Groundwater/Surface-Water Interactions in the Bad River Watershed, Wisconsin

Scientific Investigations Report 2015–5162

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

Groundwater/Surface-Water Interactions in the Bad River Watershed, Wisconsin

By Andrew T. Leaf, Michael N. Fienen, Randall J. Hunt, and Cheryl Buchwald

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

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2015

Contents

Figures

Tables

Conversion Factors

Multiply	By	To obtain
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
square mile (mi^2)	2.590	square kilometer (km^2)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32)/1.8$$

Datum

Horizontal coordinate information is referenced to the North American Datum of 1927 (NAD 27).
Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88).

Acknowledgments

Groundwater/Surface-Water Interactions in the Bad River Watershed, Wisconsin

By Andrew T. Leaf, Michael N. Fienen, Randall J. Hunt, and Cheryl A. Buchwald

Abstract

A groundwater-flow model was developed for the Bad River Watershed and surrounding area by using the U.S. Geological Survey (USGS) finite-difference code MODFLOW–NWT. The model simulates steady-state groundwater-flow and base flow in streams by using the streamflow routing (SFR) package. The objectives of this study were to: (1) develop an improved understanding of the groundwater-flow system in the Bad River Watershed at the regional scale, including the sources of water to the Bad River Band of Lake Superior Chippewa Reservation (Reservation) and groundwater/surface-water interactions; (2) provide a quantitative platform for evaluating future impacts to the watershed, which can be used as a starting point for more detailed investigations at the local scale; and (3) identify areas where more data are needed. This report describes the construction and calibration of the groundwater-flow model that was subsequently used for analyzing potential locations for the collection of additional field data, including new observations of water-table elevation for refining the conceptualization and corresponding numerical model of the hydrogeologic system.

The study area can be conceptually divided into three primary hydrogeologic environments. The first encompasses the southern uplands with relatively low topographic relief, where groundwater flow is unconfined and occurs primarily in sandy till and glacial outwash overlying Archean-aged crystalline bedrock. The second includes a transitional area of higher topographic relief and shallow depth to bedrock, in the vicinity of ridges formed by steeply dipping, early-Proterozoic aged metasedimentary units of the Marquette Range Supergroup (including the Ironwood Formation), and late-Proterozoic igneous units associated with the Midcontinent Rift System (MRS). Groundwater flow in this area likely occurs primarily through connected networks of bedrock fractures that are not well characterized, and also in isolated pockets of Quaternary deposits. The third and last hydrogeologic environment includes lowlands along Lake Superior where a deep sandstone aquifer is confined by thick deposits of clay-rich till.

Model input was compiled by using both published and unpublished data. Constant flux boundary conditions for the model perimeter were developed from a regional analytic element model described in appendix 1 of this report. Pumping from 26 high-capacity wells within the model area was included. The SFR stream network was developed from the National Hydrography Dataset (NHDPlus Version 2) and hydrography from the Wisconsin Department of Natural Resources (WDNR). Hydraulic conductivity values were determined for each model cell by interpolation from a network of pilot points, within zones representing major hydrogeologic units.

Recharge to the groundwater system was estimated on a cell-by-cell basis by using the Soil Water Balance code (SWB), with gridded daily temperature and precipitation data for the period 1980–2011, and GIS coverages of soil and land-surface conditions. Estimated recharge varies considerably, following spatial patterns in the precipitation and soil hydrologic group inputs. The lowest recharge values occur in the Superior Lowlands, whereas the highest values occur in the upland areas, especially those underlain by sandy soils, and in the vicinity of bedrock hills.

The model was calibrated to groundwater levels and base flows obtained from the USGS National Water Information System (NWIS) database, and groundwater levels obtained from the WDNR and Bad River Band well-construction databases. Calibration was performed via nonlinear regression by using the parameter-estimation software suite PEST. Groundwater levels and base-flow observations in the calibration dataset were well simulated by the calibrated model, with reasonable values of hydraulic conductivity. The pilot-point parameters that were most constrained by observations during model calibration coincided with the locations containing the most wells (head observations)—especially the population centers of Ashland, Mellen, and other communities along the major highway corridors.

Results from the calibrated model illustrate differences in the nature of groundwater-flow within the watershed. In the southern part of the watershed, where bedrock is shallow, groundwater flow paths are relatively short, extending from local recharge areas to adjacent first and second-order streams. In contrast, laterally continuous deposits of clay-rich till covering the Superior Lowlands isolate most smaller streams from the sandstone aquifer, allowing for longer flow paths toward larger streams such as the Bad, Marengo, and White Rivers. Approximately three-quarters of all first-order stream cells were dry in the Superior Lowlands, compared to only half of first-order stream cells in the southern bedrock uplands.

The model was used to delineate the groundwatershed for the Bad and Kakagon Rivers. "Groundwatershed" is defined as the area contributing groundwater discharge to one of these streams and their tributaries. The groundwatershed was found to align closely with the surface-watershed, with the most notable exception occurring along the southwestern half of Birch Hill, where surface water drains southwest towards the Potato River, and groundwater flows north and east towards Lake Superior. Similarly, the contributing area of groundwater-flow to the Reservation was delineated. Results indicate the off-Reservation groundwater contributing area to be limited in comparison to the extent of the watershed, extending southward into the highlands underlain by MRS igneous rock units, but not further into the area underlain by the Marquette Range Supergroup.

Stable isotope samples were collected from 54 wells within the watershed, to investigate sources of groundwater. Oxygen-18 (δ^{18}O) values lower than -13.0 per mil were documented in the sampling, and likely indicate the presence of recharge water from the last glacial period (>9,500 years old) beneath the northern portion of the Reservation, in the vicinity of Odanah, Wisconsin.

Finally, a new data-worth analysis of potential new monitoring-well locations was performed by using the model. The relative worth of new measurements was evaluated based on their ability to increase confidence in model predictions of groundwater levels and base flows at 35 locations, under the condition of a proposed open-pit iron mine. Results of the new data-worth analysis, and other inputs and outputs from the Bad River model, are available through an online dynamic web mapping service at (http://wim.usgs.gov/BadRiver/).

Introduction

The Bad River Watershed drains approximately 1,000 square miles (mi^2) spanning Ashland, Iron and Bayfield counties of northwestern Wisconsin, extending from forested uplands along the continental divide to Lake Superior's southern shore (fig. 1). The Bad River Watershed encompasses most of the Bad River Reservation (Reservation), the cultural homeland of the Bad River Band of Lake Superior Chippewa. Near Lake Superior, the Bad River empties into the Bad River Sloughs, part of a wetland complex containing the largest remaining wild rice beds in the Great Lakes Basin and designated as a Ramsar Wetland of International Importance (www.ramsar.org). Near the headwaters, the Bad River Watershed also encompasses the western end of the informally named Gogebic Range, an approximately 80-mi-long ridge of steeply dipping Paleoproterozoic strata that extends eastward into Michigan's Upper Peninsula. The western end of the Gogebic, often referred to locally as the Penokee Hills (these terms will be used interchangeably in this report), contains undeveloped resources of taconite (low-grade iron ore) (Cannon and others, 2007).

The groundwater system in the Bad River Watershed provides the primary source of drinking water to its inhabitants and also sustains numerous cold and cool-water habitats. The hydrogeology of the area is complex and has not been well-studied to date. Understanding of the regional groundwater-flow system and groundwater/surface-water interactions is important to the Bad River Band and other stakeholders, as they seek to protect and manage the watershed in the face of climate change, a proposed open-pit iron mine in the Penokee Hills (Wisconsin Department of Natural Resources, 2014a), and other concerns.

Weidman and Schultz (1915) provide one of the earliest and most thorough descriptions of the study area hydrogeology, including flowing artesian wells, springs and findings from the construction of a 3,095-foot (ft)-deep well drilled near Ashland. More recent hydrogeologic studies include Batten and Lidwin (1995), who summarized hydrologic data collected between 1983 and 1987, and proposed a conceptual model for the regional flow system. Dunning (2005) studied the deep sandstone aquifer beneath the northwest corner of the Reservation in greater detail. Although these studies provide basic information on the extent and character of the major aquifers beneath the Bad River Watershed, they did not investigate flow paths or quantitatively characterize groundwater/surface-water interactions.

Groundwater-flow models provide a quantitative basis for investigating flow paths, groundwater/surface-water interactions, and water availability by formally integrating knowledge of the groundwater system (including aquifer and stream-network geometries, subsurface hydraulic properties, and measurements of water levels and flow) with the well-developed mathematical theory of
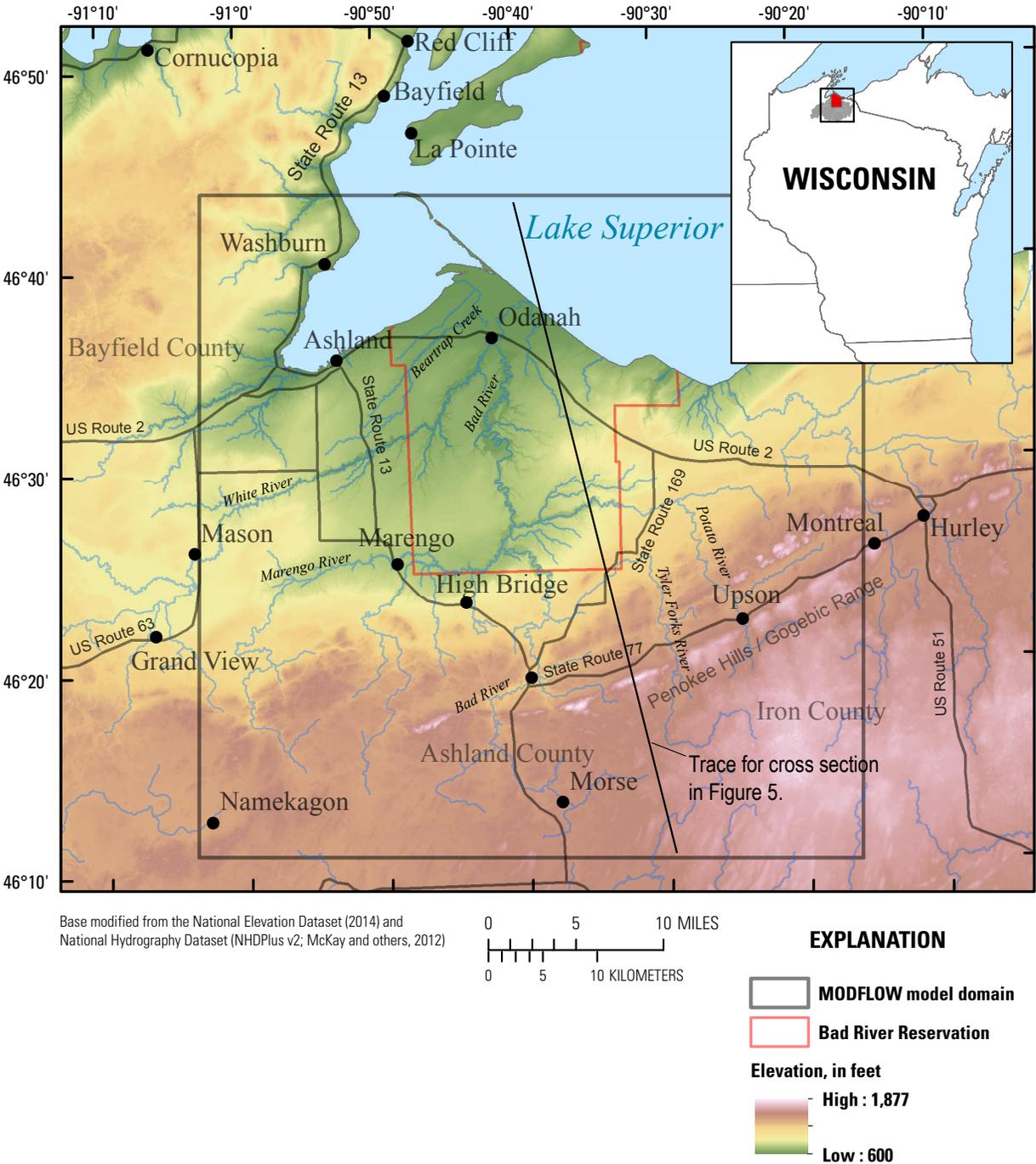
Figure 1.   Location of the study area.

4   Groundwater/Surface-Water Interactions in the Bad River Watershed, Wisconsin

groundwater flow. Investigation of uncertainty in groundwater-model inputs provides insight into where additional data are needed to better answer questions about the simulated system. Previously there were no published groundwater models for this area.

In this study, a groundwater-flow model was developed for the Bad River Watershed and surrounding area. The geographic extent of the model domain is shown in figure 1. The model synthesizes existing data, including information on the thicknesses, extent and hydraulic properties of the major hydrogeologic units, the distribution of soil types and land cover, historical records of precipitation and temperature, and historical measurements of groundwater levels and stream base flow. Additional field data were collected in support of this study, including stable isotope samples for evaluating sources of groundwater, and synoptic base-flow measurements for evaluating the spatial distribution of groundwater discharge to streams.

Purpose and Scope

This report presents the results of an investigation of the groundwater system and its interactions with surface water in the Bad River Watershed, including the development of a groundwater-flow model and the collection and interpretation of additional supporting data. The goals of this study were to (1) improve understanding of the groundwater-flow system in the Bad River Watershed, including the sources of Reservation water and groundwater/surface-water interactions; (2) provide a physically based platform for evaluating future impacts to the watershed, which can be used as a starting point for more detailed investigations at the local scale; and (3) identify areas where more data are needed.

The report includes (1) an overview of the hydrologic setting and conceptual model, (2) a brief description of construction and calibration of the groundwater model, (3) presentation and discussion of model results and applications, (4) discussion of existing data gaps, (5) application of the model to investigate additional areas for future data collection, (6) collection and interpretation of stable isotope data, and (7) discussion of limitations and assumptions. Appendixes are included to provide more detailed documentation of model construction and calibration approaches (appendixes 1–3), model calibration results (appendix 4), stable isotope sampling results (appendix 5) and new data-worth analysis (appendix 6).

Geologic History

The geology of the Bad River Watershed is complex, encompassing unconsolidated, sedimentary and crystalline units that were deposited and altered over 2.5 Ga (billion years) of earth history. The topography, bedrock geology, and drainage of the study area are influenced by the Midcontinent Rift, a 1,200-mile (mi)-long fissure that formed in the earth's crust approximately 1.1 Ga. Extension of the crust during rifting resulted in extensive volcanic activity and the creation of a large valley that was infilled by thick sequences of mafic igneous and sedimentary rock units. The historical rift valley and its deposits are together referred to as the Midcontinent Rift System (MRS; for example, Ojakangas and others, 1997). The geographic trace of the MRS extends in an arc northward from Kansas to Lake Superior, and then southward into Michigan. Most of the MRS deposits are deeply buried, except in the vicinity of Lake Superior, which is situated within the rift basin. The Bad River Watershed straddles the southern edge of MRS deposits (figs. 2 and 5). The strata in this area dip steeply north, exposing increasingly older units to the south. The rift deposits are bounded to the south by the Marquette Range Supergroup, which consists of metamorphosed sedimentary rocks deposited in a marine environment prior to the rift (approximately 1.9 Ga). The Marquette is in turn underlain by older igneous units (2.7 Ga) associated with the formation of an island arc in the Archean period (Cannon and others, 2007). The bedrock geology is described in this report following the nomenclature of Cannon and others (1999).

More recent deposits associated with glaciations during the Quaternary period drape MRS and older bedrock units, and exert an additional imprint on the study area topography and drainage (figs. 2 and 3). Most of these deposits are from the most recent glaciation, which occurred between approximately 20,000 and 10,000 years ago (Clayton, 1984). Given the large expanse of time between the deposition of the bedrock and glacial deposits, it is likely that the bedrock surface has undergone extensive weathering in places and also contains erosional features (such as buried valleys) that may have an effect on groundwater flow, but are not evident at the land surface. The Quaternary geology is described in this report following the nomenclature of Syverson and others (2011).
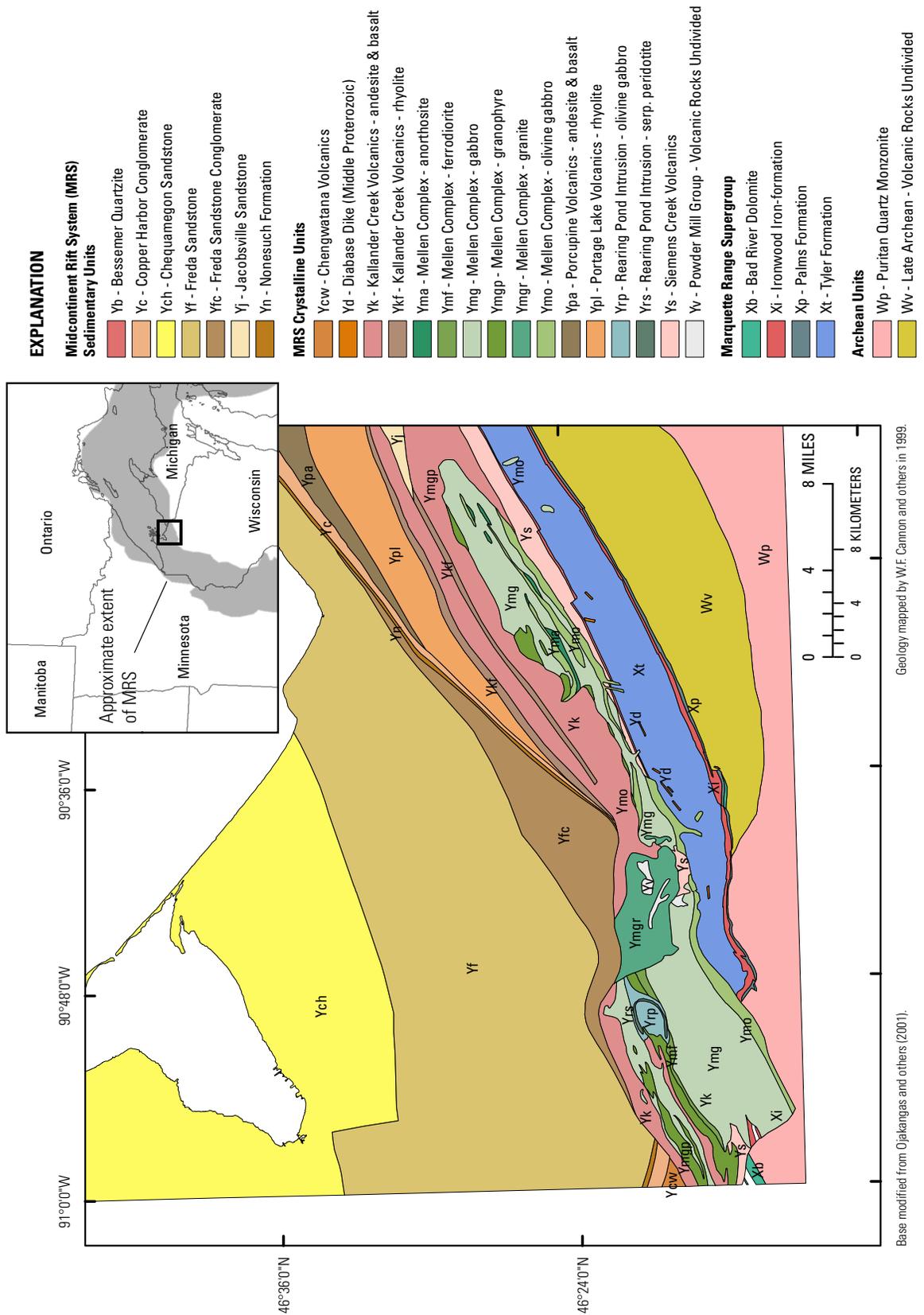
Figure 2. Bedrock geology in the study area.
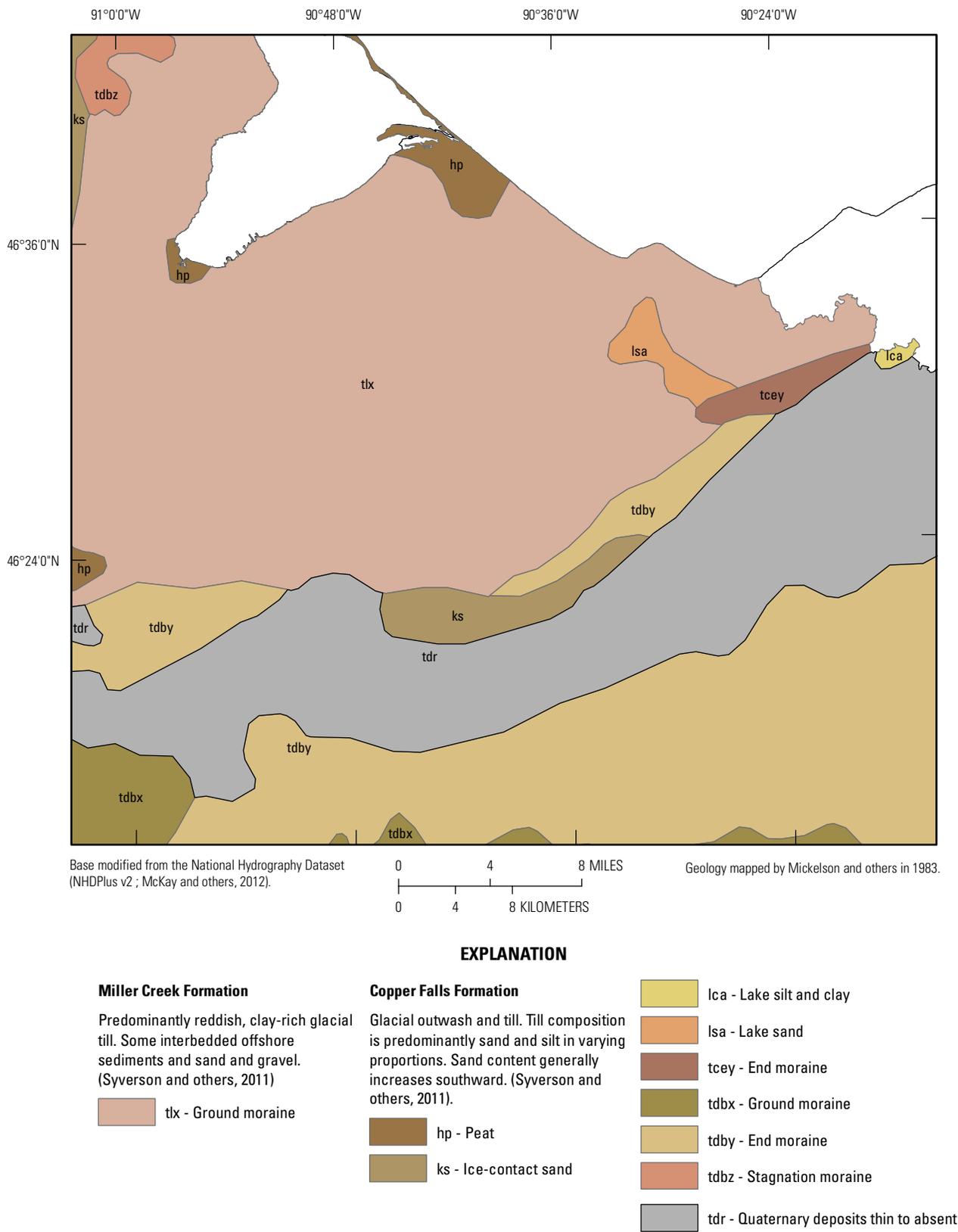
Figure 3.   Quaternary deposits in the study area.
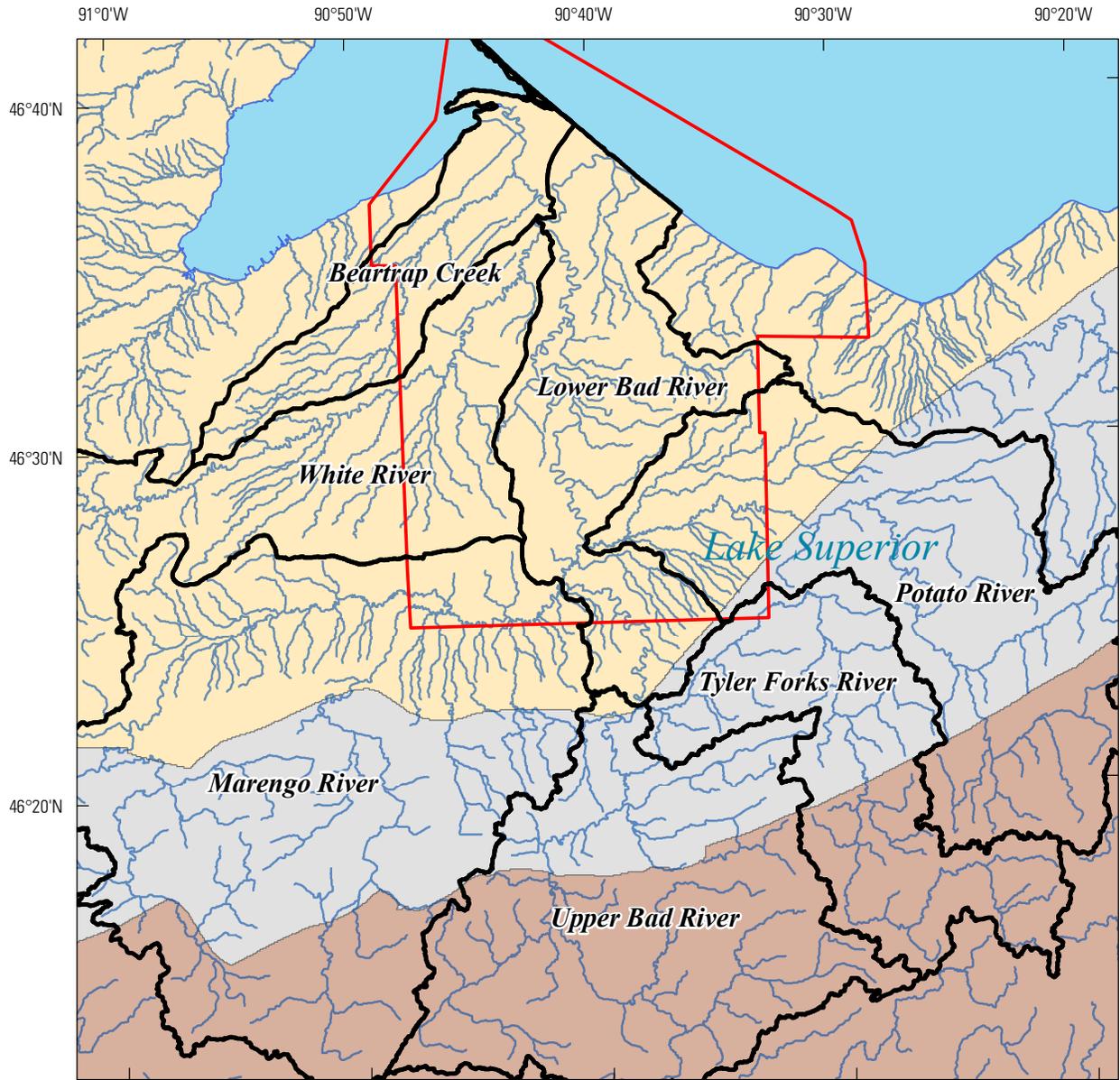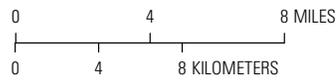
Figure 4.   Drainage in the study area with conceptual hydrologic regions based on the underlying geology.

Hydrogeologic Setting and Conceptual Model of the Flow System

Most water enters the groundwater system after infiltration through the land surface, although smaller amounts are contributed locally by leakage from losing stream reaches. Conversely, the vast majority of discharge from the subsurface system occurs as groundwater return to streams and wetlands, with a smaller amount (about 6 percent) discharging to Lake Superior and the smallest amount being captured by pumping (<1 percent). Groundwater-flow paths depend on the distribution of recharge, the locations of discharge features, and the water-bearing properties of the aquifer materials. The geologic history described above influences each of these factors.

Conceptually, the study area can be divided into three primary hydrologic regions (Wheeler and Bodette, 2011) associated with the underlying geology: (1) southern bedrock uplands underlain by Archean units; (2) a transitional area in the vicinity of ridges formed by the Marquette Range Supergroup and MRS deposits (of which the Penokee Hills is a subset); and (3) Superior Lowlands underlain by clay-rich till, which confines a deep sandstone aquifer. A conceptual discussion of the groundwater-flow systems in the three regions follows, and is illustrated in figure 5.

Southern Bedrock Uplands

The southern portion of the study area consists of poorly drained, relatively flat uplands underlain by Archean crystalline basement rock that predates the MRS. Depth to bedrock in this area ranges from approximately 100 ft to zero in areas near the Penokee ridge. Surficial deposits consist mostly of sandy tills and glacial outwash of the Copper Falls Formation (Clayton, 1984). The drainage pattern is primarily dendritic but also reflects geological drivers that resulted in a highly eroded bedrock surface overlain by recently deposited glacial till. Groundwater flow occurs largely in a single surficial aquifer consisting of shallow fracture networks in the Archean bedrock, and intergranular flow in the overlying sandy tills and outwash. Groundwater recharge is expected to be moderate to high for a humid temperate climate in areas where the water table is appreciably below the surface (that is, topographic highs). Lower recharge values are expected for wetland areas, where precipitation runs off or is stored on the surface. Groundwater-flow paths in this area are relatively short, with local recharge discharging to local streams. Base flows are low relative to

total streamflow, as indicated by lower base flow factors (Gebert and others 2007; see also appendix 3). This is likely due to limited groundwater storage capacity in the fractured bedrock and thin glacial deposits.

Transition Area

The Archean bedrock uplands are bounded to the north by the northeast-southwest-oriented ridge of the Gogebic range, which is formed by the steeply north-dipping, metasedimentary Palms and Ironwood Formations of the Marquette Range Supergroup (early Proterozoic; fig. 2). The Ironwood Formation contains a large undeveloped resource of taconite (low-grade iron ore) (Cannon and others, 2007). The Ironwood is overlain by the less resistant Tyler Formation, which forms a valley between the Gogebic/Penokee ridge and hills associated with resistant MRS igneous rocks to the north. Quaternary deposits in the transition area are thin to absent. Variable resistance to weathering in the bedrock units, as well as faulting and fracturing from past tectonic activity associated with the MRS and prior events, has contributed to a trellised or rectangular drainage pattern in this area. The stream patterns in this area may therefore provide insight into the geometries of fractures in the underlying crystalline units.

Groundwater flow in the transition area is likely dominated by connected fracture networks. The location, extent, and connectivity of these fracture networks are poorly known. The spatial distribution of aquifer transmissivity provided by these fracture networks is an important consideration for evaluating dewatering effects in the bedrock. Conditions in the transition area, including lake-effect precipitation, B-soil types, and mixed forest land cover, are favorable to high infiltration rates (see figs. 2.3–2.8 in appendix 2). However, a substantial portion of infiltration may be rejected, due to thin soil cover and limited aquifer storage. Similar to the southern bedrock uplands, much of the groundwater flow in the transition area is likely local, from local topographic highs to nearby perennial streams. Longer flow paths may occur in the part of the groundwater system sloping down towards the Potato River, where many of the tributaries run parallel to the slope in the water table. Stream base flows in the transition area are generally low relative to total streamflow, likely due to limited groundwater storage and low average aquifer transmissivities. However, numerous springs and perennial headwater streams are known to exist, probably as a result of the variable surface topography and preferential flow through fractures (Weidman and Schultz, 1915).
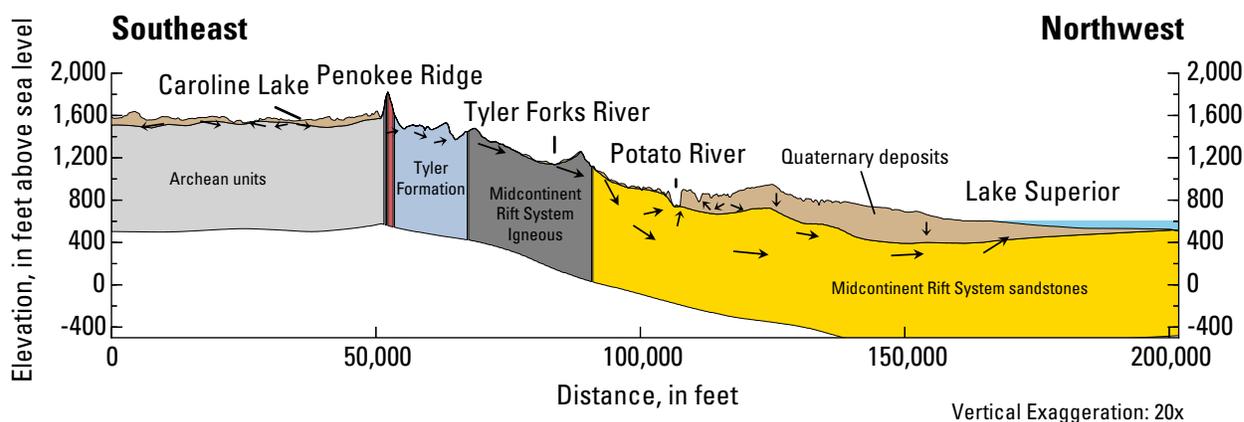
Figure 5.   Conceptual model of the groundwater flow system.

Superior Lowlands

The Superior Lowlands make up the area downslope, as part of the larger topographic trend, from the transition area. Bedrock geology north of the MRS igneous units consists of rift-fill sandstones and conglomerates that are steeply dipping near the igneous units, and subhorizontal in the vicinity of Lake Superior. The thickness of the sedimentary sequence increases northward for at least 3 mi beyond the Lake Superior shoreline (Cannon and others, 1996). The sandstones are overlain by up to several hundred feet of surficial deposits from the Quaternary period, primarily the clay-rich tills of the Miller Creek formation, which were derived from offshore sediments deposited by proglacial lakes occupying the Superior Basin. Sandy shoreline deposits associated with the most recent of these lakes ring the upper margins of the Superior Lowlands at elevations ranging from approximately 890 to 1,080 ft. Additional sandy shoreline and glacial meltwater deposits are interbedded within the clayey tills, as well as outwash deposits beneath the Miller Creek that correlate with the Copper Falls Formation, but none are well characterized or mapped (Clayton, 1984).

Localized surficial aquifers exist in the clays; however, the majority of aquifer transmissivity in the Superior Lowlands is likely attributable to the sandstone units. The sandstone aquifer provides a large reservoir of groundwater storage, which is confined in many areas beneath the Miller Creek clays, and extends to unknown but variable depths, where the quality of the water transitions to saline (for example, Young and Skinner, 1974; Weidman and Schultz, 1915). Bulk recharge rates through the clays are relatively low—on the order of inches per year (in/yr) or less. Higher rates of recharge occur in areas where sand bodies provide a connection between the sandstone aquifer and the land surface. The previously mentioned sandy shoreline deposits may

provide such a connection.

Recharge at higher elevations and the confining effects of the clays create artesian conditions, allowing for naturally flowing wells in many areas. Likewise, larger rivers in valleys that perforate the clay (such as the White and Marengo shown in fig. 4) have high rates of base flow (for example, approximately 70 percent of total flow for the White River) that are sustained through dry periods. In contrast, smaller streams that are underlain by clay are flashy with low levels of base flow (for example, approximately 10 percent of total flow for Beartrap Creek). Low recharge rates through the clays, and large distances between penetrating rivers allow for long flow paths in the sandstone aquifer, and long residence times in the northern portions of the groundwater system.

Water Use

The estimated population of the model area is 19,361 persons (based on 2010 census data; Wisconsin Department of Administration-Demographic Services Center, 2010). Of that amount, 21 percent (or 4,158 persons) use approximately 0.39 million gallons per day (Mgal/d) of publicly supplied groundwater, which includes some minimal business and industrial use. This also includes approximately 0.056 Mgal/d of publicly supplied groundwater from 13 wells on the Reservation (table 1). About 36 percent of the model area's population (6,987 persons) obtains their water from private supplies, mostly private wells. Assuming a per capita domestic use of 49 gallons per day (gal/d), this amounts to an estimated 0.34 Mgal/d. The City of Ashland accounts for 42 percent (8,216 persons) of the model area population and withdraws about 0.64 Mgal/d of surface water from Lake Superior.

Eight municipalities within the model area use groundwater, including the cities of Mellen and Washburn, limited service within the towns of Barksdale and Knight (which includes Iron Belt) and four communities within the Reservation (Birch Hill, Diaperville (Old Odanah), New Odanah, and Frank's Field). The Odanah water system, which serves the Bad River Lodge & Casino and Health Service Department, is the largest community water system on the Reservation. The remainder of the model area is served by private domestic supply wells and approximately 50 other low capacity, noncommunity wells associated with places such as churches, eating and drinking establishments, parks, motels, and campgrounds.

Estimates of groundwater pumping from high-capacity wells were developed for the model area based on pumping records and the methods described by Buchwald (2009). For some water users, directly reported values are available. However, other water uses must be imputed from available data and converted into water use values. The minimum reported water user is defined as a "property." A high-capacity property is defined by the WDNR as capable of withdrawing more than 70 gallons per minute (gal/min) from its combined sources of water supply. Only 26 wells within the model domain were associated with properties that meet this definition. Average pumping rates were developed for the periods of 1990–2010 and 2011–13. The 1990–2010 period, which corresponds to the bulk of groundwater-level observations in the calibration dataset, was used to calibrate the model to be consistent with the time period over which observations of heads and flows were measured. The 2011–13 rates are included in the final version of the model. Pumping records were not available for wells on the Reservation, so rates for Reservation community wells were estimated by population, assuming a per-capita water use of 49 gal/d (Buchwald, 2009). Three other non-community Reservation wells were each assigned default rates of 1 million gallons per year (Mgal/yr) (366 cubic feet per day [ft^3/d]).

Estimated pumping rates for the 26 high-capacity wells represented in the model are shown in table 1. The rates shown are those that were used in the model. Withdrawals from smaller wells were not considered, under the assumption that they are negligible at the scale of the model and return most water through on-site septic systems (resulting in little to no net withdrawal). Net high-capacity groundwater withdrawals within the model domain are estimated to total approximately 0.43 Mgal/d, or about 0.66 cubic feet per second (ft^3/s), a flow rate roughly similar to that of a first-order stream. This amount of pumping represents approximately 0.1 percent of total recharge, and has a minimal effect on the regional groundwater-flow system.

Groundwater/Surface Water Modeling Approach

A finite-difference groundwater-flow model of the study area was constructed using MODFLOW with the Newton Raphson solver (MODFLOW–NWT: Niswonger and others, 2011). A summary of the construction and calibration of the MODFLOW model, including the underlying data sources, is given in the Construction section. A more detailed description of the model can be found in appendixes 1–3.

Construction

The MODFLOW grid consists of 800 rows and 800 columns of uniform cells that are 250 ft on a side. Two layers are used to represent the surficial Quaternary deposits, and three layers are used to represent bedrock, extending downward to a depth of approximately 1000 ft below the bedrock surface. Hydraulic conductivity (K) was specified by using a network of 1,352 horizontal K and 1,352 collocated vertical K pilot points (Doherty, 2003) grouped into 13 zones representing the major hydrogeologic units (figs. 2.12 and 2.13). Hydraulic conductivity values at individual model cells were interpolated from nearby pilot points in the same zone and same layer via kriging. Pilot points were spaced every 44 cells, with additional pilot points added to the thin zones representing the Ironwood and Palms formations (see appendix 2). Precalibration values for the pilot points were based on geometric mean hydraulic conductivity values for the major hydrogeologic units in the model domain, estimated by K. Bradbury (Wisconsin Geological and Natural History Survey, unpub. data, 2013), by using specific capacity data from well construction reports (table 4.1).

Steady-state recharge to the model was estimated for each cell, using the Soil Water Balance code (SWB) (Westenbroek and others, 2010). SWB calculates net deep drainage, accounting for infiltration and evapotranspiration through the root zone. At this regional scale, all deep drainage is assumed to enter the groundwater as recharge. To obtain representative steady-state values, recharge was simulated for the period of 1980–2011, on daily timesteps, using gridded temperature and precipitation estimates, and GIS coverages of land surface aspect, available soil water capacity, land cover and soil hydrologic group as inputs (see appendix 2). A single grid of temporally averaged values was then computed from the transient SWB results. Rejected recharge in areas where the water table is near or at the land surface was accounted for with the Unsaturated Zone Flow (UZF) package (Niswonger and others, 2006), by routing the excess water to the stream network. The

Table 1. High-capacity wells represented in the model.

[ft³/d, cubic feet per day; Mgal/d, million gallons per day; Mgal/year, million gallons per year; WDNR, Wisconsin Department of Natural Resources. Average pumping rates for the period 1990–2010 were used to calibrate the model. This time period coincided with the majority of water-level observations; Current pumping rates in the model are averages for the period 2011–13]

Entity	Type	Number of representative wells in model	Calibration pumping rate (ft³/d)[1,3]	Current pumping rate (ft³/d)[2,3]	Current pumping rate (Mgal/d)[3]
Hayward Bait & Tackle aquaculture facility	Commercial / Industrial	3	0	1,831	0.0137
			0	1,831	0.0137
			0	1,818	0.0136
Milestone Materials	Commercial / Industrial	1	0	243	0.0018
Iron Belt Sanitary District	Municipal supply	2	1,402	1,255	0.0094
			1,033	1,258	0.0094
City of Washburn	Municipal supply	2	12,156	10,049	0.0752
			12,156	10,129	0.0758
City of Mellen	Municipal supply	2	1,060	0	0.0000
			19,852	21,367	0.1598
Northern States Power Company	Commercial / Industrial	1	0	123	0.0009
Copper Falls State Park (WDNR)	Other	3	1	1	0.0000
			32	43	0.0003
			34	45	0.0003
Birch Hill (Bad River Reservation)	Municipal supply	2	960	960	0.0072
			960	960	0.0072
Diaperville (Old Odanah; Bad River Reservation)	Municipal supply	2	117	117	0.0009
			117	117	0.0009
New Odanah (Bad River Reservation)	Municipal supply	4	794	794	0.0059
			794	794	0.0059
			794	794	0.0059
			199	794	0.0059
Frank's Field (Bad River Reservation)	Municipal supply	2	504	504	0.0038
			378	504	0.0038
Bad River Reservation—unspecified	Other	1	183	366	0.0027
Bad River Reservation Recycling center	Other	1	183	366	0.0027
Bad River Reservation—Old School	Other	1	366	366	0.0027

[1]Calibration pumping rates based on 1990–2010.
[2]Current pumping rates based on 2011–2013.
[3]It should be noted that the number of digits displayed does not reflect measurement precision. For example, the values of 366 ft³/d for the three non-community Reservation wells are the result of a unit conversion of 1 Mgal/yr to ft³/d. Precision varies for the metered municipal wells, depending on the volume metered, and how that volume was reported. For consistency with the model, the values are shown here as they were entered into the model, rounded to the nearest ft³/d. Two significant figures are used when discussing the high-capacity pumping totals in the report text, for consistency with other large-scale water use numbers that are reported to two digits, such as the estimated total pumpage by private wells.

Totals:	Current pumping rate (Mgal/d)	Percentage of total
Total commerical and industrial	0.0437	10.2
Total municipal (off reservation)	0.3296	76.7
Copper Falls State Park	0.0007	0.2
Total Reservation high-capacity pumpage	0.0556	13.0
Total pumpage from high-capacity wells	**0.4296**	

entire recharge array generated by SWB was allowed to vary uniformly during calibration of the MODFLOW model by using a single recharge multiplier to achieve the best match between measured baseflows in streams, and those simulated by the model. The calibrated recharge field is shown in figure 12 with low values over the Superior Lowlands and highly variable values throughout the rest of the domain.

Other boundary conditions were simulated as follows:

- Streams and connected inland lakes were represented as head-dependent flux boundaries with the Streamflow Routing (SFR2) package (Niswonger and Prudic, 2005), which tracks base flow accumulated from stream leakage. Input to the SFR2 package, including routing, stream geometry, and elevations for each model cell, was developed from the National Hydrography Dataset (NHDPlus v2; McKay and others, 2012) by using an automated Python program.

- Lake Superior was represented as a head-dependent flux boundary using the General Head Boundary (GHB) package.

- Regional groundwater flow across the MODFLOW grid perimeter boundary (fig. 4) is accounted for with constant flux boundary cells, which were developed by using an analytic element model that covers a larger area (see appendix 1).

- Pumping from high-capacity wells (those withdrawing greater than 70 gpm) was simulated with the Multi-node Well (MNW2) package.

The model was solved using the MODFLOW–NWT solver (Niswonger and others, 2011; Hunt and Feinstein, 2012), which offers improved handling of solution nonlinearities associated with the simulation of the transition of wet-dry cells in unconfined flow and groundwater/surface-water interactions (Hunt and Feinstein, 2012), making it ideal for this study.

Calibration

A calibration dataset was developed to compare steady-state model outputs with field measurements of the system. Historical water-level measurements were obtained from Wisconsin Department of Natural Resources well construction reports (WCRs), the National Water Information System (NWIS), and the Bad River Band. Where present, multiple measurements of head were averaged to develop a single, steady-state value. Stream base flows were obtained from NWIS, annual average values published in Gebert and others (2011), and measurements collected for this study. Average annual

base flows were computed from continuous records for the USGS gages 04027000 (Bad River near Odanah, Wisconsin), 04026561 (Tyler Forks River at Stricker Road), and 04026390 (Beartrap Creek at Highway 2), by using the Institute of Hydrology (1980a; 1980b) method implemented in the base flow Index (BFI) software (Wahl and Wahl, 2013). Base flow measurements collected for this study were adjusted to be consistent with the average annual values published in Gebert and others (2011), and those estimated by base-flow separation. This adjustment is necessary to make single-time measurements representative of long-term conditions. For example, current base-flow measurements may be biased due to short-term climate fluctuations. The measurements were first adjusted by using an empirical relationship developed for Wisconsin (Gebert and others 2007, 2011), and then additional correction factors were applied for the Upper Bad and Tyler Forks/Potato Basins, based on an apparent bias in these areas attributed to the recent drought, and shallow, fractured-rock aquifers (see appendix 3).

The calibration was completed by using the parameter-estimation computer code (PEST, Doherty, 2010a, 2014). For parameter estimation, observations in the calibration dataset were weighted to address differences in information content and measurement uncertainty related to measurement quality, location uncertainty, and temporal variability. The overall goal of the weighting was to maximize the transfer of information from the observation dataset to the estimated parameters (for example, Doherty and Hunt, 2010). Initial weights were assigned to be inversely proportional to estimated measurement error associated with the observations and the application to steady-state model calibration. Observations were then grouped into heads and base flows, and also by information source and quality. Additional weight multipliers were assigned to the base-flow groups to ensure similar influence on the calibration process by both heads and base flows. The locations of calibration data sets are shown in figure 6, with the points representing the observations sized according to their relative weight. A total of 1,224 weighted observations were included in the calibration process; weighting for the observation groups is given in appendix table 3.1.

The model was calibrated to the weighted observations by automated adjustment of model inputs until a satisfactory level of fit between the model outputs and field observations was obtained while also honoring prior knowledge about field conditions as expressed through parameter values. Following the guidelines of Doherty and Hunt (2010), the calibration used SVD–Assist, Tikhonov (preferred homogeneity) regularization, and singular value decomposition (see appendix 3). A number of K pilot points were tied during

parameter estimation to protect against extreme values of overfitting and to reduce the computational burden. Of the 2,710 possible calibration parameters, a total of 1,715 were adjusted during the calibration process. Parameters that were adjusted during calibration included horizontal hydraulic conductivity, vertical anisotropy for a subset of the pilot points, global vertical conductivity (SFR input) and conductance (GHB cells), and a global multiplier on the recharge array.

Calibration Results and Discussion

Observed heads and base flows were well simulated by the calibrated model (figs. 7–11), with reasonable parameter values. Table 2 shows metrics of calibration performance for the various observation groups that are described in appendix 4. Root mean squared error (RMSE) is calculated as

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n} \right)} \qquad (1)$$

where

n is the number of observations in a group;

$\hat{y}_i$ is the modeled value of the ith observation;

y_i is the ith measured observation value.

The mean absolute error is calculated as

$$MAE = \frac{1}{n} \left| \sum_{i=1}^{n} (\hat{y}_i - y_i) \right|. \qquad (2)$$

In addition to these metrics, the spatial distribution of residuals is depicted in figures 10 and 11.

The calibrated hydraulic conductivity fields are within the ranges of hydraulic conductivity values estimated by K. Bradbury (Wisconsin Geological and Natural History Survey, unpub. data, 2013), for wells in the simulated units (appendix table 4.1 and appendix figures 4.2?4.5). Summary statistics of final hydraulic conductivity parameter values are compared to the precalibration parameter values in Table 4.1 in appendix 4. The calibrated multiplier for the recharge array was 1.24, which translates to a calibrated range of spatial values spanning from 0 to 38.5 in/yr, and an areal average steady-state rate of 6.7 in/yr.

Parameter identifiability (Doherty and Hunt 2009) is a qualitative statistic from parameter estimation that reflects how well a given parameter can be constrained by a given or hypothetical observation dataset. Identifiability values

indicate that many of the pilot points, especially those near the major highways and population centers in the model, were informed by observations in the calibration dataset. The lowest identifiabilities occur in the areas of the model domain that have few wells, and in the pilot point parameters that represent vertical hydraulic conductivity (appendix fig. 4.1).

Figures 10 and 11 show the spatial distribution of residuals (measured-simulated, as reported by PEST) in heads and base flows across the model domain. Negative residuals indicate simulated values higher than the observed equivalent. The positive and negative residuals are generally evenly distributed throughout the model domain, although in some areas of higher topographic relief and shallow depth to bedrock—especially in the vicinity of the MRS and Penokee ridges—the model may be biased towards positive residuals (simulated values lower than observed). This might indicate aquifer-system depths that are shallower in reality than represented in the model. Local highs in the bedrock topography or locally shallow fracture networks may exist that were not captured in the bedrock surface used in the model. The magnitudes of the residuals in these areas are also generally higher than elsewhere in the model domain, possibly due to localized variability in the depth and permeability of bedrock fracture networks or in the unconsolidated deposits draping the bedrock surface that is absent from or incorrectly represented in the model.

In addition to potential simulation artifacts, the calibration also reflects the quality of observation data. Almost all of the head observations in the vicinity of the MRS and Penokee ridges are single measurements obtained from WCRs, which may be poor representatives of long-term average conditions due to considerable variability in water levels owing to the low storage in fractured rock aquifers. Improving the model in these areas would require the collection of additional field data to support refinement of the bedrock surface and hydraulic conductivity parameterization, and the collection of higher quality, vertically distributed, long-term head data.

Similarly, the largest residuals between simulated and observed fluxes are also likely due to temporal variability and data quality issues. Many of the base-flow observations in the calibration dataset were one-time measurements collected during the summer of 2011, the end of a multiyear period of below-average recharge (see appendix fig. 2.9). As a result, base-flow observations are likely biased low relative to average annual values such as those published in Gebert and others (2011). Adjustments made to these measurements (described in appendix 3) were intended to mitigate the adverse effects of this bias at the subbasin scale. However, individual measurements at different locations in their respective subbasins may have been affected differently by the drought. Therefore these
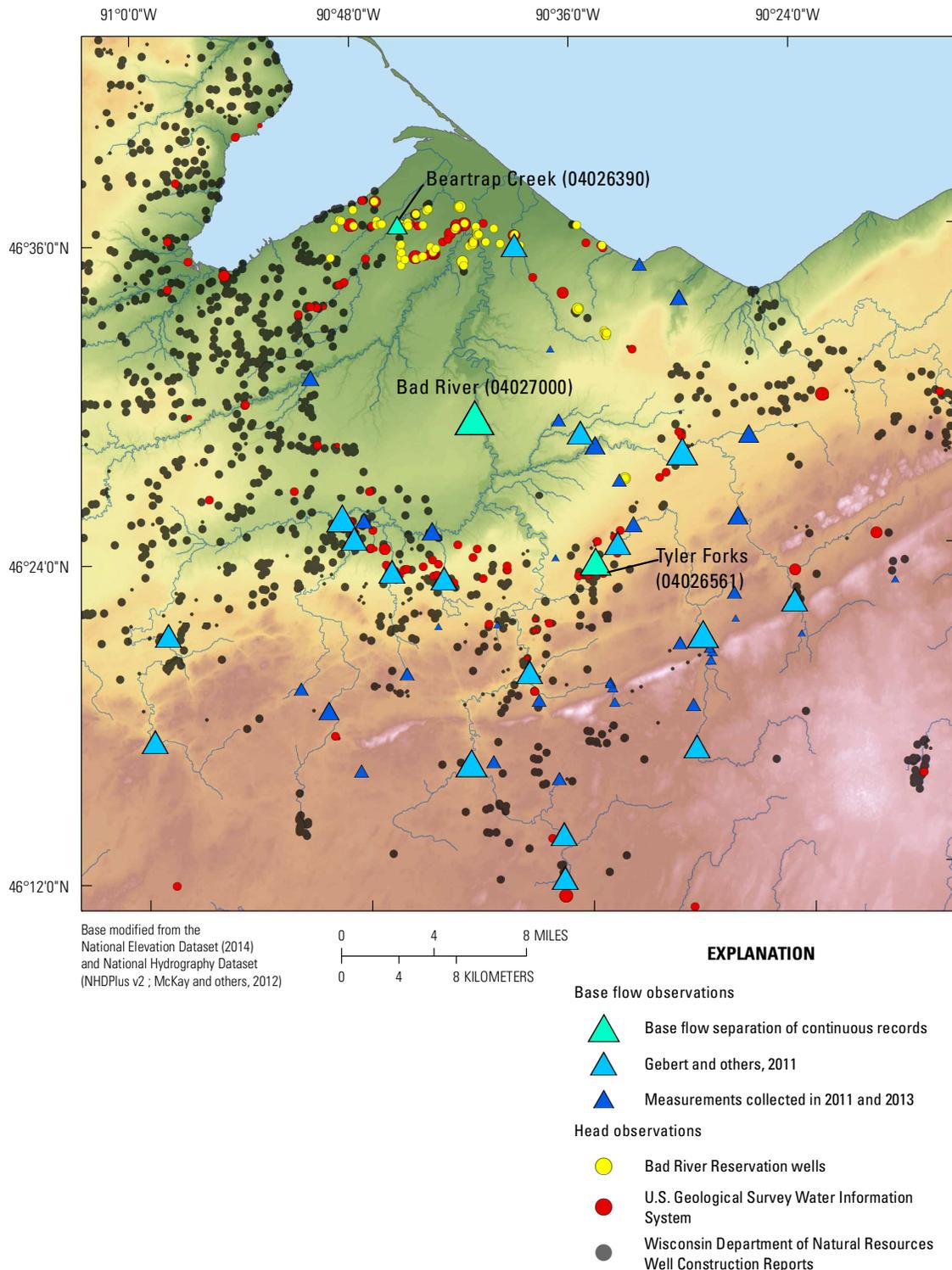
Figure 6. Observations used for model calibration. The markers representing the head and base-flow observations are sized proportional to the observation weights, relative to other observations in the respective categories. Weights were generally assigned to be inversely proportional to the estimated uncertainty in the observation. The smallest markers represent 231 head observations that were assigned weights of 0 (no influence on the calibration process), due to large uncertainties in their reported wellhead elevations. In general, the WCRs were the least reliable source of head data, due to uncertainties in their locations and wellhead elevations, and other factors including potential disequilibrium of water levels in the well following drilling. Similarly, single base-flow measurements, especially those for small streams, have higher uncertainties due to the error inherent in their conversion to average annual values (see discussion in appendix 3).

"observed" values are subject to error stemming not only from the streamgaging process, but also from estimation of average annual flows from single values.

Model Results and Discussion

The simulation of the groundwater system produces quantifiable values of inflows and outflows based on horizontal and vertical head gradients and stresses imposed within the model. Table 3 shows the simulated distribution of those inflows and outflows in the model. As would be expected, areal recharge from infiltration at the land surface is the dominant contributor of water (83.1 percent), with lesser amounts leaking from streams (12.7 percent). Most water leaves the groundwater system as base flow in streams (72.2 percent, much of it ultimately through the Bad River, which carries the bulk of the total flow shown in figure 16), with a minor amount of deep flow discharging to Lake Superior (5.7 percent). A discussion of the simulated spatial distribution of flows within these components follows.

Recharge

Estimated recharge values in each model cell (fig. 12), following adjustment by a multiplier during calibration of the MODFLOW model, was consistent with other studies in the area (for example, Lenz and others (2003), Gebert and others, 2011), and results in simulated base flows that correspond closely with the observed base flows (fig. 8). The spatial distribution of estimated recharge reflects the combined influence of factors including precipitation, soil type, and land use. Influence of soil type and precipitation is especially apparent, with high recharge values occurring in sandier (A- and B-type) soils at higher elevations, especially on Birch Hill, the Bayfield Peninsula, and in the vicinity of the Penokee ridge and MRS hills, which in addition to sandier soils have higher precipitation, including lake-effect snow (appendix figs. 2.3, 2.4, and 2.8). In the Superior Lowlands, localized areas of higher recharge are also apparent in areas mapped as having sandier soils.

Groundwater Flow

Steady-state (long-term average) water-table elevations from the calibrated model are shown in figure 13. The black lines are contours of equal water-table elevation at 20-ft intervals. Groundwater flow occurs in the direction of steepest decent, perpendicular to the equal elevation contour lines. Simulated groundwater-flow directions and magnitudes (normalized to the size of the

model cells), for every 20th cell in model Layer 4, which corresponds to the deeper groundwater-flow systems, are shown in figure 14. Together, these two figures show smaller, localized flow paths in the southern and eastern parts of the model domain (indicated by the intricate contour patterns and scattered arrow directions), where the aquifers are shallow, and intersected by many streams. Flow paths in the deep sandstone aquifer beneath the Superior Lowlands are longer, extending to regional discharge areas in the major river valleys and Lake Superior.

These phenomena are also illustrated in figure 15, which shows vertical differences in simulated heads between model layers. The blue areas in this figure symbolize areas of relatively strong upward flow of groundwater towards discharge features. The red areas symbolize areas of strong downward flow from local recharge areas. Areas of white indicate model cells that are dry (that is, above the simulated water table) in one or both of the layers being compared. In general, the southern part of the model domain has a greater local density of recharge and discharge areas, supporting the concept of small flow systems in this area. Larger swaths of blue (upward flow) in the major river valleys of the Superior Lowlands and along the shoreline of Lake Superior illustrate regional discharge in these areas.

Groundwater/Surface-Water Interactions

Simulated interactions between the groundwater system and streams are shown in figures 16 and 17. Together these figures show where the groundwater system is discharging water to streams, and where streams lose water back to the groundwater system. In general, the southern portion of the watershed, which has a shallow, unconfined aquifer system with a high water table, has a greater portion of perennial headwater streams. Approximately 50 percent of the stream cells representing first-order (headwater) streams in this area were simulated as having no base flow. Conversely, in the Superior Lowlands, where small streams are isolated from the sandstone aquifer by the clay-rich till of the Miller Creek Formation, only 25 percent of first-order stream cells were simulated as having base flow. These results are consistent with the perennial and ephemeral stream classifications in the NHDPlus v2 database (see appendix 2). Many of the perennial streams in the Superior Lowlands are in the vicinity of Marengo and Potato Rivers, which are situated near the southern margin of the Miller Creek Formation, where the clay is thin, and immediately adjacent to recharge areas at higher elevations. In the Superior Lowlands, the thick clay with few streams fully penetrating it contributes to artesian conditions in the
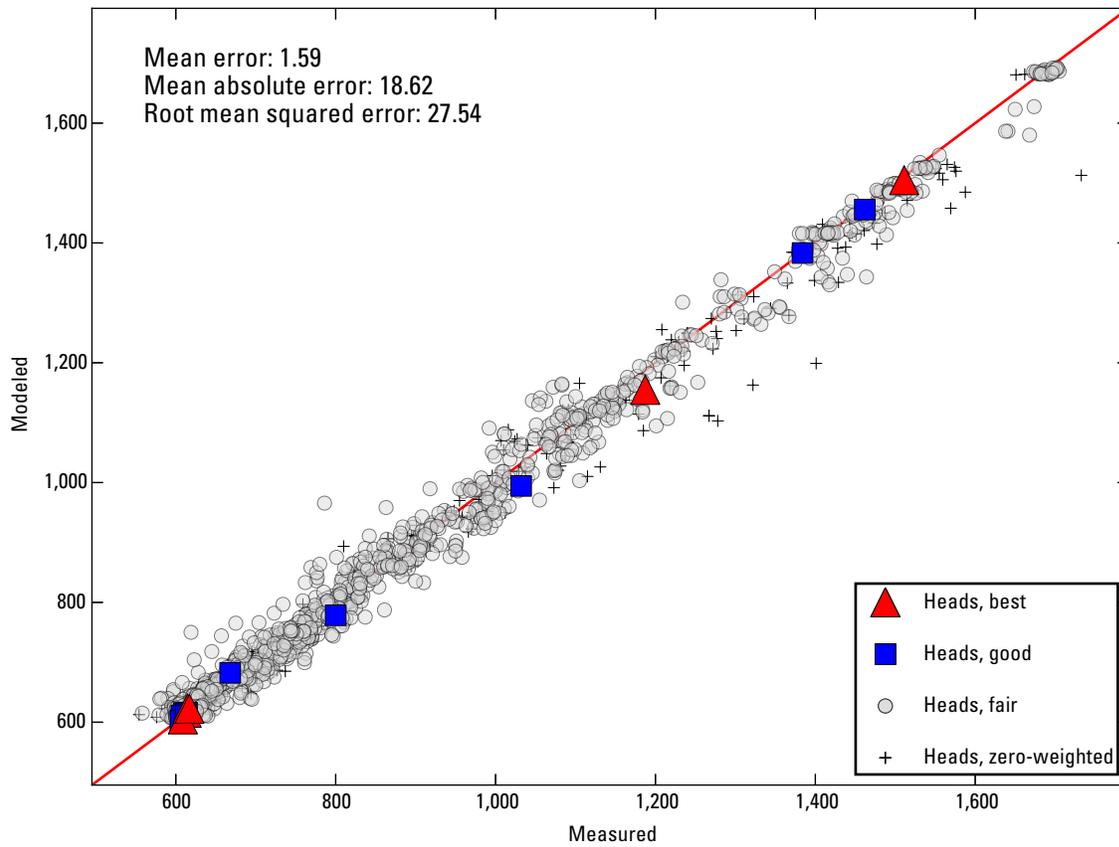
Figure 7.   Comparison of simulated versus observed heads. All units are in feet. Diagonal red 1:1 line indicates perfect fit. The group definitions of best, good, fair, and zero-weighted are explained in appendix 2.
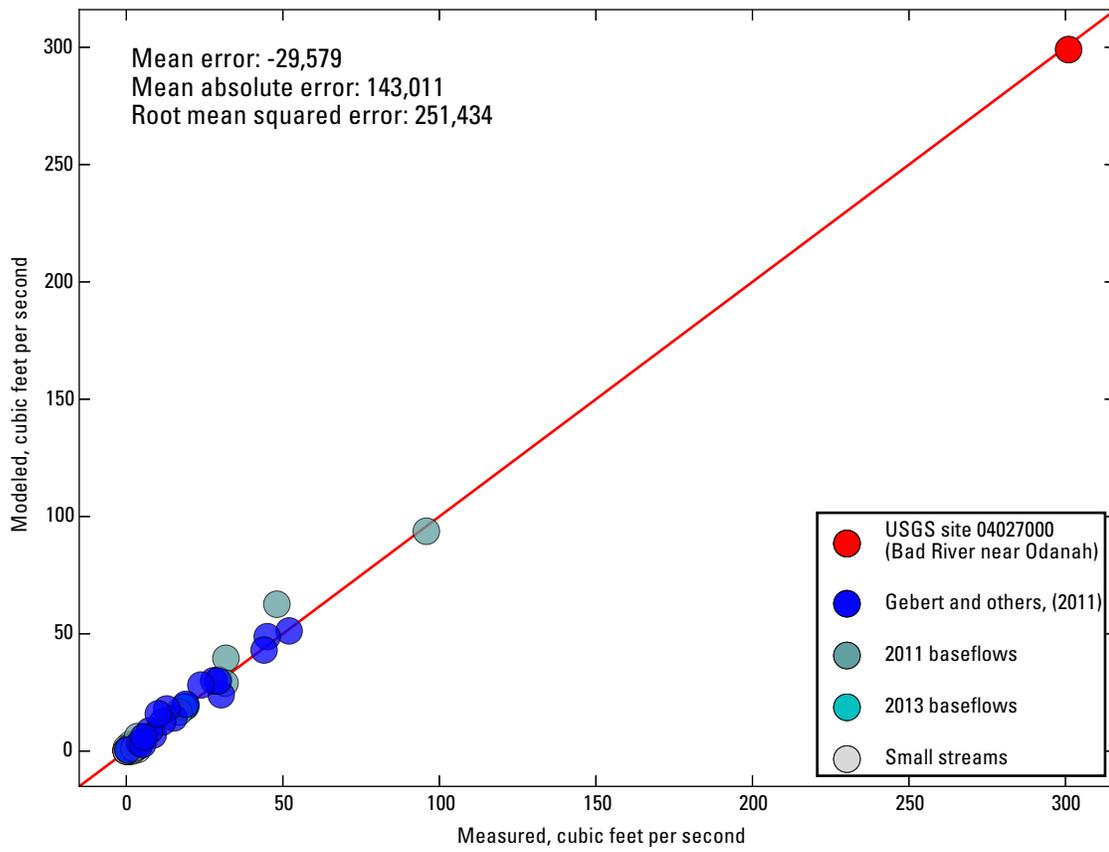
Figure 8.   Comparison of simulated versus observed fluxes. Diagonal red 1:1 line indicates perfect fit.
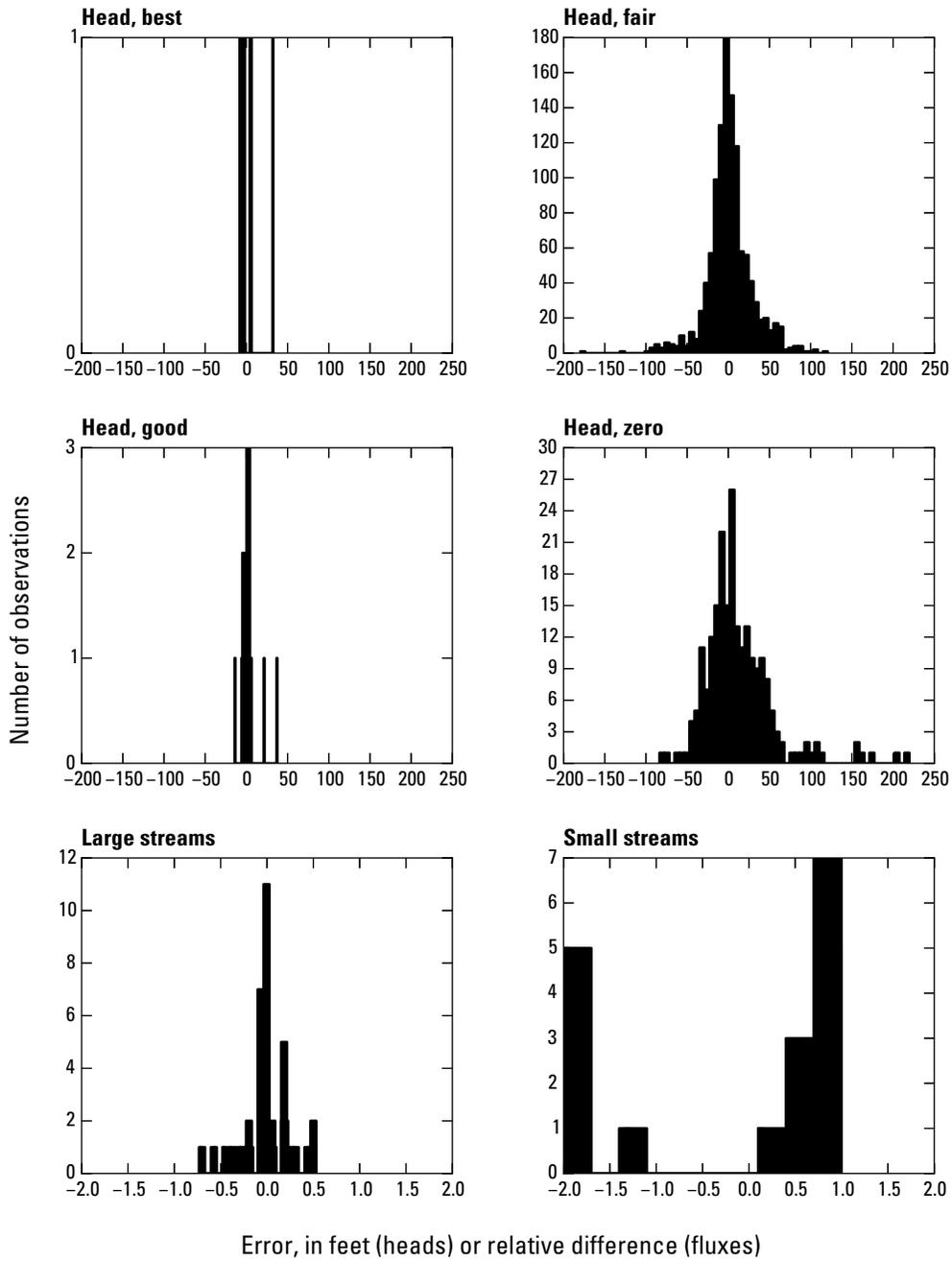
Figure 9.   Histograms of all head and flux group calibration results showing the general central tendency.
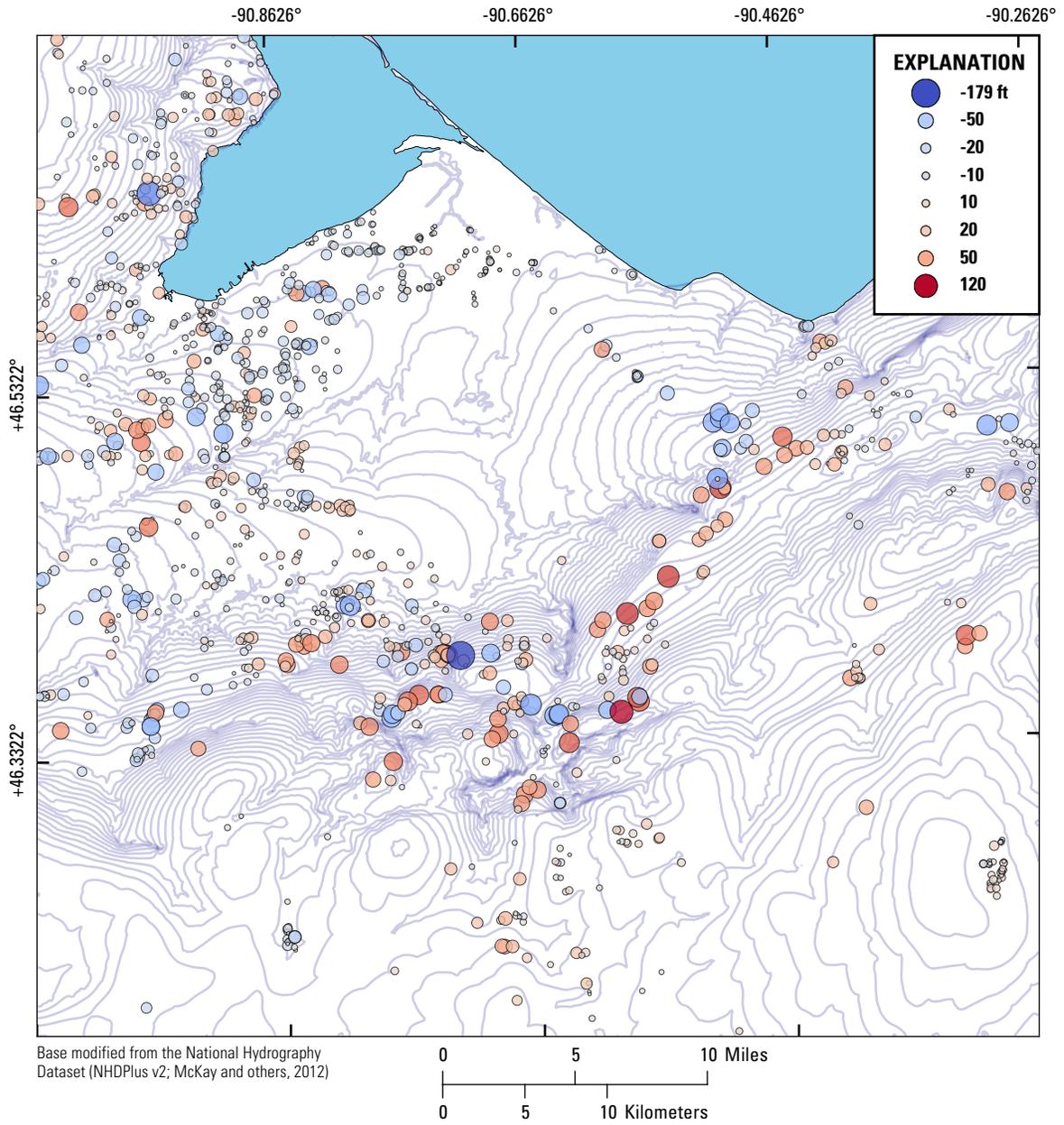
Figure 10.   Spatial distribution of head residuals (measured versus simulated) for the calibrated model. Blue indicates simulated values that are higher than observed; residuals are shown in feet.
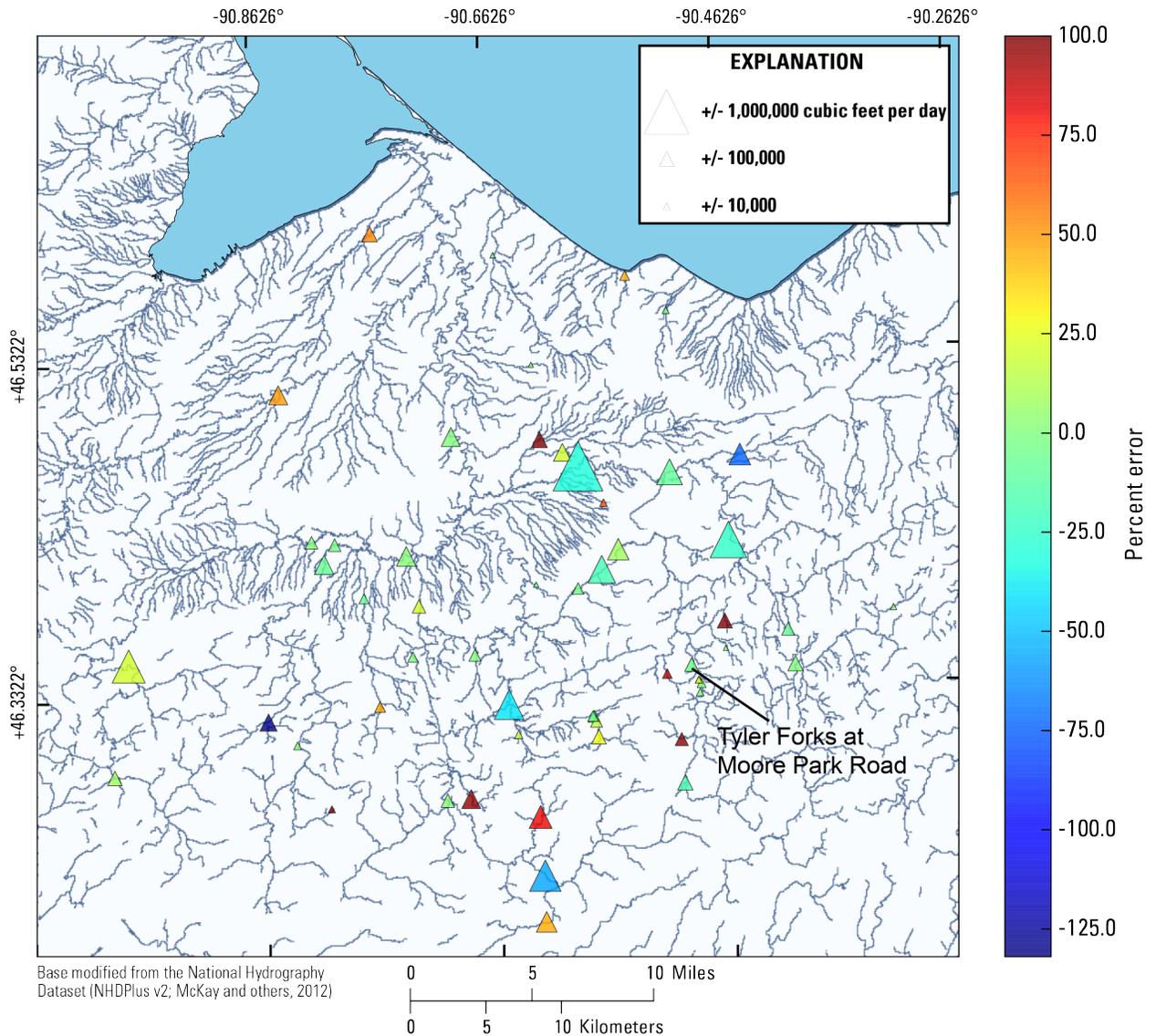
Figure 11.   Spatial distribution of flux residuals (measured versus simulated) for the calibrated model. The colors indicate the relative degree of mismatch between the simulated and observed values (as a percent difference, with negative values indicating simulated values that are higher than observed), while the size of the triangles indicates the absolute quantity of the flux error.

Table 2.   Metrics of calibration performance.

[NWIS, National Water Information System; --, not applicable]

Type	Data Source	Quality group	Estimated uncertainty	Group weight multiplier	Observation weights	Number of observations in group	Phi contribution
Heads	NWIS	Best	5 (feet)	1	0.200	6	169.1
Heads	NWIS and Bad River Band	Good	15	1	0.067	17	80.4
Heads	All sources	Fair	35	1	0.029	1,144	1,618.4
Heads	All sources	Zero-weighted	--	--		231	0.0
Base flows	NWIS	Bad River near Odanah	0.01 (CV[1])	3	3 / (0.01 x flow)	1	7.2
Base flows	Gebert and others (2011)	--	0.126–0.19	3	3 / (0.126–0.19 x flow)	18	1,951.3
Base flows	NWIS	2011 base-flow measurements	0.5	5	5 / (0.5 x flow)	11	251.1
Base flows	NWIS	2013 base-flow measurements	1	2	2 / (1 x flow)	8	3.3
Base flows	NWIS	2011 base-flow measurements, small streams[2]	1	3	3 / (1 x flow)	19	137.2
		Total weighted observations				**1,224**	**4,218.0**

[1]Coefficient of variation. Values of 0.126 reflect the standard error in the relation-line approach of Gebert and others (2011), while 0.19 reflects a standard error of 14 percent in the statewide equation approach, and an estimated 5 percent error for a "good" quality rating in NWIS. Higher values reflect increasing uncertainty inherent in measuring small streams.

[2]Streams measured at 0.5 cubic feet per second or less. Streams with no measured flow were assigned weights of 1 x 10[-5].

Base modified from the National Hydrography
Dataset (NHDPlus v2; McKay and others, 2012)

Figure 12.   Estimated annual recharge to the groundwater system, following smoothing of the Soil Water Balance results, and adjustment during calibration of the groundwater model.
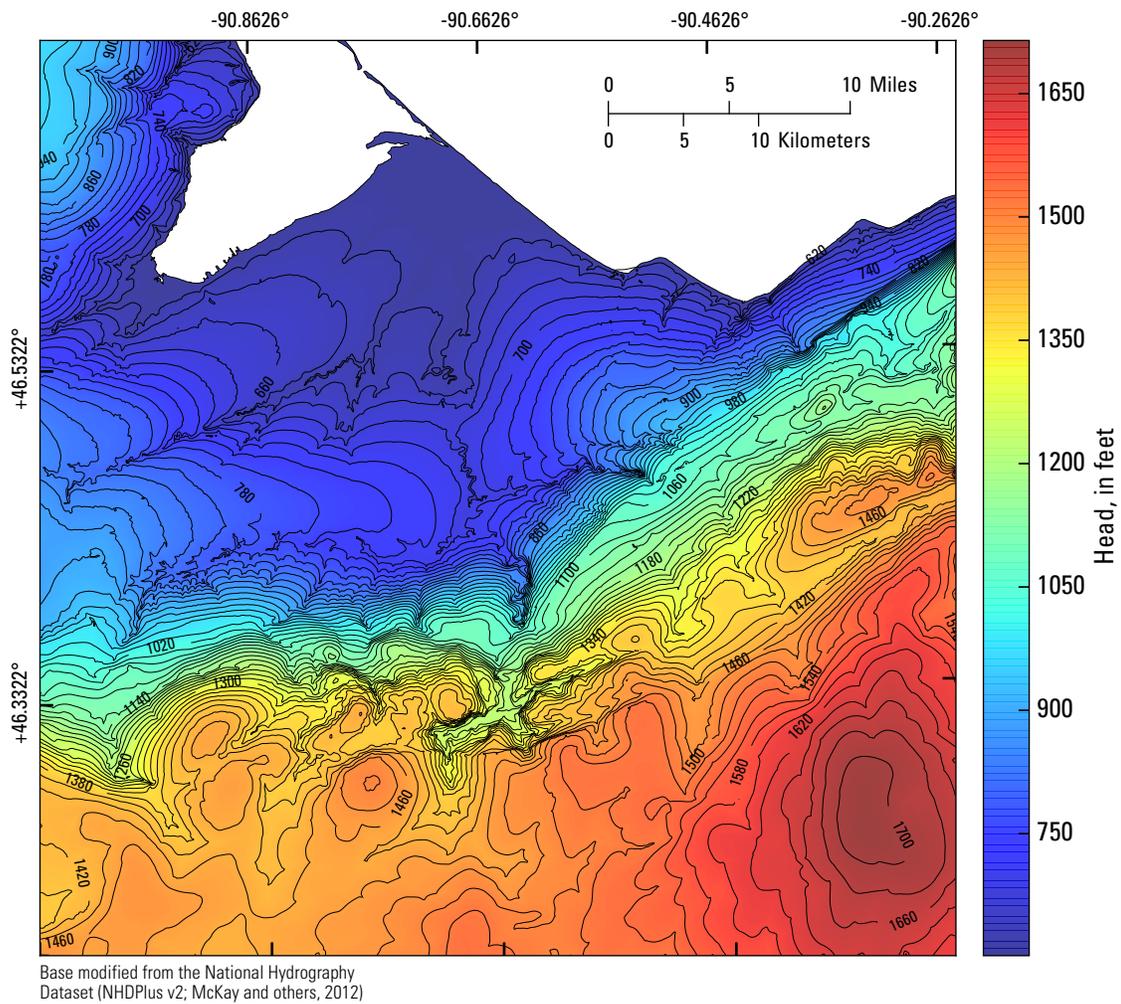
Figure 13.   Simulated water-table elevation. The black lines are contours of equal water-table elevation at 20 foot intervals. Groundwater flow occurs in the direction of steepest decent, perpendicular to the equal elevation contour lines.

EXPLANATION

Velocities (arrow size indicates magnitude)

Down

Up

Stream-aquifer interactions

State

Gaining

Losing

Dry cells
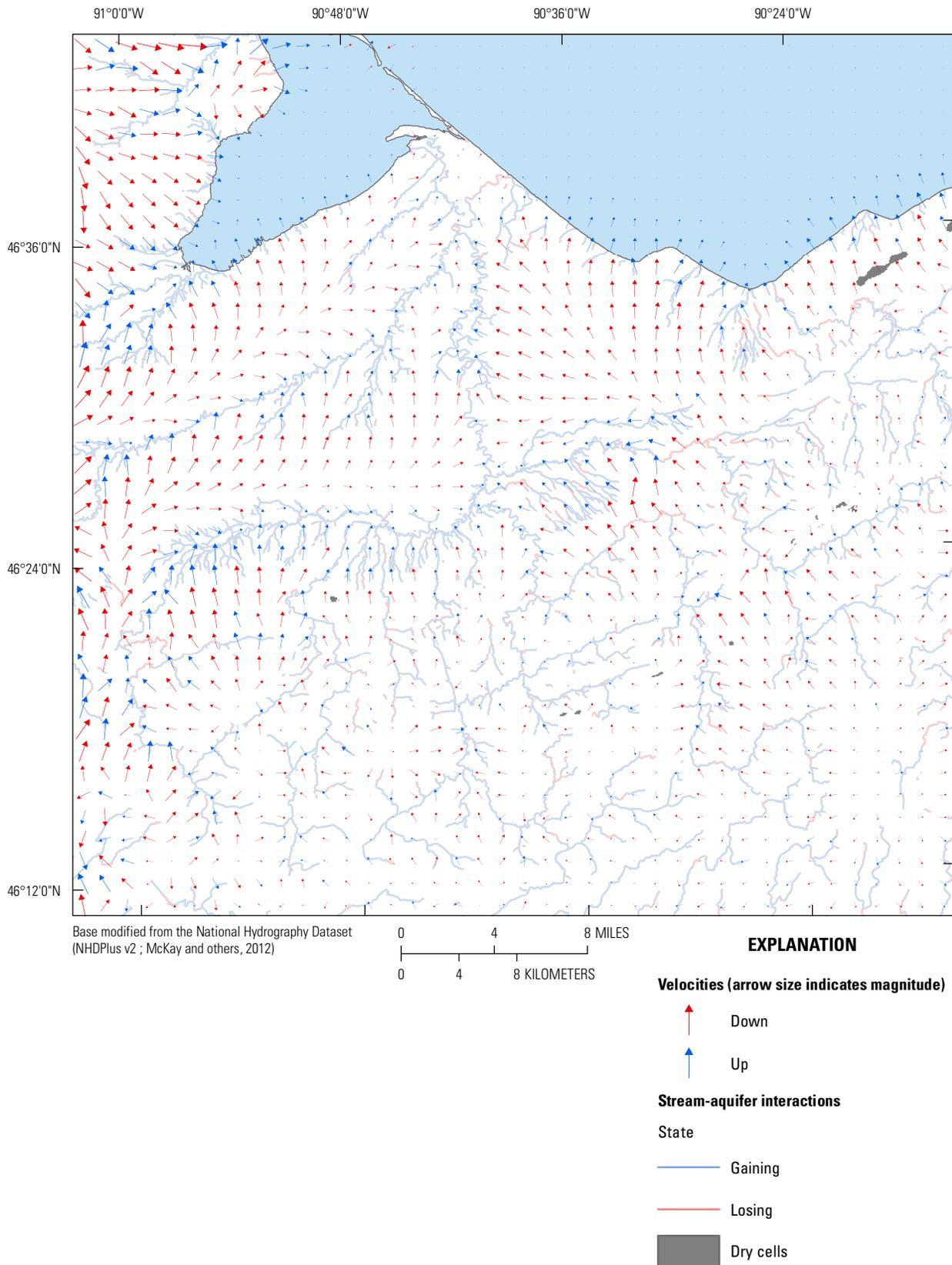
Figure 14. Simulated specific discharge of groundwater in model layer 4. The arrows represent the net direction and magnitude (normalized to model cell cross-sectional area) of simulated groundwater flow at 20 cell intervals. Smaller, scattered arrows in the southern and eastern part of the model domain indicate shallow, localized flow systems in the southern bedrock uplands, whereas larger, more consistent arrows closer to Lake Superior indicate longer regional flow paths in the deep sandstone system.
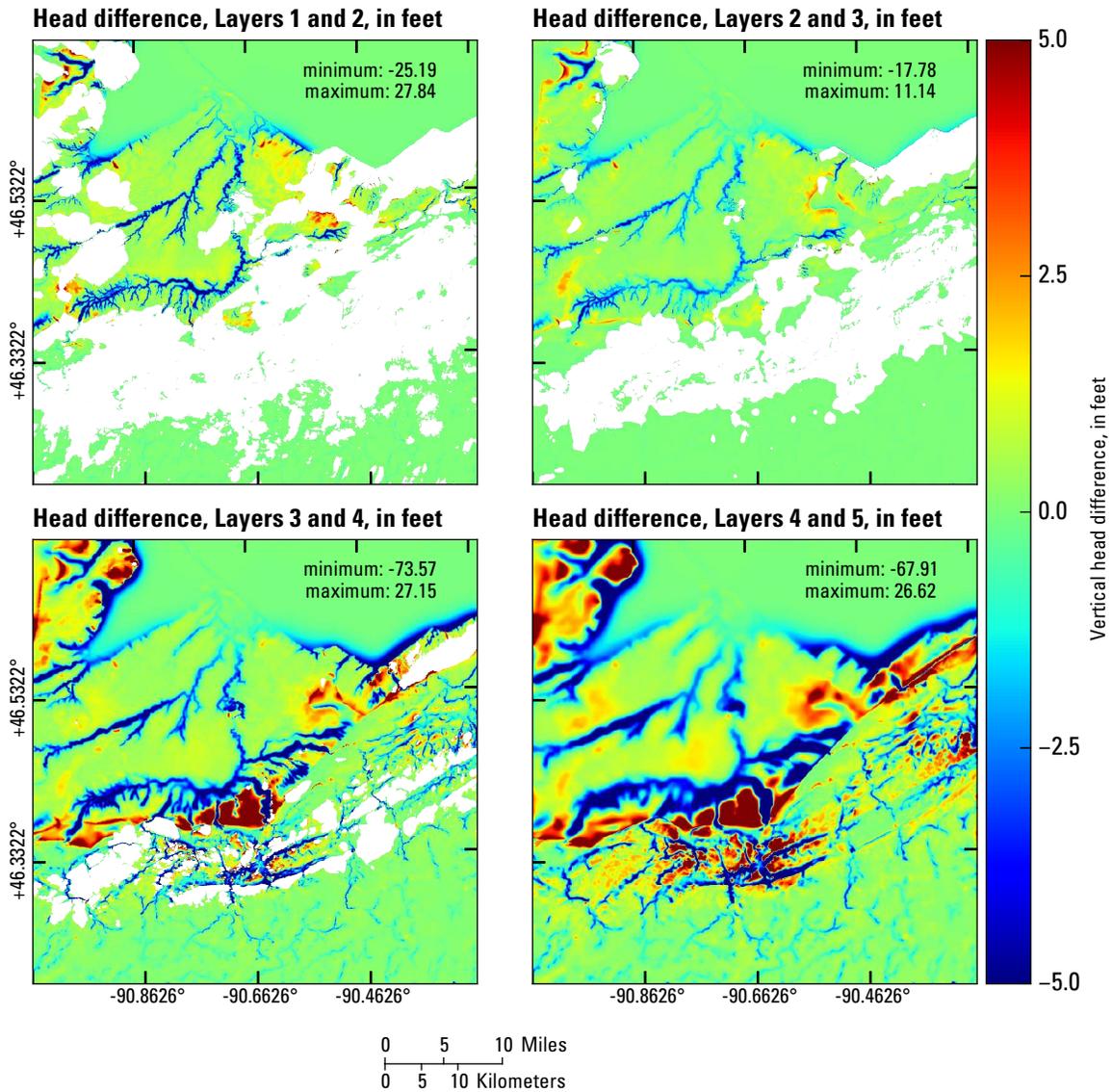
Figure 15.   Vertical differences in simulated heads between model layers. The blue areas symbolize decreases in heads in the upward direction, indicating the upward flow of groundwater towards discharge features. The red areas symbolize higher heads is overlying layers, indicating downward flow from recharge areas. Areas of white indicate model cells that are dry (that is, have a bottom elevation above the simulated water table) in one or both of the layers being compared. The strongest vertical differences in head indicate a greater potential for flow due to both higher fluxes of (focused) recharge and discharge (for example, near the Marengo and Potato Rivers), and (or) areas where low vertical hydraulic conductivity in local hydrogeologic units create a greater resistance to groundwater flow (as in the volcanic units along the Bad River, near Copper Falls State Park, Mellen, Wisconsin).

Table 3.   Mass balance for model domain.

[ft^3/s, cubic feet per second]

	Inflow (ft^3/s)	Percent
Recharge	526.2	83.1
Leakage from streams	80.7	12.7
Lateral flow across model boundaries	26.7	4.2
Total	**633.6**	

	Outflow (ft^3/s)	Percent
Discharge to streams	457.7	72.2
Disharge to land surface[1]	118.1	18.6
Discharge to Lake Superior	36.2	5.7
Lateral flow across model boundaries	19.8	3.1
Pumping from high-capacity wells	0.6	0.1
Total	**632.5**	

[1]Groundwater discharge simulated by UZF package where the water table elevation is above land-surface. In reality this represents groundwater discharge to wetlands, seeps and small streams that are not represented in the SFR stream network.

northern part of the Reservation, where heads in the deep sandstone aquifer are above the land surface.

Losing Reaches Along the Tyler Forks River

Of particular interest is the northernmost reach of the Tyler Forks River (located on fig. 17) where it bends from its northward course back towards the southwest, passing through the southeast corner of the Reservation. Most of this river is gaining, but this reach is simulated as losing (fig. 17). Field observations were not definitive, but the calibration found good agreement between simulated and measured values for the upstream and downstream base-flow observations (although the closest upstream observation is several miles away, at Moore Park Road; see figure 11). Consistency of this trend over successive calibration iterations indicated that losing reaches near the bend in the Tyler Forks River may be required to match both the base flows near the Penokee ridge, and those downstream of the bend, while balancing residuals from other observations.

Conceptually, losing reaches in this area are plausible, given the proximity of the Tyler Forks River to much lower discharge points in the Potato River subbasin (which could act as competing sinks, lowering the water table beneath the Tyler Forks River). In addition, this portion of the Tyler Forks River is underlain by outwash sands of the Copper Falls Formation, which are relatively high in permeability, and therefore would tend to promote the transmission of effects from competing sinks. Synoptic base-flow measurements collected in 2011 and 2012 showed no net increase in base flow between stations 04026559 (near Highway 169) and 04026561 (at Stricker

Road), indicating at least a neutral (or possibly losing and then gaining) state in the Tyler Forks River downstream of the bend (see USGS NWIS data, http://waterdata.usgs.gov/wi/nwis/uv?site_no=04026561). However, there are limited to no observation data to constrain the model closer to and upstream of the bend.

Substantial leakage from the Tyler Forks River into the groundwater system is potentially important for the water resources of the Reservation because it forms a potential source for contamination of the groundwater system, with implications for the Potato River subbasin. Reduction of base flow in the Tyler Forks River during dry periods could also result in a decrease or cession of leakage, and a corresponding decrease in groundwater levels near the river. More field investigation could lead to better understanding of the relation between the Tyler Forks River and the groundwater system in this area.

Delineation of Groundwater Contributing Areas

Just as a surface watershed is defined by the topography draining surface runoff to a single, downstream location, a similar concept can be defined for the area contributing groundwater. In this report, the groundwatershed is defined as the area contributing groundwater to the surface-water system within the watershed. In many systems dominated by topography (such as montane areas) and (or) those where the groundwater-flow system is shallow, the surface watersheds and groundwatersheds can be similar. But this can vary locally even in shallow groundwater systems (for example, Hunt and others, 1998). In other places, such as
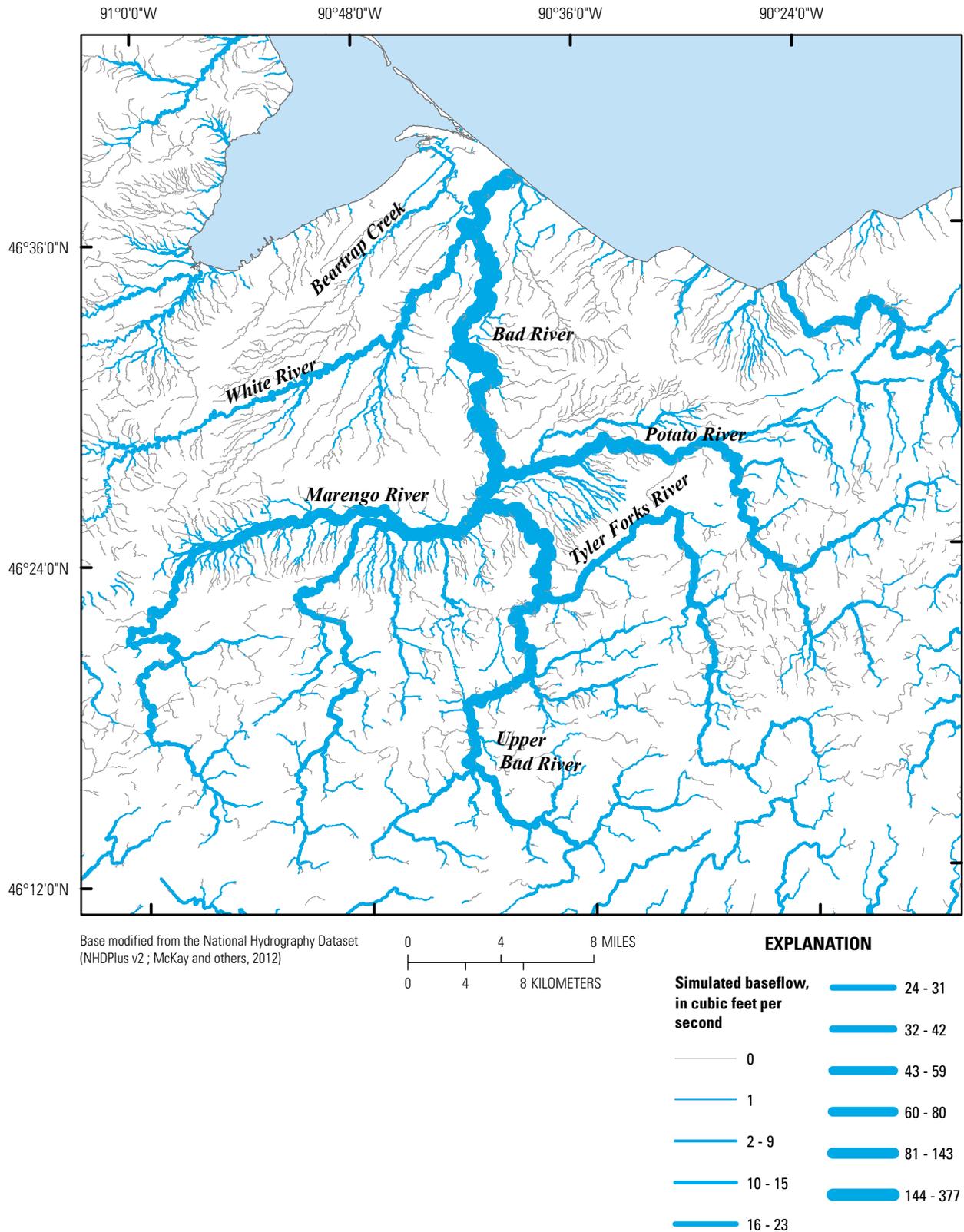
Figure 16. Simulated base flows in streams scaled to display flow magnitude. The blue lines indicate streams simulated as flowing, with the line thickness proportional to the quantity of flow. Grey lines indicate streams simulated as dry during the average conditions used for calibration.
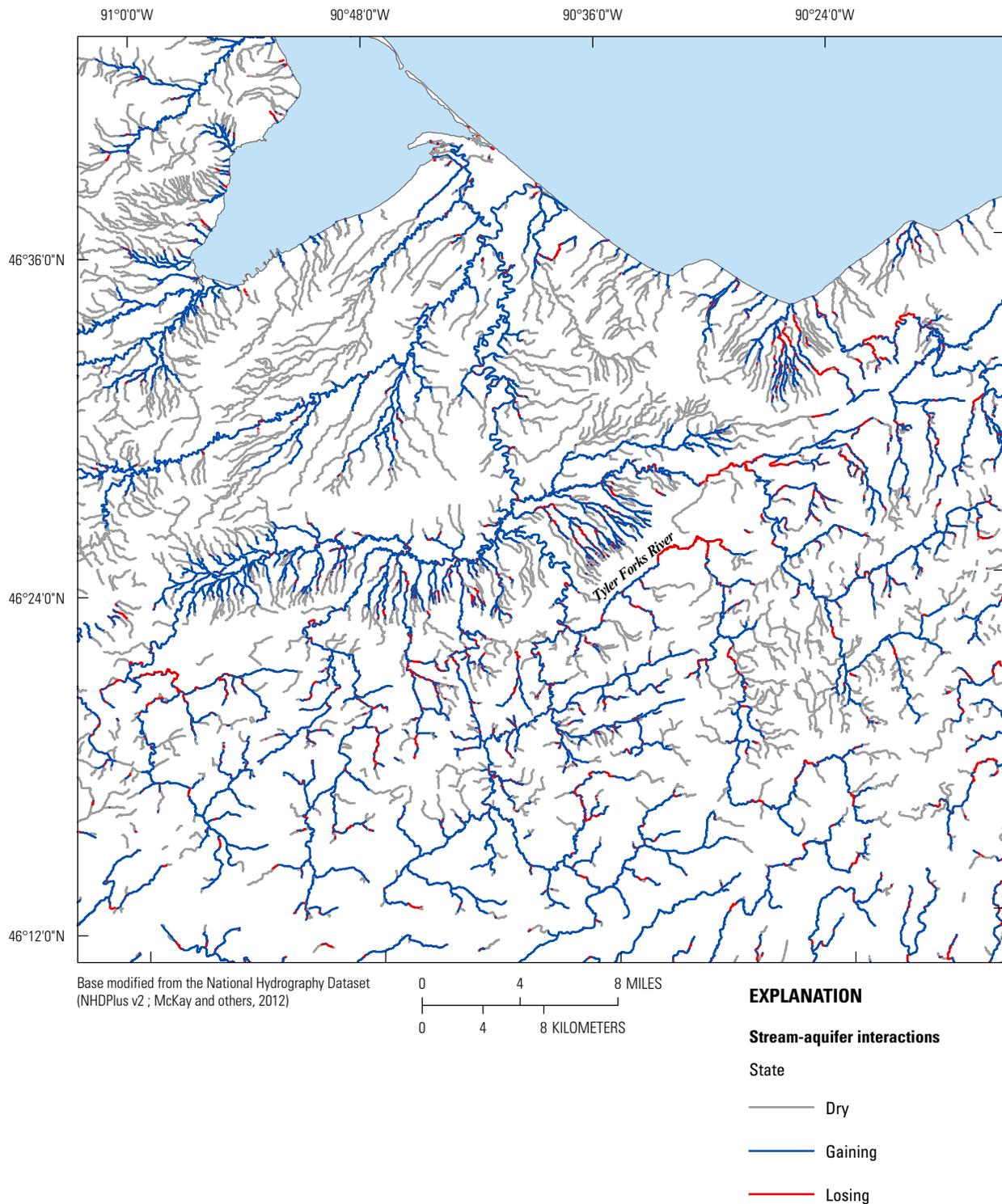
Figure 17.   Simulated groundwater/surface-water interactions. The blue lines indicate portions of streams that are simulated as gaining (receiving water from the groundwater system); the red lines indicate streams that are losing water to the groundwater system. The grey lines, although represented in the MODFLOW model, are dry due to simulated water-table elevations below the bottom of the stream bed.

southeastern Wisconsin, with a deep, confined flow system and extensive pumping, the groundwater and surface watersheds are different (Feinstein and others, 2005).

The program FlowSource (Black and Foley, 2013) was used to delineate groundwater contributing areas. FlowSource reads MODFLOW budget output and traces groundwater flow backwards from a defined set of "destination cells," calculating the volume and proportion of water in each upstream cell that ultimately reaches the destination cells (Black and Foley, 2013). Cells where 100 percent of flow reaches the destination cells are within the groundwater contributing area (groundwatershed) for those destination cells. Cells where no flow reaches the destination belong to a different contributing area in the simulated groundwater system (their flow goes to a different sink). Cells with percent contributing values between 0 and 100 define the fringe of the groundwater contributing area.

Figure 18 shows the groundwatershed for the Bad River and Beartrap Creek, in comparison to the surface watershed (denoted by the black line). Red denotes complete contribution, where 100 percent of cell-by-cell flow in the model reaches a surface-water feature within the Bad River-Beartrap Creek Watershed. Areas colored blue denote the fringe of the groundwatershed, where a minor percentage of water ends up in the Bad River-Beartrap Creek surface-water system, but most discharges to another sink, such as Lake Superior. Similar to an area of Vilas County described by Hunt and others (1998), the groundwater and surface watersheds in figure 18 generally match but do not perfectly coincide. The largest deviation is beneath Birch Hill, where surface drainage is towards Vaughn Creek, but groundwater flows towards Lake Superior, which provides a stronger competing sink.

Figure 19 shows the area of groundwater contribution to the Reservation. The colored area in figure 19 represents the sources of groundwater within the Reservation. Groundwater outside of the colored area in this figure (but within the groundwatershed delineated in fig. 18) discharges to surface water before entering the Reservation. This indicates that under current conditions most groundwater within the Reservation originates within or near the Reservation boundary. In particular, the groundwater contributing area for the Reservation is not simulated as extending south into the area underlain by the Marquette Range Supergroup. This indicates that the primary pathway of potential contamination from sources located distant from the Reservation boundary would likely be through surface water.
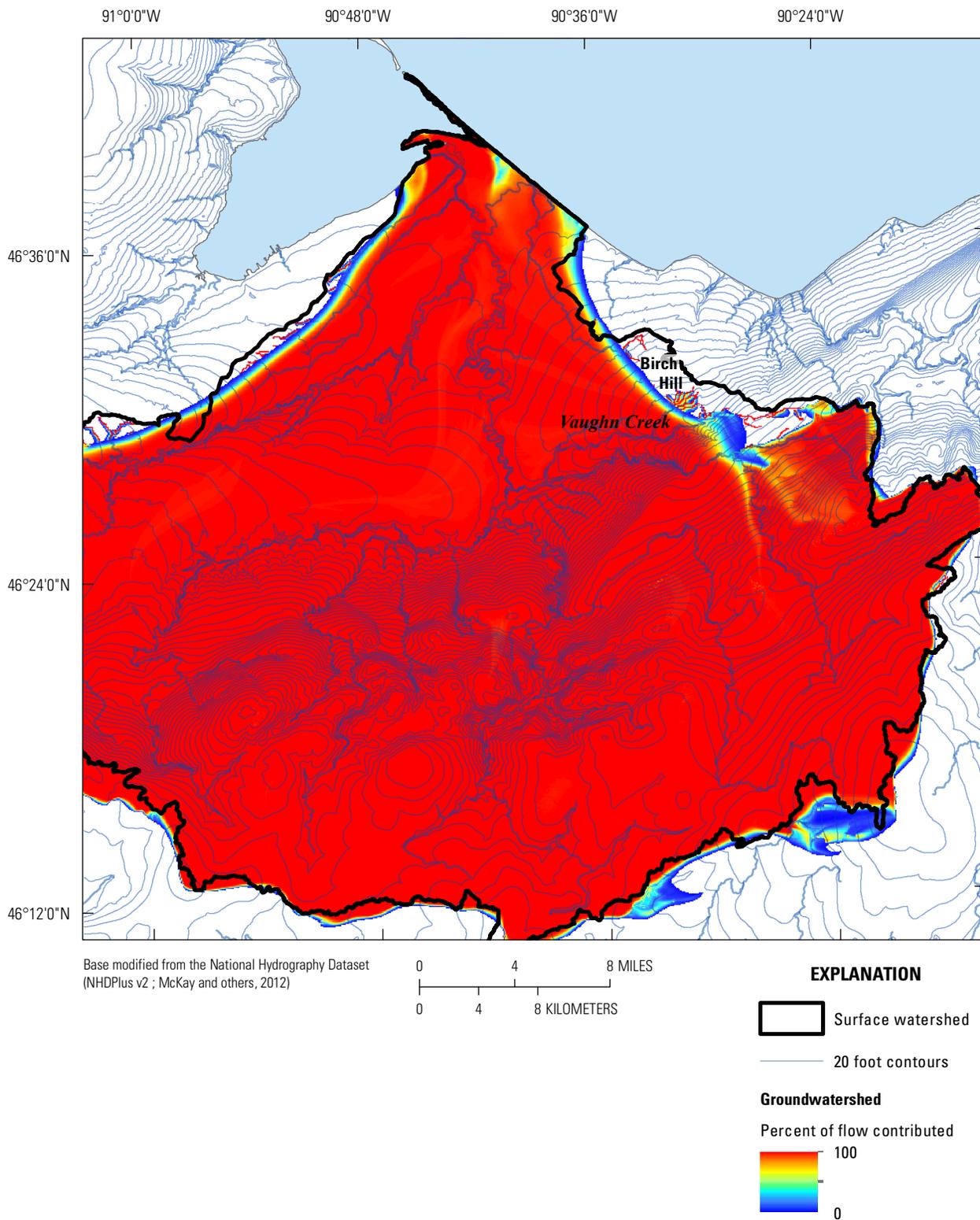
Figure 18.   Groundwatershed for the Bad River and Beartrap Creek, Wisconsin, as calculated by FLOWSOURCE (Black and Foley, 2013).
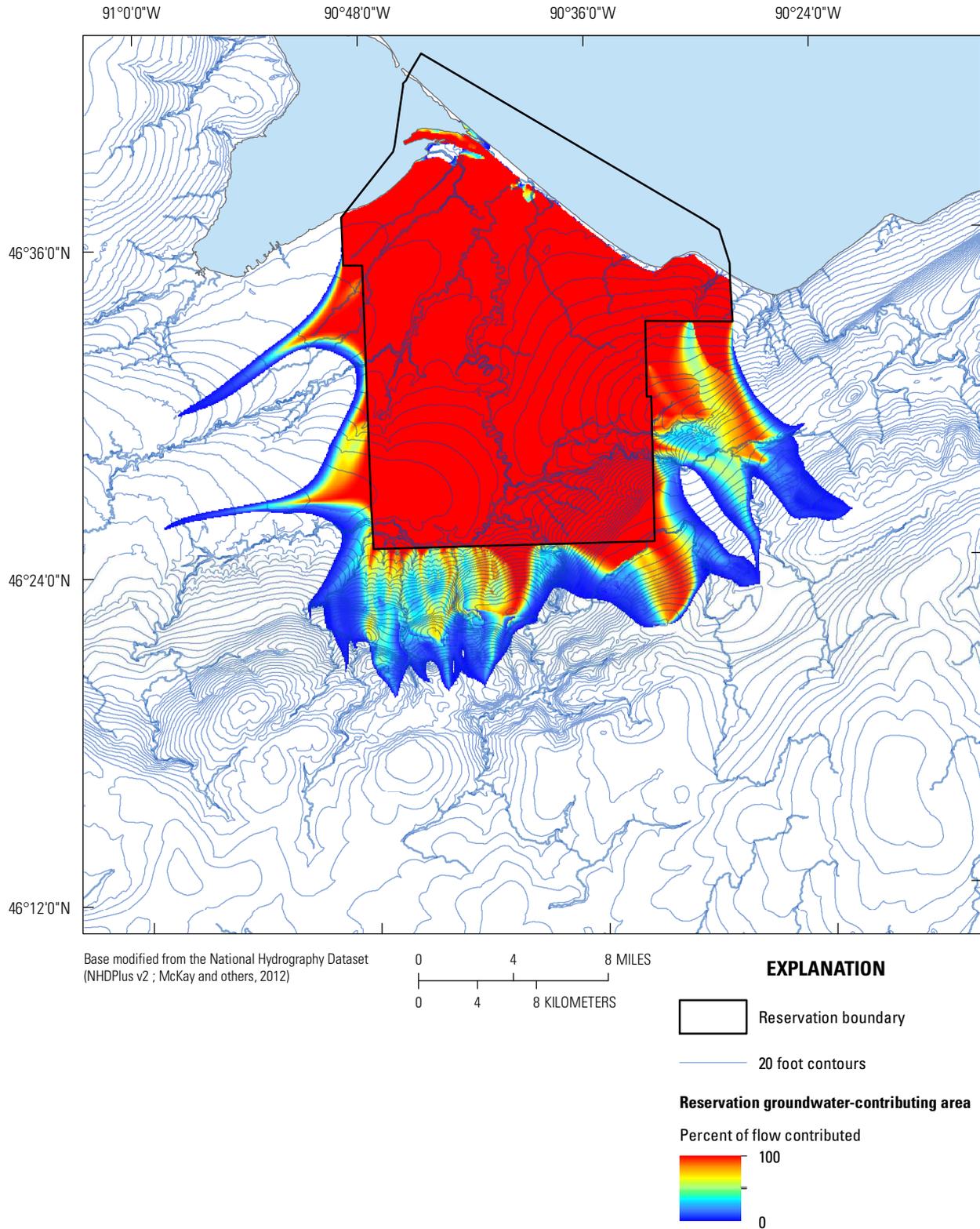
Figure 19.   Area of contribution of groundwater flow to the Reservation, as calculated by FLOWSOURCE (Black and Foley, 2013).

New Data-Worth Analysis

The model was used to evaluate potential locations for new monitoring wells though linear analysis of model prediction uncertainty following the methods of Doherty (2010b) and the design of Fienen and others (2010), as implemented in the PREDUNC5 utility of PEST (Doherty (2010a), Doherty, 2014). A regular grid of potential new monitoring-well locations was created at 1-kilometer intervals throughout the model domain. Prediction uncertainty analysis was then performed for a suite of 35 model predictions of base flow and groundwater levels (shown in fig. 20), under the hypothetical future condition of large-scale dewatering. The prediction used here is driven by changes associated with an open-pit iron mine described by Gogebic Taconite (2013) and assumes that groundwater pumping for dewatering the mine would lower water levels to an elevation of 800 ft above sea level, in a rectangular area approximately 4.2 mi long by 0.5 mi wide (Gogebic Taconite, 2013, p. 10 and 59). The mine dewatering stress was simulated using a constant head boundary condition. The disposition of pumped water for this purpose is unknown, so in these simulations it is not assumed to be discharged back into the system through streams, but simply removed. Further details about the potential use of the water for operations or other disposition were not available for this study. The 35 predictions–a combination of base flow and groundwater-level predictions–were elicited from stakeholders in the area as concerns of potential impacts from the proposed mine.

For each of the 35 predictions, metrics of data worth (indicating which new data points would most reduce uncertainty of each prediction) were calculated at each potential new monitoring-well location. These results were then contoured, resulting in a map of data worth for potential new data tied to each prediction of interest (for example, fig. 22).

Uncertainties in the model predictions were computed by linear propagation of uncertainties in model input parameters, as informed by observed data used to calibrate the model (see equation 6.1 in appendix 6). The uncertainties are transferred to the model predictions through parameter sensitivities, which describe changes in model outputs (heads and base flows) that occur with incremental changes in each model input parameter. The term "linear" is used, because sensitivities are assumed to be independent of the parameter values (that is, the relations between model outputs and input parameters are assumed to be constant across parameter space). This enables the matrix of sensitivities to represent the transfer of information from observations to parameters in a closed-form, analytical step. As a result, potential new data can be simulated as informing the parameters through

calibration without needing to know the actual values of new data at those locations.

Innate parameter variability information and uncertainty associated with each observation are needed to perform the analysis. Innate parameter variability in hydraulic conductivity values at each pilot point were based on the approximate 95 percent confidence limits of the log-space hydraulic conductivity distributions computed by K. Bradbury (Wisconsin Geological and Natural History Survey, unpub. data, 2013) for the major hydrogeologic unit types in the model domain (table 4). Uncertainties in observations were supplied as the inverse of their weights from the calibration process (table 3.1). The new monitoring points were assumed to have the same measurement uncertainties as the highest quality head measurements in the calibration dataset.

Two sets of sensitivities were computed:

1. The sensitivities of observations in the calibration dataset (referred to as the base observations) and the hypothetical new observation of groundwater levels with the model in the same conditions as for calibration (without the drawdown of the proposed mine simulated), using the parameter set from the calibrated model.

2. The sensitivities of the 35 prediction "observations" under the condition of steady-state drawdown from the proposed mine, also with the parameter set from the calibrated model.

The "data worth" of each hypothetical new observation n is then computed by comparing the prediction uncertainty values for the stressed conditions with and without the new observations after normalizing for the original prediction uncertainty (Fienen and others 2010, p 14):

$$dataworth_n = \frac{\sigma^2_{post\_calibration} - \sigma^2_{post\_calibration+n}}{\sigma^2_{post\_calibration}} \quad (3)$$

where

$\sigma^2_{post\_calibration}$ is base post-calibration variance and

$\sigma^2_{post\_calibration+n}$ is post-calibration variance with additional data point n.

A data worth of 1.0 would indicate a complete reduction of uncertainty from the addition of the new observation; a data worth of 0 indicates that the new observation has no impact on uncertainty for the prediction of interest. Several base flow predictions located near the proposed mine site did not have any uncertainty results, due to the drying up of stream cells under the prediction condition of mine dewatering. These include the predictions for
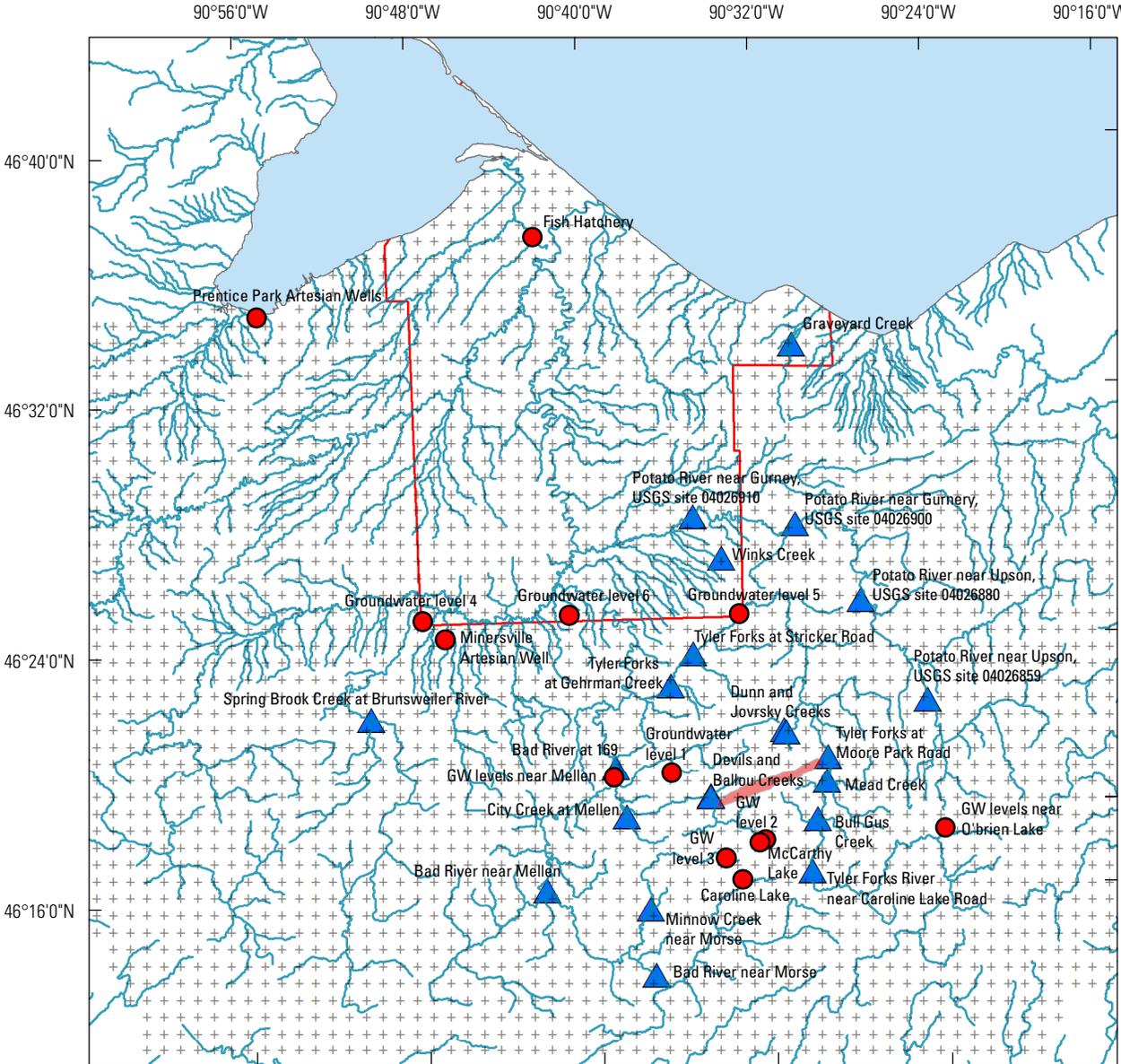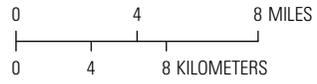
Figure 20.   Locations of model predictions of groundwater levels and base flows included in the new data-worth analysis. The mesh of gray crosses shows the network of hypothetical new monitoring-well locations that was evaluated.

Table 4.   Parameter uncertainty values used to construct the matrix of innate parameter variability (C_p matrix).

[σ, standard deviation; --, not reported or not applicable; MRS, Midcontinent Rift System; all units are in feet per day, except the global recharge multiplier, which is dimensionless]

Parameter / Pilot point zone	Model zones	Approximate +/- 2σ range in $\log_{10}$ space reported by K. Bradbury[1]	Assumed $\log_{10}$ space standard deviation used in the C(p) matrix
Global recharge multiplier	All zones	--	0.06
Global streambed hydraulic conductivity	All zones	--	0.5
Lake Superior bed hydraulic conductivity	--	--	0.5
Copper Fall Formation sand	7, 10, 11, 12	2.25	0.56
Miller Creek Formation	8	--	1.25
MRS Sedimentary	6	2.5 - 4.0	0.75
Copper Falls Formation sandy till	9	2.5	0.63
Area of shallow bedrock	13	5.0	1.25
MRS Igneous	1	5.0	1.25
Marquette Range Supergroup	2,4,5	--	1.25
Archean Units	3	5.0	1.25

[1]Wisconsin Geological and Natural History Survey, unpublished data (2013).

Javorsky, Dunn, Bull Gus, Ballou, Devils, and Mead Creeks. Dry stream cells for these predictions resulted in 0 values for their sensitivities (no change in the flow condition with changes in model input parameters). This represents a limitation in the application of this linear method, which cannot account for nonlinear processes such as drying. Data-worth values were computed for each of the 29 predictions with uncertainty results, at each hypothetical new monitoring-well location. Color flood maps of interpolated data worth were then developed for each prediction. Four example data-worth maps are shown below in figures 21 through 24; additional examples are shown in appendix 6, and the complete set of data-worth maps are accessible in the online map companion to this report (http://wim.usgs.gov/BadRiver/NewDataWorth/).

In general, predictions sensitive to the mine dewatering as simulated showed greatest data worth either in the immediate vicinity of the proposed dewatering (especially to the south), typically between the dewatering and the prediction location (for example, fig. 21). For base-flow predictions, data worth is generally concentrated in upstream contributing areas. This is illustrated in figure 22, which shows new data worth for reducing prediction uncertainty for base flow in the Tyler Forks River at Moore Park Road. The greatest data worth is concentrated south of the proposed mine, upstream of the prediction location. Drawdown in water levels in this area from the simulated mine dewatering affects base flow in the Tyler Forks River in two ways: (1) by inducing

leakage through the streambed (in areas where groundwater levels are lower than the river stage); and (2) by diverting discharge that would have otherwise gone to the Tyler Forks. Knowledge of water levels south of the proposed dewatering helps constrain the possible values of hydraulic conductivity that are estimated during model calibration, which in turn constrains the shape of the simulated drawdown in the water table, and the simulated partitioning of discharge between the dewatering activities and the Tyler Forks River.

The results for Spring Brook Creek, which is located away from the proposed mine site in the Marengo River Subbasin, reinforce the concept of greatest data worth in up-gradient areas (fig. 23). New data worth in this case is strongly concentrated near a groundwater divide between the contributing area for Spring Brook Creek and another headwater tributary to the Marengo River. Knowledge of water levels near the groundwater divide provides information about hydraulic conductivity in this area, thereby reducing uncertainty in the simulated partitioning of discharge between the two adjacent headwater streams.

New data-worth values for downstream base-flow predictions are often lower, as the addition of a single water-level observation to the calibration process does little to inform discharge accumulated from a large area. This is illustrated in the results for USGS site 04026910 (Potato River base flow near Gurney, Wisconsin; appendix fig. 6.2). An exception to this is the Tyler Forks River, which is in close proximity to the proposed dewatering.
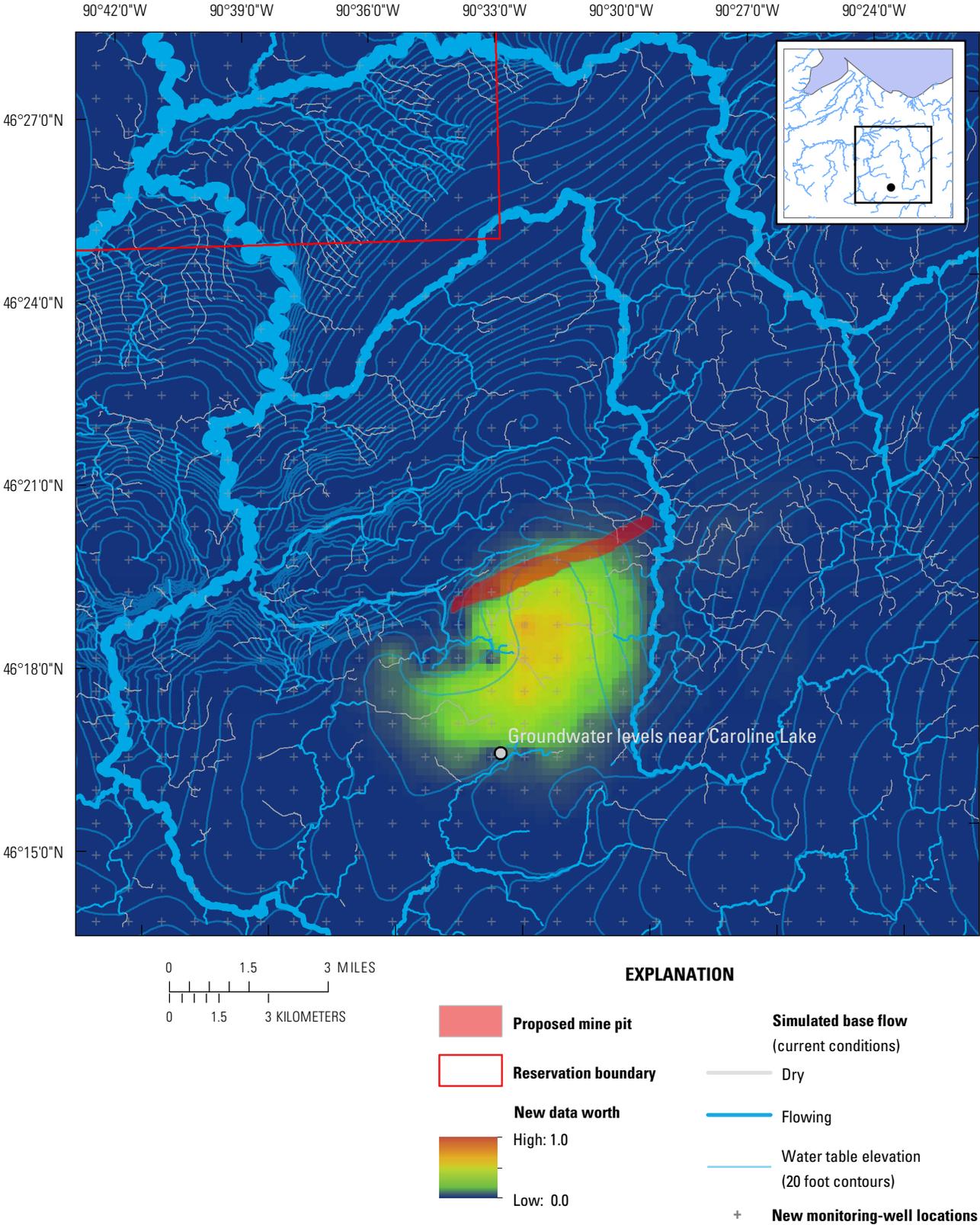
Figure 21. Normalized reductions in prediction uncertainty from potential new head measurements, for predicting water levels near Caroline Lake, Wisconsin, under the condition of large scale dewatering from a proposed open-pit mine described by Gogebic Taconite (2013).
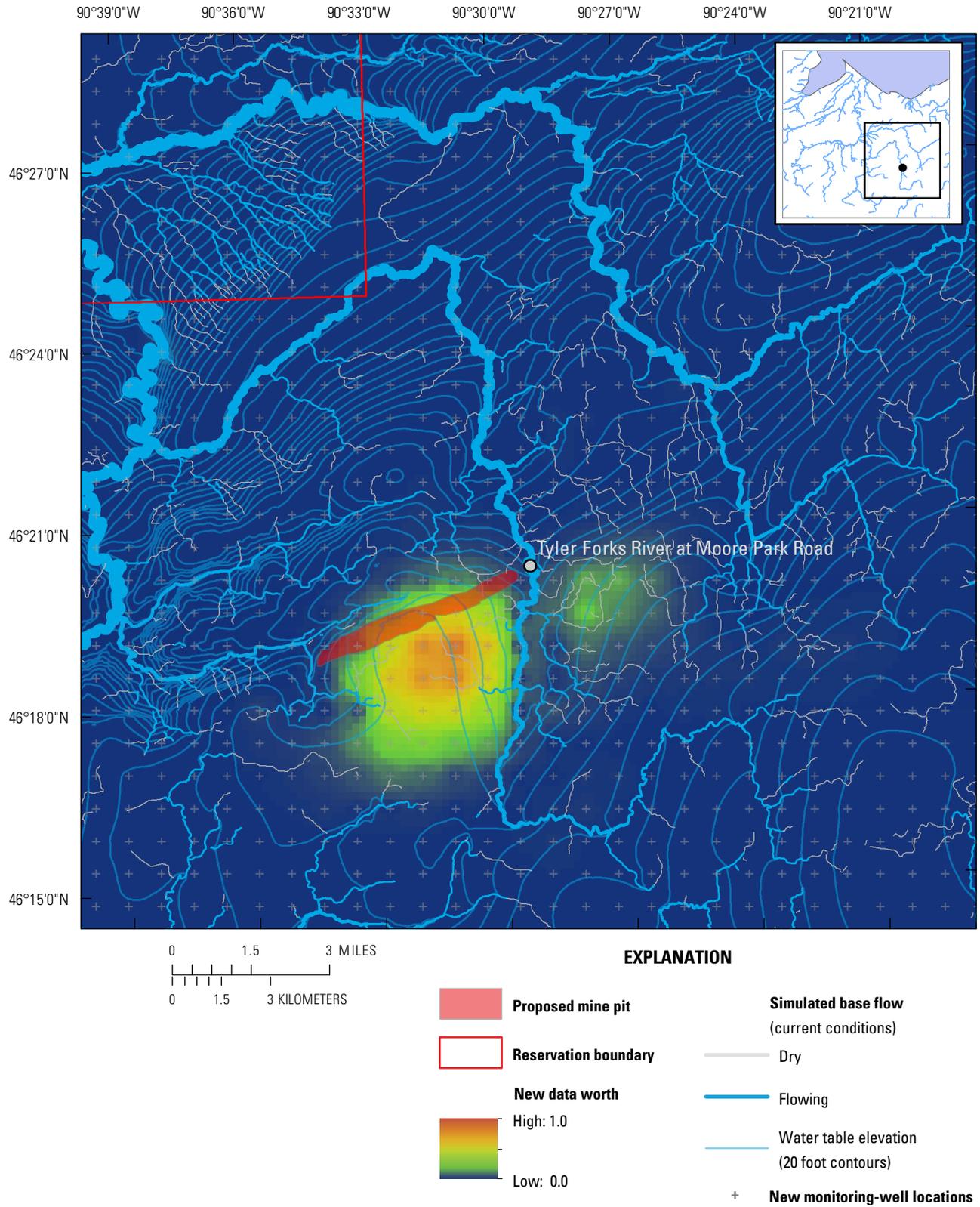
Figure 22.   Normalized reductions in prediction uncertainty from potential new head measurements, for predicting base flow in the Tyler Forks River, Wisconsin at Moore Park Road, under the condition of large scale dewatering from a proposed open-pit mine described by Gogebic Taconite (2013).
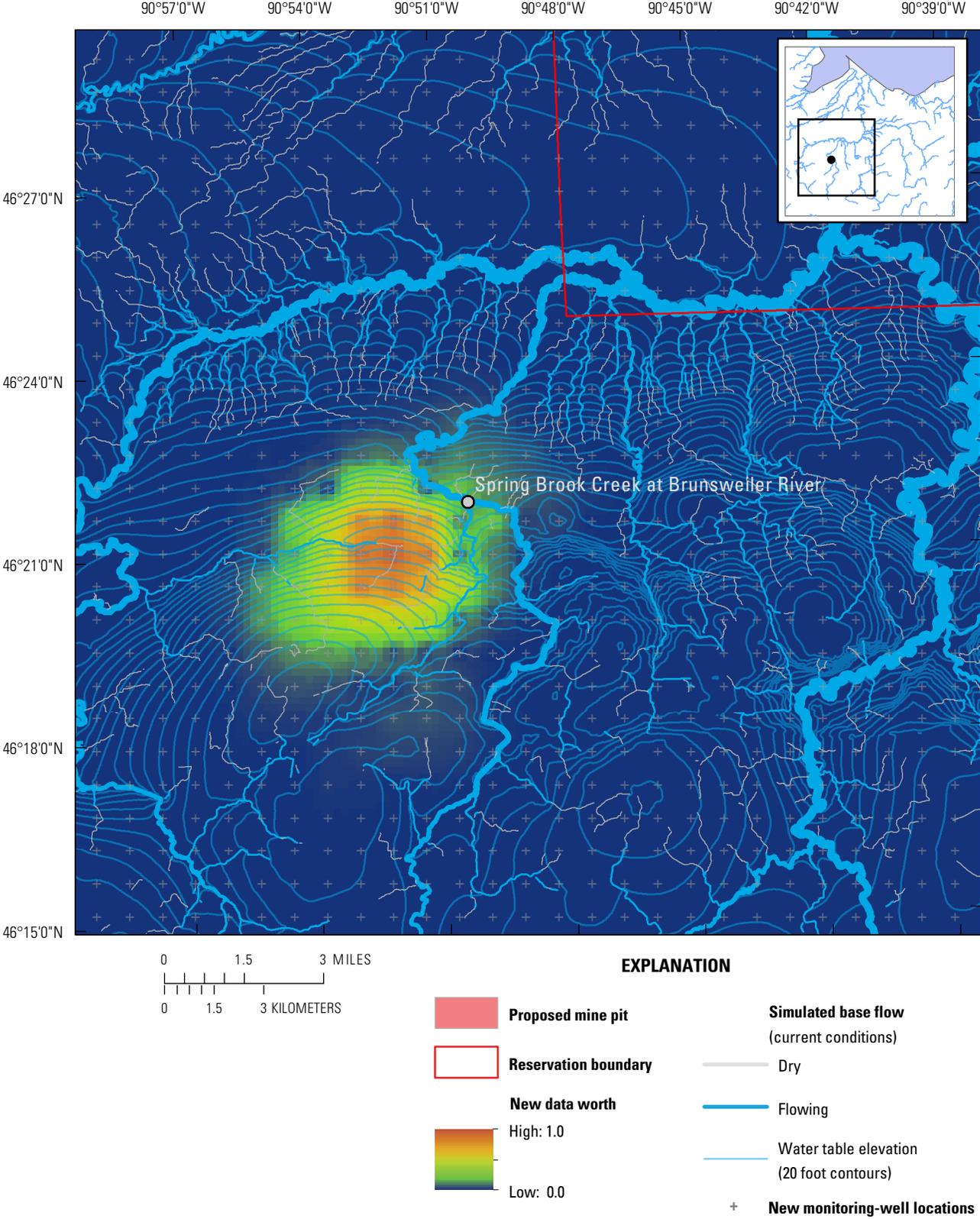
Figure 23.   Normalized reductions in prediction uncertainty from potential new head measurements, for predicting base flow in Spring Brook Creek, Wisconsin, under the condition of large scale dewatering from a proposed open-pit mine described by Gogebic Taconite (2013).
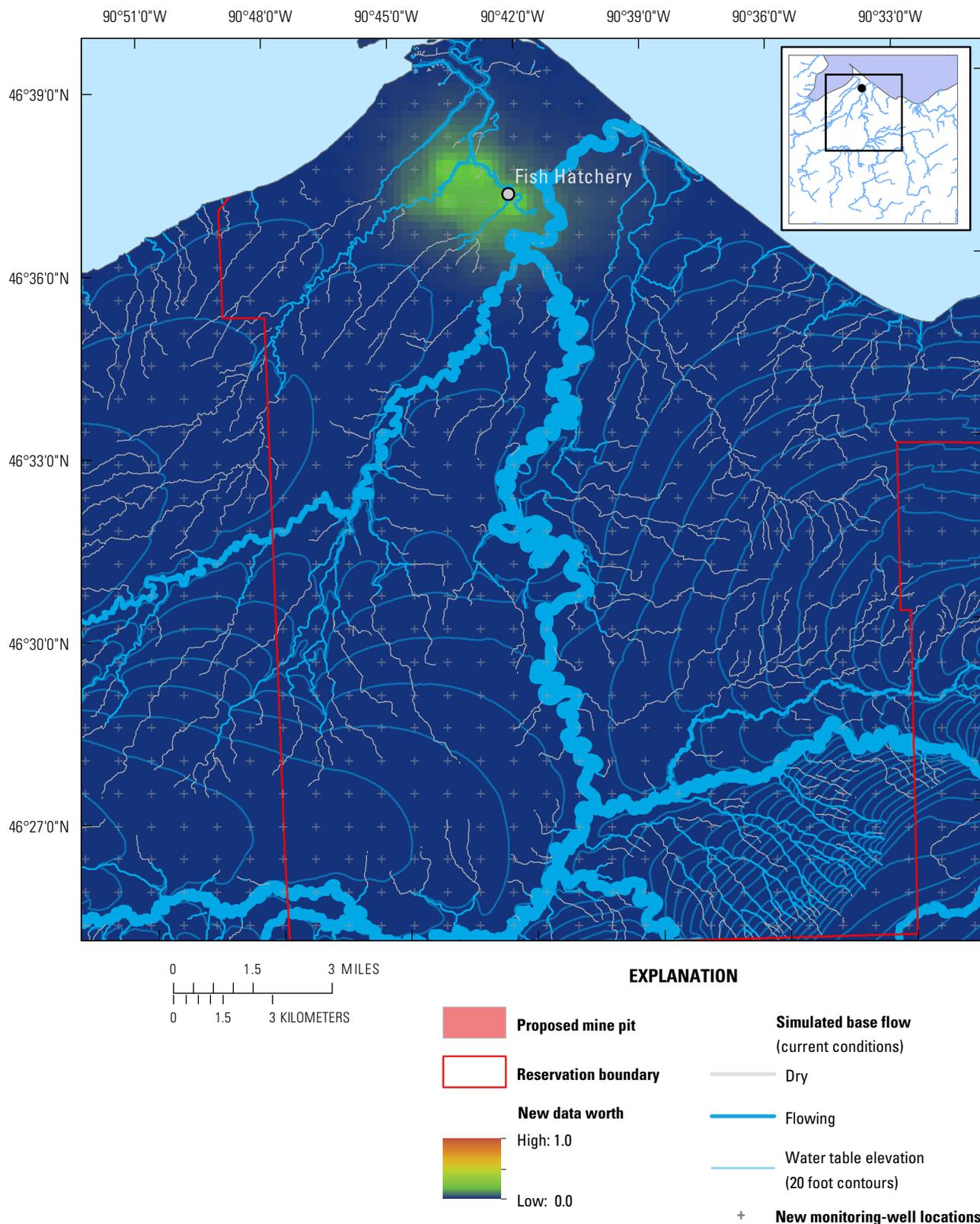
Figure 24. Normalized reductions in prediction uncertainty from potential new head measurements, for predicting water levels near the Bad River Fish Hatchery, Odanah, WI, under the condition of large scale dewatering from a proposed open-pit mine described by Gogebic Taconite (2013).

The results for USGS site 04026561 (Tyler Forks at Stricker Road) are very similar to the results for Moore Park Road, which is considerably upstream (appendix fig. 6.3). Water levels close to the dewatering activities have more worth compared to other locations in the watershed, as they provide more direct information on the partitioning of discharge between the river and dewatering. The ultimate disposition of water removed from the groundwater system for dewatering may change flow in the streams if it is discharged into them, but no specific plans about water use at a proposed mine are available, so is was assumed for this analysis that all dewatering water was removed from the system and used for processing or otherwise discharged outside the basin. More detail on this water use question could change the assumptions about reduced overall flow in the nearby streams although the base-flow-specific analysis would stand.

Boundary conditions and existing observations can also affect new data worth. Whereas Spring Brook Creek (fig. 23) is located in an area of relatively few head observations, the Fish Hatchery artesian well (fig. 24) is located near other wells, and near the boundary conditions of Lake Superior and several rivers. New data worth for this prediction is low because heads in this location are already relatively well constrained by existing data.

Stable Isotope Investigation

Stable isotopes can be valuable tools for investigating hydrologic systems (Clark and Fritz, 1997). In this study, stable isotopes of the water were used to qualitatively "fingerprint" the water on and near the Reservation. Water isotopes are ideal conservative tracers of water sources because they are part of the water molecule itself. Stable isotopes do not undergo radioactive decay, and stable isotopes of water are conservative in aquifers at low temperature, but fractionate at the surface by evaporation when humidity is less than 100 percent (Gat, 1970). The isotopes of water used here were comparing ratios of hydrogen (^{1}H) and deuterium (^{2}H) and oxygen-16 (^{16}O) and oxygen-18 (^{18}O). For example, because the vapor pressure of ^{1}H$_2^{16}$O is greater than ^{2}H$_2^{18}$O, the residual liquid is characterized by a higher ^{2}H$_2^{18}$O content after evaporation. The efficiency of this isotope fractionation is also temperature dependent, where colder temperatures increase fractionation efficiency (Clark and Fritz, 1997). Hydrogen and deuterium fractionate to a greater extent due to larger percent atomic mass difference. Thus, characteristic ^{18}O/^{16}O and ^{2}H/^{1}H ratios can identify different sources of groundwater.

Stable water isotope approaches are commonly used in Wisconsin (for example, Krabbenhoft and others,

1990; Ackerman, 1992; Krabbenhoft and others, 1994; Michaels, 1995; Hunt, 1996; Hunt and others, 1996, 1998; Hunt and Steuer, 2000; Lott and Hunt, 2001; Walker and others, 2003; Kurtz and others, 2007; Hunt and others, 2005, 2006; Fienen and others, 2009a; Hunt and others, 2010). Whereas discrete physical measurements represent the system at the point in time the measurement was taken, stable isotope compositions reflect the initial source waters entering the system (typically through recharge on land), subsequent additions and withdrawals, and processes acting within the system. Consequently, transient hydrologic events have isotopic effects proportional to their physical importance. Given a sufficient distinction between sources, water isotopes can give similar information with minimal sampling and physical measurement (Hunt and others, 1998).

Groundwater samples were collected from 56 wells for stable water isotope analysis (fig. 25). Unfiltered water was stored in 40-milliliter glass vials with inverted cone polyseal caps that were sealed with Parafilm®. Analyses of water isotopes were performed at the USGS Reston Stable Isotope Laboratory (http://isotopes.usgs.gov/). Values of ^{18}O/^{16}O (δ^{18}O) and deuterium/^{1}H (δ deuterium) are reported in standard delta notation (δ as "per mil" or ‰) relative to Vienna Standard Mean Ocean Water (VSMOW) standard. Analytic error (2σ) is estimated at ±0.1‰ and ±2.0‰ for δ^{18}O and δ deuterium, respectively. Results of the analyses show that the samples follow a single meteoric water line similar to the local meteoric water line (fig. 26). This correspondence indicates that none of the samples were appreciably evaporated before recharging to the groundwater system; rather, groundwaters sampled on or near the Reservation were not derived from infiltration from surface-water sources but rather from terrestrial groundwater recharge.

The spatial pattern in figure 25 follows a general trend from heavier stable isotope composition (that is, a relatively high ^{18}O/^{16}O ratio) in the south to lighter stable isotope composition to the north, with the most depleted δ^{18}O (most negative values) occurring near Odanah, Wisconsin. For comparison, the annual weighted isotopic δ^{18}O composition for precipitation for nearby Vilas County was around -10.9‰ (Krabbenhoft and others, 1990). Groundwater δ^{18}O composition for northern Wisconsin is around -11.3‰, with values as light as -12‰ noted in areas having increased snow deposition from the lake effect associated with Lake Superior (Michaels, 1995). Samples collected near Lake Superior at depth during this study are substantially lighter (for example, -15‰ to -17‰) than isotopic compositions reported for current northern Wisconsin precipitation and associated shallow groundwater. Michaels (1995) also noted light δ^{18}O composition in groundwater wells near Lake Superior in Iron County, Wisconsin (-13.4‰), and Little
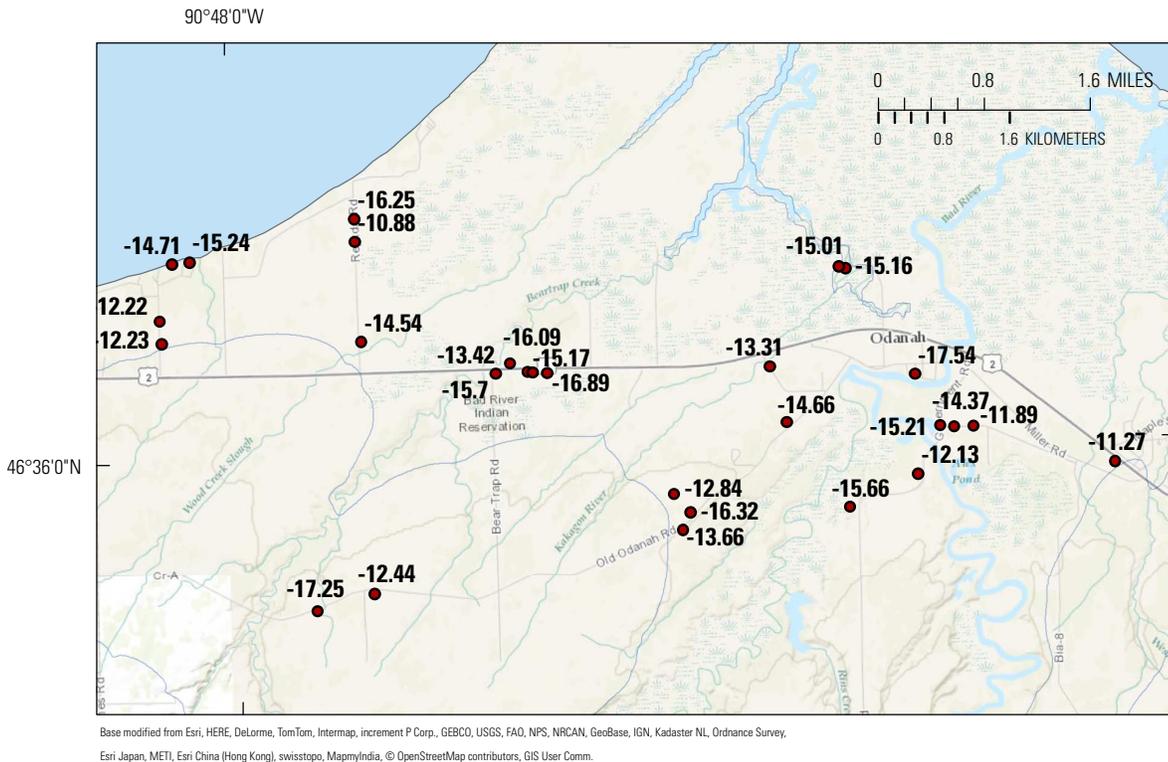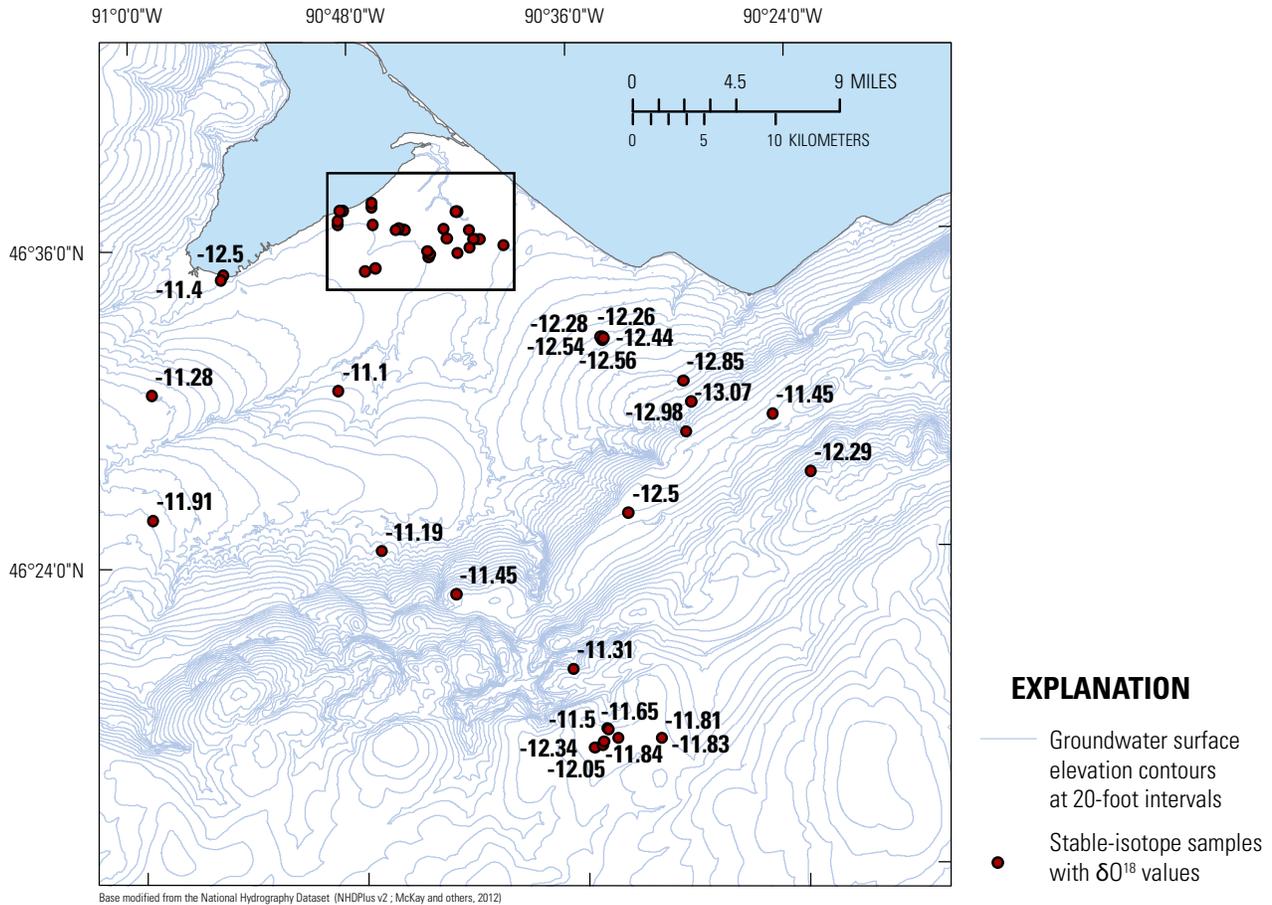
Figure 25. Stable isotope sampling locations with δ^{18}O values. Simulated water-level contours are provided for reference; inset in upper map is shown in more detail in the lower map.

Figure showing scatter plot with local meteoric water line. Axis labels: x-axis $\delta^{18}O$, y-axis $\delta^2 H$.

Text within plot:
Local meteoric water line for the
North Shore of Lake Superior, Minnesota
(from Magner and others, 2014)

$y = 8.2138x + 14.868$
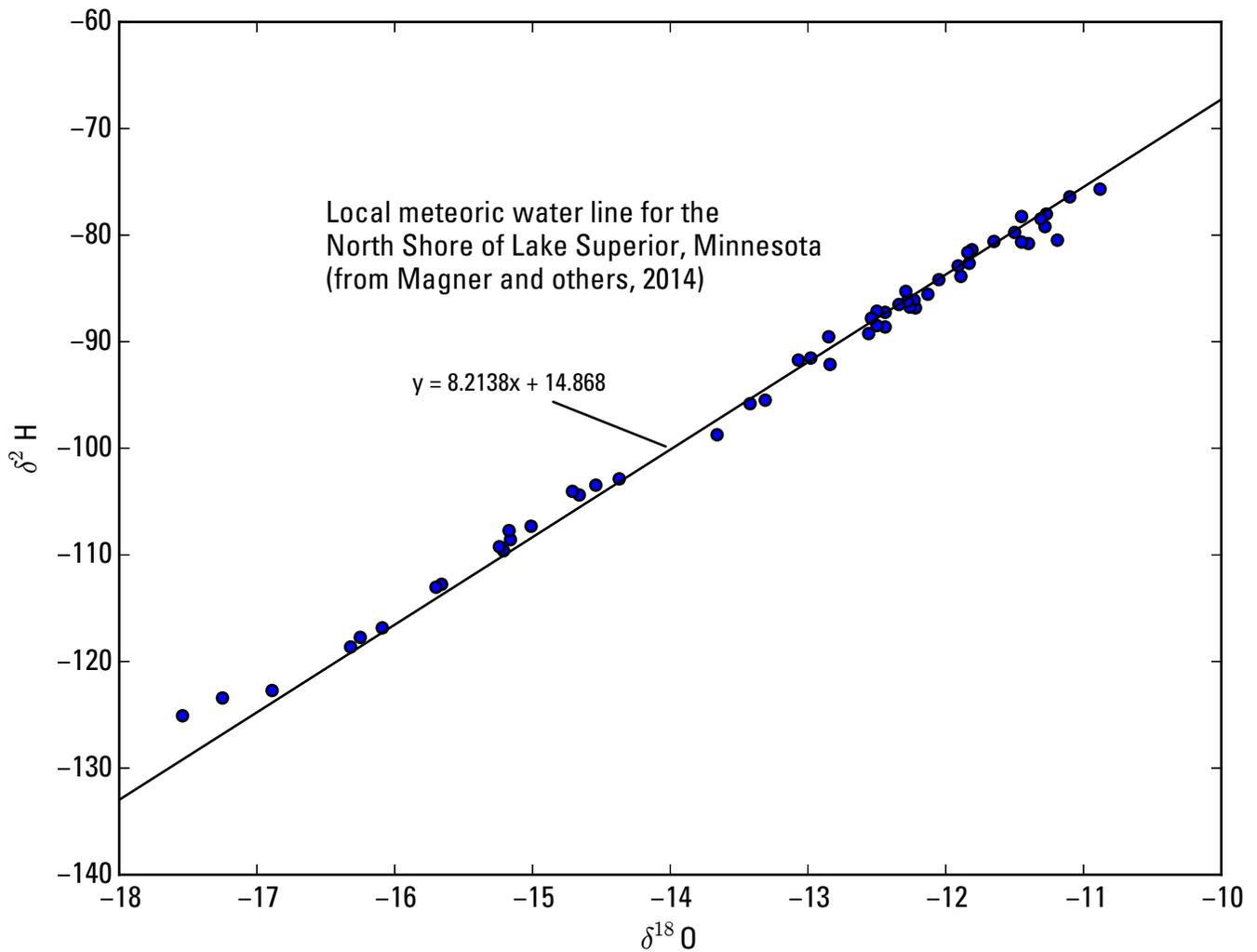
Figure 26. Local meteoric water line established by Magner and others (2014). Close correspondence of the samples to the line indicates that the source water was not appreciably evaporated prior to entering the groundwater system. This indicates terrestrial groundwater recharge to be the dominant source of groundwater as opposed to infiltration from surface water.

Girls Point, MI (-13.5‰). These extremely light δ^{18}O compositions are interpreted to have been recharged to the groundwater system during glacial/interglacial periods (as mentioned previously, colder temperatures enhance fractionation efficiency, thus increases in the efficiency illustrated by the significantly lighter δ^{18}O composition indicates recharge during a climate colder than present). Given that the last glacial retreat is estimated to have occurred approximately 9,500 years before present (Clayton, 1984), this result indicates that groundwater in the sandstone near Lake Superior is at least this old.

Assumptions and Limitations

The Bad River groundwater model, like all models, is a simplification of a natural system of ultimately unknowable complexity. Although an effort was made to include all available information in the model, there are areas of the Bad River Watershed for which there is little or no high-quality information on either the groundwater system properties (for example, hydraulic conductivity) or the system state (water levels and flows). Therefore, the model is not only simplified but also conditioned on an incomplete set of information. These two factors contribute to "structural error" in the model (Doherty and Welter, 2010), which can limit its ability to simulate reality, especially at smaller scales where omitted system detail is likely important for predictions of interest. Despite these limitations, the model is a useful quantitative framework for investigating broader questions related to the regional flow system, including the delineation of a contributing area for the Reservation, broad characterization of the recharge areas, flow paths of groundwater throughout the watershed, and the evaluation of data gaps affecting predictions such as large scale dewatering on the Penokee ridge.

Model Discretization and Parameterization

The model discretizes the groundwater system into cells that are 250 ft on a side, and up to hundreds of feet in thickness. Model inputs and outputs therefore represent average values for the volumes of their respective cells. Furthermore, hydraulic conductivity is parameterized at pilot points that are distributed every 44 cells (11,000 ft) and interpolated to cells between them. The estimated hydraulic conductivity parameters are therefore averaging properties across substantial distances, and may correspond poorly to field data collected from a single location. Estimates of uncertainty for hydraulic conductivity values in the major hydrogeologic units types are given in table 4. These uncertainties should be considered in localized applications of the model, unless there are additional field data to constrain the possible values of hydraulic conductivity.

Localized use of the model to simulate flow within the fractured bedrock (for example, in the vicinity of the Penokee ridge and other bedrock hills formed by the MRS igneous units) would benefit from the collection of addition field data and refinement of the model. Very little is known about the local groundwater system in this area, a situation which is especially problematic in that the system is likely dominated by fracture flow. Although perennial flows have been observed at springs and several headwater streams draining the Penokee ridge, including Javorsky and Bull Gus Creeks (fig. 4), these were not well reproduced by the model (fig. 11), which does not simulate a rise in head beneath the Penokee ridge that follows the land surface. The source of water to these streams is not well understood; it may be a shallow groundwater-flow system that is draped over the ridge and mostly unconnected to any deeper flows (if they exist), or it could also stem from a deeper system of connected fractures that extends throughout the ridge. More field data are needed to test the conceptualization of flow in these areas.

The vertical extent of the active groundwater-flow system in bedrock is mostly unknown. Model layers 3 and 4 represent the depths penetrated by most wells, with layer 3, which is 50 ft thick, representing a weathered zone with increased fracturing resulting from the exposure of the bedrock surface for a long period of time prior to the deposition of glacial deposits. Layer 5 is intended to represent any remaining transmissivity below layer 4. In reality, the way in which these thicknesses relate to the actual groundwater-flow system is not well known. Therefore, it is likely that the calibration flexibility provided by the pilot point hydraulic conductivity parameters is compensating for deficiencies in the model layer thicknesses. This is especially true for flow in layer 5, which may be highly variable. Therefore, the model is better thought of in terms of transmissivity, and is less suitable for simulating flow conditions at depth.

Steady-State Assumption

A fundamental assumption is that on average, conditions in the Bad River Watershed are at steady state (that is, any changes occurring in heads or the associated amount of water stored in the system represent a dynamic equilibrium). As a result, all input stresses to the model, including recharge rates, the level of Lake Superior, and flows of water across the model boundary, are assumed to represent long-term average values. Likewise, all outputs of the model, including simulated heads and base flows,

are assumed to represent long-term averages, as are the observations in the calibration dataset.

As discussed in the report sections on model calibration, the assumption of long-term average conditions is problematic for observations that represent only a single or limited number of measurements. Such observations are prone to structural error resulting from temporal variability. Substantial effort was invested in minimizing the impacts of this structural error, through adjustments to one-time base-flow measurements, observation weighting, and regularization of parameter values (appendix 3). However, some of the pilot points with relatively high K values, including the "hot spots" near the Tyler Forks and Potato Rivers (appendix figs. 4.2 through 4.3) are likely compensating for these structural errors. The effects of such compensatory parameter values are not necessarily detrimental (Doherty and Welter, 2010), but caution should be attributed to model predictions that are sensitive to them.

Boundary Conditions

There are several simplifications in the representation of surface water in the model, which may limit its ability to simulate local processes that require accurate representation of surface water. These include the following:

- Most surface flows and processes (with the exception of those related to groundwater discharge) are encapsulated in the recharge array that was simulated by using the SWB model. These include precipitation, snowmelt, canopy interception, runoff, evapotranspiration, and soil moisture storage. SWB calculations do not consider the location of the water table, so rejection of recharge from a high water table is not considered. However, when the recharge array is applied in the model, the groundwater model will reject recharge in areas with heads above land surface. SWB also does not allow for infiltration from areas mapped as open water (such as the lakes in the area south of the Penokee ridge). This limitation was, in part, addressed by using the UZF package, but some lake seepage was not simulated in the model.

- Lakes and larger wetlands are simulated as linear features in the stream network. For example, Caroline Lake is represented by a line of stream cells running through the center of the lake, which are connected to other stream cells representing the Bad River. Therefore, the lake stage is determined by the balance of water in the stream cell as described in appendix 2. Net precipitation (accounting for evapotranspiration) is considered in the recharge

value supplied by SWB (which is zero if the lake is mapped as open water). This approach greatly facilitated the construction of a routed stream network by using existing hydrography datasets (see appendix 2), and is reasonable for regional simulations of groundwater flow, as the net water balance for the area encompassing the lakes is still controlled by the surface processes accounted for in SWB and by discharge through the stream network, which is constrained by base-flow measurements. Groundwater discharge to lakes that receive water from the groundwater system is still simulated by the linear stream cells, even if the areas of the lake receiving discharge are not optimally represented. Infiltration from ephemeral lakes and wetlands to the groundwater system may be accounted for by SWB, which allows for ponding in closed depressions.

- Smaller wetlands and seeps not included in the hydrography datasets used to build the stream network are represented by the UZF package as overland flow from locations where the water table is simulated to be above the land surface. The UZF package distributes this flow evenly among all reaches in the stream segment associated with the catchment where the discharge is occurring (see appendix 2). This approach has the benefit of restricting simulated head values in Layer 1 to the land surface, thereby eliminating unrealistic "flooded" cells, or cells with a groundwater surface simulated higher than the land surface (for example, Hunt and others, 2008). This restriction results in a more realistic head solution and improved simulation of the spatial distribution of groundwater discharge in topographic low-lying areas, which should also improve the simulation of advective transport (particle tracking). In headwater catchments, however, the even application of UZF discharge to all reaches within a user-defined stream segment can be problematic for the level of stream discretization used in this modeling effort, as it can result in the transfer of land-surface discharge occurring in low areas in the catchment to headwater stream cells that should be simulated as dry. The effects of this transfer of discharge are probably small and localized to headwater areas. In the case of the headwater catchment for Graveyard Creek, which presents a typical example of this problem, the total amount of UZF discharge transferred to dry stream-cells was approximately 0.2 ft^3/s or about 5 percent of the total base flow simulated in Graveyard Creek at the headwater catchment outlet. Regardless, good matches between observed and simulated base-flow targets give confidence that the model can simulate

regional flows. Groundwater/surface-water interactions could be further improved by including additional processes to the groundwater model, as illustrated in the following examples:

- A more sophisticated approach to simulating lakes and wetlands would be to use packages such as the MODFLOW Lake Package (Merritt and Konikow, 2000), which calculates a full water balance for each lake. This approach would be recommended for a more focused investigation of the groundwater/surface-water interactions in a single lake or chain of lakes.

- The geometries of some small stream systems may be improved upon. In some cases, there are small gaps between the stream cells derived from NHDPlus and those derived from the WDNR hydrography, due to inconsistencies in the two datasets. The vast majority of these gaps are routed, and many are between headwater tributaries simulated as dry. In a few cases, however, streams in the WDNR dataset are completely isolated and thus were not routed to the stream network by using the algorithm employed here. The overall effects of these problems on the performance of the model are negligible, but they underscore that additional field data and refinement of the model are necessary for its application to localized investigations.

- The constant-flux boundary conditions obtained from the analytic element screening model are limited in that they provide the same flux, regardless of any new stresses added to the model that in reality would induce additional flow across the boundary. For example, new pumping stresses placed near the model boundary would result in simulated heads that are biased low, as all additional flow towards the stresses would have to come from within the model domain, when in reality some flow could come from across the model boundary. For pumping stresses located near the Reservation or the proposed mine site, this limitation of the constant-flux boundary condition has a negligible effect on the solution. It is worth noting that globally, the constant-flux boundary accounts for only 3-4 percent of the model mass balance (Table 3). Constant heads have their own limitation as an alternative boundary condition–to maintain heads they must allow an unlimited amount of flux, which can result in unrealistic flows across the model boundary, and heads that are biased high. In future applications of the model, the effect of the constant-flux boundary conditions can be tested by comparing heads along the boundaries with and without the added stress(es).

Data Gaps

The goal of the calibration process was to improve the model input parameter values by incorporating information contained in the calibration dataset, through the matching of model outputs to the observations. As shown in appendix figure 4.1, in many places pilot-point parameters were identifiable (a reasonable value could be estimated from the observations). In some places, however, where observation data are sparse (including the central portion of the Reservation), pilot-point identifiability values were low. In these areas the "calibrated" pilot-point parameter values reflect professional judgment as reflected in the initial (precalibration) values and (or) the values of neighboring pilot points. Identifiability values for vertical hydraulic conductivity were generally low (appendix fig. 4.1), reflecting a lack of vertical head-difference targets.

In other parts of the model, the calibration process may have estimated physically inaccurate values to pilot points, in order to compensate for structural error in the model and achieve the best fit possible with important observations such as large stream gages (in lieu of constraining local data). This may be the case with the "bulls-eyes" in the pilot point field near the Tyler Forks and Potato Rivers. In either of these cases, predictions that are sensitive to these pilot point values may be of lower quality than other predictions made by the model.

This issue again emphasizes the importance of additional field data collection and refinement of the model when using it to make localized predictions. The data collected should be relevant to the model purpose, as important parameters are defined by their importance to the prediction. The new data-worth analysis technique can articulate the worth of data to any prediction that can be simulated by the model. Additional accuracy may also be gained by recalibrating the model in the context of the desired prediction, as it is increasingly recognized that the best model calibrations are prediction-specific (White and others, 2014).

Summary and Conclusions

The Bad River Watershed encompasses the cultural homeland of the Bad River Band of Lake Superior Chippewa and the largest remaining wild rice beds in the Great Lakes Basin. The groundwater system provides the primary drinking-water source for the Bad River Band and other inhabitants, sustains numerous coldwater and coolwater habitats, and also may be important to the wild rice beds in the Bad River and Kakagon sloughs. The hydrogeology of the area reflects a long and varied geologic history and has not been well studied.

Understanding of the regional groundwater-flow system and groundwater/surface-water interactions is important to the Bad River Band and other stakeholders, as they seek to protect and manage the watershed in the face of climate change and changes in land use within and outside of the Reservation boundaries.

The groundwater-flow model described in this report was constructed to: (1) improve understanding of the groundwater-flow system in the Bad River Watershed at the regional scale, including the sources of water to the Reservation and groundwater/surface-water interactions; (2) provide a quantitative platform to allow for the evaluation of potential future impacts to specific locations within the watershed; and (3) identify areas where more data are needed. Primary conclusions from the study include the following:

- Previously published and unpublished data were used to construct and calibrate the model. Available measurements of groundwater levels and stream base flows are well-simulated by the calibrated model, with input parameters that are reasonable based on what is known about the hydrogeology of the area. As currently (2015) calibrated, the model is suitable as a regional water-management tool and provides valuable quantitative information on the groundwater system and groundwater/surface-water interactions.

- The southern part of the model domain, extending from the areas underlain by igneous units of the Midcontinent Rift System southward, is characterized by shallow, localized flow systems in mostly thin quaternary deposits and fractured bedrock, with relatively short flow paths from local recharge areas to nearby streams and wetlands. As evidenced by the model and consistent with stable isotope results, groundwater in this area appears to be relatively young, on the order of several hundred years old or less.

- The northern part of the model domain, encompassing the Superior Lowlands, is characterized by a confined flow system in the deep sandstone aquifer, which is isolated from many surface-water features by thick deposits of clay-rich till. As a result of this isolation, groundwater discharge is mostly restricted to the major river valleys and Lake Superior. Limited connection with surface water allows for longer flow paths, less turnover of groundwater by recharge, and therefore older water. Based on stable isotope results, much of the water in the sandstone aquifer beneath the northern part of the Reservation is likely at least 9,500 years old, having an oxygen-18 (δ^{18}O)

signature consistent with recharge during the cooler climatic conditions of the last glaciation.

- The model was used to delineate the groundwatershed for Bad and Kakagon Rivers, defined as the area contributing groundwater discharge to these streams and their tributaries. The groundwatershed was generally found to align with the surface watershed, with the most notable exception occurring along the southwestern half of Birch Hill, where surface water drains southwest towards the Potato River, and groundwater flows north and east towards Lake Superior. Similarly, the contributing area of groundwater flow to the Reservation was delineated. The results indicate the off-Reservation groundwater contributing area to be limited in comparison to the extent of the watershed, extending southward into the highlands underlain by Midcontinent Rift System igneous rock units, but not further into the area underlain by the Marquette Range Supergroup.

- Data-worth analysis was performed in the context of predicting responses of groundwater levels and baseflows to a hypothetical future dewatering stress. The results indicate a diversity of reductions in uncertainty for the 35 predictions investigated. In general, the predictions near the proposed mine site showed greatest data worth in new monitoring wells located either in the immediate vicinity of the proposed mine, or between the proposed mine and the prediction location. Other predictions located away from the proposed mine site showed either diffuse or no data worth, or data worth concentrated around local phenomena such as groundwater divides between competing sinks. The ability of new monitoring points to reduce prediction uncertainty appears to be greatest for predictions that are sensitive to small areas where existing data are sparse.

- In many parts of the model, observation data are sparse, and the local reliability of the model is uncertain. In general, the regional focus and scale of the model limit the usefulness in site-specific investigations, especially those focusing on small surface-water features, or where fine-scale detail in the hydraulic conductivity field such as preferential flow through fractures is important. Another potential concern is compensatory parameterization, where pilot-point parameters assume calibrated values that may not reflect field values in order to compensate for structural error and other defects in the model. Site-specific investigations should always be supported by additional data, and possibly refinement

of the model or even recalibration. The new data-worth analysis techniques can help identify the most useful locations for additional monitoring.

References Cited

Ackerman, J.A., 1992, Extending the isotope based (δ^{18}O) mass budget technique for lakes and comparison with solute based lake budgets: M.S. thesis, University of Wisconsin-Madison, Department of Geology and Geophysics, 151 p.

Batten, W.G. and Lidwin, R.A., 1995, Water resources of the Bad River Indian Reservation, northern Wisconsin: U.S. Geological Survey Water-Resources Investigations Report 95–4207, 45 p.

Black, Alastair and Foley, Christopher, 2013, FlowSource—a program to efficiently delineate volumetric capture areas, pathways and source areas in groundwater models, *in* MODFLOW and More 2013—Translating Science into Practice: Colorado School of Mines, Golden, Colo., June 2–5, 2013, accessed April 21, 2015, at http://igwmc.mines.edu/conference/abstracts/22_Black.pdf.

Bradbury, K.R., 2013, Hydraulic conductivity estimates for the CCNF Washburn–Great Divide Unit (unpublished data), Wisconsin Geological and Natural History Survey.

Buchwald, C.A., 2009, Water use in Wisconsin, 2005: U.S. Geological Survey Open-File Report 2009–1076, 76 p.

Cannon, W.F., LaBerge, G.L., Klasner, J.S., and Schulz, K.J., 2007, The Gogebic iron range—A sample of the northern margin of the Penokean fold and thrust belt: U.S. Geological Survey Professional Paper 1730, 44 p.

Cannon, W.F., Woodruff, L.G., Nicholson, S.W., and Hedgman, C.A., 1996, Digital Bedrock Geologic Map of the Ashland and the Northern Part of the Ironwood $30' \times 60'$ Quadrangles, Wisconsin and Michigan: U.S. Geological Survey Miscellaneous Investigations Series Map I-2566, scale 1:100,000.

Cannon, W.F., Woodruff, L.G., Nicholson, S.W., Hedgman, C.A., and Barber-Delach, R.D., 1999, Bedrock geologic map of the Ashland and the northern part of the Ironwood $30' \times 60'$ quadrangles, Wisconsin and Michigan: U.S. Geological Survey Open-File Report 99–546, accessed November 13, 2014, at http://pubs.usgs.gov/of/1999/of99-546/.

Clark, I.D. and Fritz, Peter, 1997, Environmental isotopes in hydrogeology: Boca Raton, Fla., Lewis Publishers, 342 p.

Clayton, Lee, 1984, Pleistocene geology of the Superior Region, Wisconsin: Wisconsin Geological and Natural History Survey Information Circular 46, 40 p., plus map (scale 1:250,000).

Doherty, John, 2003, Ground water model calibration using pilot points and regularization: Ground Water, v. 41, no. 2, 170–177, http://dx.doi.org/10.1111/j.1745-6584.2003.tb02580.x.

Doherty, John, 2010a, PEST, Model-independent parameter estimation—user manual (5th ed., with slight additions): Brisbane, Australia, Watermark Numerical Computing, [333 p.] accessed November 13, 2014, at http://www.pesthomepage.org/Downloads.php.

Doherty, John, 2010b, Methodologies and software for PEST-based model predictive uncertainty analysis: Brisbane, Australia, Watermark Numerical Computing, [160 p.] accessed November 13, 2014, at http://www.pesthomepage.org/.

Doherty, John, 2014, Addendum to the PEST manual: Brisbane, Australia, Watermark Numerical Computing [316 p.], accessed November 13, 2014, at http://www.pesthomepage.org/Downloads.php.

Doherty, John and Hunt, R.J., 2009, Two statistics for evaluating parameter identifiability and error reduction: Journal of Hydrology, v. 366, no. 1–4, p. 119–127, http://dx.doi.org/10.1016/j.jhydrol.2008.12.018.

Doherty, J.E. and Hunt, R.J., 2010, Approaches to highly parameterized inversion—a guide to using PEST for groundwater-model calibration: U.S. Geological Survey Scientific Investigations Report 2010–5169, 59 p.

Doherty, John and Welter, David, 2010, A short exploration of structural noise: Water Resources Research, v. 46, no. 5, W05525, http://dx.doi.org/10.1029/2009wr008377.

Doherty, J.E., Fienen, M.N., and Hunt, R.J., 2010, Approaches to highly parameterized inversion—pilot-point theory, guidelines, and research directions: U.S. Geological Survey Scientific Investigations Report 2010–5168, 36 p.

Dunning, C.P., 2005, Lithology, hydraulic properties, and water quality of the sandstone aquifer in the northwestern part of the Bad River Indian Reservation, Wisconsin, 1998–1999: U.S. Geological Survey-Open-File Report 2004–1425, 39 p.

Esri, 2014, ArcGIS platform accessed June 15, 2014, at http://www.esri.com/software/arcgis.

Feinstein, D.T., Eaton, T.T., Hart, D.J., Krohelski, J.T., and Bradbury, K.R., 2005, Regional aquifer model for southeastern Wisconsin, report 1—data collection, conceptual model development, numerical model construction, and model calibration: Southeastern Wisconsin Regional Planning Commission Technical Report No. 41, 81 p.

Feinstein, D.T., Hunt, R.J., and Reeves, H.W., 2010, Regional groundwater-flow model of the Lake Michigan Basin in support of Great Lakes Basin water availability and use studies: U.S. Geological Survey Scientific Investigations Report 2010–5109, 379 p.

Fienen, M., Hunt, R., Krabbenhoft, D., and Clemo, T., 2009a, Obtaining parsimonious hydraulic conductivity fields using head and transport observations—a Bayesian geostatistical parameter estimation approach: Water Resources Research, v. 45, no. 8, W08405, http://dx.doi.org/10.1029/2008wr007431.

Fienen, M.N., Muffels, C.T., and Hunt, R.J., 2009b, On constraining pilot point calibration with regularization in PEST: Ground Water, v. 47, no. 6, p. 835–844, http://dx.doi.org/10.1111/j.1745-6584.2009.00579.x.

Fienen, M.N., Doherty, J.E., Hunt, R.J., and Reeves, H.W., 2010, Using prediction uncertainty analysis to design hydrologic monitoring networks—example applications from the Great Lakes Water Availability Pilot Project: U.S. Geological Survey Scientific Investigations Report 2010–5159, 44 p.

Fry, J.A., Xian, G., Jin, S., Dewitz, J.A., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011, Completion of the 2006 National Land Cover Database for the conterminous United States: Photogrammetric Engineering and Remote Sensing, v. 77, no. 9, p. 858–864.

Gat, J.R., 1970, Environmental isotope balance of Lake Tiberias *in* Isotope Hydrology, Symposium on Use of Isotopes in Hydrology, Vienna, Austria, 1970, International Atomic Energy Agency, Vienna, Austria, p. 109–127.

Gebert, W.A., Radloff, M.J., Considine, E.J., and Kennedy, J.L., 2007, Use of streamflow data to estimate base flow/ground-water recharge for Wisconsin: Journal of the American Water Resources Association, v. 43, no. 1, p. 220–236, http://dx.doi.org/10.1111/j.1752-1688.2007.00018.x.

Gebert, W.A., Walker, J.F., and Kennedy, J.L., 2011, Estimating 1970–99 average annual groundwater recharge in Wisconsin using streamflow data: U.S. Geological Survey Open–File Report 2009–1210, 14 p., plus appendixes, available at http://pubs.usgs.gov/ofr/2009/1210/.

Gogebic Taconite, 2013, Preapplication notice, accessed June 5, 2014, at http://dnr.wi.gov/topic/mines/gogebic.html.

Haitjema, H.M., 1995, Analytic element modeling of groundwater flow: San Diego, Academic Press, 394 p.

Harbaugh, A.W., 2005, MODFLOW–2005, the U.S. Geological Survey modular ground-water model—the ground-water flow process: U.S. Geological Survey Techniques and Methods 6–A16, variously paginated.

Hunt, R., 1996, Do created wetlands replace the wetlands that are destroyed?: U.S. Geological Survey Fact Sheet 246–96, 4 p.

Hunt, R., Anderson, M., and Kelson, V., 1998, Improving a complex finite-difference ground water flow model through the use of an analytic element screening model: Ground Water, v. 36, no. 6, p. 1011–1017, http://dx.doi.org/10.1111/j.1745-6584.1998.tb02108.x.

Hunt, R.J., Borchardt, M.A., Richards, K.D., and Spencer, S.K., 2010, Assessment of sewer source contamination of drinking water wells using tracers and human enteric viruses: Environmental Science & Technology, v. 44, no. 20, p. 7956–7963, http://dx.doi.org/10.1021/es100698m.

Hunt, R.J., Bullen, T.D., Krabbenhoft, D.P., and Kendall, C., 1998, Using stable isotopes of water and strontium to investigate the hydrology of a natural and a constructed wetland: Ground Water, v. 36, no. 3, 434–443, http://dx.doi.org/10.1111/j.1745-6584.1998.tb02814.x.

Hunt, R.J., Coplen, T.B., Haas, N.L., Saad, D.A., and Borchardt, M.A., 2005, Investigating surface water-well interaction using stable isotope ratios of water: Journal of Hydrology, v. 302, no. 1–4, p. 154–172, http://dx.doi.org/10.1016/j.jhydrol.2004.07.010.

Hunt, R.J., Doherty, John, and Tonkin, M.J., 2007, Are models too simple?—Arguments for increased parameterization: Ground Water, v. 45, no. 3, p. 254–262, http://dx.doi.org/10.1111/j.1745-6584.2007.00316.x.

Hunt, R.J. and Feinstein, D.T., 2012, MODFLOW–NWT—Robust handling of dry cells using a Newton Formulation of MODFLOW–2005: Ground Water, v. 50, no. 5, p. 659–663.

Hunt, R.J., Feinstein, D.T., Pint, C.D., and Anderson, M.P., 2006, The importance of diverse data types to calibrate a watershed model of the Trout Lake Basin, northern Wisconsin, USA: Journal of Hydrology, v. 321, no. 1–4, p. 286–296, http://dx.doi.org/10.1016/j.jhydrol.2005.08.005.

Hunt, R.J., Haitjema, H.M., Krohelski, J.T., and Feinstein, D.T., 2003, Simulating ground water-lake interactions—approaches and insights: Ground Water, v. 41, no. 2, p. 227–237, http://dx.doi.org/10.1111/j.1745-6584.2003.tb02586.x.

Hunt, R., Krabbenhoft, D., and Anderson, M., 1996, Groundwater inflow measurements in wetland systems: Water Resources Research, v. 32, no. 3, p. 495–507, http://dx.doi.org/10.1029/95WR03724.

Hunt, R., Prudic, D., Walker, J., and Anderson, M., 2008, Importance of unsaturated zone flow for simulating recharge in a humid climate: Ground Water, v. 46, no. 4, p. 551–560, http://dx.doi.org/10.1111/j.1745-6584.2007.00427.x.

Hunt, R.J. and Steuer, J.J., 2000, Simulation of the recharge area for Frederick Springs, Dane County, Wisconsin: U.S. Geological Survey Water–Resources Investigations Report 00–4172, 33 p.

Hunt, R.J., Walker, J.F., Selbig, W.R., Westenbroek, S.M., and Regan, R.S., 2013, Simulation of climate-change effects on streamflow, lake water budgets, and stream temperature using GSFLOW and SNTEMP, Trout Lake Watershed, Wisconsin: U.S. Geological Survey Scientific Investigations Report 2013–5159, 118 p.

Institute of Hydrology, 1980a, Low flow studies report no. 1—Research report: Wallingford, Oxon, United Kingdom, Institute of Hydrology Report no. 1, 41 p.

Institute of Hydrology, 1980b, Low flow studies report no. 3—Research report: Wallingford, Oxon, United Kingdom, Institute of Hydrology Report no. 3, p. 12-19

Krabbenhoft, D.P., Bowser, C.J., Anderson, M.P., and Valley, J.W., 1990, Estimating groundwater exchange with lakes—1. The stable isotope mass balance method.: Water Resources Research, v. 26, no. 10, p. 2445–2453, http://dx.doi.org/10.1029/wr026i010p02445.

Krabbenhoft, D.P., Bowser, C.J., Kendall, C., and Gat, J.R., 1994, Use of oxygen-18 and deuterium to assess the hydrology of groundwater-lake systems, *in* Baker, L., ed., Environmental chemistry of lakes and reservoirs: American Chemical Society Advances in Chemistry Series no. 237, p. 67–90.

Kurtz, A.M., Bahr, J.M., Carpenter, Q.J., and Hunt, R.J., 2007, The importance of subsurface geology for water source and vegetation communities in Cherokee Marsh, Wisconsin: Wetlands, v. 27, no. 1, p. 189–202.

Lenz, B.N., Saad, D.A., and Fitzpatrick, F.A., 2003, Simulation of ground-water flow and rainfall runoff with emphasis on the effects of land cover, Whittlesey Creek, Bayfield County, Wisconsin, 1999–2001: U.S. Geological Survey Water–Resources Investigations Report 03–4130, 47 p.

Lott, R.B. and Hunt, R.J., 2001, Estimating evapotranspiration in natural and constructed wetlands: Wetlands, v. 21, no. 4, p. 614–628.

Magner, J., Zhang, L., Engel, L., and Jasperson, J., 2014, Stable isotopes in NSLS watershed management *in* proceedings of the MN Lake Superior Watershed Stream Science Symposium, January 7, Laurentian Resource Conservation and Development Council, Duluth, MN, accessed October 17, 2014, at http://www.lrcd.org/stream-science-symposium.html.

McDonald, M.G. and Harbaugh, A.W., 1988, A modular three-dimensional finite-difference ground-water flow model: U.S. Geological Survey Techniques of Water Resources Investigations, book 6, chap. A6, 576 p.

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., and Rea, A., 2012, NHDPlus Version 2—User guide, accessed November 15, 2014, at http://www.horizon-systems.com/NHDPlus/NHDPlusV2_documentation.php.

Merritt, M.L. and Konikow, L.F., 2000, Documentation of a computer program to simulate lake-aquifer interaction using the MODFLOW ground-water flow model and the MOC3D solute-transport model: U.S. Geological Survey Water–Resources Investigations Report 00–4167, 146 p.

Michaels, S.S., 1995, Regional analysis of lakes, groundwaters, and precipitation, northern Wisconsin: A stable isotope study: M.S. thesis, University of Wisconsin-Madison, Department of Geology and Geophysics M.S. thesis, 161 p.

Mickelson, D.M., Clayton, L., Knox, J.C., and Cahow, A., 1983, Quaternary Geologic Map of the Minneapolis 4° X 6° Quadrangle, United States, Wisconsin Compilation: U.S. Geological Survey Map I-1420 (NL-15).

Moore, Catherine and Doherty, John, 2005, Role of the calibration process in reducing model predictive error: Water Resources Research, v. 41, no. 5, W05020, 14 p., http://dx.doi.org/10.1029/2004WR003501.

Moore, Catherine and Doherty, John, 2006, The cost of uniqueness in groundwater model calibration: Advances in Water Resources, v. 29, no. 4, p. 605–623, http://dx.doi.org/10.1016/j.advwatres.2005.07.003.

Moore, C. and Wöhling, T. and Doherty, J., 2010, Efficient regularization and uncertainty analysis using a global optimization methodology: Water Resources Research, v. 46, no. 8, W08527, http://dx.doi.org/10.1029/2009wr008627.

National Atmospheric and Oceanic Administration, 2013, Hydrographs: Lake Superior, accessed June 15, 2014, at http://www.glerl.noaa.gov/data/now/wlevels/lowlevels/plot/Superior.jpg.

National Oceanic and Atmospheric Administration, 2014, Bathymetry of Lake Superior, accessed June 15, 2014, at http://www.ngdc.noaa.gov/mgg/greatlakes/superior.html.

Niswonger, R.G., Panday, S., and Ibaraki, M., 2011, MODFLOW–NWT—A Newton formulation for MODFLOW–2005: U.S. Geological Survey Techniques and Methods, book 6, chap. A37, 44 p.

Niswonger, R.G. and Prudic, D.E., 2005, Documentation of the Streamflow-Routing (SFR2) Package to include unsaturated flow beneath streams—A modification to SFR1: U.S. Geological Survey Techniques and Methods book 6, chap. A13, 50 p.

Niswonger, R.G., Prudic, D.E., and Regan, R.S., 2006, Documentation of the Unsaturated-Zone Flow (UZF1) Package for modeling unsaturated flow between the land surface and the water table with MODFLOW–2005: U.S. Geological Survey Techniques and Methods book 6, chap. A19, 62 p.

Ojakangas, R.W., Dickas, A.B., and Green, J.C., 1997, Introduction—Middle Proterozoic to Cambrian rifting, central North America: Geological Society of America Special Papers, v. 312, p. 1–5, http://dx.doi.org/10.1130/0-8137-2312-4.1.

Ojakangas, R.W., Morey, G.B., and Green, J.C., 2001, The Mesoproterozoic Midcontinent Rift System, Lake Superior Region, USA: Sedimentary Geology 141–142 (2001), p. 421–442, p. 1–5, http://dx.doi.org/10.1016/S0037-0738(01)00085-9.

Prudic, D.E., Konikow, L.F., and Banta, E.R., 2004, A new Streamflow-Routing (SFR1) Package to simulate stream-aquifer interaction with MODFLOW–2000: U.S. Geological Survey Open–File Report 2004–1042, 95 p.

Runkel, A.C., Tipping, R.G., Alexander, E. C., Jr., Green, J.A., Mossler, J.H., and Alexander, S.C., 2003, Hydrogeology of the Paleozoic bedrock in southeastern Minnesota: Minnesota Geological Survey Report of Investigations 61, 105 p., 2 pls.

Schreüder, W.A., 2009, Running BeoPEST, in PEST Conference 2009, Potomac, Maryland, November 2–4, 2009, Proceedings: Bethesda, Md., S.S. Papadopulos & Associates, Inc., 304 p.

Strack, O. D.L., 1984, Three-dimensional streamlines in Dupuit-Forchheimer models: Water Resources Research, v. 20, p. 812–822, http://dx.doi.org/10.1029/WR020i007p00812.

Strack, O. D.L., 1989, Groundwater mechanics: Prentice Hall, 732 p.

Syverson, K.M., Clayton, Lee, Attig, J.W., and Mickelson, D.M., eds., 2011, Lexicon of Pleistocene Stratigraphic Units of Wisconsin: Wisconsin Geological and Natural History Survey Technical Report 1, 180 p.

Thornthwaite, C.W., and Mather, J.R., 1955, The water balance: Publications in Climatology, v. 8, p. 1–104 p.

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wilhelmi, N., Wei, Y., Devarakonda, R., and Cook, R.B., 2014, Daymet: Daily surface weather data on a 1-km grid for North America (ver. 2), data set, accessed april 24, 2015, at http://daymet.ornl.gov.

Tikhonov, A.N., 1963a, Solution of incorrectly formulated problems and the regularization method: Soviet Mathematics Doklady, v. 4, p. 1035–1038.

Tikhonov, A.N., 1963b, Regularization of incorrectly posed problems: Soviet Mathematics Doklady, v. 4, p. 1624–1637.

Tonkin, M.J. and Doherty, John, 2005, A hybrid regularized inversion methodology for highly parameterized environmental models: Water Resources Research, v. 41, no. 10, W10412, http://dx.doi.org/10.1029/2005WR003995.

U.S. Department of Agriculture-Natural Resources Conservation Service, 1986, Urban hydrology for small watersheds (2d ed.): Washington, D.C., Conservation Engineering Division, Technical Release 55, Washington, D.C., 164 p.

United States Department of Agriculture-Natural Resources Conservation Service, 2012, Soil Survey Geographic (SSURGO) Database for Black Presque Isle, Flambeau, Namekagon, Upper Chippewa, Beartrap-Nemadji, and Bad-Montreal survey areas, Wisconsin, accessed October 18, 2013, at http://www.arcgis.com/apps/OnePane/basicviewer/index.html?appid=a23eb436f6ec4ad6982000dbaddea5ea.

U.S. Geological Survey, 2014a, National Elevation Dataset, accessed June 15, 2014, at http://ned.usgs.gov/.

U.S. Geological Survey, 2014b, Online Guide to MODFLOW–2005—LAK-Lake Package, accessed April 24, 2015, at http://water.usgs.gov/ogw/modflow/MODFLOW-2005-Guide/index.html?lak.htm.

Wahl, T.W. and Wahl, K.L., 2013, BFI—a computer program for determining an index to base flow: U.S. Department of the Interior, Bureau of Reclamation, accessed june 12, 2013, at http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/.

Walker, J.F., Hunt, R.J., Bullen, T.D., Krabbenhoft, D.P., and Kendall, Carol, 2003, Spatial and temporal variability of isotope and major ion chemistry in the Allequash Basin, northern Wisconsin: Ground Water, v. 41, no. 7, p. 883–894, http://dx.doi.org/10.1111/j.1745-6584.2003.tb02431.x.

Weidman, Samuel and Schultz, A.R., 1915, The underground and surface water supplies of Wisconsin: Wisconsin Geological and Natural History Survey Bulletin No. 35, Economic Series No. 17, 664 p.

Westenbroek, S.M., Kelson, V.A., Dripps, W.R., Hunt, R.J., and Bradbury, K.R., 2010, SWB—A Modified Thornthwaite-Mather Soil-Water-Balance code for estimating groundwater recharge: U.S. Geological Survey Techniques and Methods book 6, chap. A31, 60 p.

Wheeler, M. and Bodette, C., 2011, Bad River Watershed Culvert Restoration Program, *in* Institute on Lake Superior Geology, 57th Annual Meeting: Ashland, Wis., v. 57, part 2, p. 31–46.

White, J.T., Doherty, J.E., and Hughes, J.D., 2014, Quantifying the predictive consequences of model error with linear subspace analysis: Water Resources Research, v. 50, no. 2, p. 1152–1173, http://dx.doi.org/10.1002/2013WR014767.

Wisconsin Department of Administration-Demographic Services Center, 2010, Population estimates by time series—minor civil division final population estimates by county, accessed February 28, 2012, at http://www.doa.state.wi.us/divisions/intergovernmental-relations/demographic-services-center/estimates#timeseries.

Wisconsin Department of Natural Resources, 2014a, Gogebic Taconite, LLC, potential mining project, accessed June 5, 2014, at http://dnr.wi.gov/topic/mines/gogebic.html.

Wisconsin Department of Natural Resources, 2014b, Surface water (hydrography) data, accessed June 17, 2014, at http://dnr.wi.gov/maps/gis/datahydro.html.

Wold, R., 1979, Maps showing bedrock topography of Lake Superior: U.S. Geological Survey Miscellaneous Field Studies Map 1174, scale 1:600,000.

Young, H.L. and Skinner, E.L., 1974, Water resources of Wisconsin—Lake Superior basin: U.S. Geological Survey Hydrologic Investigations Atlas HA-524, 3 sheets.

Appendix 1—Analytic Element Screening Model

An analytic element screening model based on the program GFLOW (Haitjema, 1995) was used to simulate regional groundwater flow beyond the boundaries of the MODFLOW model. Screening models are meant to be fast to construct at the expense of system detail (for example simulating only two-dimensional flow). The analytic element method is well suited for screening models because it does not require the construction of a model grid with finite boundaries (Hunt and others, 1998). Instead, surface-water bodies and other hydrologic features are digitized into the model as elements, and analytic solutions for each element are added together to obtain an overall solution for groundwater flow that is continuous and extends to infinity. The effective limits of the modeled area can be easily expanded or contracted by adding or removing elements, and interaction with distant hydraulic boundaries (for example, rivers and lakes) can be accounted for directly and with a minimum of effort compared to redeveloping a finite-difference model. In the areas of interest, additional detail can be easily added by refining the elements in those locations. Detailed descriptions of the analytic element method can be found in Strack (1989) and Haitjema (1995).

The screening model has the following simplifying assumptions:

- the groundwater system is simulated as an areal two-dimensional system. Vertical resistance to flow and vertical variability in geologic materials are neglected; vertical dimensions of flow are solved by mass balance rather than Darcy's law (Strack, 1984).

- Recharge and hydraulic conductivity parameters are represented with piecewise-constant zones (inhomogeneities) that extend over large areas.

- A single value for the aquifer bottom elevation was assumed for the entire model. Since GFLOW simulates an areal two-dimensional groundwater flow, aquifer transmissivities (hydraulic conductivity multiplied by aquifer thickness) are ultimately used to solve the groundwater-flow governing equation; therefore, changes in thickness are implicitly addressed during hydraulic conductivity calibration.

- The geometries of surface-water bodies are approximately represented by strings of "linesink" elements.

- A steady-state is assumed (heads and flows can change in space but not in time).

Due to these simplifying assumptions, the screening model may not be suitable for simulating flow at smaller scales, where spatial variation in recharge and detail in surface-water geometry become important, or where vertical gradients play an important role. The screening model also is not well suited for parts of the Bad River Watershed where large changes in aquifer thickness and bottom topography over short distances violate the assumption of a piecewise-constant bottom elevation. Finally, it can not be used to address future questions involving rates of change in water levels or flows over time due to the steady-state assumption. Despite these limitations, the GFLOW model is well suited to provide perimeter hydraulic boundary conditions for the more detailed MODFLOW model, and its ease and flexibility for refining hydrogeological features at a variety of scales make it also well suited for quick exploratory modeling at smaller scales.

Construction

The approximate east-west extent of the screening model domain is from the Gile Flowage near Ironwood, Michigan, to Lake Owen, near Cable, Wisconsin. In the north-south direction, the model extends from Lake Superior to near the southern limits of Ashland and Iron Counties, on the opposite side of the Bad River Watershed divide (fig. 1.1). Surface-water features were digitized into the model as strings of linesink elements with heads specified based on elevations determined from topographic maps. Streams in the nearfield area of the model were represented with stream elements, represented as greater linesink resolution head-dependent flux boundaries, with hydraulic resistance values to account for resistance to flow caused by lower conductivity material in the streambed sediments. The streamflow routing functionality of GFLOW's stream elements was used to simulate the accumulation of base flow along consecutive linesinks. Streams in the farfield were represented in coarser detail as constant-head boundaries, without routing or streambed resistance. All lakes except Caroline, McCarthy and O'Brien were simulated with stream linesinks around their perimeter. Caroline, McCarthy and O'Brien were simulated by using the Lake element functionality (Hunt and others, 2003) in GFLOW, where stage is computed based on a lake water budget consisting of groundwater inflow, net precipitation, and a supplied table of outflow rates for various stages. However, since the outflows of these lakes are not well characterized, the lakes were simulated to maintain their known stages by using the outflowing stream component of the lake water budget. This was accomplished by setting very high outflow rates at stages above the known lake stages, and no outflow at stages below the known stages of the lakes.
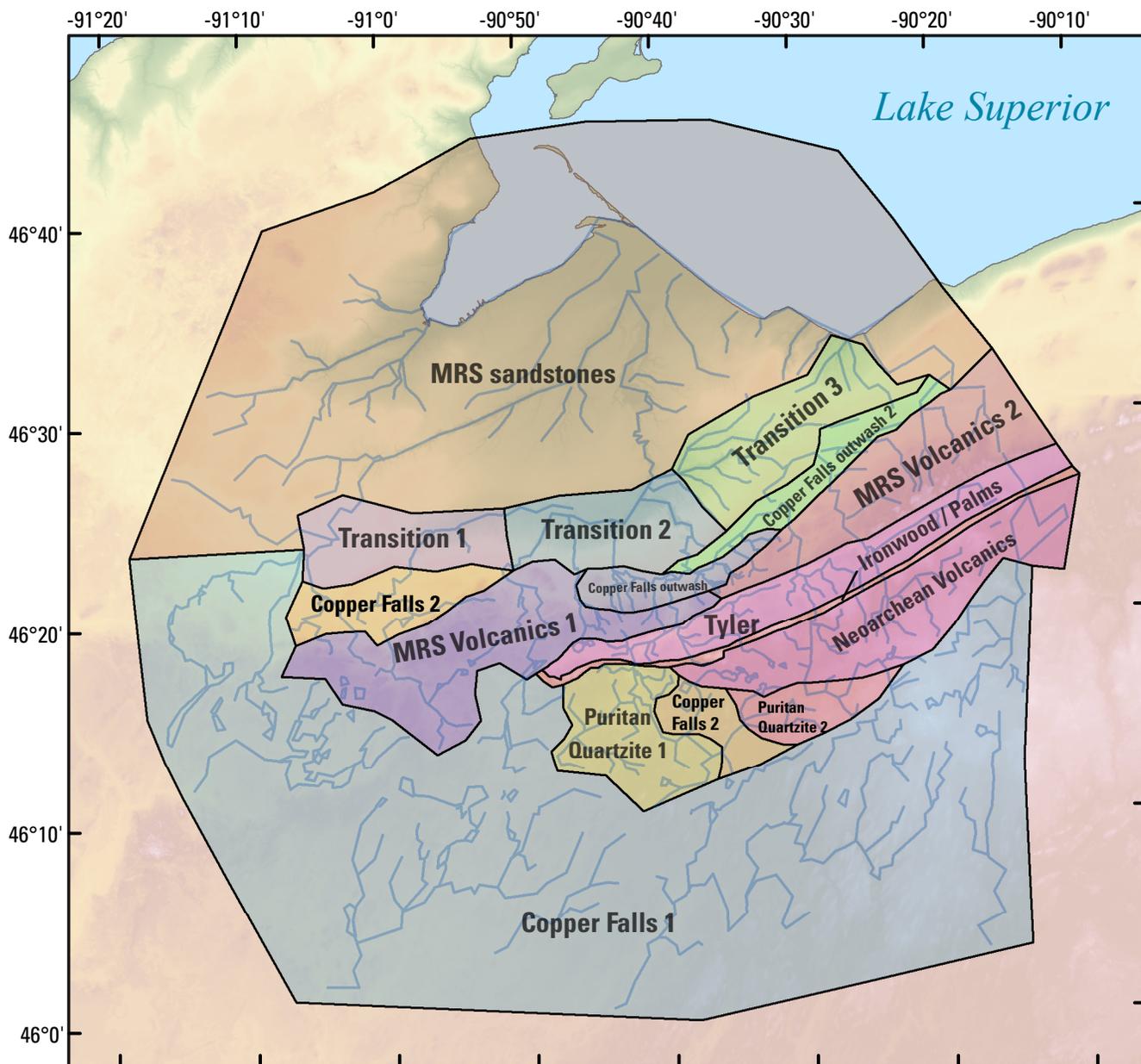
Figure 1.1. Depiction of GFLOW model showing inhomogeneities and the stream (linesink) network.

Inhomogeneity elements were developed to represent variability in hydraulic conductivity, recharge, and aquifer thickness across the modeled area (fig. 1.1). A uniform model base elevation of 500 ft above mean sea level was assumed across the model domain, and was not adjusted during calibration. Aquifer thickness was set to 2,000 ft, above the highest water levels simulated, to simulate unconfined flow (where aquifer thickness is dependent on the simulated water levels).

Parameterization and Calibration

The screening model was calibrated by using the PEST parameter-estimation software (Doherty, 2010a). PEST systematically adjusted model inputs to minimize the misfit between the observed field data and the model simulated equivalent. This fit is represented numerically by an objective function, which is the sum of squared, weighted residuals between model outputs and their equivalent "observed" values (the calibration targets). Recharge and hydraulic conductivity for each inhomogeneity was included as an adjustable parameter during the calibration process. Stream linesinks were categorized based on underlying lithology and whether or not they were in an area of high slope. The resistance linesink representing Lake Superior was included as its own category. Each linesink category was included as an adjustable parameter. During calibration, a total of 31 parameters were adjusted. The observations and calibration strategy were similar to that used for the MODFLOW model, except that preferred-value Tikhonov regularization was employed on all parameters, and a Pareto analysis was used to determine the final parameter set to use for extracting perimeter boundary conditions for the MODFLOW model (appendix 3).

Screening Model Results

The screening model provided initial insight into the groundwater system and connections to surface water. The simplification of the hydrogeologic conditions resulted in spatial bias in the model's ability to fit the field measured data expressed through analysis of the residuals. Alternative parameterization schemes were not evaluated as the purpose of the screening model was to provide perimeter boundary conditions for a three-dimensional and more spatially heterogeneous MODFLOW numerical model—an approach better suited for evaluating calibration residuals. The screening model provided constant-flux boundary conditions to the perimeter of a subsequent inset MODFLOW model. This allowed for efficient simulation of large-scale regional drivers of groundwater flow while retaining nearfield resolution in the numerical modeling.

Appendix 2—MODFLOW Model Construction

Discretization/Layering

The Bad River groundwater model is discretized by a grid of 800 rows by 800 columns, with a uniform spacing of 250 ft. This resolution is intended to be fine enough for the accurate representation of stream geometry and groundwater/surface-water interactions at the regional scale, while limiting the computational burden associated with solving the model. The model is discretized vertically into five contiguous layers. The geometry of the layering is controlled by the land surface, at the top of Layer 1, and the bedrock surface, at the top of Layer 3. Layers 1 and 2 represent equal portions of the interval between the land surface and the top of bedrock. Layer 3 represents shallow bedrock, to a uniform depth of 50 ft below the bedrock surface, where fracturing may be more abundant due to weathering or phenomena such as uplift (including isostatic unloading from recent glaciations), or erosion of younger bedrock (e.g., Runkel and others, 2003). Layer 4 was assigned a uniform thickness of 300 ft. The bottom of Layer 5 is smoothed relative to the bottom of Layer 4 (see step 6 in the next section), and extends approximately 650 ft below Layer 4 to a depth of 1,000 ft below the bedrock surface. The true bottom of the flow system is unknown, so Layer 5 is intended to simulate the remaining transmissivity in the bedrock below Layer 4. Approximately 95 percent of bedrock pumping wells intersect Layers 3 and 4; only 5 percent are open exclusively to Layer 5.The model is steady-state, representing long-term conditions in dynamic equilibrium.

Creation of the Layer Surfaces

Layer 1 top elevations (excluding cells representing Lake Superior) were developed from a 10-meter digital elevation model (DEM) (U.S. Geological Survey, 2014a), by using the arithmetic means of the DEM elevation values contained within each MODFLOW cell. Layer 1 top elevations for the cells representing Lake Superior were computed from bathymetric data for the lake (National Oceanic and Atmospheric Administration, 2014). The bedrock elevation surface was developed from bedrock surface elevations in WCRs, outcrop locations from Cannon and others (1999), and seismic surveys by Batten and Lidwin (1995). The goal in constructing the bedrock surface was to be consistent with these data without imposing extraneous detail from the modern land surface in areas covered by thick quaternary deposits.

1. Point values of depth to bedrock obtained from well construction reports (WCRs), bedrock outcrop data (Cannon and others, 1999), seismic reflection studies (Batten and Lidwin, 1995), and depth to bedrock below Lake Superior (Wold, 1979), were merged and then interpolated by using the Natural Neighbor algorithm in ArcMap (Esri, 2014), to create a continuous surface of depth to bedrock.

2. The depth to bedrock surface was subtracted from the land surface to create a preliminary bedrock elevation surface imprinted with details from the land surface.

3. The preliminary bedrock elevation surface was sampled at the locations of all WCRs, and compared to WCR bottom depths. Some WCRs that were completed in the Quaternary deposits were intersected by the preliminary bedrock surface, indicating overestimates in the bedrock surface elevation at those locations. The depth to bedrock dataset was updated at these locations with elevations set to 1 ft below the bottom elevations of the WCRs, and new depth to bedrock and bedrock elevation surfaces were created by repeating steps 1 and 2.

4. Land-surface detail resulting from subtraction of the depth to bedrock surface from the land surface was then removed from the areas not mapped as shallow bedrock (e.g., Quaternary deposits thin to absent in fig. 3). Every 13th model row and column was sampled in the area not mapped as shallow bedrock, corresponding to a resolution of approximately 0.6 miles. In the area mapped as shallow bedrock, every grid cell was sampled. The rationale behind this strategy is that the bedrock surface closely mimics land surface in the areas mapped as Quaternary deposits thin to absent, but likely does not follow the land surface as closely in areas covered by thick Quaternary deposits. The resampled bedrock elevation points were combined with the bedrock outcrop points, point estimates of depth from the seismic studies, and point estimates of depth to bedrock below Lake Superior, and interpolated by using the Natural Neighbor method to create a final bedrock surface (fig. 2.2).

5. Discrepancies between the land surface and bedrock surface were reconciled during creation of the MODFLOW grid, by setting depth to bedrock at 2 ft below land surface, in cells where it was higher than 2 ft below land surface (effectively enforcing a minimum thickness of 1 ft in layers 1 and 2, which represent equal portions of the unconsolidated deposits, fig. 2.1).

6. The bottoms of layers 3 and 4 were created by subtracting uniform elevations of 50 and 350 ft directly from the bedrock surface elevations. The bottom of layer 5 (model bottom) was further smoothed, by sampling the bedrock surface every 52 row and columns, subtracting 1,000 ft, and then reinterpolating with the Natural Neighbor method.

Boundary Conditions

Recharge Model (Soil Water Balance)

Recharge to the groundwater system was estimated for each model grid cell by using the Soil Water Balance code(SWB)(Westenbroek and others, 2010). SWB estimates recharge by calculating a Thornthwaite-Mather water balance for each grid cell, at daily time steps:

$$R = (P + M + Q_{in}) - (Int + Q_{out} + ET) - \Delta S \qquad (2.1)$$

where

R is recharge;

P is daily precipitation;

M is water made available on days when temperatures are high enough to melt accumulated snowpack;

Q_{in} is daily surface runoff entering the cell from neighboring cells;

Int is the amount of daily rainfall intercepted/trapped by vegetation as a function of land-use type and season;

Q_{out} is daily surface runoff from a cell according to a curve number rainfall-runoff relation (U.S. Department of Agriculture-Natural Resources Conservation Service, 1986) related to soil type, land use, surface condition, and antecedent runoff condition;

ET is daily evapotranspiration from the root zone of the soil as a function of temperature and vegetation; and

ΔS is the change in the amount of water stored in the root zone calculated according to the method of Thornthwaite (Thornthwaite and Mather, 1955).

Inputs to SWB include daily temperature and precipitation, and maps for land surface flow direction, land cover, available soil water capacity, and soil hydrologic group. Gridded datasets at the 1 kilometer (km) resolution from the Daymet model (Thornton and others, 2014) were used for precipitation and temperature. The Daymet model provides estimates of daily weather by using a DEM and observations of maximum temperature, minimum temperature, and precipitation from ground-based meteorological stations. This approach is well suited for the Bad River Watershed, as it can account for differences in precipitation driven by topography, and localized "lake effect" snowfall (figs. 2.3 and 2.4). Input datasets and their sources for the application of SWB to the Bad River model are shown in table 2.1, and figures 2.3 through 2.8.

The SWB simulation was run for the period of 1980–2011. Recharge values for this period were then averaged to develop a single array of estimated steady-state recharge corresponding one-to-one to the model cells (fig. 12). The steady-state recharge array was then allowed to vary uniformly by adjusting a single array multiplier during calibration, to allow the model to best match measured base flows in streams (which represent the bulk of discharge from the groundwater system). Individual inputs to SWB were not included in the calibration process, due to the known non-uniqueness given the field data available, and the long runtimes of the recharge model (approximately 1.5 days for the period of 1980–2011). The calibrated MODFLOW model resulted in a recharge multiplier of 1.24, or 1.24 times the original SWB calculated recharge rate for each node.

Figure 2.9 shows the variability of annual mean recharge estimates for the model area for the period of 1980–2011, prior to calibration of the groundwater model. Figure 12 shows the steady-state, post-calibration recharge for each model cell.
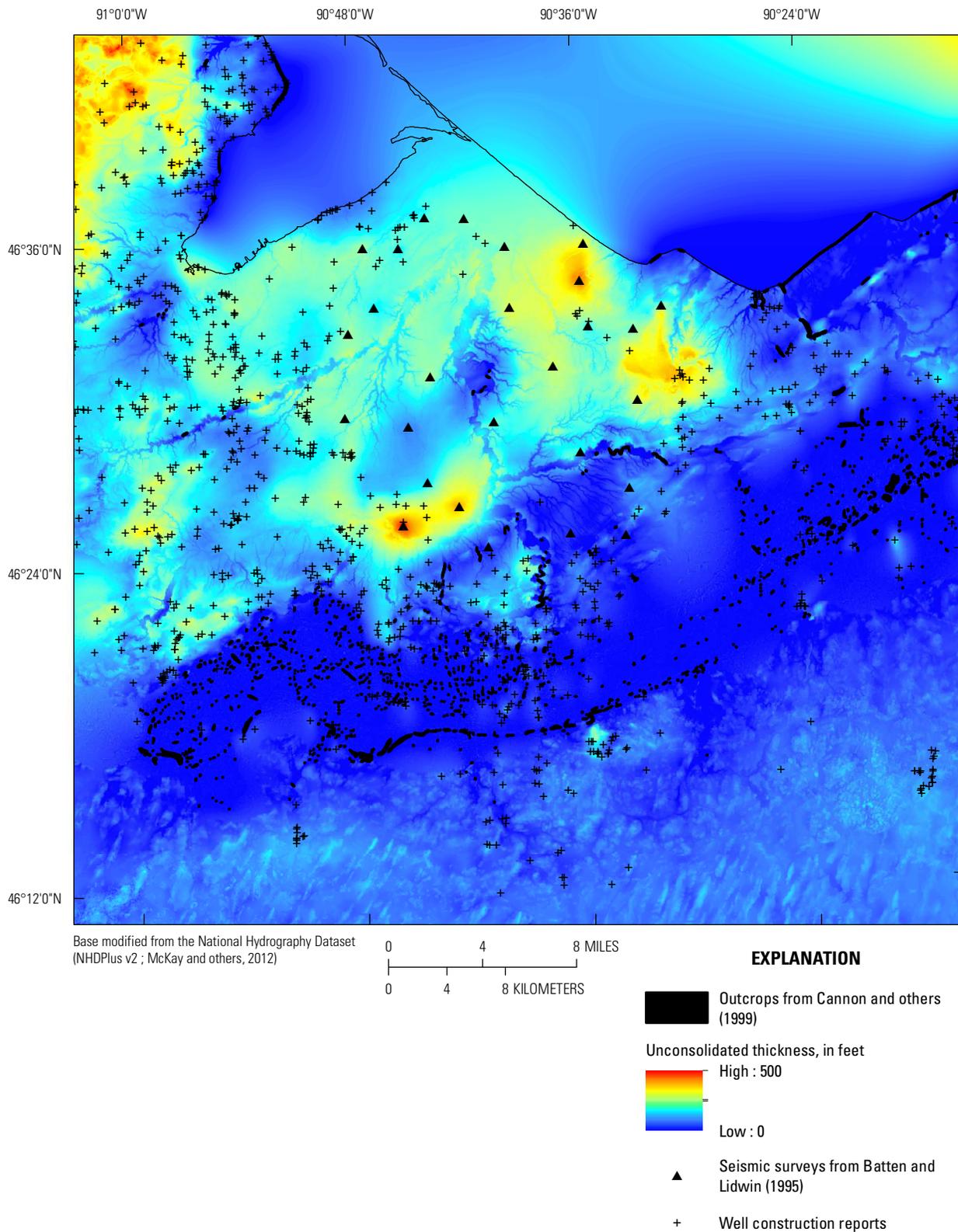
Figure 2.1.   Thickness of Quaternary deposits represented by model layers 1 and 2.
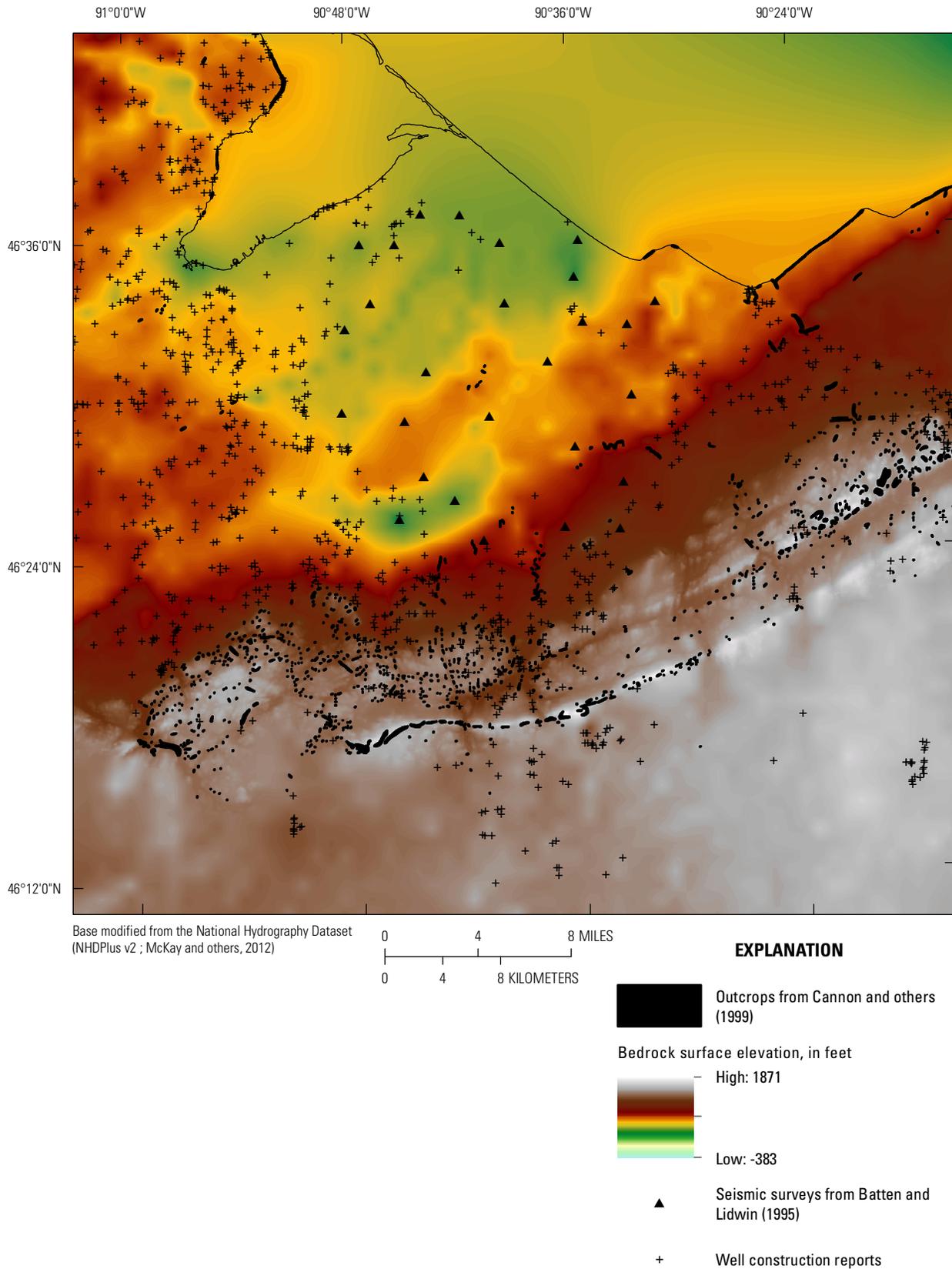
Figure 2.2. Bedrock-elevation surface represented by the bottom of model layer 2, which serves as the top of layer 3.
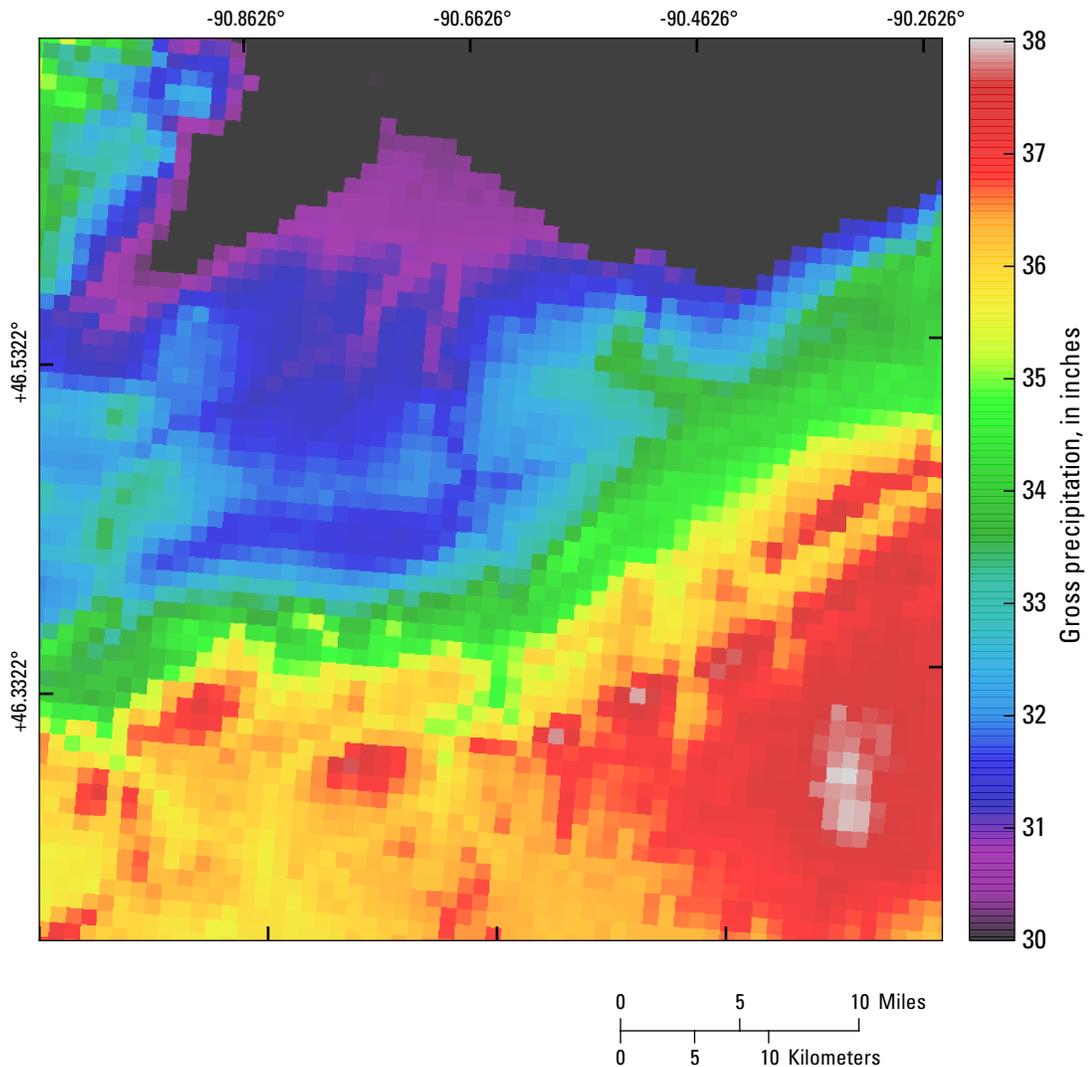
Figure 2.3.   Average annual gross precipitation (defined in Soil Water Balance as net precipitation plus interception) for 1980–2011, as estimated by the Daymet model (Thornton and others, 2014). The Daymet model estimates daily weather by using a digital elevation model and observations of maximum temperature, minimum temperature, and precipitation from ground-based meteorological stations.
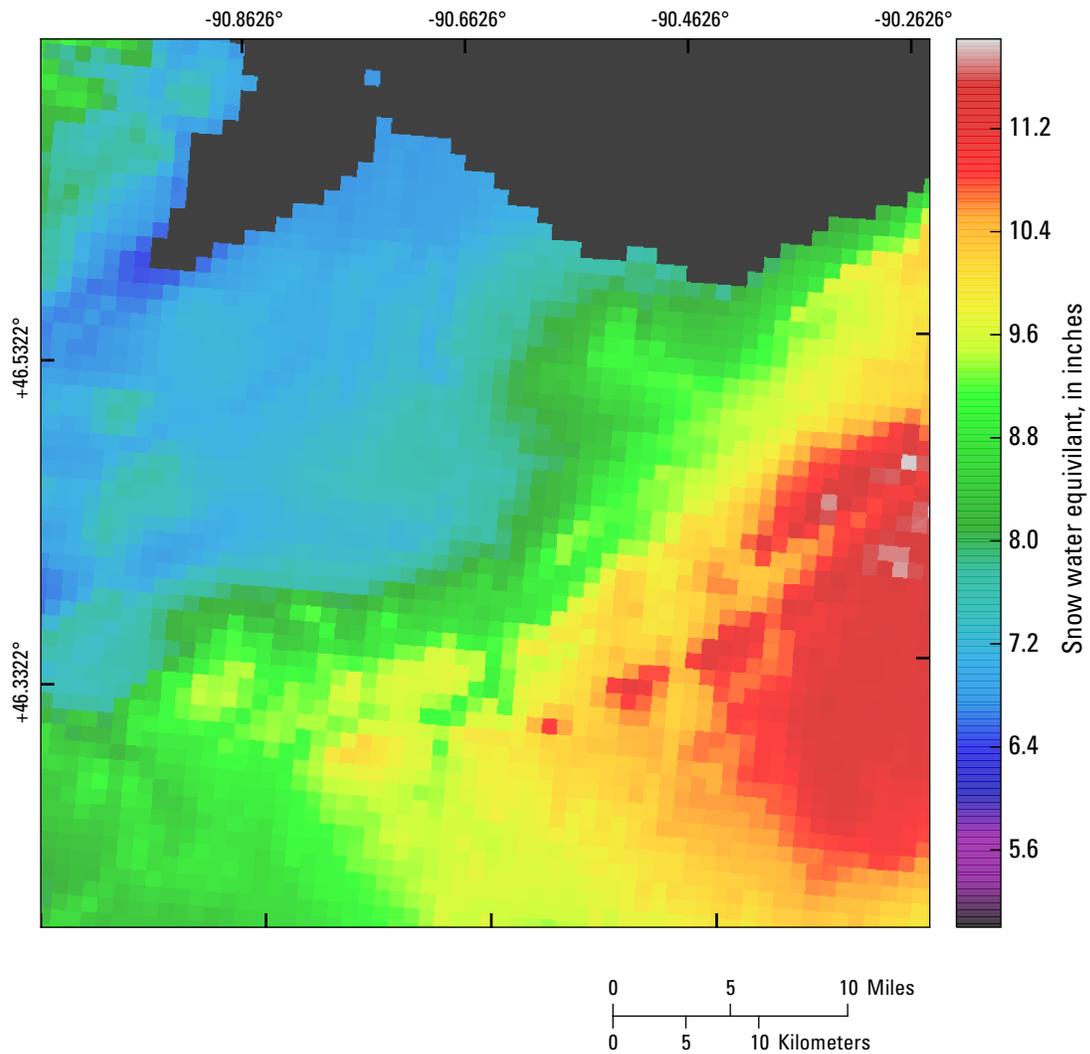
Figure 2.4.   Average annual snowfall (as snow water equivalent), estimated by the Daymet model (Thornton and others, 2014). Higher values in the southeast corner of the model domain show the strong influence of lake-effect snow.

EXPLANATION

▪ East		▪ West	
▪ Southeast		▪ Northwest	
▪ South		▪ North	
▪ Southwest		▪ Northeast	
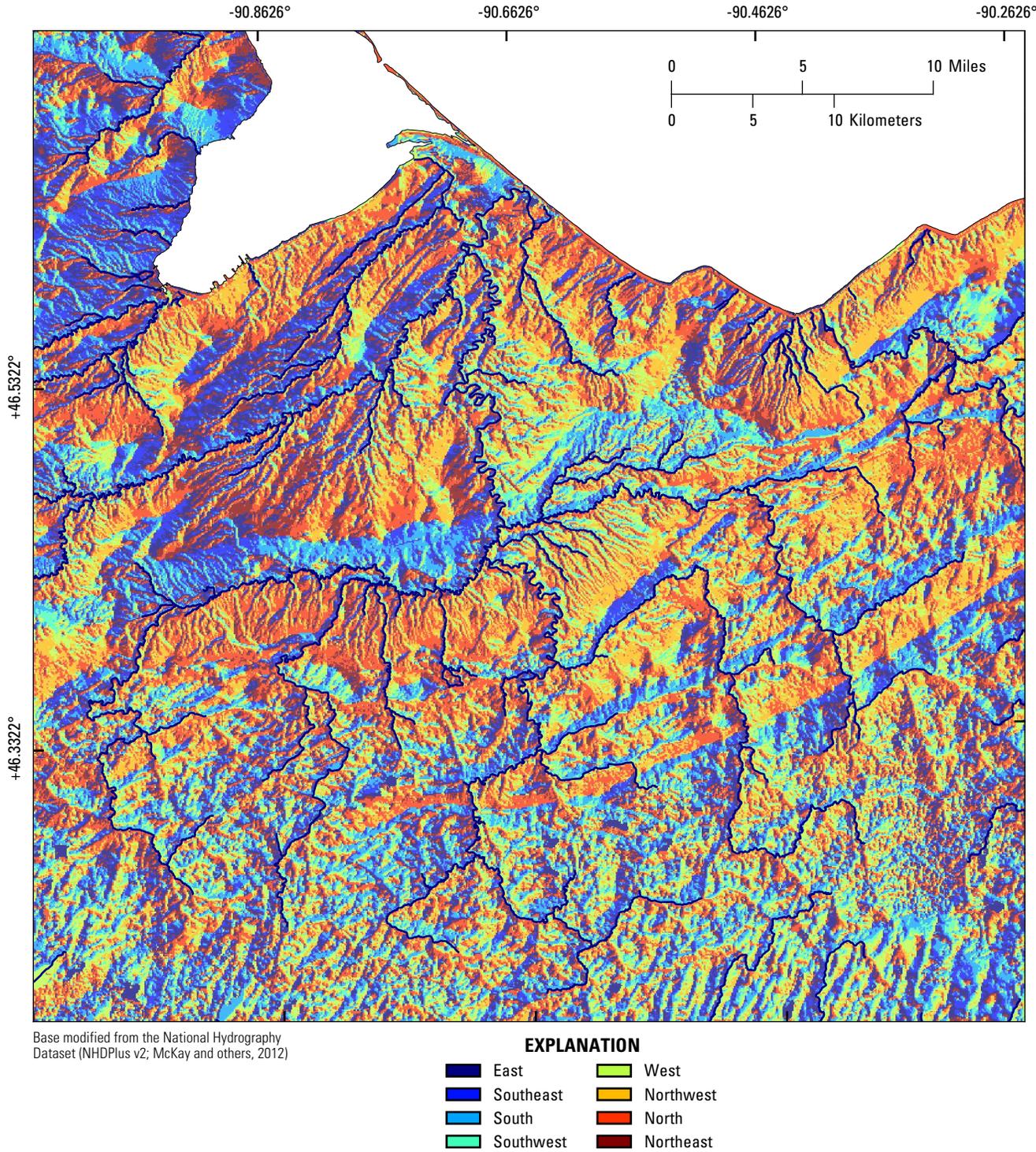
Figure 2.5.   Land surface flow direction, used to route surface runoff between cells in Soil Water Balance (U.S. Geological Survey, 2014a).
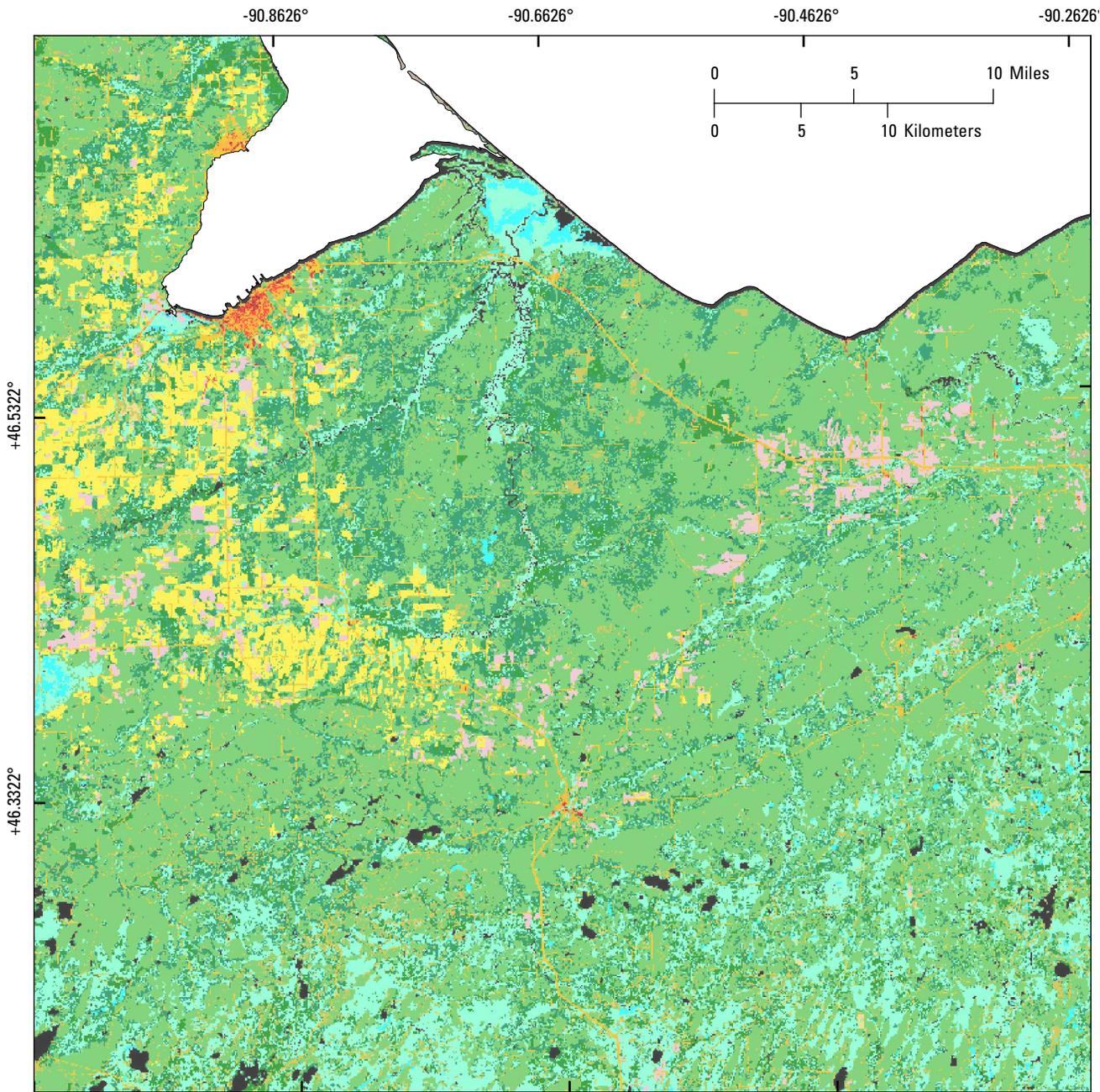
Base modified from the National Hydrography
Dataset (NHDPlus v2; McKay and others, 2012)

EXPLANATION

OPEN WATER

Developed, Open Space

Developed, Low Intensity

Developed, Medium Intensity

Developed, High Intensity

Barren Land

Deciduous Forest

Evergreen Forest

Mixed Forest

Dwarf Scrub

Shrub/Scrub

Grassland/Herbaceous (arid)

Pasture/Hay (good)

Cultivated Crops (SR+CR poor)

Woody Wetlands

Emergent Herbaceous Wetlands

Figure 2.6.   Land cover (modified from Fry and others, 2011).
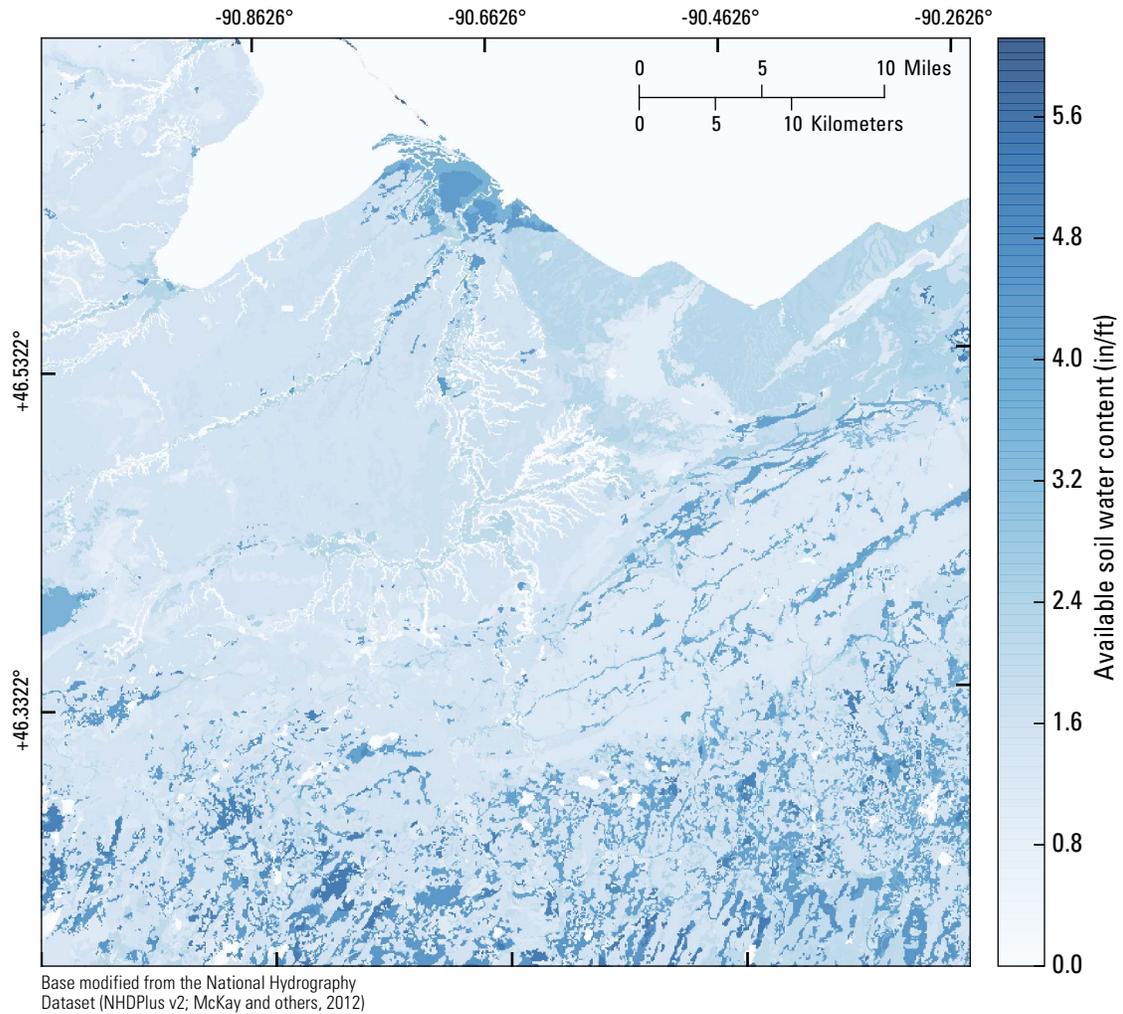
Figure 2.7.   Available soil water content modified from U.S. Department of Agriculture-Natural Resources Conservation Service (2012).
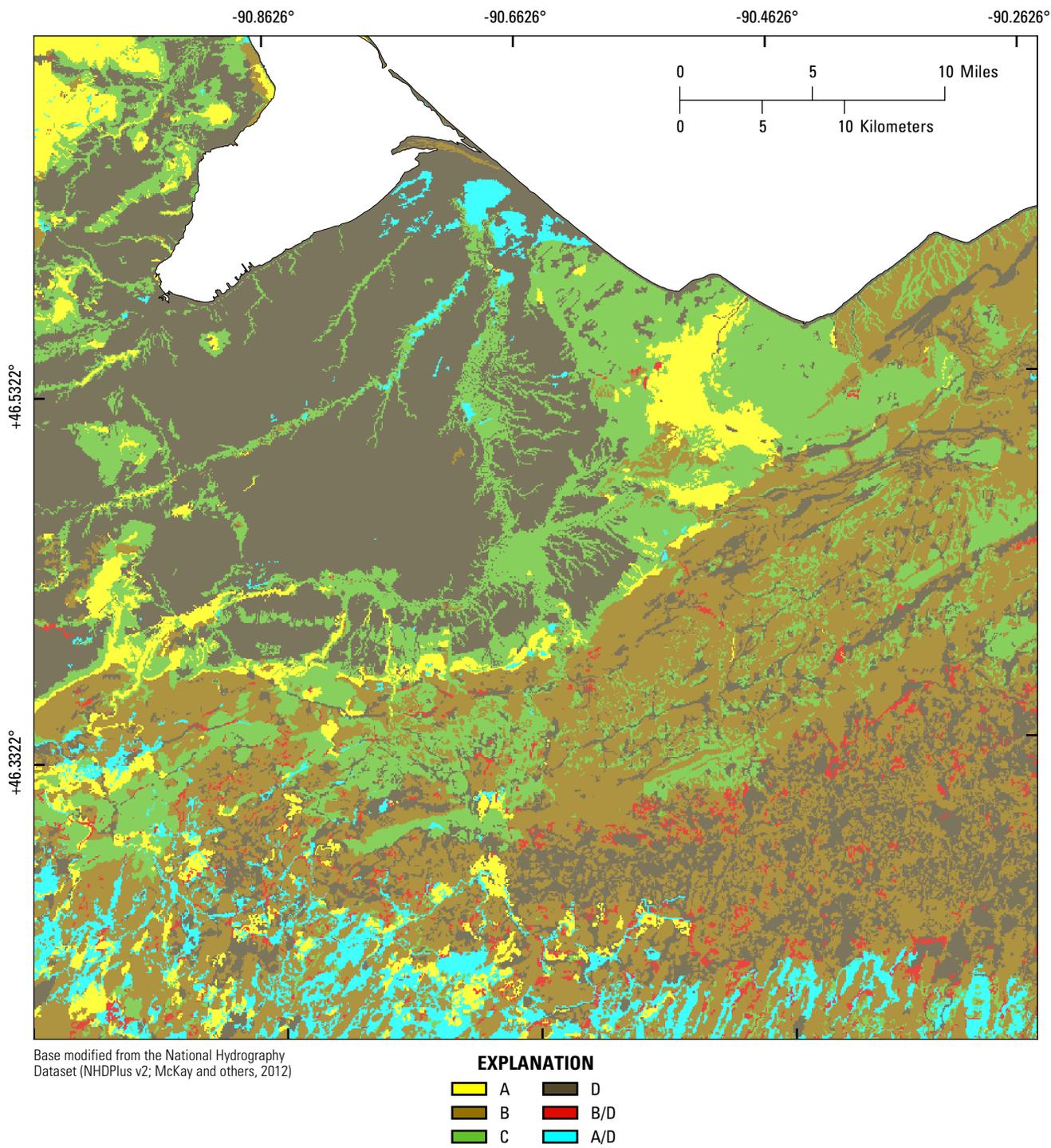
Figure 2.8.   Soil hydrologic groups (modified from U.S. Department of Agriculture-Natural Resources Conservation Service, 2012).

Table 2.1. Data sources for Soil Water Balance simulation.

[USDA, U.S. Department of Agriculture; NRCS, Natural Resources Conservation Service; USGS, U.S. Geological Survey]

Input	Source	Reference
Daily precipitation and temperature	DayMet dataset (Oak Ridge National Labs)	Thornton and others, 2014
Soil hydrologic group and available water content	Soil Survey Geographic Database (USDA-NRCS)	NRCS, 2012
Land cover	2001 National Land Cover Database	Homer and others, 2007
Surface-flow direction	National Elevation Dataset	USGS, 2014a

Table 2.2. Summary annual statistics for the original, smoothed and post-calibration recharge arrays. σ represents standard deviation.

[Max, maximum; in., inches; Min, minimum, ft^3, cubic feet]

Array	Max (in.)[1]	Min	Median	Mean	Standard deviation	Total volume (ft^3)
unsmoothed	90.5	0.0	5.0	6.0	3.6	44,948,532
smoothed	30.9	0.0	4.7	5.4	3.1	40,284,006
post-calibration	38.5	0.0	5.8	6.7	3.8	50,121,804

[1]All units in inches except total volume.

Smoothing of the Recharge Array

The routing process in SWB resulted in high rates of recharge coinciding with closed depressions in the flow direction map. It is unclear if these areas—many of which area only a single grid cell in size—are realistic or simply artifacts of the SWB input dataset. To address this issue, the recharge map was smoothed by requiring that recharge in each cell could be no greater than the average of its neighbors:

$$R_{j,i_{smoothed}} <= \frac{\sum_{i=i-1}^{i+1} \sum_{j=i-1}^{j+1} R - R_{i,j}}{8} \tag{2.2}$$

This approach assumes that the additional estimated recharge that would be simulated in these areas does not infiltrate (that is, it runs off to surface water, or is stored on the land surface and evaporated). The effects of smoothing the recharge field are summarized in table 2.2. Implementation of this method initially reduced the total SWB-calculated volume of recharge by 10 percent. However, this reduction was offset during calibration, by the use of the global multiplier parameter for the recharge array. Another option for correction of these recharge hot spots would have been experimentation with the maximum infiltration rate inputs into SWB. Considering runtimes of 1.5 days to complete a SWB simulation for 1980–2011, however, the smoothing approach was considerably faster, and produced a similar result.
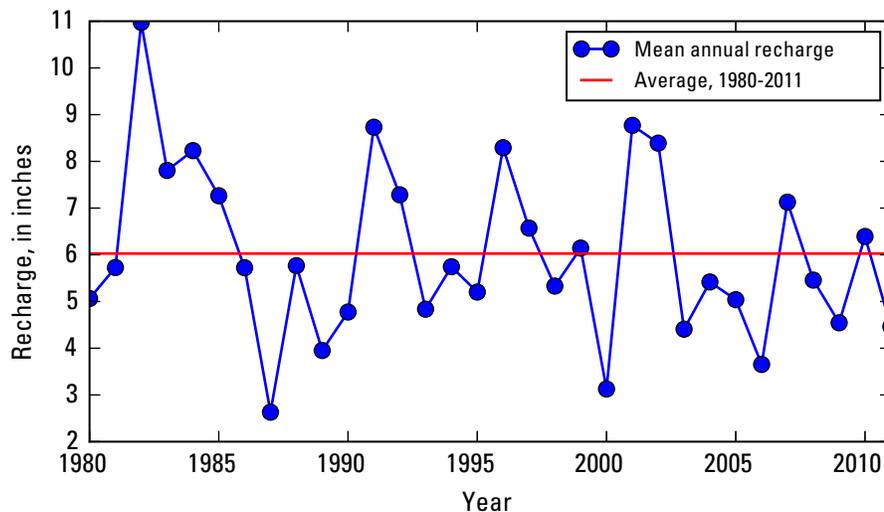
Figure 2.9.   Annual average recharge for the model domain, prior to smoothing and adjustment by the final calibration multiplier of 1.24.

Lake Superior

Lake Superior is represented by a head-dependent flux boundary condition, by using the General Head Boundary (GHB) package in MODFLOW (McDonald and Harbaugh, 1988). Discharge to the lake occurs as a function of the specified lake stage relative to the head in the aquifer, multiplied by a conductance term C_D, which is specified by the user:

$$Discharge = C_D(h_b - h_a)$$ (2.3)

$$C_D = \frac{K_v A_{cell}}{b}$$ (2.4)

where

h_b is the head specified at the boundary;

h_a is the head in the aquifer;

K_v is the vertical hydraulic conductivity of the lakebed;

A_{cell} is the area of the model cell; and

b is the assumed lakebed thickness.

For the Bad River model, Lake Superior stage was set at 601.7 ft above sea level, which represents the average value for the period of 1918-2012, and is also very close to the average value for the period of 1970-2000 (NOAA, 2013), which is representative of most of the observation data. An initial conductance value of 5 square feet per day (ft^2/d) was specified prior to calibration, for all GHB cells representing Lake Superior. Rearranging the formula above, and assuming a thickness of 5 ft, this corresponds to a vertical hydraulic conductivity of approximately 4×10^{-4} feet per day (ft/d). During the calibration process, global conductance for GHB cells was increased to 42 ft/d (corresponding to a vertical lakebed hydraulic conductivity of 3×10^{-3} ft/d). This resulted in approximately 6 percent of groundwater from the MODFLOW model area discharging to Lake Superior.

Streams and Lakes

Streams and Lakes are represented as head-dependent flux boundary conditions, by using the Streamflow Routing (SFR2) package in MODFLOW (Niswonger and Prudic, 2005; Prudic and others, 2004). Groundwater interactions with streams are represented with the same general formulation as shown in equations 2.3 and 2.4, except that the area term A represents the area of the stream channel in a cell (width x length), and a water balance, calculated independently of the groundwater budget, is used to constrain the amount of water available to leak into the aquifer in each SFR cell, and to route excess water to downstream SFR cells:

$$\sum Q_{in} = Q_{inspecified} + Q_{trb} + Q_{ro} + Q_{ppt} - Q_{gwin} \tag{2.5}$$

$$\sum Q_{out} = Q_{sro} + Q_{div} + Q_{et} + Q_{gwout} \tag{2.6}$$

where $\sum Q_{in} = \sum Q_{out}$, and

$Q_{inspecified}$ is an inflow rate specified to a headwater stream;

Q_{trb} is inflow from upstream tributaries;

Q_{ro} is overland runoff entering the cell;

Q_{ppt} is precipitation falling directly on the stream reach;

Q_{gwin} is groundwater discharge to the stream;

Q_{sro} is outflow routed downstream;

Q_{div} is diverted outflows;

Q_{et} is evapotranspiration from a stream reach;

Q_{gwout} is discharge from the stream to the aquifer.

For example, if there are no inflows to a stream reach, and the streambed is above the head in the aquifer, the stream does not interact with the aquifer. Similarly, if the computed leakage to the aquifer exceeds inflows, the leakage will be limited to the amount of inflow, and the stream will be simulated as dry. Discharge from the aquifer to the stream occurs whenever the heads in the aquifer are greater than the stream stage. The stream stage (h_b) can be a specified height above the streambed elevation, or can be estimated by using a reformulation of Manning's Equation:

$$y = \left[\frac{Qn}{CwS_0^{1/2}} \right]^{3/5} \tag{2.7}$$

where

y is the stream depth (or height of water column above the streambed);

Q is the stream discharge;

n is Manning's roughness coefficient;

C is a constant for unit conversion, which is 1.0 for units of cubic meters/second, or 1.486 for units of ft^3/s;

w is the width of the stream channel;

S_0 is the slope of the stream channel.

Required input for the SFR2 package is divided into two levels—segments and reaches. Reaches are sections of streams that simulate streamflow and groundwater/surface-water interactions at the scale of a single model cell. Segments are composed of consecutive reaches in a single stream, which share some uniform or linearly varying properties. For each segment, inputs include inflows (if any) and routing connections with downstream segments, along with Manning's roughness coefficient, and stream width. For each reach, inputs include cell information (row, column and layer), reach number (used to route consecutive reaches within a segment), streambed elevation, streambed slope, streambed thickness, and streambed vertical hydraulic conductivity.

Input for the SFR package in the Bad River model was developed from geographic information system (GIS) information in the National Hydrography Dataset (NHDPlus v2; McKay and others, 2012) and the Wisconsin Department of Natural Resources (WDNR) hydrography (WDNR, 2014b), by using a series of scripts written in the Python programming language. Only NHDPlus flowlines were used, meaning that lakes and wetlands were represented as linear features within the stream network. The NHDPlus v2 database includes 101 lakes and ponds, and 96 wetlands within the model domain. Explicit simulation of these (through the use of boundary conditions such those employed in the Lake and Evapotranspiration packages for MODFLOW) is beyond the scope of this regional modeling study. Explicit representation of these features may refine the flow solution in their immediate vicinity (Haitjema, 1995), though likely with little effect on the regional system as they are approximated by the both the SFR and UZF packages (see below).

Within the model domain, 56 percent of stream segments listed in the NHDPlus v2 database are classified as first order. Of the first-order streams, 37 percent are classified as perennial. The portion of streams classified as perennial varies greatly across the model domain, with only 20 percent of first-order streams in the Superior Lowlands classified as perennial, and 62 percent of first-order streams classified as perennial in the transition area and Southern bedrock uplands.

The steps taken to develop SFR input are summarized below.

1. The NHDPlus hydrography linework were subdivided at the model grid cell boundaries, and then spatially joined to the model grid, along with their corresponding database tables, to create preliminary reaches.

2. The original identifiers for NHD hydrography linework (Common identifiers, or COMIDs) were used to group the preliminary reaches, and associate the linework with database tables.

3. COMIDs and reaches were routed by using information in the NHDPlus database, and geographic proximity to avoid unrealistic routing among segments that are separated by large distance.

4. Minimum and maximum streambed elevations were calculated for each preliminary reach, by extracting elevation values from the 10 m DEM at the locations of vertices in the subdivided linework. Although the NHDPlus contains elevation information, the resolution (1:100,000) is too coarse for the Bad River application (only a minimum and maximum elevation are available for each COMID). Use of finer-scale elevation data allows for improved representation of the stream network geometry at the resolution of the model cells, allowing for more accurate simulation of groundwater/surface-water interactions.

5. Streambed elevations within each COMID were forced to decrease monotonically downstream (so that the streambed is never simulated as running uphill). This is necessary because of mismatch between the NHDPlus linework and DEM (where the linework meanders periodically out of the stream channel, causing a local maximum in streambed elevation), and also because of inconsistencies in the DEM streambed elevations due to the spatial averaging of DEM cells. For example, in a wide section of stream valley, the stream elevation may be well represented. If the valley narrows subsequently downstream, however, the DEM elevation value may include the effects of steep valley sides, causing a local maximum in the streambed estimate. This method therefore assumes that subsequent minima in the DEM encountered by the NHDPlus linework are closer to the true elevation of the stream, and so should be preferentially honored in the streambed estimates. The minimum and maximum elevation listed in NHDPlus were maintained at the COMID endpoints.

6. Stream widths for each COMID were estimated based on arbolate sum (total length of upstream drainage), by using the regression relation of Feinstein and others (2010) in a similar temperate environment.

7. COMIDs were converted to SFR package segments by combining consecutive COMIDs between confluences, and updating the reach numbering and routing.

8. The Wisconsin DNR hydrography contains many smaller streams not included in the NHDPlus v2 database. These were added to the SFR network by removing overlaps in the two hydrography datasets, and then incorporating the

WDNR linework, by using a similar procedure to the one above, except preexisting routing tables were not used. Segments and routing were established by grouping consecutive reaches based on geographic proximity and directional information in the linework shapefile, and then terminating the segments at confluences, or where no other reaches were found within a specified distance cutoff. The segments derived from the WDNR dataset were then routed into the existing SFR network by locating the closest 2nd order or higher SFR cell (within a specified distance). Because the additional WDNR hydrography represents small streams, a uniform width of 5 ft was assigned to all WDNR-derived SFR reaches.

UZF Package

The SWB recharge model does not account for the rejection of recharge in areas where the water table is at or above the land surface. This resulted in localized unrealistic simulated rises in the water table to above land surface in some topographic low-lying areas of the model (fig. 2.11). These areas of "flooded" cells can be problematic for accurate delineation of flow paths, as they can produce mounding of the water table (implying recharge to the groundwater system) in low areas, such as river valleys, that in reality are in the discharge portion of the flow system. Flooded cells are also potentially problematic for accurate simulation of drawdowns under future stresses (such as large scale dewatering), as they provide extra head to offset any declines. In the field, many of these areas of flooded cells represent either lakes or wetlands, where groundwater seepage discharges to the stream network, or they also may be areas that are drained by unmapped streams.

The UZF package (Niswonger and others, 2006) allows for the simulation of rejected recharge in steady-state models, by assigning a head-dependent flux boundary condition to the land surface, which routes flow as recharge to the groundwater system when the water table is below land surface, or as overland runoff to surface water, when the water table is above land surface:

$$Q = \frac{A_{cell}K_v(watertable - modeltop)}{0.5 * cellthickness} \tag{2.8}$$

where

Q is either recharge to the groundwater system, or runoff;

A_{cell} is the area of the model cell;

K_v is the vertical hydraulic conductivity of the unsaturated zone (in this application, K_v for Layer 1 was used).

Computed discharge values are decreased when groundwater levels are within a specified distance of the land surface, denoted by the SURFDEP variable in the UZF package input (see U.S. Geological Survey, 2014b). This smooths the transition to a condition of groundwater discharge when groundwater levels approach the land surface, improving numerical stability. It also accounts for the existence of small undulations in the actual land surface, which in reality would tend to dampen the onset of discharge (groundwater flooding would be expected to start in the lowest areas, not instantaneously over the entire area of a model cell). A SURFDEP value of 3 ft was used in the Bad River Model.

In the Bad River application, overland flow resulting from flooded cells was routed to nearby SFR segments. Catchment areas from NHDPlus v2 (McKay and others, 2012; shown in figure 2.10), were associated spatially with the Layer 1 cells and the SFR network, and the segment with the highest number of SFR cells in each catchment was chosen as the drainage point for the Layer 1 cells within the catchment. This resulted in approximately 20 percent of groundwater recharge being converted to overland flow contributions to streamflow. Spatial distribution of seepage is shown in figure 2.11.

Subsurface Properties

Representation of Major Hydrogeologic Units

Within the layers discussed in Section 2, major hydrogeologic units were represented by hydraulic conductivity zones. In layers 1 and 2, zones were included for outwash sand and gravel, the Miller Creek Formation, the area of shallow bedrock, and sandy tills of the Copper Falls Formation, based on the extents of these units mapped by Mickelson and others (1983) (figs. 3 and 2.12). In layers 3 through 5, zones were included to represent the sandstones and
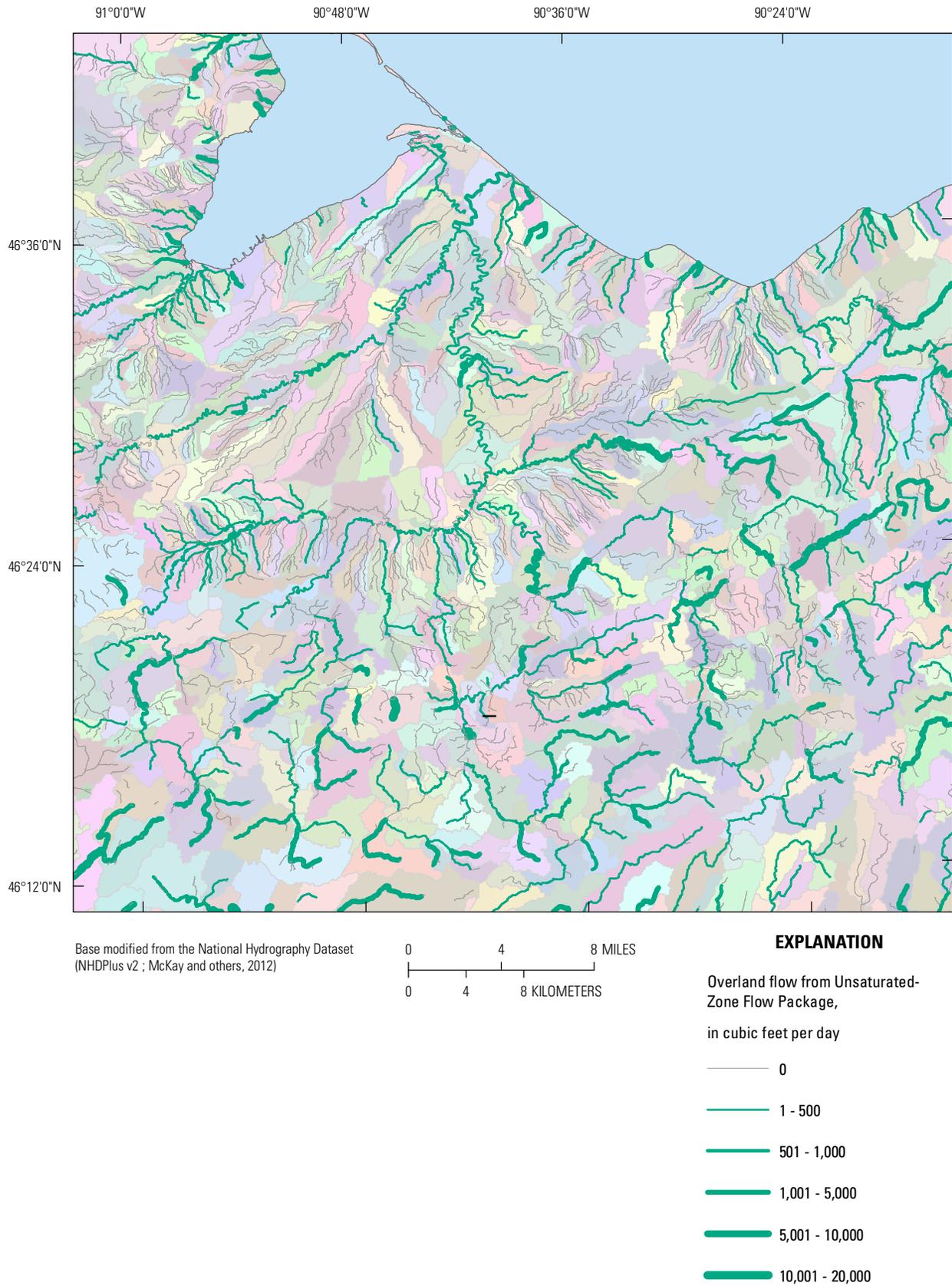
Figure 2.10.   Overland flow routing to streams simulated by the Unsaturated Zone Flow (UZF) package, with contributing catchments. Line thicknesses represent inflows from the overland component in each reach.
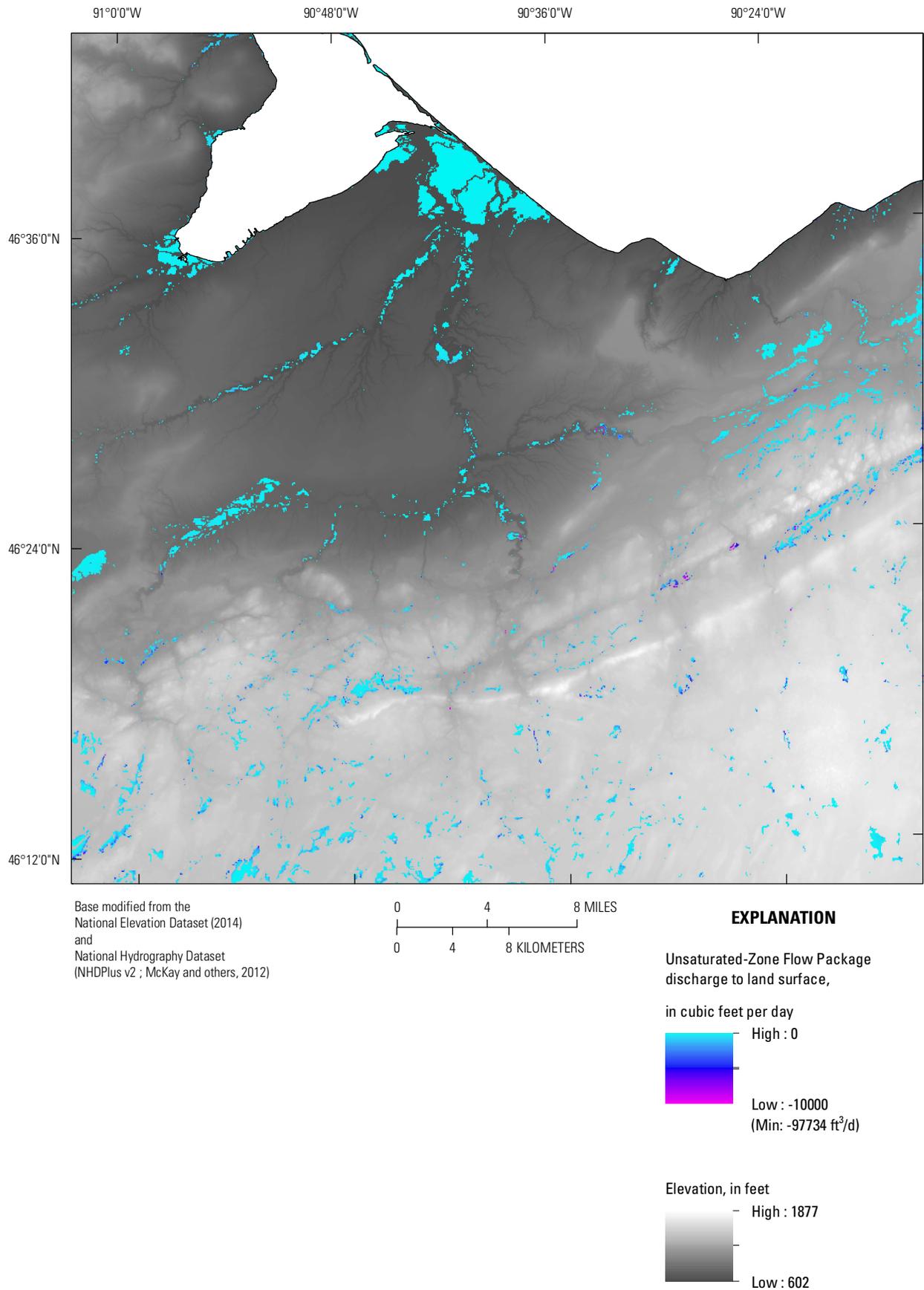
Figure 2.11.   Discharge to the land surface simulated by the UZF package; negative discharge equals flow out of the groundwater system to the surface-water system.
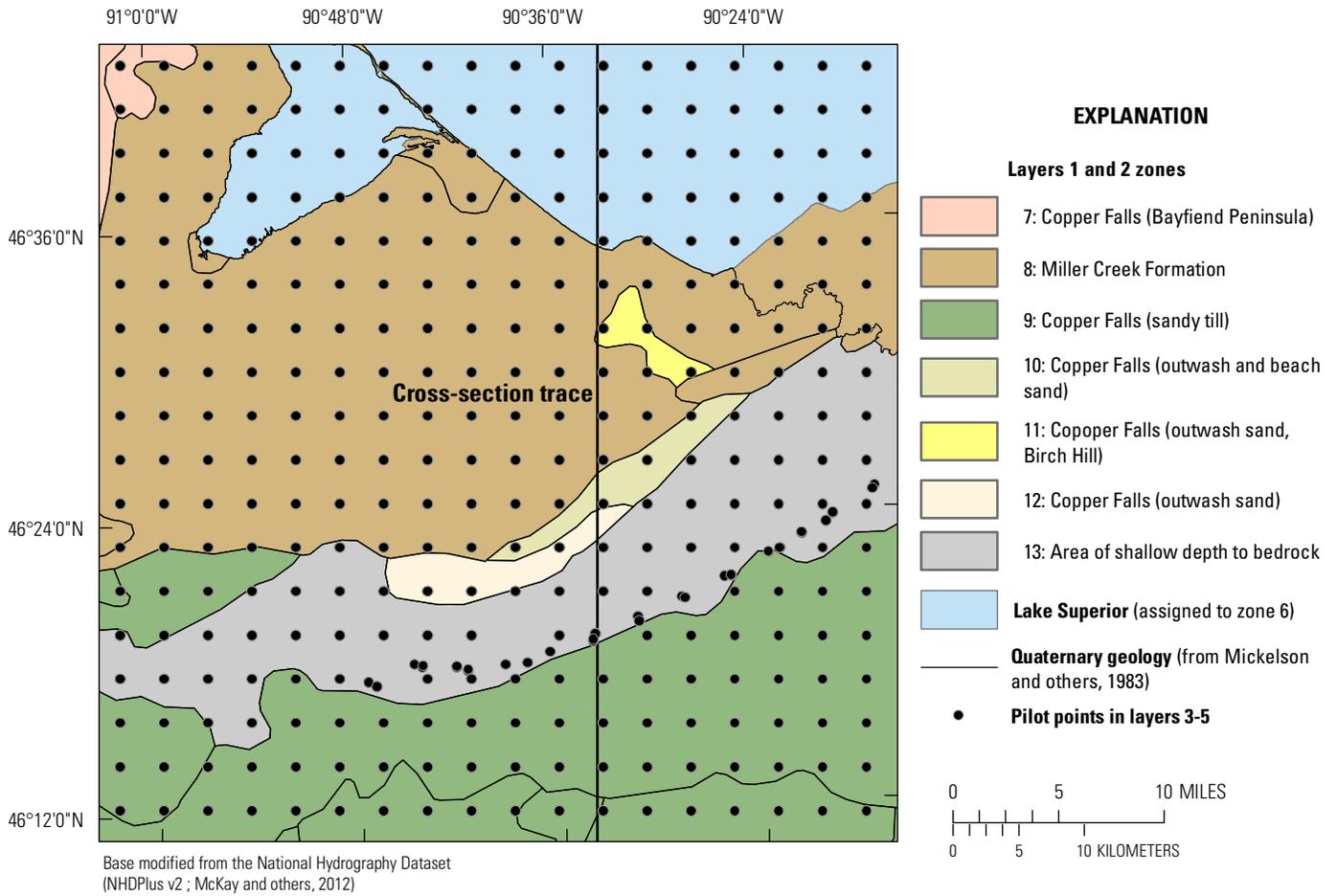
Figure 2.12.   Hydraulic conductivity pilot point network and interpolation zones for layers 1 and 2 (representing Quaternary deposits).

conglomerates of the Oronto and Bayfield Groups, the igneous units of the Keweenawan Supergroup and Mellen Complex, the Tyler, Ironwood and Palms formations, and the Archean units of the southern bedrock uplands, as mapped by Cannon and others (1999) (figs. 2 and 2.13). Figure 2.14 shows the hydraulic conductivity zonation in cross section along column 500 of the model.

Table 2.3.   Average dip values reported for each zone in the model.

Zone	Mean dip (degrees)	Number of Measurements
1	55	112
2	69	44
3	51	14
4	56	30
6	56	37
2,4,*	64	74

* Special case combining zones 2 and 4 to use in zone 5

The bedrock strata in the Bad River Watershed are characterized by steep dip toward Lake Superior. The average measurements of dip reported by Cannon and others (1999) were calculated in the footprint of the bedrock zones in the MODFLOW model. Table 2.3 reports the average dip values calculated for each zone. No dip measurements were available for zone 5, so the combined average dip in adjacent zones 2 and 4 were applied to zone 5. In the finite-volume grid cells of the model, dip does not play a role in how the cells are simulated as vertical faces are assumed bounding the model cells. However, dip in the strata can result in on offset of zones with depth. To accommodate the dip in this model, the lateral (assumed northward for simplicity) offset of zones with depth was calculated as a function of depth and dip. Mean depth was calculated for each zone in each layer as

$$MEAN\_DEPTH_{z\ell} = \frac{\sum_{i=1}^{ncells} \left(TOP_{z\ell,row_i,col_i} - BOTTOM_{z\ell,row_i,col_i} \right)}{ncells} \tag{2.9}$$

where

$z\ell$ indicates the zone and layer in which calculations are being made;

ncells is the number of cells in the zone/layer combination;

TOP is the model top, calculated in the ith cell; and

$BOTTOM$ is the model bottom, calculated in the ith cell.

The offset for each zone was calculated by using the mean depth and mean dip for each zone as

$$offset_{z\ell} = \tan(\theta_z) \times MEAN\_DEPTH_{z\ell} \tag{2.10}$$

where

θ_z is the mean dip in zone z; and

$MEAN\_DEPTH_{z\ell}$ is the mean depth for zone z and layer ℓ calculated by using equation 2.9.

For each zone, if $offset_{z\ell}$ was greater than the cell spacing in the north-south direction (dy) the zone boundary was shifted northward by the number of cells ($shift\_cells$) calculated as

$$shift\_cells_{z\ell} = floor\left(\frac{offset_{z\ell}}{dy}\right) \tag{2.11}$$

Figure 2.14 illustrates this offset.
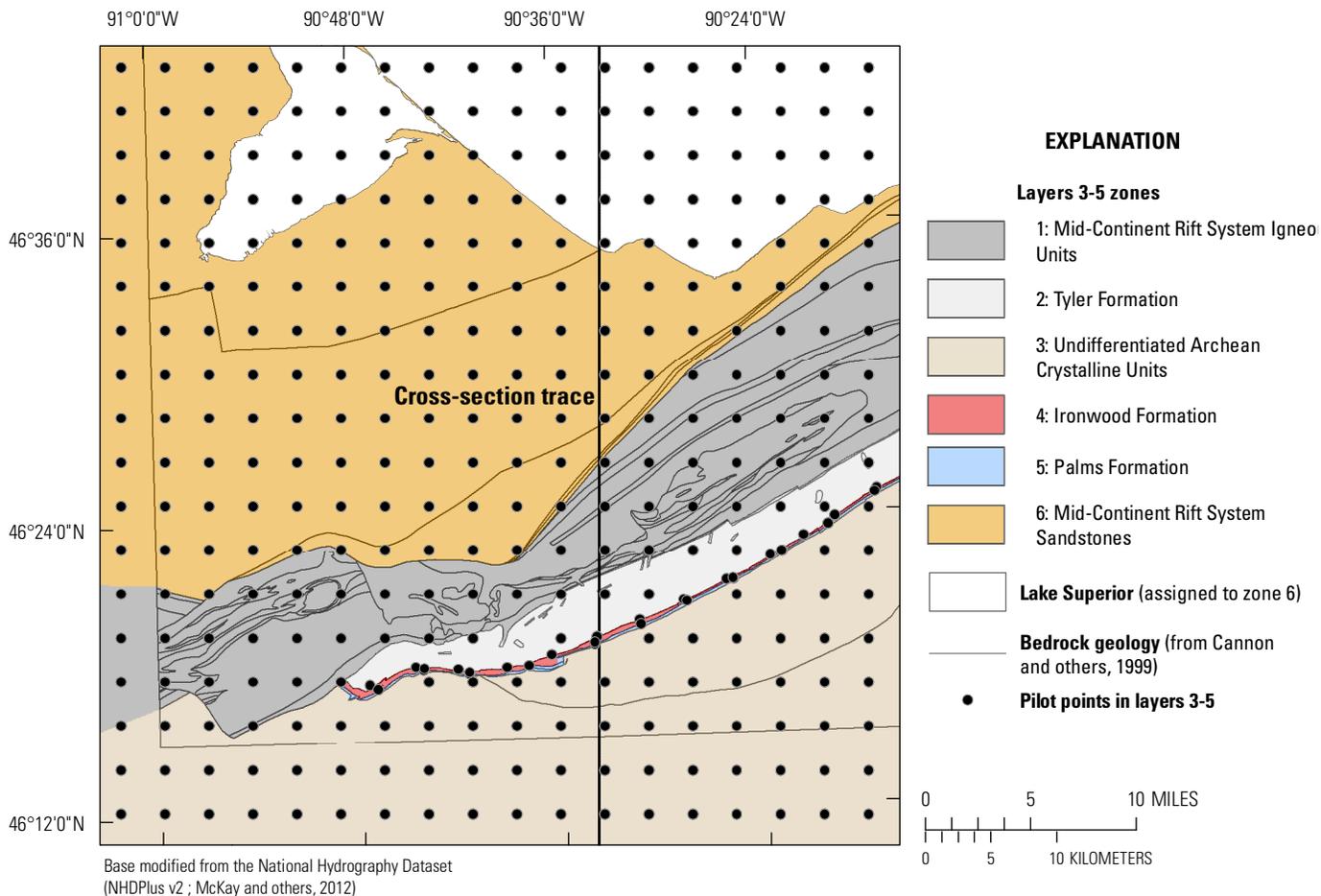
Figure 2.13.   Hydraulic conductivity pilot point network and interpolation zones for layers 3–5 (representing bedrock units).
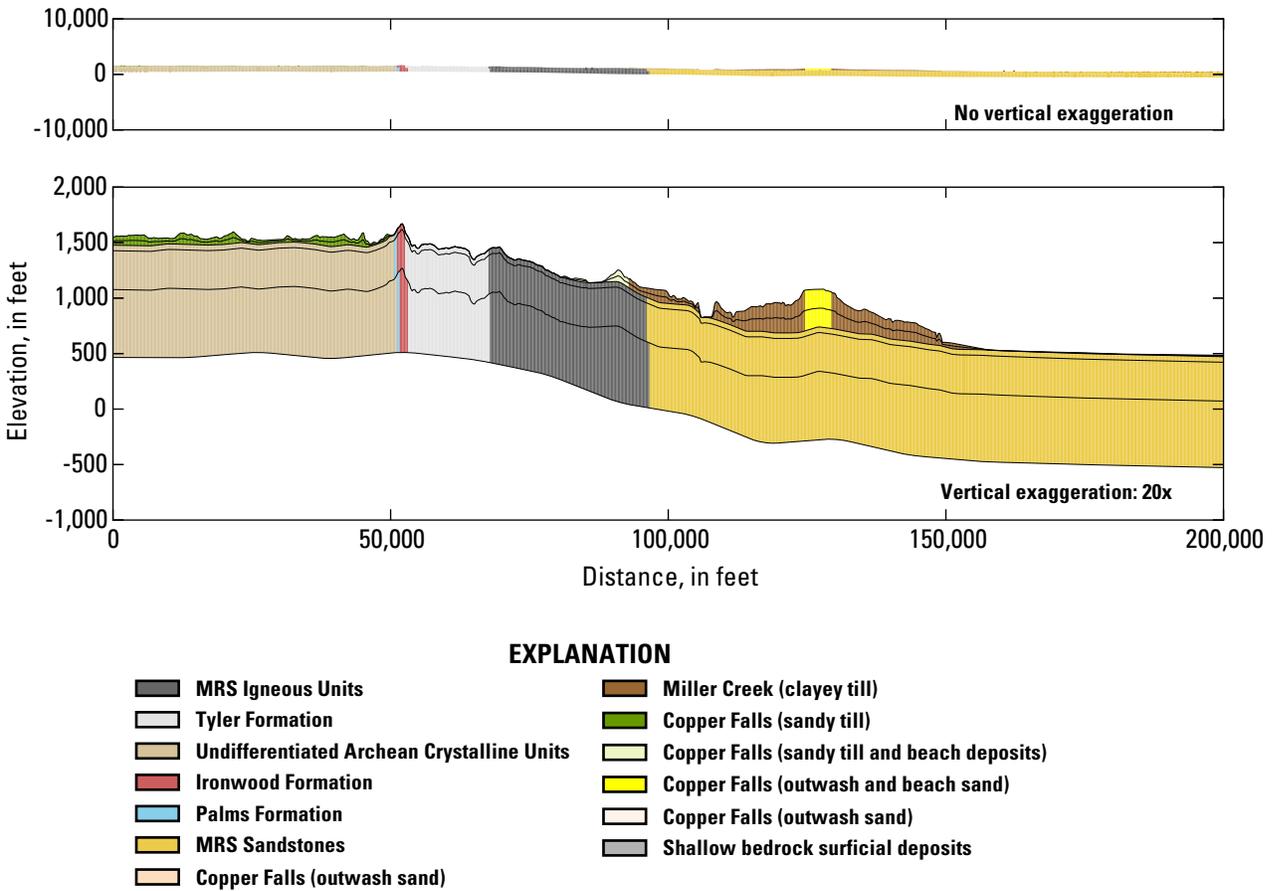
Figure 2.14. Hydraulic conductivity zones along model column 500. The top panel illustrates the actual variation (without exaggeration) in elevation and model thickness, relative to the horizontal extent of the model.

Parameterization of Hydraulic Conductivity

Hydraulic conductivity within each zone was specified by using a network of pilot points (Doherty, 2003). The general goal of pilot-point use is to provide an intermediate alternative to two end-member approaches of parameterization of cell-by-cell variability, whereby each model node can have a different value, and a priori simplification of the hydraulic conductivity field through piecewise-constant homogeneous zones. In the pilot-point approach, parameter values are estimated at a number of discrete locations distributed throughout the model domain; cell-by-cell parameterization then takes place through spatial interpolation from the pilot points to the model grid or mesh. Spatial interpolation between pilot points was applied within the hydraulic conductivity zones described above (that is, no interpolation between neighboring points separated by a zone boundary). This approach expresses the mapped extents of distinct geologic units, while allowing for realistic spatial variability in hydraulic conductivity within the major units.

Horizontal hydraulic conductivity (K_h) was specified at pilot points across the model domain in a regular grid, at 11,000 ft (44 cell) intervals, except in the zones representing the Ironwood and Palms formations, where additional points were added for parameterization flexibility. Pilot points were included in Layers 1, 3, 4 and 5 (figs. 2.12 and 2.13). Hydraulic conductivity in Layer 2 was set equal to Layer 1. In Layer 5, where zone boundaries were shifted to account for dipping strata, the pilot points were tied by setting them equal to their equivalent points in Layer 4. In all layers, pilot points in the area beneath Lake Superior were tied by setting them equal to the closest pilot point on shore (i.e., the closest pilot point informed by head observations). Vertical-anisotropy pilot points were also specified at each horizontal hydraulic conductivity (K_h) pilot-point location. This resulted in a total of 2,704 pilot points in the model domain, 1,712 of which were included as adjustable parameters in the calibration process.

Model Solver

The Bad River model was solved by using MODFLOW–NWT (version 1.08; Niswonger and others, 2011; Hunt and Feinstein, 2012), a stand-alone version of MODFLOW–2005 (Harbaugh, 2005) that incorporates a Newton-Raphson formulation for solution of the groundwater-flow equation. The robust nonlinear-solution capabilities of the Newton-Raphson scheme allow for the continuous treatment of drying and rewetting cells (caused by variations in the simulated water table during solver iterations), as well as improved handling of nonlinear-boundary conditions such as those applied by the SFR and UZF packages. These capabilities make MODFLOW–NWT well-suited for simulations focusing on groundwater/surface-water interactions and also serve to facilitate robust calibration of the model through parameter estimation, a process that requires thousands of model runs under a widely varying set of input parameters (see appendix 3). Although solutions obtained using MODFLOW–NWT are independent of the supplied starting head values, the solution times can be greatly affected by the starting heads chosen by the modeler. To obtain faster forward run times, initial head values were periodically updated to reflect the most recent converged solution; a range of solver settings were also investigated.

Appendix 3—MODFLOW Model Calibration

The calibration process updates initial estimates for selected input values (calibration parameters) in order to improve the match between simulated groundwater levels and baseflows and corresponding field observations; calibration ends when a good fit is obtained and resulting parameters are reasonable. Several steps are performed to enhance the calibration process and facilitate the estimation of independent variables (the parameters) through observed values of dependent variables (head and flow targets). For the automated parameter estimation used here, the steps include selection of parameters to estimate; development of a calibration dataset from field observations; setting of target weights; and, after preliminary adjustment of parameters by manual trial-and-error, application of nonlinear regression performed by using singular value decomposition with regularization.

For this study, the nonlinear regression was implemented with the PEST software suite (Doherty, 2010a), which systematically adjusts model inputs (via the Gauss-Levenberg-Marquardt method) to minimize an objective function, which is the sum of squared, weighted differences between model outputs and their equivalent field observations. In this way, the level of model fit is influenced by the weighting; highly weighted observations contribute more to the objective function, and so will be emphasized in the parameter estimation in its quest to minimize the objective function.

A perfect fit between the model outputs and observations is neither realistic or desired, as all models, regardless of their level of sophistication, are simplifications of natural systems that are ultimately unknowably complex. In addition, the observations themselves are uncertain, due to measurement error and other factors including temporal variability (Moore and Doherty, 2005; Hunt and others, 2007). A result of these uncertainties is that many different model configurations or sets of input parameter values may produce similar levels of fit to the calibration dataset (Doherty and Hunt, 2010; Moore and Doherty, 2006). Regularization is one process for constraining this non-uniqueness, either through ad-hoc, a priori simplification (such as the piecewise-constant zones employed in the screening model), or through mathematical techniques, discussed below. Ultimately, the result is a trade-off between the incorporation of two sources of information into the model: the observation data (hard data which are used to assess model outputs), and "soft-knowledge" of the hydrogeologic system (which inform choices of model inputs) (e.g., Hunt and others, 2007). The balancing of this tradeoff during the calibration process results in a model that reasonably reproduces field observations and also reflects what is known about the hydrogeologic system.

Observations

The following types of observations were used in the calibration:

1. Head measurements from WDNR WCRs

2. Head measurements from the National Water Information System (NWIS) database, for the area represented in the model (http://nwis.waterdata.usgs.gov)

3. Head measurements from WCRs provided by the Bad River Band

4. Average annual base flows computed by base-flow separation of continuous streamflow records from NWIS

5. Published average annual base flows for the period of 1969–1999 (Gebert and others, 2011)

6. Estimated average annual base flows from synoptic measurement events conducted as part of this work during 2011 and 2013

Observation Processing

Whereas the steady-state model simulates single values that represent long-term averages, the observations encompass heterogeneous data spread across several decades, representing variable lengths of time. To address this mismatch, and create calibration targets equivalent to model outputs, representative average values were developed for sites with multiple measurements. Mean annual base flows also were estimated from one-time base-flow measurements, by using the statewide-equation method of Gebert and others (2007), described below. Additional adjustments to the base-flow measurements, described below, were necessary to create a coherent calibration dataset that was spatially consistent with field observations. For sites with only single head measurements, target weights were adjusted to account for additional uncertainty due to temporal variability not captured by the target value.

Only 21 wells (from the NWIS database) had two or more measurements. For these wells, average head values were computed, for the period after 1969, if available. If less than two records post-1969 were available, an average of all records was used. The WCRs and Bad River Reservation wells only had single measurements.

Mean annual base flows were computed for USGS stations 04027000 (Bad River near Odanah), 04026561 (Tyler Forks River at Stricker Road) and 04026390 (Beartrap Creek at Highway 2) by using the Institute of Hydrology (1980a; 1980b) method, as implemented by the Base-Flow Index (BFI) software (Wahl and Wahl, 2013). Since streamflow data at 04026561 were only available for 2012 and 2013, a relation was developed between the annual base flows for those years and coincident annual base flows at 04027000. This relation was then used to estimate a long-term average annual base flow for 04026561, by using the average annual base flow at 04027000 for 1970-1999. For consistency with the base flows obtained from Gebert and others (2011), the average annual base flow from 1970-1999 was used for 04027000.

For the synoptic base-flow measurements taken in 2011 and 2013, the statewide-equation approach of Gebert and others (2007) was used to estimate average values. A base-flow factor was developed for each site:

$$B_f = \left(\frac{Q_m}{Q_r} \right) \left(\frac{Q_{90}}{A} \right) \tag{3.1}$$

where

Q_m is the measured discharge at the site;

Q_r is the corresponding discharge at a nearby index station;

Q_{90} is the 90-percent flow duration at the index station; and

A is the drainage area at the measurement site.

The gage on the Bad River near Odanah, Wisconsin (USGS station 04027000) was used as the index station. The base-flow factor was then used in the statewide equation to estimate annual average base flow:

$$Q_b = 0.907 A^{1.02} B_f^{0.52} \tag{3.2}$$

Reconciliation of Base-Flow Estimates from Multiple Sources

At locations with both a value in Gebert and others (2011) and a synoptic measurement, the Gebert and others (2011) value was used. The only exception to this condition was the measurement at station 4026740, on the Marengo River near Mellen. The Gebert and others (2011) value of 122.8 ft^3/s, based on a single measurement taken in 1970, was 37 ft^3/s higher than the sum of the three upstream tributaries (sites 04026600, 04026620, and 04026650), and was consistently under-simulated in the model (i.e. the modeled values were less than the measured values), by a similar amount. Two measurements were taken at this site in 2011, both rated "Good," meaning the likely error is 5 percent or less. The average of these two values, adjusted with the statewide equation, corresponded more closely to the simulated values, and the sum of the upstream tributaries. Therefore, the adjusted average of the 2011 measurements was used for this target.

Initially, only the adjusted 2011 synoptic measurements were used in the calibration. However, this resulted in an apparent regional bias in simulated base-flow error. Base-flow values throughout the Potato, Tyler Forks, and Upper Bad Watersheds were consistently over-simulated, while those in the Marengo Watershed were under-simulated. Different configurations of recharge (via adjustment of SWB input parameters, or zoned-based multiplication of the post-SWB recharge array) yielded similar results. By comparison, the error between the simulated values and the Gebert and others (2011) values showed little or no spatial bias.

Whereas the Gebert and others (2011) values (and the statewide equation on which many of them are based) represent a period of three decades, the 2011 measurements were collected following a hot summer, after several years of drought in northwest Wisconsin. In addition, the shallow, fractured crystalline rock hydrogeologic setting of the Potato, Tyler Forks, and Upper Bad Watersheds is different from most other watersheds in the State, which typically have either thicker glacial deposits, or deep sedimentary rocks with relatively high capacities for groundwater storage. In contrast, the Marengo Basin is fed by thick glacial deposits and the deep sandstone aquifer. Base flows in streams draining low-storage aquifers would be expected to be disproportionately affected by extended dry periods. It is therefore possible that the statewide equation, which represents base flows throughout Wisconsin, is not capable of accurately predicting base-flow conditions in the Potato, Tyler Forks and Upper Bad Watersheds during exceptionally dry conditions. It is also possible
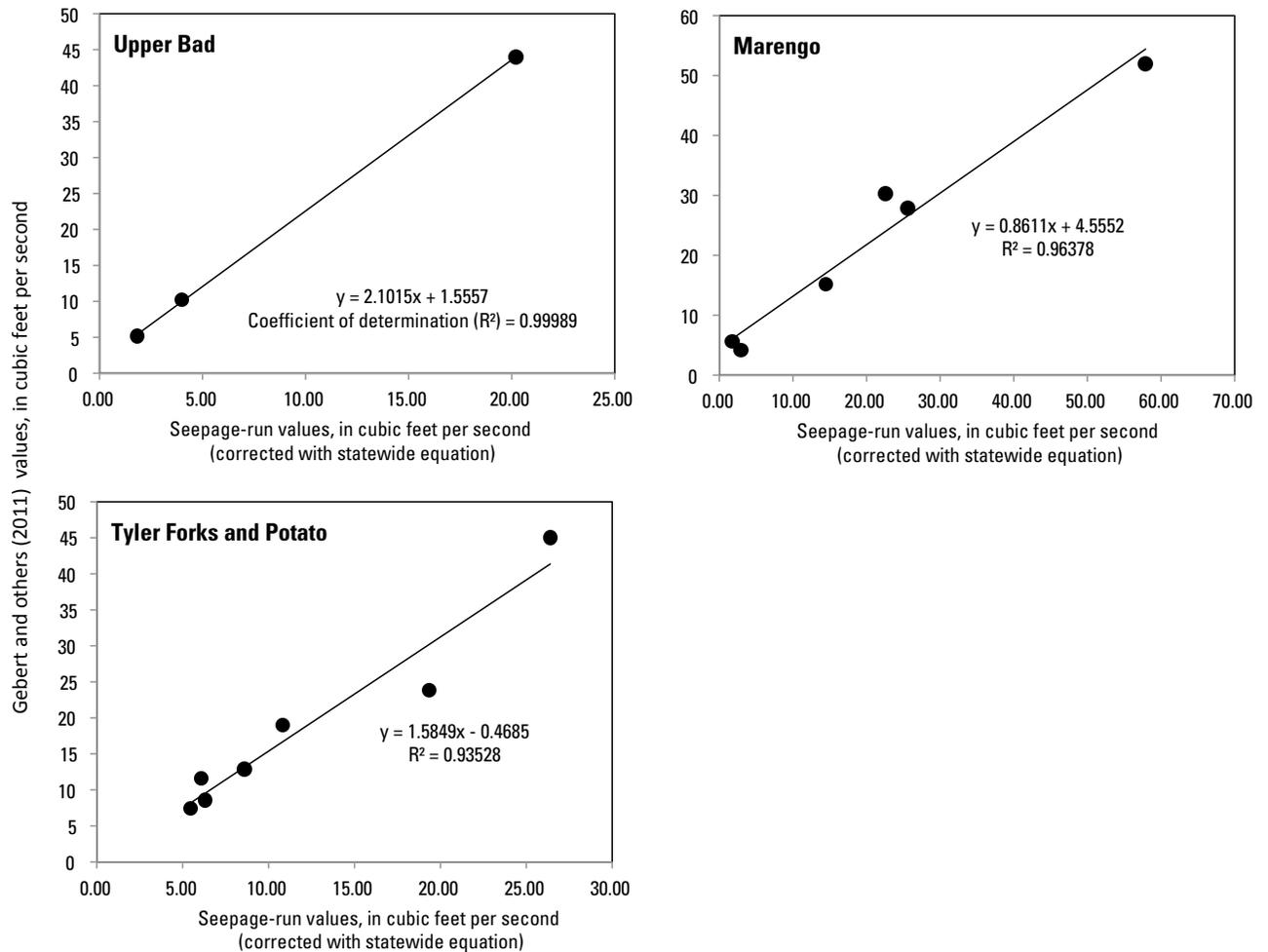
Figure 3.1.   Relations between annual base flows published in Gebert and others (2011) and annual base flows estimated from 2011 field measurements, by using the statewide equation.

that the Bad River gage near Odanah is a poor index station for streams in the bedrock uplands, as low flows at Odanah are likely buffered substantially by contributions from the sandstone aquifer, in the lower Bad and Marengo Basins. A systematic bias in the 2011 statewide equation adjusted values, compared to the Gebert and others (2011), is illustrated in figure 3.1. The relationships shown in these figures were used to normalize the statewide-equation estimates of the 2011 synoptic measurements to the annual base flows published in Gebert and others (2011). For example, base-flow values from the Tyler Forks / Potato Watersheds were multiplied by 1.58, the slope of the relationship line shown in figure 3.1.

Base-flow measurements collected in 2013 also were included in the calibration dataset, to better inform the distribution of flow around the proposed mine site. These values were made consistent with the adjusted 2011 measurements by comparing duplicates. For example, the 2013 measurement for Devils Creek immediately below the confluence with Ballou Creek was 2.62 ft^3/s, and the adjusted 2011 measurement was 5.25 ft^3/s (a factor of 2.00 greater). A multiplier of 2.00 was applied to all of the 2013 measurements collected in the Devil's and Ballou Creek Watersheds. Since there was no duplicate measurement on the Tyler Forks River, the 2011 measurement at Highway 77 was compared to the 2013 measurement at Moore Park Road, and the resulting multiplier was applied to the *difference* between the measurement upstream of the Penokee ridge and the measurement at Moore Park Road. Although this comparison is imperfect, it is useful in that it informs the calibration process that a neutral to slightly gaining condition should be simulated across the Penokee ridge. The increased uncertainty in this target value and uncertainties in the other adjusted target values discussed above were addressed by applying lower weights, as discussed in the next section.

Observation Weighting

Observations in the calibration dataset were weighted to address differences in information content and measurement uncertainty related to measurement quality, location uncertainty, and temporal variability. Generally, higher weights were assigned to the most reliable observations (those with lower uncertainty), and lower weights to the lower quality measurements (those with higher uncertainties). Observations were grouped based on the different sources listed above and then by quality. Weights were then adjusted by group, to balance the contributions of different observation types to the initial objective function. The overall goal of the weighting was to maximize the transfer of information from the observation dataset to the estimated parameters (for example, Doherty and Hunt, 2010).

Head observations from the NWIS database were categorized as "best," "good," "fair," or "poor," based on number of measurements at each location, time period, wellhead-elevation accuracy, and other metainformation. WCRs were separated into two groups based on the uncertainty in their wellhead elevations (which translates to uncertainty in their water-level measurements). Uncertainty in wellhead elevations was evaluated by using location confidence values supplied by the Wisconsin Geological and Natural History Survey (WGNHS). A buffer radius equal to the location confidence was delineated for each WCR, and then wellhead-elevation uncertainty was computed as two standard deviations of the 10-meter DEM values (U.S. Geological Survey, 2014a) within the buffer area for each well. WCRs with wellhead-elevation uncertainties <20 ft (n=987) were placed in the WCRs1 group, while the rest (n=200) were included in WCRs2. Weights within each head observation group were assigned to be inversely proportional to estimated measurement uncertainty. For example, heads in the NWIS "best" category were initially assigned weights of 0.2, which corresponds to a standard deviation of 5 ft, and a 95-percent confidence interval of approximately +/- 10 ft, while heads in the NWIS "fair" and WCRs1 categories were given initial weights of 0.029 (a 95-percent confidence interval of approximately +/- 70 ft). Observations in groups WCRs2 and NWIS "poor" were given weights of zero—meaning they had no influence on the parameter-estimation process, but were reported in the parameter-estimation output so as to be easily assessed manually during calibration.

Base-flow observations were placed into five groups, based on the different measurement categories discussed previously, and their reliability/magnitude. Measurements of larger fluxes were generally assumed to be of higher quality than measures of headwater streams. Measurements from Gebert and others (2011) were also assumed to be more reliable, for the reasons discussed previously. Weights within each base-flow observation group were initially assigned to be approximately inversely proportional to measurement uncertainty, calculated as the observation value multiplied by an estimated coefficient of variation. Table 3.1 shows head and flux targets with their estimated coefficients of variation and weights.

After the observations groups and their initial uncertainty-based weighting were developed, the objective function was "balanced" following the methodology of Doherty and Hunt (2010). Multipliers were applied to several flux groups, so that they had similar contributions to the objective function (Φ) as the head observations (table 3.1). A balanced objective function is important for ensuring that all important components of the observation dataset are "seen" by the parameter-estimation algorithm, so that they inform the parameter estimates.

Calibration Approach

The model was calibrated by regularized inversion (Hunt and others, 2007; Doherty and Hunt, 2010), which included the use of:

- pilot points (Doherty, 2003; Doherty and others, 2010) in addition to a traditional parameter zones;

- Tikhonov regularization (Tikhonov, 1963a; Tikhonov, 1963b; Doherty, 2003; Doherty and others, 2010);

- singular value decomposition (Tonkin and Doherty, 2005; Hunt and others, 2007).

The advantages of these tools are discussed by Hunt and others (2007). "Best practices" for their application are given by Doherty and Hunt (2010) and Doherty and others (2010). A total of 1,715 adjustable parameters were included in the calibration. These included a global multiplier for the recharge array, a parameter for the hydraulic conductivity of GHB cells representing Lake Superior, a parameter for the hydraulic conductivity of all SFR cells, and 1,712 adjustable pilot points. The techniques used to estimate these parameters are discussed in the following sections.

Table 3.1.   Observation weighting, with contributions to the total objective function (phi) at the start of model calibration. Final observation weights were computed as the inverse of measurement error times a group multiplier (used to balance the contributions between heads and fluxes in the objective function).

[NWIS, National Water Information System; --, not applicable]

Type	Data Source	Quality group	Estimated uncertainty	Group weight multiplier	Observation weights	Number of observations in group	Phi contribution
Heads	NWIS	Best	5 (feet)	1	0.200	6	169.1
Heads	NWIS and Bad River Band	Good	15	1	0.067	17	80.4
Heads	All sources	Fair	35	1	0.029	1,144	1,618.4
Heads	All sources	Zero-weighted	--	--		231	0.0
Base flows	NWIS	Bad River near Odanah	0.01 (CV[1])	3	3 / (0.01 x flow)	1	7.2
Base flows	Gebert and others (2011)	--	0.126–0.19	3	3 / (0.126–0.19 x flow)	18	1,951.3
Base flows	NWIS	2011 base-flow measurements	0.5	5	5 / (0.5 x flow)	11	251.1
Base flows	NWIS	2013 base-flow measurements	1	2	2 / (1 x flow)	8	3.3
Base flows	NWIS	2011 base-flow measurements, small streams[2]	1	3	3 / (1 x flow)	19	137.2
		Total weighted observations				**1,224**	**4,218.0**

[1]Coefficient of variation. Values of 0.126 reflect the standard error in the relation-line approach of Gebert and others (2011), while 0.19 reflects a standard error of 14 percent in the statewide equation approach, and an estimated 5 percent error for a "good" quality rating in NWIS. Higher values reflect increasing uncertainty inherent in measuring small streams.

[2]Streams measured at 0.5 cubic feet per second or less. Streams with no measured flow were assigned weights of 1 x 10[-5].

Pilot Points

Pilot points increase spatial flexibility in the model parameterization, allowing the calibration process to extract more information from the observation data, which can compensate for structural error in the model (Hunt and others, 2007; Doherty and Welter, 2010). The increased flexibility comes at the cost of a greatly increased number of parameters, however, which can result in parameter insensitivity and correlation, solution nonuniqueness and an ill-posed inverse problem (Hunt and others, 2013). "Regularization" constraints can address these issues by facilitating a stable and unique solution to the inverse problem. Two types of regularization were used in calibrating the Bad River MODFLOW model–soft-knowledge/Tikhonov regularization and subspace regularization, as performed by using singular value decomposition (see Hunt and others (2007) and Doherty and Hunt (2010) for more details).

Tikhonov Regularization

Tikhonov regularization allows for the incorporation of "soft" information into the calibration process while simultaneously promoting a unique solution. Soft information can be incorporated into the calibration process as preferred parameter values (for example, "the hydraulic conductivity should have a value around 1 foot per day"), or preferred homogeneity ("the hydraulic conductivity should be uniform in this area"), or preferred differences ("the hydraulic conductivity there is expected to be higher than here") (Hunt and others, 2013). The preferred conditions are implemented as an internal set of equations describing the differences between parameters and their preferred values, or preferred differences between parameter values. The equations act as pseudo-observations, allowing deviations from the preferred conditions to be penalized by increases in the objective function. For a deviation to occur, it must be balanced by an improved fit between the model outputs and field observations, resulting in an overall smaller objective function. They also enhance numerical stability of the parameter estimation by providing a fall-back position for parameters that are not constrained by the observation dataset available.

The regularization equations comprise a regularization objective function that is minimized in tandem with the measurement objective function describing the model's fit to observations, subject to the constraint that the measurement objective function adhere to a user-specified target (if this can be achieved). This target (known as the PHIMLIM variable in PEST) controls the trade-off between fitting the observation data, and achieving a parameter field that is most consistent with the soft knowledge described above. If PHIMLIM is set too low, parameter fields may become unrealistic as a result of fitting to measurement noise (the classic problem of "overfitting"); if set too high, parameter fields may fit the preferred condition too well, and the calibration process will fail to gain information from the observation data. Fienen and others (2009b) describe PHIMLIM in detail.

The Tikhonov regularization for horizontal hydraulic conductivity and vertical anisotropy pilot points in the Bad River model consisted of preferred differences between neighboring pilot points of zero, resulting in a preferred homogeneity soft-knowledge constraint, for each pilot point zone listed in appendix 2. Interregularization group weights adjustment was also activated (PEST variable IREGADJ=1; Doherty and Hunt, 2010, p. 20); thus, the soft-knowledge importance among parameter types was adjusted during the calibration to ensure a similar relative importance for each parameter group soft-knowledge preferred condition. A PHIMLIM value equal to the number of weighted observations (1,224) was used, which is consistent with the expected value of the objective function if the weights assigned to the observations are equal to the inverse of their measurement error (Fienen and others, 2009b). A Pareto analysis of the calibration, which systematically investigated the tradeoff between the regularization and measurement objective functions, indicated that this value for the measurement-objective function was not an extreme value (see "Pareto Analysis").

Singular Value Decomposition

In contrast to Tikhonov regularization, which adds pseudo-observation information to the calibration process, subspace methods remove insensitive parameters, or combinations of correlated parameters, that are inestimable on the basis of the observations. Inestimable parameters are identified through singular value decomposition (SVD) of the weighted Jacobian matrix (see Moore and Doherty, 2005; Tonkin and Doherty, 2005). The Jacobian matrix consists of one column for each adjustable model parameter, and one row for each model output that is equivalent to a field observation. Each entry in the Jacobian matrix records the sensitivity of a model output with respect to a model parameter (that is, the amount the output changes with respect to a factor change in that parameter). Parameters deemed inestimable (considered to be in the calibration null space) are left at their initial values. The remaining parameters, comprising the

calibration solution space, are estimated. An eigenvalue ratio threshold of 5×10^{-7} was used to demarcate the boundary between the solution and null space (Doherty, 2010a, p. 8-12).

SVD–Assist

Calculating the Jacobian matrix requires a model run for each adjustable parameter. In addition, multiple Jacobians (each followed by a successive upgrade of parameter values) are computed in the nonlinear-regression process. "SVD–assist" (or SVDA) is an extension of the SVD approach that can lighten this computational burden by allowing a reduced number of "superparameters" to be estimated. Superparameters are defined at the beginning of the calibration process, by using the sensitivities calculated from the full set of initial parameter values. The full parameter space is then reduced by projecting base parameters onto a reduced set of axes composed of linear combinations of base parameters (Tonkin and Doherty, 2005). This reduction is easily obtained by using an algorithm within PEST (Doherty and Hunt, 2010, p. 21). The resulting superparameters are linear combinations of the base parameters spanning the calibration solution space; their coefficients must be multiplied in order to achieve a calibrated set of "native" or "base" model parameters. Three hundred superparameters (out of 1,715 base parameters) were used in the calibration of the Bad River groundwater model.

Successive runs of the parameter-estimation process were conducted on multiple processors by using the modeling array at the USGS Wisconsin Water Science Center and a parallelized version of PEST called BeoPEST (Schreüder, 2009). The results of each run were used to evaluate adjustments and refinements to the conceptual model and parameterization, observation dataset, and observation weighting.

Pareto Analysis

Pareto analysis (Doherty, 2014, Moore and others, 2010, Doherty, 2010b) systematically explores the trade-off between closely fitting field observations, and adhering to soft information such as what parameter values are geologically reasonable. In the context of model calibration, the Pareto front consists of a locus of possible models (that is, parameter sets; shown in figure 3.2) that are optimal in the sense that neither the measurement nor regularization objective functions can be improved without a degradation of the other one (Doherty, 2010b). Space to the right of this curve represents many other possible parameter combinations (models) that are not optimal.

PEST traverses the Pareto front through the use of a weighting factor on the measurement objective function, which allows the influence of the measurements on the parameter-estimation process to be systematically adjusted. Pareto calibration of the Bad River model was initiated with a measurement weighting factor of 0 (strong regularization), with the parameters at their precalibration values. The measurement weighting factor was then increased by an increment of 0.05 every two PEST iterations (each iteration consisting of the calculation of a Jacobian matrix of superparameter sensitivities, followed by an upgrade to a new parameter set). With each increase in the measurement weighting factor, the influence of the regularization on the parameter-estimation process is reduced, allowing for the parameter values to deviate further from their preferred condition so that the observations are better matched, thereby increasing the value of the regularization objective function.

Figure 3.2 shows the Pareto front for the Bad River model. Note that there are relatively few points in the lower portion of the curve, indicating more unique solutions when the regularization constraints are highly weighted compared to the observations. Conversely, the upper portion of the curve (low regularization) is characterized by nonuniqueness, where many candidate solutions provide somewhat similar fits to the observations, but with low adherence to the preferred condition of homogeneity in the pilot point parameters. The Pareto optimum marks the modeler's view of the ideal trade-off between these two competing interests. In this case, based on comparison of the hydraulic conductivity fields and observation residuals, the optimum was subjectively chosen to be in the vicinity of a value of 1,200 for the measurement objective function, near PEST iteration 20 in figure 3.2. This indicates that the chosen PHIMLIM value of 1,224 in the final model calibration represents a good balance away from the two extremes of best possible model fit and overly strict adherence to soft information.
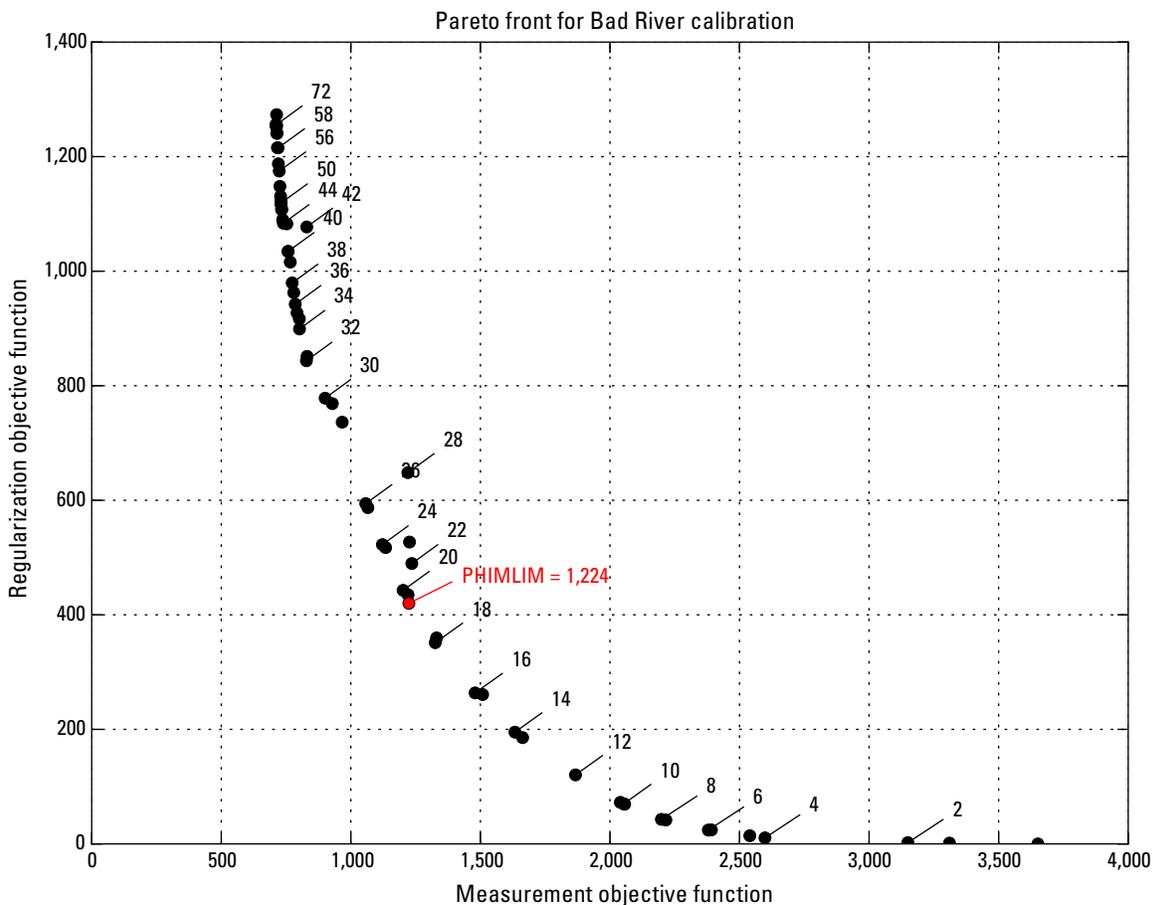
Figure 3.2. The Pareto front displaying the tradeoff of model fit and soft-information applied to the Bad River groundwater-flow model. Within the software, the front was traversed by increasing the weighting factor on the measurement objective function by 0.05 every two PEST iterations. The labels indicate the PEST iteration number. The location of the final model chosen on the front is seen at a Measurement Objective Function of 1224.

Appendix 4—Calibration Results
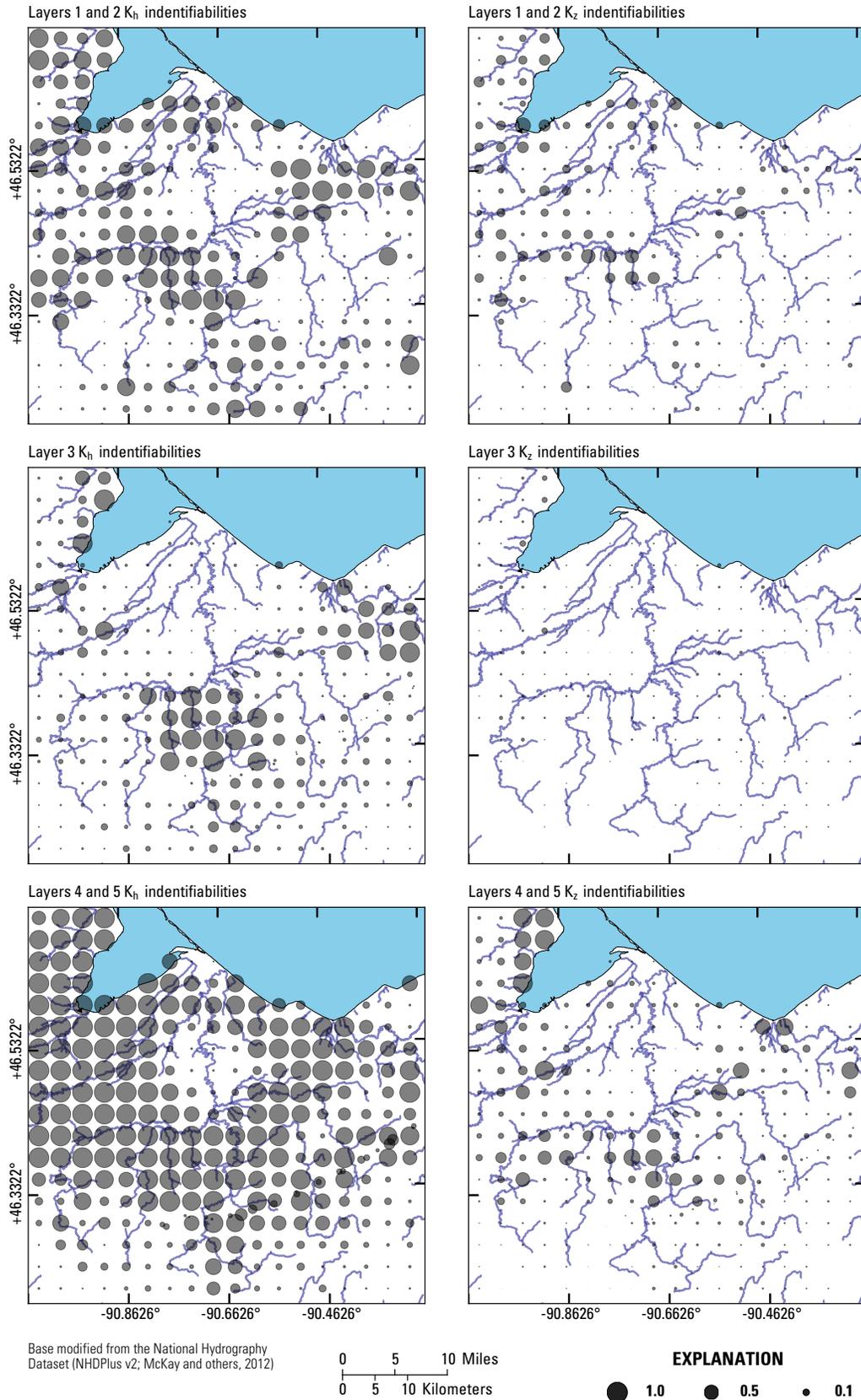
Figure 4.1. Identifiabilities of pilot point parameters. Higher values occur in areas where observations (especially groundwater levels) help constrain the possible values of hydraulic conductivity. High identifiabilities can be seen along the major highway corridors and near population centers. Identifiabilities are lower in areas such as the central part of the Bad River Reservation, Wisconsin, where data are sparse. Low identifiabilities for vertical hydraulic conductivity indicate that the calibration dataset does not provide much information on these parameters.
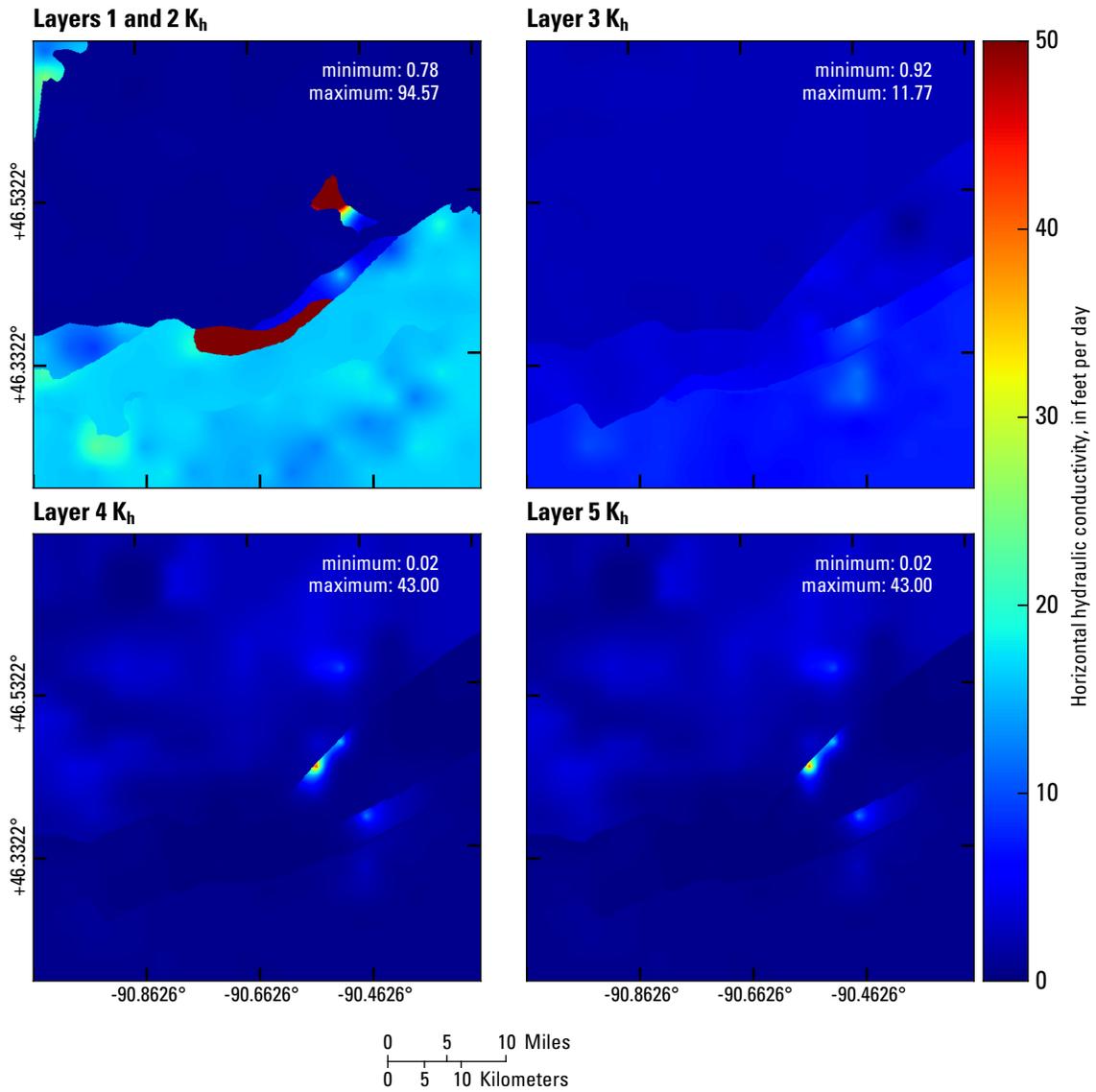
Figure 4.2.   Calibrated horizontal hydraulic conductivity (K_h).
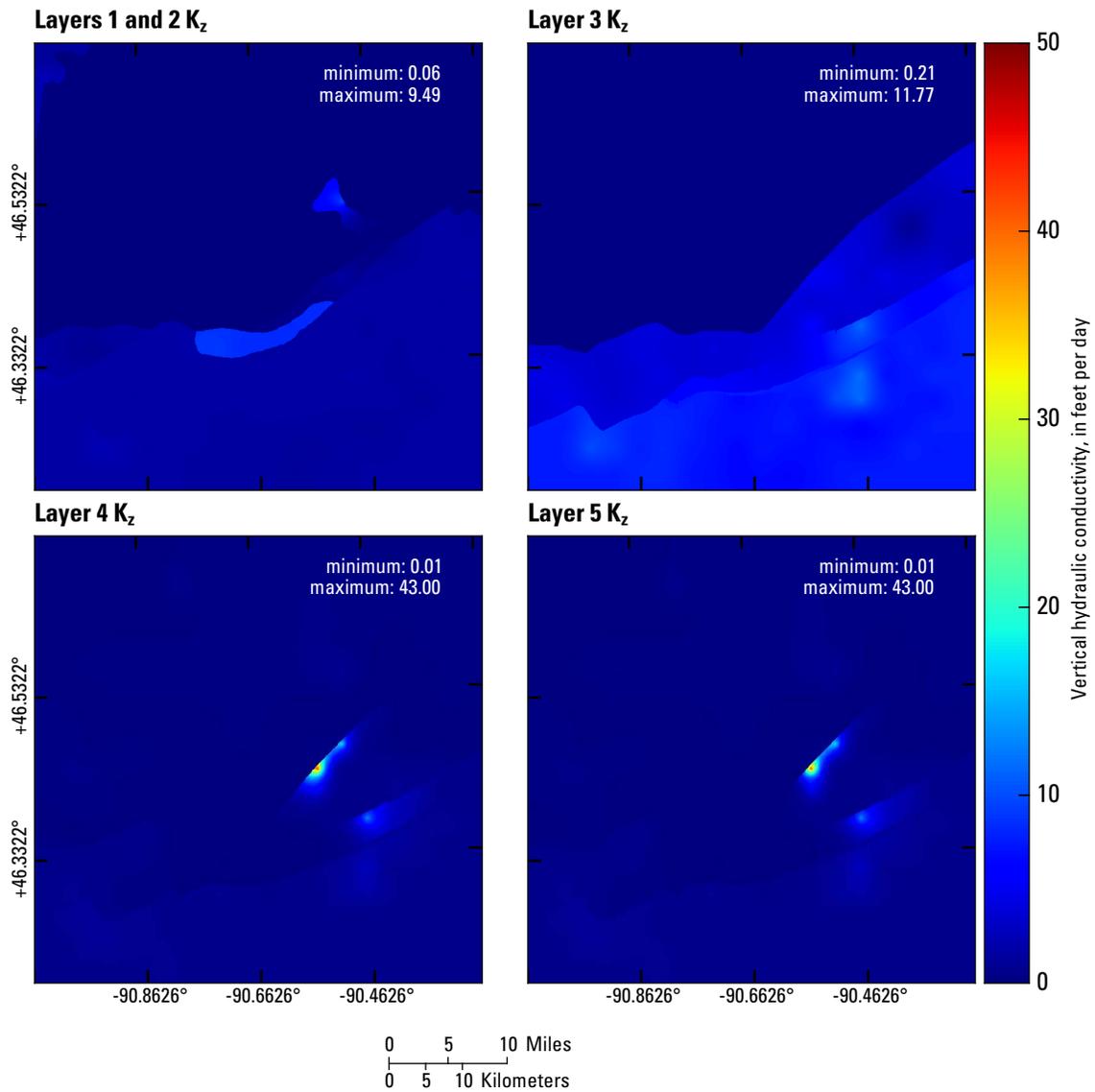
Figure 4.3.   Calibrated vertical hydraulic conductivity (K_z).
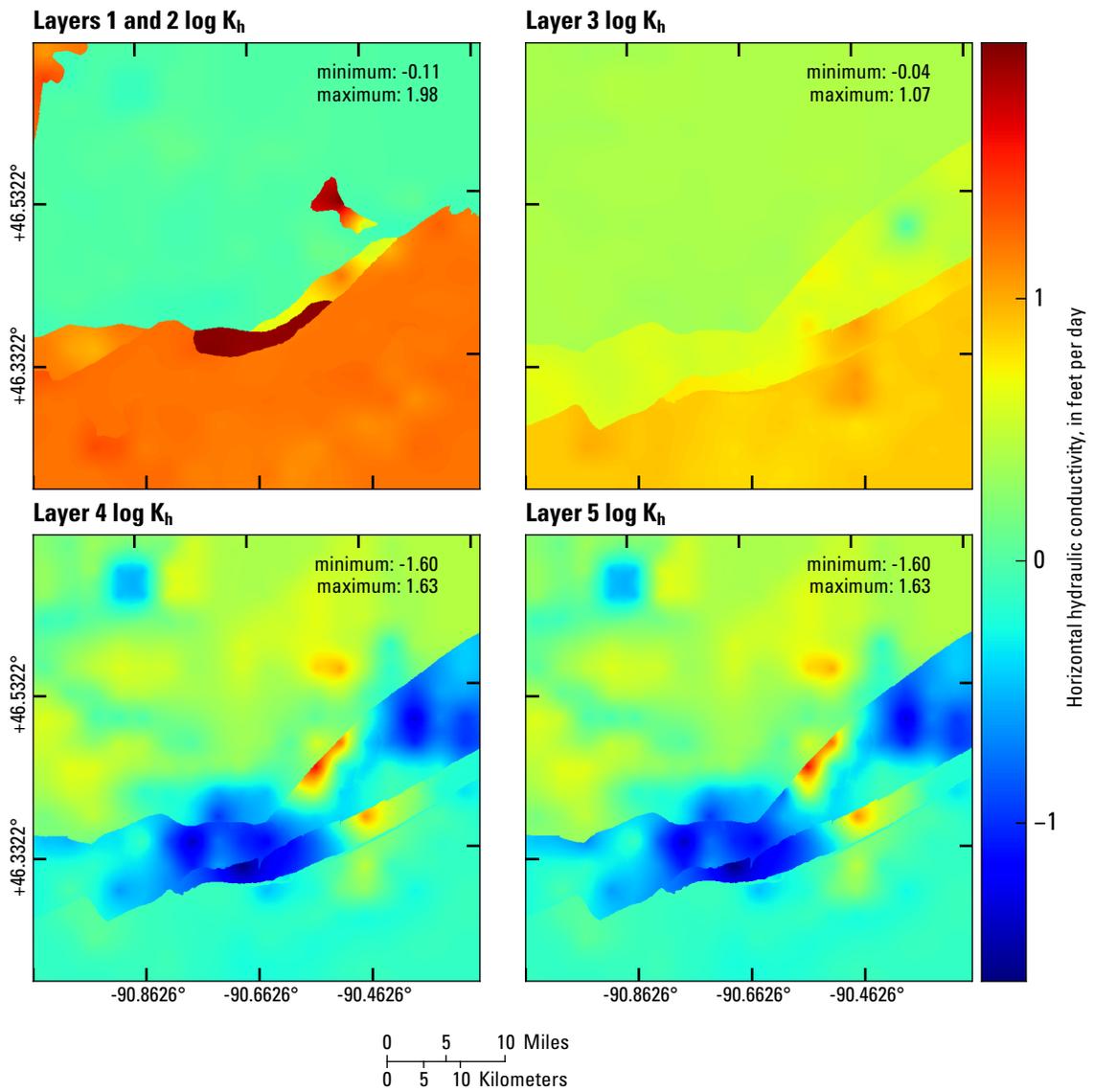
Figure 4.4.   Calibrated log horizontal hydraulic conductivity (K_h).
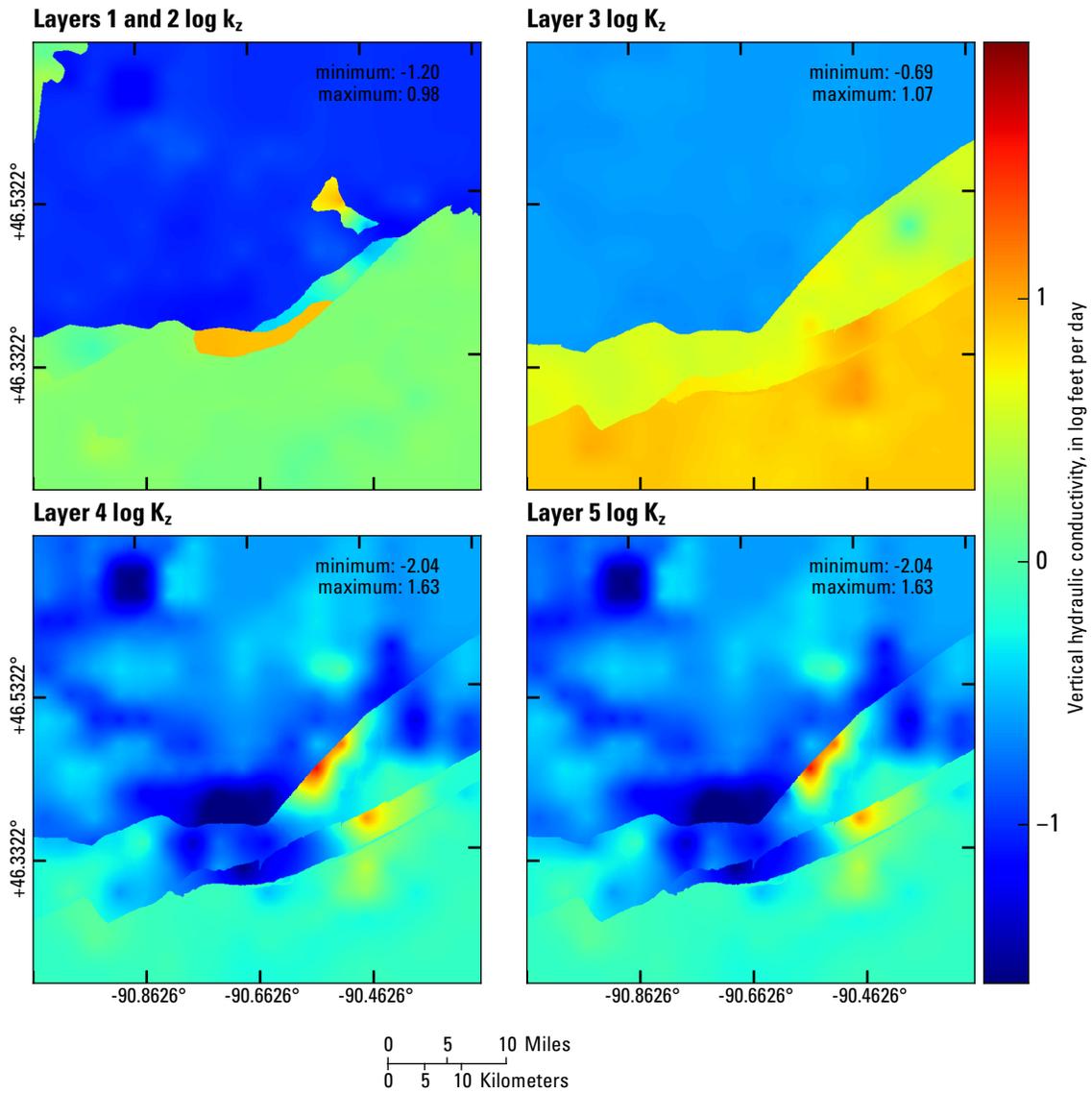
Figure 4.5.   Calibrated log vertical hydraulic conductivity (K_z).

Table 4.1.   Summary of hydraulic conductivity parameterization before and after calibration.

[ft/d, feet per day; --, Insufficient values or not reported; MRS, Midcontinent Rift System]

Values reported by K. Bradbury[1]

Hydrogeologic Unit	Model Zones	Geometric mean Kh (ft/d)	Range (ft/d)
Copper Falls Fm. sand	7, 10, 11, 12	41.2	0.4 - 1,060
Miller Creek Fm.	8	--	--
MRS sedimentary	6	2.5 - 3	0.028 - 1330
Copper Falls Fm. sandy till	9	16.5	0.016 - 473
Area of shallow bedrock	13	--	--
MRS igneous	1	3.9	0.0026 - 9130
Marquette Supergroup	2,4,5	--	--
Archean Units	3	7.5	0.007 - 1190

Calibrated Kh pilot point values

Hydrogeologic Unit	Precalibration (starting) values (ft/d)	Geometric mean (ft/d)	Minimum (ft/d)	[3]- 2σ (ft/d)	[3]+ 2σ (ft/d)	Maximum (ft/d)	Number of pilot points[4]
Copper Falls Fm. sand	41.2	36.0	2.95	--	--	95	8
Miller Creek Fm.	1	1.00	0.78	0.84	1.2	1.3	179
MRS sedimentary	2.5	2.2	0.105	0.81	5.9	9.8	354
Copper Falls Fm. sandy till	16.5	15.1	3.49	8.7	26.1	23.3	83
Area of shallow bedrock	16.5	16.4	9.2	13	20.0	22.0	53
MRS igneous	[2]3.9, 0.39	1.17	0.038	0.066	21	43	108
Marquette Supergroup	[2]7.5, 0.75	2.1	0.025	0.126	35	13	70
Archean Units	[2]7.5, 0.75	2.4	0.33	0.233	25	12	154

Calibrated Kv pilot point values

Hydrogeologic Unit	Precalibration (starting) values (ft/d)	Geometric mean (ft/d)	Minimum (ft/d)	[3]- 2σ (ft/d)	[3]+ 2σ (ft/d)	Maximum (ft/d)	Number of pilot points[4]
Copper Falls Fm. sand	4.12	3.6	0.29	--	--	9.5	8
Miller Creek Fm.	0.1	0.1	0.06	0.08	0.1	0.2	179
MRS sedimentary	0.25	0.2	0.01	0.08	0.6	1.0	354
Copper Falls Fm. sandy till	1.65	1.5	0.31	0.83	2.7	2.4	83
Area of shallow bedrock	1.65	1.6	0.98	1.36	2.0	2.2	53
MRS igneous	[2]3.9, 0.39	1.2	0.04	0.07	20.5	43.0	108
Marquette Supergroup	[2]7.5, 0.75	2.1	0.02	0.13	35.4	12.8	70
Archean Units	[2]7.5, 0.75	2.4	0.33	0.23	24.6	11.8	154

[1]Wisconsin Geological and Natural History Survey, unpublished data (2013).

[2]For the crystalline bedrock, Layer 3 pilot points were initially given the geometric mean values from K. Bradbury[1]; Layers 4 and 5 were set one order of manitude lower.

[3]Assuming a log-normal distribution.

[4]Excluding Layer 5.

Appendix 5—Table of Stable Isotope Sampling Results

Table 5.1.   Results of stable isotope sampling.

[USGS, U.S. Geological Survey; δ18O, 18O/16O; δD, deuterium/1H; in units of "per mil" or ‰ relative to Vienna Standard Mean Ocean Water (VSMOW) standard. Analyses of water isotopes were performed at the USGS Reston Stable Isotope Laboratory (http://isotopes.usgs.gov/).]

USGS station number	Longitude	Latitude	δ18O	δD	Sample depth	Bottom material
463725090464901	-90.78	46.62	-10.88	-75.69	58.5	sand
463030090485501	-90.82	46.51	-11.10	-76.41	[1]254	gravel
462423090464801	-90.78	46.41	-11.19	-80.47	[1]209.9	sandstone
463550090393801	-90.66	46.60	-11.27	-78.02	[1]200.5	Sand
463030090590801	-90.99	46.51	-11.28	-79.20	[1]215	sand
461944090363001	-90.61	46.33	-11.31	-78.47	150.0	likely crystalline
463448090551201	-90.92	46.58	-11.40	-80.79	[1]208.6	sandstone
462910090251301	-90.42	46.49	-11.45	-78.25	190.0	crystalline
462241090424801	-90.71	46.38	-11.45	-80.65	147.5	crystalline
461728090344501	-90.58	46.29	-11.50	-79.75	65.0	sand and gravel
461725090344301	-90.58	46.29	-11.65	-80.59	48.0	sand and gravel
461704090341201	-90.57	46.28	-11.81	-81.37	117.0	sand and gravel
461701090314801	-90.53	46.28	-11.83	-82.64	32.0	gravel
461649090350301	-90.58	46.28	-11.84	-81.62	87.5	crystalline
463605090405901	-90.68	46.60	-11.89	-83.87	[1]200	sandstone
462547090591401	-90.99	46.43	-11.91	-82.89	[1]325.9	sand
461644090352901	-90.59	46.28	-12.05	-84.18	81.0	sand and gravel
463547090413101	-90.69	46.60	-12.13	-85.55	190.0	sandstone
463655090484201	-90.81	46.62	-12.22	-86.83	[1]88.8	sandstone
463647090484201	-90.81	46.61	-12.23	-86.11	162.0	sandstone
463216090342401	-90.57	46.54	-12.26	-86.76	183.0	sand
463217090343001	-90.57	46.54	-12.28	-86.23	[1]183.5	sand
462657090231301	-90.39	46.45	-12.29	-85.28	20.0	gravel
461658090345801	-90.58	46.28	-12.34	-86.51	110.0	crystalline
463506090464401	-90.78	46.58	-12.44	-88.62	204.0	sand
463214090342001	-90.57	46.54	-12.44	-87.25	[1]183.5	sand
462535090331601	-90.55	46.43	-12.50	-87.14	81.5	crystalline
463458090550201	-90.92	46.58	-12.50	-88.49	[1]221.1	sandstone
463211090342601	-90.57	46.54	-12.54	-87.80	158.5	sand
463209090342501	-90.57	46.54	-12.56	-89.24	176.0	sand
463542090435101	-90.73	46.59	-12.84	-92.13	190.0	sand and gravel
463030090300301	-90.50	46.51	-12.85	-89.54	[1]345	sandstone
462835090295901	-90.50	46.48	-12.98	-91.53	159.0	sandstone
462944090293901	-90.49	46.50	-13.07	-91.72	160.0	sand
463631090425401	-90.72	46.61	-13.31	-95.48	[1]150.4	sandstone
463635090452301	-90.76	46.61	-13.42	-95.81	83.0	sand
463528090434601	-90.73	46.59	-13.66	-98.73	131.0	sand
463605090411001	-90.69	46.60	-14.37	-102.87	[1]208	sandstone
463646090464701	-90.78	46.61	-14.54	-103.46	143.0	sandstone
463609090424601	-90.71	46.60	-14.66	-104.38	175.0	sand and sandstone
463718090483401	-90.81	46.62	-14.71	-104.04	73.5	sandstone
463710090421301	-90.70	46.62	-15.01	-107.31	[1]220	sandstone
463709090420901	-90.70	46.62	-15.16	-108.57	258.0	sandstone
463631090451001	-90.75	46.61	-15.17	-107.72	242.0	sandstone
463606090411801	-90.69	46.60	-15.21	-109.61	217.0	sandstone
463719090482501	-90.81	46.62	-15.24	-109.23	[1]76	sandstone
463535090421101	-90.70	46.59	-15.66	-112.76	268.0	likely sandstone
463631090453101	-90.76	46.61	-15.70	-113.03	212.0	sandstone
463632090451301	-90.75	46.61	-16.09	-116.85	[1]274.5	sandstone
463734090465001	-90.78	46.63	-16.25	-117.74	77.5	sandstone
463534090434201	-90.73	46.59	-16.32	-118.64	[1]282	sandstone
463631090450201	-90.75	46.61	-16.89	-122.72	292.0	sandstone
463460090471701	-90.79	46.58	-17.25	-123.42	208.0	likely sandstone
463627090413101	-90.69	46.61	-17.54	-125.09	[1]210.8	sandstone

[1]Sample depth estimated from nearby well logs

Appendix 6—Additional Results from New Data-Worth Analysis

Description of Linear Uncertainty Analysis

Prediction uncertainties were computed by using PREDUNC5, which is implemented in the equation (Doherty, 2014, p. 198):

$$\sigma_{post-calibration}^2 = \mathbf{y}^T \mathbf{C_p} \mathbf{y} - \mathbf{y}^T \mathbf{C_p} \mathbf{X}^T \left[\mathbf{X} \mathbf{C_p} \mathbf{X}^T + \mathbf{C}_\varepsilon \right]^{-1} \mathbf{X} \mathbf{C_p} \mathbf{y} \tag{6.1}$$

where

$\sigma_{post-calibration}^2$ is the prediction uncertainty after model calibration;

$\mathbf{C_p}$ is a covariance matrix containing the variance in the input parameter values on the diagonal (no off-diagonal terms were considered in this work);

$\mathbf{X}$ is the Jacobian matrix containing sensitivities of observations to incremental changes in model parameters;

$\mathbf{C}_\varepsilon$ is the uncertainty (error) in observations used to calibrate the model;

$\mathbf{y}$ is a vector containing sensitivities of the model predictions to incremental changes in the parameters, under the future stress condition; and

$(\cdot)^T$ indicates a matrix or vector transpose.

PREDUNC5 returns three sets of uncertainty values:

1. Precalibration uncertainty ($\sigma_{pre-calibration}^2 = \mathbf{y}^T \mathbf{C_p} \mathbf{y}$ in equation 6.1), which is only based on uncertainty in the model-input parameters.

2. Post-calibration uncertainty ($\sigma_{post-calibration}^2$), as calculated in equation 6.1 using only the base observations.

3. For each hypothetical new observation n, a post-calibration uncertainty ($\sigma_{post-calibration+n}^2$) computed with that observation added to the base observations (in $\mathbf{X}$)

The "data worth" (or equivalently, normalized uncertainty reduction) of each hypothetical new observation n is then computed by comparing the prediction uncertainty values for the stressed conditions with and without the new observations after normalizing for the original prediction uncertainty (Fienen and others, 2010, p. 14):

$$data\,worth_n = \frac{\sigma_{post-calibration}^2 - \sigma_{post-calibration+n}^2}{\sigma_{post-calibration}^2} \tag{6.2}$$
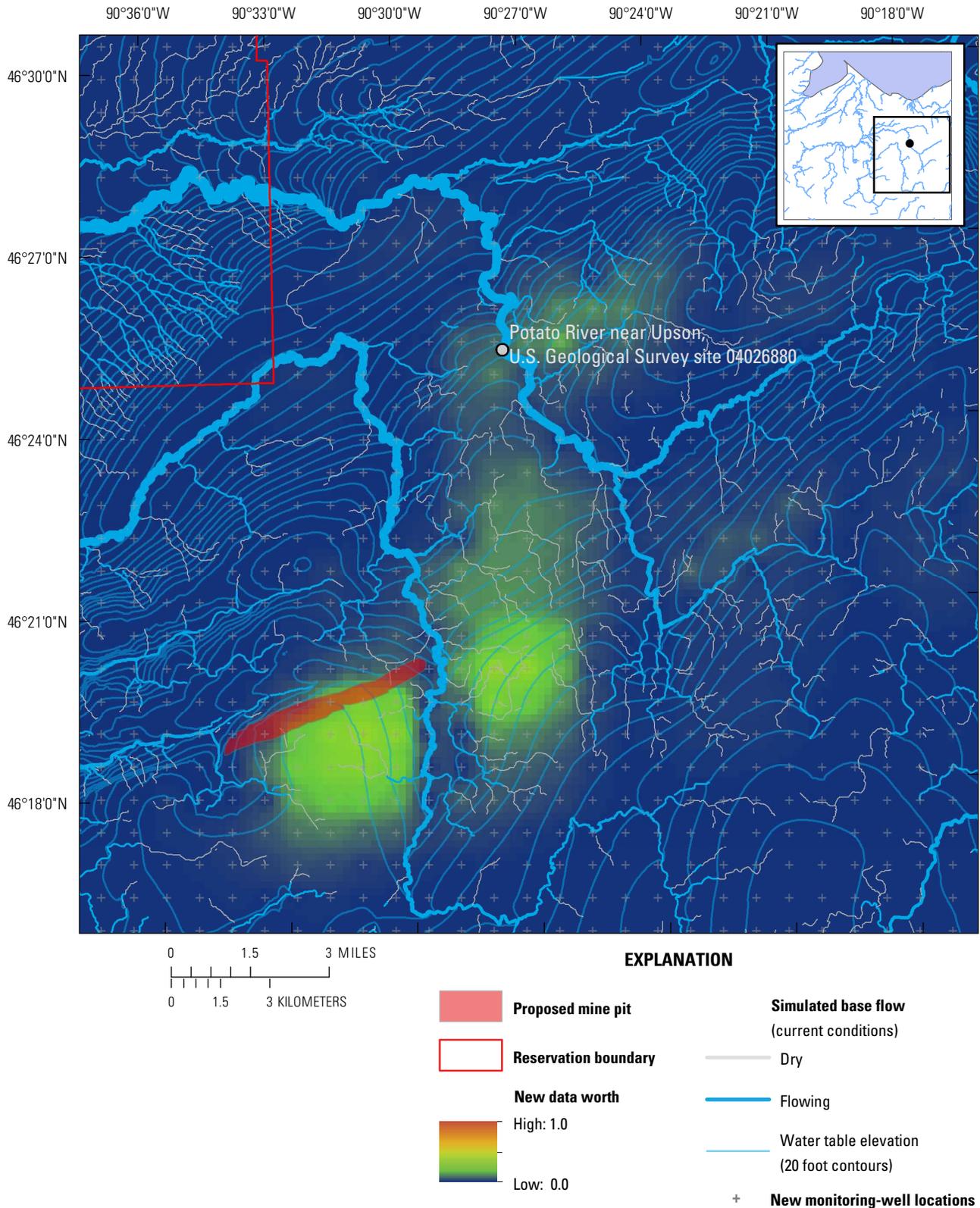
Figure 6.1.   Normalized reductions in prediction uncertainty from potential new head measurements, for predicting base flow in the Potato River near Upson, Wisconsin, at U.S. Geological Survey site 4026880, under the condition of dewatering pumping from the proposed open-pit mine described by Gogebic Taconite (2013).
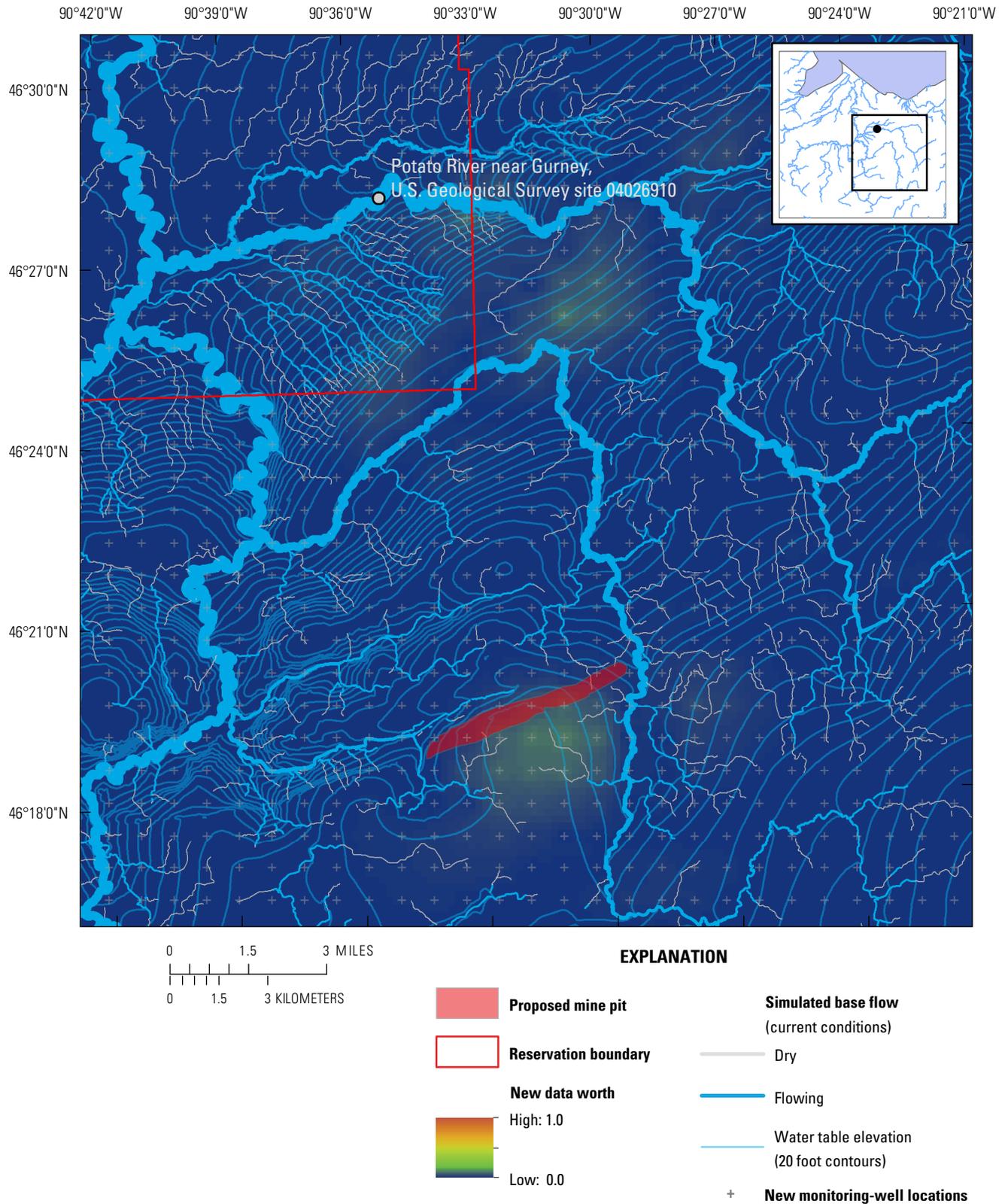
Figure 6.2.   Normalized reductions in prediction uncertainty from potential new head measurements, for predicting base flow in the Potato River near Gurney, Wisconsin, at U.S. Geological Survey site 4026910, under the condition of dewatering pumping from a proposed open-pit mine described by Gogebic Taconite (2013).
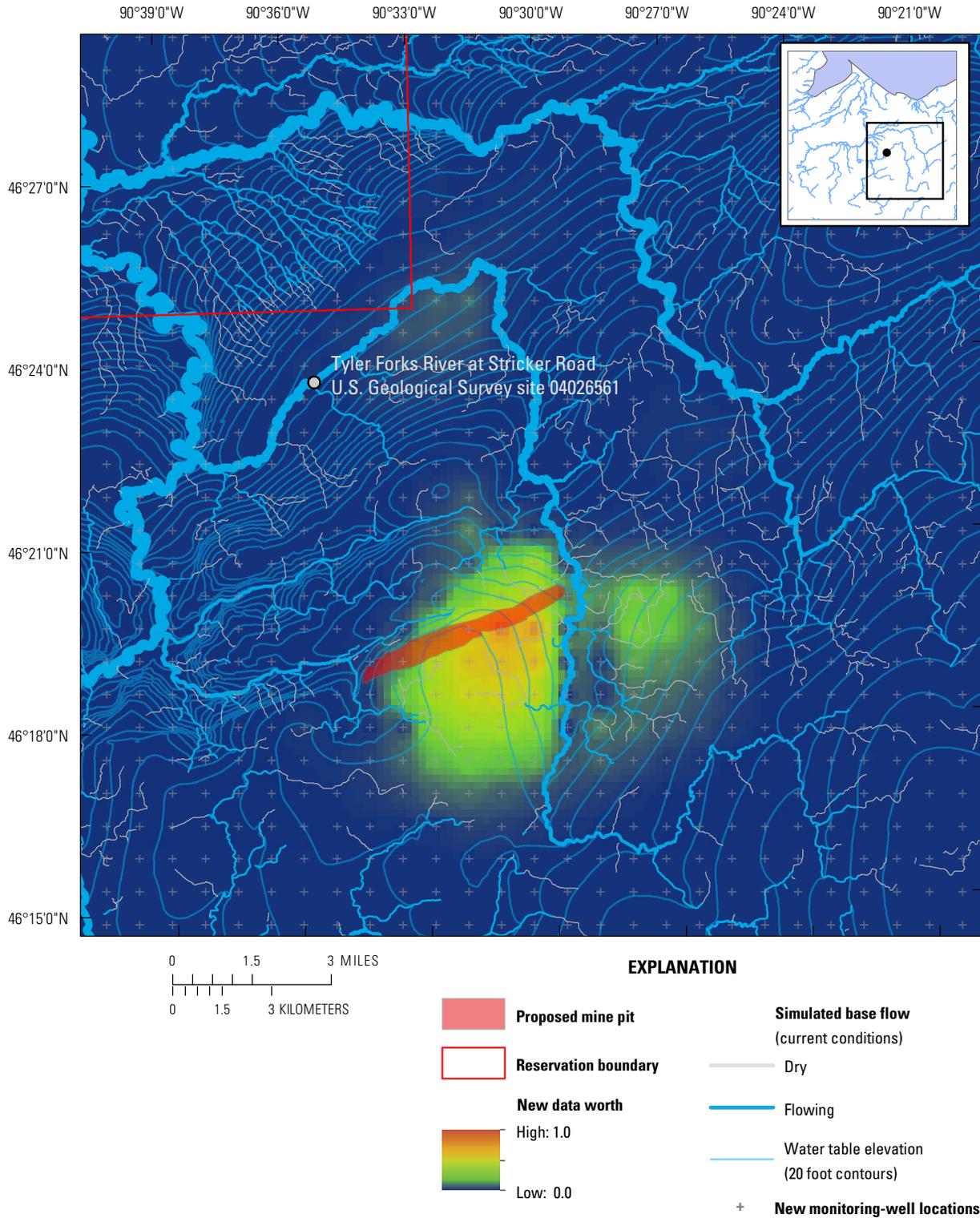
Figure 6.3.   Normalized reductions in prediction uncertainty from potential new head measurements, for predicting base flow in the Tyler Forks River at Stricker Road, at U.S. Geological Survey site 4026561, under the condition of large scale pumping from a proposed open-pit mine described by Gogebic Taconite (2013).