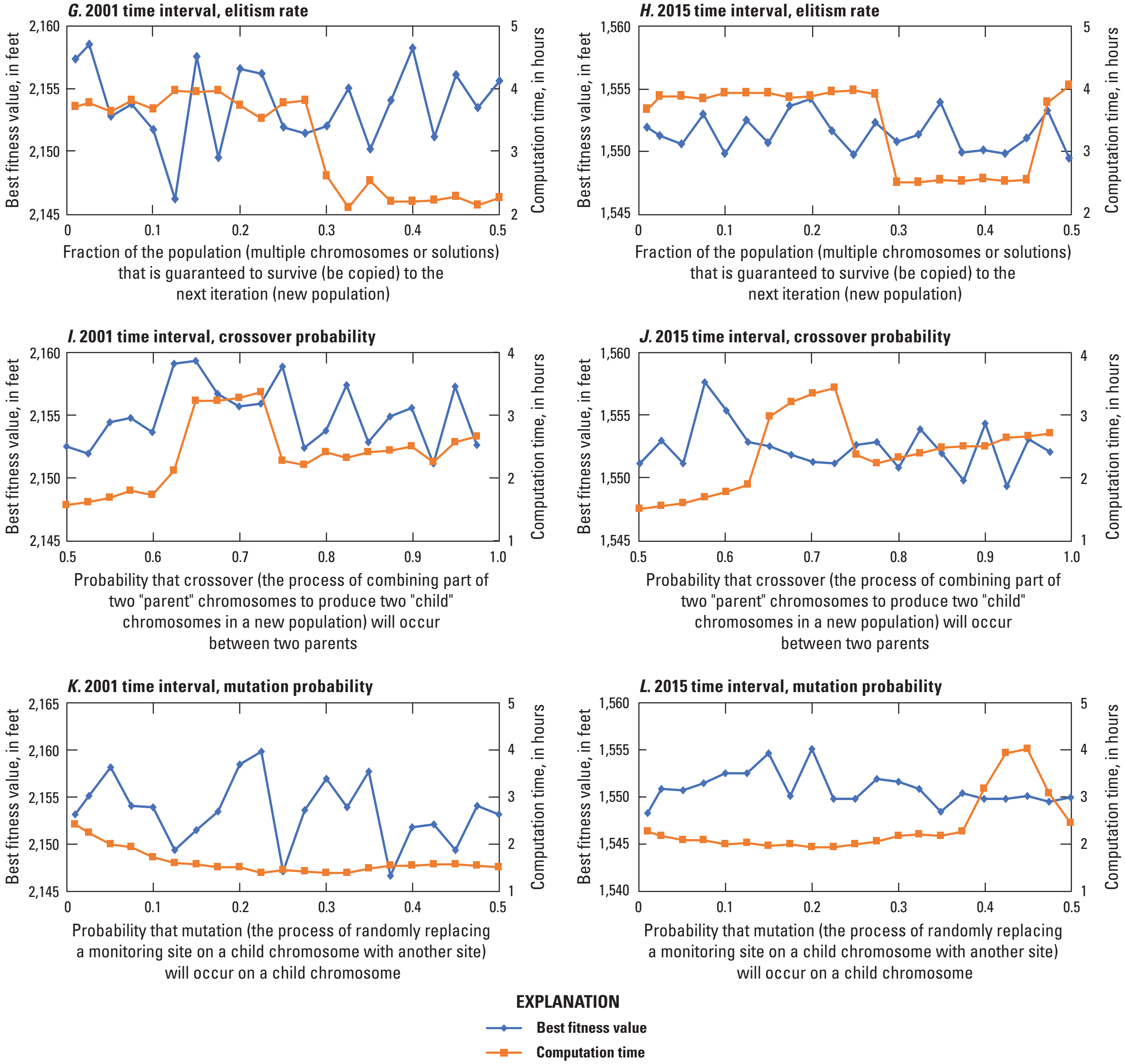


Figure 13. Best fitness value and computation time response curves of genetic algorithm sensitivity analysis for the control parameters: number of sites removed for the *A*, 2001 time interval and *B*, 2015 time interval; kriging grid resolution for the *C*, 2001 time interval and *D*, 2015 time interval; population size for the *E*, 2001 time interval and *F*, 2015 time interval; elitism rate for the *G*, 2001 time interval and *H*, 2015 time interval; crossover probability for the *I*, 2001 time interval and *J*, 2015 time interval; and mutation probability for the *K*, 2001 time interval and *L*, 2015 time interval.—Continued

Parameter Selection

The confidence interval of the local regression of daily mean depth-to-groundwater data was calculated at the 95-percent confidence level. The confidence level represents the probability that the confidence interval contains the true trend estimated using local regression. The degree of smoothing in the local regression is controlled by the span parameter, which was held constant at 0.5 for the temporal assessment. The degree of the polynomials used in the local regression is controlled by the degree parameter, which was held constant at 1 for the temporal assessment. The method of trend fitting is controlled by the family parameter, which was held constant at “gaussian” (indicating fitting by least squares) for the temporal assessment.

Principal Component Analysis Methods

A PCA approach, like that of Winter and others (2000), was used to identify groups of monitoring sites that had statistically similar hydrographs (groundwater elevation as a function of time). This study addressed the same objective by combining PCA and k-means clustering. This methodology began by assembling daily mean depth-to-groundwater measurements at each of the 15 continuously monitored sites classified as being in the aquifer production zone (see discussion in “Monitoring Sites Used for Optimization Assessment” section; “yes” in “used for PCA” column of table 1). These measurements were aggregated into monthly median data to best fulfill PCA data requirements and allow analysis of a long time span of data. PCA and k-means clustering were then utilized to identify groups of monitoring sites with similar characteristics. This information can inform water manager decisions regarding the spatial and temporal assessment results by distinguishing spatially distributed regions that have similar hydrographs. For example, water managers may decide to maintain a minimum number of continuous and discrete wells within each of these regions.

Background

PCA seeks a low-dimensional representation of a dataset by reprojecting its variables as weighted linear combinations of all variables to ease pattern recognition (Hastie and others, 2009; James and others, 2013). These linear combinations are called PCs and are computed such that they capture a large proportion of the information in the dataset by using fewer dimensions. PCs are ordered by the amount of dataset variability for which they account, meaning that the first PC explains the most dataset variance, followed by the second PC, then the third PC, and so on. There are as many PCs as there are variables, and they have the following mathematical form:

$$PC_i = \lambda_1 V_1 + \lambda_2 V_2 + \ldots + \lambda_{nvar} V_{nvar} \quad (2)$$

where

- PC_i is principal component score (i = 1 to $nvar$);
- λ is weight (also known as loading) derived from PCA, where subscripts correspond to variable number;
- V is original variable value (groundwater elevation measured at a site as a function of time in this study), where subscripts correspond to variable number; and
- $nvar$ is total number of variables considered.

Loadings in PCA are mathematically determined by using eigen decomposition and are related to the correlation between variables (Hastie and others, 2009; James and others, 2013). Winter and others (2000) used well hydrographs as variables and then plotted the loadings from the first and second PCs against each other to identify sites with similar hydrographs, because hydrographs of like shape obtain resembling loadings in PCA. They used visual inspection of this plot to identify groups of similar monitoring sites. This study instead applied Euclidean distance-based k-means clustering to the PCA results to quantitatively group the hydrographs. This approach was advantageous because more than two PCs could easily be considered for group identification. K-means clustering identifies groups of data by randomly assigning each data point to a cluster, calculating the resulting cluster centroids, and reassigning data points to their nearest centroid (Everitt and others, 2011; James and others, 2013). This procedure is done iteratively and is repeated at least hundreds of times to determine the most prevalent cluster assignments. The hydrographs assigned to each cluster were then visually inspected to evaluate the prominent patterns that led to their grouping. These groupings may result for numerous reasons including aquifer characteristics, proximity to recharge sources, and nearby water use.

Input Data Description

Input data for the PCA were a matrix of monthly median groundwater-elevation data spanning from February 2008 through July 2014, where columns corresponded to individual monitoring sites and each row coincided with a month within the analyzed time interval (table 5). The available period of record for the continuously monitored sites used for the PCA was variable with an overall range from November 2, 1988, through September 30, 2017 (table 1). PCA requires a complete matrix of data (meaning no missing or “not available” [NA] values). If one monitoring site lacks data for a given month, then that month’s data must be removed from the input data matrix for all monitoring sites. This removal leads to temporal gaps in the input data matrix. Previous research suggests that these gaps should be restricted to compose no more than 20 percent of the overall dataset to avoid appreciably skewing the PCA results (Smirnov and Egbert, 2012; Severson and others, 2017). Daily mean depth-to-groundwater measurements were aggregated into monthly median values to reduce data gap percentage and thereby expand the overall time span that could be analyzed. Figure 14 illustrates the percentage of monitoring sites lacking groundwater-level data (“offline” sites) when using daily mean, weekly median, and monthly median values; time intervals with the lowest overall percentage are more ideal for PCA. Several time periods were available that met the 20-percent or less data gap guideline; however, the longest interval spanned from February 2008 through July 2014 (gray shaded interval in fig. 14) and was selected for PCA to include as much data as possible. This interval begins about 10 months before the SJCDWP was implemented and includes 6.4 years of data with data gap percentages of 29, 22, and 9 percent for the daily mean, weekly median, and monthly median values, respectively. The monthly aggregated data were used for PCA because they were well below the 20-percent data gap guideline. A total of 1,065 monthly median groundwater elevations are represented in the analysis, with the longest data gap spanning 2 months.

The influence of removing NA values from the input dataset (9-percent data gaps) on the PCA results was interrogated by replacing all NA values with an interpolated mean groundwater elevation rather than omitting the corresponding months. Additionally, the implications of aggregating daily data into monthly data were evaluated by performing PCA on daily and weekly data over the same time interval that was used in the monthly analysis. Hydrograph clusters identified in the PCA analysis remained unaltered for all these tests, thereby indicating that the monthly aggregation and data gap percentage have a negligible influence on the PCA results. Furthermore, PCA clusters were not affected by using depth-to-groundwater values instead of groundwater elevations, nor were they affected by using mean values in place of median values.

Implementation and Parameter Selection

R (version 3.5.3; R Core Team, 2019) was used in the PCA and clustering analysis. Boxplots were first constructed for the groundwater elevations at each monitoring site to identify potential outlier data that could have biased the PCA results. These plots were constructed using the “far out” value criteria proposed by Tukey (1977), which considers potential outliers to be any point that is outside three times the interquartile range of the data. This approach is often considered to be a conservative test for outliers and does not require assumptions about the distribution of the data (Iglewicz and Hoaglin, 1993). No potential outliers were identified when using this criterion; therefore, outlier adjustment techniques that are sometimes applied in PCA were not needed.

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Isleta	Montessa Park	Westgate Heights Park	San Jose 9	Armijo
USGS site number	**345650106415902**	**350056106370102**	**350244106450202**	**350256106390801**	**350307106410602**
Beginning day of month	**Monthly median groundwater elevation (feet above NAVD 88)**				
2/1/2008	4,889.87	4,877.77	4,886.19	4,912.58	4,889.19
3/1/2008	4,889.39	4,877.17	4,886.13	4,912.21	4,888.7
4/1/2008	4,888.61	4,875.805	4,886.065	4,911.955	4,885.655
5/1/2008	4,888.37	4,874.44	4,886.27	4,911.5	4,889.24
6/1/2008	4,887.495	4,871.525	4,886.3	4,910.34	4,889.84
7/1/2008	4,887.2	4,871.97	4,886.3	4,910.98	4,889.27
8/1/2008	4,887.445	4,871.89	4,886.44	4,910.61	4,894.68
10/1/2008	4,889.205	4,873.56	4,886.73	4,911.47	4,897.07
11/1/2008	4,889.17	4,874.45	4,887.215	4,912.085	4,897.855
12/1/2008	4,889.89	4,875.82	4,888.795	4,912.41	4,899.4
1/1/2009	4,890.25	4,876.35	4,889.045	4,912.74	4,900.24
2/1/2009	4,890.185	4,876.435	4,889.805	4,912.37	4,900.945
3/1/2009	4,889.76	4,875.98	4,889.74	4,912.5	4,901.22
4/1/2009	4,889.15	4,875.505	4,889.45	4,912.38	4,901.315
5/1/2009	4,888.6	4,874.21	4,889.2	4,912.12	4,901.34
6/1/2009	4,889.12	4,873.83	4,889.06	4,911.905	4,901.36
7/1/2009	4,888.55	4,873.26	4,888.75	4,911.75	4,901.36
8/1/2009	4,888.11	4,872.79	4,888.76	4,911.53	4,901.53
9/1/2009	4,888.6	4,872.9	4,888.415	4,911.76	4,901.505
10/1/2009	4,889.29	4,874.56	4,888.57	4,912.48	4,901.66
11/1/2009	4,890.13	4,876.57	4,888.64	4,913.27	4,902
12/1/2009	4,890.46	4,878.205	4,888.8	4,913.71	4,902.06
1/1/2010	4,890.61	4,879.305	4,889.08	4,914.01	4,902.24
2/1/2010	4,890.36	4,879.415	4,890.48	4,914.185	4,903.45
3/1/2010	4,889.58	4,879.26	4,891.5	4,914.22	4,904.63
4/1/2010	4,889.175	4,878.245	4,892.06	4,913.895	4,905.19
5/1/2010	4,889.13	4,877.82	4,892.69	4,913.99	4,905.67
6/1/2010	4,888.51	4,876.615	4,893.25	4,913.675	4,906
7/1/2010	4,887.77	4,875.9	4,893.56	4,913.23	4,906.25
8/1/2010	4,888.55	4,876.19	4,893.93	4,913.78	4,906.625
9/1/2010	4,888.2	4,875.855	4,894.285	4,913.665	4,906.73
10/1/2010	4,889.22	4,875.98	4,894.59	4,914.04	4,907.24
11/1/2010	4,890.055	4,877.615	4,894.91	4,914.76	4,907.565
1/1/2011	4,890.35	4,879.47	4,895.39	4,915.035	4,908.24
3/1/2011	4,889.07	4,879.63	4,895.83	4,914.9	4,908.52
5/1/2011	4,888.08	4,878.49	4,896.14	4,914.56	4,908.9
6/1/2011	4,887.41	4,877.495	4,896.235	4,914.01	4,909.115
7/1/2011	4,887.05	4,876.36	4,896.3	4,913.66	4,909.37

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Isleta	Montessa Park	Westgate Heights Park	San Jose 9	Armijo
USGS site number	**345650106415902**	**350056106370102**	**350244106450202**	**350256106390801**	**350307106410602**
Beginning day of month	**Monthly median groundwater elevation (feet above NAVD 88)**				
10/1/2011	4,888.91	4,875.73	4,896.59	4,914.06	4,909.6
11/1/2011	4,889.325	4,877.54	4,896.73	4,914.825	4,909.75
12/1/2011	4,890.26	4,879.91	4,896.87	4,915.39	4,909.99
1/1/2012	4,890.66	4,881.1	4,897.03	4,915.76	4,910.46
2/1/2012	4,890.26	4,881.69	4,897.26	4,915.82	4,910.71
3/1/2012	4,890.08	4,881.43	4,897.4	4,915.83	4,910.78
4/1/2012	4,889.215	4,880.42	4,897.52	4,915.495	4,910.89
5/1/2012	4,888.43	4,879.13	4,897.59	4,915.15	4,911.11
6/1/2012	4,887.155	4,878.14	4,897.61	4,914.5	4,911.135
7/1/2012	4,887.06	4,877.89	4,897.51	4,914.5	4,911.13
8/1/2012	4,886.9	4,878.58	4,897.43	4,914.69	4,911.09
9/1/2012	4,887.575	4,877.035	4,897.47	4,914.755	4,911
10/1/2012	4,888.22	4,877.15	4,897.54	4,914.86	4,910.94
11/1/2012	4,889.115	4,878.325	4,897.53	4,915.93	4,910.76
12/1/2012	4,890.18	4,880.39	4,897.69	4,916.03	4,910.77
1/1/2013	4,890.69	4,881	4,897.79	4,916.15	4,910.71
2/1/2013	4,890.595	4,881.385	4,897.955	4,916.175	4,911.145
3/1/2013	4,889.44	4,881.82	4,898.12	4,916.16	4,911.44
4/1/2013	4,888.57	4,881.455	4,898.435	4,915.79	4,911.77
5/1/2013	4,887.73	4,880.68	4,898.52	4,915.56	4,911.82
6/1/2013	4,887.295	4,879.96	4,898.41	4,915.38	4,907.115
7/1/2013	4,887.22	4,878.54	4,898.24	4,915.42	4,908.58
8/1/2013	4,887.79	4,877.48	4,898.14	4,915.42	4,910.39
9/1/2013	4,888.01	4,877.775	4,898.355	4,915.94	4,911.015
11/1/2013	4,889.705	4,881.105	4,898.45	4,916.585	4,910.99
12/1/2013	4,890.37	4,881.48	4,898.53	4,916.82	4,910.93
1/1/2014	4,890.36	4,882.18	4,898.52	4,916.88	4,911.03
2/1/2014	4,890.055	4,882.74	4,898.65	4,916.87	4,911.23
3/1/2014	4,889.71	4,882.78	4,898.64	4,916.9	4,911.35
4/1/2014	4,888.7	4,882.255	4,898.66	4,916.555	4,911.53
5/1/2014	4,888.37	4,881.76	4,898.75	4,916.47	4,911.56
6/1/2014	4,887.89	4,881.03	4,898.74	4,916.18	4,911.64
7/1/2014	4,887.71	4,880.34	4,898.64	4,916.02	4,911.18

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Del Sol Divider	Jerry Cline Park	West Bluff Nest 1	Garfield Park	Montaño Nest 6
USGS site number	350534106354702	350545106335902	350638106413702	350706106390302	350836106395401
Beginning day of month	Monthly median groundwater elevation (feet above NAVD 88)				
2/1/2008	4,866.37	4,848.03	4,942.41	4,922.38	4,937.22
3/1/2008	4,864.46	4,848.12	4,936.79	4,921.79	4,938.26
4/1/2008	4,861.69	4,845.82	4,930.35	4,921.35	4,937.2
5/1/2008	4,857.81	4,842.93	4,935.58	4,920.17	4,933.96
6/1/2008	4,852.22	4,839.615	4,936.435	4,917.985	4,931.37
7/1/2008	4,852.45	4,839.93	4,942.14	4,918.47	4,931.95
8/1/2008	4,851.16	4,838.85	4,937.27	4,918.34	4,932.86
10/1/2008	4,854.98	4,840.78	4,948.73	4,919.35	4,932.98
11/1/2008	4,858.275	4,842.11	4,949.82	4,920.735	4,936.31
12/1/2008	4,861.69	4,844.47	4,950.56	4,921.3	4,936.56
1/1/2009	4,863.5	4,845.665	4,951.21	4,922.53	4,939.74
2/1/2009	4,863.59	4,846.28	4,951.175	4,922.39	4,939.835
3/1/2009	4,863.05	4,845.95	4,950.69	4,922.2	4,939.51
4/1/2009	4,862.115	4,845.38	4,950.975	4,921.935	4,937.595
5/1/2009	4,858.97	4,843.64	4,947.99	4,921.08	4,937.96
6/1/2009	4,857.46	4,842.4	4,948.345	4,920.805	4,937.32
7/1/2009	4,856.55	4,839.93	4,940.75	4,919.25	4,934.54
8/1/2009	4,853.65	4,837.96	4,943.67	4,918.74	4,932.99
9/1/2009	4,856.02	4,839.37	4,949.27	4,919.17	4,933.99
10/1/2009	4,861.46	4,843.91	4,951.22	4,921.37	4,938.33
11/1/2009	4,865.77	4,846.62	4,952.2	4,922.875	4,939.84
12/1/2009	4,869.25	4,849.83	4,952.48	4,923.59	4,940.7
1/1/2010	4,870.71	4,851.28	4,950.61	4,924.18	4,941.49
2/1/2010	4,871.935	4,852.17	4,951.41	4,924.65	4,941.305
3/1/2010	4,871.69	4,851.85	4,951.24	4,924.8	4,940.62
4/1/2010	4,868.81	4,851.86	4,949.675	4,923.655	4,937.925
5/1/2010	4,867.25	4,850.32	4,941.32	4,923.62	4,938.17
6/1/2010	4,862.97	4,847.655	4,939.96	4,923.185	4,937.13
7/1/2010	4,862.07	4,847.01	4,938.54	4,922.37	4,936.48
8/1/2010	4,862.42	4,847.27	4,949.67	4,922.96	4,937.77
9/1/2010	4,861.785	4,846.28	4,951.29	4,922.91	4,938.115
10/1/2010	4,863.54	4,847.33	4,952.46	4,923.54	4,939.4
11/1/2010	4,866.615	4,850.515	4,952.745	4,924.755	4,941.22
1/1/2011	4,872.855	4,854.17	4,951.96	4,925.81	4,941.98
3/1/2011	4,873.31	4,855.43	4,949.21	4,925.86	4,940.88
5/1/2011	4,868.05	4,854.26	4,950.48	4,925.77	4,942.37
6/1/2011	4,866.675	4,851.69	4,949.35	4,925.455	4,942.025
7/1/2011	4,864.85	4,850.555	4,940.83	4,924.83	4,941.52

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Del Sol Divider	Jerry Cline Park	West Bluff Nest 1	Garfield Park	Montaño Nest 6
USGS site number	**350534106354702**	**350545106335902**	**350638106413702**	**350706106390302**	**350836106395401**
Beginning day of month	**Monthly median groundwater elevation (feet above NAVD 88)**				
10/1/2011	4,863.35	4,846.95	4,938.95	4,924.22	4,937.81
11/1/2011	4,871.27	4,854.2	4,935.97	4,925.63	4,942
12/1/2011	4,876.07	4,857.08	4,945.01	4,927	4,943.22
1/1/2012	4,878.27	4,859.22	4,951.09	4,927.65	4,943.78
2/1/2012	4,878.21	4,860.43	4,944.31	4,927.62	4,943.27
3/1/2012	4,877.06	4,860.32	4,942.09	4,927.39	4,942.59
4/1/2012	4,874.815	4,859.4	4,937.575	4,926.955	4,942.58
5/1/2012	4,871.38	4,857.01	4,930.91	4,926.67	4,940.61
6/1/2012	4,870.075	4,856.655	4,933.675	4,926.375	4,941.69
7/1/2012	4,870.92	4,858.06	4,937.29	4,926.63	4,941.85
8/1/2012	4,872	4,858.31	4,937.03	4,926.85	4,941.25
9/1/2012	4,863.36	4,850.28	4,925.13	4,925.07	4,938.59
10/1/2012	4,865.13	4,850.3	4,932.79	4,925.28	4,937.42
11/1/2012	4,871.84	4,856.13	4,936.595	4,926.2	4,940.46
12/1/2012	4,875.58	4,858.89	4,940.99	4,927.91	4,943.7
1/1/2013	4,876.33	4,859.8	4,942.38	4,928.37	4,944.04
2/1/2013	4,878.51	4,861.155	4,943.755	4,928.575	4,944.175
3/1/2013	4,878.97	4,861.58	4,950.01	4,928.7	4,944.2
4/1/2013	4,877.24	4,861.65	4,940.39	4,928.26	4,943.1
5/1/2013	4,877.5	4,863.04	4,948.39	4,928.32	4,943.78
6/1/2013	4,874.66	4,861.2	4,942.66	4,926.395	4,935.005
7/1/2013	4,863.93	4,852.62	4,930.85	4,925.03	4,933.82
8/1/2013	4,862.23	4,850.69	4,928.62	4,924.89	4,934.12
9/1/2013	4,862.095	4,854.335	4,932.54	4,925.34	4,937.555
11/1/2013	4,874.735	4,861.455	4,946.91	4,928.19	4,940.935
12/1/2013	4,875.03	4,859.67	4,934.53	4,928.04	4,941.99
1/1/2014	4,879.92	4,863.37	4,937.79	4,928.75	4,942.19
2/1/2014	4,881.475	4,864.65	4,946.745	4,928.955	4,943.575
3/1/2014	4,881.56	4,864.9	4,941.75	4,929.33	4,944.07
4/1/2014	4,881.94	4,865.96	4,937.46	4,929.375	4,944.24
5/1/2014	4,880.73	4,865.53	4,935.66	4,929.3	4,944.21
6/1/2014	4,878.79	4,864.43	4,933.25	4,928.725	4,943.5
7/1/2014	4,876.04	4,863.23	4,932.57	4,927.66	4,940.94

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Sister Cities	Sierra Vista	Nor Este	Hunters Ridge Nest 1	Sandia Pueblo
USGS site number	350908106344402	350910106414802	351114106330602	351201106400502	351357106323002
Beginning day of month	Monthly median groundwater elevation (feet above NAVD 88)				
2/1/2008	4,888.73	4,949.39	4,910.63	4,947.54	4,952.03
3/1/2008	4,888.15	4,949.24	4,910.97	4,946.92	4,952.13
4/1/2008	4,887.5	4,948.86	4,910.9	4,945.48	4,952.2
5/1/2008	4,886.33	4,949.16	4,910.74	4,943.92	4,952.21
6/1/2008	4,884.465	4,949.325	4,909.72	4,942.755	4,952.185
7/1/2008	4,884.42	4,949.34	4,908.85	4,942.25	4,952.14
8/1/2008	4,883.85	4,949.28	4,910.26	4,942.07	4,952.14
10/1/2008	4,884.09	4,949.71	4,910.8	4,945.14	4,952.16
11/1/2008	4,884.89	4,950.05	4,911.34	4,945.78	4,952.225
12/1/2008	4,885.66	4,950.29	4,911.84	4,946.47	4,952.3
1/1/2009	4,886.69	4,950.58	4,912.19	4,946.61	4,952.34
2/1/2009	4,887.29	4,950.77	4,912.18	4,946.49	4,952.44
3/1/2009	4,887.75	4,950.95	4,912.7	4,945.28	4,952.55
4/1/2009	4,886.75	4,951.025	4,912.93	4,943.78	4,952.6
5/1/2009	4,886.31	4,950.98	4,912.72	4,943.08	4,952.63
6/1/2009	4,886.27	4,951.04	4,912.73	4,941.435	4,952.67
7/1/2009	4,884.97	4,950.81	4,912.46	4,941.1	4,952.69
8/1/2009	4,883.98	4,950.69	4,912.24	4,941.07	4,952.68
9/1/2009	4,884.675	4,950.79	4,912.31	4,941.32	4,952.68
10/1/2009	4,886.53	4,951.05	4,912.88	4,943.56	4,952.73
11/1/2009	4,887.705	4,951.305	4,913.53	4,944.69	4,952.76
12/1/2009	4,888.93	4,951.52	4,913.94	4,945.98	4,952.84
1/1/2010	4,889.65	4,951.74	4,914.51	4,945.76	4,952.93
2/1/2010	4,890.32	4,951.88	4,914.885	4,946.43	4,953.045
3/1/2010	4,890.7	4,951.94	4,915.19	4,947.09	4,953.2
4/1/2010	4,890.515	4,951.78	4,915.23	4,944.83	4,953.3
5/1/2010	4,890.77	4,951.47	4,915.17	4,943.51	4,953.33
6/1/2010	4,889.535	4,951.28	4,914.745	4,942.3	4,953.345
7/1/2010	4,889.59	4,950.97	4,914.67	4,942.2	4,953.37
8/1/2010	4,889.18	4,951.12	4,914.88	4,941.49	4,953.4
9/1/2010	4,889.225	4,951.345	4,914.855	4,941.195	4,953.44
10/1/2010	4,889.48	4,951.41	4,915.05	4,941.03	4,953.55
11/1/2010	4,891.325	4,951.64	4,915.535	4,944.375	4,953.57
1/1/2011	4,892.76	4,952.04	4,916.52	4,944.92	4,953.67
3/1/2011	4,893.32	4,951.79	4,917.04	4,944.27	4,953.9
5/1/2011	4,893.44	4,952.22	4,917.14	4,943.22	4,954.02
6/1/2011	4,892.035	4,951.91	4,916.89	4,942.06	4,954.03
7/1/2011	4,890.73	4,951.61	4,916.58	4,942.6	4,954.04

Table 5. Input data matrix for the principal component analysis of monthly median groundwater-elevation data spanning from February 2008 through July 2014.—Continued

[USGS, U.S. Geological Survey; NAVD 88, North American Vertical Datum of 1988]

Site name	Sister Cities	Sierra Vista	Nor Este	Hunters Ridge Nest 1	Sandia Pueblo
USGS site number	**350908106344402**	**350910106414802**	**351114106330602**	**351201106400502**	**351357106323002**
Beginning day of month	**Monthly median groundwater elevation (feet above NAVD 88)**				
10/1/2011	4,889.6	4,950.86	4,916.74	4,944.15	4,954.08
11/1/2011	4,893.05	4,950.86	4,917.435	4,945.525	4,954.22
12/1/2011	4,894.51	4,951.5	4,918.21	4,947.18	4,954.48
1/1/2012	4,895.62	4,951.62	4,918.8	4,946.54	4,954.67
2/1/2012	4,896.03	4,951.96	4,919.02	4,946.13	4,954.69
3/1/2012	4,895.76	4,951.75	4,919.12	4,945.59	4,954.7
4/1/2012	4,895.545	4,951.31	4,918.965	4,945.29	4,954.785
5/1/2012	4,894.84	4,950.83	4,918.74	4,945.36	4,954.8
6/1/2012	4,895.19	4,950.92	4,918.56	4,944.06	4,954.85
7/1/2012	4,895.58	4,950.76	4,918.47	4,943.34	4,954.88
8/1/2012	4,895.83	4,950.57	4,918.64	4,942.75	4,954.91
9/1/2012	4,891.465	4,949.805	4,918.45	4,942.51	4,954.9
10/1/2012	4,891.11	4,949.79	4,918.77	4,942.65	4,954.92
11/1/2012	4,894.37	4,950.125	4,919.085	4,945.685	4,954.96
12/1/2012	4,895.74	4,950.56	4,919.67	4,946.49	4,955.16
1/1/2013	4,896.03	4,950.84	4,920.11	4,946.81	4,955.19
2/1/2013	4,897.455	4,950.99	4,920.44	4,947.215	4,955.39
3/1/2013	4,897.85	4,951.21	4,920.73	4,946.6	4,955.43
4/1/2013	4,897.745	4,951.07	4,920.69	4,944.355	4,955.53
5/1/2013	4,898.28	4,951.35	4,920.48	4,943.3	4,955.59
6/1/2013	4,897.235	4,951.205	4,920.82	4,942.455	4,955.58
7/1/2013	4,893.59	4,950.3	4,919.9	4,942.63	4,955.63
8/1/2013	4,892.04	4,950.11	4,919.91	4,942.85	4,955.73
9/1/2013	4,894.505	4,950.02	4,920.09	4,943.15	4,955.865
11/1/2013	4,897.645	4,950.95	4,921.23	4,945.8	4,955.97
12/1/2013	4,896.72	4,950.57	4,921.525	4,946.47	4,956.12
1/1/2014	4,898.72	4,950.64	4,921.945	4,946.63	4,956.2
2/1/2014	4,899.62	4,951.065	4,922.325	4,947.175	4,956.305
3/1/2014	4,900.21	4,951.2	4,922.52	4,947.06	4,956.34
4/1/2014	4,900.51	4,951.065	4,922.605	4,946.79	4,956.37
5/1/2014	4,900.41	4,950.9	4,922.54	4,946.38	4,956.37
6/1/2014	4,900.235	4,950.78	4,922.405	4,944.88	4,956.41
7/1/2014	4,899.96	4,950.52	4,922.29	4,944.45	4,956.39

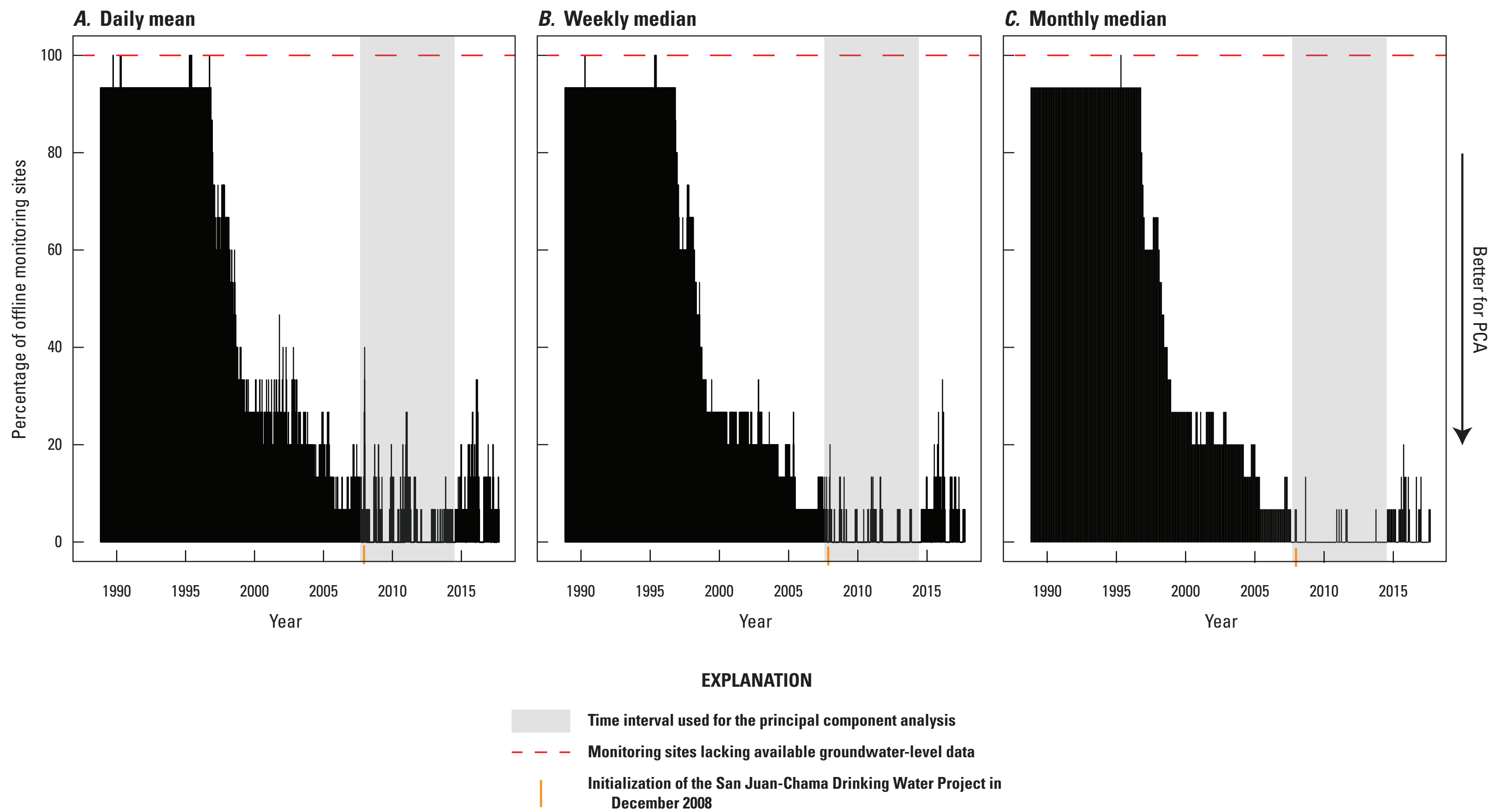


Figure 14. Percentage of monitoring sites lacking available groundwater-level data for the principal component analysis (PCA) when using *A*, daily mean; *B*, weekly median; and *C*, monthly median values. Time intervals with the lowest overall percentage are more ideal for PCA.

PCA was performed using the "prcomp" function within the "stats" (version 3.5.3) package that is provided in base R. Within this function, the variables were normalized to have a zero mean and scaled to have a unit variance, which is standard practice in PCA (James and others, 2013). The scaling permitted focusing the analysis more on hydrograph shape than on groundwater-elevation magnitude. A scree plot showing the percent of variance explained (PVE) by each PC was constructed to determine the number of PCs to include in the clustering analysis (fig. 15). A pronounced elbow in this plot at the second PC indicated that the first two PCs accounted for a large majority of variation in the dataset (PC1 PVE = 67.9, PC2 PVE = 16.6). A second elbow at the fourth PC, however, indicated that additional pertinent information was contained in the third and fourth PCs (PC3 PVE = 9.4, PC4 PVE = 2.3). All together these four PCs accounted for 96.2 percent of the dataset's variance, which was an improvement on the 84.5 percent cumulatively explained by the first two PCs. Therefore, four PCs were included in the clustering analysis, which effectively reduced the number of variables/dimensions to consider from 15 to 4.

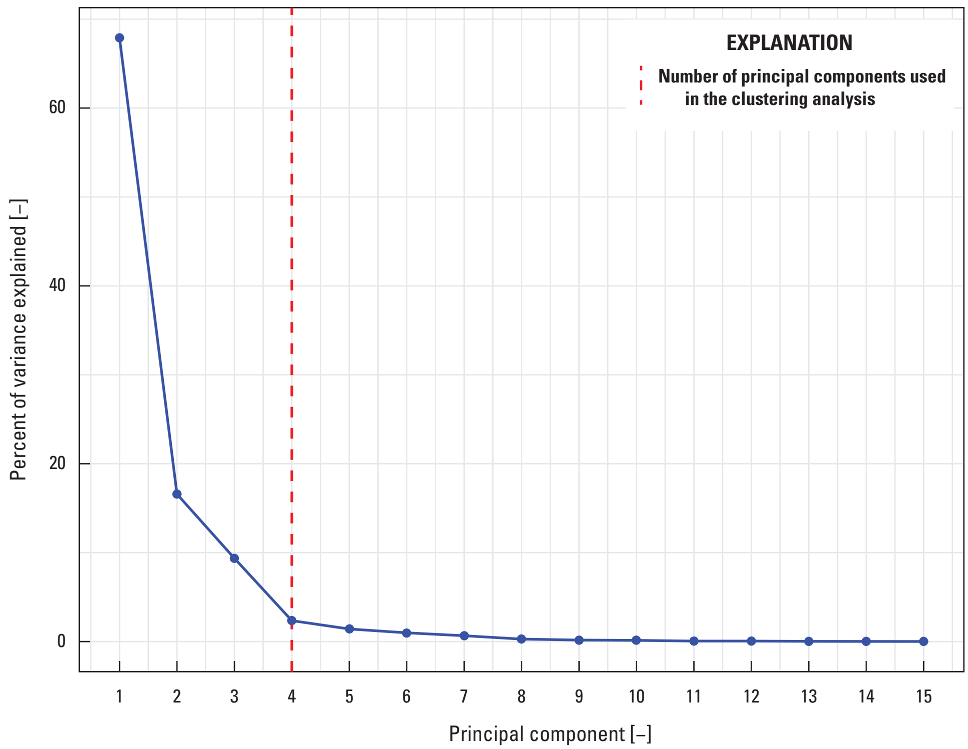


Figure 15. The amount of dataset variance for which each principal component (PC) accounted.

K-means clustering analysis was performed using the Hartigan and Wong (1979) algorithm as implemented by the “kmeans” function within the “stats” (version 3.5.3) package included in base R. This algorithm was applied to a Euclidean-distance matrix computed from the loadings of PCs 1 through 4. The distance matrix was calculated using the “dist” function within the base-R “stats” (version 3.5.3) package. The “kmeans” function requires user specification of both the number of random starts to consider before converging on a clustering solution and the number of clusters to seek. Everitt and others (2011) noted that it is best to initialize the algorithm several thousand times to promote finding a global clustering solution. The algorithm was therefore initiated 100,000 times in this study. The final solution was reproducible even when only five initializations were performed, which is strong evidence that the final solution is in fact the global solution rather than a local optimum. The number of clusters to seek was specified by examining plots of the within-group sum of squares (WSS) and mean silhouette width (MSW) metrics, as computed by the “stats” (version 3.5.3; base R) and “cluster” (version 2.0.9; Maechler and others, 2019) R packages, respectively. The WSS metric aims to illustrate the overall tightness of all clusters as an indicator of strong groupings, whereas the MSW metric conveys the likeness of data assigned to the same clusters (Everitt and Hothorn, 2010; Everitt and others, 2011). An elbow in the WSS plot (fig. 16*A*) implied that the data support a four-cluster solution. The MSW (fig. 16*B*), and therefore the mean similarity of data within each cluster, was maximized by a five-cluster solution (MSW = 0.4798) but had a very similar value for a four-cluster solution (MSW = 0.4790, percent difference = −0.17). Therefore, these metrics combined indicated that a four-cluster solution was supported by the data and appropriate to specify in the k-means clustering algorithm. The grouped hydrographs from the resulting clustering solution were examined to identify persistent characteristics of each cluster. Additionally, the silhouette width was calculated for each monitoring site by using the “cluster” (version 2.0.9; Maechler and others, 2019) R package to further interrogate the validity of the determined clusters.

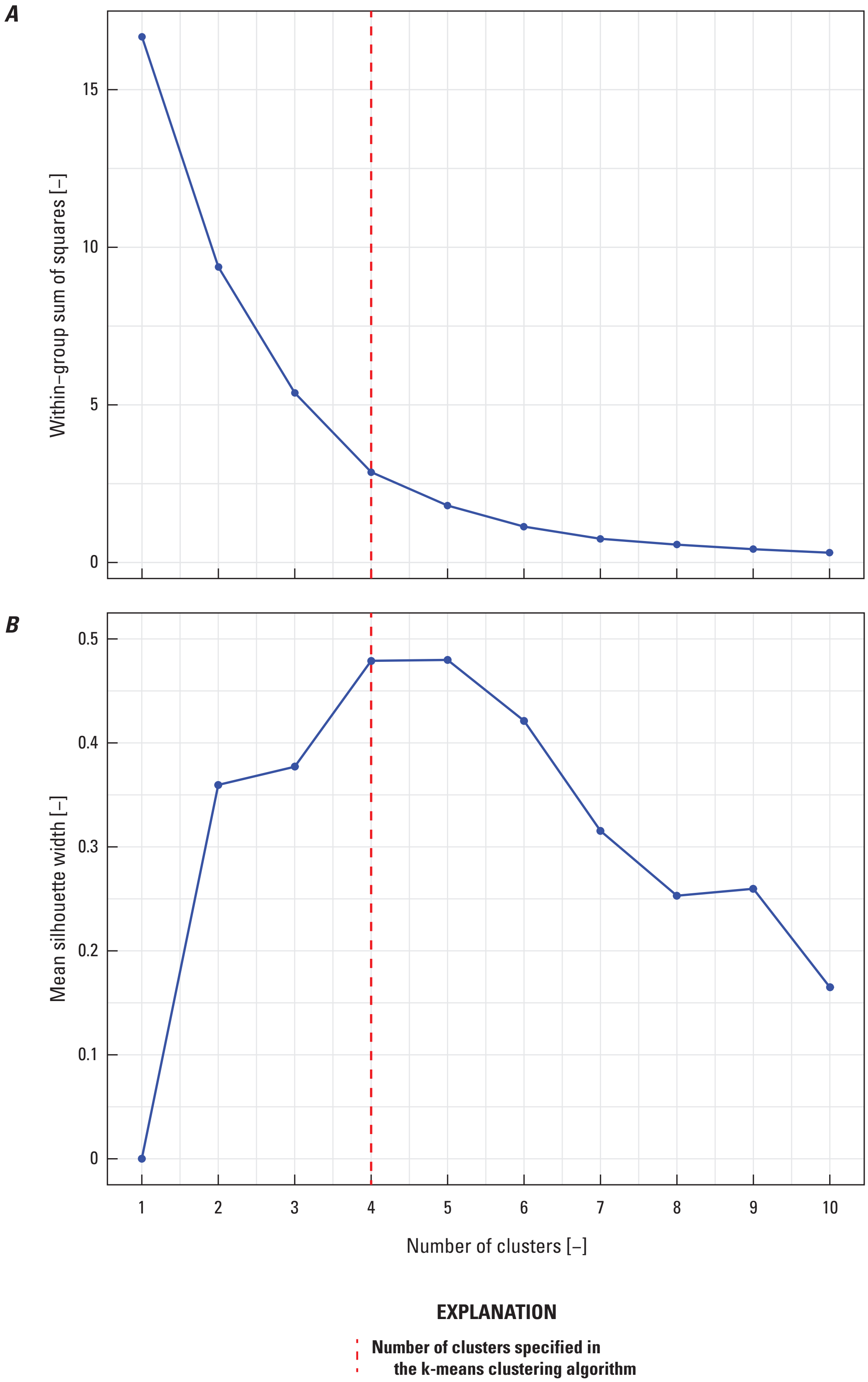


Figure 16. The *A*, within-group sum of squares and *B*, mean silhouette width metrics.

Optimization Assessment of the Observation Network

Spatial Assessment Results

Results from the spatial assessment (figs. 17–19, tables 6–8) can be used to identify sites that are particularly useful to create a groundwater-elevation surface and those that would have minimal effect on that surface if they were removed from the observation network. The spatial assessment used median groundwater elevations of the production zone over two winter periods: the 2001 time interval and 2015 time interval. Therefore, the results are most suited to identifying sites that could be removed from the observation network with minimal effect on the generation of winter groundwater-elevation contour maps of the production zone. Results from the temporal assessment and PCA can also be consulted when deciding which sites to remove entirely from the network, especially for sites that are monitored more frequently than annually.

Optimal networks were identified for the 2001 time interval and 2015 time interval by using the GA for 11 reduced network sizes: removing 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, and 40 sites from the observation network. The control parameter values used in the GA runs are listed in table 3 (“final value” column). The GA iteration stopping criteria were set at 1,000,000 for the maximum number of iterations and 500 for the maximum number of consecutive iterations without any improvement in the best fitness value (table 3). Table 6 summarizes the range of weighted objective function values for the GA runs, indicating the relative influence of each objective in determining the optimal solution (Fisher, 2013). The performance of each GA run (best fitness value, number of iterations, computation time, root-mean-square deviation [RMSD], and the percent local error [PLE]) is summarized in table 7. The RMSD is a measure of the difference between the kriged surfaces generated using the original and reduced networks (Fisher, 2013). The PLE is the maximum absolute difference in the kriged surfaces generated using the original and reduced networks divided by the range in groundwater elevations across the kriged surface generated using the original network. Figure 17 plots response curves for best fitness value, RMSD, and PLE for the GA runs for the 2001 time interval and 2015 time interval. Sites identified for removal for the GA runs are summarized in table 8, wherein “times identified” indicates the number of times the site was identified for removal in the 11 GA runs for the 2001 time interval and 2015 time interval. The difference in the kriged surfaces generated using the original and reduced networks and the sites identified for removal when removing 10, 20, 30, and 40 sites are shown spatially for the 2001 time interval in figure 18 and for the 2015 time interval in figure 19.

Weighted Objective Values

The spatial assessment placed an emphasis on observation network spatial coverage by applying a larger weighting coefficient to the first term (weighting coefficient of 100 applied to the mean standard error) and applying smaller weighting coefficients to the remaining terms (weighting coefficient of 1 applied to the second term, RMSE; to the third term, mean standard deviation; and to the fourth term, mean measurement uncertainty) in the multiobjective function (table 3). The weighted objective values are summed to produce a fitness value; therefore, the weighted objective values with the largest range during the GA search will dominate the evolution of the GA (Fisher, 2013). As indicated by the range in weighted objective values (table 6), the first and second terms in the multiobjective function had the largest effect on GA evolution, with the first term generally dominating the GA as intended.

Genetic Algorithm Search Performance and Sites Identified for Removal

The best fitness value, RMSD, and PLE generally increase for the 2001 time interval and 2015 time interval as the number of sites identified for removal from the observation network increases (fig. 17, table 7). The solution space for a GA run is proportional to the number of sites removed (Fisher, 2013); therefore, the number of iterations needed to converge on an optimal solution generally increased as the number of sites removed increased (table 7). Computation time was more variable (table 7), reaching a maximum at 15 and 10 sites removed for the 2001 time interval and 2015 time interval, respectively. Best fitness values generally decreased between GA runs implemented using base-case control parameter values (table 4) and the final control parameter values (table 7), which indicates that the GA performance was improved by the control parameter sensitivity analysis.

The response curves for RMSD and PLE show a rapid worsening (increase) starting at between 10 and 25 sites removed for the 2001 time interval and 2015 time interval (fig. 17). These response curves indicate that up to about 20 sites can be removed with a relatively small degradation in the ability to represent the winter kriged surfaces that were generated using the existing observation network. Ultimately, selection of an RMSD and a PLE that are acceptable is a management decision that depends on the amount of data degradation that is permissible while still being able to meet ABCWUA network objectives.

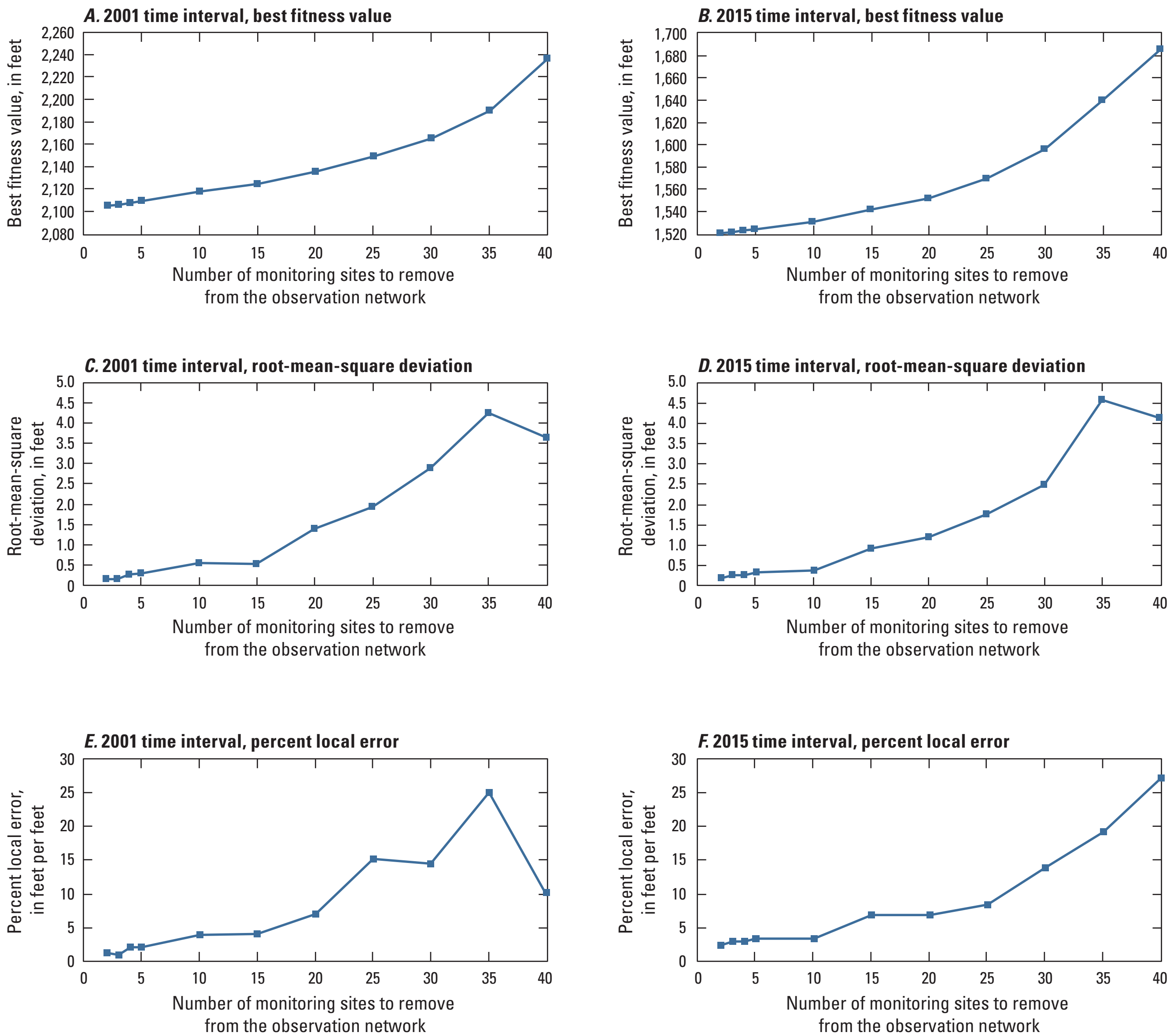


Figure 17. Response curves of genetic algorithm search performance when removing sites from the observation network: best fitness value for the *A*, 2001 time interval and *B*, 2015 time interval; root-mean-square deviation for the *C*, 2001 time interval and *D*, 2015 time interval; and percent local error for the *E*, 2001 time interval and *F*, 2015 time interval.

A. **2001 time interval, 10 sites removed**

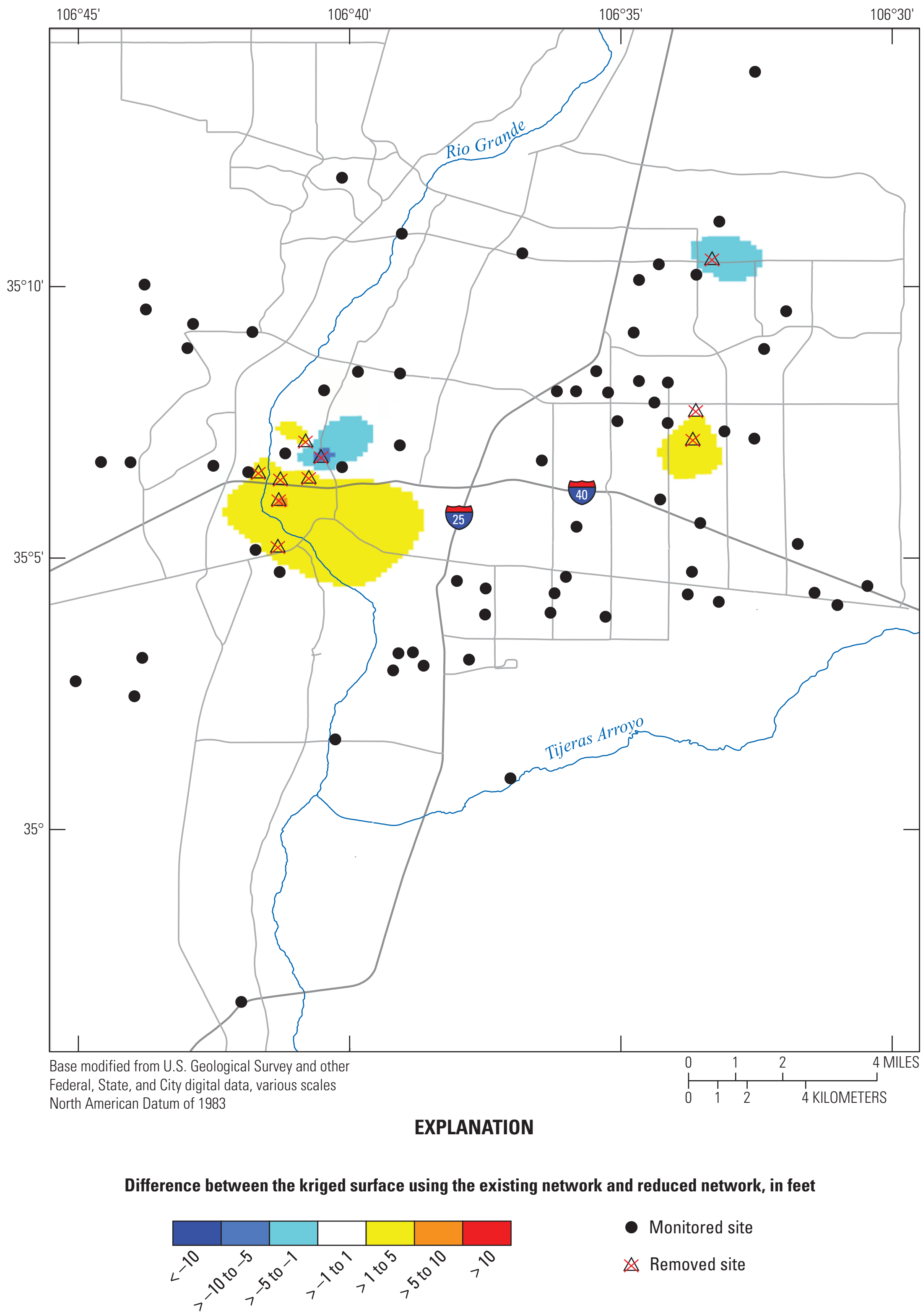


Figure 18. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2001 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.

B. **2001 time interval, 20 sites removed**

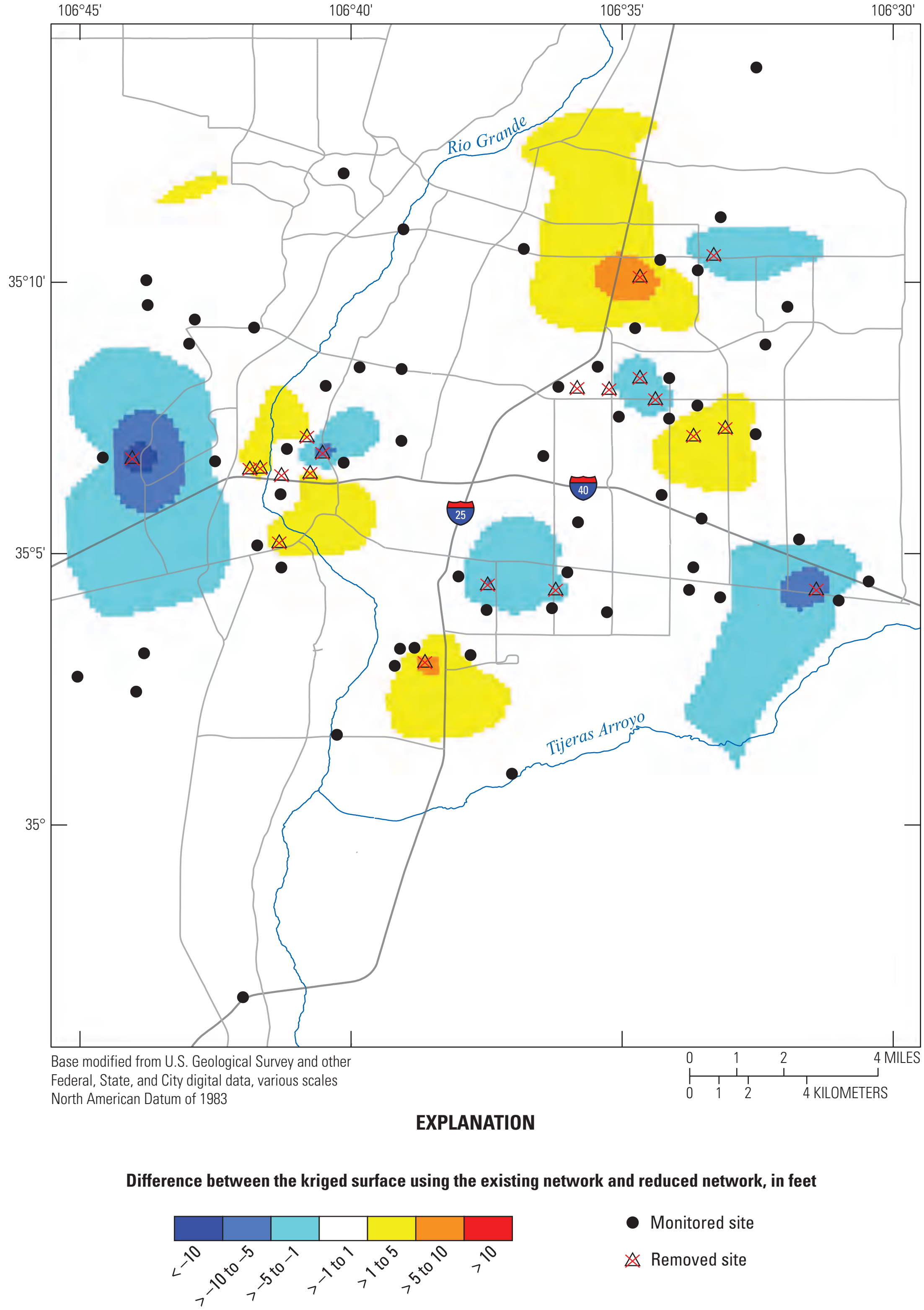


Figure 18. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2001 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

C. **2001 time interval, 30 sites removed**

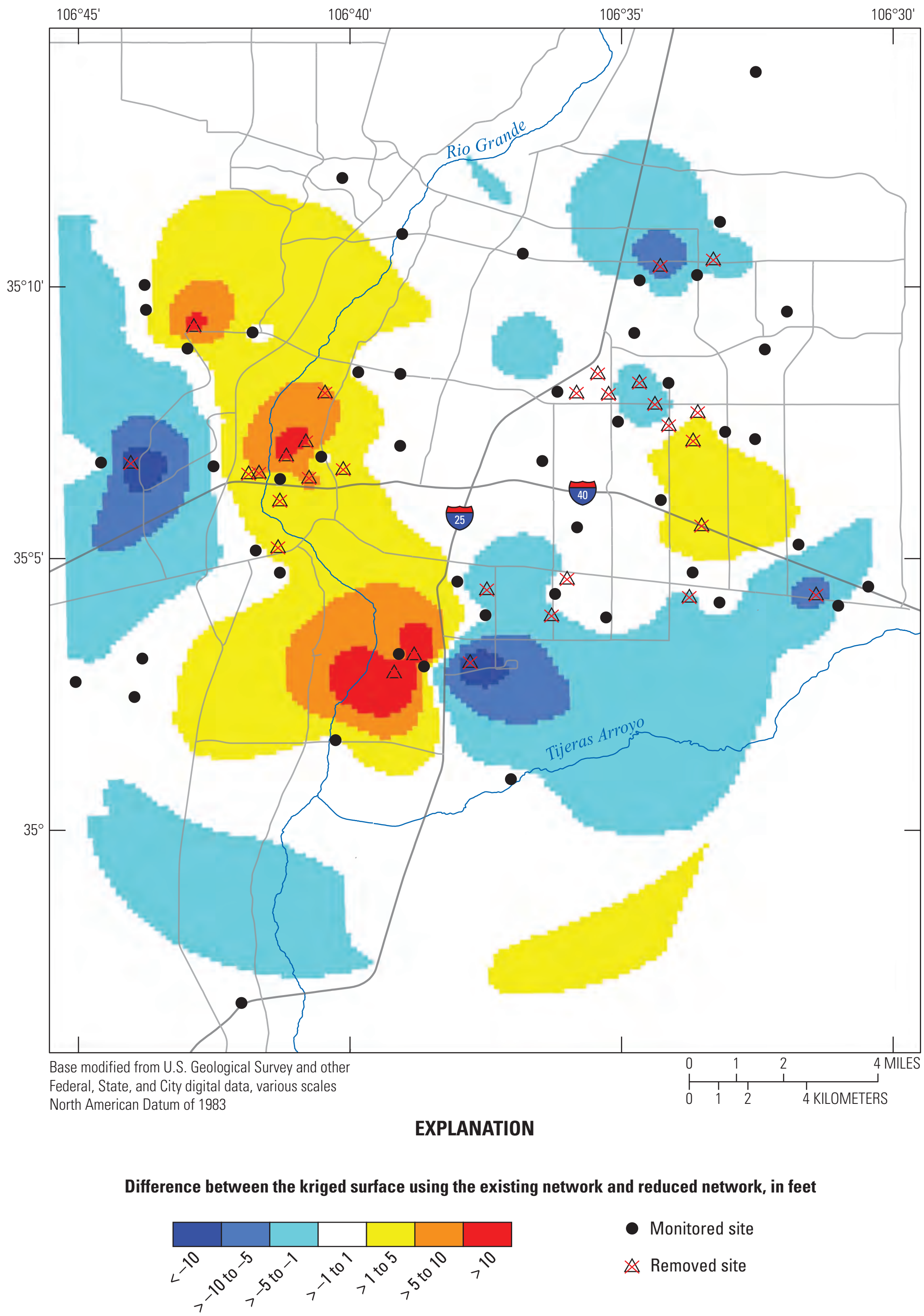


Figure 18. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2001 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

D. 2001 time interval, 40 sites removed

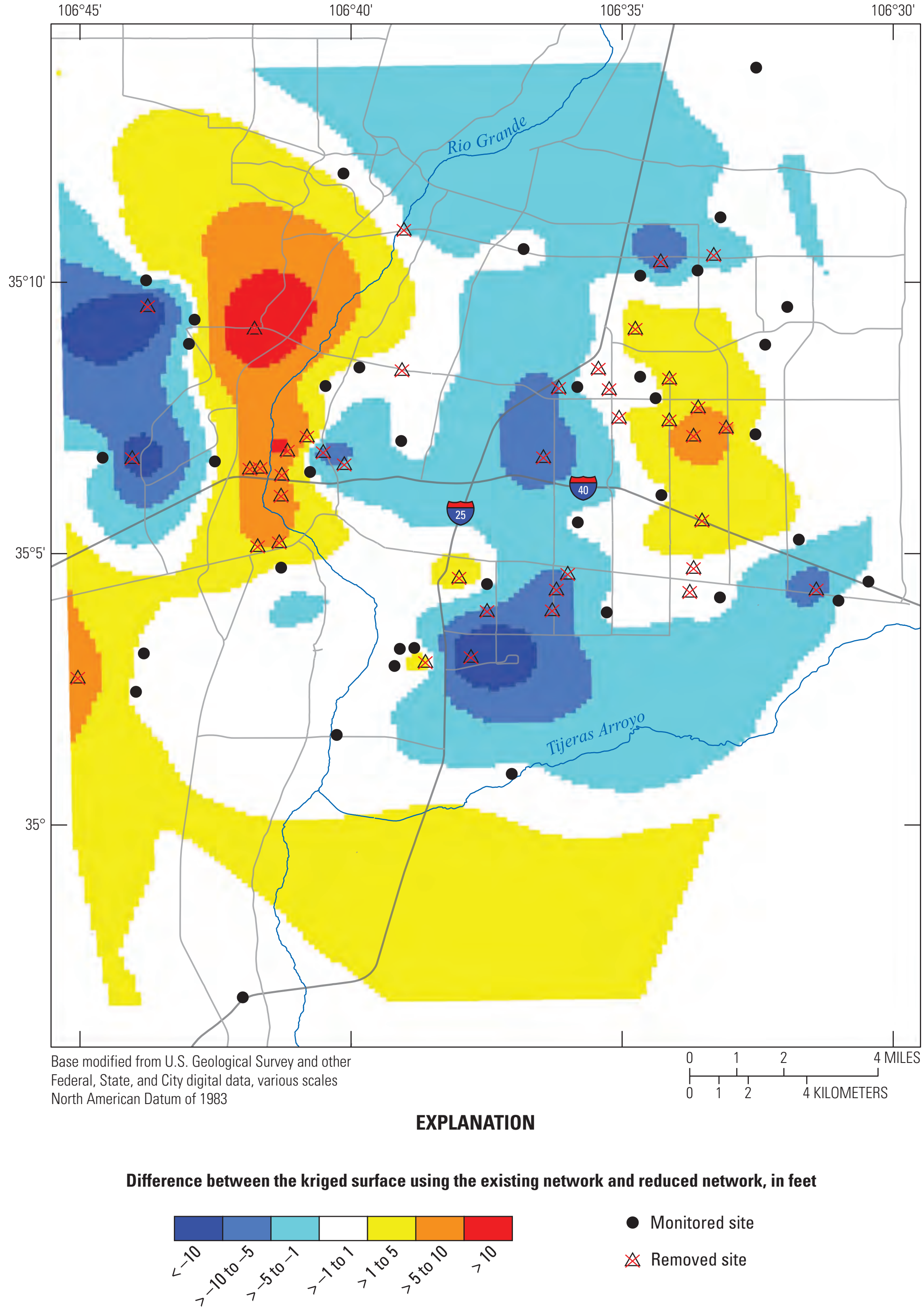


Figure 18. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2001 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

A. 2015 time interval, 10 sites removed

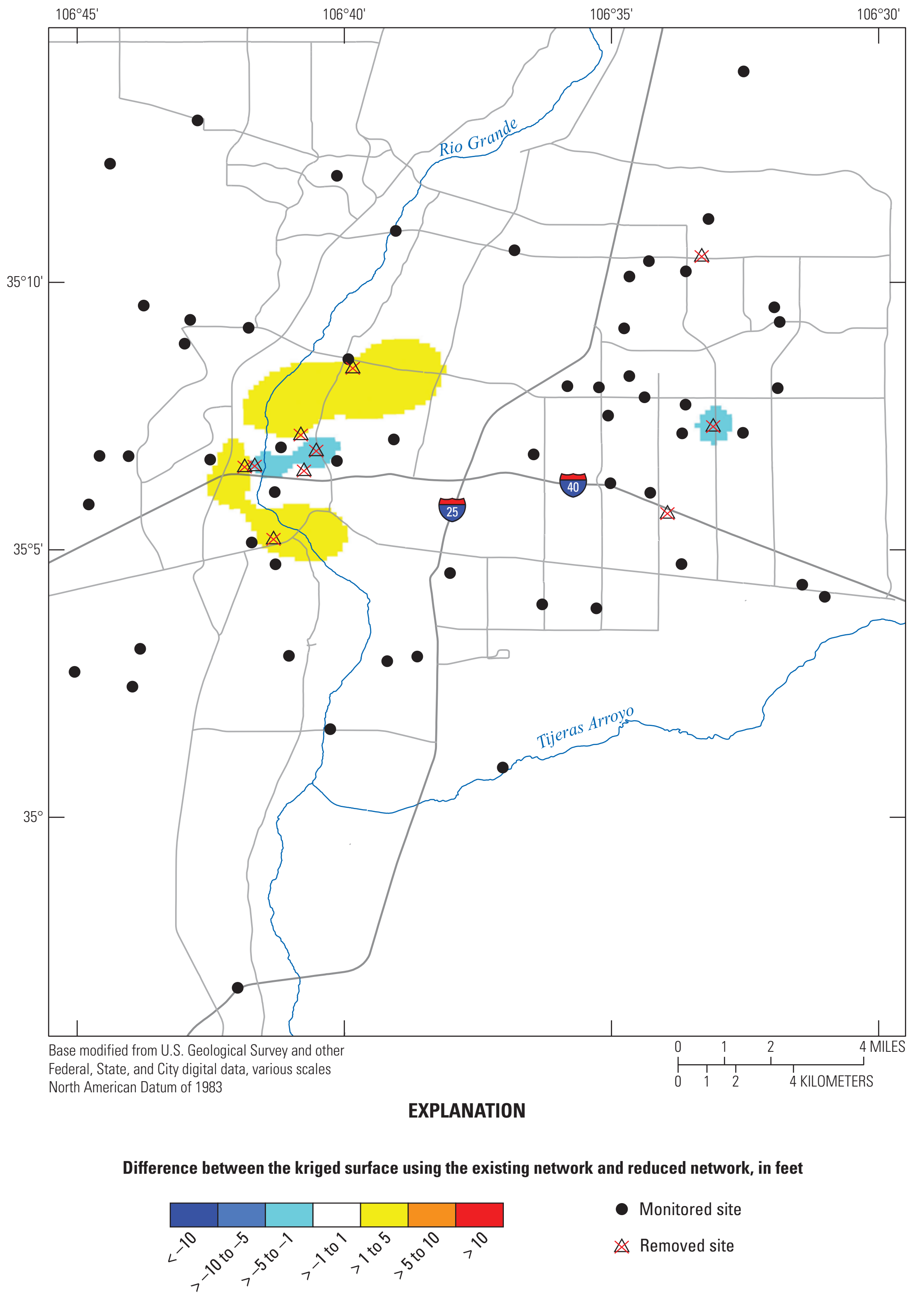


Figure 19. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2015 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.

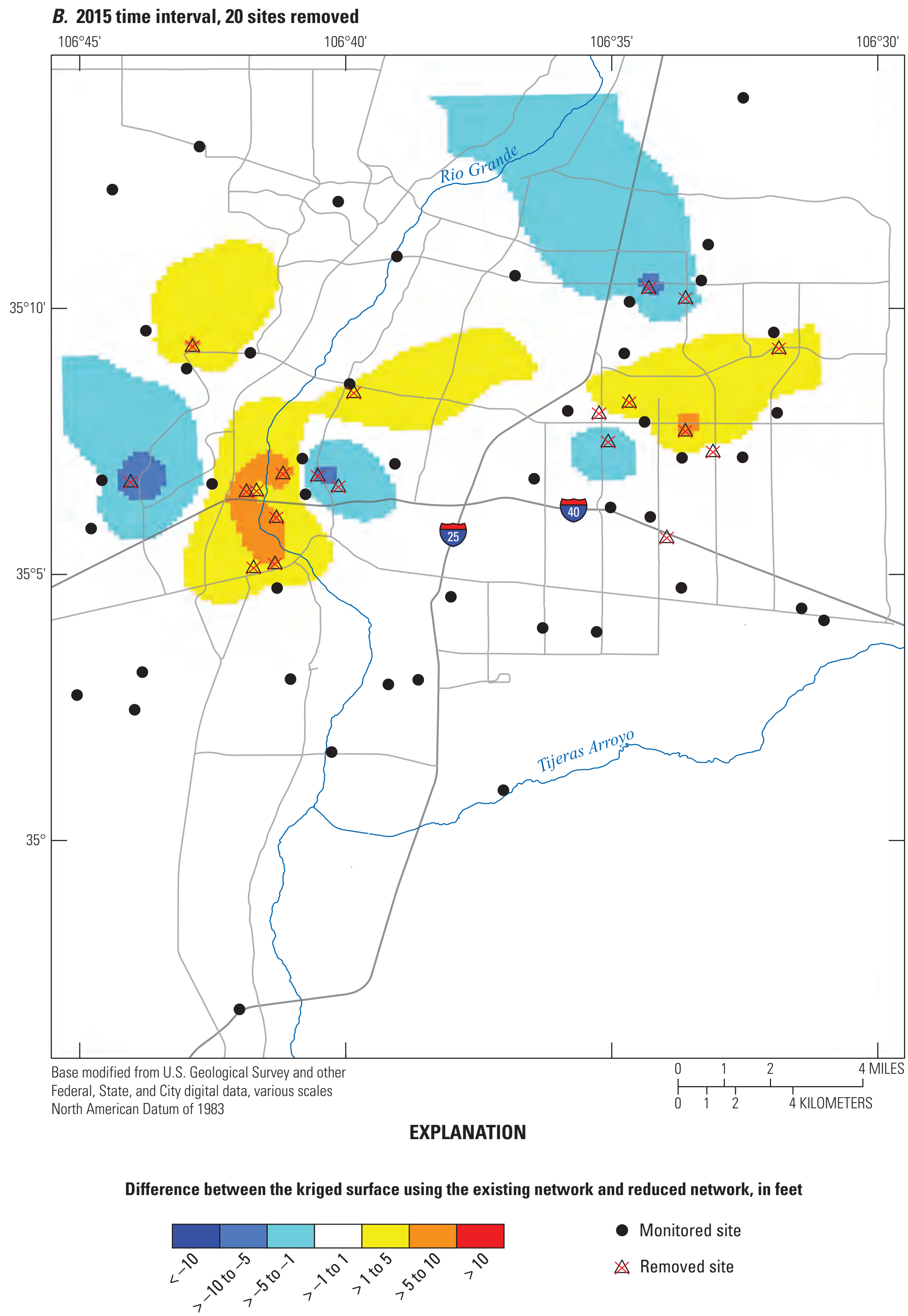


Figure 19. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2015 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

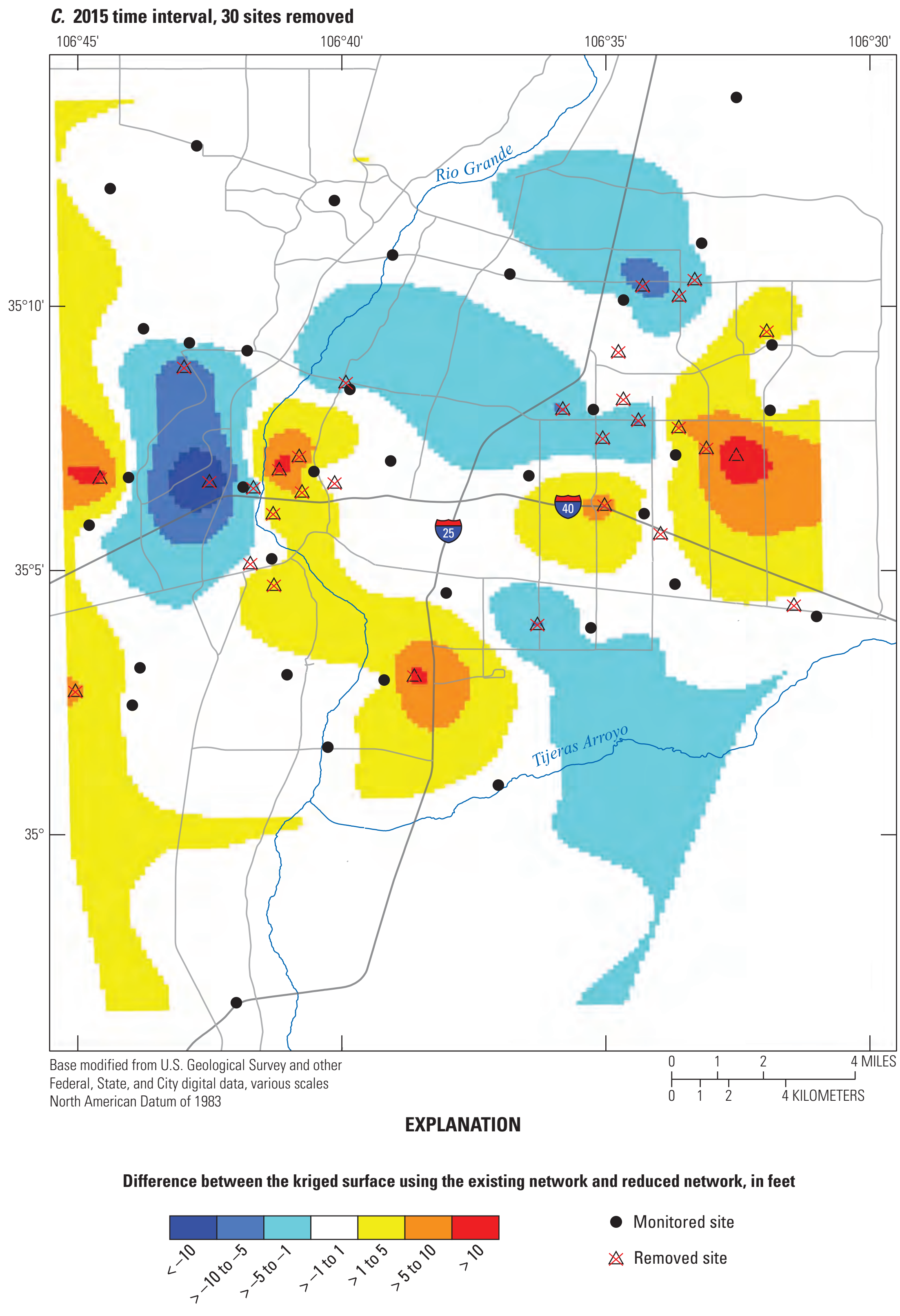


Figure 19. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2015 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

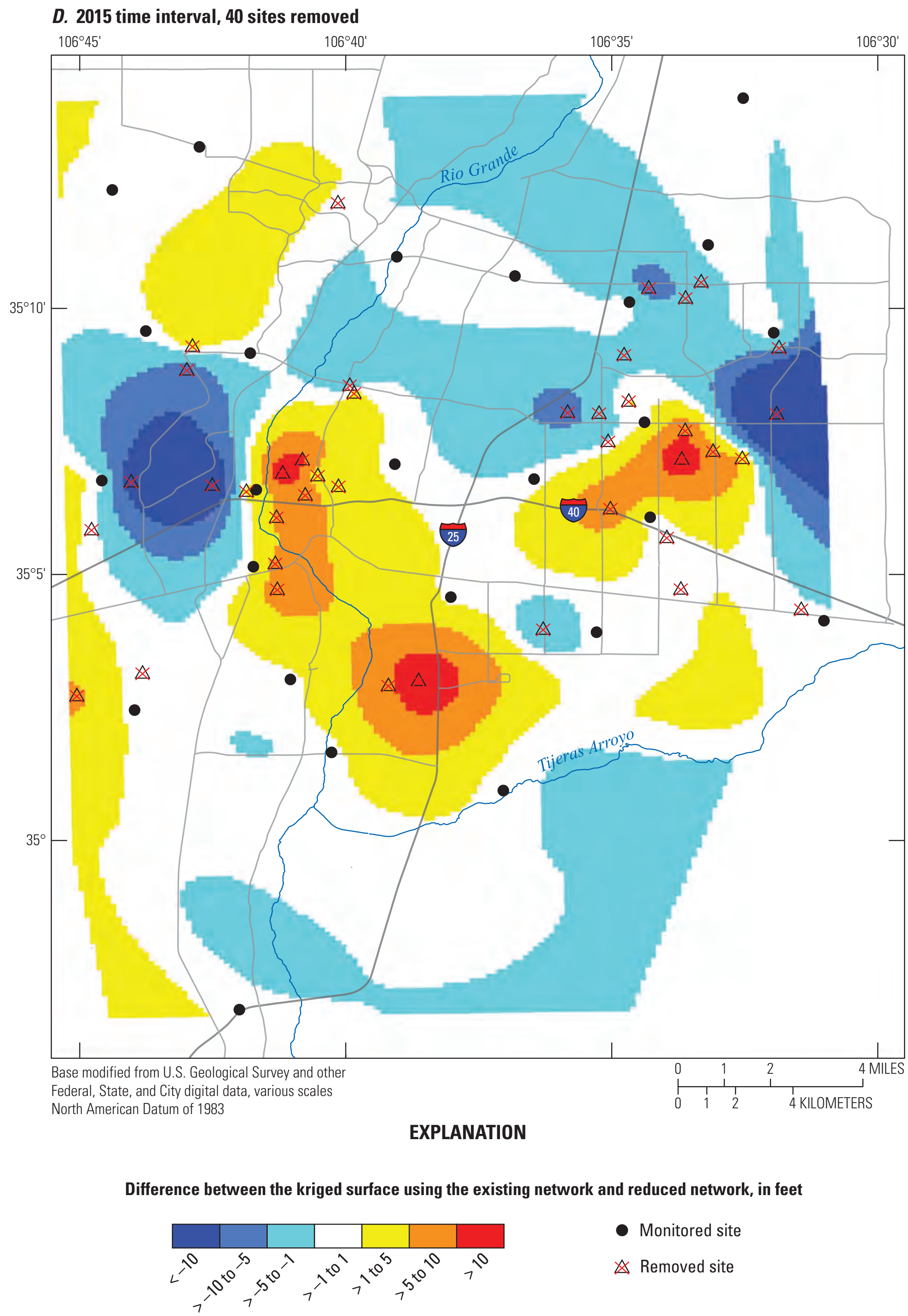


Figure 19. The kriged surface generated using the original network minus the kriged surface generated using the reduced networks for the 2015 time interval, and the sites identified for removal when removing *A*, 10; *B*, 20; *C*, 30; and *D*, 40 sites.—Continued

Table 6. Range of weighted objective values for genetic algorithm runs performed by removing sites from the observation network for the 2001 time interval and 2015 time interval.

Number of sites removed	Objective function	2001 time interval			2015 time interval		
		Minimum (feet)	Maximum (feet)	Range (feet)	Minimum (feet)	Maximum (feet)	Range (feet)
2	Standard error	2,097	2,323	227	1,513	1,727	214
	Root-mean-square error	0.11	131	131	0.033	75	75
	Standard deviation	0.79	72	71	0.47	38	37
	Measurement uncertainty	1.9	2.1	0.25	1.0	1.0	0.016
3	Standard error	2,097	2,395	298	1,513	1,788	274
	Root-mean-square error	0.18	111	111	0.086	71	71
	Standard deviation	0.94	59	58	0.73	34	33
	Measurement uncertainty	1.8	2.2	0.37	1.0	1.0	0.016
4	Standard error	2,097	2,411	314	1,514	1,803	289
	Root-mean-square error	0.22	101	101	0.11	64	64
	Standard deviation	1.3	53	52	1.1	31	30
	Measurement uncertainty	1.7	2.2	0.43	1.0	1.0	0.016
5	Standard error	2,098	2,471	373	1,515	1,844	329
	Root-mean-square error	0.30	93	93	0.25	70	69
	Standard deviation	1.3	47	45	1.2	29	28
	Measurement uncertainty	1.7	2.2	0.49	1.0	1.0	0.017
10	Standard error	2,103	2,580	477	1,519	2,007	488
	Root-mean-square error	0.92	71	70	1.3	53	52
	Standard deviation	2.8	33	31	2.3	23	21
	Measurement uncertainty	1.6	2.3	0.71	1.0	1.0	0.018
15	Standard error	2,109	2,716	607	1,527	2,058	532
	Root-mean-square error	2.1	68	66	2.3	45	43
	Standard deviation	3.8	28	24	3.7	18	15
	Measurement uncertainty	1.5	2.4	0.89	1.0	1.0	0.020
20	Standard error	2,113	2,794	680	1,536	2,175	639
	Root-mean-square error	3.3	58	55	2.9	41	38
	Standard deviation	4.2	24	20	4.5	17	12
	Measurement uncertainty	1.4	2.5	1.1	1.0	1.0	0.022
25	Standard error	2,126	2,818	692	1,552	2,377	825
	Root-mean-square error	4.3	56	51	4.3	40	36
	Standard deviation	5.6	21	16	5.4	16	10
	Measurement uncertainty	1.3	2.6	1.3	1.0	1.0	0.025
30	Standard error	2,141	2,949	808	1,578	2,426	848
	Root-mean-square error	5.2	52	47	5.9	38	32
	Standard deviation	6.7	20	13	6.1	14	7.7
	Measurement uncertainty	1.3	2.7	1.4	1.0	1.0	0.028

Table 6. Range of weighted objective values for genetic algorithm runs performed by removing sites from the observation network for the 2001 time interval and 2015 time interval.—Continued

Number of sites removed	Objective function	2001 time interval			2015 time interval		
		Minimum (feet)	Maximum (feet)	Range (feet)	Minimum (feet)	Maximum (feet)	Range (feet)
35	Standard error	2,164	3,007	843	1,619	2,527	908
	Root-mean-square error	6.7	51	45	6.5	42	35
	Standard deviation	7.0	18	11	6.8	13	6.2
	Measurement uncertainty	1.3	2.9	1.7	1.0	1.0	0.033
40	Standard error	2,208	3,091	883	1,665	2,845	1,180
	Root-mean-square error	7.2	49	42	7.4	41	33
	Standard deviation	8.0	17	9.3	7.1	13	5.5
	Measurement uncertainty	1.2	3.1	1.9	1.0	1.0	0.040

Table 7. Summary of genetic algorithm search performance when removing sites from the observation network for the 2001 time interval and 2015 time interval.

[RMSD, root-mean-square deviation; PLE, percent local error]

Number of sites removed	2001 time interval					2015 time interval				
	Best fitness value (feet)	Number of iterations	Computation time (hours)	RMSD (feet)	PLE	Best fitness value (feet)	Number of iterations	Computation time (hours)	RMSD (feet)	PLE
2	2,105	501	12	0.10	1.1	1,520	500	17	0.19	2.2
3	2,106	529	25	0.10	1.0	1,521	511	20	0.25	2.8
4	2,107	577	30	0.22	2.1	1,523	528	17	0.25	2.8
5	2,109	616	19	0.25	2.1	1,524	1,250	37	0.33	3.2
10	2,117	690	25	0.51	3.8	1,531	2,341	71	0.37	3.3
15	2,124	1,778	50	0.49	4.0	1,541	1,063	37	0.91	6.7
20	2,136	1,456	13	1.4	7.0	1,552	2,139	19	1.2	6.8
25	2,149	1,631	6.9	1.9	15	1,569	1,718	9.5	1.8	8.3
30	2,165	3,169	12	2.9	14	1,596	2,111	8.7	2.5	14
35	2,190	2,753	10	4.2	25	1,640	1,544	9.9	4.6	19
40	2,236	1,026	4.4	3.6	10	1,685	1,807	8.6	4.1	27

Table 8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval.

[USGS, U.S. Geological Survey. “Times identified” indicates the number of times the site was identified for removal in the 11 genetic algorithm runs for each interval. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3]

Number of sites removed	2001 time interval			2015 time interval		
	USGS site number	Site name	Times identified	USGS site number	Site name	Times identified
2	350628106411501	Duranes #4 (DUR-4)	9	350508106411901	Atrisco #3 (ATR-3)	10
	350605106411801	Duranes #5 (DUR-5)	5	350629106405101	Duranes #3 (DUR-3)	8
3	350638106413702	West Bluff Nest 1	10	350508106411901	Atrisco #3 (ATR-3)	10
	350508106411901	Atrisco #3 (ATR-3)	10	350708106405801	Duranes #2 (DUR-2)	7
	350628106411501	Duranes #4 (DUR-4)	9	350629106405101	Duranes #3 (DUR-3)	8
4	350638106413702	West Bluff Nest 1	10	350508106411901	Atrisco #3 (ATR-3)	10
	350508106411901	Atrisco #3 (ATR-3)	10	350642106401101	Duranes #1 (DUR-1)	7
	350629106405101	Duranes #3 (DUR-3)	8	350708106405801	Duranes #2 (DUR-2)	7
	350628106411501	Duranes #4 (DUR-4)	9	350629106405101	Duranes #3 (DUR-3)	8
5	350638106413702	West Bluff Nest 1	10	350508106411901	Atrisco #3 (ATR-3)	10
	350508106411901	Atrisco #3 (ATR-3)	10	350708106405801	Duranes #2 (DUR-2)	7
	350629106405101	Duranes #3 (DUR-3)	8	350629106405101	Duranes #3 (DUR-3)	8
	350628106411501	Duranes #4 (DUR-4)	9	350653106403001	Duranes #6 (DUR-6)	7
	351029106332001	Webster #1 (WEB-1)	8	350827106395001	Griegos #1 (GRI-1)	6
10	350638106413702	West Bluff Nest 1	10	350545106335902	Jerry Cline Park	5
	350508106411901	Atrisco #3 (ATR-3)	10	350638106413702	West Bluff Nest 1	6
	350708106405801	Duranes #2 (DUR-2)	6	350508106411901	Atrisco #3 (ATR-3)	10
	350629106405101	Duranes #3 (DUR-3)	8	350708106405801	Duranes #2 (DUR-2)	7
	350628106411501	Duranes #4 (DUR-4)	9	350629106405101	Duranes #3 (DUR-3)	8
	350605106411801	Duranes #5 (DUR-5)	5	350653106403001	Duranes #6 (DUR-6)	7
	350653106403001	Duranes #6 (DUR-6)	6	350635106415001	Gonzales #2 (GON-2)	6
	350744106333501	Thomas #5 (THO-5)	6	350827106395001	Griegos #1 (GRI-1)	6
	350719106333401	Thomas #7 (THO-7)	6	350720106330401	Thomas #6 (THO-6)	6
	351029106332001	Webster #1 (WEB-1)	8	351029106332001	Webster #1 (WEB-1)	6
15	350638106413702	West Bluff Nest 1	10	350638106413702	West Bluff Nest 1	6
	350508106411901	Atrisco #3 (ATR-3)	10	350508106411901	Atrisco #3 (ATR-3)	10
	350359106362401	Burton #1 (BUR-1)	4	351007106343801	Coronado #2 (COR-2)	1
	350708106405801	Duranes #2 (DUR-2)	6	350642106401101	Duranes #1 (DUR-1)	7
	350629106405101	Duranes #3 (DUR-3)	8	350605106411801	Duranes #5 (DUR-5)	6
	350628106411501	Duranes #4 (DUR-4)	9	350653106403001	Duranes #6 (DUR-6)	7
	350653106403001	Duranes #6 (DUR-6)	6	350655106395001	Duranes #7 (DUR-7)	6
	350635106415001	Gonzales #2 (GON-2)	6	350635106415001	Gonzales #2 (GON-2)	6
	350301106383601	San Jose #3 (SAJ-3)	4	350827106395001	Griegos #1 (GRI-1)	6
	350744106333501	Thomas #5 (THO-5)	6	350744106333501	Thomas #5 (THO-5)	5
	350720106330401	Thomas #6 (THO-6)	4	350720106330401	Thomas #6 (THO-6)	6
	350805106354901	Vol Andia #1 (VOA-1)	4	350803106351101	Vol Andia #4 (VOA-4)	5
	350803106351101	Vol Andia #4 (VOA-4)	6	350918106315401	Walker #4 (WAL-4)	3
	351029106332001	Webster #1 (WEB-1)	8	351029106332001	Webster #1 (WEB-1)	6
	350358106372901	Yale #2 (YAL-2)	3	351013106333501	Webster #2 (WEB-2)	5

Table 8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval.—Continued

[USGS, U.S. Geological Survey. "Times identified" indicates the number of times the site was identified for removal in the 11 genetic algorithm runs for each interval. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3]

Number of sites removed	2001 time interval			2015 time interval		
	USGS site number	Site name	Times identified	USGS site number	Site name	Times identified
20	350638106413702	West Bluff Nest 1	10	350545106335902	Jerry Cline Park	5
	350508106411901	Atrisco #3 (ATR-3)	10	350638106413702	West Bluff Nest 1	6
	350421106361001	Burton #2 (BUR-2)	2	350508106411901	Atrisco #3 (ATR-3)	10
	350647106440001	College #2 (COL-2)	5	350509106414401	Atrisco #4 (ATR-4)	3
	351007106343801	Coronado #2 (COR-2)	3	350647106440001	College #2 (COL-2)	4
	350708106405801	Duranes #2 (DUR-2)	6	351025106341601	Coronado #1 (COR-1)	5
	350629106405101	Duranes #3 (DUR-3)	8	350642106401101	Duranes #1 (DUR-1)	7
	350628106411501	Duranes #4 (DUR-4)	9	350605106411801	Duranes #5 (DUR-5)	6
	350653106403001	Duranes #6 (DUR-6)	6	350653106403001	Duranes #6 (DUR-6)	7
	350635106415001	Gonzales #2 (GON-2)	6	350655106395001	Duranes #7 (DUR-7)	6
	350752106342101	Leyendecker #1 (LEY-1)	4	350635106415001	Gonzales #2 (GON-2)	6
	350819106344001	Leyendecker #3 (LEY-3)	4	350827106395001	Griegos #1 (GRI-1)	6
	350422106312401	Lomas #5 (LOM-5)	5	350819106344001	Leyendecker #3 (LEY-3)	5
	350301106383601	San Jose #3 (SAJ-3)	4	350744106333501	Thomas #5 (THO-5)	5
	350720106330401	Thomas #6 (THO-6)	4	350720106330401	Thomas #6 (THO-6)	6
	350719106333401	Thomas #7 (THO-7)	6	350732106350101	Vol Andia #2 (VOA-2)	4
	350805106354901	Vol Andia #1 (VOA-1)	4	350803106351101	Vol Andia #4 (VOA-4)	5
	350803106351101	Vol Andia #4 (VOA-4)	6	350918106315401	Walker #4 (WAL-4)	3
	351029106332001	Webster #1 (WEB-1)	8	351013106333501	Webster #2 (WEB-2)	5
	350426106372601	Yale #1 (YAL-1)	4	350918106425401	Zamora #1 (ZAM-1)	4
25	350638106413702	West Bluff Nest 1	10	350256106390801	San Jose 9	3
	350508106411901	Atrisco #3 (ATR-3)	10	350638106413702	West Bluff Nest 1	6
	350359106362401	Burton #1 (BUR-1)	4	350508106411901	Atrisco #3 (ATR-3)	10
	350440106355801	Burton #3 (BUR-3)	4	350509106414401	Atrisco #4 (ATR-4)	3
	350647106440001	College #2 (COL-2)	5	350606106341101	Charles #2 (CHA-2)	2
	351007106343801	Coronado #2 (COR-2)	3	350647106440001	College #2 (COL-2)	4
	350629106405101	Duranes #3 (DUR-3)	8	351025106341601	Coronado #1 (COR-1)	5
	350628106411501	Duranes #4 (DUR-4)	9	350642106401101	Duranes #1 (DUR-1)	7
	350653106403001	Duranes #6 (DUR-6)	6	350629106405101	Duranes #3 (DUR-3)	8
	350655106395001	Duranes #7 (DUR-7)	4	350605106411801	Duranes #5 (DUR-5)	6
	350635106415001	Gonzales #2 (GON-2)	6	350653106403001	Duranes #6 (DUR-6)	7
	350802106402901	Griegos #3 (GRI-3)	2	350655106395001	Duranes #7 (DUR-7)	6
	350752106342101	Leyendecker #1 (LEY-1)	4	350635106415001	Gonzales #2 (GON-2)	6
	350819106344001	Leyendecker #3 (LEY-3)	4	350827106395001	Griegos #1 (GRI-1)	6
	350422106312401	Lomas #5 (LOM-5)	5	350752106342101	Leyendecker #1 (LEY-1)	3
	350445106334001	Ridgecrest #4 (RID-4)	3	350819106344001	Leyendecker #3 (LEY-3)	5
	350316106384801	San Jose #1 (SAJ-1)	3	350422106312401	Lomas #5 (LOM-5)	3
	350744106333501	Thomas #5 (THO-5)	6	350744106333501	Thomas #5 (THO-5)	5
	350720106330401	Thomas #6 (THO-6)	4	350719106333401	Thomas #7 (THO-7)	3

Table 8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval.—Continued

[USGS, U.S. Geological Survey. “Times identified” indicates the number of times the site was identified for removal in the 11 genetic algorithm runs for each interval. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3]

Number of sites removed	2001 time interval			2015 time interval		
	USGS site number	Site name	Times identified	USGS site number	Site name	Times identified
25—Continued	350719106333401	Thomas #7 (THO-7)	6	350711106323101	Thomas #8 (THO-8)	4
	350805106354901	Vol Andia #1 (VOA-1)	4	350803106351101	Vol Andia #4 (VOA-4)	5
	350732106350101	Vol Andia #2 (VOA-2)	2	350931106315501	Walker #3 (WAL-3)	3
	350803106351101	Vol Andia #4 (VOA-4)	6	351029106332001	Webster #1 (WEB-1)	6
	351029106332001	Webster #1 (WEB-1)	8	350918106425401	Zamora #1 (ZAM-1)	4
	350426106372601	Yale #1 (YAL-1)	4	350852106425701	Zamora #2 (ZAM-2)	3
30	350256106390801	San Jose 9	1	350244106450202	Westgate Heights Park	2
	350638106413702	West Bluff Nest 1	10	350545106335902	Jerry Cline Park	5
	350508106411901	Atrisco #3 (ATR-3)	10	350638106413702	West Bluff Nest 1	6
	350359106362401	Burton #1 (BUR-1)	4	350836106395401	Montaño Nest 6	3
	350440106355801	Burton #3 (BUR-3)	4	350908106344402	Sister Cities	3
	350647106440001	College #2 (COL-2)	5	350445106411501	Atrisco #2 (ATR-2)	3
	351025106341601	Coronado #1 (COR-1)	2	350509106414401	Atrisco #4 (ATR-4)	3
	350642106401101	Duranes #1 (DUR-1)	3	350359106362401	Burton #1 (BUR-1)	3
	350708106405801	Duranes #2 (DUR-2)	6	350615106345901	Charles #5 (CHA-5)	2
	350629106405101	Duranes #3 (DUR-3)	8	350646106443201	College #1 (COL-1)	1
	350605106411801	Duranes #5 (DUR-5)	5	351025106341601	Coronado #1 (COR-1)	5
	350655106395001	Duranes #7 (DUR-7)	4	350642106401101	Duranes #1 (DUR-1)	7
	350635106415001	Gonzales #2 (GON-2)	6	350708106405801	Duranes #2 (DUR-2)	7
	350802106402901	Griegos #3 (GRI-3)	2	350629106405101	Duranes #3 (DUR-3)	8
	350752106342101	Leyendecker #1 (LEY-1)	4	350605106411801	Duranes #5 (DUR-5)	6
	350727106340801	Leyendecker #2 (LEY-2)	3	350655106395001	Duranes #7 (DUR-7)	6
	350819106344001	Leyendecker #3 (LEY-3)	4	350642106422801	Gonzales #1 (GON-1)	3
	350422106312401	Lomas #5 (LOM-5)	5	350752106342101	Leyendecker #1 (LEY-1)	3
	350538106333001	Love #8 (LOV-8)	2	350819106344001	Leyendecker #3 (LEY-3)	5
	350308106374601	Miles Road #1 (MIL-1)	2	350422106312401	Lomas #5 (LOM-5)	3
	350420106334401	Ridgecrest #5 (RID-5)	2	350301106383601	San Jose #3 (SAJ-3)	3
	350316106384801	San Jose #1 (SAJ-1)	3	350744106333501	Thomas #5 (THO-5)	5
	350744106333501	Thomas #5 (THO-5)	6	350720106330401	Thomas #6 (THO-6)	6
	350719106333401	Thomas #7 (THO-7)	6	350711106323101	Thomas #8 (THO-8)	4
	350805106354901	Vol Andia #1 (VOA-1)	4	350805106354901	Vol Andia #1 (VOA-1)	2
	350803106351101	Vol Andia #4 (VOA-4)	6	350732106350101	Vol Andia #2 (VOA-2)	4
	350828106352101	Vol Andia #6 (VOA-6)	3	350931106315501	Walker #3 (WAL-3)	3
	351029106332001	Webster #1 (WEB-1)	8	351029106332001	Webster #1 (WEB-1)	6
	350426106372601	Yale #1 (YAL-1)	4	351013106333501	Webster #2 (WEB-2)	5
	350918106425401	Zamora #1 (ZAM-1)	2	350852106425701	Zamora #2 (ZAM-2)	3

Table 8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval.—Continued

[USGS, U.S. Geological Survey. "Times identified" indicates the number of times the site was identified for removal in the 11 genetic algorithm runs for each interval. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3]

Number of sites removed	2001 time interval			2015 time interval		
	USGS site number	Site name	Times identified	USGS site number	Site name	Times identified
	350534106354702	Del Sol Divider	1	350256106390801	San Jose 9	3
	350638106413702	West Bluff Nest 1	10	350545106335902	Jerry Cline Park	5
	350508106411901	Atrisco #3 (ATR-3)	10	350638106413702	West Bluff Nest 1	6
	350440106355801	Burton #3 (BUR-3)	4	350706106390302	Garfield Park	1
	350355106351501	Burton #5 (BUR-5)	1	350836106395401	Montaño Nest 6	3
	350647106440001	College #2 (COL-2)	5	350908106344402	Sister Cities	3
	351007106343801	Coronado #2 (COR-2)	3	350910106414802	Sierra Vista	1
	350642106401101	Duranes #1 (DUR-1)	3	350445106411501	Atrisco #2 (ATR-2)	3
	350708106405801	Duranes #2 (DUR-2)	6	350508106411901	Atrisco #3 (ATR-3)	10
	350629106405101	Duranes #3 (DUR-3)	8	350359106362401	Burton #1 (BUR-1)	3
	350605106411801	Duranes #5 (DUR-5)	5	350606106341101	Charles #2 (CHA-2)	2
	350653106403001	Duranes #6 (DUR-6)	6	350647106440001	College #2 (COL-2)	4
	350655106395001	Duranes #7 (DUR-7)	4	351025106341601	Coronado #1 (COR-1)	5
	350635106415001	Gonzales #2 (GON-2)	6	350642106401101	Duranes #1 (DUR-1)	7
	350827106395001	Griegos #1 (GRI-1)	1	350708106405801	Duranes #2 (DUR-2)	7
	350752106342101	Leyendecker #1 (LEY-1)	4	350605106411801	Duranes #5 (DUR-5)	6
	350727106340801	Leyendecker #2 (LEY-2)	3	350653106403001	Duranes #6 (DUR-6)	7
35	350819106344001	Leyendecker #3 (LEY-3)	4	350655106395001	Duranes #7 (DUR-7)	6
	350422106312401	Lomas #5 (LOM-5)	5	350642106422801	Gonzales #1 (GON-1)	3
	350445106334001	Ridgecrest #4 (RID-4)	3	350635106415001	Gonzales #2 (GON-2)	6
	350316106384801	San Jose #1 (SAJ-1)	3	350223106435401	Leavitt #3 (LEA-3)	1
	350315106390401	San Jose #2 (SAJ-2)	1	350752106342101	Leyendecker #1 (LEY-1)	3
	350301106383601	San Jose #3 (SAJ-3)	4	350819106344001	Leyendecker #3 (LEY-3)	5
	350744106333501	Thomas #5 (THO-5)	6	350408106310101	Lomas #6 (LOM-6)	1
	350719106333401	Thomas #7 (THO-7)	6	350301106383601	San Jose #3 (SAJ-3)	3
	350711106323101	Thomas #8 (THO-8)	1	350648106362501	Santa Barbara #1 (SAB-1)	1
	350803106351101	Vol Andia #4 (VOA-4)	6	350720106330401	Thomas #6 (THO-6)	6
	350809106360901	Vol Andia #5 (VOA-5)	2	350719106333401	Thomas #7 (THO-7)	3
	350828106352101	Vol Andia #6 (VOA-6)	3	350711106323101	Thomas #8 (THO-8)	4
	350950106434001	Volcano Cliffs #1 (VOC-1)	2	350732106350101	Vol Andia #2 (VOA-2)	4
	351029106332001	Webster #1 (WEB-1)	8	350803106351101	Vol Andia #4 (VOA-4)	5
	351013106333501	Webster #2 (WEB-2)	1	350931106315501	Walker #3 (WAL-3)	3
	350426106372601	Yale #1 (YAL-1)	4	351029106332001	Webster #1 (WEB-1)	6
	350358106372901	Yale #2 (YAL-2)	3	351013106333501	Webster #2 (WEB-2)	5
	350918106425401	Zamora #1 (ZAM-1)	2	350918106425401	Zamora #1 (ZAM-1)	4

Table 8. Sites identified for removal by genetic algorithm searches for the 2001 time interval and 2015 time interval.—Continued

[USGS, U.S. Geological Survey. “Times identified” indicates the number of times the site was identified for removal in the 11 genetic algorithm runs for each interval. Site name abbreviations listed in parentheses are used as labels for production wells in figure 3]

Number of sites removed	2001 time interval			2015 time interval		
	USGS site number	Site name	Times identified	USGS site number	Site name	Times identified
40	350244106450202	Westgate Heights Park	1	350244106450202	Westgate Heights Park	2
	350638106413702	West Bluff Nest 1	10	350256106390801	San Jose 9	3
	350908106344402	Sister Cities	1	350545106335902	Jerry Cline Park	5
	350910106414802	Sierra Vista	1	350836106395401	Montaño Nest 6	3
	351059106385903	Paseo del Norte Nest 1	1	350908106344402	Sister Cities	3
	350508106411901	Atrisco #3 (ATR-3)	10	351201106400502	Hunters Ridge Nest 1	1
	350509106414401	Atrisco #4 (ATR-4)	1	350552106444601	Arroyo Vista	1
	350359106362401	Burton #1 (BUR-1)	4	350445106411501	Atrisco #2 (ATR-2)	3
	350421106361001	Burton #2 (BUR-2)	2	350508106411901	Atrisco #3 (ATR-3)	10
	350440106355801	Burton #3 (BUR-3)	4	350359106362401	Burton #1 (BUR-1)	3
	350647106440001	College #2 (COL-2)	5	350615106345901	Charles #5 (CHA-5)	2
	351025106341601	Coronado #1 (COR-1)	2	350647106440001	College #2 (COL-2)	4
	350642106401101	Duranes #1 (DUR-1)	3	351025106341601	Coronado #1 (COR-1)	5
	350708106405801	Duranes #2 (DUR-2)	6	350642106401101	Duranes #1 (DUR-1)	7
	350628106411501	Duranes #4 (DUR-4)	9	350708106405801	Duranes #2 (DUR-2)	7
	350605106411801	Duranes #5 (DUR-5)	5	350629106405101	Duranes #3 (DUR-3)	8
	350653106403001	Duranes #6 (DUR-6)	6	350605106411801	Duranes #5 (DUR-5)	6
	350655106395001	Duranes #7 (DUR-7)	4	350653106403001	Duranes #6 (DUR-6)	7
	350635106415001	Gonzales #2 (GON-2)	6	350655106395001	Duranes #7 (DUR-7)	6
	350821106390101	Griegos #4 (GRI-4)	1	350642106422801	Gonzales #1 (GON-1)	3
	350727106340801	Leyendecker #2 (LEY-2)	3	350635106415001	Gonzales #2 (GON-2)	6
	350815106340601	Leyendecker #4 (LEY-4)	1	350827106395001	Griegos #1 (GRI-1)	6
	350422106312401	Lomas #5 (LOM-5)	5	350309106434501	Leavitt #1 (LEA-1)	1
	350538106333001	Love #8 (LOV-8)	2	350819106344001	Leyendecker #3 (LEY-3)	5
	350308106374601	Miles Road #1 (MIL-1)	2	350422106312401	Lomas #5 (LOM-5)	3
	350445106334001	Ridgecrest #4 (RID-4)	3	350800106315001	Ponderosa #2 (PON-2)	1
	350420106334401	Ridgecrest #5 (RID-5)	2	350445106334001	Ridgecrest #4 (RID-4)	1
	350301106383601	San Jose #3 (SAJ-3)	4	350301106383601	San Jose #3 (SAJ-3)	3
	350648106362501	Santa Barbara #1 (SAB-1)	1	350744106333501	Thomas #5 (THO-5)	5
	350744106333501	Thomas #5 (THO-5)	6	350720106330401	Thomas #6 (THO-6)	6
	350720106330401	Thomas #6 (THO-6)	4	350719106333401	Thomas #7 (THO-7)	3
	350719106333401	Thomas #7 (THO-7)	6	350711106323101	Thomas #8 (THO-8)	4
	350732106350101	Vol Andia #2 (VOA-2)	2	350805106354901	Vol Andia #1 (VOA-1)	2
	350803106351101	Vol Andia #4 (VOA-4)	6	350732106350101	Vol Andia #2 (VOA-2)	4
	350809106360901	Vol Andia #5 (VOA-5)	2	350803106351101	Vol Andia #4 (VOA-4)	5
	350828106352101	Vol Andia #6 (VOA-6)	3	350918106315401	Walker #4 (WAL-4)	3
	350950106434001	Volcano Cliffs #1 (VOC-1)	2	351029106332001	Webster #1 (WEB-1)	6
	351029106332001	Webster #1 (WEB-1)	8	351013106333501	Webster #2 (WEB-2)	5
	350358106372901	Yale #2 (YAL-2)	3	350918106425401	Zamora #1 (ZAM-1)	4
	350435106380101	Yale #3 (YAL-3)	1	350852106425701	Zamora #2 (ZAM-2)	3

Each of the 11 GA runs provides a unique, nonsequential solution to the GA search; therefore, sites identified for removal from the GA run removing 10 sites are not required to be removed from the GA run removing 15 sites (Fisher, 2013). The number of times that a site was identified for removal in the GA runs (table 8) provides a measure of the influence of that site on the kriged surface; larger values imply that the site provides little to no beneficial added information (Fisher, 2013). Atrisco no. 3 (site no. 350508106411901) was identified for removal in 10 GA runs for the 2001 and 2015 time intervals. Other sites with relatively large numbers of times identified include West Bluff Nest 1 (site no. 350638106413702, 10 and 6 times for the 2001 time interval and 2015 time interval, respectively); Duranes no. 3 (site no. 350629106405101, 8 times for the 2001 and 2015 time intervals); and Duranes no. 2 and Duranes no. 6 (site nos. 350708106405801 and 350653106403001, 6 and 7 times for the 2001 time interval and 2015 time interval, respectively).

Some sites considered for the spatial assessment (table 1) did not have a groundwater-level measurement in the 2001 time interval and 2015 time interval (“NA” entries in table 2); therefore, sites could be identified for removal in one time interval but not the other time interval solely because the site did not have a groundwater-level measurement in both time intervals. For example, Duranes no. 4 (site no. 350628106411501) was identified for removal in 9 of the 11 GA runs for the 2001 time interval but was not analyzed for the 2015 time interval (table 8).

The spatial distribution of the difference in the kriged surfaces generated using the original and reduced networks for the 2001 time interval and 2015 time interval with 10 and 20 sites removed indicates relatively small differences (magnitudes less than 1 foot) throughout most of the study area, with the largest differences (magnitudes ranging from greater than 1 foot to about 14 feet) surrounding the removed sites (figs. 18*A* and *B* and 19*A* and *B*). For 30 and 40 sites removed, differences with magnitudes greater than 1 foot are widespread with the largest differences (magnitudes larger than 10 feet) extending 1–2 miles away from the removed sites (figs. 18*C* and *D* and 19*C* and *D*). The spatial expansion and increase in the magnitude of differences between 20 and 30 sites removed correspond to the rapid worsening seen in the RMSD and PLE response curves above about 20 sites removed for the 2001 time interval and 2015 time interval (fig. 17).

Sites identified for removal are in areas with high network spatial density (along the Rio Grande within about 1–2 miles north and south of Interstate 40, and the region within about 4 miles east of Interstate 25 and 4 miles north of Interstate 40; figs. 18 and 19). These regions have redundant groundwater-elevation data that can be removed with minimal effect on the kriged surfaces. Areas where sites were not identified for removal include regions of low network spatial density (margins of the study area; figs. 18 and 19) and regions of substantial change in the kriged groundwater elevation (east margin of the study area for the 2001 time interval, fig. 8*A*; northwest corner of the study area for the 2015 time interval, fig. 8*B*).

The spatial assessment aimed to determine the spatial importance of monitoring sites by using a kriging-based GA approach to analyze median groundwater elevations from 87 production zone monitoring sites over two winter periods: the 2001 time interval and 2015 time interval. The spatial assessment results indicate that as many as 20 sites can be removed with a relatively small degradation in the ability to represent the winter kriged surfaces that were generated by using the existing observation network.

Temporal Assessment Results

Results from the temporal assessment can be used to inform minimum monitoring frequency of groundwater levels at the observation network required to capture the trends shown in higher frequency monitoring. Figure 20 and table 9 summarize the median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data over all iterations (hereafter called the fraction within confidence intervals) and the corresponding mean interval between monitoring, in days, tested over all iterations. Reducing the frequency of monitoring is ultimately a management decision that depends on the amount of data loss that is acceptable while still meeting the network objectives of the ABCWUA. Fractions within confidence intervals greater than 0.5 indicate that most of the trend estimates on the reduced dataset were within the confidence interval for the trend on the monthly data.

Results of the temporal assessment indicate that reduced monitoring frequencies could be adopted at many monitoring sites while still adequately characterizing seasonal and long-term trends. For example, if 0.8 (80 percent) fraction within confidence intervals was deemed an acceptable amount of data loss while still meeting network objectives, the results from the temporal assessment indicate that a monitoring interval of 98 days (approximately a quarterly monitoring frequency) is adequate to characterize trends at the shallow Isleta piezometer (site no. 345650106415904; fig. 20*A*). Using the same fraction within confidence intervals criteria, the results from the temporal assessment indicate a monitoring interval of 63 days (approximately a bimonthly monitoring frequency) is adequate to characterize trends at the shallow Montessa Park piezometer (site no. 350056106370103; fig. 20*B*).

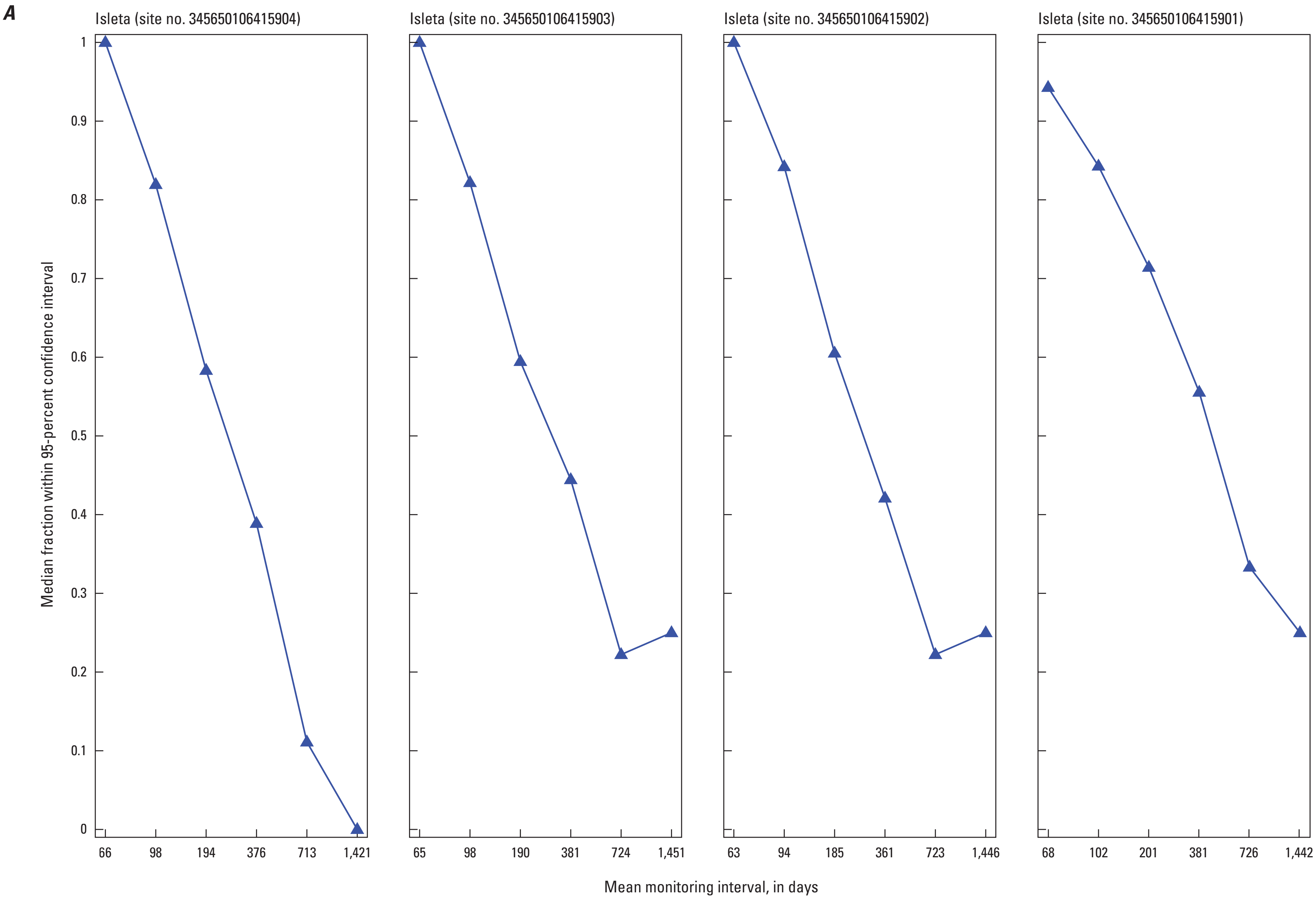


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.

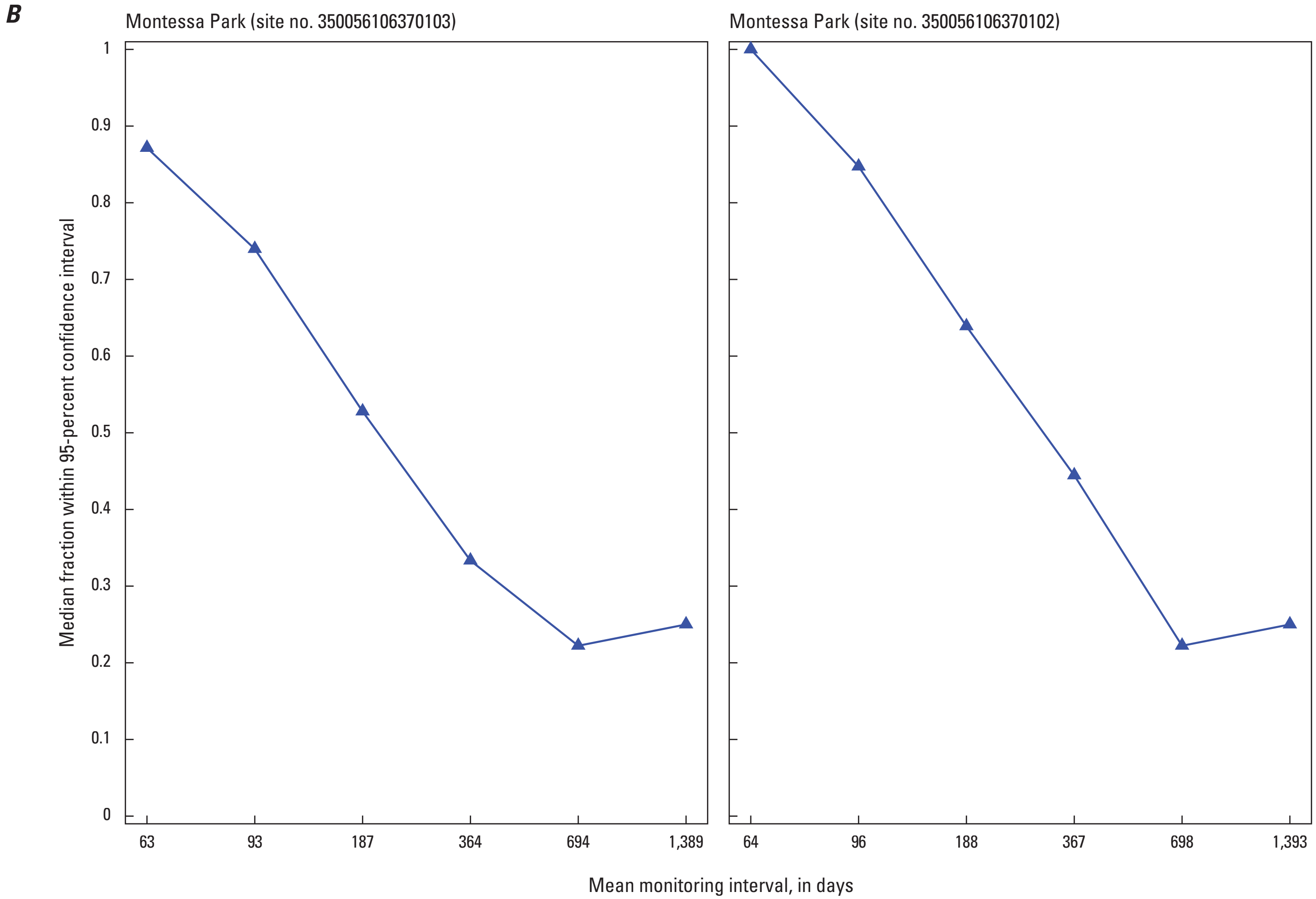


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

C

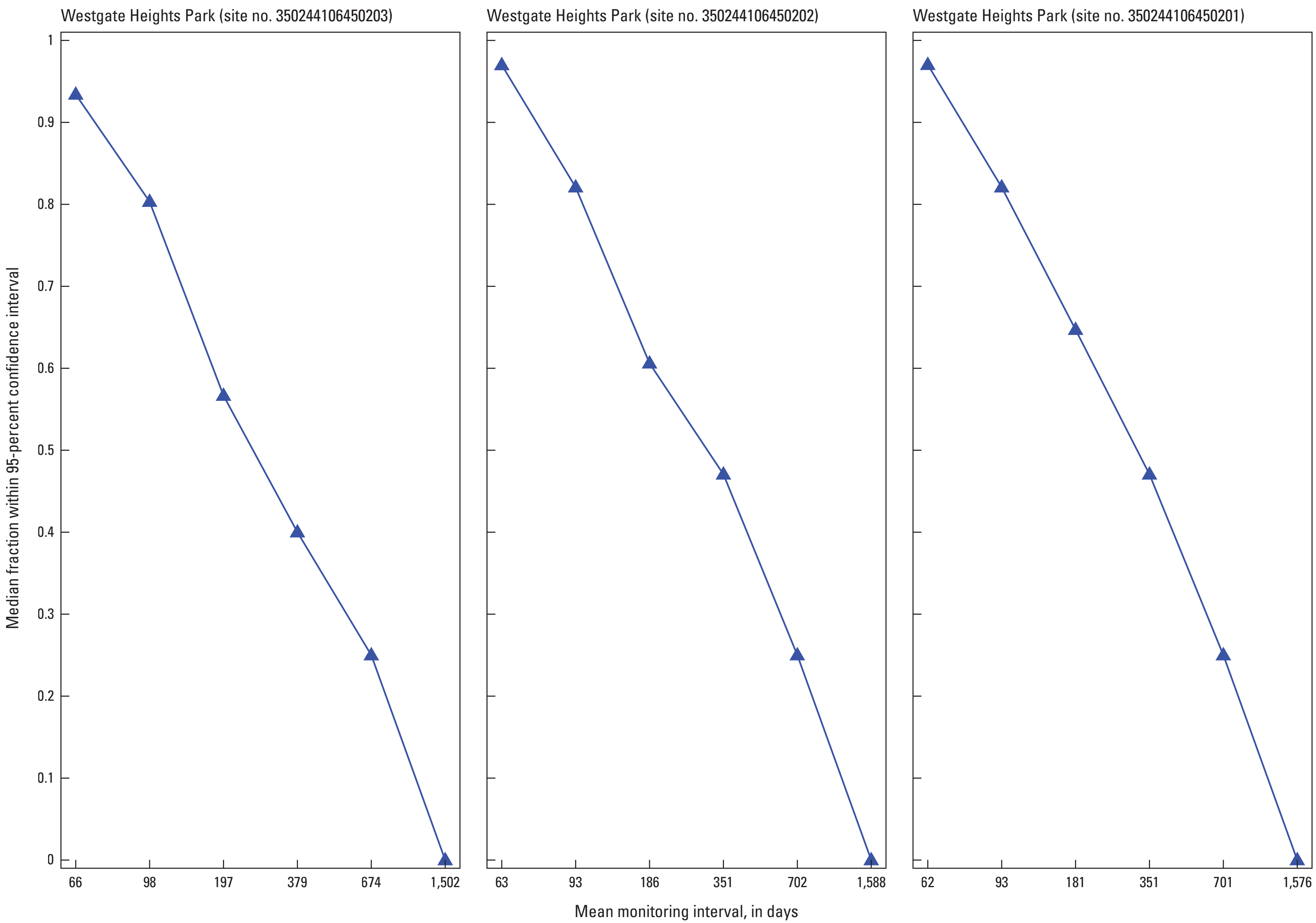


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

D

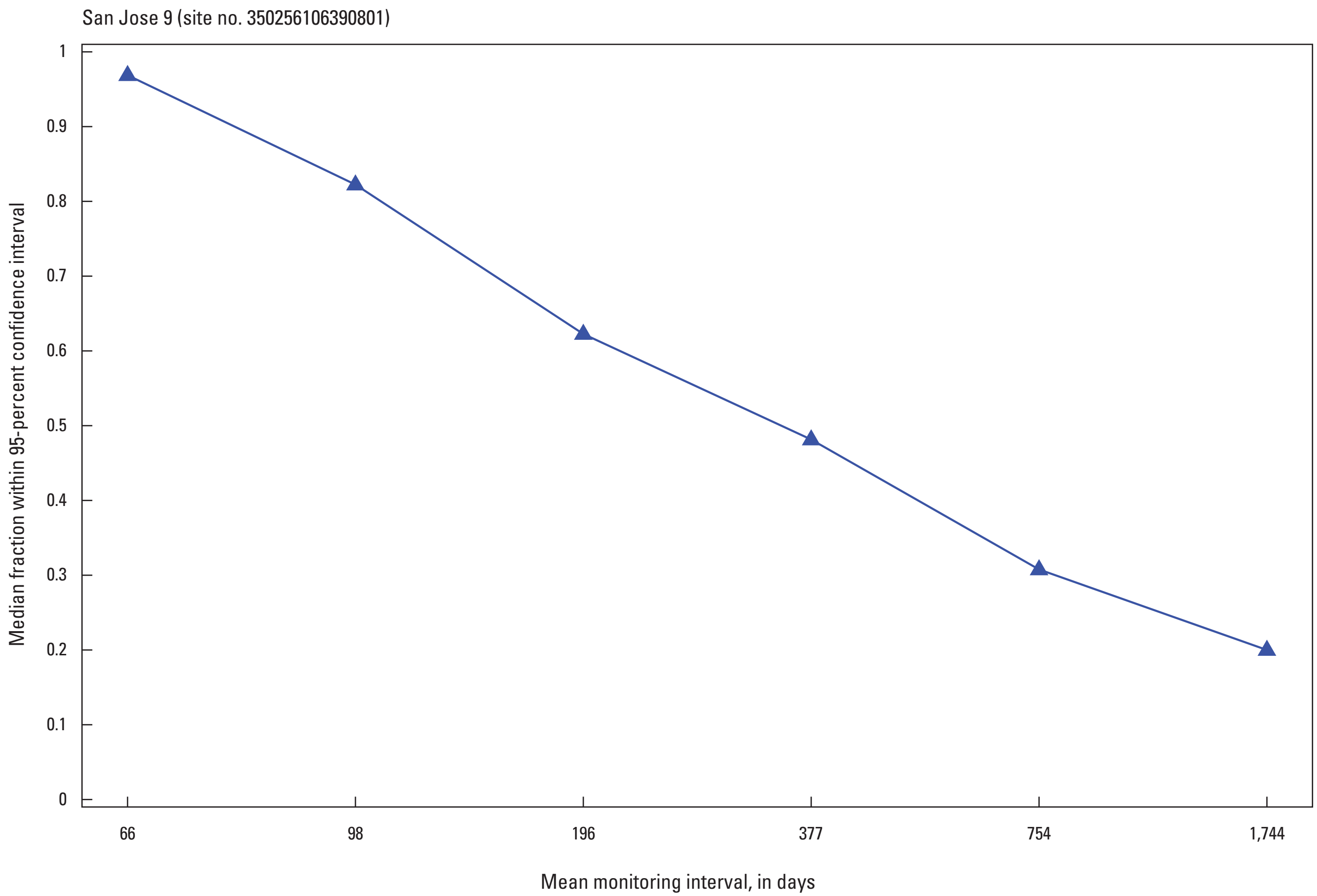


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

E

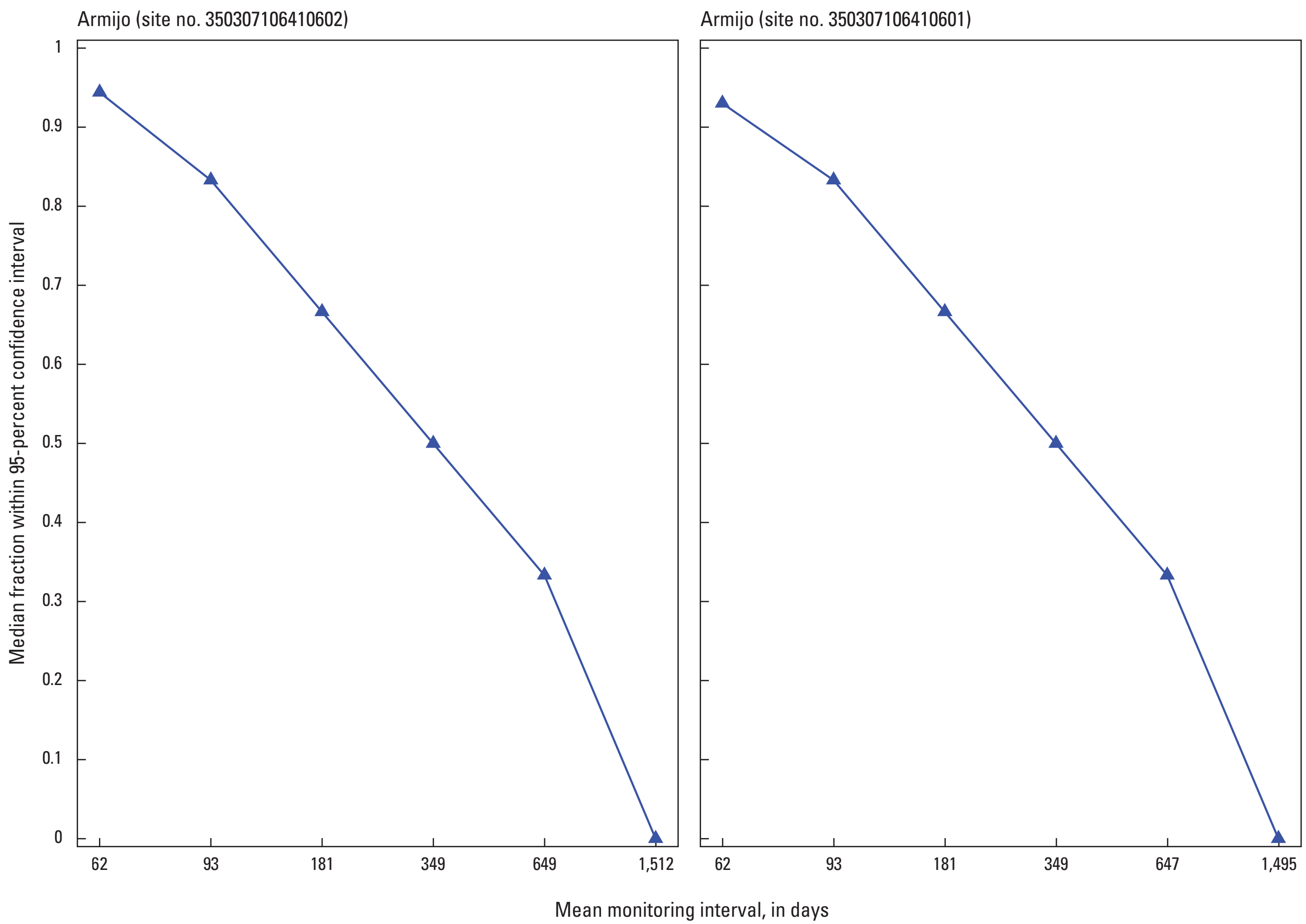


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

F

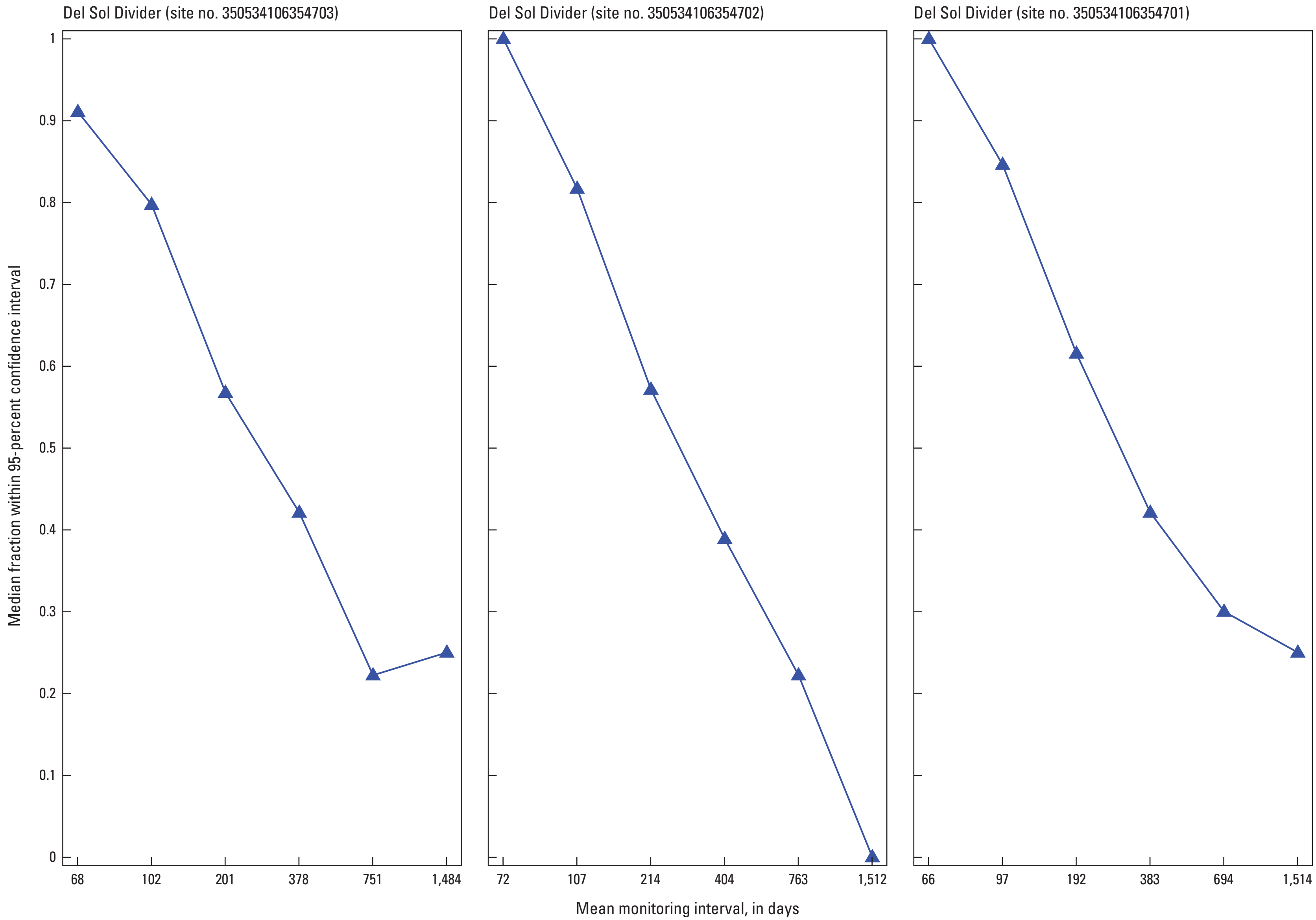


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

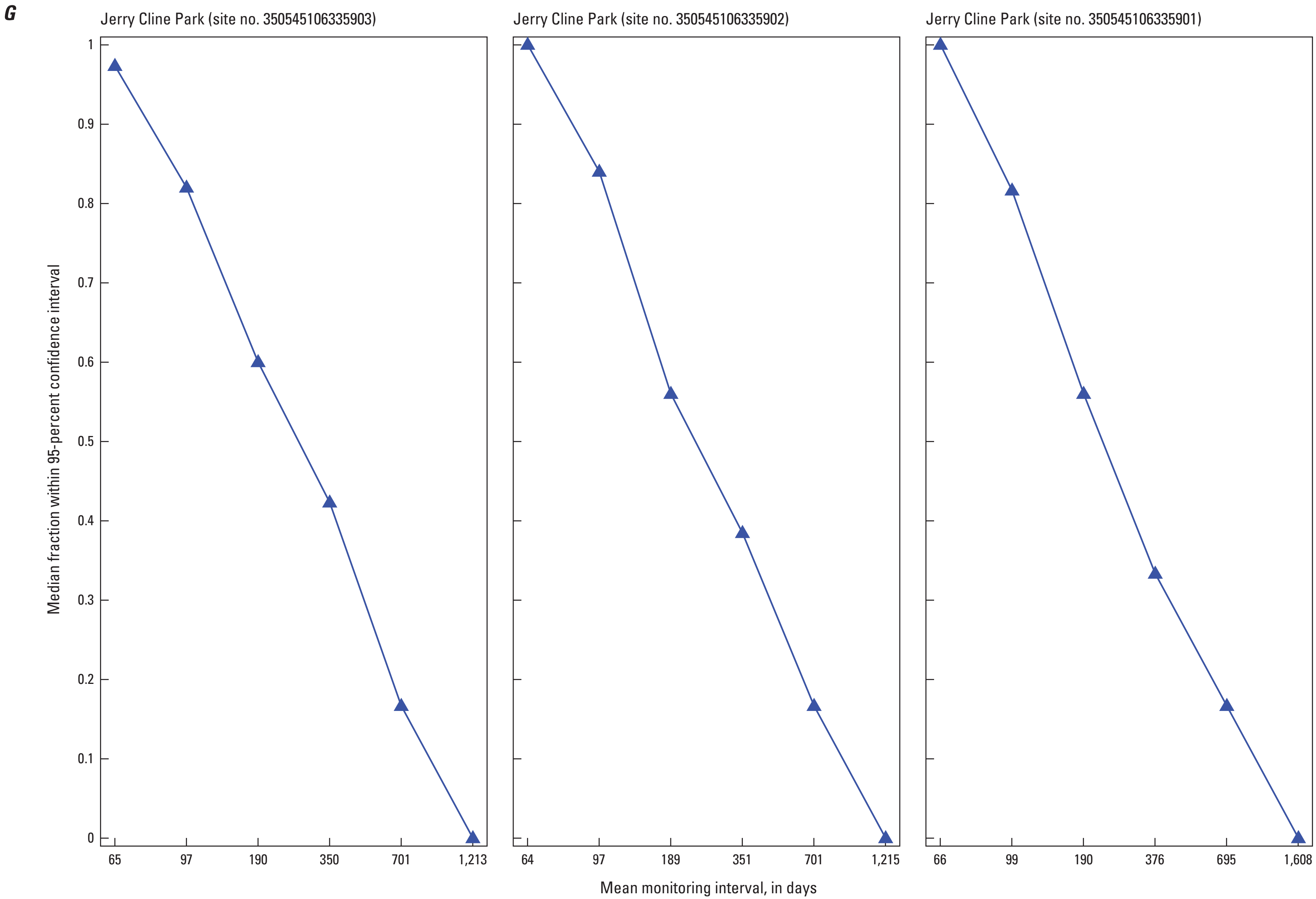


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

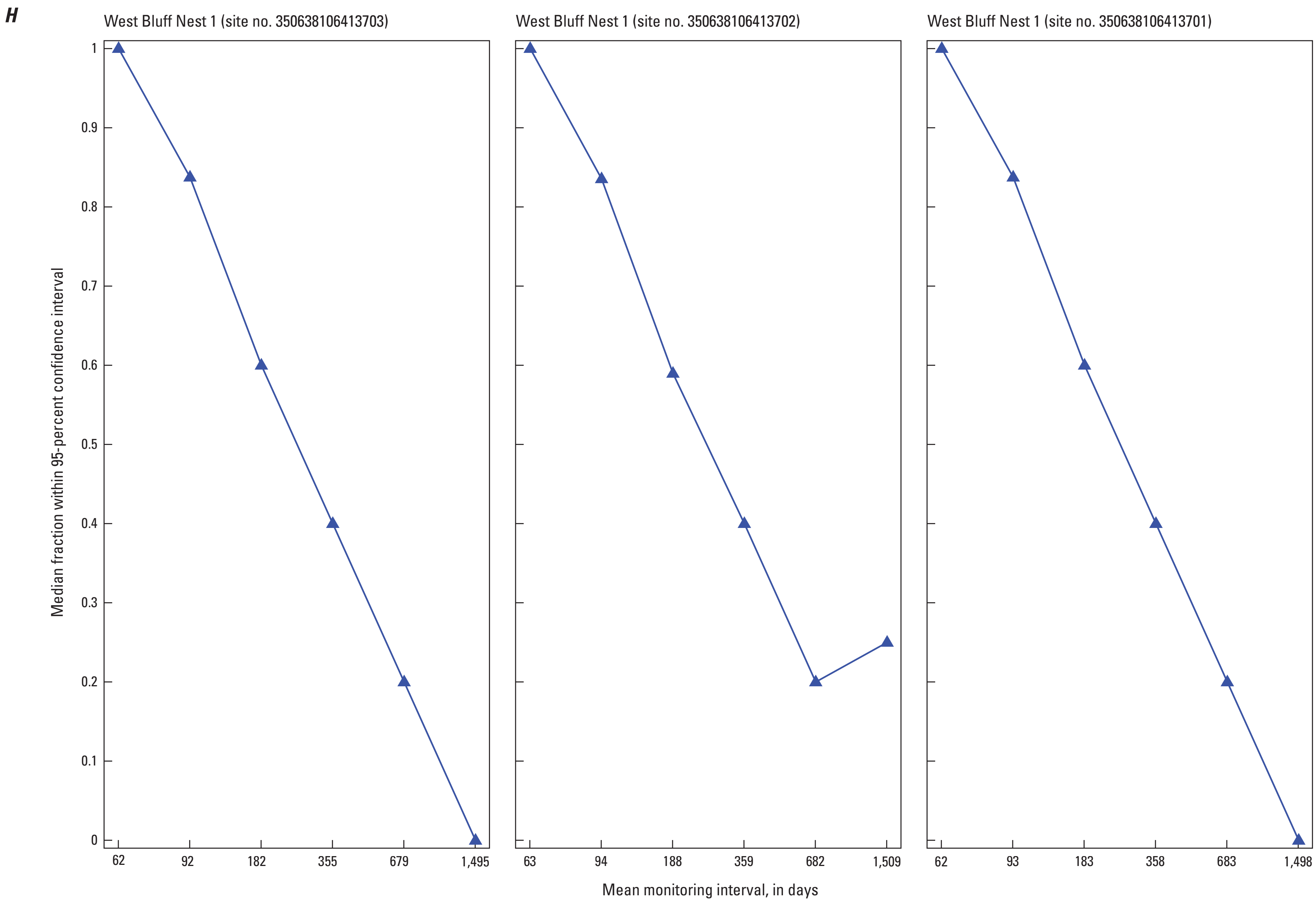


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

I

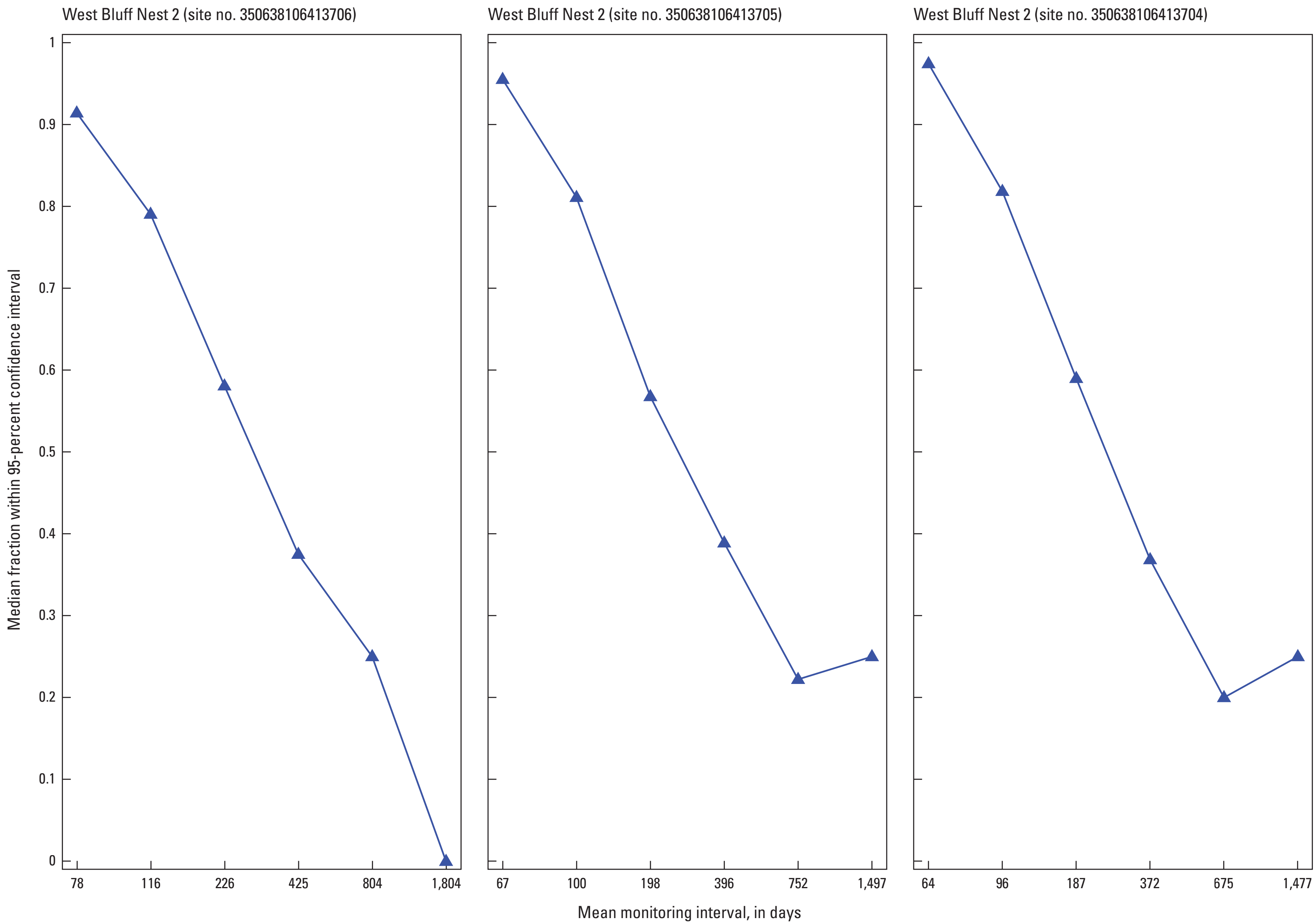


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

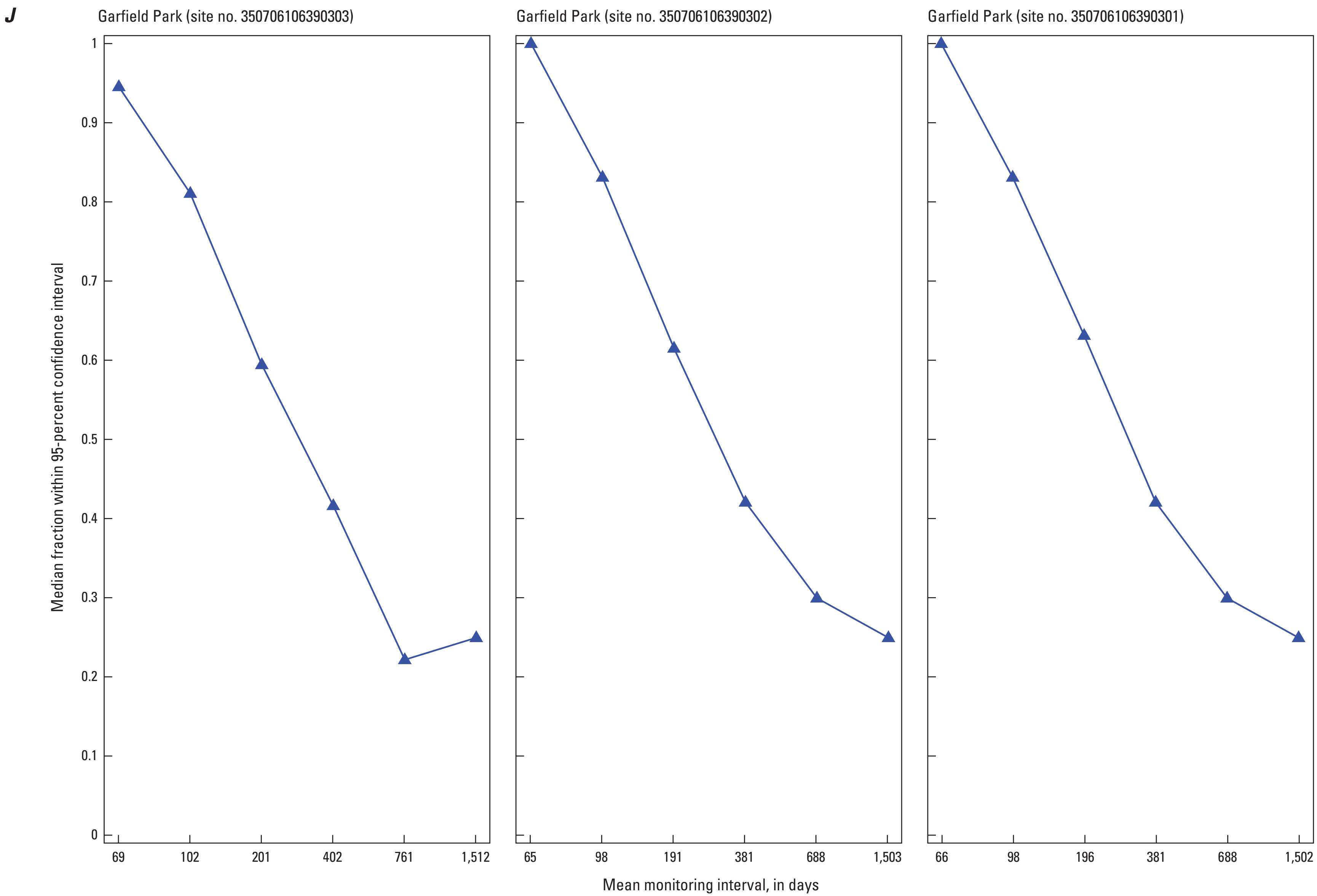


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

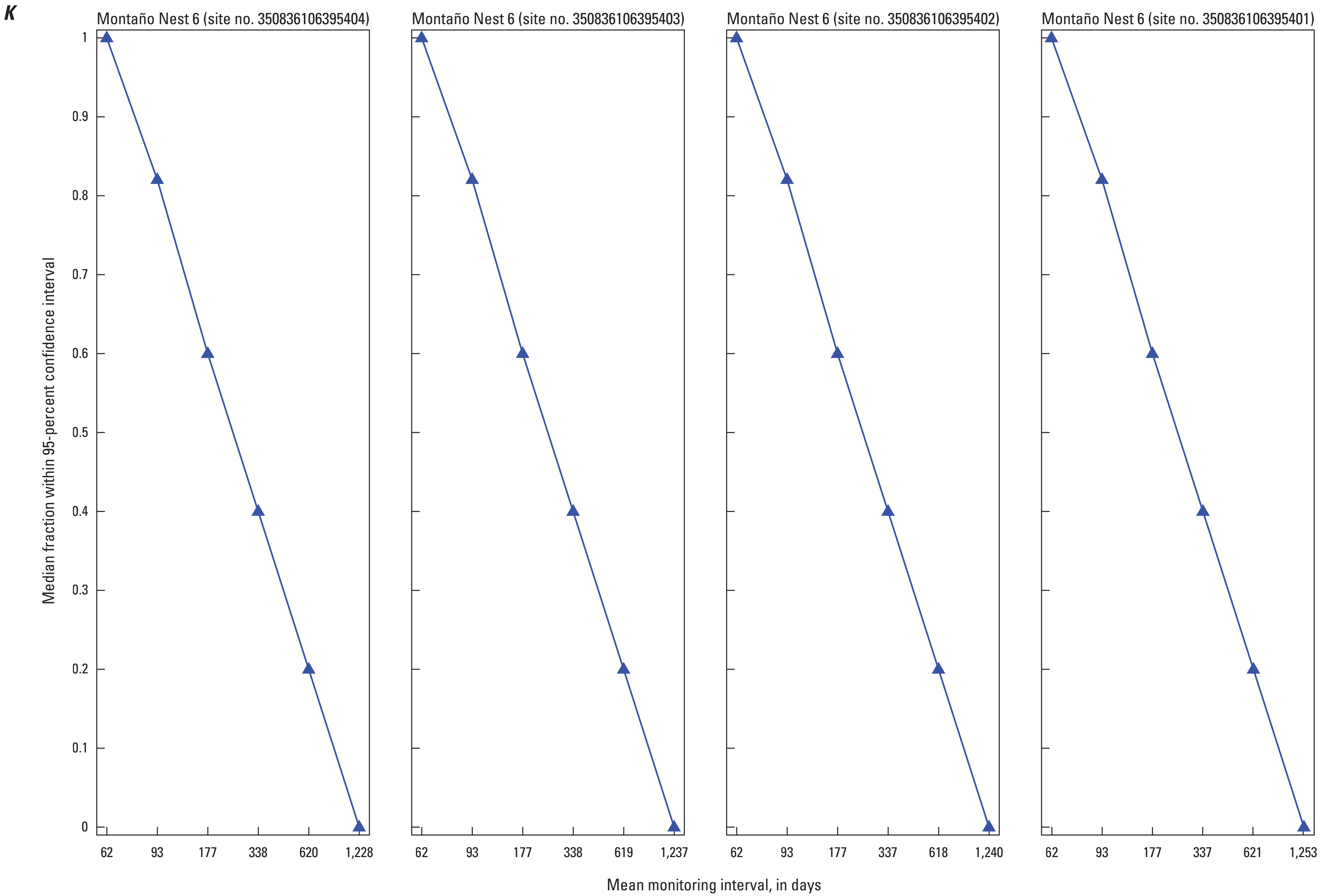


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F,* Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

L

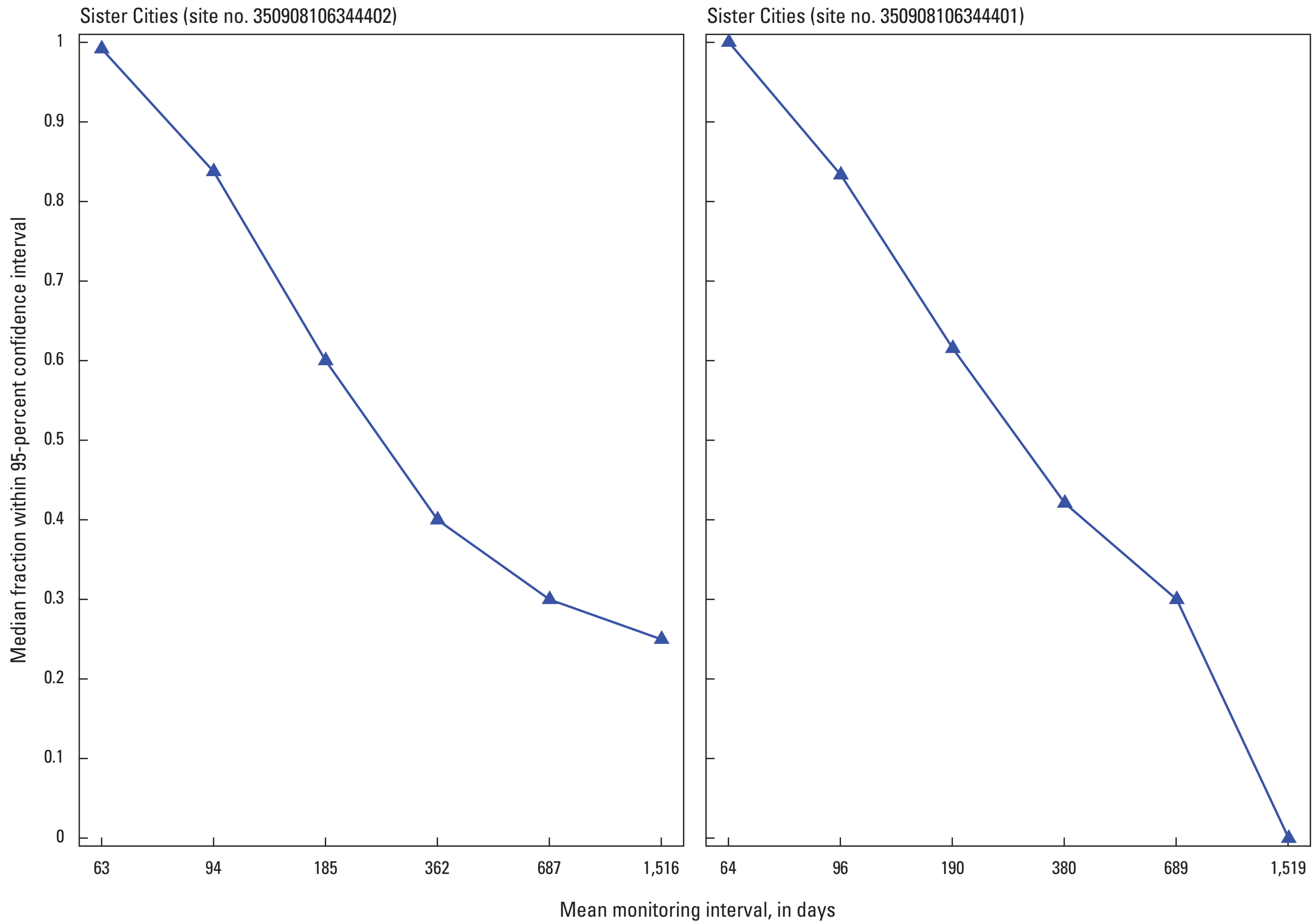


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

M

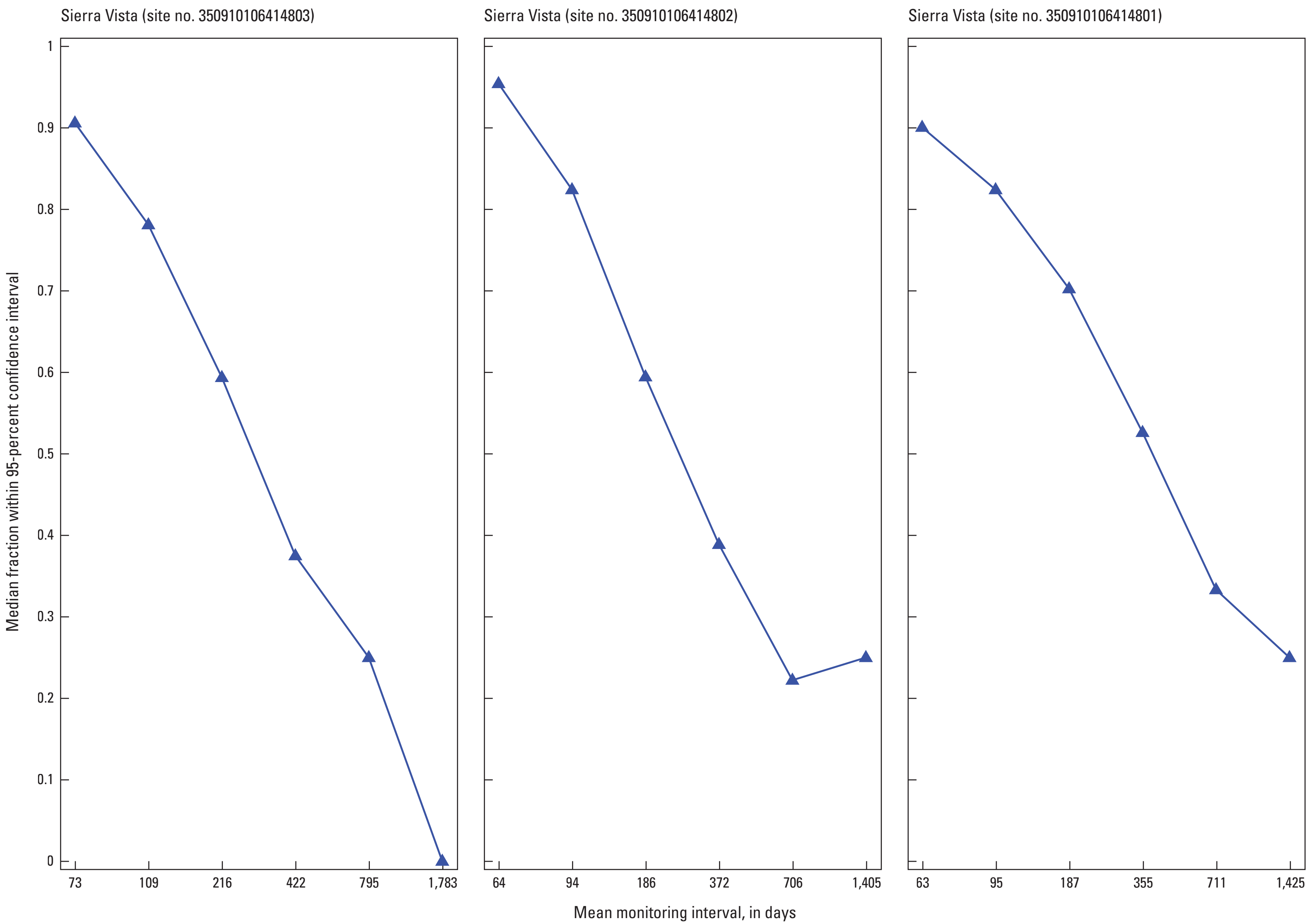


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

N

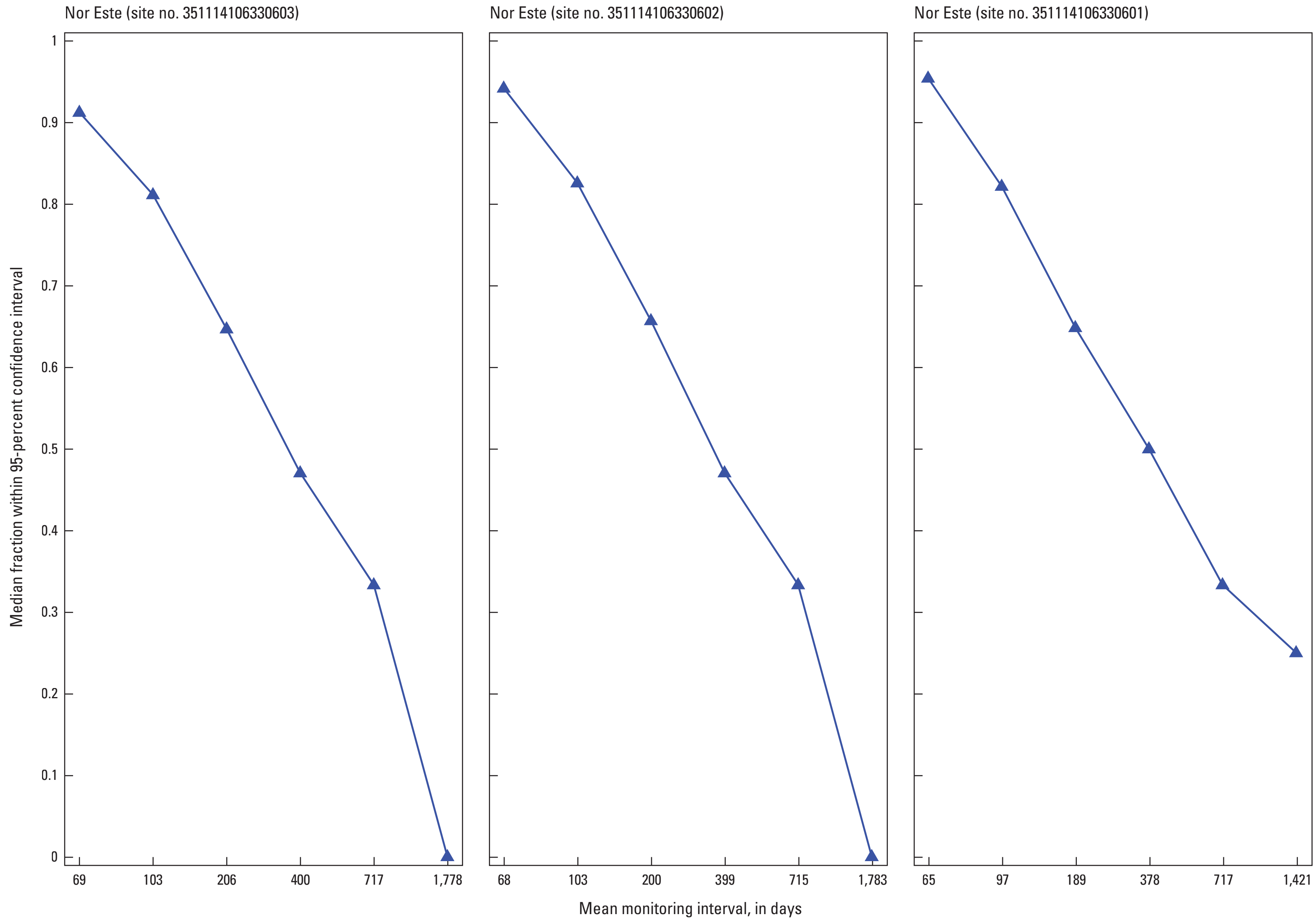


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

O

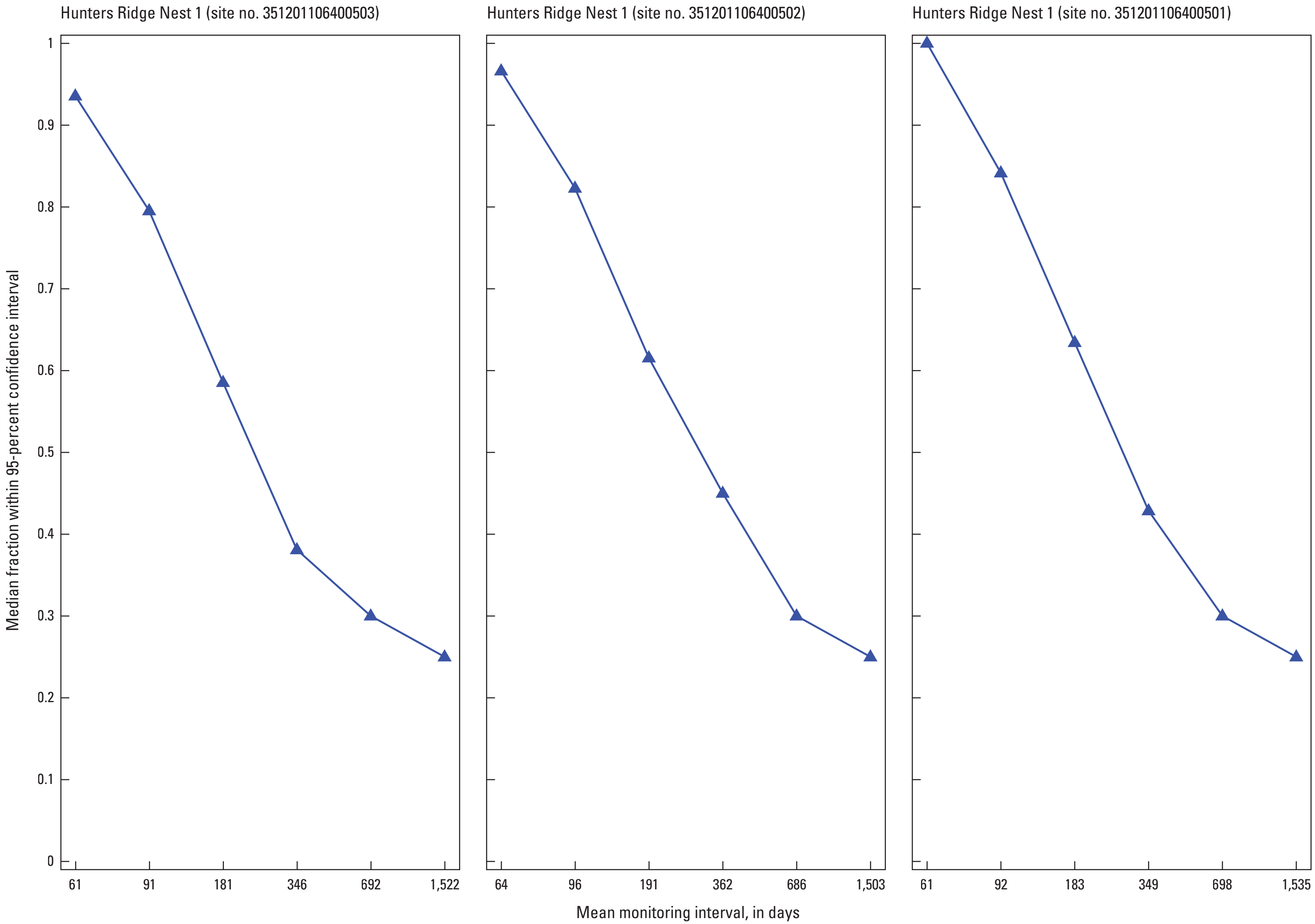


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

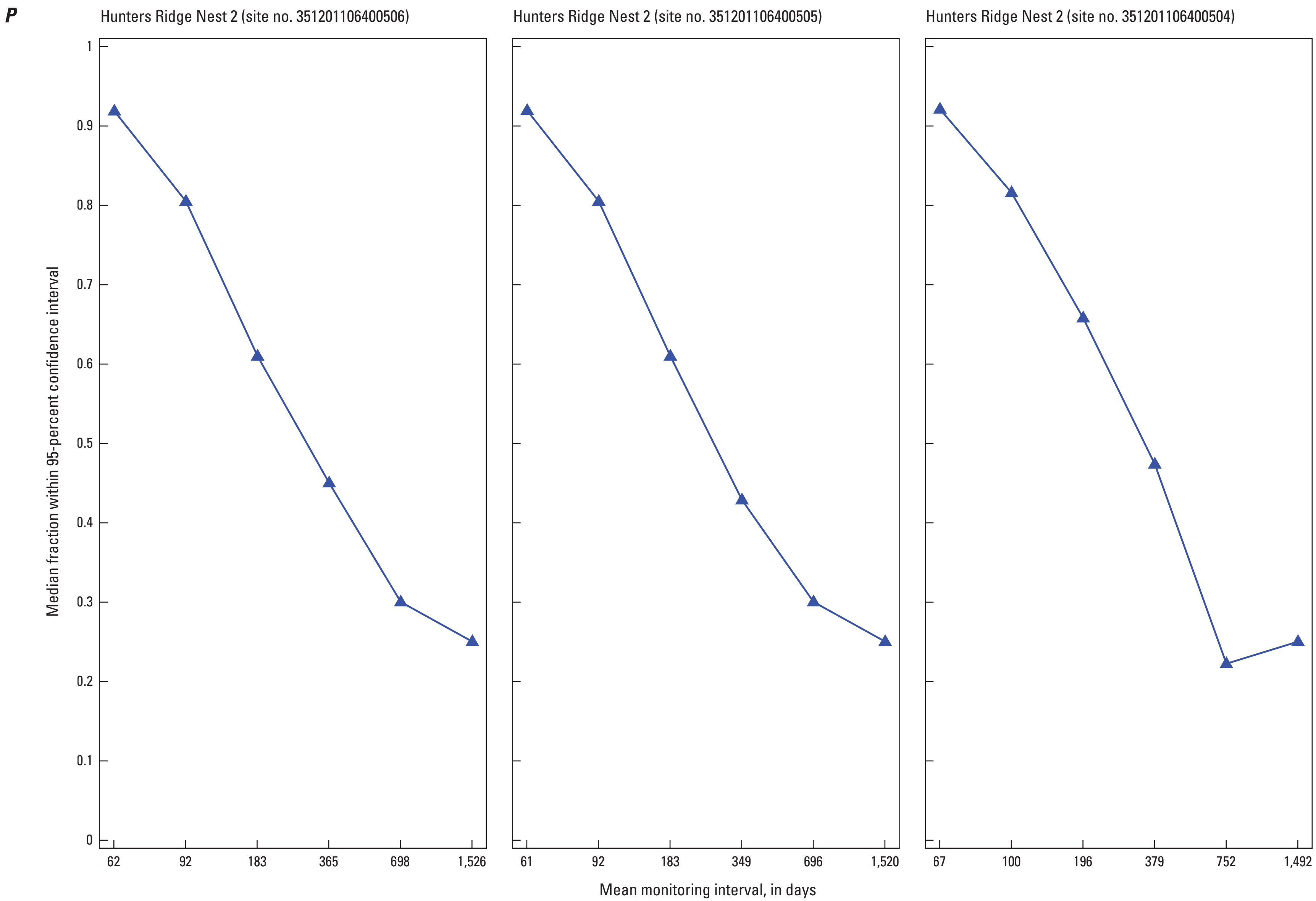


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

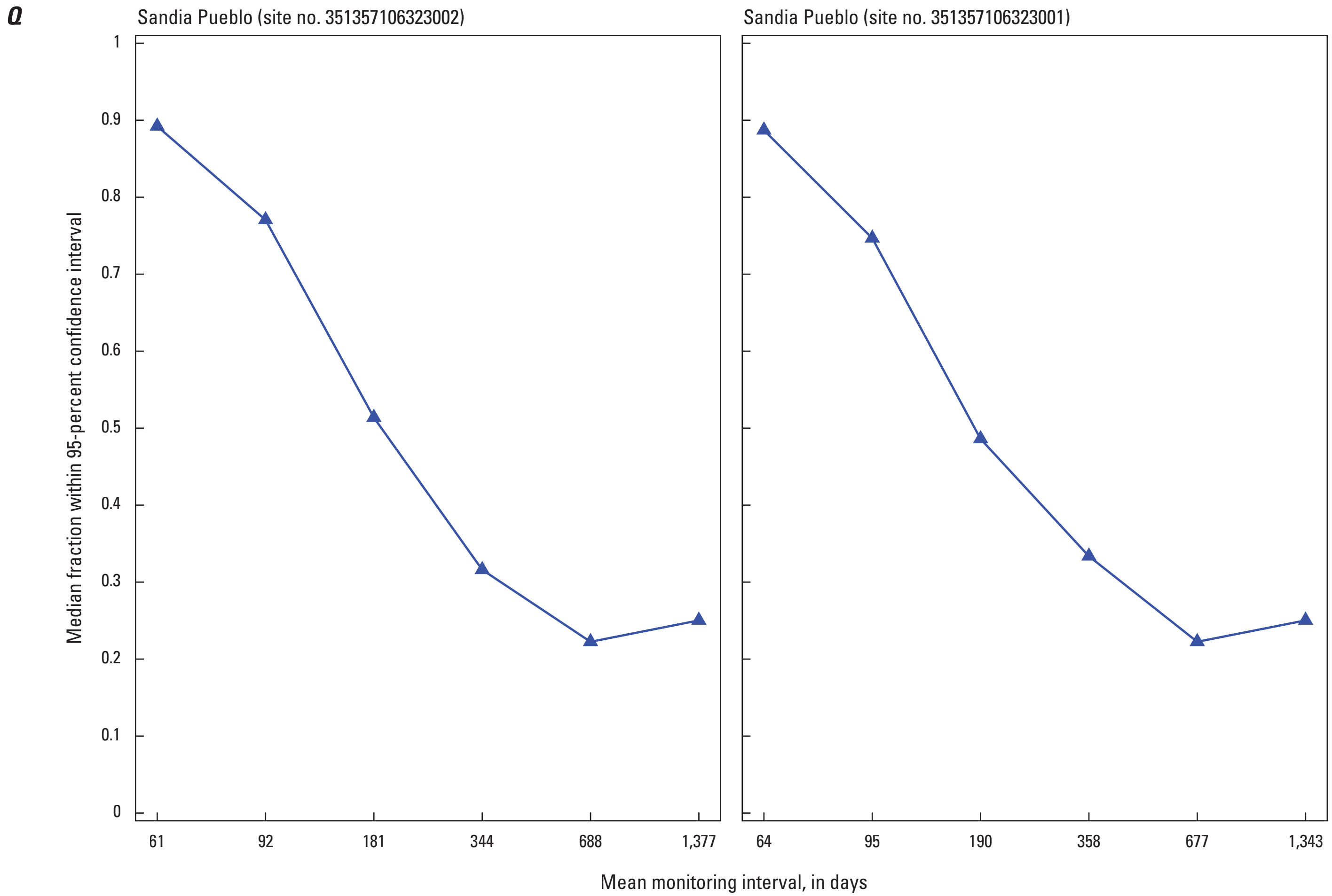


Figure 20. Response curves of median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data versus mean interval between monitoring, in days, tested for each piezometer within the nest or monitoring well at the *A*, Isleta; *B*, Montessa Park; *C*, Westgate Heights Park; *D*, San Jose 9; *E*, Armijo; *F*, Del Sol Divider; *G*, Jerry Cline Park; *H*, West Bluff Nest 1; *I*, West Bluff Nest 2; *J*, Garfield Park; *K*, Montaño Nest 6; *L*, Sister Cities; *M*, Sierra Vista; *N*, Nor Este; *O*, Hunters Ridge Nest 1; *P*, Hunters Ridge Nest 2; and *Q*, Sandia Pueblo sites.—Continued

Table 9. Summary of mean interval between monitoring, in days, and median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data at each site.

[USGS, U.S. Geological Survey. “Monitoring interval” indicates the mean over all iterations of the mean interval between monitoring in the reduced data for each iteration, in days. “Fraction within 95% CI” indicates the median over all iterations of the fraction of estimates from local regression on reduced data for each iteration within a 95-percent confidence interval on estimates from local regression on daily data]

Site name	USGS site number	Statistic	Temporal optimization summary					
Isleta	345650106415904	Monitoring interval	66	98	194	376	713	1,421
		Fraction within 95% CI	1	0.82	0.58	0.39	0.11	0
	345650106415903	Monitoring interval	65	98	190	381	724	1,451
		Fraction within 95% CI	1	0.82	0.59	0.44	0.22	0.25
	345650106415902	Monitoring interval	63	94	185	361	723	1,446
		Fraction within 95% CI	1	0.84	0.61	0.42	0.22	0.25
	345650106415901	Monitoring interval	68	102	201	381	726	1,442
		Fraction within 95% CI	0.94	0.84	0.71	0.56	0.33	0.25
Montessa Park	350056106370103	Monitoring interval	63	93	187	364	694	1,389
		Fraction within 95% CI	0.87	0.74	0.53	0.33	0.22	0.25
	350056106370102	Monitoring interval	64	96	188	367	698	1,393
		Fraction within 95% CI	1	0.85	0.64	0.44	0.22	0.25
Westgate Heights Park	350244106450203	Monitoring interval	66	98	197	379	674	1,502
		Fraction within 95% CI	0.93	0.8	0.57	0.4	0.25	0
	350244106450202	Monitoring interval	63	93	186	351	702	1,588
		Fraction within 95% CI	0.97	0.82	0.61	0.47	0.25	0
	350244106450201	Monitoring interval	62	93	181	351	701	1,576
		Fraction within 95% CI	0.97	0.82	0.65	0.47	0.25	0
San Jose 9	350256106390801	Monitoring interval	66	98	196	377	754	1,744
		Fraction within 95% CI	0.97	0.82	0.62	0.48	0.31	0.2
Armijo	350307106410602	Monitoring interval	62	93	181	349	649	1,512
		Fraction within 95% CI	0.94	0.83	0.67	0.5	0.33	0
	350307106410601	Monitoring interval	62	93	181	349	647	1,495
		Fraction within 95% CI	0.93	0.83	0.67	0.5	0.33	0
Del Sol Divider	350534106354703	Monitoring interval	68	102	201	378	751	1,484
		Fraction within 95% CI	0.91	0.8	0.57	0.42	0.22	0.25
	350534106354702	Monitoring interval	72	107	214	404	763	1,512
		Fraction within 95% CI	1	0.82	0.57	0.39	0.22	0
	350534106354701	Monitoring interval	66	97	192	383	694	1,514
		Fraction within 95% CI	1	0.85	0.62	0.42	0.3	0.25
Jerry Cline Park	350545106335903	Monitoring interval	65	97	190	350	701	1,213
		Fraction within 95% CI	0.97	0.82	0.6	0.42	0.17	0
	350545106335902	Monitoring interval	64	97	189	351	701	1,215
		Fraction within 95% CI	1	0.84	0.56	0.38	0.17	0
	350545106335901	Monitoring interval	66	99	190	376	695	1,608
		Fraction within 95% CI	1	0.82	0.56	0.33	0.17	0
West Bluff Nest 1	350638106413703	Monitoring interval	62	92	182	355	679	1,495
		Fraction within 95% CI	1	0.84	0.6	0.4	0.2	0
	350638106413702	Monitoring interval	63	94	188	359	682	1,509
		Fraction within 95% CI	1	0.84	0.59	0.4	0.2	0.25
	350638106413701	Monitoring interval	62	93	183	358	683	1,498
		Fraction within 95% CI	1	0.84	0.6	0.4	0.2	0

Table 9. Summary of mean interval between monitoring, in days, and median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data at each site.—Continued

[USGS, U.S. Geological Survey. "Monitoring interval" indicates the mean over all iterations of the mean interval between monitoring in the reduced data for each iteration, in days. "Fraction within 95% CI" indicates the median over all iterations of the fraction of estimates from local regression on reduced data for each iteration within a 95-percent confidence interval on estimates from local regression on daily data]

Site name	USGS site number	Statistic	Temporal optimization summary					
West Bluff Nest 2	350638106413706	Monitoring interval	78	116	226	425	804	1,804
		Fraction within 95% CI	0.91	0.79	0.58	0.38	0.25	0
	350638106413705	Monitoring interval	67	100	198	396	752	1,497
		Fraction within 95% CI	0.95	0.81	0.57	0.39	0.22	0.25
	350638106413704	Monitoring interval	64	96	187	372	675	1,477
		Fraction within 95% CI	0.97	0.82	0.59	0.37	0.2	0.25
Garfield Park	350706106390303	Monitoring interval	69	102	201	402	761	1,512
		Fraction within 95% CI	0.95	0.81	0.59	0.42	0.22	0.25
	350706106390302	Monitoring interval	65	98	191	381	688	1,503
		Fraction within 95% CI	1	0.83	0.62	0.42	0.3	0.25
	350706106390301	Monitoring interval	66	98	196	381	688	1,502
		Fraction within 95% CI	1	0.83	0.63	0.42	0.3	0.25
Montaño Nest 6	350836106395404	Monitoring interval	62	93	177	338	620	1,228
		Fraction within 95% CI	1	0.82	0.6	0.4	0.2	0
	350836106395403	Monitoring interval	62	93	177	338	619	1,237
		Fraction within 95% CI	1	0.82	0.6	0.4	0.2	0
	350836106395402	Monitoring interval	62	93	177	337	618	1,240
		Fraction within 95% CI	1	0.82	0.6	0.4	0.2	0
	350836106395401	Monitoring interval	62	93	177	337	621	1,253
		Fraction within 95% CI	1	0.82	0.6	0.4	0.2	0
Sister Cities	350908106344402	Monitoring interval	63	94	185	362	687	1,516
		Fraction within 95% CI	0.99	0.84	0.6	0.4	0.3	0.25
	350908106344401	Monitoring interval	64	96	190	380	689	1,519
		Fraction within 95% CI	1	0.83	0.62	0.42	0.3	0
Sierra Vista	350910106414803	Monitoring interval	73	109	216	422	795	1,783
		Fraction within 95% CI	0.91	0.78	0.59	0.38	0.25	0
	350910106414802	Monitoring interval	64	94	186	372	706	1,405
		Fraction within 95% CI	0.95	0.82	0.59	0.39	0.22	0.25
	350910106414801	Monitoring interval	63	95	187	355	711	1,425
		Fraction within 95% CI	0.9	0.82	0.7	0.53	0.33	0.25
Nor Este	351114106330603	Monitoring interval	69	103	206	400	717	1,778
		Fraction within 95% CI	0.91	0.81	0.65	0.47	0.33	0
	351114106330602	Monitoring interval	68	103	200	399	715	1,783
		Fraction within 95% CI	0.94	0.83	0.66	0.47	0.33	0
	351114106330601	Monitoring interval	65	97	189	378	717	1,421
		Fraction within 95% CI	0.95	0.82	0.65	0.5	0.33	0.25
Hunters Ridge Nest 1	351201106400503	Monitoring interval	61	91	181	346	692	1,522
		Fraction within 95% CI	0.94	0.8	0.59	0.38	0.3	0.25
	351201106400502	Monitoring interval	64	96	191	362	686	1,503
		Fraction within 95% CI	0.97	0.82	0.62	0.45	0.3	0.25
	351201106400501	Monitoring interval	61	92	183	349	698	1,535
		Fraction within 95% CI	1	0.84	0.63	0.43	0.3	0.25

Table 9. Summary of mean interval between monitoring, in days, and median fraction of estimates from the local regression fit on the reduced dataset that are within the 95-percent confidence interval for the regression fit on the monthly data at each site.—Continued

[USGS, U.S. Geological Survey. "Monitoring interval" indicates the mean over all iterations of the mean interval between monitoring in the reduced data for each iteration, in days. "Fraction within 95% CI" indicates the median over all iterations of the fraction of estimates from local regression on reduced data for each iteration within a 95-percent confidence interval on estimates from local regression on daily data]

Site name	USGS site number	Statistic	Temporal optimization summary					
Hunters Ridge Nest 2	351201106400506	Monitoring interval	62	92	183	365	698	1,526
		Fraction within 95% CI	0.92	0.8	0.61	0.45	0.3	0.25
	351201106400505	Monitoring interval	61	92	183	349	696	1,520
		Fraction within 95% CI	0.92	0.8	0.61	0.43	0.3	0.25
	351201106400504	Monitoring interval	67	100	196	379	752	1,492
		Fraction within 95% CI	0.92	0.82	0.66	0.47	0.22	0.25
Sandia Pueblo	351357106323002	Monitoring interval	61	92	181	344	688	1,377
		Fraction within 95% CI	0.89	0.77	0.51	0.32	0.22	0.25
	351357106323001	Monitoring interval	64	95	190	358	677	1,343
		Fraction within 95% CI	0.89	0.75	0.49	0.33	0.22	0.25

Principal Component Analysis Results

The PCA results (figs. 21–24) can be used to identify regions of the observation network that have similar hydrographs. This information can inform the management decisions that are made when using the spatial and temporal assessment results. For example, water managers may decide to continue monitoring a subset of sites at a high sampling frequency within each of the regions that have similar hydrographs, where the subset of sites would serve as index monitoring sites for those regions. Furthermore, water managers may opt to maintain a minimum number of, or add more, discrete sampling sites within each of the identified regions to promote accurately describing the groundwater-elevation surface.

The grouped distribution of monitoring sites and corresponding silhouette width for the final cluster solution are shown in figure 21. Clusters 1 and 3 are the smallest clusters with only two hydrographs per cluster, whereas clusters 2 and 4 are larger and consist of five and six hydrographs, respectively. The silhouette width metric ranges from −1 to 1 and provides a means to evaluate how well each hydrograph has been clustered (Everitt and others, 2011; Maechler and others, 2019). Values closer to 1 indicate that an object is well matched to its own cluster and is poorly matched to all other clusters. Alternatively, values near –1 imply that an object would be better clustered in another group. Values near zero show that it is unclear whether the object should be assigned to its current cluster or to a neighboring cluster. Clustering solutions with an MSW above 0.5 are typically considered to be reasonable classifications, whereas solutions with means below 0.2 should be interpreted as being poor classifications (Everitt and others, 2011). The overall clustering solution had an MSW of 0.48 (approximately 0.50), indicating that the clustering solution is well supported by the data (fig. 21). The individual clusters had MSWs as follows: cluster 1 was −0.02, cluster 2 was 0.54, cluster 3 was 0.52, and cluster 4 was 0.58. These values indicate that clusters 2 through 4 are well classified but that cluster 1 is a weakly supported cluster. The MSW metric was initially used to aid in determining the number of appropriate clusters to seek in the k-means clustering algorithm and implied that a five-cluster solution is also supported by the data (fig. 16*B*). Upon testing a five-cluster solution with otherwise identical implementation, the clustering algorithm split the hydrographs from cluster 1 into their own single-hydrograph clusters to form the fifth cluster. This result, along with the low silhouette widths of cluster 1 (fig. 21), indicates that these two hydrographs are unique and that cluster 1 would not be well represented by fewer monitoring sites. The four-cluster solution was still utilized for the analysis, because these two hydrographs have very similar curvature, the MSW for the four and five cluster solutions are nearly identical, and the WSS metric better supports the four-cluster solution.

Figure 22 presents PC loadings plots of the final clustered solution. The loadings for each PC are related to the correlation between hydrographs; sites that consistently group together on these plots possess similar characteristics. Previous researchers who used a similar approach on another network derived their groupings by visually inspecting a plot of PC1 versus PC2 loadings (Winter and others, 2000). The overlap between clusters 2 and 4 in this plot (fig. 22*A*) indicates that the additional information provided by including PCs 3 and 4 in the analysis is important for defining clusters. The influence of this additional information is further evidenced by the separation of these clusters in figures 22*C*, *E*, and *F*. The ability to account for these additional PCs is a clear advantage of using quantitative clustering methods to find data-supported clusters rather than using qualitative approaches. The overlap of clusters 2 and 4 in figures 22*A*, *B*, and *D* indicates that the hydrographs in these clusters share characteristics, but the separation of these clusters in figures 22*C*, *E*, and *F* makes them different enough to form their own clusters. The monitoring sites within clusters 1 and 3 are consistently separated from most other data in figure 22. This is evidence that these two clusters are distinctly different from the other two clusters.

Figure 23 shows the hydrographs sorted into their clustered groups, and figure 24 displays the spatial distribution of the corresponding monitoring sites. Plotting all hydrographs on the same fixed y-axis scale and examining their spatial distribution allows cluster characteristics to become more recognizable (figs. 23*A* and 24). The upward trend in groundwater elevations observed in clusters 2 and 4 (fig. 23) is consistent with groundwater-level recovery throughout the Albuquerque area and has been linked to conservation efforts and the implementation of surface water as a drinking water source (Driscoll and Brandt, 2017; Galanter and Curry, 2019; Ritchie and others, 2019). This similarity is likely why clusters 2 and 4 overlap in figures 22*A*, *B*, and *D*. Despite the similar long-term increase in groundwater elevations of clusters 2 and 4, cluster 2 has a dampened seasonal signal relative to cluster 4 (fig. 23*A*). This difference may be a product of well screen depth variation between the monitoring sites in each cluster, as the median midpoint depth of the screened interval for the monitoring sites in cluster 2 is 192 feet deeper (median = 1,005 feet) than that of cluster 4 (median = 813 feet). This pattern is consistent with evidence presented earlier that short-term surface processes are attenuated at depth.

The hydrographs in cluster 1, despite having the aforementioned low silhouette width, both share a pronounced seasonal signal and an approximately stationary mean groundwater elevation over the period of record. The monitoring sites corresponding to this cluster are separated by a large distance of 17.5 miles and are both on the margins of the study area (fig. 24). Therefore, it is not surprising that these hydrographs are different enough to produce a low silhouette width (fig. 21) while also not displaying a prominent groundwater recovery signal.

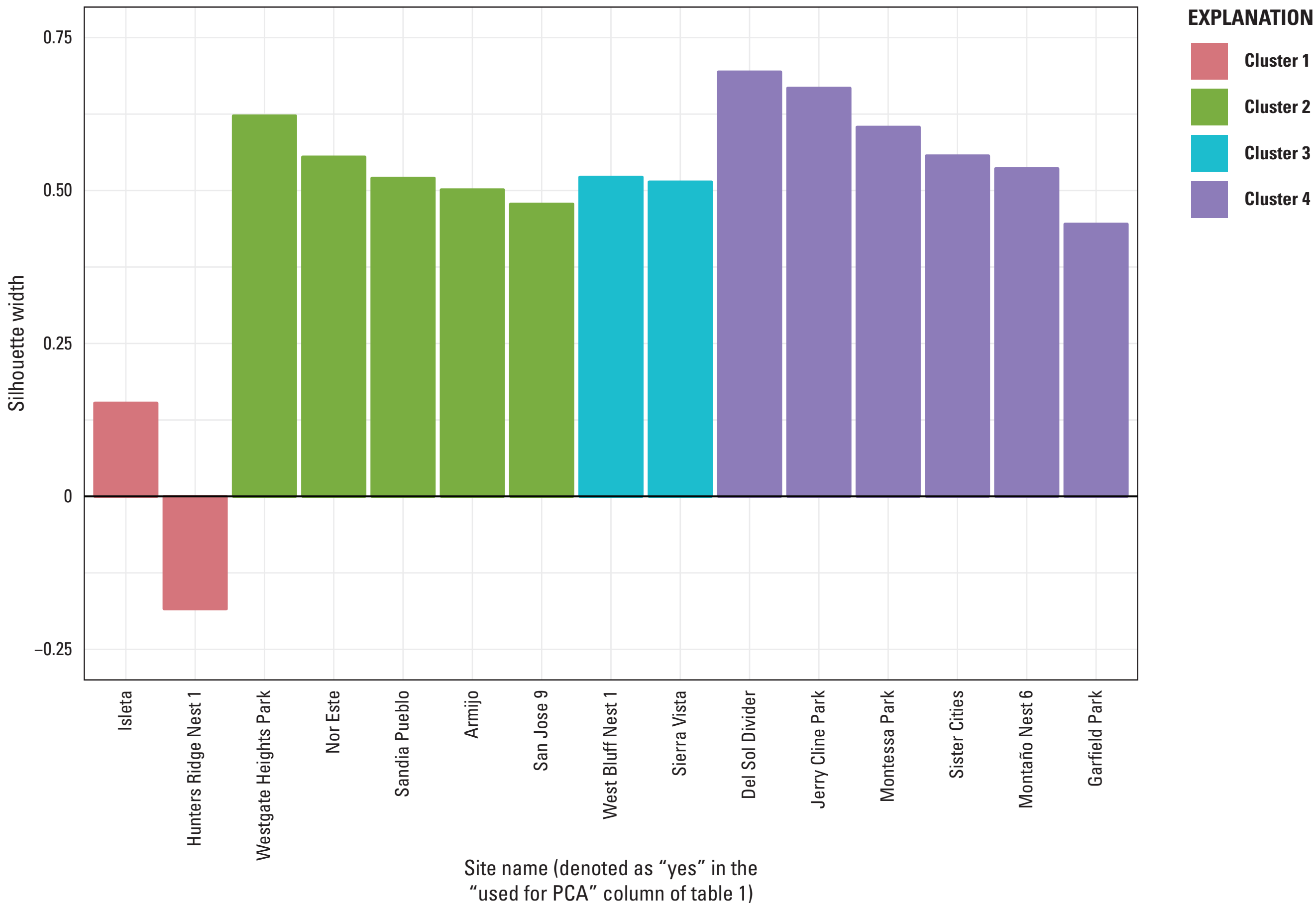


Figure 21. Calculated silhouette width for each clustered monitoring site.

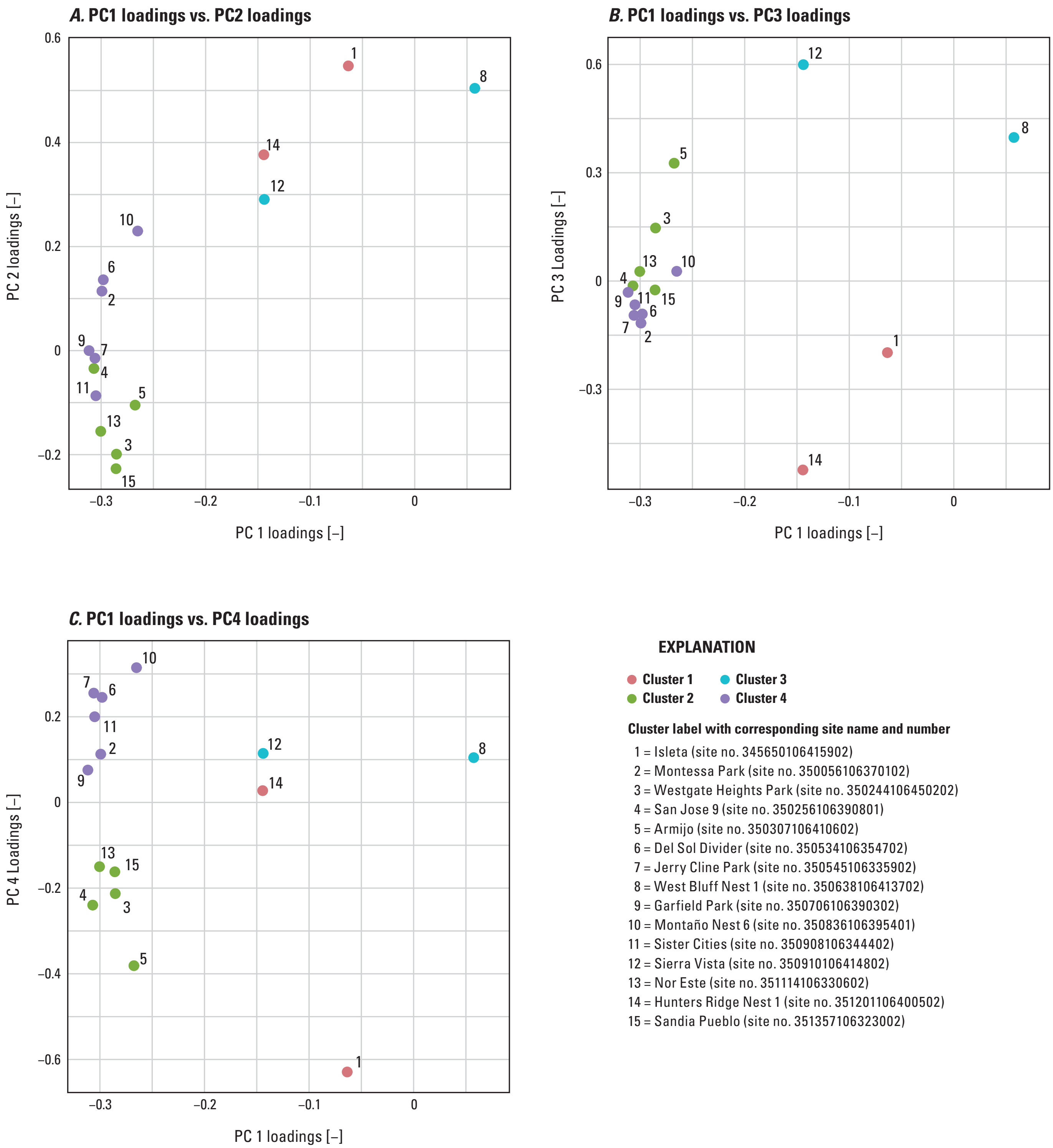


Figure 22. Principal component (PC) loadings plots for PCs 1 through 4 showing the final clustered solution: *A*, PC1 loadings versus PC2 loadings; *B*, PC1 loadings versus PC3 loadings; *C*, PC1 loadings versus PC4 loadings; *D*, PC2 loadings versus PC3 loadings; *E*, PC2 loadings versus PC4 loadings; and *F*, PC3 loadings versus PC4 loadings.

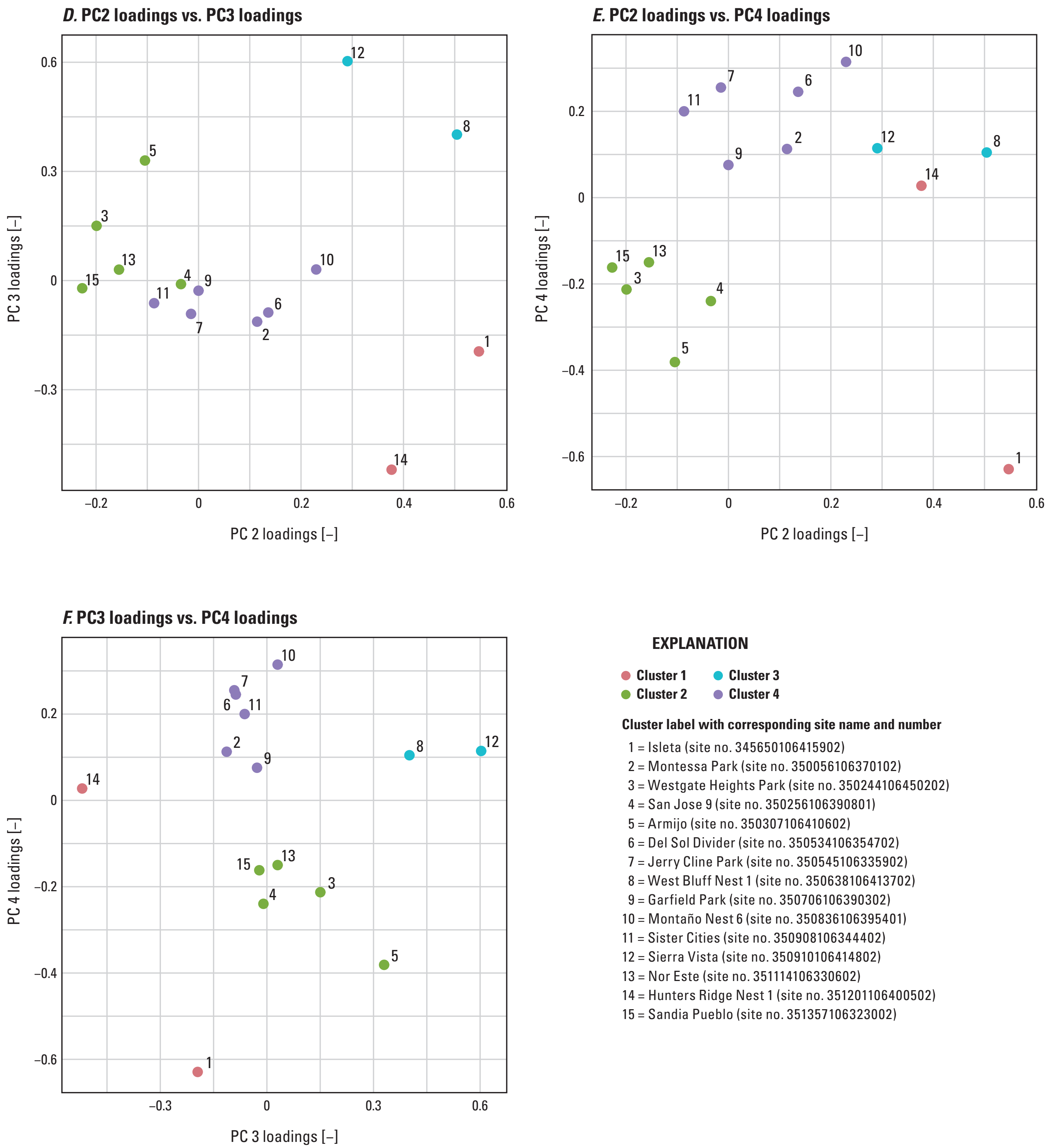


Figure 22. Principal component (PC) loadings plots for PCs 1 through 4 showing the final clustered solution: *A*, PC1 loadings versus PC2 loadings; *B*, PC1 loadings versus PC3 loadings; *C*, PC1 loadings versus PC4 loadings; *D*, PC2 loadings versus PC3 loadings; *E*, PC2 loadings versus PC4 loadings; and *F*, PC3 loadings versus PC4 loadings.—Continued

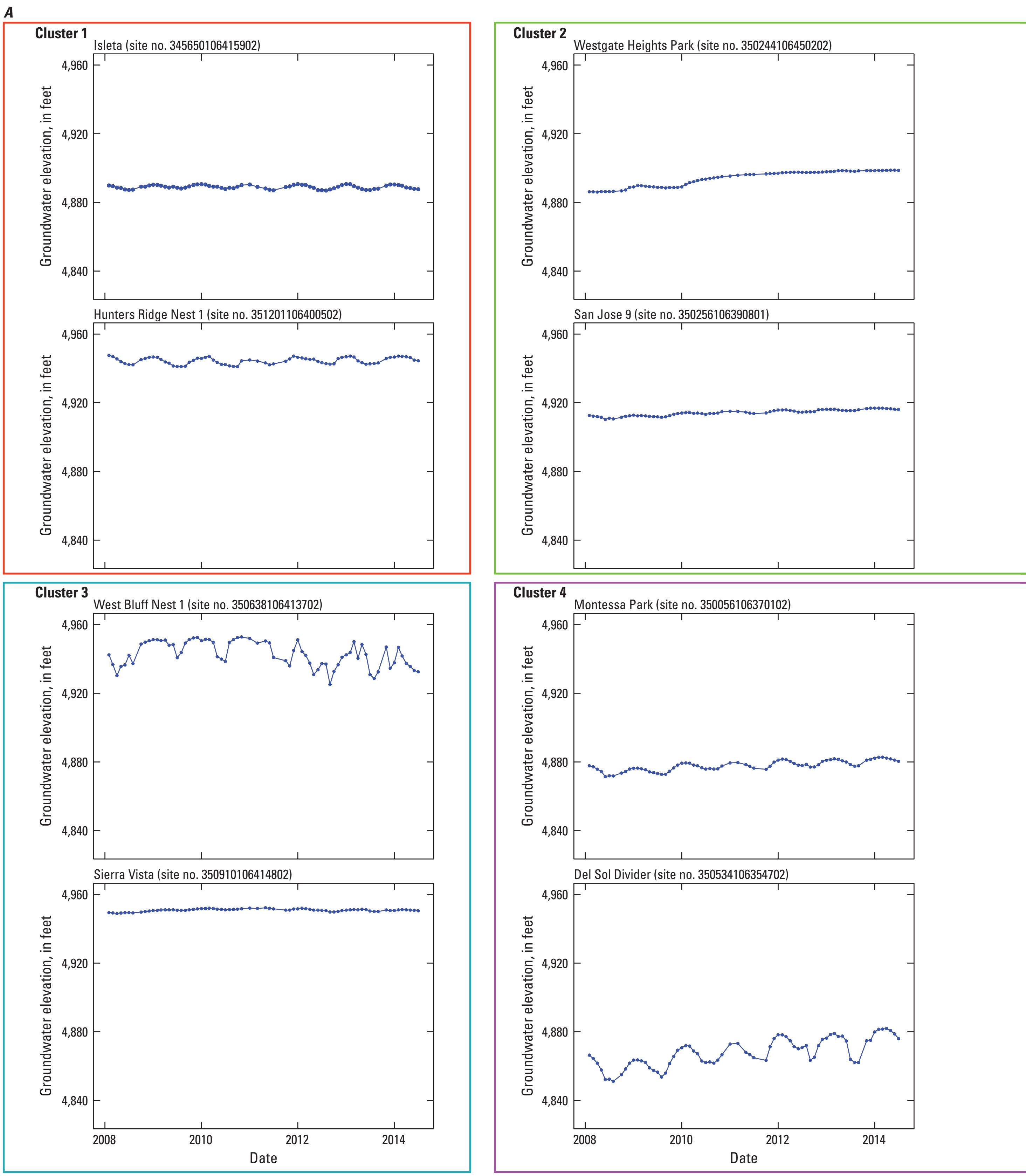
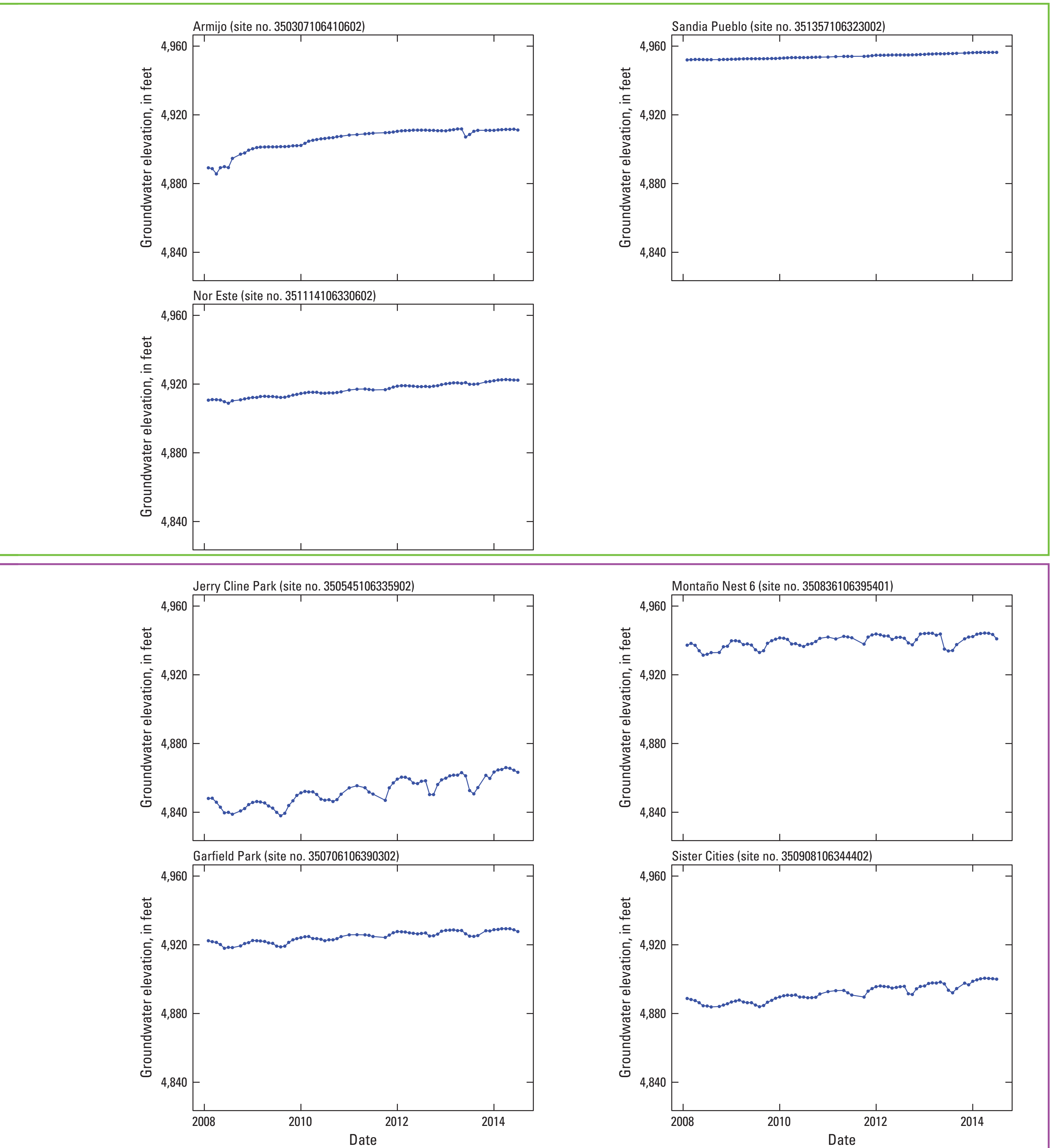


Figure 23. Clustered monthly median hydrographs plotted using *A*, the same fixed y-axis range and *B*, an independent y-axis range for all plots.


Armijo (site no. 350307106410602)
Sandia Pueblo (site no. 351357106323002)
Nor Este (site no. 351114106330602)
Jerry Cline Park (site no. 350545106335902)
Montaño Nest 6 (site no. 350836106395401)
Garfield Park (site no. 350706106390302)
Sister Cities (site no. 350908106344402)
Groundwater elevation, in feet
4,960
4,920
4,880
4,840
2008
2010
2012
2014
Date

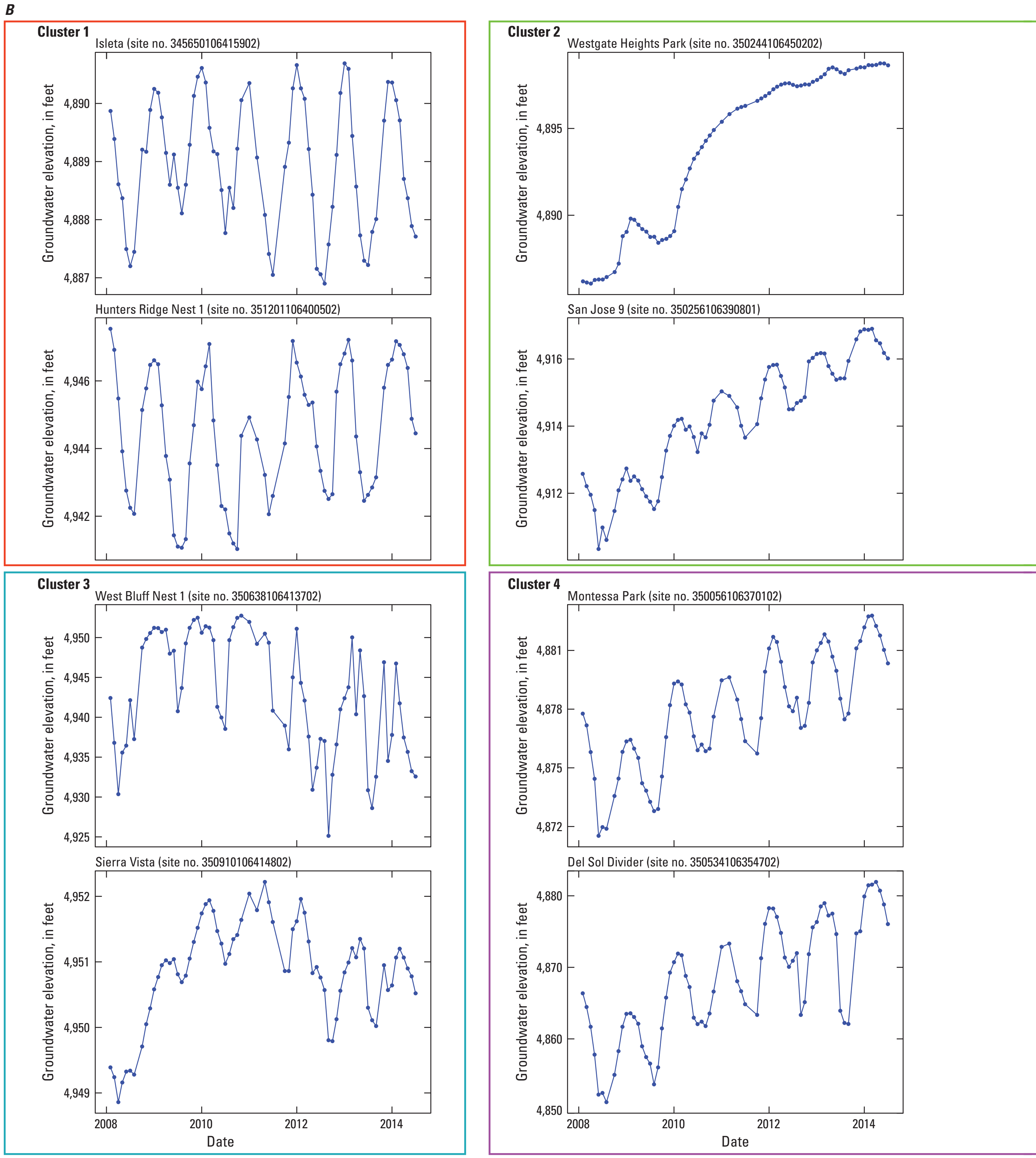
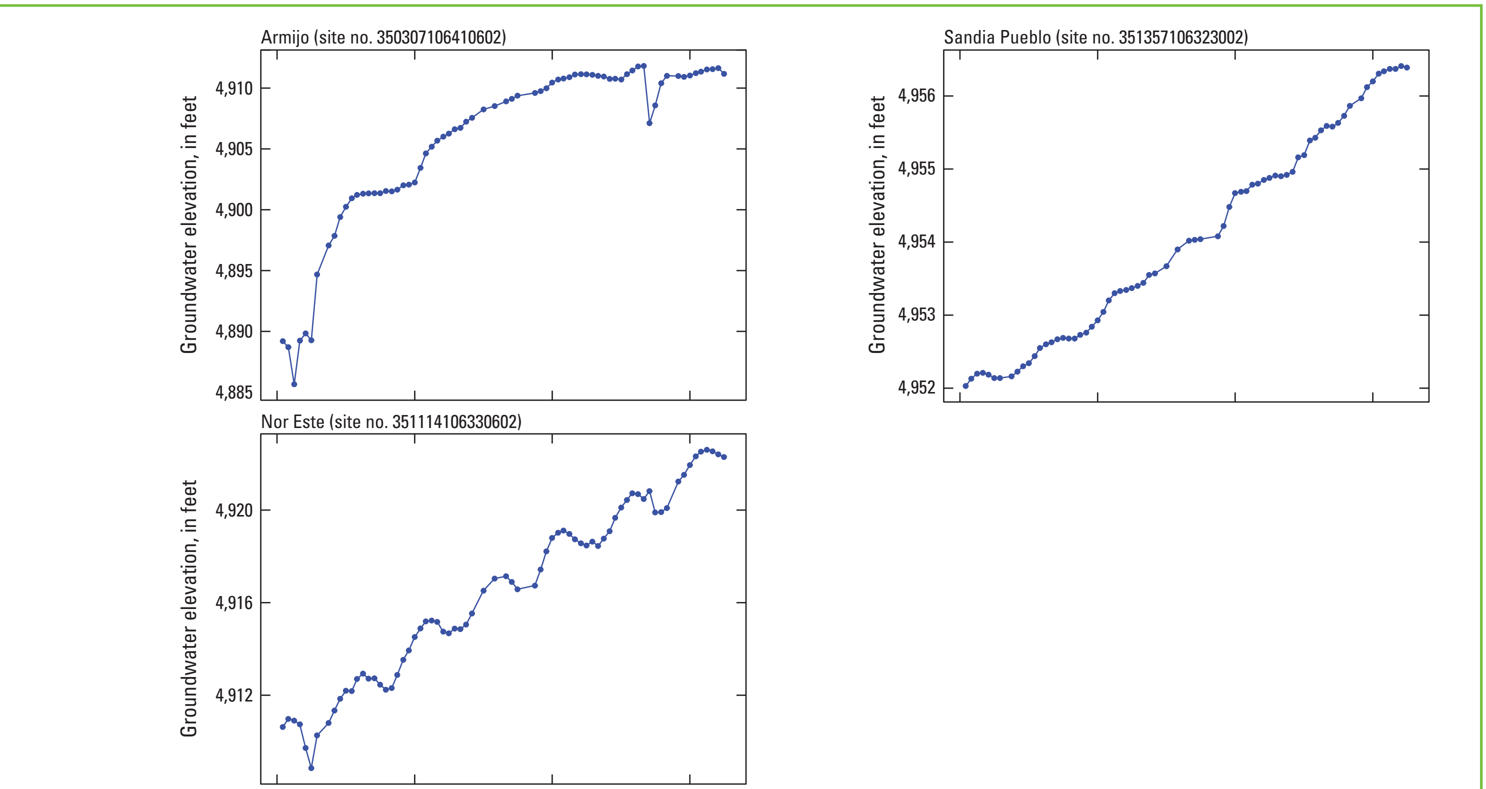
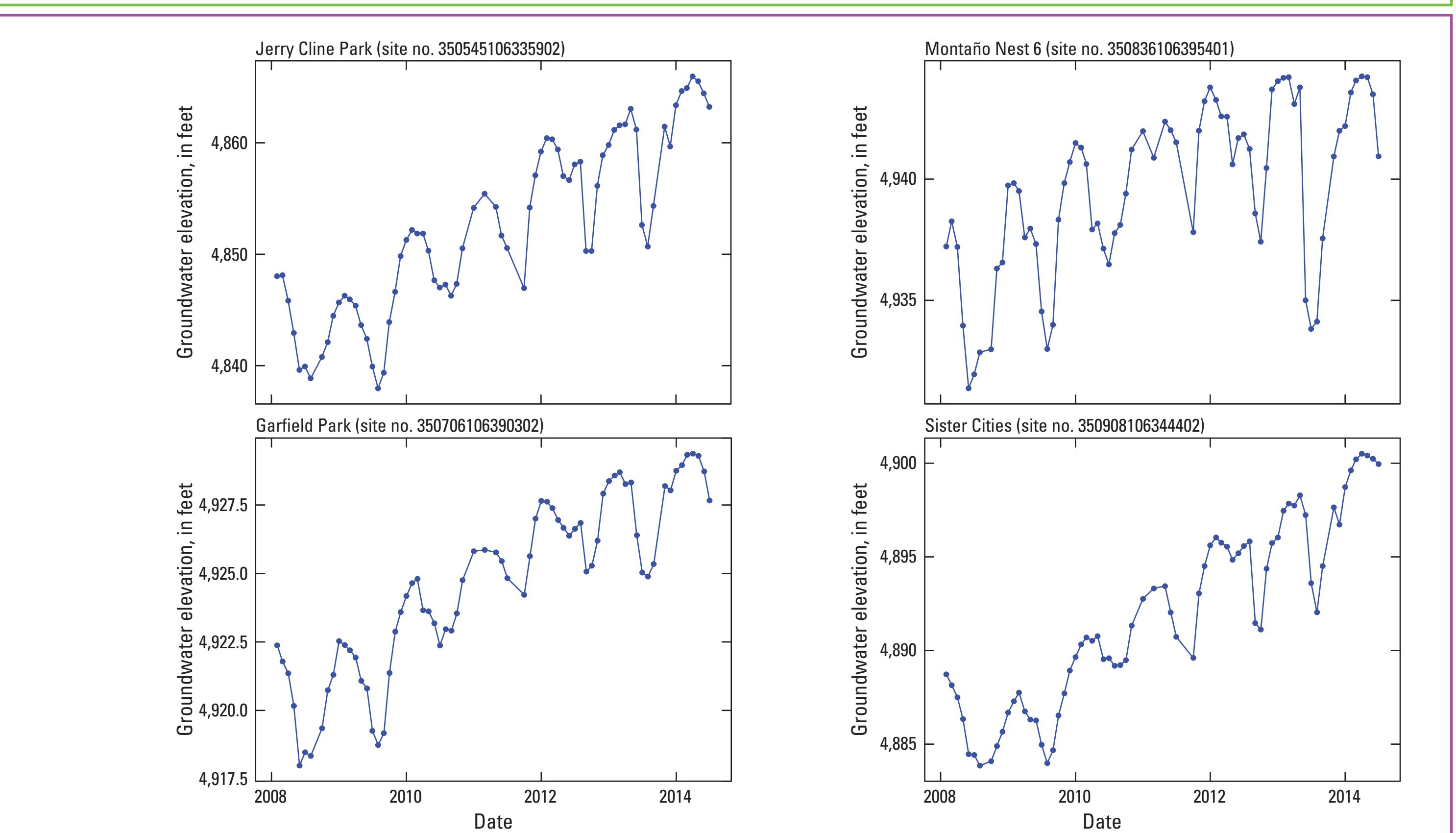


Figure 23. Clustered monthly median hydrographs plotted using *A*, the same fixed y-axis range and *B*, an independent y-axis range for all plots.—Continued


Armijo (site no. 350307106410602)
Sandia Pueblo (site no. 351357106323002)
Nor Este (site no. 351114106330602)
Groundwater elevation, in feet
4,885
4,890
4,895
4,900
4,905
4,910
4,952
4,953
4,954
4,955
4,956
4,912
4,916
4,920


Jerry Cline Park (site no. 350545106335902)
Montaño Nest 6 (site no. 350836106395401)
Garfield Park (site no. 350706106390302)
Sister Cities (site no. 350908106344402)
Groundwater elevation, in feet
4,840
4,850
4,860
4,935
4,940
4,917.5
4,920.0
4,922.5
4,925.0
4,927.5
4,885
4,890
4,895
4,900
2008
2010
2012
2014
Date

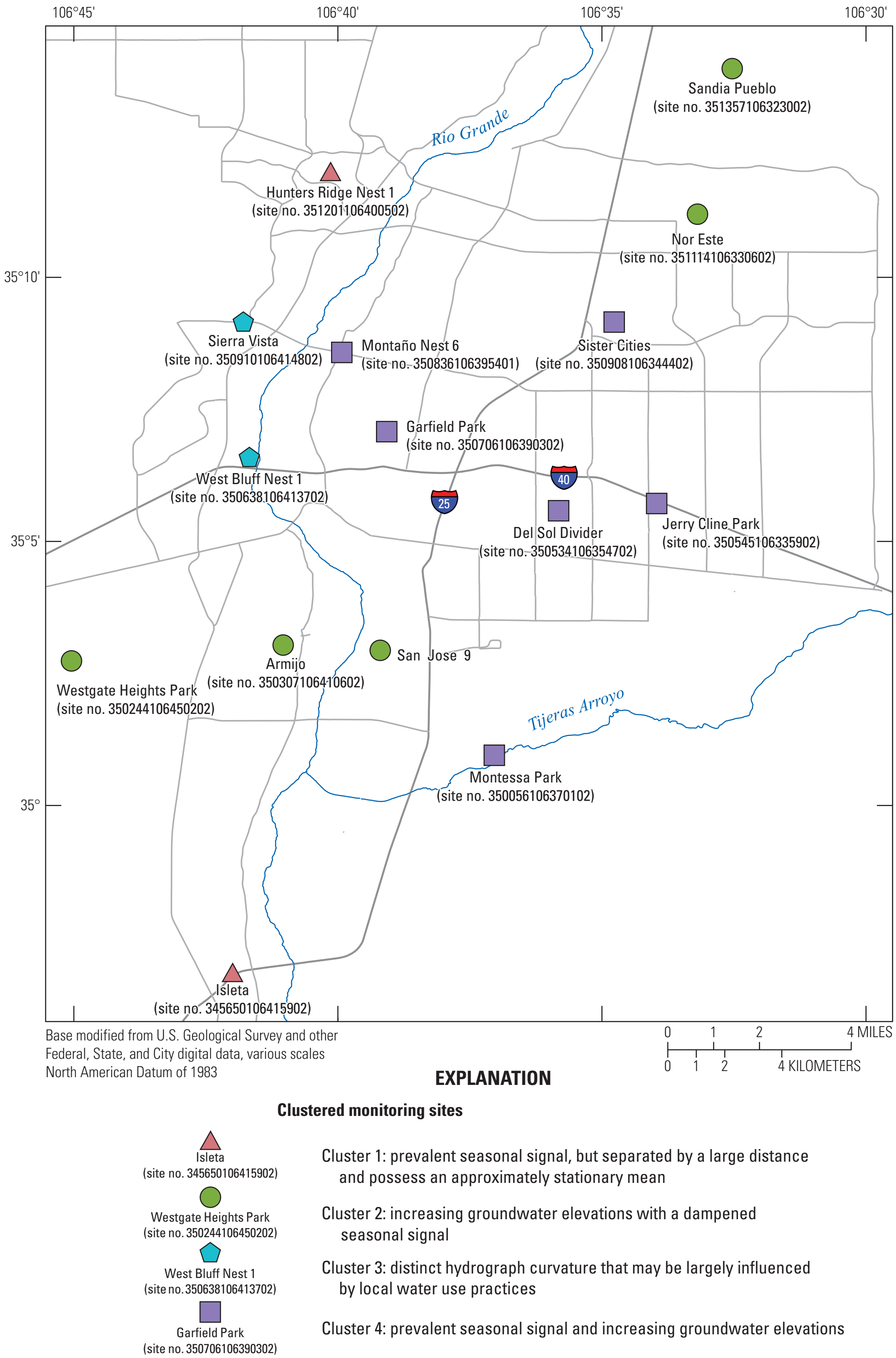


Figure 24. The spatial distribution of clustered monitoring sites.

The two hydrographs in cluster 3 are located near each other (fig. 24) but appear to be quite different when plotted on the same fixed y-axis scale (fig. 23*A*). However, inspection of the hydrographs on independent y-axis scales (fig. 23*B*) shows a sharp increase in groundwater elevation early in the hydrograph that is followed by a period of relative stability that is then followed by a decline. This curvature is unique to this cluster and is likely why these two hydrographs are grouped with each other and separate from the other hydrographs. The timing of the groundwater-elevation decline in these two plots coincides with increased pumping rates at several nearby production wells, thereby indicating that this unique curvature likely results from local water use practices. In this cluster, the seasonal signal in the hydrograph of West Bluff Nest 1 (site no. 350638106413702) is more pronounced than that in the hydrograph of Sierra Vista (site no. 350910106414802) (fig. 23). The midpoint depth of the screened intervals at these two sites differs by 239 feet (West Bluff Nest 1 [site no. 350638106413702] = 682 feet; Sierra Vista [site no. 350910106414802] = 921 feet). This pattern is therefore consistent with the greater susceptibility of the shallower monitoring sites to seasonal variations as compared to the deeper sites.

Overall, the clustered-PCA results indicate that these hydrographs can be categorized into four groups of differing characteristics (figs. 23 and 24). Hydrographs in cluster 1 have a prevalent seasonal signal but are separated by a large distance and possess an approximately stationary mean. Cluster 2 is a group of hydrographs that show increasing groundwater elevations without a pronounced seasonal signal relative to the majority of analyzed hydrographs. This difference is likely related to these monitoring sites being screened about 20 percent deeper than the other sites considered in the PCA. Cluster 3 is a unique group of hydrographs that have distinct curvature that may be largely influenced by local water use practices. Cluster 4 is the largest group of hydrographs, has a prevalent seasonal signal, and shows increasing groundwater elevations. The pronounced seasonal signal of this cluster differentiates it from cluster 2. Except for cluster 1, all clusters have the potential to be well represented by fewer index monitoring sites. These findings represent an additional resource to inform management decisions that are made when using the spatial and temporal assessment results.

Summary

The Albuquerque Bernalillo County Water Utility Authority (ABCWUA) provides water and wastewater services to more than 600,000 water users within the Middle Rio Grande Basin. Groundwater levels are periodically measured by the U.S. Geological Survey, in cooperation with the ABCWUA, at an observation network of wells and piezometers in the City of Albuquerque and adjacent areas, New Mexico, that are managed by the ABCWUA to provide a long-term hydrologic dataset that is heavily relied upon to make water management decisions. The desire to upgrade and perform maintenance on this observation network initiated this study to assess the spatial and temporal importance of measurements towards optimization of monitoring the observation network to reduce or redirect monitoring costs. Two monitoring reduction techniques are considered: (1) removing monitoring sites that have a minimal effect on the overall spatial characterization of the groundwater-elevation surface at a particular time and (2) decreasing the frequency of measurements at monitoring sites that indicate statistically similar temporal trends at lower measurement frequencies. The optimization assessment included separate spatial and temporal methodologies and an evaluation using principal component analysis (PCA). The spatial assessment and PCA both considered only monitoring sites that were screened at depths within the production zone of the Santa Fe Group aquifer system.

The spatial assessment aimed to determine the spatial importance of monitoring sites by using a kriging-based genetic algorithm approach to analyze median groundwater elevations from 87 production zone monitoring sites over two winter periods: the 2001 time interval and 2015 time interval. These results are most suited to identifying sites that could be removed with minimal effect on the generation of winter groundwater-elevation contour maps of the production zone. Results from the temporal assessment and PCA can also be consulted when deciding which sites to remove entirely from the network, especially for sites that are monitored more frequently than annually. The spatial assessment results indicate that as many as about 20 sites can be removed with a relatively small degradation in the ability to represent the winter kriged surfaces that were generated by using the existing observation network.

The temporal assessment considered 47 hourly monitored sites and utilized an iterative thinning technique coupled with trend statistical analyses to identify sites that may be sampled less frequently without significant information loss. Iterative thinning of the monthly median data at each monitoring site was performed for a range of reduced monitoring frequencies generally corresponding to bimonthly (every 2 months), quarterly (every 3 months), semiannually (every 6 months), annually (every 12 months), every 2 years, and every 5 years. The trend of the reduced frequency data was computed and compared to the trend of the monthly data. Results from the temporal assessment can be used to inform minimum monitoring frequency at the observation network required to capture the trends shown in higher frequency monitoring. At many sites, trends were well represented for at least one of the reduced monitoring frequencies tested, indicating that a reduction in frequency may be sufficient to adequately characterize seasonal and long-term trends.

PCA and k-means clustering analysis were applied to the hydrographs (groundwater elevation over time) of 15 hourly monitored sites that are screened within the production zone

of the aquifer to identify groups of monitoring sites that have similar hydrographs. This information can be utilized to inform the management decisions that are made when using the spatial and temporal assessment results. The PCA results indicate that the hydrographs can be categorized into four groups of differing characteristics. Hydrographs in cluster 1 have a prevalent seasonal signal but are separated by a large distance and possess an approximately stationary mean. Cluster 2 is a group of hydrographs that show increasing groundwater elevations without a pronounced seasonal signal relative to the majority of analyzed hydrographs. This difference is likely related to these monitoring sites being screened about 20 percent more deeply than the other sites considered in the PCA. Cluster 3 is a unique group of hydrographs that have distinct curvature that may be largely influenced by local water use practices. Cluster 4 is the largest group of hydrographs, has a prevalent seasonal signal, and shows increasing groundwater elevations. The pronounced seasonal signal of cluster 4 differentiates it from cluster 2. Except for cluster 1, all clusters have the potential to be well represented by fewer index monitoring sites.

This study was conducted to increase understanding of how reductions in monitoring fidelity might affect the existing observation network. Reducing the temporal frequency or spatial density of monitoring is ultimately a management decision that depends on the amount of data loss or degradation that is deemed acceptable while still meeting the network objectives of the ABCWUA. However, this study could also be used to guide ongoing efforts to understand local aquifer characteristics and responses, as well as guide enhanced monitoring as needs arise or in advance of future water management decisions.

Acknowledgments

The authors wish to thank the numerous USGS hydrologists and hydrologic technicians who established, maintained, and collected high-quality data at the network of wells and piezometers used in this report.

References Cited

Albuquerque Bernalillo County Water Utility Authority [ABCWUA], 2016, Water 2120: Securing our water future: Volume I: Main text and appendices: 309 p., accessed December 7, 2017, at https://www.abcwua.org/uploads/files/Water_2120_Volume_I.pdf.

Asefa, T., Kemblowski, M.W., Gilberto, U., McKee, M., and Khalil, A., 2004, Support vectors-based groundwater head observation networks design: Water Resources Research, v. 40, no. 11, accessed December 10, 2019, at https://doi.org/10.1029/2004WR003304.

Aziz, J.J., Ling, M., Rifai, H.S., Newell, C.J., and Gonzales, J.R., 2003, MAROS—A decision support system for optimizing monitoring plans: Ground Water, v. 41, no. 3, p. 355–367.

Bartolino, J.R., and Cole, J.C., 2002, Ground-water resources of the Middle Rio Grande Basin: U.S. Geological Survey Circular 1222, 132 p., accessed October 23, 2018, at https://doi.org/10.3133/cir1222.

Bexfield, L.M., and Anderholm, S.K., 2002, Estimated water-level declines in the Santa Fe Group aquifer system in the Albuquerque area, central New Mexico, predevelopment to 2002: U.S. Geological Survey Water-Resources Investigations Report 02–4233, accessed May 5, 2017, at https://doi.org/10.3133/wri024233.

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.

Cunningham, W.L., and Schalk, C.W., comps., 2011, Groundwater technical procedures of the U.S. Geological Survey: U.S. Geological Survey Techniques and Methods 1–A1, 151 p., accessed March 5, 2020, at https://doi.org/10.3133/tm1A1.

Dhar, A., and Datta, B., 2010, Logic-based design of groundwater monitoring network for redundancy reduction: Journal of Water Resources Planning and Management, v. 136, no. 1, p. 88–94.

Driscoll, J.M., and Brandt, J.T., 2017, Land subsidence and recovery in the Albuquerque Basin, New Mexico, 1993–2014: U.S. Geological Survey Scientific Investigations Report 2017–5057, 31 p., accessed October 23, 2018, at https://doi.org/10.3133/sir20175057.

Everitt, B.S., and Hothorn, T., 2010, A handbook of statistical analyses using R (2d ed.): Boca Raton, Fla., Chapman & Hall/CRC, 348 p.

Everitt, B.S., Landau, S., Leese, M., and Stahl, D., 2011, Cluster analysis (5th ed.): West Sussex, United Kingdom, John Wiley & Sons, Ltd., 330 p.

Falk, S.E., Bexfield, L.M., and Anderholm, S.K., 2011, Estimated 2008 groundwater potentiometric surface and predevelopment to 2008 water-level change in the Santa Fe Group aquifer system in the Albuquerque area, central New Mexico: U.S. Geological Survey Scientific Investigations Map 3162, accessed October 23, 2018, at https://doi.org/10.3133/sim3162.

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., accessed October 23, 2018, at https://doi.org/10.3133/sir20135120.

Galanter, A.E., and Curry, L.T.S., 2019, Estimated 2016 groundwater level and drawdown from predevelopment to 2016 in the Santa Fe Group aquifer system in the Albuquerque area, central New Mexico: U.S. Geological Survey Scientific Investigations Map 3433, 1 sheet, 13-p. pamphlet, accessed May 24, 2019, at https://doi.org/10.3133/sim3433.

Gangopadhyay, S., Gupta, A.D., and Nachabe, M.H., 2001, Evaluation of ground water monitoring network by principal component analysis: Ground Water, v. 39, no. 2, p. 181–191.

Hartigan, J.A., and Wong, M.A., 1979, Algorithm AS 136: A *K*-means clustering algorithm: Journal of the Royal Statistical Society, Series C (Applied Statistics), v. 28, no. 1, p. 100–108.

Hastie, T., Tibshirani, R., and Friedman, J., 2009, The elements of statistical learning—Data mining, inference, and prediction (2d ed.): New York, Springer Science+Business Media, LLC, 745 p.

Hawley, J.W., and Haase, C.S., 1992, Hydrogeologic framework of the northern Albuquerque Basin: New Mexico Bureau of Mines and Mineral Resources Open-File Report 387, variously paged, accessed April 23, 2019, at https://geoinfo.nmt.edu/publications/openfile/details.cfml?Volume=387.

Heath, R.C., 1976, Design of ground-water level observation-well programs: Ground Water, v. 14, no. 2, p. 71–77.

Herrera, G.S., and Pinder, G.F, 2005, Space-time optimization of groundwater quality sampling networks: Water Resources Research, v. 41, no. 12, accessed December 10, 2019, at https://doi.org/10.1029/2004WR003626.

Huang, C., and Mayer, A.S., 1997, Pump-and-treat optimization using well locations and pumping rates as decision variables: Water Resources Research, v. 33, no. 5, p. 1001–1012.

Hudson, M.R., and Grauch, V.J.S., 2013, Introduction, *in* Hudson, M.R., and Grauch, V.J.S., eds., New perspectives on Rio Grande Rift basins—From tectonics to groundwater: Geological Society of America Special Paper 494, p. v–xii, accessed December 10, 2019, at https://doi.org/10.1130/2013.2494(00).

Iglewicz, B., and Hoaglin, D.C., 1993, Volume 16: How to detect and handle outliers: Milwaukee, Wis., ASQC Quality Press, 87 p.

James, G., Witten, D., Hastie, T., and Tibshirani, R., 2013, An introduction to statistical learning with applications in R: New York, Springer Science+Business Media, 426 p., accessed December 10, 2019, at http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf.

Jones, M.A., Melis, E.A., and Baggerman, J., 2013, Progress report, updated evaluation of the deep bedrock aquifers in and around the Middle Rio Grande Basin, New Mexico: John Shomaker & Associates, Inc., 86 p., accessed October 23, 2018, at http://commongroundrising.org/wp-content/uploads/2018/07/ISC-JSAI_deepAquifModel_rpt_2013June.pdf.

Kernodle, J.M., 1996, Hydrogeology and steady-state simulation of ground-water flow in the San Juan Basin, New Mexico, Colorado, Arizona, and Utah: U.S. Geological Survey Water-Resources Investigations Report 95–4187, 117 p., accessed December 11, 2019, at https://doi.org/10.3133/wri954187.

Khan, S., Chen, H.F., and Rana, T., 2008, Optimizing ground water observation networks in irrigation areas using principal component analysis: Ground Water Monitoring & Remediation, v. 28, no. 3, p. 93–100.

Li, Y., and Hilton, A.B.C., 2007, Optimal groundwater monitoring design using an ant colony optimization paradigm: Environmental Modelling & Software, v. 22, no. 1, p. 110–116.

Loaiciga, H.A., Charbeneau, R.J., Everett, L.G., Fogg, G.E., Hobbs, B.F., and Rouhani, S., 1992, Review of ground-water quality monitoring network design: Journal of Hydraulic Engineering, v. 118, no. 1, p. 11–37.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., Studer, M., and Roudier, P., 2019, Package 'cluster': Comprehensive R Archive Network (CRAN), accessed June 6, 2019, at https://cran.r-project.org/web/packages/cluster/cluster.pdf.

Massmann, J., and Freeze, R.A., 1987a, Groundwater contamination from waste management sites—The interaction between risk-based engineering design and regulatory policy—1. Methodology: Water Resources Research, v. 23, no. 2, p. 351–367.

Massmann, J., and Freeze, R.A., 1987b, Groundwater contamination from waste management sites—The interaction between risk-based engineering design and regulatory policy—2. Results: Water Resources Research, v. 23, no. 2, p. 368–380.

McAda, D.P., and Barroll, P., 2002, Simulation of ground-water flow in the Middle Rio Grande Basin between Cochiti and San Acacia, New Mexico: U.S. Geological Survey Water-Resources Investigations Report 02–4200, 81 p., accessed October 23, 2018, at https://doi.org/10.3133/wri20024200.

Meyer, P.D., and Brill, E.D., Jr., 1988, A method for locating wells in a groundwater monitoring network under conditions of uncertainty: Water Resources Research, v. 24, no. 8, p. 1277–1282.

Nunes, L.M., Cunha, M.C., and Ribeiro, L., 2004a, Groundwater monitoring network optimization with redundancy reduction: Journal of Water Resources Planning and Management, v. 130, no. 1, p. 33–43.

Nunes, L.M., Paralta, E., Cunha, M.C., and Ribeiro, L., 2004b, Groundwater nitrate monitoring network optimization with missing data: Water Resources Research, v. 40, no. 2, accessed December 10, 2019, at https://doi.org/10.1029/2003WR002469.

Olea, R.A., 1984, Sampling design optimization for spatial functions: Mathematical Geology, v. 16, no. 4, p. 369–392.

Pebesma, E., and Graeler, B., 2019, Package 'gstat': Comprehensive R Archive Network (CRAN), accessed March 11, 2019, at https://cran.r-project.org/web/packages/gstat/gstat.pdf.

Plummer, L.N., Bexfield, L.M., Anderholm, S.K., Sanford, W.E., and Busenberg, E., 2012, Geochemical characterization of ground-water flow in the Santa Fe Group aquifer system, Middle Rio Grande Basin, New Mexico (ver. 1.2, November 20, 2012): U.S. Geological Survey Water-Resources Investigations Report 03–4131, 395 p., CD-ROM in pocket. [Also available at https://doi.org/10.3133/wri034131.]

Powell, R.I., and McKean, S.E., 2014, Estimated 2012 groundwater potentiometric surface and drawdown from predevelopment to 2012 in the Santa Fe Group aquifer system in the Albuquerque metropolitan area, central New Mexico: U.S. Geological Survey Scientific Investigations Map 3301, accessed April 7, 2017, at https://doi.org/10.3133/sim3301.

R Core Team, 2019, R: A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed August 13, 2019, at https://www.R-project.org/.

Rankin, D.R., McCoy, K.J., Moret, G.J.M., Worthington, J.A., and Bandy-Baldwin, K.M., 2013, Groundwater hydrology and estimation of horizontal groundwater flux from the Rio Grande at selected locations in Albuquerque, New Mexico, 2003–9: U.S. Geological Survey Scientific Investigations Report 2012–5007, 66 p., accessed October 23, 2018, at https://doi.org/10.3133/sir20125007.

Rice, S.E., Oelsner, G.P., and Heywood, C.E., 2014, Simulated and measured water levels and estimated water-level changes in the Albuquerque area, central New Mexico, 1950–2012: U.S. Geological Survey Scientific Investigations Map 3305, accessed October 23, 2018, at https://doi.org/10.3133/sim3305.

Ritchie, A.B., and Galanter, A.E., 2019, Water-level data for the Albuquerque Basin and adjacent areas, central New Mexico, period of record through September 30, 2018: U.S. Geological Survey Data Series 1116, 42 p., accessed August 16, 2019, at https://doi.org/10.3133/ds1116.

Ritchie, A.B., Galanter, A.E., and Curry, L.T.S., 2019, Groundwater-level change for the periods 2002–8, 2008–12, and 2008–16 in the Santa Fe Group aquifer system in the Albuquerque area, central New Mexico: U.S. Geological Survey Scientific Investigations Map 3435, 1 sheet, 16-p. pamphlet, accessed July 3, 2019, at https://doi.org/10.3133/sim3435.

Rouhani, S., 1985, Variance reduction analysis: Water Resources Research, v. 21, no. 6, p. 837–846.

Rouhani, S., and Hall, T.J., 1988, Geostatistical schemes for groundwater sampling: Journal of Hydrology, v. 103, nos. 1–2, p. 85–102.

Sanford, W.E., Plummer, L.N., McAda, D.P., Bexfield, L.M., and Anderholm, S.K., 2004, Use of environmental tracers to estimate parameters for a predevelopment-ground-water-flow model of the Middle Rio Grande Basin, New Mexico: U.S. Geological Survey Water-Resources Investigations Report 03–4286, 102 p., accessed October 23, 2018, at https://doi.org/10.3133/wri034286.

Scheibe, T.D., and Lettenmaier, D.P., 1989, Risk-based selection of monitoring wells for assessing agricultural chemical contamination of ground water: Groundwater Monitoring & Remediation, v. 9, no. 4, p. 98–108.

Sen, P.K., 1968, Estimates of the regression coefficient based on Kendall's tau: Journal of the American Statistical Association, v. 63, no. 324, p. 1379–1389.

Severson, K.A., Molaro, M.C., and Braatz, R.D., 2017, Principal component analysis of process datasets with missing values: Processes, v. 5, no. 38, accessed December 10, 2019, at https://doi.org/10.3390/pr5030038.

Smirnov, M.Y., and Egbert, G.D., 2012, Robust principal component analysis of electromagnetic arrays with missing data: Geophysical Journal International, v. 190, p. 1423–1438.

Taylor, C.J., and Alley, W.M., 2001, Ground-water-level monitoring and the importance of long-term water-level data: U.S. Geological Survey Circular 1217, 68 p., accessed November 1, 2018, at https://doi.org/10.3133/cir1217.

Thakur, J.K., 2015, Optimizing groundwater monitoring networks using integrated statistical and geostatistical approaches: Hydrology, v. 2, no. 3, p. 148–175.

Thorn, C.R., McAda, D.P., and Kernodle, J.M., 1993, Geohydrologic framework and hydrologic conditions in the Albuquerque Basin, central New Mexico: U.S. Geological Survey Water-Resources Investigations Report 93–4149, 106 p., accessed October 23, 2018, at https://doi.org/10.3133/wri934149.

Tukey, J.W., 1977, Exploratory data analysis: London, Pearson, 712 p.

U.S. Census Bureau, 2010, 2010 Census: U.S. Census Bureau American FactFinder, accessed October 23, 2018, at https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml.

U.S. Geological Survey [USGS], 2019a, Groundwater daily data, USGS water data for the nation help system: U.S. Geological Survey National Water Information System Help System, accessed December 10, 2019, at https://help.waterdata.usgs.gov/faq/groundwater/groundwater-daily-data.

U.S. Geological Survey [USGS], 2019b, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 23, 2019, at https://doi.org/10.5066/F7P55KJN.

Wagner, B.J., and Gorelick, S.M., 1987, Optimal groundwater quality management under parameter uncertainty: Water Resources Research, v. 23, no. 7, p. 1162–1174.

Wagner, B.J., and Gorelick, S.M., 1989, Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: From data to design: Water Resources Research, v. 25, no. 10, p. 2211–2225.

Winter, T.C., Mallory, S.E., Allen, T.R., and Rosenberry, D.O., 2000, The use of principal component analysis for interpreting ground water hydrographs: Ground Water, v. 38, no. 2, p. 234–246.

For more information about this publication, contact
Director, New Mexico Water Science Center
U.S. Geological Survey
6700 Edith Blvd. NE
Albuquerque, NM 87113

For additional information visit
https://www.usgs.gov/centers/nm-water

Publishing support provided by
Lafayette Publishing Service Center

USGS

Ritchie and Pepin—**Optimization Assessment of a Groundwater-Level Observation Network in the Middle Rio Grande Basin, New Mexico**—SIR 2020–5007, ver. 2

ISBN 978-1-4113-4384-9

9 781411 343849

ISSN 2328-031X (print)
ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20205007