USGS
science for a changing world

National Water-Quality Assessment Program

Hydrochemical Regions of the Glacial Aquifer System, Northern United States, and Their Environmental and Water-Quality Characteristics

Scientific Investigations Report 2008–5015

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

Cover. Hydrochemical regions of the glacial aquifer system, Northern United States, described in this report.

Hydrochemical Regions of the Glacial Aquifer System, Northern United States, and Their Environmental and Water-Quality Characteristics

By Terri L. Arnold, Kelly L. Warner, George E. Groschen, James P. Caldwell, and Stephen J. Kalkhoff

National Water-Quality Assessment Program

Scientific Investigations Report 2008–5015

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

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2008

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:
Arnold, T.L., Warner, K.L., Groschen, G.E., Caldwell, J.P., and Kalkhoff, S.J., 2008, Hydrochemical regions of the glacial aquifer system, northern United States, and their environmental and water-quality characteristics: U.S. Geological Survey Scientific Investigations Report 2008–5015, 83 p. with appendixes.

Foreword

The U.S. Geological Survey (USGS) is committed to providing the Nation with credible scientific information that helps to enhance and protect the overall quality of life and that facilitates effective management of water, biological, energy, and mineral resources (*http://www.usgs.gov/*). Information on the Nation's water resources is critical to ensuring long-term availability of water that is safe for drinking and recreation and is suitable for industry, irrigation, and fish and wildlife. Population growth and increasing demands for water make the availability of that water, now measured in terms of quantity and quality, even more essential to the long-term sustainability of our communities and ecosystems.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program in 1991 to support national, regional, State, and local information needs and decisions related to water-quality management and policy (*http://water.usgs.gov/nawqa*). The NAWQA Program is designed to answer: What is the condition of our Nation's streams and ground water? How are conditions changing over time? How do natural features and human activities affect the quality of streams and ground water, and where are those effects most pronounced? By combining information on water chemistry, physical characteristics, stream habitat, and aquatic life, the NAWQA Program aims to provide science-based insights for current and emerging water issues and priorities. From 1991–2001, the NAWQA Program completed interdisciplinary assessments and established a baseline understanding of water-quality conditions in 51 of the Nation's river basins and aquifers, referred to as Study Units (*http://water.usgs.gov/nawqa/studyu.html*).

In the second decade of the Program (2001–2012), a major focus is on regional assessments of water-quality conditions and trends. These regional assessments are based on major river basins and principal aquifers, which encompass larger regions of the country than the Study Units. Regional assessments extend the findings in the Study Units by filling critical gaps in characterizing the quality of surface water and ground water, and by determining status and trends at sites that have been consistently monitored for more than a decade. In addition, the regional assessments continue to build an understanding of how natural features and human activities affect water quality. Many of the regional assessments employ modeling and other scientific tools, developed on the basis of data collected at individual sites, to help extend knowledge of water quality to unmonitored, yet comparable areas within the regions. The models thereby enhance the value of our existing data and our understanding of the hydrologic system. In addition, the models are useful in evaluating various resource-management scenarios and in predicting how our actions, such as reducing or managing nonpoint and point sources of contamination, land conversion, and altering flow and (or) pumping regimes, are likely to affect water conditions within a region.

Other activities planned during the second decade include continuing national syntheses of information on pesticides, volatile organic compounds (VOCs), nutrients, selected trace elements, and aquatic ecology; and continuing national topical studies on the fate of agricultural chemicals, effects of urbanization on stream ecosystems, bioaccumulation of mercury in stream ecosystems, effects of nutrient enrichment on stream ecosystems, and transport of contaminants to public-supply wells.

The USGS aims to disseminate credible, timely, and relevant science information to address practical and effective water-resource management and strategies that protect and restore water quality. We hope this NAWQA publication will provide you with insights and information to meet your needs, and will foster increased citizen awareness and involvement in the protection and restoration of our Nation's waters.

The USGS recognizes that a national assessment by a single program cannot address all water-resource issues of interest. External coordination at all levels is critical for cost-effective management, regulation, and conservation of our Nation's water resources. The NAWQA Program, therefore, depends on advice and information from other agencies—Federal, State, regional, interstate, Tribal, and local—as well as nongovernmental organizations, industry, academia, and other stakeholder groups. Your assistance and suggestions are greatly appreciated.

Matthew C. Larsen
Acting Associate Director for Water

Contents

Foreward iii
Abstract 1
Introduction 1
Purpose and Scope 3
Previous Studies 3
Conceptual Regional Framework 4
Description of the Study Area 4
Methods and Results of Data Analysis 6
Unsupervised Classification 8
Environmental Characteristics Related to Unsupervised Classification Results 11
Supervised Classification 12
Confidence in the Supervised Classification Results 14
Comparing Results of Supervised and Unsupervised Classifications 14
Classification of Hydrochemical Regions in Alaska 18
Hydrochemical Regions 18
Environmental Characteristics of Hydrochemical Regions 19
The Midwestern Agricultural Region 20
The Urban-Influenced Region 21
The Northern and Great Lakes Forested Region 22
The Western Agriculture and Grassland Region 22
The Mountain and Coastal Forested Region 24
Water-Quality Characteristics of Hydrochemical Regions 26
Selected Trace Elements and Nitrite Plus Nitrate in Hydrochemical Regions 28
Local- to Regional-Scale Comparisons of Water Quality 29
Summary 30
Acknowledgements 31
References 31
Glossary 35
Appendix A Tables 39
Appendix B Spatial Data Layers Used to Map Hydrochemical Regions 73

Figures

1. Map showing location of wells in the glacial aquifer system, northern United States 2
2. *A*, flowchart of framework layers, *B*, framework areas, and conceptual diagrams showing *C1*, valley-fill, *C2*, buried valley, and *C3*, layer or lens deposit 5
3. Map showing approximate location of glacial lobes, ice flow divides, and select glacial lakes approximately 14,000 years before present (2007), and direction of ice flow 6

4–6. Graphs showing—

4. Major- and minor-ion concentrations at the medoid and the median of the clusters, indicating water type of each cluster 9

5. Distribution of major- and minor-ion concentrations by cluster membership 10
6. Explanation of a box plot graph.....12
7. Diagram showing criteria for classifying water-quality characteristics of cluster groups (overall, 88-percent correct classification of clusters) 13
8. Graphs showing environmental characteristics calculated within 1,640 feet around the well that varied significantly by cluster membership 14
9. Diagram showing criteria for classifying environmental characteristics of cluster groups (overall, 70-percent correct classification of clusters) 15
10. Maps showing *A*, location and *B*, classification confidence, graphs showing *C*, mean and standard deviation of select chemical characteristics, and *D*, diagram showing generalized environmental and water-quality characteristics of the hydrochemical regions in the glacial aquifer system, northern United States.....16
11–16. Graphs Showing—
11. Distribution of major and minor ions in clusters and corresponding hydrochemical regions 19
12. Median measured major- and minor-ion concentrations, indicating general water type of each hydrochemical region 22
13. Distribution of major-ion concentrations by hydrochemical region in the glacial aquifer system, northern United States 23
14. Distribution of selected constituents by hydrochemical region in the glacial aquifer system, northern United States 25
15. Water type in the hydrochemical regions of the glacial aquifer system is dominated by calcium and bicarbonate ions; however, there are variations in the relative concentrations of other ions 27
16. Ratios of *A*, sodium to potassium, and *B*, calcium to magnesium, magnesium to sulfate, and sodium to chloride in hydrochemical regions 28

Appendix A Tables

A1. Reference, methods, variables, significant findings, and scale of selected previous studies to group, map, or define water quality 40
A2. Types of well networks used for this study, well depth, and casing diameter.....41
A3. Characteristics of data used in analyses for this report: number of samples with detections, estimated samples, and censored samples.....42
A4. Cluster medoid, number of wells, dissimilarity and separation measures, and percent poorly and well-classified wells.....43
A5. Results of Kruskal-Wallis rank sum test for differences among clusters.....44
A6. Regression tree for analysis of ion concentrations to characterize the water-quality cluster groups 45
A7. Spearman correlations of clusters and environmental characteristics within 1,640 feet of the well.....46
A8. Regression tree for analysis of environmental characteristics (within 1,640 feet of the well) to characterize the water-quality cluster groups 47
A9. The signature file used for maximum likelihood classification of hydrochemical regions in the glacial aquifer system 49
A10. Environmental characteristics used in supervised classification and their significance for classifying hydrochemical regions.....51
A11. Tabulation of clusters and hydrochemical regions to compare results of unsupervised and supervised classifications.....51

A12. Results of Kruskal-Wallis rank sum test for differences among hydrochemical regions 52
A13. Spearman correlation coefficients and probability results from z-test for equal proportions to compare the fit of hydrochemical regions to corresponding clusters for 12 different maximum likelihood classification models 54
A14. Environmental characteristics and general geographic location of each hydrochemical region 55
A15. Descriptive statistics by hydrochemical region 57
A16. Descriptive statistics for subsurface contact time and estimated ground-water recharge date for wells in each hydrochemical region 62
A17. Percent of total wells in each hydrochemical region by type of NAWQA sampling network 62
A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region 63
A19. Comparison of select water-quality characteristics as described by regional-scale hydrochemical region analysis, and a local-scale study in Michigan 72
A20. Comparison of select water-quality characteristics as described by regional-scale hydrochemical region analysis, and a local-scale study in Connecticut 72

Appendix B Spatial Data Layers Used to Map Hydrochemical Regions

Table

B1. Descriptions and sources of environmental spatial data layers used to map hydrochemical regions 74

Figures

B1–B9. Maps Showing—
B1. Average annual precipitation (1980–1997) across the glacial aquifer system in the northern United States 75
B2. Average annual temperature (1980–1997) across the glacial aquifer system in the northern United States 76
B3. Land-surface slope across the glacial aquifer system in the northern United States 77
B4. Land use (1990s) across the glacial aquifer system in the northern United States 78
B5. Vertical soil permeability in the glacial aquifer system in the northern United States 79
B6. Soil clay content in the glacial aquifer system in the northern United States 80
B7. Texture of surficial deposits in the glacial aquifer system in the northern United States 81
B8. Type of surficial deposits in the glacial aquifer system in the northern United States 82
B9. Estimated annual natural ground-water recharge in the glacial aquifer system in the northern United States 83

Conversion Factors and Datums

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
	Area	
acre	4,047	square meter (m^2)
acre	0.4047	hectare (ha)
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
inch per hour (in/h)	0 .0254	meter per hour (m/h)
inch per year (in/yr)	25.4	millimeter per year (mm/yr)

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

$$°C=(°F-32)/1.8$$

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

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

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25 °C).

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

Hydrochemical Regions of the Glacial Aquifer System, Northern United States, and Their Environmental and Water-Quality Characteristics

By Terri L. Arnold, Kelly L. Warner, George E. Groschen, James P. Caldwell, and Stephen J. Kalkhoff

Abstract

The glacial aquifer system in the United States is a large (953,000 square miles) regional aquifer system of heterogeneous composition. As described in this report, the glacial aquifer system includes all unconsolidated geologic material above bedrock that lies on or north of the line of maximum glacial advance within the United States. Examining ground-water quality on a regional scale indicates that variations in the concentrations of major and minor ions and some trace elements most likely are the result of natural variations in the geologic and physical environment. Study of the glacial aquifer system was designed around a regional framework based on the assumption that two primary characteristics of the aquifer system can affect water quality: intrinsic susceptibility (hydraulic properties) and vulnerability (geochemical properties). The hydrochemical regions described in this report were developed to identify and explain regional spatial variations in ground-water quality in the glacial aquifer system within the hypothetical framework context. Data analyzed for this study were collected from 1991 to 2003 at 1,716 wells open to the glacial aquifer system.

Cluster analysis was used to group wells with similar ground-water concentrations of calcium, chloride, fluoride, magnesium, potassium, sodium, sulfate, and bicarbonate into five unique groups. Maximum Likelihood Classification was used to make the extrapolation from clustered groups of wells, defined by points, to areas of similar water quality (hydrochemical regions) defined in a geospatial model. Spatial data that represented average annual precipitation, average annual temperature, land use, land-surface slope, vertical soil permeability, average soil clay content, texture of surficial deposits, type of surficial deposit, and potential for ground-water recharge were used in the Maximum Likelihood Classification to classify the areas so the characteristics of the hydrochemical regions would resemble the characteristics of the clusters. The result of the Maximum Likelihood Classification is a map showing five hydrochemical regions of the glacial aquifer system.

Statistical analysis of ion concentrations (calcium, chloride, fluoride, magnesium, sodium, potassium, sulfate, and bicarbonate) in samples collected from wells completed in the glacial aquifer system illustrates that variations in water quality can be explained, in part, by related environmental characteristics that control the movement of ground water through the aquifer system. A comparison of median concentrations of chemical constituents in ground water among the five hydrochemical regions indicates that ground water in the Midwestern Agricultural Region, the Urban-Influenced Region, and the Western Agriculture and Grassland Region has the highest concentrations of major and minor ions, whereas ground water in the Northern and Great Lakes Forested Region and the Mountain and Coastal Forested Region has the lowest concentrations of these ions. Median concentrations of barium, arsenic, lithium, boron, strontium, and nitrite plus nitrate as nitrogen also are significantly different among the hydrochemical regions.

Introduction

The glacial aquifer system is one of the principal aquifers in the United States studied by the National Water-Quality Assessment (NAWQA) Program of the U.S. Geological Survey (USGS) (Lapham and others, 2005). In this report, the term “glacial aquifer system” refers to all unconsolidated geologic material above bedrock that lies on or north of the line of maximum glacial advance within the United States (Warner and Arnold, 2005). The glacial aquifer system (fig. 1) is a large (953,000 square miles) regional aquifer system of heterogeneous composition. Many water-chemistry reactions are studied on a local scale; however, regional-scale studies can provide a broad understanding of general differences in water quality across an aquifer (Alley, 1993). If the glacial aquifer system is studied on a local scale, it is a complex and irregular series of layered coarse- and fine-grained glacial deposits. The complexity of the system must be simplified to understand the system and the quality of ground water on a regional scale. On a regional scale, the glacial aquifer system is, in general

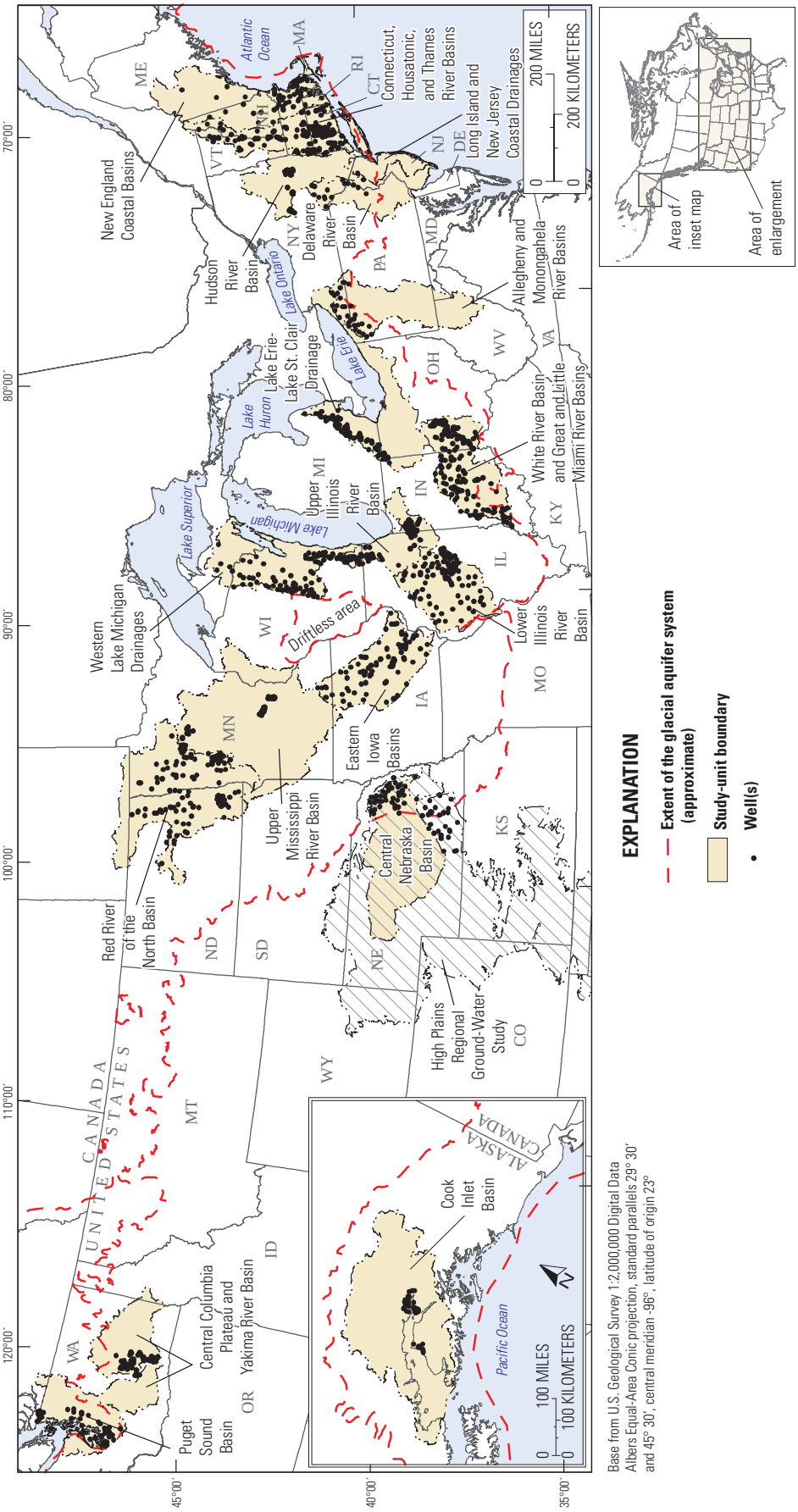


Figure 1. Location of wells in the glacial aquifer system, northern United States.

terms, ‘similar’ across its spatial extent because of the overall heterogeneity of the glacial deposits. Identifying water-quality variations within a regional context allows visualization of the broad geochemical and environmental mechanisms at work in the glacial aquifer system and facilitates an understanding of the system as a whole.

Subtle differences in geology and physical environmental characteristics, including human land use, across such a large expanse as the northern United States can affect water chemistry in this aquifer system. Examining regional-scale variations in ground-water quality reveals that the variations in major- and minor-ion and trace-element concentrations commonly are the result of natural differences in the geologic and physical environment that affect rock-water interactions and weathering processes; however, natural water types may be modified by the addition of anthropogenic contaminants.

Concentrations of major and minor ions in ground water typically are indicative of geological materials and other factors encountered along a flow path and can be used to describe broad types of water. Relative concentrations of specific anions and cations are important in how they chemically react with other constituents. By identifying combined variations in major- and minor-ion concentrations and identifying related environmental characteristics to aid in explaining those variations, extrapolations can be used to map areas where water quality is similar. In this report, these areas of similar water quality are referred to as hydrochemical regions. The hydrochemical regions described in this report were delineated on the basis of ground-water-quality data collected from the glacial aquifer system.

Purpose and Scope

The purpose of this report is to describe the statistical and spatial-analysis methods used to identify and map hydrochemical regions in the glacial aquifer system in the northern United States, and to describe the spatial variations in environmental and water-quality characteristics across the glacial aquifer system based on mapped hydrochemical regions. The hydrochemical regions described in this report were developed to identify and explain regional spatial variations in ground-water quality in the glacial aquifer system.

The analysis described in this report is based on data collected from 1,716 wells in 19 NAWQA study units across the glacial aquifer system in the northern United States (the study area) from 1991 to 2003 (fig. 1). Only data collected as part of the NAWQA program were used to map hydrochemical regions of the glacial aquifer system. This analysis is based on the initial conceptual framework that was established to assess water quality in the glacial aquifer system and expands on the explanatory aspect of that framework (Warner and Arnold, 2005). The hydrochemical regions of the glacial aquifer system presented in this paper divide the system on the basis of physical environmental and measured chemical characteristics.

Previous Studies

A number of ground-water classification systems based on physical environmental characteristics such as physiography or aquifer lithology have been proposed (table A1—all tables are in Appendix A, at the back of this report) during the past 100 years (Chapelle, 2004). One of the first ground-water classifications was that by Meinzer (1923), which classified 21 ground-water provinces based on rock units and physiography across the United States. Thomas (1952) refined Meinzer’s provinces and reduced them to 10 regions by eliminating areas in which differences in ground-water chemistry were minor. Heath (1984) expanded the geographic scope to include nonconterminous areas of the United States (Alaska, Hawaii, and Puerto Rico) and characterized ground-water regions using additional aquifer properties and conditions such as mineral composition of rocks, structure of water-bearing openings, arrangement of recharge and discharge areas, and permeability, all of which affect the occurrence and availability of ground water (table A1).

Regional-scale studies to classify water resources, including ground water, have focused on physical characteristics of the entire aquifer system including bedrock aquifers (Heath, 1984), or on factors predominantly useful for surface-water quality (Wolock and others, 2004). There have been various efforts to group the entire United States into regions with similar environmental characteristics for various purposes (Wolock, 2003a; Wolock and others, 2004; Heath, 1984; Heath, 1988); most of these efforts are based only on physical characteristics of the area (table A1).

Classifications based on changes in water chemistry along a flow path were proposed by Cederstrom (1946) and Chebotarev (1955). Cederstrom classified ground water in the Coastal Plain of Virginia; Chebotarev had a broad classification of natural waters in the upper part of the lithosphere that covered both surface and ground water. These and related classifications laid the ground work for understanding the most important water-rock interactions that ultimately determine the chemistry of ground water. Back (1966) first proposed the concept of “hydrochemical facies” for the Atlantic Coastal Plain ground-water system. Back examined the variations and evolution of water-chemistry types from calcium-bicarbonate-dominated freshwater to sodium-chloride-dominated saline water and identified the most important water-rock interactions that affect water chemistry in a Coastal Plain aquifer. Back’s work has application to most other aquifers because the water-rock interactions he described are common.

More recently, ground-water regions based primarily or exclusively on statistical classification systems using multivariate techniques have been developed. Many of these statistically driven studies concluded that geologic source material affects ground-water chemistry (Güler and Thyne, 2004; Güler and others, 2002; Pacheco, 1999; Stetzenbach and others, 2001; and Stetzenbach and others, 1999). Most of these studies have been conducted on a local or small regional scale (table A1).

Most classification schemes were designed to identify large patterns in water-chemistry changes in specific aquifers. The utility of any regional ground-water or aquifer-system classification scheme is determined by how well it organizes disparate regional ground-water data, creates a framework that enhances the transferability of results, and identifies regional patterns in water quality. The conceptual glacial framework design (fig. 2a) of Warner and Arnold (2005) was a preliminary classification scheme to identify areas of the glacial aquifer system that may have different water-quality conditions.

Conceptual Regional Framework

Study of the glacial aquifer system initially was designed around a conceptual regional framework that provided a mechanism for identifying the most general and important characteristics of the glacial aquifer system as a whole (fig. 2a). The hypothesis for the framework was that two primary characteristics affect water quality in the glacial aquifer system: intrinsic susceptibility and vulnerability (Warner and Arnold, 2005). Intrinsic susceptibility is related to the sources of water and to the hydraulic properties affecting the flow of water into and through the aquifer system (Focazio and others, 2002). Vulnerability depends, in part, on the intrinsic aquifer susceptibility and is a function of the geochemical properties of the geologic materials composing the aquifer system and the characteristics and availability of any potential sources of contaminants (Focazio and others, 2002).

Many studies at various scales have been done on glacial deposits to determine differences in susceptibility based on matrix composition and texture (Hitt and Nolan, 2005; Arnold and Friedel, 2000; Berg and others, 1984; Freeze and Cherry, 1979). On a regional scale, intrinsic susceptibility is represented in the glacial-aquifer system framework by the spatial distribution of fine- or coarse-grained material at the land surface, and the physical setting of the aquifer system (Warner and Arnold, 2005). The physical setting refers to the structure or geometry of the aquifer as either a “layer or lens” deposit or as a “buried valley” deposit (fig. 2).

Natural vulnerability is a function of both intrinsic susceptibility and the proximity and characteristics of the sources of the chemical constituents, such as the mineralogy of the aquifer material or the geochemical conditions within the aquifer system. In this report, chemical constituents in ground water that potentially could have been added to the natural ground-water system from weathering and breakdown of glacial deposits are called “natural contaminants.” Weathering and chemical breakdown can mobilize many constituents from the aquifer matrix that will affect the natural ground-water chemistry, but the presence of a chemical constituent in the aquifer matrix is not always directly related to its presence in ground water. For the conceptual framework, natural vulnerability is represented by regional-scale features such as physiography and general direction of travel of the glacial ice lobes (fig. 3). The direction of travel of the glacial ice lobes was used to indicate possibly different geologic source materials. On the basis of these regional-scale features, the glaciated area in the United States was divided into four areas—East, Central, West-Central, and West (fig. 2) (Warner and Arnold, 2005). It was hypothesized that these four framework areas would have different intrinsic susceptibility and natural vulnerability. Vulnerability to anthropogenic contaminants could be assessed by adding land use to the preliminary framework. Land use can be indicative of locations and types of anthropogenic contaminants to which the aquifer may be vulnerable. Land use was not included in the preliminary conceptual framework because assessing vulnerability to anthropogenic contaminants was not the purpose of the preliminary framework.

Analysis and mapping of hydrochemical regions presented in this report include land use as an indicator of anthropogenic vulnerability. The hydrochemical regions are based on the two following conceptual framework premises: hydraulic and geochemical properties of the aquifer system vary across the spatial extent of the aquifer system, and variations in these properties (and in land use) affect water quality in the glacial aquifer system. The goal of the analysis used to map the hydrochemical regions was to verify differences in water quality across the glacial aquifer system, identify specific combinations of environmental characteristics related to those water-quality differences, and extrapolate measured water-quality characteristics from areas where data were collected to areas with similar hydrochemistry where no data were collected.

Description of the Study Area

The study area includes parts of 26 states in the northern United States that were at one time partly or completely covered by continental glaciers (fig. 1). The glacial history of the study area is complex and includes repeated episodes of ice advance (glaciation) and retreat (deglaciation). During the Pleistocene Epoch, which began approximately 1.8 million years before present (2007) and ended about 11,000 years ago (U.S. Geological Survey Geologic Names Committee, 2007), two large continental ice-sheet complexes, the Laurentide ice sheet and the Cordilleran ice sheet, episodically covered parts of Canada, the northern conterminous United States, and Alaska. The Laurentide ice sheet (fig. 2b) originated in the Hudson Bay area of Canada. At its maximum extent, Laurentide ice extended as far south as the Missouri and Ohio Rivers and covered most of the north-central and northeastern United States east of the Rocky Mountains (Mickelson and others, 1983). The Cordilleran ice sheet (fig.2b) advanced southward from source areas in British Columbia and terminated in the northwestern United States between the Continental Divide and the Pacific Ocean (Waitt and Thorson, 1983). During episodes of maximum glacial advance, the eastern margin of the Cordilleran ice connected with the western margin of Laurentide ice near the continental divide (Porter and others, 1983).

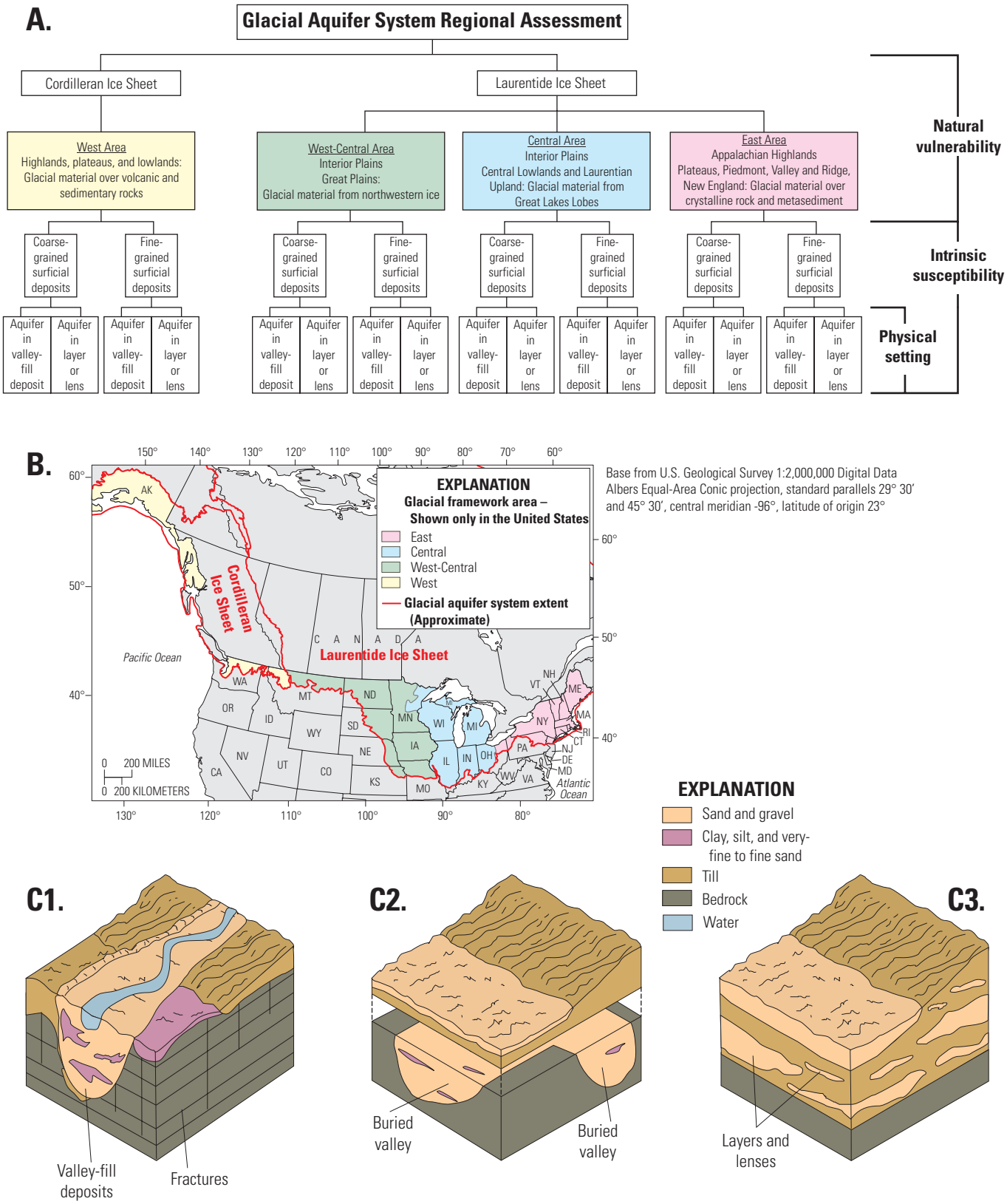
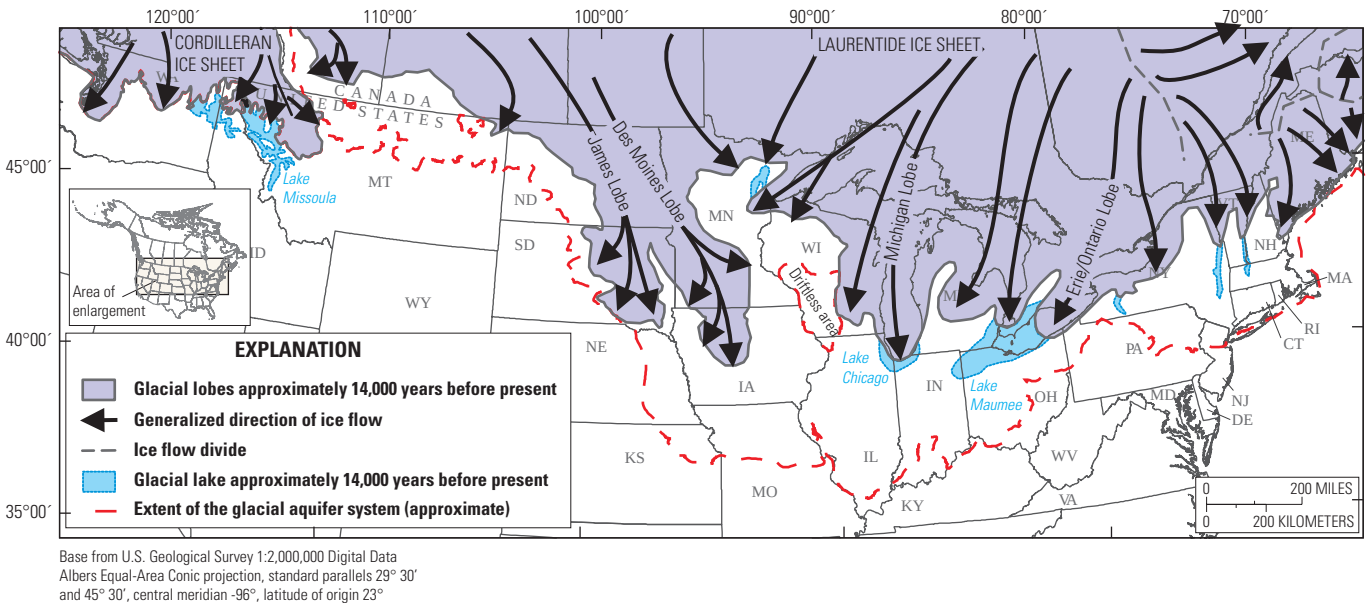


Figure 2. *A,* flowchart of framework layers (modified from Warner and Arnold, 2005), *B,* framework areas (modified from Warner and Arnold, 2005), and conceptual diagrams showing *C1,* valley-fill, *C2,* buried valley, and *C3,* layer or lens deposit.



Base from U.S. Geological Survey 1:2,000,000 Digital Data
Albers Equal-Area Conic projection, standard parallels 29° 30′ and 45° 30′, central meridian -96°, latitude of origin 23°

Figure 3. Approximate location of glacial lobes, ice flow divides, and select glacial lakes approximately 14,000 years before present (2007), and direction of ice flow (modified from Geological Survey of Canada, 1989).

In Alaska, Cordilleran ice covered piedmont and valley areas within the coastal ranges (Clague, 1989).

Episodes of Pleistocene glacial advances were interrupted by sudden climate shifts to warmer and dryer periods (interglacial stages), which lasted from several thousands to many tens of thousands of years. During interglacial stages, the continental ice sheets shrank in size or completely vanished. Periodic deglaciation produced terminal and ground moraines; extensive meltwater deposits as large accumulations of outwash; and other deposits of stratified and unstratified unconsolidated sediments. During interglacial and post-glacial periods, wind reworked fine-grained surficial glacial sediment and redeposited it as sand dunes and extensive blankets of loess (Porter and others, 1983). The geologic materials deposited during the numerous glacial and deglacial stages are heterogeneous and have complex distributions controlled largely by physiography and glacial history (Clague, 1989).

The environmental setting of the glacial aquifer system reflects its natural history as well as human effects. Land-use practices have become an integral part of the environment by modifying the land surface, potentially affecting ground-water recharge, and by providing potential sources of contamination. No single environmental characteristic can account for the variability of water quality across the glacial aquifer system because all of the environmental characteristics are related.

Methods and Results of Data Analysis

Only data collected as part of the NAWQA program of the USGS were used to map the hydrochemical regions of the glacial aquifer system. The NAWQA program is a primary source of long-term, nationwide water-quality data from ground water and surface water (*http://water.usgs.gov/nawqa/*). Data collected as part of the NAWQA program typically are based on surface-water basins called “study units”. The data analyzed for the glacial aquifer system study described here were collected from 1991 to 2003 at 1,716 wells in 19 NAWQA study units across the system (fig. 1). Data were collected following standard USGS protocols (Lapham and others, 1995; Koterba and others, 1995; Koterba, 1998; and Squillace and Price, 1996). In the NAWQA program, wells are grouped in networks designed to facilitate analysis of data that were collected for a specific purpose. Two primary types of such networks are designed to facilitate analysis of land-use effects on shallow ground water or analysis of drinking-water quality. Wells in networks designed for analysis of land-use effects on water quality typically are shallow to allow sampling of recently recharged ground water that may exhibit chemical characteristics indicative of the surrounding land use. Wells in networks designed for analysis of drinking-water quality generally are deeper than land-use-network wells and are assumed to yield water that is representative of the primary aquifer in the study unit area. Because the glacial aquifer system comprises irregular layers of fine- and coarse-grained deposits, wells in both types of networks may be either unconfined or confined. Generally, ground water pumped from shallow wells in land-use networks may have shorter flow paths than does ground water from deeper wells in drinking-water networks. In addition to the two major network types described above, other networks are designed for special ground-water studies, such as those of ground-water flow paths.

Data from all available well networks were included in the analysis and results of the study described in this report. In

this study, 574 wells were in agricultural row-crop or orchard land-use networks; 55 wells were in forest land-use networks; 322 wells were in urban, residential, and/or commercial land-use networks; 550 were in drinking-water networks; 211 wells were in flow-path or other special studies networks; and 4 wells were in reference networks that were presumed to be minimally affected by human activities and land-use practices (table A2).

Several well networks lie at least partly outside of the glacial boundary as defined in this report, but are included in this analysis—two networks in the Central Columbia and Yakima area of Washington, two networks in the Puget Sound area of Washington and Canada, and one network in the High Plains area of Nebraska (fig. 1). The networks in the Central Columbia and Yakima area of Washington are completely outside the glacial boundary (75 wells); however, the wells in these networks are completed either in glaciofluvial materials that were distributed by a catastrophic flood when glacial lake Missoula broke through an ice dam or in wind-blown loess (Sandra Embrey, U.S. Geological Survey, oral commun., 2003). The networks in the Puget Sound area of Washington and Canada fall either inside (80 wells in Washington) or outside (13 wells in Canada) of the glacial boundary and the wells are completed in glaciofluvial, coarse sand and gravel deposited during the Fraser Glaciation (Tesoriero and others, 2001; Tesoriero and others, 2000). The network in Nebraska consists of wells that fall either inside (28 wells) or outside (17 wells) of the glacial boundary; however, these wells are completed in deposits composed of reworked glacial material (Sharon Qi, U.S. Geological Survey, oral commun., 2003). For the characterization of hydrochemical regions of the glacial aquifer system, the chemical character of the samples collected from the wells in these five networks is assumed to be similar to the chemical character of ground water in other parts of the glacial aquifer system; therefore, data from these wells are included in this analysis.

Sometimes chemical constituents may be present in ground water at concentrations that are less than the laboratory equipment's ability to detect them. Data resulting from the laboratory's analysis of these very low concentrations are referred to as "censored". The concentration below which the laboratory cannot detect the constituent is called the "censored level". Data for some constituents were censored at multiple levels because laboratory analysis methods have changed with time. There are many different methods for handling censored data for use in statistical analysis. For the statistical analysis described in this report, the data for individual constituents were handled using a uniform method of simple substitution. Substitution means that a quantitative value is substituted for the censored value. Substitution is not always the best method for handling censored data because it may slightly skew the reported descriptive statistics (Helsel and Hirsch, 1995); however, for the statistical analysis used for this report, substitution yields adequate results in that the results of the analysis depend on relative, not exact, concentrations.

For constituents that had data censored at multiple levels, the most common censored level was identified. For example, if there were five ground-water samples where a constituent was censored at 0.5 milligrams per liter (mg/L) and one sample where the same constituent was censored at 0.2 mg/L, the most common censored level for that constituent was 0.5 mg/L. Data that were censored at the most common censored level were treated as quantitative values (the same as a measured concentration). Data that were censored above the most common censored level for a given constituent were rare, and they were discarded from the data set. Data reported as an estimated value above the most common censored level were treated as quantitative values (the same as a measured concentration). The value of the most common censored level for each constituent was substituted for any data where measured, censored, or estimated data were below the most common censored value; these substituted values were treated as a quantitative value (the same as a measured concentration). For all constituents, the quantitative values and the substituted values were used in the analyses for this report including: descriptive statistics, correlations, supervised and unsupervised classifications, regressions, and graphical representations. The number of measured, censored, and estimated data for the constituents used in this report, along with the most frequent censored levels and substituted values are shown in table A3.

Spearman rank correlation coefficients (Helsel and Hirsch, 1995) were calculated to evaluate pair-wise relations between select constituents and/or environmental characteristics. All correlations stated in this report are significant at the 95 percent confidence level (p-value < 0.05) or higher. Although some correlation coefficients are low values, they are significant because of the large sample sizes.

Kruskal-Wallis Rank Sum tests (Helsel and Hirsch, 1995) were used to evaluate differences among groups. Compared groups were either clusters or hydrochemical regions and are noted in appropriate tables, figures, and text. Differences were evaluated based on the mean of the ranks for the various groups. The null hypothesis was that the difference in the means of the group ranks was equal to zero. If the probability (p-value) resulting from the test was less than or equal to the alpha level (0.05), indicating the evaluated groups were different at a 95 percent confidence level, then the null hypothesis was rejected.

Bicarbonate concentrations used for this report were calculated (not measured). If the ground-water sample was analyzed for bicarbonate (1,058 samples), then the resulting measurement was used. If the ground-water sample was analyzed for alkalinity (398 samples), then the bicarbonate concentration was calculated from alkalinity by dividing alkalinity (reported in mg/L as calcium carbonate) by 0.8202 (Hem, 1992). If neither bicarbonate nor alkalinity were reported (261 samples) then bicarbonate concentration for those samples was not used in further analysis for this report.

Dissolved solids concentrations used for this report also were calculated (Fishman and Friedman, 1989). Dissolved solids concentration was calculated by summing concen-

trations of the following: calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), carbonate (CO_3), chloride (Cl), nitrate (NO_2 plus NO_3 as N), sulfate (SO_4), and silica (SiO_2). If the sample was analyzed for carbonate (836 samples), then the resulting measurement was used; otherwise, carbonate concentration was calculated by multiplying alkalinity by 0.60 (Fishman and Friedman, 1989) where an alkalinity measurement was available (616 samples).

Subsurface contact time (of water with geologic materials) is used in this report to interpret how ion concentrations may vary among hydrochemical regions but was not directly used in analysis and mapping of hydrochemical regions. An index of subsurface contact time was computed by Wolock and others (1997) from surface-water basin topography and soil hydraulic conductivity. The source for the topography data was the HYDRO1K Elevation Derivative Database (Verdin and Greenlee, 1996). The soil characteristics (Wolock, 1997) were derived from the U.S. Department of Agriculture's State Soil Geographic (STATSGO) database. For use in this report, subsurface contact time as determined by Wolock and others (1997) was summarized as an average within 1,640 feet (ft) of a well.

The analysis of this data set was a three-step process. The first step was to verify and identify any variations in water quality across the glacial aquifer system. Water-quality variations were identified by using an unsupervised classification method (cluster analysis) to group wells with similar water-quality characteristics based on major- and minor-ion concentrations. Unsupervised classification guarantees that the groups of wells (clusters) are as distinct as possible with respect to major- and minor-ion concentrations. All wells with enough major- and minor-ion data (1,316 wells) were used in the unsupervised classification. The second step was to identify environmental characteristics related to the variability of water quality that was identified using unsupervised classification. These environmental characteristics are mapped for the glacial aquifer system (excluding Alaska) and were used to extrapolate water-quality characteristics of ground-water from areas where ground-water samples were collected to areas of the glacial aquifer system where no ground-water samples were collected. The third step was to perform a supervised classification of environmental characteristics using Maximum Likelihood Classification (MLC) spatial analysis. In this classification, the range of possible environmental characteristics (result of step 2) was defined for each cluster group (result of step 1). The supervised classification resulted in an extrapolation of characteristics defined at points into hydrochemical regions.

Unsupervised Classification

A water type describes the chemical composition (characteristics) of a water sample based on the relative concentrations of major cations and anions (Hiscock, 2005). The ions sodium (Na), calcium (Ca), magnesium (Mg), chloride (Cl), bicarbonate (HCO_3), and sulfate (SO_4) generally make up about 90 percent of the dissolved ions in ground water (Hiscock, 2005). These major ions, along with the minor ions potassium (K) and fluoride (F), commonly are used to classify water types. Because major and minor ions can be used to classify different water types, the ion concentrations were used to identify water-quality variations across the glacial aquifer system.

Cluster analysis was used to group wells from which samples contained similar concentrations of Ca, Cl, F, Mg, K, Na, SO_4, and HCO_3. Only 1,316 of the wells in the glacial aquifer system (77 percent) had all of the water-quality data needed for the analysis. There are many different clustering algorithms and each algorithm produces slightly different results. The algorithm used for this analysis was clustering by medoids (centers of the clusters) using Euclidean distance as a measure of dissimilarity among and within groups. A medoid is the data point closest to the center of the cluster in multidimensional space. The medoid may be different from the median of values (fig. 4). The Euclidean distance is the root sum-of-squares of differences for all of the variables for each well in the data set. Cluster analysis is a method of unsupervised classification that can be used as a descriptive tool, a data simplification tool, or a relation identification tool (Rees, 2001; Hair and others, 1998). For this report, cluster analysis was used for all three purposes: to examine relations among many constituents (data simplification and relation identification), and to describe variations in water quality across a broad regional area (description).

An unsupervised classification method, such as cluster analysis, provides a means of grouping wells based on the actual chemical concentration data (variables) rather than on pre-defined group characteristics. In this study, the grouped objects are wells and the variables used to form the groups are major- and minor-ion concentrations (Ca, Cl, F, Mg, K, Na, SO_4, and HCO_3). The goal is to identify groups of wells that are more similar to each other, with respect to ion concentrations, than they are to wells in any other group; in other words, the goal is minimizing within group differences while maximizing between group differences (Massart and Kaufman, 1983; Romesburg, 1984; Everitt, 1980; Kachigan, 1986; Hair and others, 1998).

Within- and between-group differences are measured by the Euclidean distance between points (distance in multidimensional space). In the case of this analysis, the "multidimensional space" has eight dimensions, one for each variable used to define the clusters. Because distance measures are sensitive to variations in measurement scales and magnitudes among the variables being clustered, the variables were standardized by subtracting the mean and dividing by the standard deviation for each variable. This type of standardization (z-score) has a mean of zero (0) and standard deviation of one (1) which allows the variables to be examined on a similar scale (Hair and others, 1998).

Five groups of wells were judged to be the best cluster solution judging from the silhouette width, average and maxi-

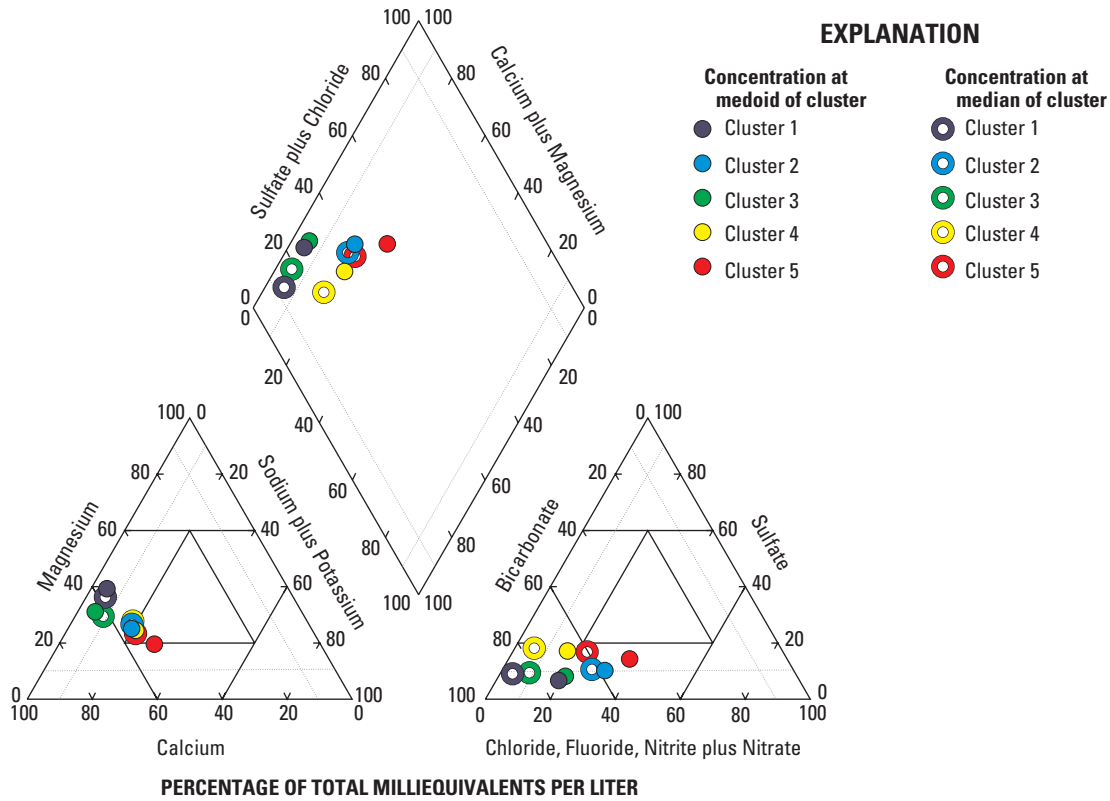


Figure 4. Major- and minor-ion concentrations at the medoid and the median of the clusters, indicating water type of each cluster.

mum dissimilarity among groups, and the number of poorly classified wells. The silhouette width measures the "goodness" of the classification and is calculated using the following equation:

$$\text{Silhouette width} = (B\text{-}A)/(\text{the greater of either } A \text{ or } B)$$

where

B is the minimum dissimilarity of an object to those objects in other clusters, and

A is the average dissimilarity of the object in its own cluster.

Silhouette widths greater than zero indicate well-classified objects, whereas silhouette widths less than zero indicate poorly classified objects, and silhouette widths near zero indicate an object's classification lies between clusters. Silhouette widths closer to one are better classified than widths closer to zero. According to the overall average silhouette width, 89 percent of the wells were well classified and 11 percent were poorly classified. The clusters generally have low individual average silhouette widths ranging from 0.08 (cluster 1) to 0.24 (cluster 2) as shown in table A4. The low silhouette width of cluster 1 indicates that some of the wells assigned to cluster 1 (21 percent) may have been similar to wells assigned to a different cluster. Cluster 3 had the highest percentage of well-classified wells (100 percent) followed by cluster 2 (97 percent); cluster 1 had the lowest percentage of well-classified wells (79 percent) (table A4).

Clustering results are highly subjective and depend on which variables and how many groups are specified for the analysis. Clustering algorithms tend to form roughly round-shaped groups of approximately equal size. Cluster analysis always forms clusters (groups), regardless of whether or not there is a natural structure in the data (Everitt, 1980; Massart and Kaufman, 1983; Romesburg, 1984). Resulting clusters are contingent on the variables selected to compare the objects and can change substantially if the variables change.

As a result of this cluster analysis, each cluster contains between 237 and 286 wells (table A4). The clusters generally overlap in some characteristics and are completely different in others. The overlapping clusters indicate that there is not a strong structure in the data; however, most clusters have statistically significant differences (p-value < 0.05) in the constituent concentrations among the five different groups (figs. 5, 6; table A5). Pacheco (1999) found a similar result, and hypothesized that his clusters overlapped because of highly similar water chemistries among the water samples being analyzed. The water types of the glacial aquifer system are overwhelmingly Ca-Mg-HCO_3, with minor variations in the other ions, and this is reflected in the overlapping clusters. The results of this unsupervised classification indicated that there is variability in water quality across the glacial aquifer system. Although the clusters overlap in multidimensional space, the cluster-

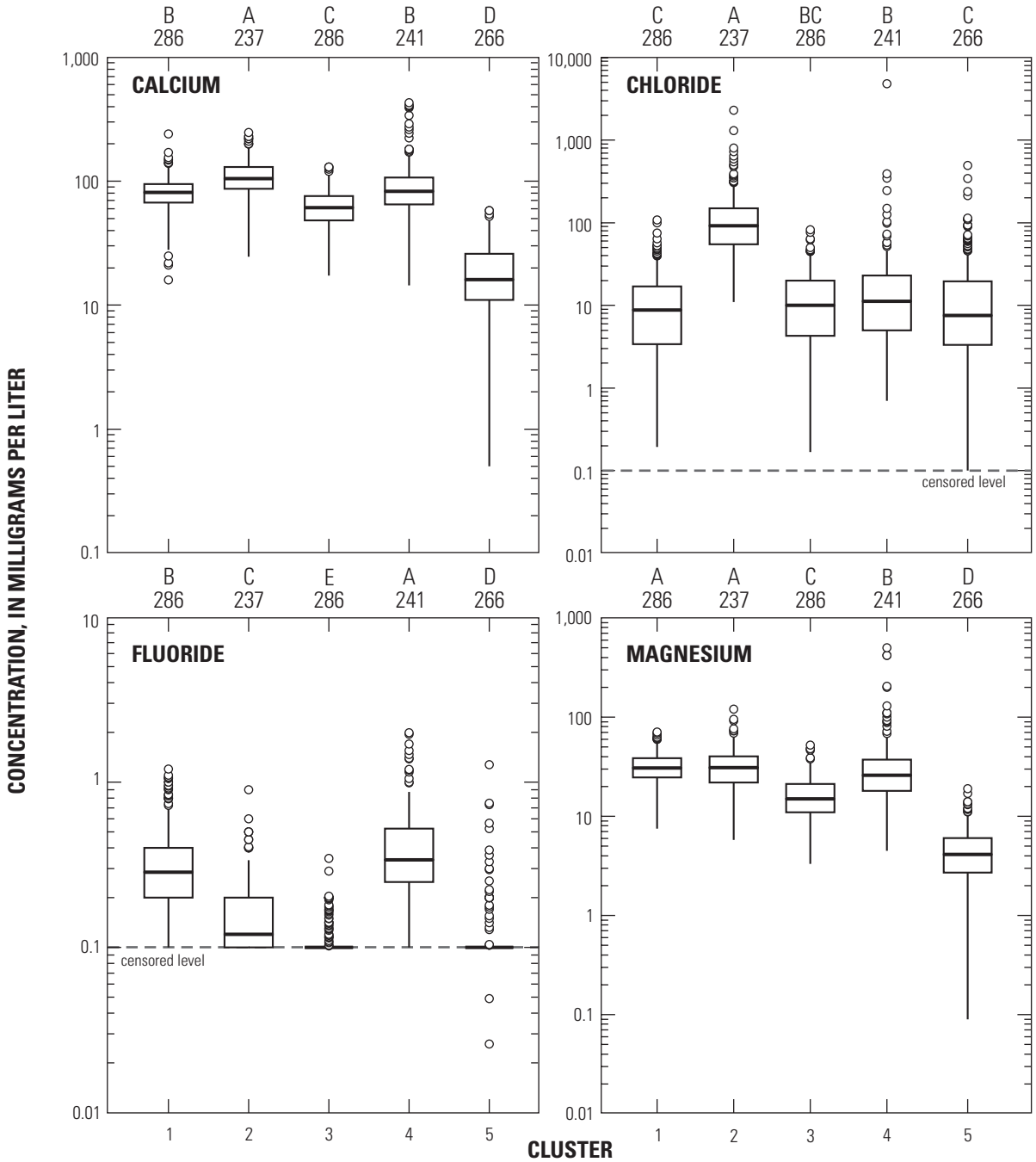


Figure 5. Distribution of major- and minor-ion concentrations by cluster membership. Refer to figure 6 for an explanation of box plot graphs.

ing results typically formed significantly different groups (table A5) and the analysis progressed to the next step.

Regression Tree Analysis (also called Decision Tree Analysis) was used to evaluate basic water-quality characteristics of the cluster groups. Regression Tree Analysis commonly is used for examining characteristics of categorical data instead of logistic regression (Hair and others, 1998). For this report, Regression Tree Analysis was used to find divisions in the ion concentrations that yield the desired classification of the cluster groups (fig. 7; table A6). The output is a decision flow chart (dendrogram) that shows sequential splits of the major- and minor-ion concentrations that end in the cluster group classifications (fig. 7). When unlimited, the regression tree resulted in 33 combinations of ion concentrations to classify the wells into the five clusters. This unlimited regression tree had a misclassification error rate of 0.09, which means that 91 percent of the wells were classified correctly into their corresponding cluster based on the decision rules of the regression analysis. To reduce the complexity of the regression tree, the analysis was limited to 18 ion concentration combinations

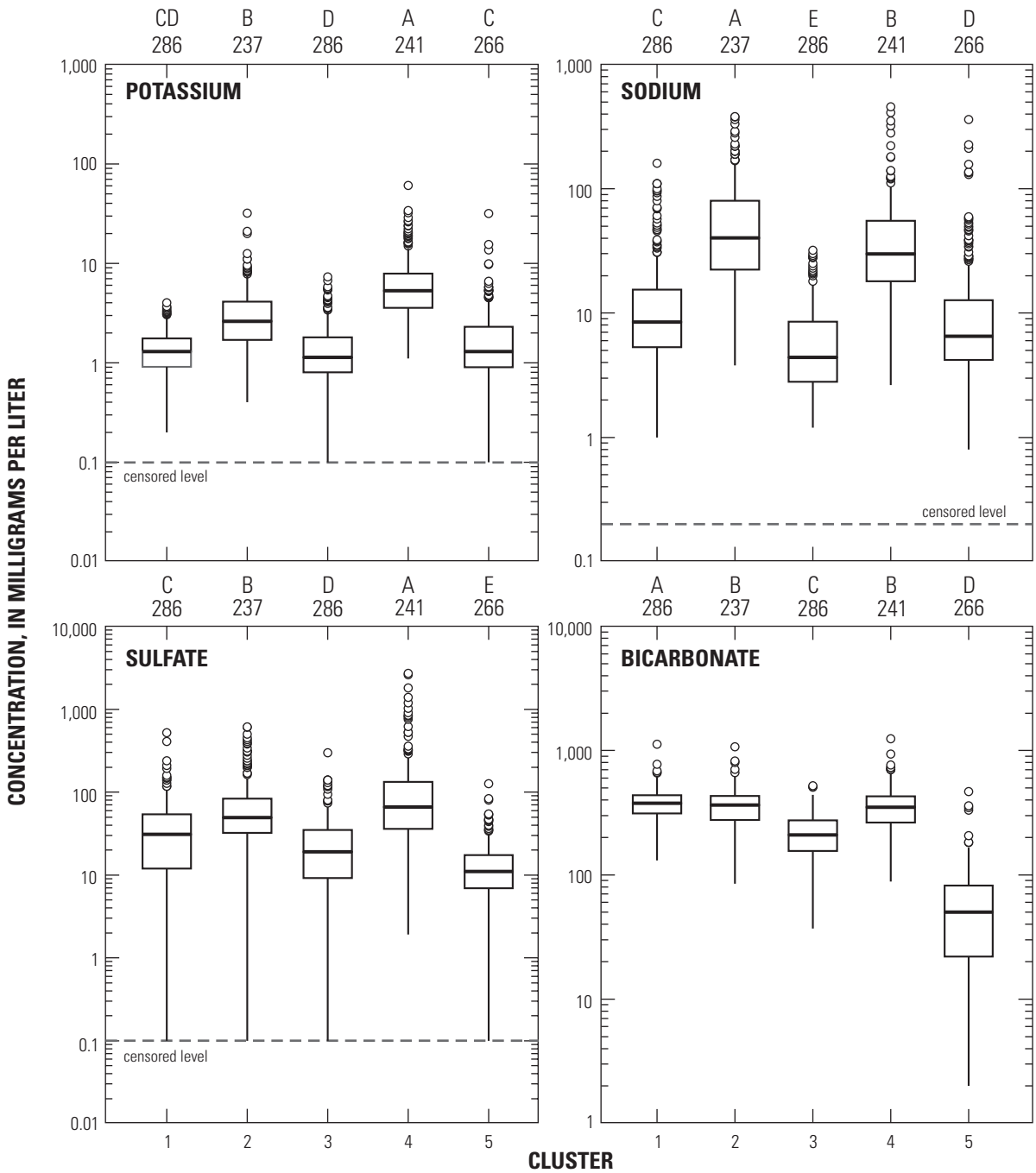


Figure 5 Distribution of major- and minor-ion concentrations by cluster membership. Refer to figure 6 for an explanation of box plot graphs.—Continued

while minimizing the misclassification rate as much as possible. The misclassification error rate for this limited regression tree increased only 3 percent to 0.12, which means 88 percent of the wells were classified correctly (fig. 7; table A6).

In Regression Tree Analysis, the same variables are not equally important for classifying all possible outcomes of the dependant variable (Hair and others, 1998). The Regression Tree Analysis indicates which variables are important for differentiating clusters from each other. Calcium concentration was important only for differentiating between clusters 5 and 3. Similarly, sulfate was important only for differentiating between clusters 1 and 4 (fig. 7; table A6).

Environmental Characteristics Related to Unsupervised Classification Results

Environmental characteristics (variables) that describe soil properties (including permeability, texture, and erod-

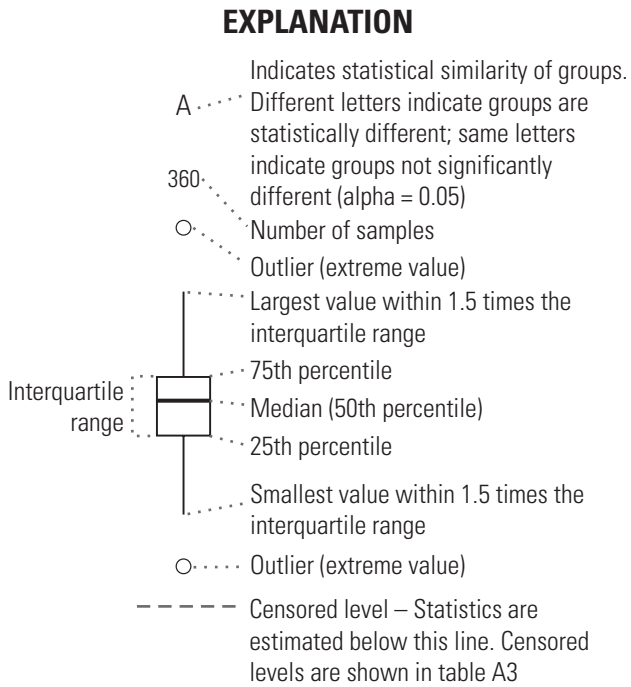


Figure 6. Explanation of a box plot graph. Number of samples and indication of similarity among groups may not always be included in the graph.

ibility), land use, land-surface slope, climate (temperature and precipitation), texture of surficial deposits, and modeled recharge/runoff conditions were calculated for the area within 1,640 ft of sampled wells. All of these variables were numeric and on a continuous scale. Spearman rank correlation coefficients (r) were calculated for each variable pair to identify environmental characteristics significantly correlated (p-value < 0.05) with the cluster groups (table A7). For the purpose of calculating correlations, cluster groups were indicated by five binary variables, one variable for each cluster. The binary variables used a “1” to indicate a well was in the cluster or a ”0” to indicate a well was not in the cluster. Each of these correlated variables was plotted by cluster on box plots to visually determine which variables had the most variation among clusters. The variables percent land-surface slope; percent forest, agriculture, and urban land use; average annual precipitation and temperature; percent clay in soil; percent fine-grained surficial material; vertical soil permeability; and estimated ground-water recharge appeared to vary the most between clusters (fig. 8). In addition, a categorical variable describing type of surficial deposit was evaluated for differences among clusters by tabulating the number of clustered wells with each type of surficial deposit. These 11 environmental characteristics were retained for further analysis.

The 11 environmental characteristics then were used in Regression Tree Analysis to see how well they might classify the wells into their corresponding water-quality clusters. In an unlimited regression tree, 77 combinations of environmental characteristics were identified to classify the wells into clusters. In this unlimited regression tree, the 11 environmental characteristics yielded a 74-percent correct classification of wells. When the regression tree was limited to 28 possible combinations, the fine-grained surficial materials characteristic was eliminated from the tree, and the correct classification rate dropped to 70 percent (table A8; fig. 9). The percentage of correctly classified wells indicated that these environmental characteristics could be used to predict membership in each cluster. The fine-grained surficial material characteristic was retained for later analysis because it was identified as an important variable when the regression tree was unlimited.

After determining that the environmental characteristics within 1,640 ft around a well could be used to predict membership in each cluster, these characteristics were used as spatial data layers in a Geographic Information System (GIS) to extrapolate the clusters and delineate hydrochemical regions. The data layers used in the GIS were not restricted to the area within 1,640 ft around each well; instead, they were spatially continuous layers where a point representing the well was overlaid on the layer to identify the environmental characteristic at the point. The spatial data layers used were: 1) precipitation, 2) temperature, 3) land use, 4) land-surface slope, 5) vertical permeability, 6) clay in soil, 7) texture and 8) type of surficial deposits, and 9) ground-water recharge (Appendix B, figs. B1–B9; table B1). All of the spatial layers are raster data resampled to a 1,640-ft cell size. The Digital Elevation Model and type of surficial deposits were the only layers of environmental data that were available for the entire glaciated area (Alaska and the Conterminous United States). Further analysis for Alaska based on the selected environmental characteristics was not possible because not all of the spatial data layers were available. In addition, some spatial data were not available for at least one of the environmental layers for small areas of the conterminous United States so that these areas could not be classified into hydrochemical regions; this lack of data affected the hydrochemical region classification for 10 wells in the Puget Sound area.

Supervised Classification

Maximum Likelihood Classification (MLC) was used to make the extrapolation from clusters defined by points to areas of hydrochemical regions. Maximum Likelihood Classification is a geostatistical method of supervised classification, commonly used in remote sensing (Rees, 2001), in which characteristics of the clusters are known in some locations before the analysis begins. This set of known characteristics and locations is called a “training sample” and is used to define a range of characteristic values for each cluster. The MLC algorithm classifies spatial areas based on the maximum probability of an area being similar to a particular cluster based on the training sample.

As a result of the previously described unsupervised classification, 1,316 wells were assigned to clusters. Each cluster represented a different hydrochemical region into which the glacial aquifer system would be classified. For the

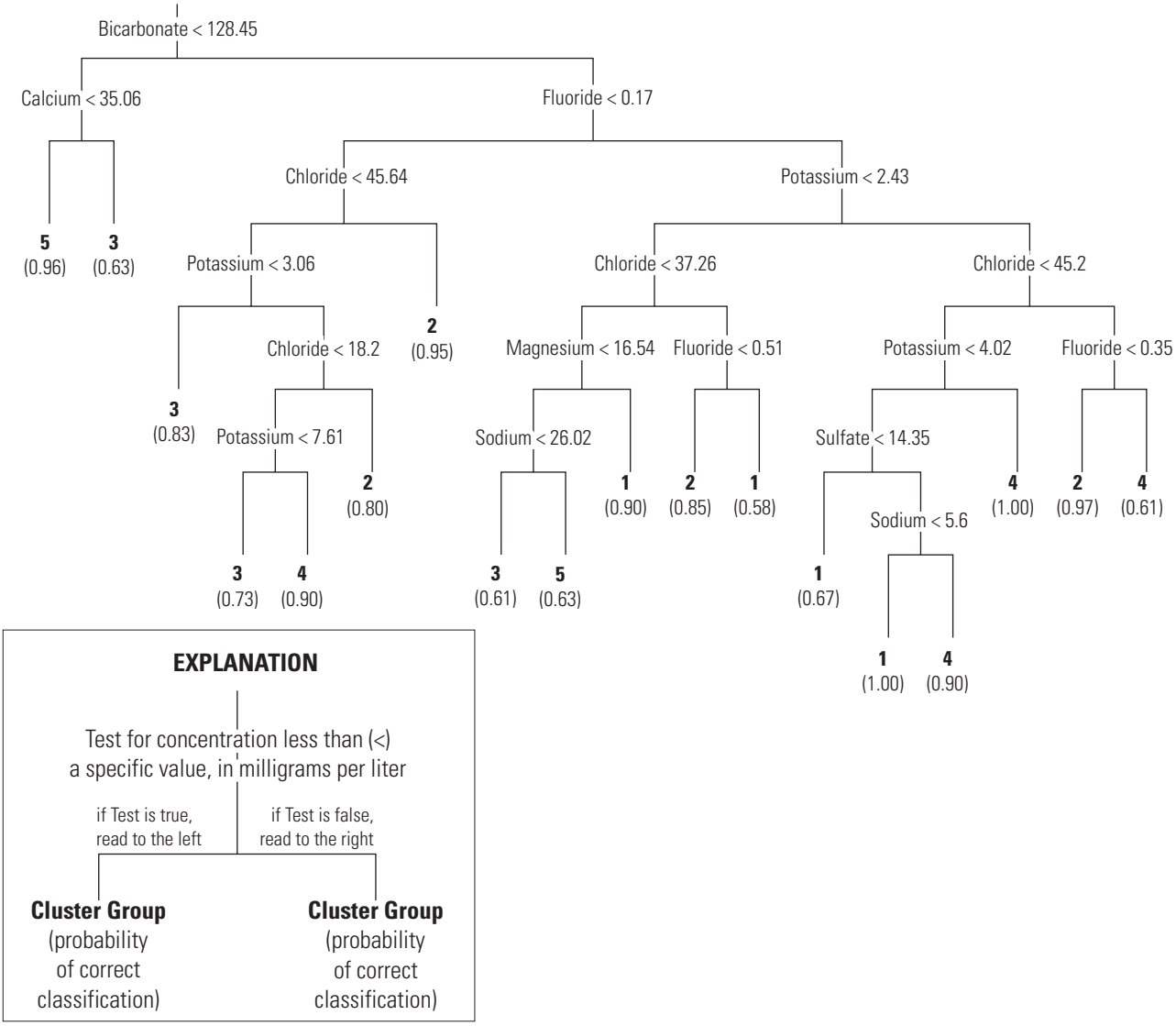


Figure 7. Criteria for classifying water-quality characteristics of cluster groups (overall, 88-percent correct classification of clusters). See Table A6.

analysis described in this report, the environmental characteristics were spatially combined in the GIS. Environmental characteristics of each cluster were identified when the wells were spatially overlaid with the combined layers of environmental characteristics, creating the training sample. Multivariate statistics describing environmental characteristics for the training sample were stored in a text file called a signature or index file (table A9) that the MLC then used to classify the glacial aquifer system into hydrochemical regions. The hydrochemical regions have the highest probability of matching the environmental characteristics of the corresponding cluster groups (the training sample). The result of the MLC is a map showing five hydrochemical regions of the glacial aquifer system (listed in the order of corresponding cluster groups 1 through 5): (1) Midwestern Agricultural Region (MAR), (2) Urban-Influenced Region (UIR), (3) Northern and Great Lakes Forested Region (NGLFR), (4) Western Agriculture and Grassland Region (WAGR), and (5) Mountain and Coastal Forested Region (MCFR) (fig. 10). The hydrochemical region for 1,660 wells was identified when the wells were spatially overlaid with the hydrochemical regions map. This spatial overlay allowed an extrapolation of information to 400 wells not originally included in the classifications because the wells lacked spatial and/or chemical data. The hydrochemical region could not be identified for 56 of the 1,716 wells in the glacial aquifer system because these wells overlaid areas that could not be classified by the MLC. Thirty-three wells were in Alaska, and 23 wells were in the Puget Sound area (10 wells in Washington and 13 wells in Canada).

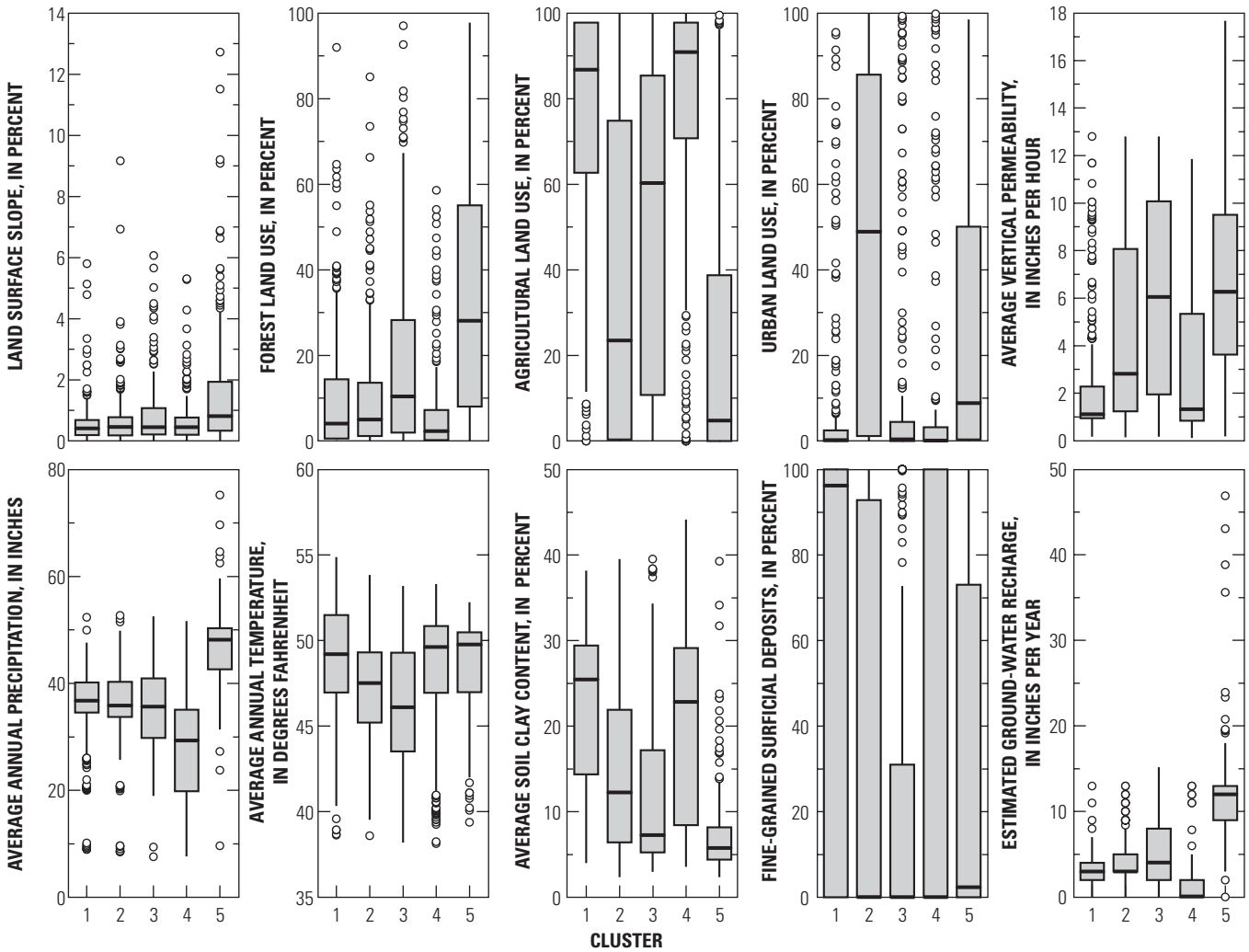


Figure 8. Environmental characteristics calculated within 1,640 feet around the well that varied significantly by cluster membership. Refer to figure 6 for an explanation of box plot graphs.

Confidence in the Supervised Classification Results

The delineation of hydrochemical regions of an aquifer system, as defined in this report, is appropriate for use only in large regional-scale studies. Confidence in the water-quality characteristics that are extrapolated to unsampled areas of the hydrochemical regions is necessarily low; in contrast, confidence generally is high where data were actually collected (fig. 10). A confidence map was made along with the hydrochemical regions map as an optional output from the MLC analysis. The confidence map shows locations where hydrochemical region classifications have a higher or lower probability of being correct based on the environmental characteristics at the location, as compared to characteristics stored in the signature file for the training sample.

The environmental characteristics were tested for significance (at the 95-percent confidence level) as linear predictors of hydrochemical region using logistic regression. Results of the logistic regression analysis indicated that all analyzed characteristics were important for predicting membership in the regions (table A10). Cumulatively, the nine environmental characteristics used in the MLC describe how much water can enter the ground-water system and provide opportunities for water to remain in contact with soil and rock material. This is consistent with the conclusions of Güler and Thyne (2004), Güler and others (2002), Pacheco (1999), and Stetzenbach and others (2001).

Comparing Results of Supervised and Unsupervised Classifications

The goal of the supervised classification was to identify hydrochemical regions of the glacial aquifer system that had unique and distinguishable water-quality characteristics. Hydrochemical regions were classified based on the probability of matching environmental characteristics identified at the 1,316 wells used for the unsupervised classification. Data from

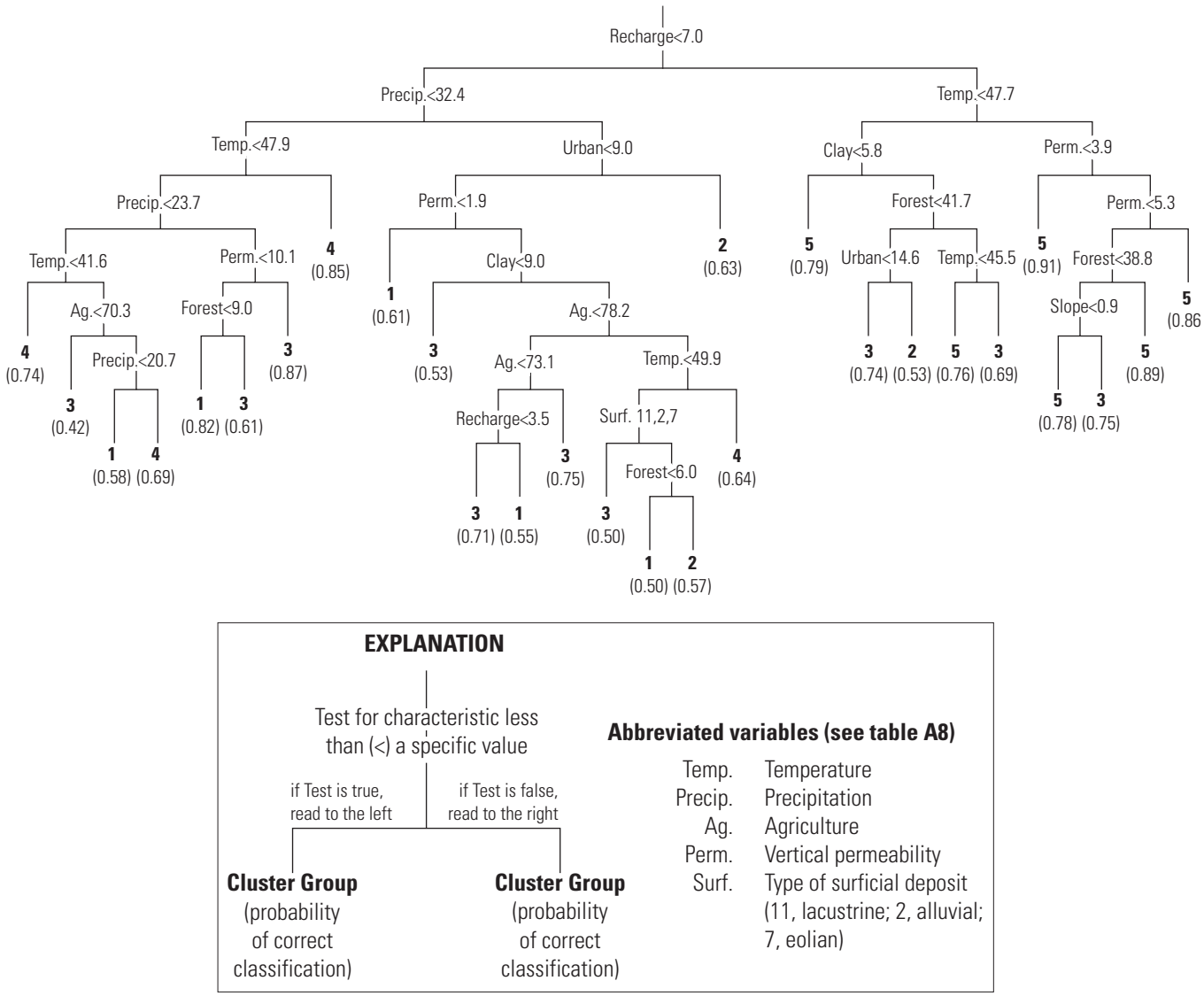


Figure 9. Criteria for classifying environmental characteristics of cluster groups (overall, 70-percent correct classification of clusters) See table A8.

1,260 wells were used in both the unsupervised and supervised classifications (table A11). Because the hydrochemical regions were classified based on probability, the results of the unsupervised and supervised classifications may not correspond exactly. For example, 18 percent of the wells assigned to cluster 5 spatially overlaid the NGLFR, the hydrochemical region corresponding to cluster 3 (table A11). However, the highest percentage of wells in each cluster was correctly classified as the corresponding hydrochemical region.

At least three comparisons between the results of the unsupervised classification (cluster analysis) and the supervised classification (MLC analysis) indicate that the hydrochemical regions mapped by the MLC adequately reflect areas of unique and distinguishable water-quality characteristics as defined by ion chemistry (clusters). The first comparison indicates that the concentrations of Ca, Cl, F, Mg, Na, K, SO_4, and HCO_3 vary significantly among the hydrochemical regions, just as the concentrations varied significantly among clusters. The second comparison indicates that the water-quality characteristics of the hydrochemical regions closely correspond to those of the cluster groups. The third comparison indicates that the wells assigned to hydrochemical regions closely correspond to the appropriate clusters. Hydrochemical regions were identified for all of the 1,660 wells that spatially overlay the regions map, and data from those wells were used for the three comparisons.

For the first comparison, a series of Kruskal-Wallis rank sum tests were used to determine if the concentrations of Ca, Cl, F, Mg, Na, K, SO_4, and HCO_3 varied significantly among the hydrochemical regions. Test results indicated that concentrations of major and minor ions varied significantly (95-percent confidence) between regions (table A12), and observed

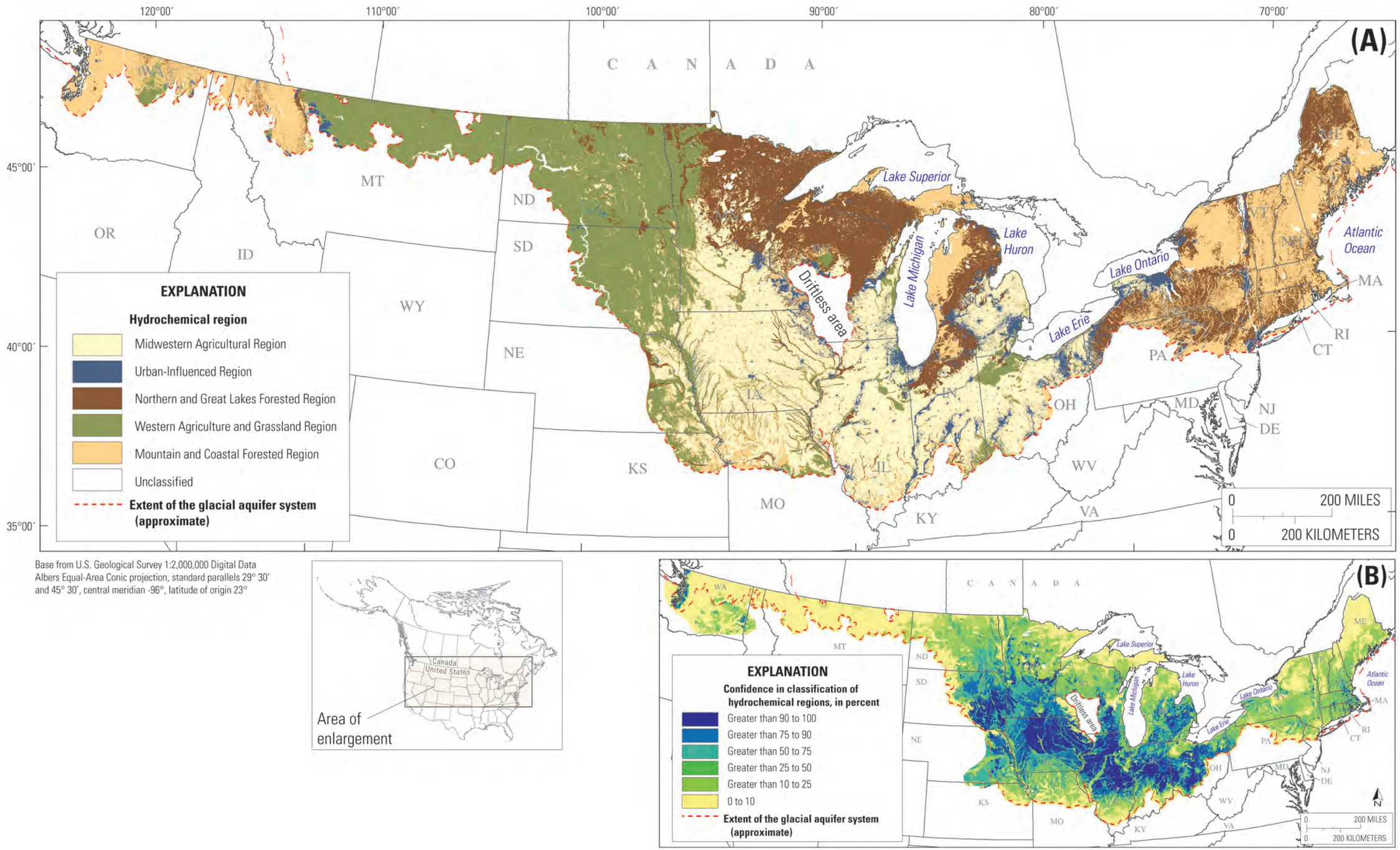
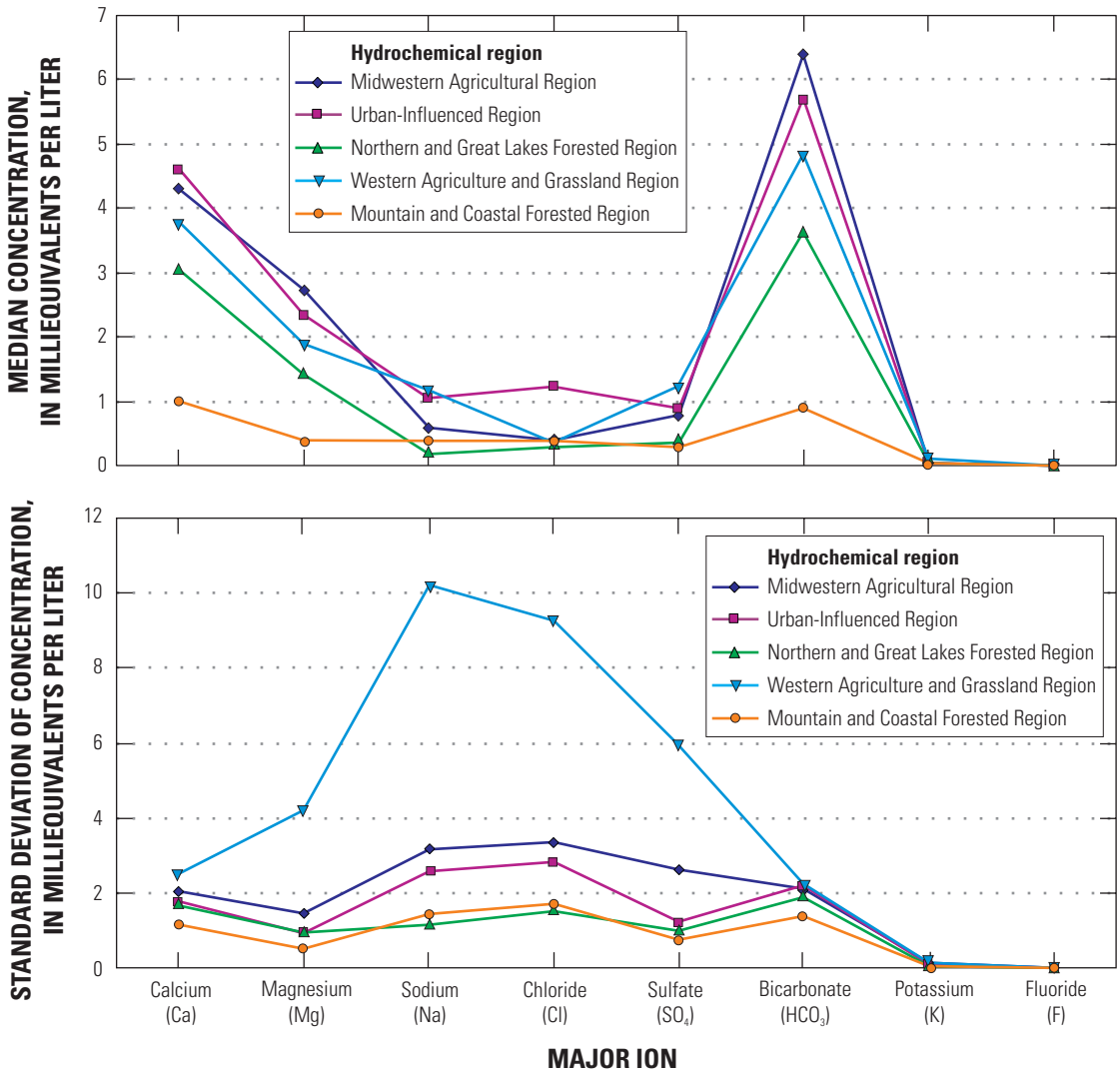


Figure 10. *A,* Location, *B,* classification confidence, *C,* median and standard deviation of select chemical characteristics, and *D,* generalized environmental and water-quality characteristics of the hydrochemical regions in the glacial aquifer system, northern United States.

(C)



(D)

Hydrochemical Region	Generalized Environmental Characteristics		Generalized Water-Quality Characteristics	
Urban-Influenced Region	Mostly urban areas with population greater than 70,000; low permeability	Long Contact Times	High in Major and Minor Ions	Highest median Ca, Cl; Second highest median Mg, K, Na, SO_4, HCO_3
Midwestern Agricultural Region	Warm and humid, tile-drained, clayey agricultural areas; low permeability; low recharge			Highest median Mg, HCO_3; Second highest median Ca, Cl, F
Western Agriculture and Grassland Region	Warm and dry, clayey, irrigated agricultural and grassland areas; low permeability; very low recharge			Highest median F, K, Na, So_4; Second lowest median Cl
Northern and Great Lakes Forested Region	Sandy, cool, forested, areas with high organic content; high permeability	Short Contact Times	Low in Major and Minor Ions	Lowest median Cl, F, Na; Second lowest median Ca, Mg, K, So_4, HCO_3
Mountain and Coastal Forested Region	Sandy and wet mountainous and coastal forested areas; high permeability; high recharge			Lowest median Ca, F, Mg, K, SO_4, HCO_3; Second lowest median Na

Figure 10. *A,* Location, *B,* classification confidence, *C,* median and standard deviation of select chemical characteristics, and *D,* generalized environmental and water-quality characteristics of the hydrochemical regions in the glacial aquifer system, northern United States.—Continued

differences in concentrations were more distinct among hydrochemical regions than among the cluster groups (table A5). Distributions of Cl, K, and Na concentrations were similar in some regions, but generally the concentration distributions were unique among hydrochemical regions (table A12), which implies that the delineation adequately describes regions of unique water-quality characteristics.

For the second comparison, statistical distributions of ion concentrations, grouped by cluster and hydrochemical region, were plotted as box plots (fig. 11). Compared visually, the statistical distributions of ions in the cluster groups mostly were similar to those in the hydrochemical regions; however, Kruskal-Wallis Rank Sum tests indicated differences in some distributions. Because the distributions of the ion concentrations in each hydrochemical region are more distinct than those in each cluster, differences between the cluster and hydrochemical region major ion concentrations are acceptable.

For the third comparison, proportions of wells classified in each hydrochemical region and in each cluster were calculated. If the hydrochemical regions adequately represent the clusters, then the hydrochemical regions and clusters should have similar proportions of wells assigned to them. A z-test for equal proportions (Ott, 1993) was used to compare the proportion of samples assigned to each hydrochemical region to the proportion of wells assigned to each cluster (table A13). The z-test results indicate that all except one of the hydrochemical regions compare favorably to the clusters. Cluster 1 was the least well-defined cluster and the hydrochemical region corresponding to cluster 1 reflects this in a significant z-score (95-percent confidence). The significant z-score indicates that cluster 1 and the corresponding hydrochemical region have different proportions of wells assigned to them. In an effort to improve the correspondence of the supervised and unsupervised classification results for cluster 1, different combinations of the nine environmental layers were tested in alternative Maximum Likelihood Classifications (table A13). The alternative MLC trials did not improve the correspondence. Additionally, Spearman correlation coefficients indicate that wells assigned to the individual hydrochemical regions by the original MLC in which all nine environmental layers were used are significantly correlated to those wells assigned to corresponding clusters (table A13); therefore, the original MLC was confirmed to have the best correspondence to the unsupervised classification.

Considering all three of the comparisons between results of the unsupervised classification (cluster analysis) and the supervised classification (MLC analysis), the hydrochemical regions mapped by the MLC adequately reflect areas of unique and distinguishable water-quality characteristics as defined by major ion chemistry. A detailed description of each hydrochemical region in the glacial aquifer system is given in the following sections.

Classification of Hydrochemical Regions in Alaska

Ground-water-quality data from Alaska were included in the unsupervised classification, but they were not included in the supervised classification because of the lack of spatial environmental data for Alaska. When the unsupervised classification was repeated without the Alaska data, the results were not significantly changed, which indicates that data from Alaska did not bias the results of the unsupervised classification. It was determined, as previously discussed, that the results of the supervised classification adequately represented the results of the unsupervised classification. Unsupervised classification placed Alaska wells into clusters 2 (1 well), 3 (17 wells), and 5 (15 wells), which correspond to hydrochemical regions UIR, NGLFR, and MCFR, respectively. Alaska wells were in areas of forest (12 wells), agriculture (5 wells), wetland (3 wells), and urban (13 wells) land use (Glass, 2001). The environmental characteristics of Alaska include lowlands and rugged mountains (indicating a wide range of slope); 15 inches (in.) average annual precipitation; 42 degrees Fahrenheit (°F) average annual temperature; glacial, glacio-estuarine, and alluvial surficial deposits; clay, silt, sand, and boulders (implying a mix of textures); and urban, agriculture, forest, and wetland land use (Glass, 2001). These environmental characteristics are similar to the environmental characteristics of the UIR, NGLFR, and MCFR discussed later in this report.

Hydrochemical Regions

Five hydrochemical regions were identified for the glacial aquifer system: Midwestern Agricultural Region (MAR), Urban-Influenced Region (UIR), Northern and Great Lakes Forested Region (NGLFR), Western Agriculture and Grassland Region (WAGR), and Mountain and Coastal Forested Region (MCFR) (fig. 10). These five hydrochemical regions have distinct water-quality characteristics based on major- and minor-ion concentrations (figs. 10, 12) and have environmental characteristics that fit into the conceptual regional framework used to design the glacial aquifer study (fig. 2). The hydrochemical regions expand on the conceptual framework by including land use as an indicator of potential anthropogenic contaminant sources. Spatial distributions of the hydrochemical regions do not exactly match the distribution of framework areas, which are based on natural vulnerability.

The two primary differences between the hydrochemical regions (fig. 10) and the framework areas (fig. 2) are that the East and West coast areas appear to be more similar than were anticipated by the framework (likely because of similar forested land use), and the central framework area is divided into a northern (NGLFR) and a southern (MAR) section (likely a distinction between agricultural and forested land use). The division of the central framework area is supported

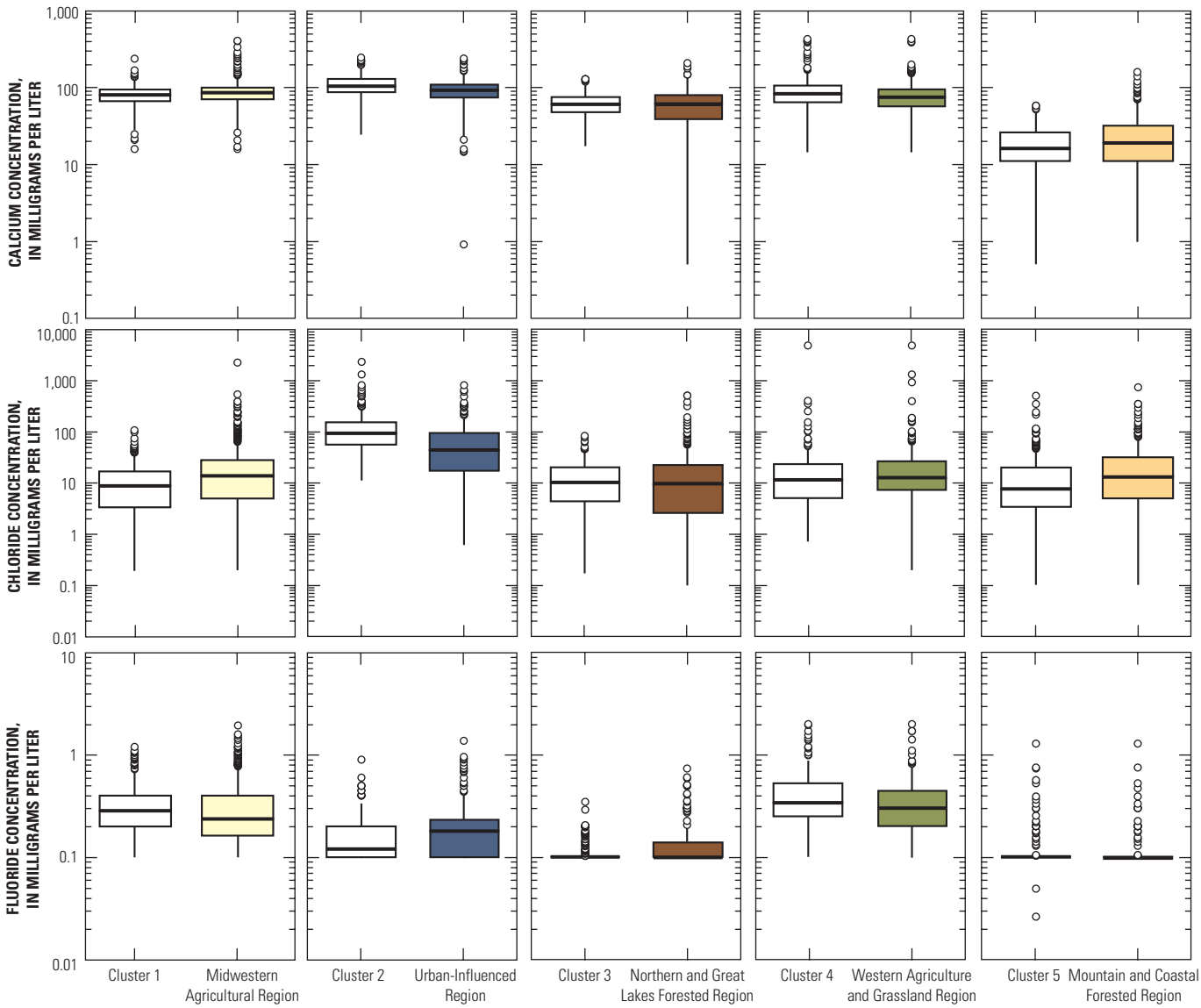


Figure 11. Distribution of major and minor ions in clusters and corresponding hydrochemical regions. Refer to figure 6 for an explanation of box plot graphs.

by recent regional studies on arsenic (Thomas, 2007) and on radionuclides (Ayotte and others, 2007) in the glacial aquifer system that divided the central framework area into northern and southern sections; however, their divisions were based on the occurrence and distribution of those constituents that were studied. Similar divisions of the central framework area that are based on different lines of reasoning indicate the aquifer system may be vulnerable to both natural and anthropogenic contaminants in this area.

The NGLFR and the MCFR (fig. 10) cover most of the East and West framework areas (fig. 2). The MAR corresponds well to the southern part of the Central framework area except for Iowa, Missouri, and southern Minnesota. The WAGR corresponds well to the West-Central framework area. The UIR was unexpected by the framework design and is a result of including land use in the analysis of hydrochemical regions. The UIR is spatially distributed such that it is interior to the other hydrochemical regions. Although there are primary areas defining each of the hydrochemical regions, all of the regions are interspersed with each other. This interspersion indicates that different combinations of the various environmental characteristics can produce similar water-quality conditions, and reflects the heterogeneity of the glacial aquifer system.

Environmental Characteristics of Hydrochemical Regions

Precipitation and dissolution are responsible for most of the important ground-water chemical reactions (Alley, 1993). Generally, the longer water is in contact with geologic and soil minerals (subsurface contact time), the higher will be

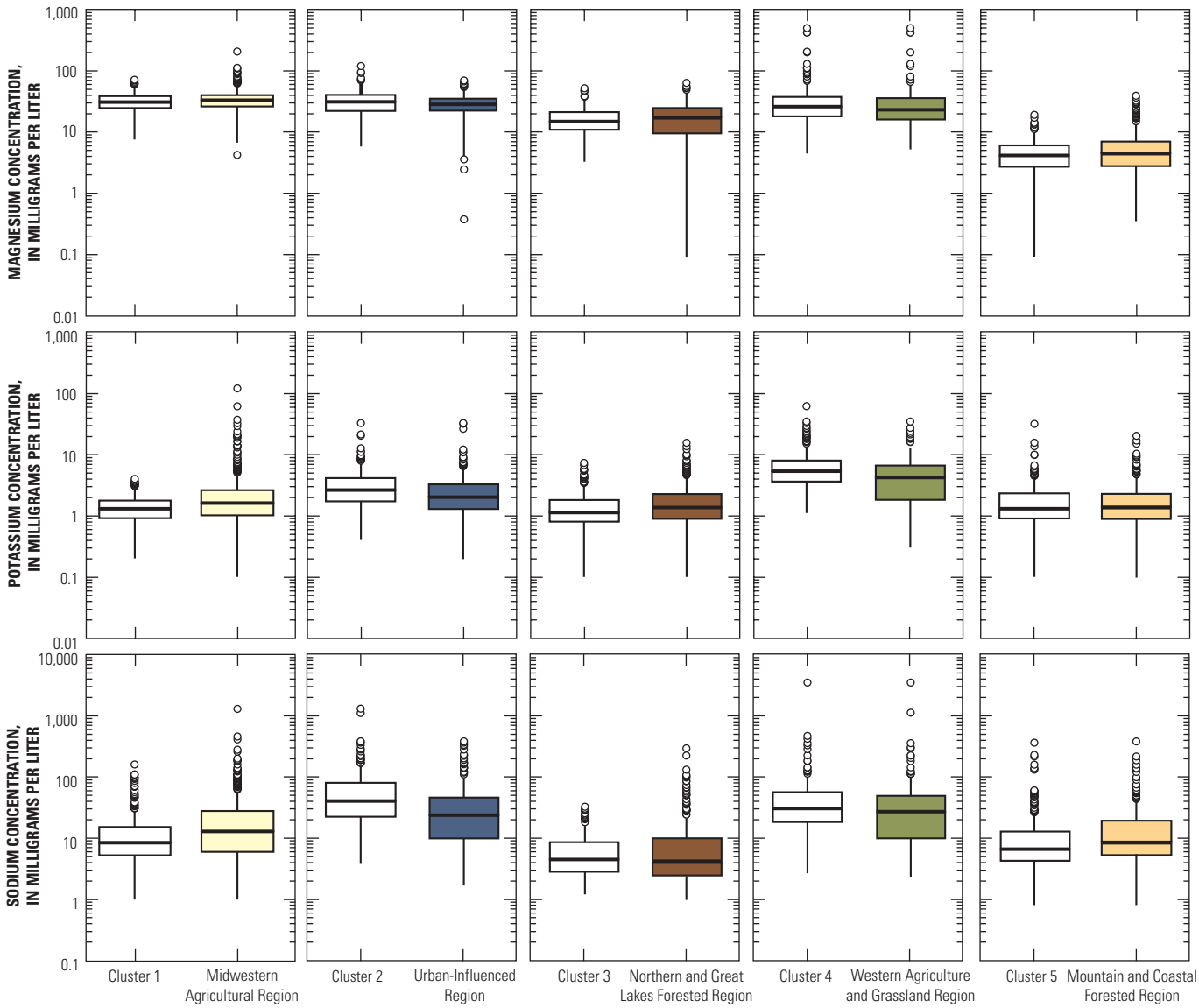


Figure 11. Distribution of major and minor ions in clusters and corresponding hydrochemical regions. Refer to figure 6 for an explanation of box plot graphs.—Continued

the dissolved solids and ionic concentrations under natural conditions. This result is evident when evaluating hydrochemical regions in the glacial aquifer system with respect to the environmental characteristics that define them.

The NGLFR and the MCFR have the lowest percent clay in soil and generally have the lowest ion concentrations in ground water (fig. 13; tables A14 and A15). In addition, both of these hydrochemical regions have environmental characteristics that facilitate the movement of water through the system (high natural ground-water recharge, moderately low percentage of fine-grained surficial and soil material, and high vertical permeability). Because water can move quickly through the system, the amount of time water is in contact with soil materials is relatively short (median subsurface contact time 85 and 49 days, respectively) (table A16); ion concentrations in the water are low (fig. 13; table A15). In comparison, waters in the MAR, the UIR, and the WAGR have substantially higher ion concentrations (fig. 13; table A15) and environmental characteristics in these regions retard the flow of water through the aquifer system, which increases the amount of time water is in contact with soil materials (median subsurface contact time 475, 207, and 456 days, respectively) (table A16). The MAR, UIR, and WAGR have high percent clay and fine-grained or mixed texture surficial deposits, low natural ground-water recharge, and low vertical permeability (table A14).

The Midwestern Agricultural Region

Located in the central part of the glacial aquifer system, the MAR (fig. 10) is a relatively flat landscape that is covered primarily by agricultural land, a potential source of agrichemi-

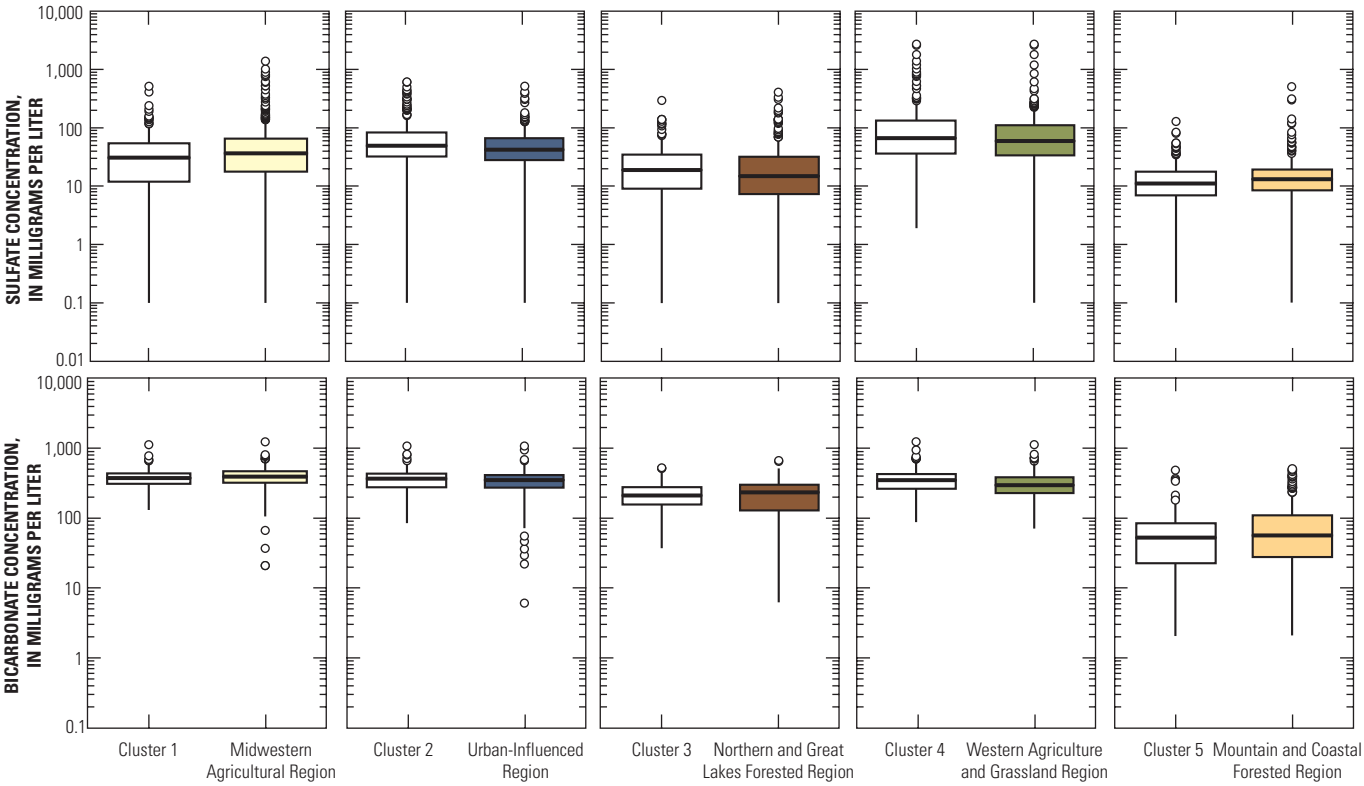


Figure 11. Distribution of major and minor ions in clusters and corresponding hydrochemical regions. Refer to figure 6 for an explanation of box plot graphs.—Continued

cals to ground water (table A14). Ninety-two percent of the wells identified in the MAR were in predominantly (greater than 50 percent of the area within 1,640 ft of the well) agricultural areas. Forty-eight percent of wells in the region were in networks designed to facilitate study of agricultural land-use effects on shallow ground water (table A17). The climate in the MAR is warm and humid relative to the other hydrochemical regions. The mean precipitation for the region is 37 inches per year (in./yr). Fine textured, silty and clayey glacial till in this area decreases potential for recharge to shallow aquifer material. The surficial deposits in the MAR have relatively low vertical permeability [mean of 1 inch per hour (in/h)] and low annual recharge (mean of 3 in./yr) (table A14), and thus ground water has the longest median subsurface contact time (475 days) of all the hydrochemical regions (table A16). Clayey soils may increase subsurface contact time of water, which may consequently increase the potential to dissolve natural chemicals from the soil materials. Tile drains are a common feature in agricultural areas of this region where surficial materials are clayey. Where present, tile drains divert excess overland flow to surface water bodies (streams and drainage ditches) instead of recharging ground water; this diversion may decrease the likelihood of anthropogenic contaminants reaching ground water.

The Urban-Influenced Region

The UIR has a low to moderate slope compared to the other hydrochemical regions and has a multitude of land uses. About one-third of the region is urban land (30 percent). Most (56 percent) of the wells identified as being in the UIR are in predominantly (greater than 50 percent of the area within 1,640 ft of the well) urban land-use areas. Urban areas of the UIR spatially correspond to metropolitan statistical areas with populations between 70,000 and 7,000,000 people in 1990 (U.S Bureau of the Census, 1991). Forty-seven percent of the wells in the UIR are in networks designed to facilitate study of urban land-use effects on shallow ground water (table A17). One example of an urban land-use area in this region is the area in northeastern Illinois that roughly corresponds to the city of Chicago. Similarly, Detroit in southeastern Michigan, and the cities of Buffalo and Rochester in northwestern New York can be identified by looking at the urban land-use areas of this hydrochemical region (fig. 10). However, not all urban areas are represented here because urban land use is not the only environmental factor contributing to this hydrochemical region.

Forest/shrub and agricultural also are land uses in the UIR (36 and 26 percent, respectively) (fig. 10; table A14). Forested and agricultural areas in this region commonly are intermingled. Areas overlain by forest/shrub in the UIR generally have fine- or coarse-grained till, alluvial, or glaciofluvial surficial deposits. Agricultural areas in the UIR generally are

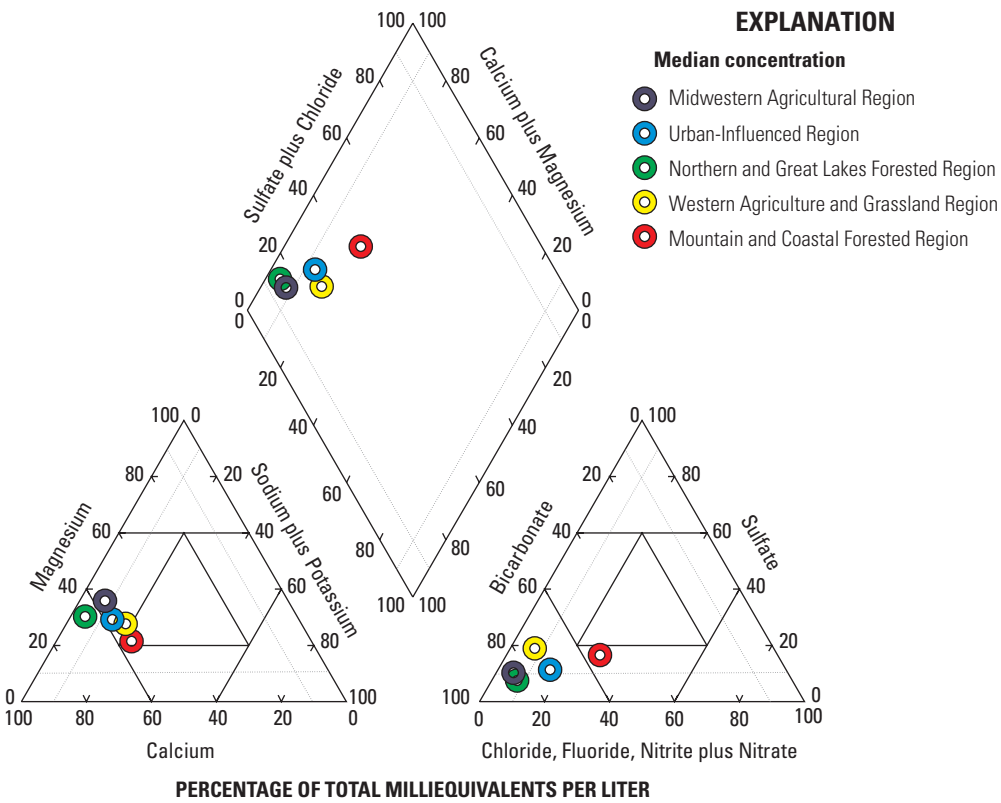


Figure 12. Median measured major- and minor-ion concentrations, indicating general water type of each hydrochemical region.

near streams and have surficial deposits that are clayey, fine-grained alluvium or colluvium.

Compared to the other hydrochemical regions, the UIR has a relatively high mean annual temperature (48 °F) and relatively moderate to high mean annual precipitation (38 in.). The wide range of climatic conditions probably is indicative of the spatially dispersed locations of UIR areas. Soils in this hydrochemical region are moderately clayey, overall. When compared to the other hydrochemical regions, however, the vertical soil permeability is moderate, with a mean of 2 in./h, and estimated ground-water recharge also is moderate, with a mean of around 6 in./yr (table A14).

The Northern and Great Lakes Forested Region

The NGLFR mostly is a mix of forest and agriculture land use, but it also has a large percentage of wetland area (50 percent, 28 percent, and 17 percent, respectively) (fig. 10; table A14). Wetland and forested areas, such as those in northern Wisconsin and Minnesota, have soils with high organic content (Wolock, 1997) that can have a mitigating effect on potential contaminants. Surficial materials in this hydrochemical region are almost dichotomously fine- or coarse-grained, a consequence of their glacial, alluvial, or lacustrine origins (figs. B7 and B8). Because of its northern location, the NGLFR has a cool, moderately humid climate, with a mean annual temperature of 43 °F, and a moderately dry mean annual precipitation of 36 in. (table A14). Soils in the NGLFR have the second lowest average percent clay, and therefore are likely to be sandy; the relatively high mean vertical permeability of the soils (4 in./h) reflects this sandiness. Furthermore, the substantially higher estimated ground-water recharge in the region—mean of 7 in./yr—is consistent with a sandy soil texture. These characteristics, combined, indicate that water has shorter subsurface contact time in soil materials in the NGLFR (median of 85 days) than in some of the other regions (table A16).

The Western Agriculture and Grassland Region

Most of the WAGR is in the west-central part of the glacial aquifer system, and consists mostly of flat plains covered with agricultural fields and grasslands (fig. 10, table A14). Many fields in the WAGR are irrigated for agriculture, which is a distinguishing characteristic of this region. Surficial materials in the area are mostly fine-grained glacial till, alluvium/colluvium, and lacustrine deposits. Examples of these types of deposits are in northern Montana and eastern South Dakota (fig. B8). The soils are very clayey compared to those in other hydrochemical regions. Vertical soil permeability is relatively small, with a mean of about 1 in./ hr. Ground-water recharge is estimated to range between less than 1 and 70 in./yr, with a mean around 1 in./yr, which is the lowest of all hydrochemi-

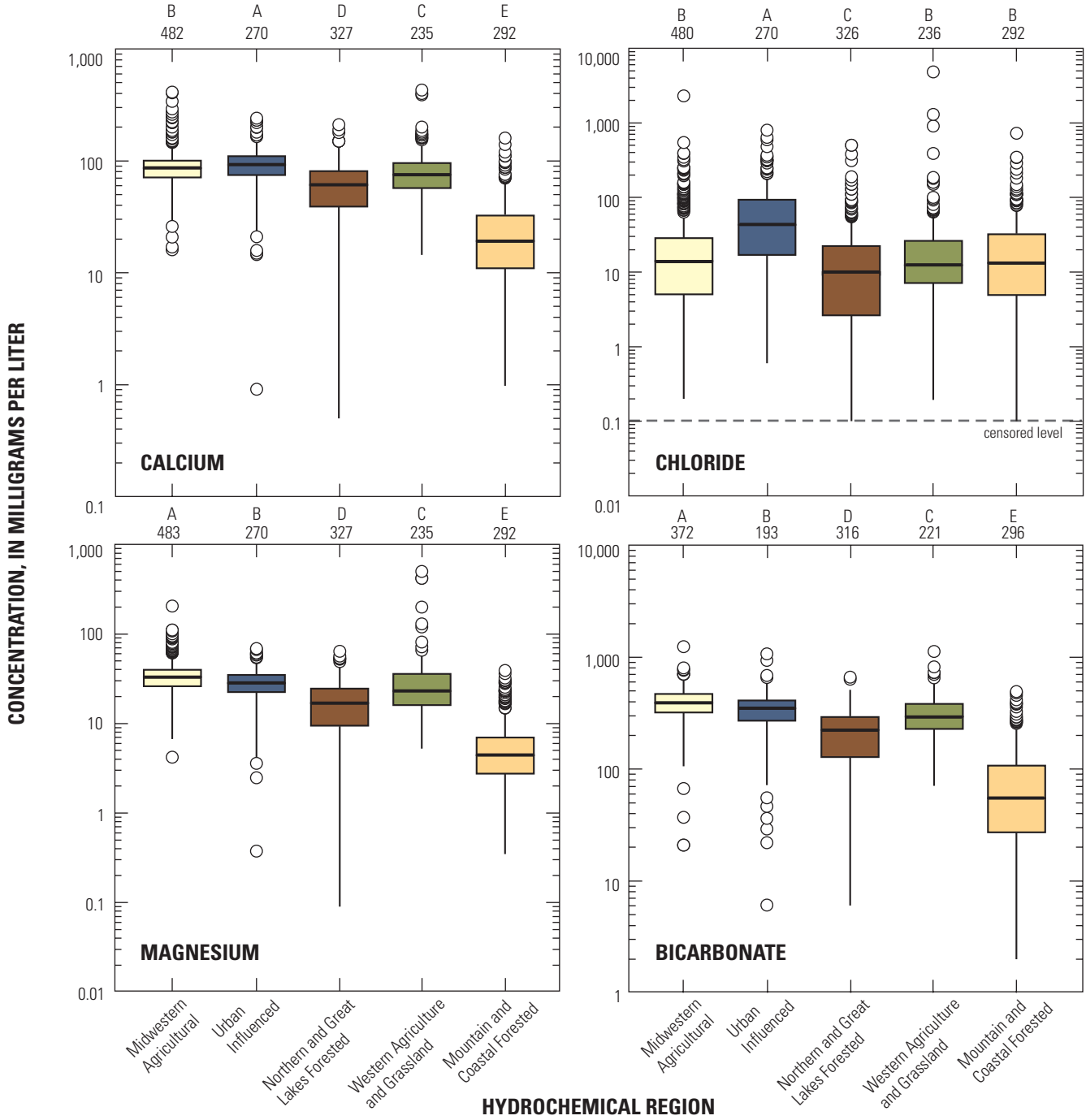


Figure 13. Distribution of major-ion concentrations by hydrochemical region in the glacial aquifer system, northern United States. Refer to figure 6 for an explanation of box plot graphs.

cal regions. The climate of this region is dry, with a mean annual precipitation of 21 in./yr, and warm, with a mean annual temperature of 45 °F (table A14). Parts of the WAGR are interspersed within other hydrochemical regions because of the environmental characteristics that define the regions. In particular, parts of the WAGR are in northwestern Ohio, coincident with the location of former glacial Lake Maumee (figs. B3, B8), and in north central Illinois, an area that also is coincident with a former glacial lake (Geological Survey of Canada, 1989). More than 75 percent of the measured silica concentrations in the WAGR are greater than 39 mg/L, which is much greater than most of the measured silica concentra-

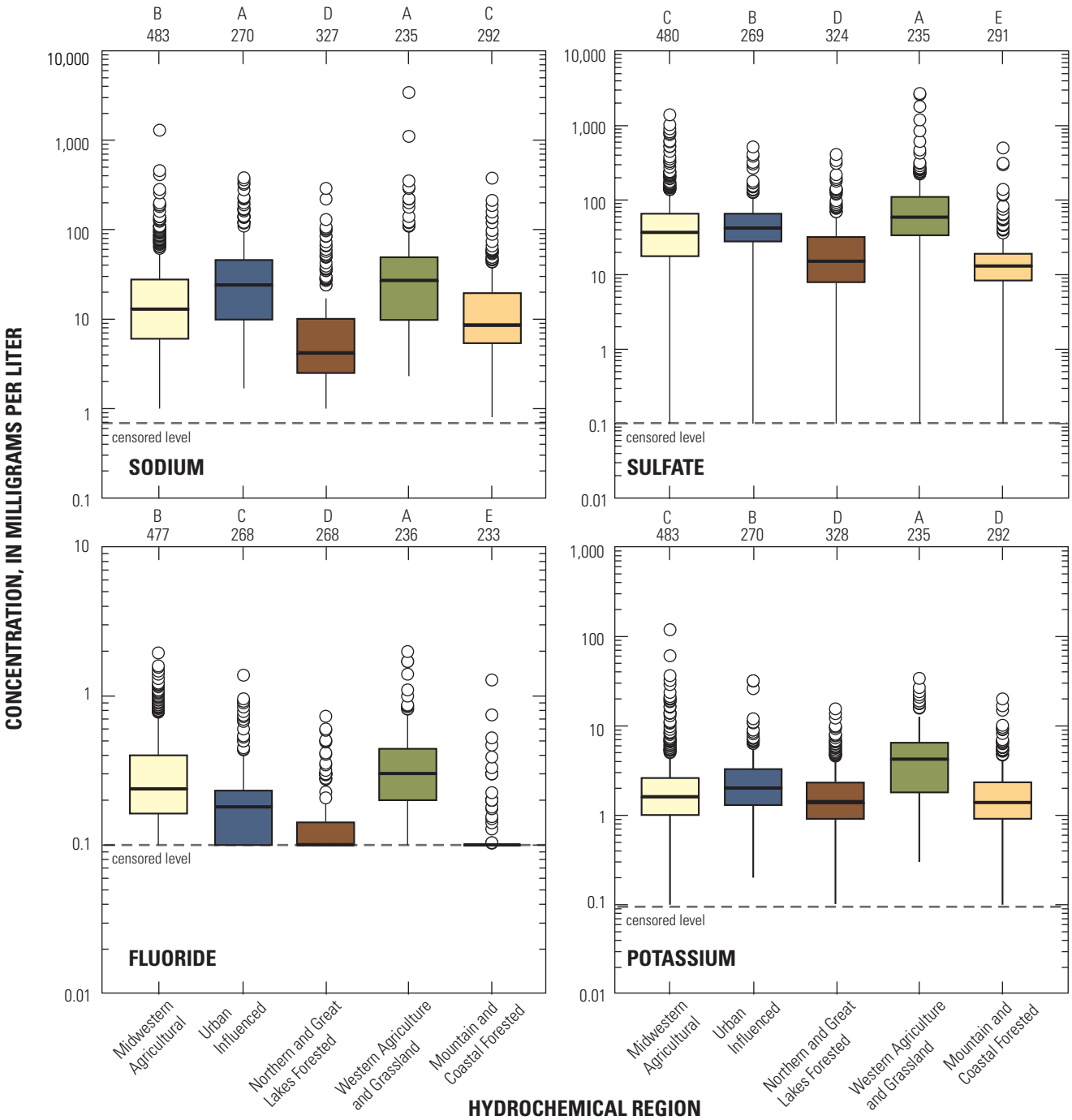


Figure 13. Distribution of major-ion concentrations by hydrochemical region in the glacial aquifer system, northern United States. Refer to figure 6 for an explanation of box plot graphs.—Continued

tions in other hydrochemical regions (fig. 14; table A15). The high silica concentrations in ground water are likely a result of the relatively long time that ground water is in contact with the geologic source material.

The Mountain and Coastal Forested Region

Most of the MCFR encompasses the eastern and western mountain and coastal areas of the glacial aquifer (fig. 10; table A14). This hydrochemical region has the most variable topography of all the regions, and consists of areas of relatively steep slopes (up to 107 percent) in Washington and western Montana (fig. B3). The land cover/land use in this hydrochemical region primarily is forest (71 percent of the area) and agriculture (13 percent of the area). Surficial deposits are mostly mixed texture, and soils are very low in clay content relative to soils in the other regions. Vertical permeability is moderately high compared to that in the other

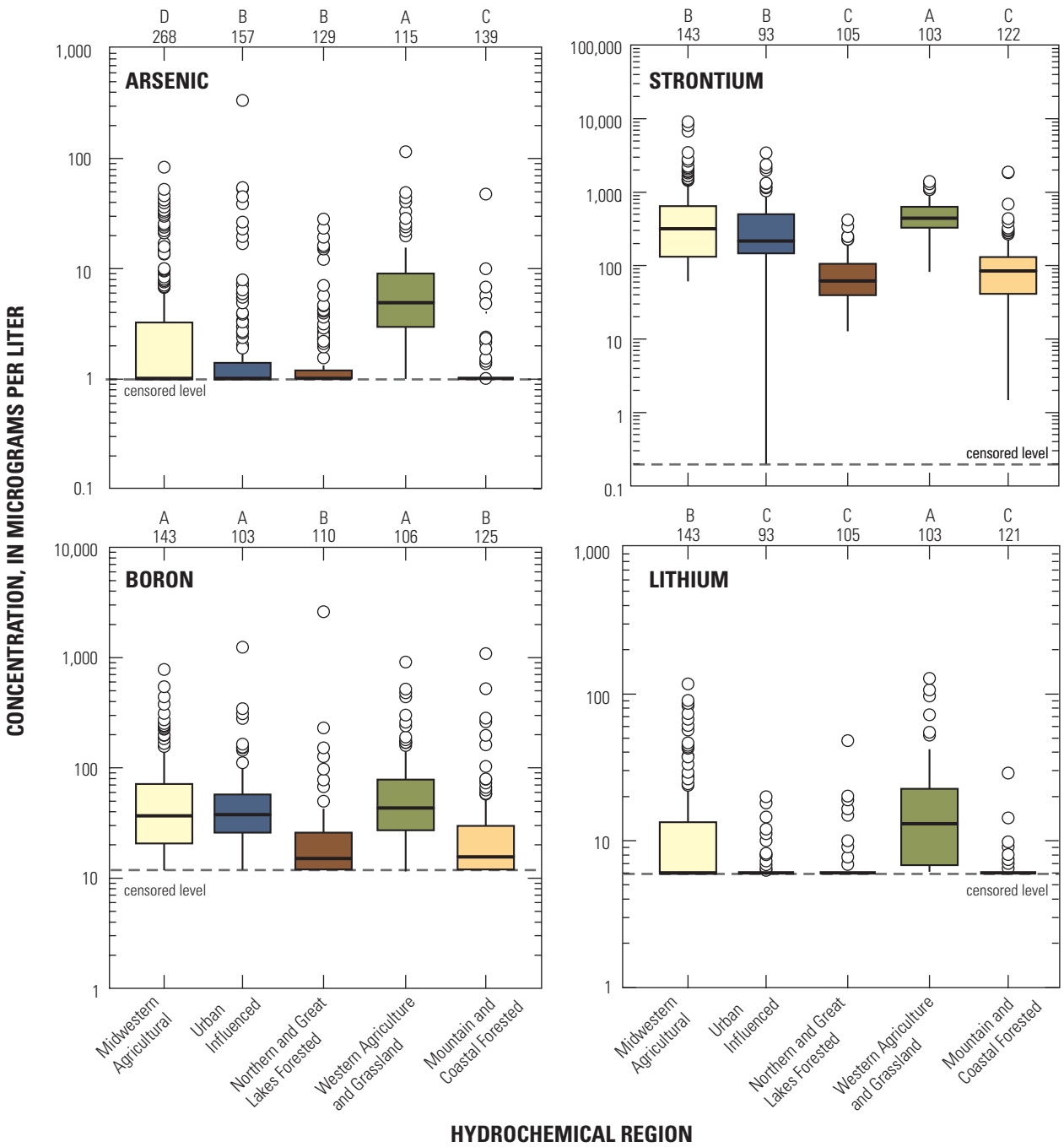


Figure 14. Distribution of selected constituents by hydrochemical region in the glacial aquifer system, northern United States. Refer to figure 6 for an explanation of box plot graphs.

hydrochemical regions, with an estimated maximum around 18 in./h. Estimated ground-water recharge also is high in the MCFR (mean about 13 and maximum about 85 in./yr). The MCFR has the wettest climate of the hydrochemical regions and is comparatively cool (figs. B1, B2). Average annual precipitation is 45 in. and the average annual temperature is around 44 °F. The cool, wet climate provides relatively greater recharge to ground water and the high permeability and low clay content of soils indicates that subsurface contact time may be short. Ground water in the MCFR has a very high median concentration of dissolved oxygen (4.9 mg/L) relative to that in other hydrochemical regions. This high median concentration of dissolved oxygen may be another indication of recently recharged ground water and short subsurface contact times (median of 49 days) (table A16). Most of the wells in the eastern part of this region—in New England and parts of New York and New Jersey, for example—were in valley-fill deposits where the glacial mate-

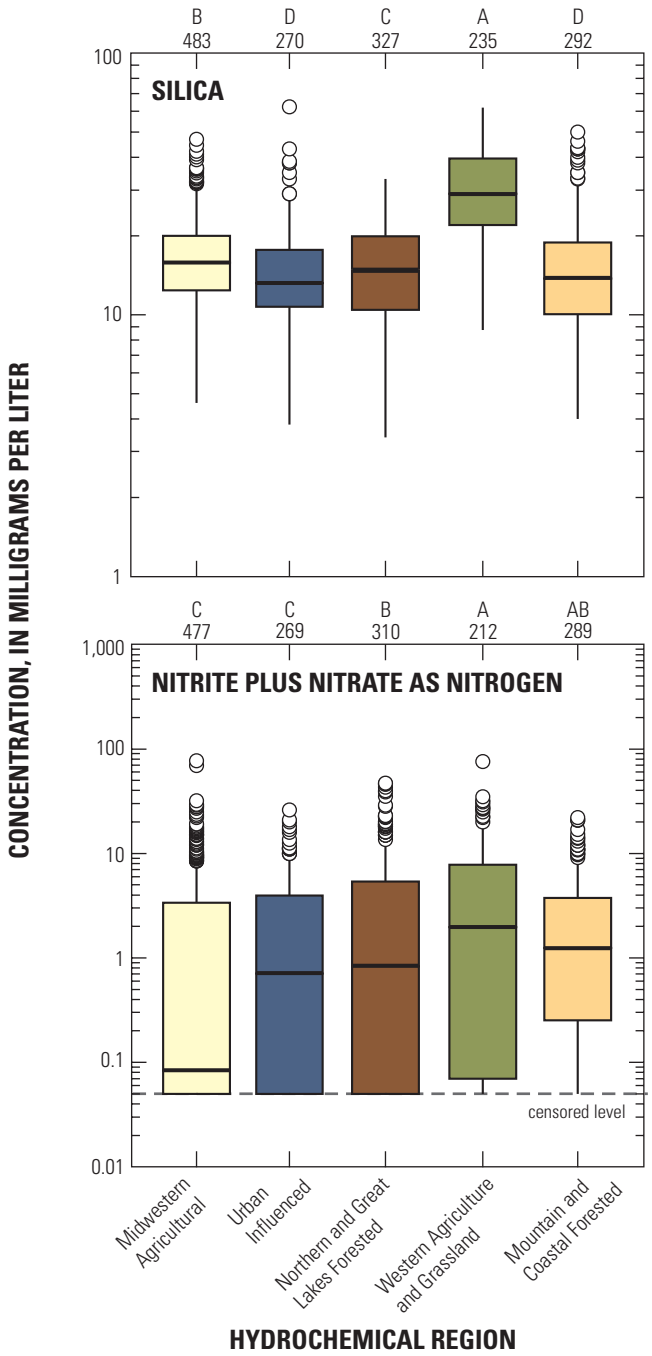


Figure 14. Distribution of selected constituents by hydrochemical region in the glacial aquifer system, northern United States. Refer to figure 6 for an explanation of box plot graphs.—Continued

rial is sandy and relatively thick. Other parts of the glacial aquifer system in the eastern part of this region are not considered a major source of water because those parts have thin till that is less sandy; although, the till can be locally thick in some areas. The hydrochemical region map of the eastern part of this region mostly represents the characteristics of the valley-fill deposits (fig. 10).

Water-Quality Characteristics of Hydrochemical Regions

Water-quality characteristics can be affected by the environmental characteristics of the aquifer material as well as by land cover/land use; this is the premise for the delineation of hydrochemical regions. Each of the ions Ca, Cl, F, Mg, K, Na, SO_4, and HCO_3 occur naturally in the environment, but they also may be added to the environment by human activities, such as water treatment or fossil-fuel combustion (Hem, 1992). Generally, ground water in the NGLFR and the MCFR has the lowest ion concentrations (figs. 10 and 13; tables A14 and A15); in comparison, ground water in the MAR, the UIR, and the WAGR has higher ion concentrations (figs. 10 and 13; table A15).

Most of the water types in the glacial aquifer system are dominated by calcium-magnesium, and bicarbonate with variations in the other ions. The WAGR has the highest percent (10 percent) of magnesium-bicarbonate dominated waters (fig.15). The median concentration of dissolved solids in the WAGR (405 mg/L) is related to the relatively high ion concentrations in the region. Recycling water through irrigation in the area may be responsible for concentrating these constituents in the water. Other, minor anthropogenic sources of magnesium are nonferrous metal production, underground pipes and storage tanks, and water heaters (Kramer, 2001).

Ground water in the MAR is dominated by calcium-bicarbonate (83 percent), similar to water in the other hydrochemical regions; however, in contrast to the other hydrochemical regions, this region has no significant sodium-chloride dominated waters (fig.15). Comparing concentrations among hydrochemical regions, the MAR (fig. 10) has the highest median concentration of magnesium (33 mg/L). The MAR has the lowest median calcium to magnesium ratio (1.59), perhaps indicating a higher percent of dolomitic material in the area (table A15).

The UIR and the MCFR have the greatest percentages of calcium-chloride (4 and 6 percent, respectively) and sodium-chloride (3 and 12 percent, respectively) dominated waters compared to the other hydrochemical regions (fig. 15). Chloride has natural sources such as atmospheric deposition, weathering of some soils and glacial deposits, volcanic activity, and saline ground water (John Mullaney, U.S. Geological Survey, oral commun., 2007). Saline springs in parts of Michigan, Illinois, and New York indicate that saline water is present at shallow depths in some parts of the glacial aquifer system (John Mullaney, U.S. Geological Survey, oral commun., 2007). Chloride also can be an anthropogenic contaminant. Some anthropogenic sources of chloride include oil and gas production, dissolution of road-deicing salts, leachate from landfills, byproducts of drinking-water treatment, and discharge of wastewater from treatment facilities or septic systems (John Mullaney, U.S. Geological Survey, oral commun., 2007). Higher chloride concentrations are significantly correlated (p-value < 0.05) to higher percentages of urban land use within 1,640 ft of a well in each of the hydrochemical

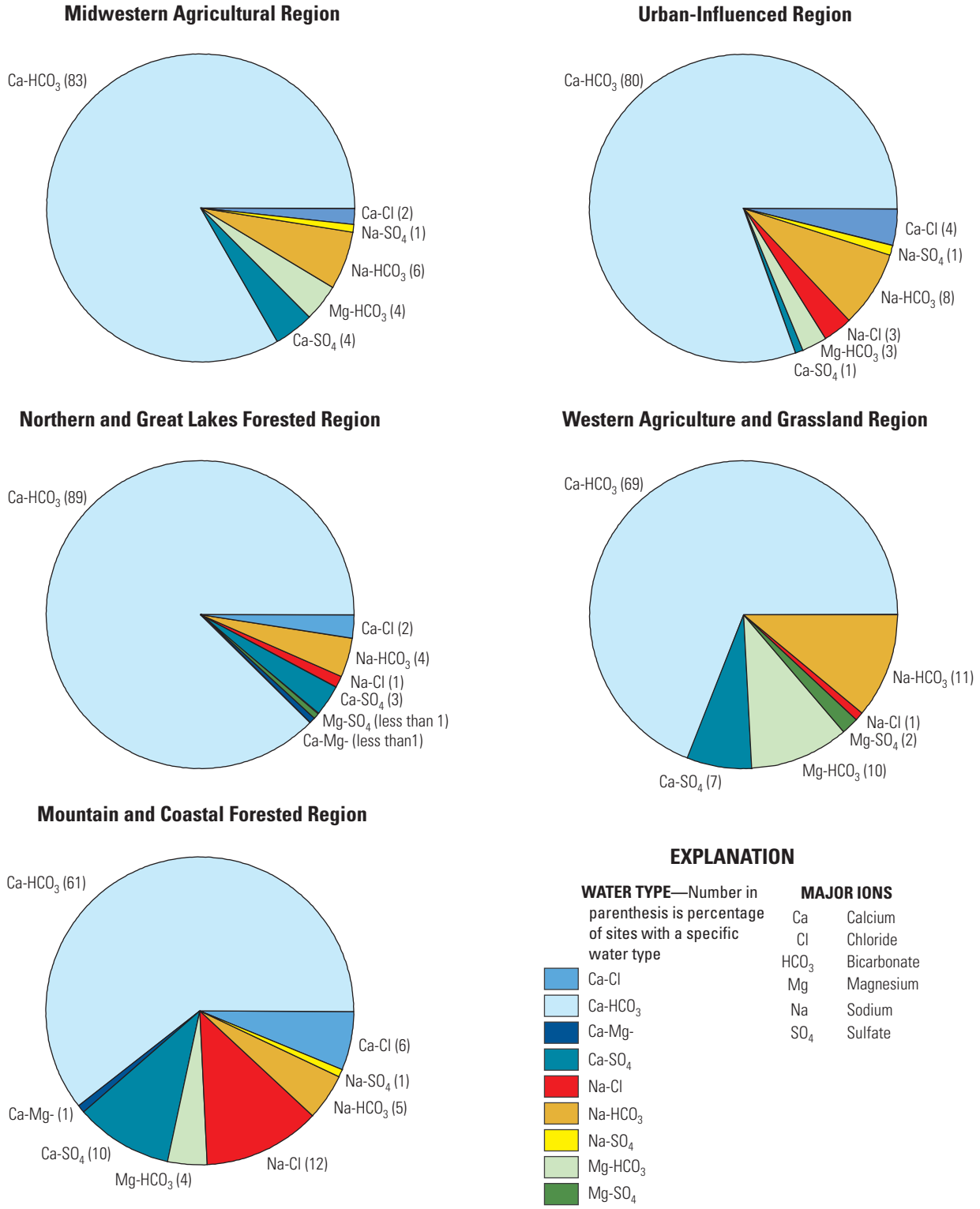


Figure 15. Water type in the hydrochemical regions of the glacial aquifer system is dominated by calcium and bicarbonate ions; however, there are variations in the relative concentrations of other ions.

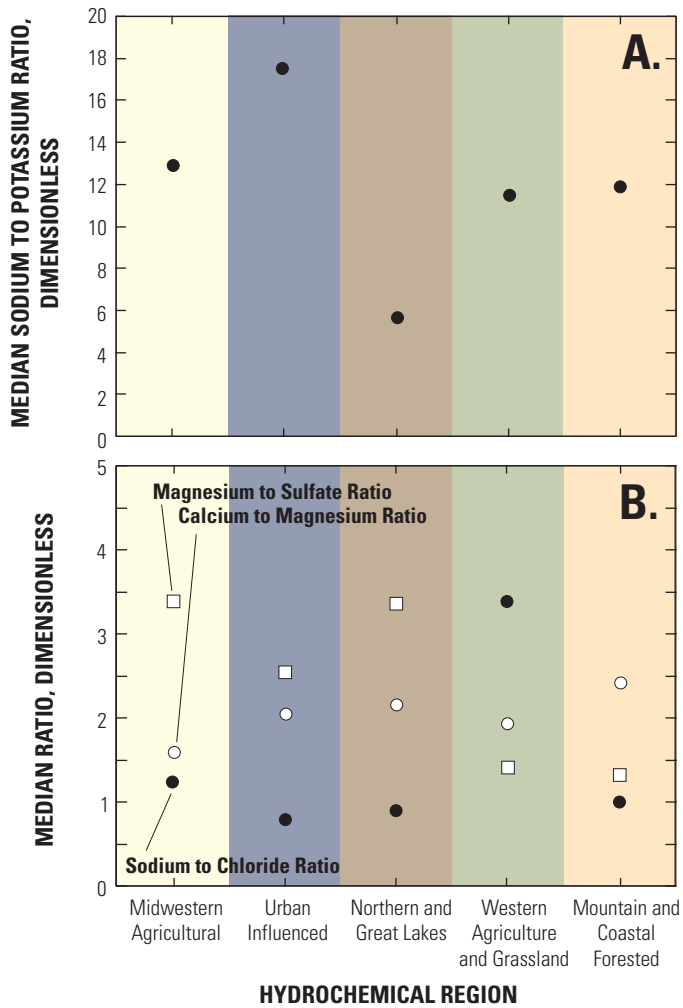


Figure 16. Ratios of *A,* sodium to potassium, and *B,* calcium to magnesium, magnesium to sulfate, and sodium to chloride in hydrochemical regions.

regions (table A18). The UIR has the lowest median sodium to chloride ratio (0.79) (fig. 16; table A15), which indicates the presence of a higher concentration of chloride than sodium. The median ratio of chloride to bromide in the UIR is 720, which is the highest median in any of the hydrochemical regions (table A15). High concentrations of chloride relative to sodium and to bromide in the UIR may be related to the urban setting of some of the UIR areas (John Mullaney, U.S. Geological Survey, oral commun., 2005). Ratios of chloride to bromide greater than 600 can be an indication that the chloride is from road salt dissolution (Risch and Robinson, 2001). Use of road salt is extensive, especially in northern urban areas.

The NGLFR has 89 percent calcium-bicarbonate dominated waters and has the lowest percentage of water-types dominated by other ions (fig. 15). In contrast, the MCFR has the lowest percent (61 percent) of calcium-bicarbonate dominated waters and the highest percent (12 percent) of sodium-chloride dominated waters (fig. 15). The high percent sodium-chloride waters in the MCFR may reflect urban land use because 35 percent of the wells in this region originally were selected to evaluate the effects of urban land use on ground-water quality (table A17). Waters in the NGLFR and the MCFR are softer than waters in the other hydrochemical regions based on median hardness (226 mg/L and 69 mg/L, respectively) (table A15).

Many ions can be depleted in ground water under mature forests (Federer and others, 1989). Concentrations of Ca, Cl, Mg, K, Na, SO_4, and HCO_3 are significantly and negatively correlated with the percentage of forested land in the UIR, NGLFR, and MCFR. Results of correlations are shown in table A18.

Potassium is a highly soluble element in water. The NGLFR and the MCFR have similar low measured potassium concentrations in ground water (fig. 13; table A12). Potassium concentrations generally are greater in the WAGR (median of 4.19 mg/L); median potassium concentrations in ground water in the other hydrochemical regions are about 2 mg/L or less (fig. 13).

Dissolved solids concentrations are significantly higher in the UIR (median of 459 mg/L) compared to concentrations in the NGLFR (median of 256 mg/L) or the MCFR (median of 124 mg/L). The MAR has dissolved solids concentrations ranging between those in the UIR and the WAGR and, therefore, is not significantly different from either of those regions (tables A12 and A15). Dissolved solids concentrations are significantly lower (95 percent confidence) in the MCFR than in any of the other hydrochemical regions. This finding is similar to that of Güler and Thyne (2004), who used inverse geochemical modeling to determine that the concentrations of dissolved solids were related to the spatial location of the water-quality groups along a flow path and that concentrations of dissolved solids were a distinguishing characteristic of their clustered water-quality groups.

Selected Trace Elements and Nitrite Plus Nitrate in Hydrochemical Regions

Concentrations of 20 trace elements and nutrients in ground water were evaluated using Kruskal-Wallis Rank Sum tests to determine if the concentrations varied by hydrochemical region. Concentrations of some trace elements and nutrients varied significantly among two or more hydrochemical regions (table A12). In another study of water quality in the glacial aquifer system, which examined the occurrence and distribution of trace elements, the concentrations of some trace elements including barium, arsenic, strontium, and lithium varied significantly among the hydrochemical regions and among the framework areas (George Groschen, U.S. Geological Survey, oral commun., 2006). Many of the trace elements and nutrients that were evaluated have both natural and anthropogenic sources; however, natural sources are the most likely explanation for variations of these concentrations among the hydrochemical regions. Variations of select trace elements are discussed below.

Potential natural sources of arsenic in ground water from the glacial aquifer system are weathering of black shale or dark limestone, mobilization of adsorbed or co-precipitated arsenic, or ground-water recharge to the glacial aquifer system from arsenic bearing bedrock (Warner and others, 2003). Arsenic is of interest because of its potential toxicity to mammals. The UIR and the NGLFR have similar predominantly low concentrations of arsenic, where 75 percent of the measured concentrations are less than 1.4 micrograms per liter (µg/L) in each region (tables A12 and A15). Median arsenic concentration (4.88 µg/L) is significantly higher in the WAGR than in the other hydrochemical regions, where median arsenic concentrations are 1 µg/L.

Strontium may be derived from weathering of sandstone or igneous rocks as there are no known major anthropogenic sources of strontium. In Minnesota, strontium concentrations in ground water were determined to be variable, indicating that geology is a controlling factor for strontium concentrations (Minnesota Pollution Control Agency, 1999). Strontium was highly correlated with arsenic in ground water from glacial valley-fill deposits in Ohio (Thomas, 2003), which is mostly in the MAR. In this report, analysis indicates that higher arsenic concentrations are significantly correlated to higher strontium concentrations in the MAR (p-value < 0.0001). Median strontium concentrations are highest in the WAGR (median of 442 µg/L), the MAR (median of 315 µg/L), and the UIR (median of 217 µg/L), and lowest in the NGLFR (median of 62 µg/L) and the MCFR (median of 86 µg/L) (table A15).

Boron normally is derived from weathering of boron-containing silicate rocks (for example, tourmaline). Anthropogenic sources of boron are the process of sewage treatment and industrial waste products (Hem, 1992). Boron is of interest because it is important for promoting plant growth; however, too much boron can be toxic to plants. The WAGR has the highest median boron concentration (43 µg/L), followed by the UIR (median of 38 µg/L) and the MAR (median of 37 µg/L). Boron is present in the bedrock aquifers underlying parts of Illinois in the MAR (Warner, 2000). These bedrock aquifers may be in contact with the glacial aquifer system. Boron is present as connate water in bedrock aquifers under the Michigan Basin upwelling to glacial deposits (Robert Kay, USGS, written commun., 2007). The southeastern part of the Michigan Basin predominantly is in the MAR. The NGLFR and the MCFR have similar low measured boron concentrations (medians of 16 µg/L and 15 µg/L, respectively) (table A15, fig. 14).

Lithium may be present in clay minerals (Starkey, 1982). Once lithium is dissolved, it usually remains in the dissolved state and may not be removed from water by adsorption to clays because more preferably adsorbed cations usually are present in ground water (Hem, 1992). The UIR, NGLFR, and the MCFR have similar moderately low mean lithium concentrations (tables A12, A15). Other than among these hydrochemical regions, lithium concentrations are significantly different. The MAR and the WAGR have higher mean and higher maximum lithium concentrations than the other hydrochemical regions, and lithium concentrations and clay are significantly and positively correlated in these regions (p-value < 0.05); lithium concentrations and clay are not correlated in the other hydrochemical regions. The maximum lithium concentration, 126 µg/L, was measured in the WAGR (table A15; fig. 14).

The MAR and the UIR have similar distributions of measured nitrite plus nitrate as nitrogen (NO_2 plus NO_3 as N) concentrations and have lower median NO_2 plus NO_3 as N concentrations (0.08 mg/L and 0.71 mg/L, respectively) than the other hydrochemical regions (table A15). Unexpectedly, the MAR has the lowest median NO_2 plus NO_3 as N concentration (0.08 mg/L) despite the high percentage of agricultural land use in the area (tables A14, A15). The low median NO_2 plus NO_3 as N concentration in ground water may be related to precipitation runoff that is shunted through subsurface drains and drainage ditches directly to surface water in much of the region. Diverting runoff from agricultural fields to surface-water drainage would prevent the infiltration of NO_2 plus NO_3 as N into shallow ground water. NO_2 plus NO_3 as N has a maximum of 77 mg/L in the MAR—similar to the WAGR (maximum of 76 mg/L), which also has a high percentage of agricultural land use. However, the WAGR has the highest median NO_2 plus NO_3 as N. This region is highly irrigated, and the recycling of irrigation water may be one reason why NO_2 plus NO_3 as N is higher in this region than in the other hydrochemical regions. An alternate explanation for the differences in NO_2 plus NO_3 as N concentrations in the MAR and WAGR could be different oxidation-reduction conditions; however, not enough data are available to determine regional oxidation-reduction conditions.

Local- to Regional-Scale Comparisons of Water Quality

Hydrochemical regions of the glacial aquifer system, as defined in this report, are appropriate for use with large regional-scale studies, as previously mentioned. The effect of extrapolating data collected within relatively small areas (local-scale studies) to the large area of the glacial aquifer system, virtually the entire northern United States, is to generalize the small-scale variations out of the data set. Only large-scale variations are evident when the water-quality characteristics of the individual hydrochemical regions are compared. Confidence in the water-quality characteristics that are extrapolated to unsampled areas of the hydrochemical regions is necessarily low; in contrast, confidence is higher where data actually were collected (fig. 10).

One example of where the hydrochemical regions generalize the local variability in water quality is in the Saginaw Lowland area of Michigan (fig. B9). Local-scale studies in the Saginaw Lowland reported high concentrations of chloride (between 100 and 1,000 mg/L) in water from the glaciofluvial aquifer, which is part of the glacial aquifer system (Wahrer and others, 1996). The high concentrations of chloride in

ground water are only partly indicated by the water-quality characteristics of the hydrochemical regions of that area (table A19). The MAR and the UIR, which contain parts of the Saginaw Lowland, have the two highest median concentrations of chloride in ground water (14 mg/L and 43 mg/L, respectively) of all the hydrochemical regions; however, the water-quality characteristics modeled at the regional scale do not adequately indicate the magnitude of the chloride concentrations in ground water as measured on the local scale. High chloride concentrations in ground water in the Saginaw Lowland area are most likely a consequence of local-scale geologic environmental characteristics that facilitate the upwelling of saline ground water into the glacial aquifer system (Westjohn and others, 1994; Wahrer and others, 1996). These local-scale geologic characteristics are not represented in the generalized regional-scale environmental data used to map the hydrochemical regions. Additionally, no water-quality data from the Saginaw Lowland area of Michigan were used to map the hydrochemical regions because only data collected as part of the NAWQA program were used to map the regions.

One example of where hydrochemical regions do adequately reflect the local-scale water-quality conditions is in Connecticut, which is in the NGLFR and MCFR. Grady (1988) reports median concentrations of various constituents measured in water from the aquifer system underlying different land uses in Connecticut. The range of median concentrations reported by Grady for selected constituents (including calcium, sulfate, magnesium, boron, lithium, and strontium) are mostly similar to the median concentrations of the same constituents in ground-water in the MCFR (table A20). The data reported by Grady (1988) were not used in the mapping of hydrochemical regions described in this report because only data collected as part of the NAWQA program were used to map the regions. However, some of the water-quality data that were used in the mapping of the hydrochemical regions for this report were collected from Connecticut (1992–2003). This example indicates that the hydrochemical regions more accurately indicate water quality in areas where actual data were used for the mapping (not extrapolated).

Summary

The glacial aquifer system, as defined in this report, is a large regional aquifer of heterogeneous composition that includes all unconsolidated geologic material above bedrock that lies on or north of the line of maximum glacial advance within the United States. The hydrochemical regions defined in this report were delineated and analyzed to allow an extrapolation of ground-water-quality characteristics to unsampled areas of the glacial aquifer system. A geographic information system was used to spatially analyze environmental characteristics that were statistically related to measured ground-water-quality data. Delineation of the hydrochemical regions confirmed a hypothesis based on the conceptual framework that was developed for water-quality analysis of ground water in the glacial aquifer system. Five hydrochemical regions of the glacial aquifer system were mapped based on the results of spatial and statistical analyses: Midwestern Agricultural Region (MAR), Urban-Influenced Region (UIR), Northern and Great Lakes Forested Region (NGLFR), Western Agriculture and Grassland Region (WAGR), and Mountain and Coastal Forested Region (MCFR).

Ground-water quality in the glacial aquifer system is affected by natural and anthropogenic characteristics. Natural environmental characteristics include climate, topography, estimated ground-water recharge, soil clay content and vertical permeability, and type and texture of surficial deposits. Land use is the primary anthropogenic environmental factor affecting ground-water quality in the aquifer system. Interaction between water and glacial deposits in different environmental settings can result in differing water chemistry. Statistical analysis of the concentrations of major and minor ions and selected trace elements in water samples collected from wells in the glacial aquifer indicates that variations in water quality can, in part, be explained by environmental characteristics.

No single environmental characteristic can account for the variability of water quality across the glacial aquifer system because many of these characteristics are related. Environmental characteristics that control the flow of water through the aquifer system are the most important factors defining the hydrochemical regions and their water quality. The length of time water is in contact with soil materials can affect chemical reactions and dissolution of aquifer materials that may add or remove naturally occurring major and minor ions and trace elements in the ground water. Longer subsurface contact time, as indicated by environmental characteristics that slow the flow of water through the system, is associated with hydrochemical regions in which the ground water has the highest concentrations of major and minor ions. Generally, the MAR, the UIR, and the WAGR have the highest ion concentrations in ground water, and the NGLFR and the MCFR have the lowest concentrations.

Most of the water types in the glacial aquifer system are dominated by calcium-magnesium-bicarbonate with variations in the relative concentrations of other ions. Many of the ions can be depleted beneath mature forests. Concentrations of calcium, chloride, magnesium, sodium, sulfate, and bicarbonate are significantly and negatively correlated with forested land use in the UIR, NGLFR, and MCFR.

Data analyzed for this study were collected from 1991 to 2003 at 1,716 wells from 19 NAWQA study units across the glacial aquifer system. Cluster analysis was used to group wells with similar concentrations of calcium, chloride, fluoride, magnesium, potassium, sodium, sulfate, and bicarbonate. Five distinct groups of wells with similar major- and minor-ion water chemistry were identified.

Maximum Likelihood Classification was used to extrapolate from these clusters, defined by points, to areas of hydrochemical regions. Combinations of average annual precipitation, average annual temperature, land use, slope, vertical soil

permeability, average clay soil content, texture of surficial deposits, type of surficial deposit, and estimated ground-water recharge were used to classify the areas to match environmental characteristics of the five groups identified by the cluster analysis. The result is a map of five hydrochemical regions of the glacial aquifer system. Although there are large areas defining each of the hydrochemical regions, the hydrochemical regions are interspersed with each other. This interspersion indicates that different combinations of environmental characteristics can produce similar water-quality conditions. Results of a logistic regression analysis indicated all analyzed environmental characteristics were important for predicting membership in the hydrochemical regions. Cumulatively, these environmental characteristics are related to subsurface contact time between ground water and the geologic materials. The hydrochemical regions could be used for regional-scale analysis but would need to be further refined for local-scale analysis. Findings of the large regional study presented in this report may or may not be consistent with findings in previous local and small-scale regional studies.

The spatial distribution of the five hydrochemical regions approximately follows the conceptual glacial aquifer system framework areas based on intrinsic susceptibility and natural vulnerability. The analyses described in this report determined that the characteristics of ground-water quality in the East and West Coast areas are similar. An additional finding of this analysis is that the original central framework area should be divided into northern and southern sections. The extents of the hydrochemical regions were affected by land use, which is an anthropogenic vulnerability factor not included in the conceptual framework design.

Concentrations of barium, arsenic, lithium, boron, strontium, and nitrite plus nitrate as nitrogen are significantly different among the hydrochemical regions. Similar to the major and minor ions, all of these trace metals have natural sources and some have potential anthropogenic sources.

The delineation of hydrochemical regions of an aquifer system, as defined in this report, is appropriate for use only in large regional-scale studies. The effect of extrapolating data collected in relatively small areas (local-scale studies) to the large area of the glacial aquifer system is to generalize the small-scale variations out of the data set. Only large-scale variations are evident when comparing water-quality characteristics of the individual hydrochemical regions. Confidence in the water-quality characteristics that are extrapolated to unsampled areas of the hydrochemical regions is necessarily low; in contrast, confidence is generally higher in areas where data were collected.

Acknowledgements

The authors thank Robert Kay, USGS Hydrologist; Scott Ator, USGS Research Hydrologist; Laurel Woodruff, USGS Geologist; Katherine Laub, USGS Editor; Chester Zenone, USGS Technical Reviewer; Dorothy Tepper, USGS Eastern Region Bureau Approving Official; and Angel Martin, Jr., USGS Volunteer for Science, for their thorough and thoughtful reviews of the manuscript. Their comments and contributions greatly improved the content of the manuscript.

References

Alley, W.M., 1993, Regional ground-water quality: New York, Van Nostrand Reinhold, 634 p.

Arnold, R.L., and Friedel, M.J., 2000, Effects of land cover on recharge potential of surficial and shallow bedrock aquifers in the Upper Illinois River Basin: U.S. Geological Survey Water-Resources Investigations Report 00–4027, 18 p.

Ayotte, J.D., Flanagan, S.M., and Morrow, W.J., 2007, Occurrence of uranium and ^{222}Radon in glacial and bedrock aquifer systems in the northern United States, 1993–2003: U.S. Geological Survey Scientific Investigations Report 2007–5037, 84 p.

Back W., 1966, Hydrochemical facies and groundwater flow patterns in northern part of Atlantic Coastal Plain: U.S. Geological Survey Professional Paper 498–A, 42 p.

Berg, R.C., Kempton, J.P., and Cartwright, K., 1984, Potential for contamination of shallow aquifers in Illinois: Illinois State Geological Survey Circular 532, 30 p.

Cannon, W.F., Kress, T.H., Sutphin, D.M., Morey, G.B., and Meints, J., 1997, Digital geologic map and mineral deposits of the Lake Superior Region: U.S. Geological Survey Open-File Report 97–455, 1:1,000,000 scale , accessed July 2006 from URL *http://pubs.usgs.gov/of/1997/of97-455/*.

Cederstrom, D. J., 1946, Genesis of groundwaters in the coastal plain of Virginia: Economic Geology, vol. 41, no. 3, p. 218 –245.

Chapelle, Francis P., 2004, Geochemistry of Groundwater, *in* Holland, H.D., and Turekian, K.K., eds., Treatise on Geochemistry: Elsevier/Pergamon, chap. 5.14.

Chebotarev I. I., 1955, Metamorphism of natural waters in the crust of weathering: Geochimica et Cosmochimica Acta, vol. 8 (pt. 1), p. 22–48; vol. 8 (pt. 2), p. 137–170; vol. 8 (pt. 3), p. 198 –212.

Clague, John J., 1989, Quaternary Geology of the Canadian Cordillera, *in* Fulton, R.J., ed., DNAG Quaternary Geology of Canada and Greenland, 839 p.

Clawges, Rick, and Price, Curtis, 1999, Digital data set describing surficial geology in the conterminous US: U.S. Geological Survey Open-File Report 99–77, digital data,

1:7,500,000 scale, accessed December 3, 2007, from URL *http://water.usgs.gov/lookup/getspatial?ofr99-77_geol75m*.

Daymet, 2005, Numerical Terradynamic Simulation Group: University of Montana, digital data, accessed December 2007 at URL: *http://www.daymet.org*.

Everitt, Brian, 1980, Cluster Analysis, (2nd ed.): New York, Halstead Press, 136 p.

Federer, C.A., Hornbeck, J.W., Tritton, L.M., Martin, C.W., Pierce, R.S., and Smith, C.T., 1989, Long-term depletion of calcium and other nutrients in Eastern Forests: Environmental Management, vol. 13, no. 5, p 593–601.

Fishman, M.J., and Friedman, L.C., 1989, Methods for determination of inorganic substances in water and fluvial sediments: U.S. Geological Survey Techniques of Water-Resources Investigations, (3d ed.), book 5, chap. A1, p. 459–460.

Focazio, M.J., Reilly, T.E., Rupert, M.G., and Helsel, D.R., 2002, Assessing the ground-water vulnerability to contamination—providing scientifically defensible information for decision makers: U.S. Geological Survey Circular 1224, 33 p.

Freeze, R.A., and Cherry, J.A., 1979, Groundwater: Englewood Cliffs, N. J., Prentice Hall, 604 p.

Fullerton, D.S., Bush, C.A., Penell, J.N., 2004, Surficial deposits and materials in the Eastern and Central United States (East of 102 Degrees West Longitude): U.S. Geological Survey Geologic Investigation Series I–2789, 1:2,000,000 scale map, accessed February 13, 2008, from URL *http://pubs.usgs.gov/imap/i-2789/*.

Geological Survey of Canada, 1989, Paleogeography of Northern North America 18,000–12,000 years ago: *in* Fulton, R.J. ed., Quaternary Geology of Canada and Greenland: Geological Survey of Canada, Map 1703A, sheet 1 of 3, scale 1:12,500,000, 839 p.

Gebert, W.A., Graczyk, D.J., and Krug, W.R., 1987, Average annual runoff in the United States, 1951–80: U.S. Geological Survey Hydrologic Investigations Atlas HA–710, 1 sheet, scale 1:7,500,000.

Glass, R.L., 2001, Ground-water quality in the Cook Inlet Basin, Alaska, 1999: U.S. Geological Survey Water-Resources Investigations Report 01–4208, 58 p.

Grady, S.J., 1988, Statistical comparison of ground-water quality in four land-use areas of stratified-drift aquifers in Connecticut *in* U.S. Geological Survey Toxic Substances Hydrology Program—Proceedings of the Technical Meeting, Phoenix, Arizona September 26–30, 1988: U.S. Geological Survey Water-Resources Investigations Report 88–4220, 1989, p. 473–481, 1 fig, 2 tab, 19 ref.

Güler, C., and Thyne, G.D., 2004, Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering: Water Resources Research, vol. 40, no. 12, December 2004, p 1–11.

Güler, C., Thyne, G.D., McCray, J.E., and Turner, A.K., 2002, Evaluation of graphical and multivariate statistical methods for classification of water chemistry data: Hydrogeology Journal, vol. 10, no. 4, August 2002, p. 455–474.

Hair, J.F. Jr., Anderson, R.E., Tatham, R.L., and Black, W.C., 1998, Multivariate data analysis (5th ed.): Upper Saddle River, N. J., Prentice Hall, 742 p.

Heath, R.C., 1988, Hydrogeologic map of North America showing the major rock units that underlie the surficial layer, *in* Back, W., Rosenshein, J.S., and Seaber, P.R., eds., Hydrogeology: Boulder, Colorado, Geological Society of America, The Geology of North America, vol. O–2, pl. 1.

Heath, R.C., 1984, Ground-water regions of the United States: U.S. Geological Survey Water-Supply Paper 2242, 78 p.

Helsel, D.R., and Hirsch, R.M., 1995, Statistical methods in water resources, (3d printing): Reston, Va., U.S. Geological Survey, Studies in Environmental Science 49, 532 p.

Hem, J.D., 1992, Study and interpretation of the chemical characteristics of natural water: U.S. Geological Survey Water-Supply Paper 2254, 264 p.

Hiscock, Kevin M., 2005, Hydrogeology principles and practice: United Kingdom, Blackwell Publishing, 389 p.

Hitt, K.J. and Nolan, B.T., 2005, Nitrate in ground water—using a model to simulate the probability of nitrate contamination of shallow ground water in the conterminous United States: U.S. Geological Survey Scientific Investigations Map 2881, 2 p., accessed December 3, 2007, from URL *http://pubs.usgs.gov/sim/2005/2881/*.

Kachigan, S.K., 1986, Statistical analysis: New York, Radius Press, 589 p.

Koterba, M.T., 1998, Ground-water data-collection protocols and procedures for the National Water-Quality Assessment Program—Collection and documentation, and compilation of required site, well, subsurface, and landscape data for wells: U.S. Geological Survey Water-Resources Investigations Report 98–4107, 91 p.

Koterba, M.T., Wilde, F.D., Lapham, W.W., 1995, Ground-water data-collection protocols and procedures for the National Water-Quality Assessment Program—Collection and documentation of water-quality samples and related data: U.S. Geological Survey Open-File Report 95–399, 113 p.

Kramer, D.A., 2001, Magnesium recycling in the United States in 1998: U.S. Geological Survey, Open-File Report

01–166, accessed December 3, 2007, from URL *http://pubs.usgs.gov/of/2001/of01-166.*

Lapham, W.L., Wilde, F.D, and Koterba, M.T., 1995, Ground-water data-collection protocols and procedures for the National Water-Quality Assessment Program—Selection, installation, and documentation for wells, and collection of related data: U.S. Geological Survey Open-File Report 95–398, 69 p.

Lapham, W.L., Hamilton, P.A., and Myers, D.N., 2005, National Water-Quality Assessment Program—Cycle II-Regional assessments of aquifers: U.S. Geological Survey Fact Sheet 2005–3013, 4 p.

Massart, D.L., and Kaufman, L., 1983, The interpretation of analytical chemical data by the use of cluster analysis: John Wiley and Sons, New York, 237 p.

Meinzer, O.E., 1923, The occurrence of ground water in the United States, with a discussion of principles: U.S. Geological Survey Water-Supply Paper 489, 321 p.

Mickelson, D.M., Clayton, L., Fullerton, D.S., and Borns, H.W., Jr., 1983, The Late Wisconsin glacial record of the Laurentide Ice Sheet in the United States *in* Porter, S.C., ed., Late Quaternary environments of the United States, the late Pleistocene: Minneapolis, University of Minnesota Press, v. 1, p. 3–37.

Minnesota Pollution Control Agency, 1999, Barium, beryllium, calcium, magnesium, and strontium in Minnesota's ground water: Minnesota Pollution Control Agency, Ground Water Monitoring and Assessment Program, 2 p.

National Park Service, 1999, State surficial geology map of Alaska: National Park Service Alaska Support Office, digital data, 1:1,584,000 scale.

Ott, R.L., 1993, An introduction to statistical methods and data analysis, (4th ed.): Belmont, Calif., Wadsworth Publishing Company, 1,051 p. plus appendices.

Pacheco, Fernando Antonio Leal, 1999, Finding the number of natural clusters in groundwater data sets using the concept of equivalence class: Computers and Geosciences, vol. 24, no. 1, 29 January 1998, p. 7–15.

Porter, S.C., Pierce, K. L., and Hamilton, T. D., 1983, Late Pleistocene mountain glaciation in western United States *in* Porter, S.C., ed., Late Quaternary environments of the United States, the late Pleistocene: Minneapolis, University of Minnesota Press, v. 1, p. 71–111.

Rees, W.G., 2001, Physical principals of remote sensing, second edition: Cambridge, United Kingdom, Cambridge University Press, 343 p.

Risch, M.R., and Robinson, B.A., 2001, Use of borehole and surface geophysics to investigate ground-water quality near a road-deicing salt-storage facility, Valparaiso, Indiana: U.S. Geological Survey Water-Resources Investigations Report 2000–4070, 63 p.

Romesburg, H.C., 1984, Cluster analysis for researchers: Belmont, Calif., Wadsworth, Inc., 334 p.

Soller, D.R., and Packard, P.H., 1998, Digital representation of a map showing the thickness and character of Quaternary sediments in the glaciated Untied States east of the Rocky Mountains: U.S. Geological Survey Digital Data Series DDS–38.

Squillace, P.J., and Price, C.V., 1996, Urban land-use study plan for the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 96–217, 19 p.

Starkey, H.C., 1982, The role of clays in fixing lithium: U.S. Geological Survey Bulletin 1278–F, 20 p.

Stetzenbach, K.J., Farnham, I.M., Hodge, V.F., and Johannesson, K.H., 1999, Using multivariate statistical analysis of groundwater major cation and trace element concentrations to evaluate groundwater flow in a regional aquifer: Hydrological Processes, v. 13, no. 17, Dec. 1999, p. 2,655–2,673.

Stetzenbach, K.J., Hodge, V.F., Guo, C., Farnham, I.M, and Johannesson, K.H., 2001, Geochemical and statistical evidence of deep carbonate groundwater within overlying volcanic rock aquifers/aquitards of southern Nevada, USA: Journal of Hydrology, v. E 243, no. 3–4, March 2001, p. 254–271.

Tesoriero, A.J., Loffler, F.E., and Liebscher, H., 2001, Fate and origin of 1,2 dichloropropane in an unconfined shallow aquifer: Environmental Science Technology, 2001, v. 35, no. 3, pp. 455–461.

Tesoriero, A.J., Liebscher, H., and Cox, S.E., 2000, Mechanism and rate of denitrification in an agricultural watershed: Electron and mass balance along groundwater flow paths: Water Resources Research, v. 36, no. 6, p. 1,545–1,559.

The Geological Institute, 1984, Dictionary of geological terms, (3d ed.), Bates, R.L., and Jackson, J.A., eds.: New York, Anchor Books, 571 p.

Thomas, H.E., 1952, Ground-water regions of the United States—their storage facilities, v. 3 *of* The physical and economic foundation of natural resources: U.S. 83rd Congressional House Committee on Interior and Insular Affairs, p. 3–78.

Thomas, M.A., 2007, The association of arsenic with redox conditions, depth, and ground-water age in the glacial aquifer system of the Northern United States: U.S. Geological Survey Scientific Investigations Report 2007–5036, 26 p.

Thomas, M.A., 2003, Arsenic in Midwestern glacial deposits—Occurrence and relation to selected hydrogeologic

and geochemical factors: U.S. Geological Survey Water-Resources Investigations Report 03–4228, 36 p.

Thyne, G., Güler, C., Poeter, E., 2004, Sequential analysis of hydrochemical data for watershed characterization: Ground Water, v. 42, no. 5, September 2004, p. 711–723.

U.S Bureau of the Census, 1991, Census of population and-housing, 1990—Public Law 94–171 data: U.S. Bureau of the Census, digital data.

U.S. Department of Agriculture, 1994, State Soil Geographic Database: Natural Resources Conservation Service, 1:250,000 scale digital data, accessed April 11, 2008, from URL: *http://soils.usda.gov/survey/geography/statsgo/*

U.S. Geological Survey, 2001, National Elevation Data Set, digital data, accessed December 3, 2007, from URL http://ned.usgs.gov/

U.S. Geological Survey Geologic Names Committee, 2007, Divisions of geologic time—Major chronostratigraphic and geochronologic units: U.S. Geological Survey Fact Sheet 2007–3015, 2 p.

Verdin, K.L., and Greenlee, S.K., 1996, Development of continental scale digital elevation models and extraction of hydrographic features *in* Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January 21–26, 1996: Santa Barbara, Calif., National Center for Geographic Information and Analysis, accessed December 3, 2007, from URL *http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/papers/santafe3.html*.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., and Van Driel, N., 2001, Completion of the 1990s National land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources: Photogrammetric Engineering and Remote Sensing 67, p. 650–662.

Wahrer, M.A., Long, D.T., and Lee, R.W., 1996, Selected geochemical characteristics of ground water from the glaciofluvial aquifer in the central Lower Peninsula of Michigan: U.S. Geological Survey Water-Resources Investigations Report 94–4017, 21 p.

Waitt and Thorson, 1983, Late Quaternary environments of the United States, the late Pleistocene *in* Porter, S.C., ed.: Late Quaternary environments of the United States, the late Pleistocene: Minneapolis, University of Minnesota Press, v. 1, 407 p.

Warner, K.L., 2000, Analysis of nutrients, selected inorganic constituents, and trace elements in water from Illinois community-supply wells, 1984–91: U.S. Geological Survey Water-Resources Investigations Report 99–4152, p. 22–23.

Warner, K.L. and Arnold, T.L., 2005, Framework for regional synthesis of water-quality data for the glacial aquifer system in the United States: U.S. Geological Survey Scientific Investigations Report 2005–5223, revised May 2006, 6 p.

Warner, K.L., Martin, A.M., Jr., and Arnold, T.L., 2003, Arsenic in Illinois ground-water—Community and private supplies: U.S. Geological Survey Water-Resources Investigations Report 03–4103, 12 p.

Westjohn, D.B., Weaver, T.L., Zacharias, K.F., 1994, Hydrogeology of Pleistocene glacial deposits and Jurassic “red beds” in the central lower peninsula of Michigan: U.S. Geological Survey Water-Resources Investigations Report 93–4152, 14 p.

Wolock, D.M, 1997, STATSGO soil characteristics for the conterminous United States: U.S. Geological Survey Open-File Report 97–656, 1:250,000 scale digital data, accessed December 3, 2007, from URL *http://water.usgs.gov/lookup/getspatial?muid.*

Wolock, D.M., 2003a, Hydrologic landscape regions of the United States: U.S. Geological Survey Open-File Report 2003–145, digital data, accessed December 3, 2007, from URL *http://water.usgs.gov/GIS/metadata/usgswrd/XML/hlrus.xml*.

Wolock, D.M., 2003b, Estimated mean annual natural groundwater recharge in the conterminous United States: U.S. Geological Survey Open-File Report 03–311, digital data, accessed December 3, 2007, from URL *http://water.usgs.gov/GIS/metadata/usgswrd/XML/rech48grd.xml.*

Wolock, D.M., Fan, J., and Lawrence, G.B., 1997, Effects of basin size on low-flow stream chemistry and subsurface contact time in the Neversink River watershed, New York: Hydrological Processes, v. 11, p. 1,273–1,286.

Wolock, D.M., Winter, T.C., and McMahon, G., 2004, Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses: Springer Science and Business Media, Inc., Environmental Management, v. 34, suppl. 1, p. S71–S88.

Glossary

(Terms in definitions that are defined elsewhere in the Glossary are shown in boldface within the definition. Most definitions are described as they apply to this report)

A

anthropogenic vulnerability Vulnerability to contaminants resulting from human activities.

B

binary variable A variable containing only two values — 0 and 1 — to numerically indicate absence or presence of a categorical characteristic. Binary variables commonly are used in statistical analysis when the analysis method requires a continuous variable.

C

Ca Calcium.

censored data Data results from the laboratory's analysis of water samples that have chemical concentrations too low for the laboratory equipment to reliably detect and quantify.

censored level The concentration value below which the laboratory cannot reliably detect and quantify the chemical.

Cl Chloride.

cluster Groups of similar data resulting from **cluster analysis**.

cluster analysis A method of **unsupervised classification** that groups data (variables) such that within group differences are minimized while between group differences are maximized. Cluster analysis can be used as a descriptive tool, a data simplification tool, or a relation identification tool (Rees, 2001; Hair and others, 1998).

CO_3 Carbonate.

conceptual framework A preliminary classification scheme that identified the most general and important regional-scale characteristics of the glacial aquifer system (Warner and Arnold, 2005). Four framework areas were defined East, Central, West-central, and West.

confidence map A map that shows locations where **supervised classifications** have a higher or lower probability of being correct based on the group characteristics defined in the **signature file** for the **training sample**.

D

E

Euclidean distance The root sum-of-squares of differences for all of the variables in the data set.

F

F Fluoride.

G

GIS Geographic Information System.

glacial aquifer system As defined in this report, the glacial aquifer system consists of all unconsolidated geologic material above bedrock that lies on or north of the line of maximum glacial advance within the United States.

H

HCO_3 Bicarbonate.

hydrochemical region A categorization of the aquifer system on the basis of physical environmental, and measured chemical characteristics. The five hydrochemical regions described in this report have distinct water-quality characteristics based on **major-** and **minor-ion** concentrations and have environmental characteristics that fit into the **conceptual framework** used to design the glacial aquifer studies.

I

intrinsic susceptibility A function of the hydraulic properties of the aquifer that affect how easily water enters and moves through the aquifer system, and of the sources of water (Focazio and others, 2002).

J

K

K Potassium.

Kruskal-Wallis Rank Sum test A statistical test used to evaluate differences among groups. **Statistically significant** differences are indicated if the probability (p-value) resulting from the test is less than 0.05 (Helsel and Hirsch, 1995).

L

logistic regression analysis A statistical method of identifying the importance of categorical data for predicting a **binary variable**.

M

major ions Sodium, calcium, magnesium, chloride, bicarbonate, and sulfate, which generally make up about 90 percent of the dissolved ions in ground water (Hem, 1992).

MAR Midwestern Agricultural Region.

Maximum Likelihood Classification (MLC) A geostatistical method of **supervised classification** that is commonly used in remote sensing (Rees, 2001). Statistical characteristics of the classification groups are known in some locations before the analysis begins. This set of known characteristics and locations is called a "**training sample**". The MLC algorithm classifies spatial areas based on the maximum probability that an area has similar characteristics as the group characteristics defined in the **training sample**.

MCFR Mountain and Coastal Forested Region.

median The data value where 50 percent of the data lie above and below it when all the data are ranked in numerical order.

medoid The data point closest to the center of the **cluster**, in multidimensional space.

Mg Magnesium.

minor ions Potassium and fluoride.

MLC Maximum Likelihood Classification.

multivariate statistics Statistical procedures used for simultaneously analyzing multiple variables in relation to each other (Kachigan, 1986).

N

Na Sodium.

natural contaminants In this report, those chemical constituents in ground water that potentially could have been added to the natural ground-water system from weathering and breakdown of glacial deposits.

natural vulnerability Dependant, in part, on the **intrinsic susceptibility** of the aquifer and is a function of the geochemical properties of the geologic materials composing the aquifer system and the characteristics and availability of any potential sources of contaminants (Focazio and others, 2002).

NAWQA National Water-Quality Assessment Program.

network A group of wells from which water samples were collected during a short time period (weeks to months) for a specific purpose.

NGLFR Northern and Great Lakes Forested Region.

NO_2 plus NO_3 as N Nitrite plus nitrate as nitrogen. The concentration of nitrogen present in the nitrate and nitrite anions (negatively charged ions), expressed as elemental nitrogen.

O

P

Q

R

Regression Tree Analysis (Decision Tree Analysis) A statistical method used to identify sequential splits in the values of a set of variables that yield a classification into discrete groups. It is commonly used for examining characteristics of categorical data instead of **logistic regression** (Hair and others, 1998). The output of this analysis is a decision flow chart (dendrogram) that shows the sequential splits.

S

signature file (index file) A text file that contains **multivariate statistics** that define group characteristics for the **training sample**. It is used in a **supervised classification** method, such as **Maximum Likelihood Classification**, to classify the data to fit predefined group characteristics as closely as possible.

silhouette width The silhouette width measures the "goodness" of the **cluster analysis** classifications and is calculated using the following equation
Silhouette width = (B-A)/(the greater of either A or B),
where
B = the minimum dissimilarity of an object to those objects in other clusters, and
A = the average dissimilarity of the object in its own cluster.
Silhouette widths greater than zero indicate well-classified objects, whereas widths less than zero indicate poorly classified objects, and widths near zero indicate an object's classification lies between clusters. Silhouette widths closer to one are better classified than widths closer to zero.

SiO_2 Silica.

SO_4 Sulfate.

Spearman rank correlation coefficients Coefficients calculated to evaluate pair-wise linear relations between select variables. In this report, they are **statistically significant** if the probability (p-value) of the coefficient is less than 0.05. Statistical significance indicates a linear relation between the variables. If the coefficient is negative, the value of one variable increases as the other variable decreases. If the coefficient is positive, both variables increase in relation to each other (Helsel and Hirsch, 1995).

statistically significant For this report, a statistically significant result is defined as a probability (p-value) less than 0.05, indicating that the statistic is significant at the 95% confidence level. The probability value of less than 0.05 indicates there is less than 5 percent probability the statistic could have been obtained through random chance and the tested hypothesis could be incorrectly rejected.

study unit A term used by **NAWQA** to indicate a geographic area or river basin of an individual **NAWQA** study. For detailed information about the **NAWQA** study units, go to web page http://water.usgs.gov/nawqa/studies/study_units.html.

STATSGO U.S. Department of Agriculture's State Soil Geographic database.

subsurface contact time The time ground water is in contact with geologic materials. This report uses subsurface contact time as modeled by Wolock and others (1997) to interpret how ion concentrations may vary among **hydrochemical region**s but was not directly used in analysis and mapping of **hydrochemical regions**.

supervised classification A statistical method for classifying data based on pre-defined group characteristics. **Maximum Likelihood Classification** is a method of supervised classification.

T

training sample A set of pre-defined characteristics and specific locations that is used with a **supervised classification** method.

U

UIR Urban-Influenced Region.

unsupervised classification A statistical method for grouping data based on the statistical distribution of the data rather than on pre-defined group characteristics. The goal is to identify groups of data that are more similar to each other than they are to data in any other group—to minimize within group differences while maximizing between group differences. **Cluster analysis** is a method of unsupervised classification.

USGS United States Geological Survey.

V

W

WAGR Western Agriculture and Grassland Region.

water type Describes the chemical composition (characteristics) of a water sample based on the relative concentrations of **major** and **minor ions** (Hiscock, 2005).

X

Y

Z

z-score Is used to standardize measurement scales and magnitudes among the variables being analyzed. The variables are standardized by subtracting the mean and dividing by the standard deviation for each variable (Kachigan, 1986). This type of standardization has a mean of zero (0) and standard deviation of one (1) which allows the vari-

ables to be examined on a similar scale (Hair and others, 1998).

z-test for equal proportions The z-test for equal proportions (Ott, 1993) was used to compare the proportion of samples assigned to each **hydrochemical region** with the proportion of wells assigned to each corresponding **cluster**. A significant z-score would indicate the proportion of samples in the cluster was different from the proportion of samples in the **hydrochemical region**.

Appendix A—Tables

Table A1. Reference, methods, variables, significant findings, and scale of selected previous studies to group, map, or define water quality.

Reference for previous study	Methods	Variables	Summary of result or significant finding	Study area scale
Meinzer (1923)	Not specified	Rock units and physiography	Map of 21 ground-water provinces	Regional
Thomas (1952)	Not specified	Rock units and physiography	Map of 10 ground-water provinces (refined from Meinzer, 1923)	Regional
Heath (1984)	Not specified	Mineral composition of rocks, structure of water-bearing openings, arrangement of recharge and discharge areas, and transmissibility	Map of 15 ground-water regions	Regional
Wolock and others (2004) and Wolock (2003a)	Cluster analysis	Bedrock permeability, percent sand, relief, percent of total flatland, percent flatland in upland, percent flatland in lowland, precipitation, and potential evapotranspiration	Map of hydrologic landscape regions based on surface-water drainage basins. Hydrologic landscape regions reflect basic hydrologic processes that are thought to affect water quality. An “hydrologic landscape region” is defined as an upland adjacent to a lowland separated by a valley side.	Regional
Güler and Thyne (2004)	Cluster analysis, kriging, and inverse geochemical modeling	Major ions, pH, specific conductivity, silica	Clusters are related to geologic source material, recharge areas, and flow paths where total dissolved solids concentration increased along a flow path.	Local
Güler and others (2002)	Graphical techniques and multivariate methods	Bicarbonate, calcium, chloride, magnesium, pH, potassium, silica, sodium, and sulfate	Major ions were used as natural tracers to delineate flow paths in aquifers. Waters with similar characteristics often have similar history, recharge areas, climate, mineralogy, and residence time. Spatial variability was determined to be more important than temporal variability.	Local
Pacheco (1999)	Cluster analysis, principal components analysis	Group environmental characteristics and water-quality data (bicarbonate, calcium, chloride, magnesium, nitrate, potassium, sodium, and sulfate)	When chemistry is similar, clusters will be close or overlapping in multidimensional space. When clusters are completely distinct, they are likely to indicate water quality that is controlled by completely different processes. Although significant, differences in water quality in this study area were too narrow to detect with principal components analysis. Water chemistry is related to minerals, climate variations, and land use.	Local
Stetzenbach and others (2001)	Principal components analysis	Calcium, magnesium, potassium, sodium, and 24 trace elements	Groups of data collected in close proximity plot together on principal components graphs. Ground water inherits chemical signatures from rocks with which they interact.	Local
Stetzenbach and others (1999)	Principal components analysis	Major cations and trace elements	Oxyanion-forming trace elements may behave conservatively in some ground-water systems and some rare-earth elements may reflect the rocks through which the water has flowed. Geologic source materials affect ground-water chemistry.	Local
Thyne and others (2004)	Principal components analysis	Bicarbonate, calcium, chloride, magnesium, nitrate, pH, and sulfate	Similar water chemistry indicates similar flow paths and sources.	Local

Table A2. Types of well networks used for this study, well depth, and casing diameter.

Network type	Number of wells	Well depth range, in feet (median)	Casing diameter range, in inches (median)
Agricultural row crop or orchard land use	574	8–220 (25)	2–20 (2)
Forest land use	55	7–54 (17)	2–2 (2)
Urban residential and/or commercial land-use	322	8–142 (26)	2–12 (2)
Drinking water	550	8.38–425 (75)	1–45 (6)
Flow-path study	143	0.8–153 (22)	0–6 (2)
Special study	68	8–264 (51)	2–48 (4)
Reference	4	15.5–52 (26)	2–2 (2)

Table A3. Characteristics of data used in analyses for this report: number of samples with detections, estimated samples, and censored samples.

[µg/L, micrograms per liter; na, not applicable; mg/L, milligrams per liter; N, nitrogen; °C, degrees Celsius; $CaCO_3$, calcium carbonate]

Parameter name	Reporting unit	Number of samples collected	Most frequent censored level	Number of samples censored at most frequent level	Substituted censored level	Samples less than the detection limit		Samples with estimated values	
						Number of samples	Range of concentrations	Number of samples	Range of concentrations
Arsenic; filtered water	µg/L	851	1	232	1	336	0.18–2	69	0.09–0.90
Barium; filtered water	µg/L	863	1	1	1	1	1	0	na
Boron; filtered water	µg/L	587	12	19	12	19	7–12	44	12–14.95
Bromide; filtered water	mg/L	1,640	0.01	136	0.01	170	0.01–0.15	42	0.01–0.03
Calcium; filtered water	mg/L	1,666	127.6	1	na	1	127.6	0	na
Chloride; filtered water	mg/L	1,663	.1	3	.1	4	0.1–0.29	0	na
Dissolved oxygen; unfiltered water	mg/L	1,645	.1	102	.1	110	0–0.9	12	1.48–7.4
Fluoride; filtered water	mg/L	1,665	.1	471	.1	578	0.1–0.2	41	0.06–0.16
Iron; filtered water	µg/L	1,667	10	305	10	523	3–30	62	4.25–9.71
Lithium; filtered water	µg/L	566	6	39	6	59	0.3–11.7	51	0.16–5.89
Magnesium; filtered water	mg/L	1,667	32.37	1	na	1	32.37	0	na
Manganese; filtered water	µg/L	1,668	1	162	1	224	0.1–4	21	0.06–0.15
Molybdenum; filtered water	µg/L	863	1	169	1	254	0.2–30	42	0.11–0.33
Nickel; filtered water	µg/L	863	1	94	1	167	0.03–30	3	0.04–0.06
Nitrite plus nitrate; filtered water	mg/L as N	1,645	.05	460	.05	508	0.05–0.06	16	0.02–10.86
Potassium; filtered water	mg/L	1,667	.01	3	.1	3	0.1	0	na
Silica; filtered water	mg/L	1,667	14.66	1	na	1	14.66	0	na
Sodium; filtered water	mg/L	1,667	69.83	1	na	1	69.83	0	na
Strontium; filtered water	µg/L	566	.2	1	.2	1	0.2	0	na
Sulfate; filtered water	mg/L	1,665	.1	40	.1	48	0.1–0.31	0	na
Water temperature	°C	1,662	na	na	na	35	11.72–17.89	90	9.62–31
pH; unfiltered water; field determined	standard units	1,674	na	na	na	0	na	1	6.5
Alkalinity; filtered water; field determined	mg/L as $CaCO_3$	1,452	na	na	na	0	na	0	na
Bicarbonate (calculated for this report)	mg/L as $CaCO_3$	1,454	na	na	na	0	na	0	na

Table A4. Cluster medoid, number of wells, dissimilarity and separation measures, and percent poorly and well-classified wells.

[%, percent]

Identification number of well nearest the center of the cluster (medoid)	Cluster number	Number of wells in cluster	Maximum dissimilarity within the cluster	Average dissimilarity within the cluster	Diameter of cluster in 8-dimensional space	Separation from other clusters	Average silhouette width	Percent of wells that are well-classified (silhouette width greater than 0)	Percent of wells that are poorly classified (silhouette width less than 0)
401203089203201	1	286	6.60	2.38	9.42	0.37	0.08	79%	21%
423958083331001	2	237	6.38	2.12	9.36	.41	.24	97%	3%
452324093541601	3	286	5.68	2.14	7.98	.42	.22	100%	0%
421303097011601	4	241	8.80	2.41	11.63	.37	.14	88%	12%
470732122252801	5	266	8.43	2.73	11.66	.65	.18	83%	17%

Table A5. Results of Kruskal-Wallis rank sum test for differences among clusters. The NULL hypothesis is that the difference in means of the ranks is equal to zero; the alternate hypothesis is that the difference in means of the ranks is not equal to zero (p-value < 0.05).

[<, less than; bold text indicates rejection of the NULL hypothesis (p-value was greater than 0.05); A, B, C, D, E statistical population where the same letter across the row indicates the populations are not significantly different from each other (p-value < 0.05)]

Constituent	Cluster pairwise comparison p-value										Statistical population of cluster				
	1	1	1	1	2	2	2	3	3	4					
	2	3	4	5	3	4	5	4	5	5	1	2	3	4	5
Calcium	<0.0001	<0.0001	**0.23**	<0.0001	<0.0001	<0.0001	<0.0001	<.0001	<0.0001	<0.0001	B	A	C	B	D
Chloride	<.0001	**.12**	.002	**.97**	<.0001	<.0001	<.0001	**.09**	**.17**	0.01	C	A	BC	B	C
Fluoride	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.007	<.0001	B	C	E	A	D
Magnesium	**.96**	<.0001	.0002	<.0001	<.0001	.006	<.0001	<.0001	<.0001	<.0001	A	A	C	B	D
Potassium	<.0001	**.14**	<.0001	**.08**	<.0001	<.0001	<.0001	<.0001	.0034	<.0001	CD	B	D	A	C
Sodium	<.0001	<.0001	<.0001	.002	<.0001	.001	<.0001	<.0001	<.0001	<.0001	C	A	E	B	D
Sulfate	<.0001	.0001	<.0001	<.0001	<.0001	.004	<.0001	<.0001	<.0001	<.0001	C	B	D	A	E
Bicarbonate	.03	<.0001	.012	<.0001	<.0001	**.87**	<.0001	<.0001	<.0001	<.0001	A	B	C	B	D

Table A6. Regression tree for analysis of ion concentrations to characterize the water-quality cluster groups (fig. 7).

[>, greater than; <, less than; --, variable not tested for the terminal branch; bold text indicates probability of final classification]

Test of water-quality characteristic[1]								Cluster group[2]	Terminal branch number[3]	Number of wells classified by the terminal branch	Proportion of wells in the cluster group classified by the terminal branch	Probability of correct classification in the clusters[4]				
Bicarbonate	Calcium	Fluoride	Chloride	Potassium	Magnesium	Sodium	Sulfate					Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
>128.5	--	>0.17	<37.3	<2.4	>16.5	--	--	1	57	240	0.87	**0.90**	0	0.03	0.06	0
>128.5	--	>0.17	<45.2	2.4–4.0	--	--	<14.4	1	120	18	.06	**.67**	0	.17	.11	.06
>128.5	--	>0.51	>37.3	<2.4	--	--	--	1	59	12	.04	**.58**	0.08	0	.33	0
>128.5	--	>0.17	<45.2	2.4–4.0	--	<5.6	>14.4	1	242	7	.03	**1.0**	0	0	0	0
>128.5	--	<0.17	>45.6	--	--	--	--	2	13	129	.58	0	**.95**	.04	0	.02
>128.5	--	0.17–0.51	>37.3	<2.4	--	--	--	2	58	41	.19	.15	**.85**	0	0	0
>128.5	--	0.17–0.35	>45.2	>2.4	--	--	--	2	62	36	.16	0	**.97**	0	.03	0
>128.5	--	<0.17	18.2– 45.6	>3.1	--	--	--	2	51	15	.07	0	**.80**	.07	.13	0
>128.5	--	<0.17	<45.6	<3.1	--	--	--	3	24	237	.71	.08	.06	**.83**	0	.03
<128.5	>35.1	--	--	--	--	--	--	3	5	48	.14	0	.13	**.63**	.02	.23
>128.5	--	>0.17	<37.3	<2.4	<16.5	<26.0	--	3	112	28	.08	.39	0	**.61**	0	0
>128.5	--	<0.17	<18.2	3.1–7.6	--	--	--	3	100	22	.07	0	.05	**.73**	.23	0
>128.5	--	>0.17	<45.2	>4.0	--	--	--	4	61	136	.61	0	0	0	**1.**	0
>128.5	--	>0.17	<45.2	2.4–4.0	--	>5.6	>14.4	4	243	60	.27	.07	.03	0	**.90**	0
>128.5	--	>0.35	>45.2	>2.4	--	--	--	4	63	18	.08	.11	.28	0	**.61**	.00
>128.5	--	<0.17	<18.2	>7.6	--	--	--	4	101	10	.04	0	0	0	**.90**	.10
<128.5	<35.1	--	--	--	--	--	--	5	4	251	.97	0	0	.04	0	**.96**
>128.5	--	>0.17	<37.3	<2.4	<16.5	>26.0	--	5	113	8	.03	.13	.13	0	.13	**.63**

[1] The tests indicate if the water-quality characteristics are satisfied for that branch of the regression tree.

[2] The cluster group is the final classification based on the test of water-quality characteristics concentrations.

[3] A terminal branch is one of the ending branches of the regression tree in which a classification is made based on the highest probability of a correct classification.

[4] Final terminal branch classifications are determined on the highest probability of a correct classification. A well may have a probability of being classified into more than one cluster.

Table A7. Spearman correlations of clusters and environmental characteristics within 1,640 feet of the well.

[ns, indicates an insignificant correlation; N, indicates negative significant correlation (p-value < 0.05); P, indicates positive significant correlation (p-value < 0.05); GIS, Geographic Information System; na, correlation not calculated]

Environmental characteristic	Cluster				
	1	2	3	4	5
Average annual precipitation (1980–97)	ns	ns	ns	N	P
Average annual temperature (1980–97)	P	N	N	P	P
Land surface slope	N	ns	ns	ns	P
Urban land use (1990s)	N	P	N	N	P
Agriculture land use (1990s)	P	N	ns	P	N
Forest land use (1990s)	N	N	P	N	P
Vertical soil permeability	N	ns	P	N	P
Soil clay content	P	ns	N	P	N
Percent coarse-grained surficial deposits	N	P	P	N	P
Percent fine-grained surficial deposits	P	N	N	ns	ns
Estimated annual natural ground-water recharge potential	N	ns	P	N	P
Type of surficial deposit (only used as layer in GIS; not used to test significance as a linear predictor for clusters)	na	na	na	na	na

Table A8. Regression tree for analysis of environmental characteristics (within 1,640 feet of the well) to characterize the water-quality cluster groups (fig. 9).

[in/yr, inches per year; in., inches; °F, degrees Fahrenheit; %, percent; in/h, inches per hour; <, less than; >, greater than;--, indicates variable not tested for the terminal branch; bold text indicates probability of final classification]

Test of environmental characteristic[1]												Cluster group[2]	Terminal branch number[3]	Number of wells classified by the terminal branch	Proportion of wells in the cluster group classified by the terminal branch	Probability of correct classification in the clusters[4]				
Estimated annual natural ground-water recharge (in/yr)	Average annual precipitation (in.)	Average annual temperature (°F)	Agricultural land use (%)	Urban land use (%)	Forest land use (%)	Vertical permeability (in/h)	Soil clay content (%)	Land surface slope (%)	Coarse-grained surficial material (%)	Fine-grained surficial material (%)	Type of surficial deposit					Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
<7.0	>32.4	--	--	<9.0	--	<1.9	--	--	--	--	--	1	20	270	0.82	**0.61**	0.11	0.12	0.15	0.01
3.5–7.0	>32.4	--	<73.1	<9.0	--	>1.9	>9.0	--	--	--	--	1	345	20	.06	**.55**	.20	.20	0	.05
<7.0	>32.4	<49.9	>78.2	<9.0	<6.0	>1.9	>9.0	--	--	--	alluvial ice contact, glacial till, glaciofluvial	1	698	16	.05	**.50**	.13	.13	.25	0
<7.0	<20.7	41.6–47.9	>70.3	--	--	--	--	--	--	--	--	1	134	12	.04	**.58**	.17	0	.25	0
<7.0	23.7–32.4	<47.9	--	--	<9.0	<10.1	--	--	--	--	--	1	68	11	.03	**.82**	.00	.18	0	0
<7.0	>32.4	--	--	>9.0	--	--	--	--	--	--	--	2	11	202	.83	.19	**.63**	.08	.07	.02
>7.0	--	<47.7	--	>14.6	<41.7	--	>5.8	--	--	--	--	2	53	34	.14	0	**.53**	.32	0	.15
<7.0	>32.4	<49.9	>78.2	<9.0	>6.0	>1.9	>9.0	--	--	--	alluvial ice contact, glacial till, glaciofluvial	2	699	7	.03	0	**.57**	.43	0	0
<7.0	23.7–32.4	<47.9	--	--	--	>10.1	--	--	--	--	--	3	35	60	.25	.03	.08	**.87**	0	.02
<7.0	>32.4	--	--	<9.0	--	>1.9	<9.0	--	--	--	--	3	42	51	.21	.10	.06	**.53**	.02	.29
>7.0	--	<47.7	--	<14.6	<41.7	--	>5.8	--	--	--	--	3	52	38	.16	.08	.13	**.74**	.03	.03
<7.0	23.7–32.4	<47.9	--	--	>9.0	<10.1	--	--	--	--	--	3	69	28	.12	.29	.07	**.61**	0	.04
>7.0	--	>47.7	--	--	<38.8	3.9–5.3	--	>0.9	--	--	--	3	121	16	.07	0	0	**.75**	.06	.19
>7.0	--	45.5–47.7	--	--	>41.7	--	>5.8	--	--	--	--	3	55	13	.05	0	.15	**.69**	0	.15
<7.0	<23.7	41.6–47.9	<70.3	--	--	--	--	--	--	--	--	3	66	12	.05	.33	.08	**.42**	.17	0
<7.0	>32.4	--	73.1–78.2	<9.0	--	>1.9	>9.0	--	--	--	--	3	173	8	.03	0	0	**.75**	0	.25
<3.5	>32.4	--	<73.1	<9.0	--	>1.9	>9.0	--	--	--	--	3	344	7	.03	0	.29	**.71**	0	0
<7.0	>32.4	<49.9	>78.2	<9.0	--	>1.9	>9.0	--	--	--	lacustrine, alluvial, eolian	3	348	6	.03	0	.17	**.50**	0	.33

Table A8. Regression tree for analysis of environmental variables (within 1,640 feet of the well) to characterize the water-quality cluster groups.—Continued

[in/yr, inches per year; in., inches; °F, degrees Fahrenheit; %, percent; in/h, inches per hour; <, less than; >, greater than; --, indicates variable not tested for the terminal branch; bold text indicates probability of final classification]

Test of environmental characteristic[1]																Probability of correct classification in the clusters[4]				
Estimated annual natural ground-water recharge (in/yr)	Average annual precipitation (in.)	Average annual temperature (°F)	Agricultural land use (%)	Urban land use (%)	Forest land use (%)	Vertical permeability (in/h)	Soil clay content (%)	Land surface slope (%)	Coarse-grained surficial material (%)	Fine-grained surficial material (%)	Type of surficial deposit	Cluster group[2]	Terminal branch number[3]	Number of wells classified by the terminal branch	Proportion of wells in the cluster group classified by the terminal branch	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
<7.0	<32.4	>47.7	--	--	--	--	--	--	--	--	--	4	9	138	0.67	0.07	0.06	0.01	**0.85**	0.01
<7.0	<23.7	<41.6	--	--	--	--	--	--	--	--	--	4	32	42	.20	.05	.12	.10	**.74**	0
<7.0	20.7–23.7	41.6–47.9	>70.3	--	--	--	--	--	--	--	--	4	135	16	.08	.31	0	0	**.69**	0
<7.0	>32.4	>49.9	>78.2	<9.0	--	>1.9	>9.0	--	--	--	--	4	175	11	.05	.09	0	.27	**.64**	0
>7.0	--	>47.7	--	--	--	>5.3	--	--	--	--	--	5	31	105	.49	0	.04	.10	.01	**.86**
>7.0	--	<47.7	--	--	--	--	<5.8	--	--	--	--	5	12	38	.18	0	.16	.03	.03	**.79**
>7.0	--	>47.7	--	--	--	<3.9	--	--	--	--	--	5	14	34	.16	.03	0	0	.06	**.91**
>7.0	--	<45.5	--	--	>41.7	--	>5.8	--	--	--	--	5	54	21	.10	0	0	.24	0	**.76**
>7.0	--	>47.7	--	--	>38.8	3.9–5.3	--	--	--	--	--	5	61	9	.04	0	.11	0	0	**.89**
>7.0	--	>47.7	--	--	<38.8	3.9–5.3	--	<0.9	--	--	--	5	120	9	.04	0	0	.22	0	**.78**

[1] The tests indicate if the environmental characteristics are satisfied for that branch of the regression tree.

[2] The cluster group is the final classification based on the test of environmental characteristics.

[3] A terminal branch is one of the ending branches of the regression tree in which a classification is made based on the highest probability of a correct classification.

[4] Final terminal branch classifications are determined on the highest probability of a correct classification. A well may have a probability of being classified into more than one cluster.

Table A9. The signature file used for maximum likelihood classification of hydrochemical regions in the glacial aquifer system.

```
#  Signatures  Produced by ClassSig from
#  Class-Grid          temp\raster2
#  and Stack           temp\tmp0000
#  Number of           selected   grids
/* 9
#  Layer-Number    Grid-name
/* 1        prcpclp   --        average annula precipitation
/* 2        tempclp   --        average annual temperature
/* 3        nlcde     --        land cover (1990s)
/* 4        slope500m           --        percent land surface slope
/* 5        pergclp   --        vertical permeability
/* 6        clayclp   --        percent clay in soil
/* 7        surftxt2  --        texture of surficial deposits
/* 8        surfdep2  --        type of surficial deposits
/* 9        rchclp    --        recharge

#  Type    Number of classes  Number of layers    Number of parametric layers
   1       5          9          9
#  ==================================================================

#  Class ID           Number of Cells     Class Name
   1       262
#  Layers 1           2          3          4          5          6          7          8          9
#  Means
           90.04635 9.15681   7.23282  0.82109   2.34872  22.0534   1.45802  7.72137   3.45074
#  Covariance
   1       354.79029          21.90598 -1.45289  -1.60016 -22.63965          94.94445 -5.26156  -3.58408  19.48187
   2       21.90598 4.12829   0.17267  -0.10613  -3.26526 10.00033 -0.35358  -0.63549  0.31036
   3       -1.45289 0.17267   3.40536  -0.10066  -0.11184 0.97841   -0.03425  0.08045   -0.48279
   4       -1.60016 -0.10613  -0.10066 0.9287    -0.27604 -0.78155  0.09555   0.00798   0.31532
   5       -22.63965          -3.26526 -0.11184  -0.27604 7.72476   -19.01628           0.51892   0.52747   -0.51358
   6       94.94445 10.00033  0.97841  -0.78155  -19.01628          74.24506  -1.75745  -3.84587  0.25027
   7       -5.26156 -0.35358  -0.03425 0.09555   0.51892  -1.75745  0.45608   -0.41213  -0.21599
   8       -3.58408 -0.63549  0.08045  0.00798   0.52747  -3.84587  -0.41213  4.0715    0.13129
   9       19.48187 0.31036   -0.48279 0.31532   -0.51358 0.25027   -0.21599  0.13129   3.12956
#  ------------------------------------------------------------------

#  Class ID           Number of Cells     Class Name
   2       203
#  Layers 1           2          3          4          5          6          7          8          9
#  Means
           91.8937  8.74915   4.83744  1.0032    4.29985  16.02021  1.67488   7.39409   4.65653
#  Covariance
   1       366.03386          10.50815 -12.88598          2.79259  -10.57403           39.57879 -2.17348  -16.54902          35.76263
   2       10.50815 3.12962   -0.01205 0.05423   -3.23924 6.88727   -0.31766  -1.90502  0.02026
   3       -12.88598          -0.01205 8.19622   0.49213  -2.60002  4.53718   -0.28084  1.72775   -1.43348
   4       2.79259  0.05423   0.49213  2.7187    -0.34704 -2.36504  0.03671   -0.02942  1.47878
   5       -10.57403          -3.23924 -2.60002  -0.34704 16.14415  -31.29986           0.92624   2.09523   1.09795
   6       39.57879 6.88727   4.53718  -2.36504  -31.29986          90.85884  -2.4066   -5.9322   -5.68779
   7       -2.17348 -0.31766  -0.28084 0.03671   0.92624  -2.4066   0.43833   0.0446    0.22067
   8       -16.54902          -1.90502 1.72775   -0.02942 2.09523   -5.9322   0.0446    7.47759   -0.50325
   9       35.76263 0.02026   -1.43348 1.47878   1.09795  -5.68779  0.22067   -0.50325  7.65934
#  ------------------------------------------------------------------
```

Table A9. The signature file used for maximum likelihood classification of hydrochemical regions in the glacial aquifer system.—Continued

```
# Class ID          Number of Cells    Class Name
  3       236
# Layers 1          2          3          4          5          6          7          8          9
# Means
          90.9975   7.50227    6.41949    1.19491    6.43219    11.72993   1.92373    7.20339    5.55085
# Covariance
  1       485.86609            29.89292   -14.73141             13.27535   -37.87518             35.41077   -1.5361    -18.88724             70.45246
  2       29.89292   4.43281   -0.75096   0.30533    -3.39037   6.5865     0.13548    -2.66899   3.12985
  3       -14.73141            -0.75096   6.21902    -0.45379   0.29346    1.58517    -0.17638   1.01219    -3.3103
  4       13.27535   0.30533   -0.45379   2.98548    -2.2955    1.74999    -0.12481   -0.14746   2.37654
  5       -37.87518            -3.39037   0.29346    -2.2955    18.97737   -31.88564             0.42739    2.97295    -3.63091
  6       35.41077   6.5865    1.58517    1.74999    -31.88564             84.6769    -0.74756   -6.47681   -3.60399
  7       -1.5361    0.13548   -0.17638   -0.12481   0.42739    -0.74756   0.36863    -0.70783   -0.24261
  8       -18.88724            -2.66899   1.01219    -0.14746   2.97295    -6.47681   -0.70783   7.12441    -1.40234
  9       70.45246   3.12985   -3.3103    2.37654    -3.63091   -3.60399   -0.24261   -1.40234   14.10574
# ------------------------------------------------------------------

# Class ID          Number of Cells    Class Name
  4       234
# Layers 1          2          3          4          5          6          7          8          9
# Means
          66.80368   8.96241    6.95299    0.98587    3.20357    19.72603   1.91026    7.67521    1.738
# Covariance
  1       791.94663            2.25739    -10.20774             -7.10054   -31.50526             176.45764             -11.42303             -30.69444
  47.25977
  2       2.25739    5.01666   -0.2832    0.66961    -1.15104   1.48825    0.01048    -0.93      0.74232
  3       -10.20774            -0.2832    4.66302    -0.08788   -0.27175   0.582      -0.00853   0.65849    -0.90473
  4       -7.10054   0.66961   -0.08788   2.00005    -0.34767   -3.65921   0.1684     0.49459    0.38267
  5       -31.50526            -1.15104   -0.27175   -0.34767   12.12326   -30.33468             0.82024    3.02058    -0.10161
  6       176.45764            1.48825    0.582      -3.65921   -30.33468             126.15587             -3.6847    -13.64906             1.12997
  7       -11.42303            0.01048    -0.00853   0.1684     0.82024    -3.6847    0.80307    -0.23957   -0.54171
  8       -30.69444            -0.93      0.65849    0.49459    3.02058    -13.64906             -0.23957   8.47775    -1.10709
  9       47.25977   0.74232   -0.90473   0.38267    -0.10161   1.12997    -0.54171   -1.10709   5.22165
# ------------------------------------------------------------------
# Class ID          Number of Cells    Class Name
  5       223
# Layers 1          2          3          4          5          6          7          8          9
# Means
          118.16881            9.12186    4.88341    2.4524     6.39893    7.08863    2.13901    6.4843     12.02924
# Covariance
  1       318.97003            13.10541   -13.80723             4.71796    -3.95105   -16.78591             1.42455    -8.23925   77.92764
  2       13.10541   2.78854   -0.72937   -1.75976   0.50205    0.3191     0.49534    -0.73003   2.76904
  3       -13.80723            -0.72937   7.59896    0.409      -0.77844   3.24687    -0.13237   -0.35319   -3.34537
  4       4.71796    -1.75976  0.409      11.60424   -2.22633   0.11842    -0.26364   0.46168    1.17846
  5       -3.95105   0.50205   -0.77844   -2.22633   15.4963    -12.96673             -0.19431   -1.92596   -1.23957
  6       -16.78591            0.3191     3.24687    0.11842    -12.96673             31.48011   0.08576    2.02602    -4.94296
  7       1.42455    0.49534   -0.13237   -0.26364   -0.19431   0.08576    0.36347    0.063      0.64928
  8       -8.23925   -0.73003  -0.35319   0.46168    -1.92596   2.02602    0.063      9.1698     -0.45535
  9       77.92764   2.76904   -3.34537   1.17846    -1.23957   -4.94296   0.64928    -0.45535   29.17932
```

Table A10. Environmental characteristics used in supervised classification and their significance for classifying hydrochemical regions.

[GIS, Geographic Information System; %, percent; MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; mm, millimeter]

Environmental characteristic (spatial layer used in GIS)	Significant linear predictor indicated by logistic regression (95% confidence that slope coefficient is not equal to 0)				
	MAR	UIR	NGLFR	WAGR	MCFR
Average annual precipitation (1980–97), in inches	yes	yes	yes	yes	yes
Average annual temperature (1980–97), in degrees Fahrenheit	yes	yes	yes	yes	yes
Land surface slope, in percent	yes	yes	yes	yes	yes
1990s land use	yes	yes	yes	yes	yes
Vertical soil permeability, in inches per hour	yes	yes	yes	yes	yes
Soil clay content, in percent of material less than 2 mm in size	yes	yes	yes	yes	yes
Texture of surficial deposits	yes	yes	yes	yes	yes
Estimated annual natural ground-water recharge, in inches per year	yes	yes	yes	yes	yes
Type of surficial deposit	yes	yes	yes	yes	yes

Table A11. Tabulation of clusters and hydrochemical regions to compare results of unsupervised and supervised classifications.

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; --, not applicable]

Hydrochemical region	Clusters [1]					Total number of wells in each hydrochemical region that were also classified into clusters (1,260 total) [2]	Total number of wells in each hydrochemical egion extrapolated by spatial overlay (1,660 total) [3]
	1	2	3	4	5		
MAR	176	61	40	79	4	360	504
UIR	37	103	25	15	11	191	277
NGLFR	28	38	142	14	42	264	340
WAGR	43	24	20	130	2	219	237
MCFR	2	10	37	3	174	226	302
Total number of wells in each cluster that were also classified into a hydrochemical region (1,260 total)	286	236	264	241	233	--	--

[1] Cluster analysis classified a total of 1,316 wells (286, 237, 286, 241, and 266 wells each in clusters 1, 2, 3, 4, and 5, respectively).

[2] Maximum likelihood classification (MLC) classified the wells used in cluster analysis into hydrochemical regions.

[3] Hydrochemical regions were identified for 400 additional wells (not used in MLC) by overlaying points representing the wells on the hydrochemical regions map in a geographic information system. A hydrochemical region could not be identified for 56 wells.

Table A12. Results of Kruskal-Wallis rank sum test for differences among hydrochemical regions. The NULL hypothesis is that the difference in means of the ranks is equal to zero; the alternate hypothesis is that the difference in means of the ranks is not equal to zero (p-value <0.05).

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; <, less than; A, B, C, D, E statistical population where same letter across the row indicates the populations are not significantly different from each other (p-value <0.05); bold text indicates rejection of the NULL hypothesis (p-value was greater than 0.05)]

Constituent	p-value of hydrochemical region pairwise comparison										Statistical population of hydrochemical region				
	MAR UIR	MAR NGLFR	MAR WAGR	MAR MCFR	UIR NGLFR	UIR WAGR	UIR MCFR	NGLFR WAGR	NGLFR MCFR	WAGR MCFR	MAR	UIR	NGLFR	WAGR	MCFR
Constituents used in unsupervised classification															
Calcium	0.01	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	B	A	D	C	E
Chloride	<.0001	.0002	**.75**	**.84**	<.0001	<.0001	<.0001	.003	.0006	**.84**	B	A	C	B	B
Fluoride	<.0001	<.0001	.006	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	B	C	D	A	E
Magnesium	<.0001	<.0001	<.0001	<.0001	<.0001	.004	<.0001	<.0001	<.0001	<.0001	A	B	D	C	E
Potassium	.0002	.01	<.0001	.007	<.0001	<.0001	<.0001	<.0001	**.89**	<.0001	C	B	D	A	D
Sodium	<.0001	<.0001	<.0001	.0008	<.0001	**.26**	<.0001	<.0001	<.0001	<.0001	B	A	D	A	C
Sulfate	.02	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.001	<.0001	C	B	D	A	E
Bicarbonate	<.0001	<.0001	<.0001	<.0001	<.0001	.008	<.0001	<.0001	<.0001	<.0001	A	B	D	C	E
Select measured trace elements															
Arsenic	.01	.004	<.0001	<.0001	**.73**	<.0001	.0001	<.0001	.0002	<.0001	D	B	B	A	C
Barium	**.41**	<.0001	**.60**	<.0001	<.0001	**.70**	<.0001	<.0001	**.09**	<.0001	A	A	B	A	B
Boron	**.99**	<.0001	**.28**	<.0001	<.0001	**.16**	<.0001	<.0001	**.91**	<.0001	A	A	B	A	B
Bromide	.02	<.0001	<.0001	<.0001	<.0001	.0004	<.0001	<.0001	**.81**	<.0001	C	B	D	A	D
Iron	**.54**	.01	.002	<.0001	**.11**	.02	<.0001	**.17**	.0001	**.12**	A	AB	BC	CD	D
Lithium	.0001	<.0001	<.0001	<.0001	**.13**	<.0001	**.08**	<.0001	**.87**	<.0001	B	C	C	A	C
Manganese	**.19**	**.45**	.02	<.0001	**.67**	**.19**	.02	**.09**	.004	**.81**	A	AB	AB	BC	C
Molybdenum	.004	<.0001	**.10**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	A	B	D	A	C
Nitrite plus nitrate as nitrogen	**.09**	<.0001	<.0001	<.0001	.03	<.0001	.003	.04	**.26**	**.13**	C	C	B	A	AB
Silica	<.0001	.004	<.0001	.0001	**.16**	<.0001	**.84**	<.0001	**.32**	<.0001	B	C	C	A	C
Strontium	**.15**	<.0001	.005	<.0001	<.0001	<.0001	<.0001	<.0001	**.1**	<.0001	B	B	C	A	C
Select measured field parameters															
Well depth	**.28**	.0002	<.0001	**.07**	.01	<.0001	.005	<.0001	<.0001	.004	BC	C	D	A	B
Specific conductance	.0004	<.0001	.01	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	B	A	D	C	E
Dissolved oxygen	**.55**	<.0001	**.67**	<.0001	<.0001	**.42**	<.0001	.04	<.0001	<.0001	C	C	B	C	A
pH	**.83**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	**.82**	<.0001	<.0001	B	B	A	A	C
Alkalinity	<.0001	<.0001	<.0001	<.0001	<.0001	.006	<.0001	<.0001	<.0001	<.0001	A	B	D	C	E

Table A12. Results of Kruskal-Wallis rank sum test for differences among hydrochemical regions. The NULL hypothesis is that the difference in means of the ranks is equal to zero; the alternate hypothesis is that the difference in means of the ranks is not equal to zero (p-value <0.05).—Continued

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; <, less than; A, B, C, D, E statistical population where same letter across the row indicates the populations are not significantly different from each other (p-value <0.05); bold text indicates rejection of the NULL hypothesis (p-value was greater than 0.05)]

Constituent	p-value of hydrochemical region pairwise comparison										Statistical population of hydrochemical region				
	MAR UIR	MAR NGLFR	MAR WAGR	MAR MCFR	UIR NGLFR	UIR WAGR	UIR MCFR	NGLFR WAGR	NGLFR MCFR	WAGR MCFR	MAR	UIR	NGLFR	WAGR	MCFR
Calculated variables															
Calcium to Magnesium ratio	<0.0001	<0.0001	<0.0001	<0.0001	0.03	**0.06**	<0.0001	0.0005	0.004	<0.0001	D	C	B	C	A
Sodium to Chloride ratio	<.0001	.0001	<.0001	**.35**	**.12**	<.0001	<.0001	<.0001	.0005	<.0001	B	C	C	A	B
Dissolved solids	**.14**	<.0001	**.11**	<.0001	<.0001	.02	<.0001	<.0001	<.0001	<.0001	AB	A	C	B	D
Percent of dissolved solids contributed by calcium	**.19**	<.0001	.01	<.0001	<.0001	**.25**	<.0001	<.0001	<.0001	<.0001	B	BC	A	C	D
Percent of dissolved solids contributed by chloride	<.0001	.01	**.87**	<.0001	<.0001	<.0001	**.41**	.004	<.0001	<.0001	C	A	B	C	A
Perent of dissolved solids contributed by magnesium	<.0001	<.0001	<.0001	<.0001	.03	**.30**	<.0001	**.23**	<.0001	<.0001	A	C	B	BC	D
Percent of dissolved solids contributed by nitrite plus nitrate as nitrogen	**.85**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	**.07**	.0003	<.0001	B	B	C	C	A
Percent of dissolved solids contributed by potassium	.01	<.0001	<.0001	<.0001	.0005	<.0001	<.0001	.002	<.0001	<.0001	E	D	C	B	A
Percent of dissolved solids contributed by silica	**.08**	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.0003	<.0001	<.0001	D	D	C	B	A
Percent of dissolved solids contributed by sodium	<.0001	<.0001	<.0001	<.0001	<.0001	**.57**	<.0001	<.0001	<.0001	<.0001	C	B	D	B	A
Percent of dissolved solids contributed by sulfate	**.84**	.04	<.0001	**.18**	.04	<.0001	**.11**	<.0001	.0001	<.0001	B	B	C	A	B
Percent of dissolved solids contributed by calcium carbonate	<.0001	.0003	<.0001	<.0001	.005	**.94**	<.0001	.001	<.0001	<.0001	A	C	B	C	D

Table A13. Spearman correlation coefficients and probability results from z-test for equal proportions to compare the fit of hydrochemical regions to corresponding clusters for 12 different maximum likelihood classification models.

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; Italic underlined z-test probability indicates there was a poor fit of proportions]

Model	Spearman correlation coefficients between cluster membership and hydrochemical region membership [1]					Probability resulting from z-test for equal proportions [2]				
	Cluster 1, MAR	Cluster 2, UIR	Cluster 3, NGLFR [3]	Cluster 4, WAGR	Cluster 5, MCFR	Cluster 1, MAR	Cluster 2, UIR	Cluster 3, NGLFR	Cluster 4, WAGR	Cluster 5, MCFR
Original model (precipitation, temperature, slope, land use, clay, vertical permeability, recharge, surficial deposit type, and surficial deposit texture)	0.4	0.38	0.42	0.47	0.71	*0*	0.05	0.63	0.54	0.14
Modified model—without recharge	.39	.42	.38	.42	.68	*0*	*0*	*0.04*	0.92	.58
Modified model—without clay	.38	.38	.41	.44	.68	*0*	*.03*	.95	.17	.32
Modified model—without vertical permeability	.4	.37	.41	.41	.7	*0*	.16	.80	.10	.62
Modified model—without clay and vertical permeability	.33	.39	.34	.48	.68	*0*	.05	.44	*0*	.49
Modified model—without surficial deposits	.42	.38	.37	.54	.68	*0*	*.02*	.70	*0*	.52
Modified model—without temperature	.38	.39	.4	.47	.71	*0*	.10	*0*	.63	.98
Modified model—without precipitation	.38	.37	.37	.40	.7	*0*	.11	*0*	.87	.62
Modified model—without surficial texture	.4	.4	.37	.49	.7	*0*	.31	.93	.05	.45
Modified model—without slope	.4	.37	.38	.45	.68	*0*	.10	.14	.27	.48
Modified model—without land use	.41	.20	.37	.46	.71	*0*	*0*	.41	.50	.56
Only using recharge to model the hydrochemical regions	.31	.13	.06	.62	.65	*0*	*0*	*0*	.95	.98

[1] All correlation coefficients were highly significant (p-value <0.0001) except the correlation between cluster 3 and the NGLFR for the model using only recharge[3], which was significant (p-value <0.05). Although the actual correlation values are lower than are usually considered 'significant', the correlations are statistically significant because of the large sample sizes (around 200 samples for each comparison). For the purpose of calculating correlations, cluster groups and hydrochemical regions were indicated by binary variables, one variable for each cluster or region. The binary variables used a '1' to indicate a well was in the cluster or region or used a value of '0' to indicate a well was not in the cluster or region.

[2] The z-test for equal proportions compares proportions of samples assigned to each cluster with proportions of samples assigned to each hydrochemical region (following method in Ott, 1993, p 383–384). The NULL hypothesis is that the sample proportions are equal; the alternative hypothesis is that they are not equal. This test of equal proportions tests if a z-value at least as great as calculated could be observed under the NULL hypothesis by random chance. The NULL hypothesis was rejected if the probability (p-value) was less than 0.05, which indicated a poor model fit.

[3] The correlation coefficient for cluster 3 and NGLFR was significant at the 95th percent confidence level (p-value < 0.05).

Table A14. Environmental characteristics and general geographic location of each hydrochemical region.

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; <, less than]

Environmental characteristic or geographic location	Hydrochemical region				
	MAR	UIR	NGLFR	WAGR	MCFR
Average annual precipitation, in inches	37	38	36	21	46
Average annual temperature, in degrees Fahrenheit	49	48	43	45	44
1990s Land cover, in percent of the region					
Urban	<1	30	2	1	5
Bare rock	<1	<1	1	1	2
Forest/shrub	9	36	50	10	71
Grassland	3	4	2	24	4
Agriculture	84	26	28	61	13
Wetland	3	3	17	3	5
Texture of surficial deposits, in percent of the region					
Fine-grained	89	61	69	66	44
Coarse-grained	8	18	24	9	18
Mixed fine- and coarse-grained	3	21	7	25	38
Type of surficial deposit, in percent of the region					
Till	84	44	56	52	56
Glaciofluvial	6	9	16	2	5
Lacustrine	5	17	9	10	5
Alluvial/colluvial	3	16	9	22	21
Eolian	2	4	2	6	<1
Marsh/swamp deposits	<1	<1	5	<1	<1
Other	0	9	3	8	12
Land surface slope, in percent [1]					
Mean	1	2	2	1	8
Vertical soil permeability, in inches per hour					
Mean	1	2	4	1	3
Range	<1–9	<1–15	<1–15	<1–14	<1–18
Soil clay content, in percent of material less than 2 millimeters in size					
Mean	27	22	15	28	11
Range	0–44	0–61	0–69	0–69	0–61
Estimated annual natural ground-water recharge, in inches per year					
Mean	3	6	7	1	13
Range	<1–9	<1–85	<1–84	<1–70	<1–85

Table A14. Environmental characteristics and general geographic location of each hydrochemical region.—Continued

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; <, less than]

Environmental characteristic or geographic location	Hydrochemical region				
	MAR	UIR	NGLFR	WAGR	MCFR
General geographic location, in percent of region in each state[2]					
Connecticut	0	0	2	0	2
Idaho	0	1	0	0	3
Illinois	20	9	1	1	1
Indiana	10	8	3	1	1
Iowa	21	7	1	2	1
Kansas	2	0	0	2	1
Kentucky	0	0	0	<1	0
Maine	0	3	9	0	11
Massachusetts	0	0	1	0	4
Michigan	6	13	15	0	10
Minnesota	11	6	26	8	1
Missouri	8	2	1	2	2
Montana	1	6	1	18	8
Nebraska	4	1	1	7	0
New Hampshire	0	0	1	0	5
New Jersey	0	1	1	0	1
New York	1	11	11	0	16
North Dakota	1	1	2	28	0
Ohio	9	9	<1	2	2
Pennsylvania	0	7	4	0	5
Rhode Island	0	0	0	0	<1
South Dakota	1	1	<1	18	0
Vermont	0	1	1	0	5
Washington	0	7	1	10	19
Wisconsin	5	6	17	1	2

[1] Percent slope is calculated as 100 times the quotient of change in elevation divided by the change in ground distance. Percent slope approaches infinity as the measured slope angle approaches 90 degrees.

[2] Alaska is not included in this table because hydrochemical regions were not mapped for that state.

Table A15. Descriptive statistics by hydrochemical region.

[mg/L, milligrams per liter; µg/L, micrograms per liter; µS/cm, microsiemens per centimeter; °C, degrees Celsius; %, percent]

Hydrochemical region	Constituent	Calcium (mg/L)	Chloride (mg/L)	Fluoride (mg/L)	Magnesium (mg/L)	Potassium (mg/L)	Sodium (mg/L)	Sulfate (mg/L)	Bicarbonate (mg/L)
Midwestern Agricultural Region (504 wells total)	Minimum	16	0.20	0.10	4.20	0.10	1.00	0.10	21.00
	25th percentile	71	5.08	.16	26.00	1.02	6.07	17.77	321.65
	Mean	91.50	36.75	.34	35.24	3.05	30.07	66.97	393.12
	50th percentile (median)	86.26	14.05	.24	33	1.60	12.98	36.73	391.29
	75th percentile	100.56	28.24	.40	39.61	2.61	27.39	65.57	467.18
	Maximum	410.38	2,300	1.95	204.63	118.75	1,300	1,385.50	1,241.40
	Standard Deviation	41.11	118.81	0.27	17.72	7.08	73.08	126.59	128.02
	Number of wells	482	480	477	483	483	483	480	372
Urban-Influenced Region (277 wells total)	Minimum	0.91	.60	.10	0.38	.20	1.68	.10	6.10
	25th percentile	74.68	17	.10	22.51	1.30	9.91	27.86	271.60
	Mean	93.64	74.72	.21	28.46	2.88	42.29	55.48	340.09
	50th percentile (median)	92.55	43.16	.18	28.25	2.03	23.80	42.38	349
	75th percentile	110	93.35	.23	35	3.28	45.89	66	412.09
	Maximum	240	800	1.38	69	32	380	520	1,070
	Standard Deviation	35.74	99.98	.17	11.44	3.47	59.14	58.98	136.27
	Number of wells	270	270	268	270	270	270	269	193
Northern and Great Lakes Forested Region (311 wells total)	Minimum	.50	.10	.10	.09	.10	0.99	.10	6
	25th percentile	39.08	2.6	.10	9.55	.9	2.48	7.86	127.7
	Mean	62.53	24.72	.14	18.12	2.09	12.73	30.07	215.29
	50th percentile (median)	61	9.85	.10	17	1.4	4.2	15.07	221.5
	75th percentile	81	22.32	.14	24.61	2.3	9.75	31.1	290
	Maximum	210	500	.73	63.99	15.45	290	410	665
	Standard Deviation	34.03	53.83	.09	11.35	2.08	26.72	47.99	115.85
	Number of wells	327	326	268	327	328	327	324	316
Western Agriculture and Grassland Region (237 wells total)	Minimum	14.41	0.19	.10	5.22	.30	2.30	.10	71
	25th percentile	57	7.23	.20	16	1.83	9.91	33.90	228.70
	Mean	84.36	52.47	.36	33.16	5.01	58.03	120.56	322.04
	50th percentile (median)	75	12.42	.30	23	4.19	26.77	59	295
	75th percentile	95.2	25.59	.44	35.76	6.39	49	110	383
	Maximum	430	4,800	1.99	500	34	3,400	2,700	1,126
	Standard Deviation	50.54	328.47	.26	51.23	4.58	234.42	286.87	137.56
	Number of wells	235	236	236	235	235	235	235	221
Mountain and Coastal Forested Region (331 wells total)	Minimum	.98	.10	.10	0.35	.10	.80	.10	2
	25th percentile	11	4.90	.10	2.77	.9	5.36	8.15	26.82
	Mean	26.39	30.34	.12	6.5	2.04	18.55	18.89	81.82
	50th percentile (median)	19.15	13.17	.10	4.45	1.37	8.45	12.69	54.18
	75th percentile	32.48	32	.10	7.01	2.3	19.25	19	106.25
	Maximum	160	728.34	1.28	39	20	377.40	500	493.78
	Standard Deviation	23.42	60.49	.10	6.29	2.2	33.01	36.55	84.04
	Number of wells	292	292	233	292	292	292	291	296

Table A15. Descriptive statistics by hydrochemical region.—Continued

[mg/L, milligrams per liter; µg/L, micrograms per liter; µS/cm, microsiemens per centimeter; °C, degrees Celsius; %, percent]

Hydrochemical region	Constituent	Arsenic (µg/L)	Barium (µg/L)	Boron (µg/L)	Bromide (µg/L)	Iron (µg/L)	Lithium (µg/L)	Manganese (µg/L)	Molybdenum (µg/L)
Midwestern Agricultural Region (504 wells total)	Minimum	1	6	12	0.01	10	6	1	1
	25th percentile	1	56.56	20.82	.03	10	6	13.36	1.18
	Mean	4.38	147.95	71.75	.14	790.98	14.58	203.67	6.63
	50th percentile (median)	1	83.04	36.56	.05	48.36	6	49.15	2.66
	75th percentile	3.24	173	70.86	.09	1,193.70	12.97	176.03	6
	Maximum	84	1630.20	782.67	22	9,800	117.94	18,272	111.51
	Standard Deviation	9.13	182.59	102.17	1.03	1,377.38	19.55	887.76	12.71
	Number of wells	268	267	143	478	480	143	473	266
Urban-I nfluenced Region (277 wells total)	Minimum	1	1.42	12	.01	10	6	1	1
	25th percentile	1	52.09	25.88	.04	10	6	2.04	1
	Mean	5.33	111.80	63.55	.10	1,085.64	6.69	220.67	4.14
	50th percentile (median)	1	92.10	37.51	.06	28.90	6	39.21	2
	75th percentile	1.39	137.19	56.58	.09	962.03	6	223.76	3.98
	Maximum	340	590	1,243.40	2	28,000	19.76	2,660	55.76
	Standard Deviation	28.07	91.37	129.37	.20	3,040.18	2.28	421.74	7.15
	Number of wells	157	156	103	267	270	93	266	155
Northern and Great Lakes Forested Region (311 wells total)	Minimum	1	1	12	.01	10	6	1	1
	25th percentile	1	12.84	12	.02	10	6	2.08	1
	Mean	2.37	46.98	51.69	.06	724.62	6.94	253.52	1.78
	50th percentile (median)	1	29.10	15.54	.03	16.65	6	39	1
	75th percentile	1.18	55.7	26.09	.05	370.35	6	280.26	1.28
	Maximum	27.69	382.32	2,600	4.40	25,000	47.96	11,650.69	25.86
	Standard Deviation	4.18	59.79	248.46	.25	2,292.16	4.65	767.8	2.64
	Number of wells	129	136	110	303	328	105	303	126
Western Agricul-ture and Grassland Region (237 wells total)	Minimum	1	8	12	.01	10	6	1	1
	25th percentile	2.97	52.47	27.13	.04	10	6.91	1.	1.87
	Mean	8.75	116.88	79.84	.16	794.23	19	199.29	8.01
	50th percentile (median)	4.88	90.52	43.28	.07	10	12.94	28.	3.14
	75th percentile	8.97	148	76.90	.13	619.63	22.30	332	6.14
	Maximum	116.39	900	910	8.40	10,677	126.32	1,400	304.30
	Standard Deviation	13.75	107.26	120.66	.60	1,767.5	21.12	297.65	29.78
	Number of wells	115	122	106	219	234	103	229	107
Mountain and Coastal Forested Region (331 wells total)	Minimum	1	1	12	.01	10	6	1	1
	25th percentile	1	14.44	12	.02	10	6	2	1
	Mean	1.56	71.68	38.25	.05	675	6.42	293.59	1.28
	50th percentile (median)	1	30.71	15.20	.03	10	6	15.65	1
	75th percentile	1	101.25	28.84	.04	34.25	6	90.25	1
	Maximum	47.88	350.59	1,085.60	2.64	38,069	29.31	28,181	9
	Standard Deviation	4.09	84.19	108.81	0.17	3,380.41	2.32	1,746.91	1.1
	Number of wells	139	139	125	281	292	121	288	138

Table A15. Descriptive statistics by hydrochemical region.—Continued

[mg/L, milligrams per liter; µg/L, micrograms per liter; µS/cm, microsiemens per centimeter; °C, degrees Celsius; %, percent]

Hydrochemical region	Constituent	Nitrite plus nitrate (mg/L as nitrogen)	Silica (mg/L)	Strontium (µg/L)	Well depth (feet below land surface)	Specific conductance (µS/cm at 25 °C)	Dissolved oxygen (mg/L)	pH (standard units)	Alkalinity (mg/L as calcium carbonate)
Midwestern Agricultural Region (504 wells total)	Minimum	0.05	4.60	61.26	1	245	0.10	5.32	16
	25th percentile	.05	12.45	136.44	19.43	624	.15	6.91	264
	Mean	3.20	16.81	695.82	55.80	811.55	2.33	7.09	322.54
	50th percentile (median)	.08	15.81	314.57	32	710.50	.91	7.10	317.5
	75th percentile	3.39	20	632.03	60	856	4.09	7.28	381.58
	Maximum	77	46.90	9,121	380	7,460	12.96	9.47	1,018.40
	Standard Deviation	7.11	6.80	1,241.82	65.37	467.15	2.73	0.34	104
	Number of wells	477	483	143	498	476	476	486	372
Urban-Influenced Region (277 wells total)	Minimum	.05	3.80	0.20	4	4.92	.10	5.10	5
	25th percentile	.05	10.77	146.62	20.24	636	.14	6.92	223.20
	Mean	2.62	14.94	434.81	42.16	846	2.23	7.08	278.95
	50th percentile (median)	.71	13.21	216.63	30	795.50	.87	7.10	286
	75th percentile	3.96	17.64	502.81	49.03	974	3.71	7.28	338
	Maximum	26	62.43	3,437.30	280	2,880	13.66	8.38	877
	Standard Deviation	3.98	6.74	542.59	37.14	357.50	2.7	.35	111.61
	Number of wells	269	270	93	276	254	262	269	193
Northern and Great Lakes Forested Region (311 wells total)	Minimum	.05	3.40	12.82	0.80	47	.10	5.50	5
	25th percentile	.05	10.34	39.65	16.93	329	.2	6.98	104.75
	Mean	4.62	15.31	101.33	37.6	500.56	3.54	7.23	176.74
	50th percentile (median)	.8	14.64	62.37	25.69	467.5	2.32	7.28	181.5
	75th percentile	5.17	19.86	105.22	50	610.25	6.63	7.59	238.2
	Maximum	47	33	1,843.11	223	2,210	12.08	9.13	545
	Standard Deviation	8.23	6.24	185.53	32.43	276.7	3.49	.61	95.1
	Number of wells	310	327	105	314	328	323	328	316
Western Agriculture and Grassland Region (237 wells total)	Minimum	.05	8.76	82	8	214	.10	5.86	58
	25th percentile	.07	22	332.52	24.38	512.10	.10	7.02	184.25
	Mean	5.41	31.69	493.22	82.57	861.69	2.88	7.26	260.40
	50th percentile (median)	1.98	29	442.07	42.70	681	1.15	7.27	240
	75th percentile	7.74	39.27	624.95	120.25	873.50	6	7.53	314.60
	Maximum	76	62	1,400	425	15,600	14.16	8.30	674
	Standard Deviation	8.63	12.57	247.88	84.87	1,140.55	3.12	.43	104.80
	Number of wells	212	235	103	236	235	224	226	218
Mountain and Coastal Forested Region (331 wells total)	Minimum	.05	4	1.47	5	36	.10	4.50	1
	25th percentile	.26	10	41.49	22.5	157	1.1	6	21
	Mean	2.52	15.59	119.03	51	290.98	4.71	6.55	67.01
	50th percentile (median)	1.3	13.8	86.41	38	231	4.85	6.4	44.5
	75th percentile	3.7	18.82	131.21	65	348.5	7.6	7.2	87.25
	Maximum	20.66	50	1,873.85	273	2,620	13.3	8.7	405
	Standard Deviation	3.26	8.19	186.72	41.78	243.36	3.42	.8	69.01
	Number of wells	289	292	122	301	291	296	301	296

Table A15. Descriptive statistics by hydrochemical region.—Continued

[mg/L, milligrams per liter; µg/L, micrograms per liter; µS/cm, microsiemens per centimeter; °C, degrees Celsius; %, percent]

Hydrochemical region	Constituent	Sodium to Chloride ratio (dimensionless)	Calcium to magnesium ratio (dimensionless)	Sodium to potassium ratio (dimensionless)	Magnesium to sulfate ratio (dimensionless)	Chloride to bromide ratio (dimensionless)	Hardness (mg/L of calcium and magnesium as calcium carbonate)	Dissolved solids (mg/L)	Calcium (percent of dissolved solids)
Midwestern Agricultural Region (504 wells total)	Minimum	0.08	0.36	0.07	0.16	0.46	59.76	79.63	5%
	25th percentile	.60	1.29	7.05	1.90	113.58	293.33	347.76	18%
	Mean	6.45	1.76	23.51	82.87	539.25	373.71	494.85	20%
	50th percentile (median)	1.24	1.59	12.91	3.39	237.26	351.85	423.58	21%
	75th percentile	3.99	1.96	26.50	6.49	533.71	409.75	514.14	23%
	Maximum	138.78	6.62	307.04	1,699.55	9,684.70	1,386.82	4,090.65	35%
	Standard Deviation	15.80	.80	29.40	294.46	976.46	156.34	337.45	5%
	Number of wells	480	482	483	480	475	482	354	354
Urban-Influenced Region (277 wells total)	Minimum	.09	.29	.12	.01	10	3.81	72.39	0%
	25th percentile	.59	1.69	9.59	1.58	326.79	291.95	337.72	18%
	Mean	2.18	2.19	34.18	18.77	1,154.92	351.09	499.31	20%
	50th percentile (median)	.79	2.05	17.54	2.55	720	352.71	459.53	20%
	75th percentile	1.03	2.45	31.97	3.75	1,325	415.18	589.40	23%
	Maximum	96.54	9.10	1,534.71	1,897.18	15,277	883.58	1,728.40	32%
	Standard Deviation	7.93	.97	98.55	146.92	1,516.91	123.82	254.74	5%
	Number of wells	270	270	270	269	267	270	192	192
Northern and Great Lakes Forested Region (311 wells total)	Minimum	.06	.60	.26	.04	4.04	1.62	30.94	0%
	25th percentile	.51	1.6	3.01	1.48	129.72	142.43	174.85	20%
	Mean	2.92	2.57	15.49	15.02	621.92	230.77	276.15	22%
	50th percentile (median)	.9	2.16	5.67	3.36	330	225.97	255.84	23%
	75th percentile	1.83	2.83	12.38	7.43	675.19	292.97	346.99	26%
	Maximum	95.46	13.21	476.15	1,331.15	9,028	709.96	1,177.18	39%
	Standard Deviation	8.57	1.76	37.78	86.61	893.53	123.3	162.34	5%
	Number of wells	325	327	327	323	303	327	288	288
Western Agriculture and Grassland Region (237 wells total)	Minimum	.18	.14	1.06	.20	7.72	86.22	144.15	3%
	25th percentile	1.27	1.35	6.38	.94	110	224.06	313.05	15%
	Mean	5.50	2.10	19.21	18.61	218.71	347.28	491.55	19%
	50th percentile (median)	3.39	1.94	11.51	1.41	159.13	284.67	404.96	20%
	75th percentile	5.86	2.54	24.40	2.59	236.37	379.16	534.92	23%
	Maximum	42.85	5.27	170.05	1,600.74	1,537.09	3,033.33	4,568.80	29%
	Standard Deviation	7.17	1.06	23.12	132.45	217.35	307.80	452.13	5%
	Number of wells	235	235	235	234	219	235	204	204
Mountain and Coastal Forested Region (331 wells total)	Minimum	.20	.21	.38	.11	.49	5.89	26.79	0%
	25th percentile	.79	1.72	5.82	.9	219.43	42.25	88.32	11%
	Mean	1.96	3.03	20.27	3.2	847.63	92.67	158.79	16%
	50th percentile (median)	1	2.42	11.89	1.33	484.29	68.79	123.52	16%
	75th percentile	1.87	3.64	23.81	2.24	1,013.04	118.1	186.1	21%
	Maximum	76.48	26.78	391.13	124.50	8,312.53	560.23	1,348.81	33%
	Standard Deviation	4.81	2.44	32.41	10.66	1,050.52	78.75	129.84	7%
	Number of wells	292	292	292	291	281	292	282	282

Table A15. Descriptive statistics by hydrochemical region.—Continued

[mg/L, milligrams per liter; µg/L, micrograms per liter; µS/cm, microsiemens per centimeter; °C, degrees Celsius; %, percent]

Hydrochemical region	Constituent	Chloride (percent of dissolved solids)	Magnesium (percent of dissolved solids)	Nitrite plus nitrate as nitrogen (percent of dissolved solids)	Potassium (percent of dissolved solids)	Silica (percent of dissolved solids)	Sodium (percent of dissolved solids)	Sulfate (percent of dissolved solids)	Calcium carbonate (percent of dissolved solids)
Midwestern Agricultural Region (504 wells total)	Minimum	0%	1%	0%	0%	0%	0%	0%	2%
	25th percentile	1%	6%	0%	0%	3%	2%	5%	36%
	Mean	6%	7%	1%	1%	4%	6%	12%	43%
	50th percentile (median)	3%	8%	0%	0%	4%	3%	9%	45%
	75th percentile	8%	9%	1%	1%	5%	7%	15%	51%
	Maximum	56%	16%	21%	8%	16%	32%	59%	64%
	Standard Deviation	8%	2%	2%	1%	2%	6%	11%	12%
	Number of wells	354	354	354	354	354	354	354	354
Urban-Influenced Region (277 wells total)	Minimum	0%	0%	0%	0%	1%	1%	0%	1%
	25th percentile	5%	4%	0%	0%	2%	3%	6%	27%
	Mean	14%	6%	1%	1%	4%	8%	11%	36%
	50th percentile (median)	11%	6%	0%	0%	3%	6%	9%	36%
	75th percentile	21%	7%	1%	1%	5%	11%	13%	45%
	Maximum	46%	11%	13%	16%	28%	40%	49%	64%
	Standard Deviation	11%	2%	2%	1%	3%	7%	8%	12%
	Number of wells	192	192	192	192	192	192	192	192
Northern and Great Lakes Forested Region (311 wells total)	Minimum	0%	0%	0%	0%	1%	0%	0%	4%
	25th percentile	1%	5%	0%	0%	4%	1%	4%	29%
	Mean	8%	6%	2%	1%	7%	4%	10%	39%
	50th percentile (median)	5%	6%	0%	1%	6%	2%	8%	41%
	75th percentile	10%	8%	3%	1%	8%	4%	13%	50%
	Maximum	53%	13%	24%	13%	40%	39%	57%	59%
	Standard Deviation	10%	3%	3%	1%	5%	6%	9%	14%
	Number of wells	288	288	288	288	288	288	288	288
Western Agriculture and Grassland Region (237 wells total)	Minimum	0%	2%	0%	0%	1%	1%	0%	5%
	25th percentile	2%	5%	0%	0%	5%	3%	10%	32%
	Mean	4%	6%	1%	1%	8%	7%	17%	36%
	50th percentile (median)	3%	6%	1%	1%	7%	6%	15%	37%
	75th percentile	5%	7%	2%	2%	10%	9%	22%	42%
	Maximum	26%	16%	13%	5%	31%	28%	69%	62%
	Standard Deviation	4%	2%	2%	1%	5%	5%	12%	10%
	Number of wells	204	204	204	204	204	204	204	204
Mountain and Coastal Forested Region (331 wells total)	Minimum	0%	0%	0%	0%	1%	1%	0%	0%
	25th percentile	5%	2%	0%	1%	7%	5%	7%	13%
	Mean	16%	4%	2%	1%	14%	11%	12%	24%
	50th percentile (median)	11%	4%	1%	1%	11%	8%	10%	23%
	75th percentile	26%	5%	2%	2%	20%	14%	14%	35%
	Maximum	56%	19%	14%	6%	44%	40%	63%	58%
	Standard Deviation	14%	2%	3%	1%	10%	8%	9%	14%
	Number of wells	282	282	282	282	282	282	282	282

Table A16. Descriptive statistics for subsurface contact time and estimated ground-water recharge date for wells in each hydrochemical region.

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Descriptive statistics	Hydrochemical region				
	MAR	UIR	NGLFR	WAGR	MCFR
Subsurface contact time, in days					
Minimum	25	8	5	38	4
25th percentile	233	109	56	203	30
Mean	827	689	215	824	116
50th percentile (median)	475	207	85	456	49
75th percentile	1,058	580	187	846	104
Maximum	12,331	13,144	13,094	24,553	1,535
Standard Deviation	1,119	1,770	767	1,971	193
Number of wells	502	277	337	237	292
Estimated ground-water recharge date					
Minimum	1947	1940	1939	1950	1955
25th percentile	1950	1953	1953	1953	1984
Mean	1962	1970	1967	1965	1987
50th percentile (median)	1953	1959	1955	1955	1990
75th percentile	1976	1987	1984	1980	1992
Maximum	2001	2001	1999	1997	2000
Standard Deviation	16	19	17	15	9
Number of wells	311	160	117	42	63

Table A17. Percent of total wells in each hydrochemical region by type of NAWQA sampling network.

[MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region; %, percent]

Hydro-chemical region	Total number of wells in region	Percent of total wells in networks monitoring urban land-use areas	Percent of total wells in networks monitoring row crop or orchard agricultural land-use areas	Percent of total wells in networks monitoring forest land-use areas	Percent of total wells that are in drinking water networks	Percent of total wells that are reference networks	Percent of total wells that are in flow-path studies	Percent of total wells that are in special studies
MAR	504	5.4%	47.6%	0%	31.7%	0.6%	11.5%	3.2%
UIR	277	47.3%	18.1%	0%	17.3%	0%	8.3%	9%
NGLFR	340	14.4%	39.4%	8.2%	30.3%	.3%	7.4%	0%
WAGR	237	4.2%	43%	0%	38%	0%	3.4%	11.4%
MCFR	302	34.8%	11.6%	8.9%	38.1%	0%	6.6%	0%

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Calcium	Chloride	Fluoride	Magnesium	Potassium	Sodium	Sulfate	Bicarbonate	Recharge[1]	Average annual precipitation[1]	Average annual temperature[1]	Coarse-grained surficial material[1]	Fine-grained surficial material[1]	Clay in soil[1]	Vertical permeability[1]	Agricultural land use[1]
Calcium	MAR	--	P	N	P	P	ns	P	P	ns	ns	ns	ns	ns	P	ns	ns
	UIR	--	P	N	P	P	P	P	P	N	ns	ns	ns	ns	ns	ns	N
	NGLFR	--	P	P	P	P	P	P	P	N	N	N	P	N	ns	P	P
	WAGR	--	P	N	P	P	ns	P	P	ns	P	N	ns	P	P	ns	ns
	MCFR	--	P	ns	P	P	P	P	P	N	N	N	P	N	P	N	P
Chloride	MAR	--	--	N	P	P	P	P	ns	P	ns	ns	P	N	ns	ns	N
	UIR	--	--	N	P	P	P	P	P	ns	ns	ns	P	ns	N	P	N
	NGLFR	--	--	N	P	P	P	P	ns	P	P	P	ns	ns	ns	ns	ns
	WAGR	--	--	P	P	P	P	P	ns	ns	N	P	N	ns	ns	ns	ns
	MCFR	--	--	ns	ns	P	P	P	ns	ns	ns	N	P	N	N	P	N
Fluoride	MAR	--	--	--	ns	P	P	N	P	N	P	P	N	P	P	N	P
	UIR	--	--	--	P	ns	ns	ns	P	ns	ns	ns	ns	ns	P	N	ns
	NGLFR	--	--	--	P	ns	P	ns	P	N	N	ns	ns	ns	P	ns	ns
	WAGR	--	--	--	ns	P	P	ns	ns	N	N	P	N	ns	N	ns	P
	MCFR	--	--	--	ns	ns	ns	ns	ns	N	N	N	ns	ns	ns	ns	ns
Magnesium	MAR	--	--	--	--	P	P	P	P	ns	ns	ns	N	P	ns	ns	ns
	UIR	--	--	--	--	P	ns	P	P	N	ns	P	ns	ns	P	N	ns
	NGLFR	--	--	--	--	P	ns	P	P	N	N	N	P	N	ns	P	P
	WAGR	--	--	--	--	ns	P	P	P	ns	N	N	N	P	N	P	P
	MCFR	--	--	--	--	P	ns	P	P	N	ns	P	ns	ns	P	N	P
Potassium	MAR	--	--	--	--	--	P	P	P	N	N	N	ns	ns	ns	ns	ns
	UIR	--	--	--	--	--	P	P	P	N	ns	ns	P	N	N	P	N
	NGLFR	--	--	--	--	--	P	P	P	ns	ns	ns	ns	ns	ns	ns	ns
	WAGR	--	--	--	--	--	P	P	P	N	N	ns	N	P	ns	ns	P
	MCFR	--	--	--	--	--	P	P	P	N	N	N	ns	ns	N	ns	ns
Sodium	MAR	--	--	--	--	--	--	ns	P	N	ns	P	N	P	P	N	ns
	UIR	--	--	--	--	--	--	P	P	ns	ns	ns	ns	ns	ns	ns	N
	NGLFR	--	--	--	--	--	--	P	P	P	P	P	ns	ns	P	N	N
	WAGR	--	--	--	--	--	--	P	P	N	N	ns	N	P	N	ns	P
	MCFR	--	--	--	--	--	--	P	ns	ns	ns	ns	P	N	N	P	N
Sulfate	MAR	--	--	--	--	--	--	--	ns	ns	ns	ns	ns	ns	P	ns	ns
	UIR	--	--	--	--	--	--	--	P	N	ns	ns	P	N	ns	ns	N
	NGLFR	--	--	--	--	--	--	--	ns	ns	ns	P	P	N	ns	ns	P
	WAGR	--	--	--	--	--	--	--	P	N	N	ns	N	P	ns	ns	P
	MCFR	--	--	--	--	--	--	--	P	N	N	ns	P	N	ns	ns	P
Bicarbonate	MAR	--	--	--	--	--	--	--	--	N	P	P	N	P	P	N	P
	UIR	--	--	--	--	--	--	--	--	N	ns	ns	N	P	P	N	N
	NGLFR	--	--	--	--	--	--	--	--	N	N	N	P	ns	ns	P	P
	WAGR	--	--	--	--	--	--	--	--	ns	ns	N	ns	P	ns	ns	ns
	MCFR	--	--	--	--	--	--	--	--	N	ns	ns	ns	ns	P	N	ns

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Forest land use[1]	Urban land use[1]	Land surface slope[1]	Subsurface contact time[1]	Estimated year of ground-water recharge[1]	Arsenic	Barium	Boron	Bromide	Dissolved Oxygen
Calcium	MAR	ns	ns	ns	ns	ns	ns	N	ns	P	P
	UIR	N	P	ns	P	ns	N	P	ns	P	ns
	NGLFR	N	ns	N	P	N	ns	P	P	P	N
	WAGR	ns	N	N	ns	ns	N	ns	P	P	N
	MCFR	N	ns	ns	ns	ns	ns	P	P	P	N
Chloride	MAR	P	P	ns	ns	P	ns	ns	N	P	ns
	UIR	N	P	ns	ns	ns	N	ns	ns	P	ns
	NGLFR	N	P	ns	N	ns	ns	P	ns	P	ns
	WAGR	ns	P	ns	ns	ns	ns	ns	ns	P	P
	MCFR	N	P	N	ns	ns	ns	P	P	P	N
Fluoride	MAR	N	N	N	P	N	P	P	P	ns	N
	UIR	ns	ns	ns	P	ns	P	P	P	ns	N
	NGLFR	N	ns	ns	P	ns	P	P	P	ns	N
	WAGR	ns	P	ns	ns	ns	P	ns	ns	P	P
	MCFR	P	N	ns	ns	ns	P	ns	ns	ns	N
Magnesium	MAR	ns	N	ns	ns	ns	ns	N	N	P	ns
	UIR	N	P	ns	P	P	ns	P	ns	P	ns
	NGLFR	N	ns	N	P	ns	ns	P	ns	P	N
	WAGR	ns	N	ns	N	ns	ns	N	ns	P	ns
	MCFR	N	ns	ns	P	ns	ns	P	P	P	N
Potassium	MAR	N	ns	ns	ns	ns	P	P	P	P	ns
	UIR	N	P	ns	ns	ns	ns	ns	P	P	ns
	NGLFR	N	P	ns	ns	P	ns	P	P	P	N
	WAGR	N	ns	N	ns	ns	P	ns	P	P	ns
	MCFR	ns	ns	ns	ns	ns	ns	ns	P	P	N
Sodium	MAR	ns	P	ns	P	N	P	P	P	P	ns
	UIR	N	P	ns	ns	ns	ns	ns	P	P	ns
	NGLFR	N	P	P	ns	N	ns	P	P	P	N
	WAGR	ns	ns	ns	ns	ns	ns	N	P	P	P
	MCFR	N	P	N	ns	ns	ns	P	P	P	N
Sulfate	MAR	ns	ns	ns	ns	ns	N	N	ns	P	ns
	UIR	N	P	ns	ns	ns	ns	P	ns	P	N
	NGLFR	N	P	N	ns	ns	ns	P	P	P	ns
	WAGR	ns	ns	ns	ns	ns	ns	N	ns	P	N
	MCFR	N	P	N	ns	ns	ns	P	P	P	N
Bicarbonate	MAR	ns	N	N	P	N	P	P	ns	ns	N
	UIR	N	P	ns	P	ns	ns	P	ns	P	ns
	NGLFR	N	ns	N	P	N	ns	P	ns	P	N
	WAGR	ns	N	ns	ns	ns	ns	ns	P	ns	N
	MCFR	N	ns	P	ns	N	P	P	ns	P	N

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Iron	Lithium	Manganese	Molybdenum	Nitrite plus nitrate as nitrogen	Silica	Strontium	pH	Well depth
Calcium	MAR	ns	P	P	ns	P	ns	ns	N	N
	UIR	ns	ns	P	ns	ns	P	ns	N	N
	NGLFR	P	P	P	ns	ns	P	P	ns	ns
	WAGR	P	P	P	ns	ns	N	P	N	ns
	MCFR	ns	P	ns	P	P	ns	P	P	ns
Chloride	MAR	N	N	ns	ns	P	N	N	N	ns
	UIR	ns	ns	ns	N	ns	ns	ns	N	ns
	NGLFR	ns	ns	ns	ns	P	N	P	N	ns
	WAGR	N	ns	N	ns	P	ns	P	ns	ns
	MCFR	ns	ns	ns	N	P	N	P	N	ns
Fluoride	MAR	P	P	ns	P	N	P	P	P	P
	UIR	P	P	ns	P	N	P	P	ns	P
	NGLFR	P	ns	P	P	N	P	P	ns	P
	WAGR	N	ns	N	P	P	P	ns	P	P
	MCFR	P	P	P	P	N	ns	ns	P	ns
Magnesium	MAR	ns	ns	P	ns	N	ns	N	N	ns
	UIR	ns	ns	ns	P	ns	P	ns	N	P
	NGLFR	P	P	P	ns	ns	P	ns	ns	P
	WAGR	P	ns	P	ns	ns	ns	P	ns	N
	MCFR	ns	P	ns	ns	P	P	P	P	ns
Potassium	MAR	ns	P	ns	P	ns	P	P	P	P
	UIR	N	P	P	ns	P	ns	ns	N	N
	NGLFR	P	P	P	ns	ns	ns	P	N	ns
	WAGR	ns	P	ns	P	ns	P	ns	P	P
	MCFR	ns	ns	P	P	ns	ns	P	ns	N
Sodium	MAR	P	P	ns	P	N	P	P	ns	P
	UIR	ns	ns	ns	ns	ns	ns	ns	N	ns
	NGLFR	ns	ns	P	ns	ns	ns	P	N	ns
	WAGR	ns	ns	N	P	P	P	P	P	P
	MCFR	ns	ns	P	ns	ns	ns	P	N	ns
Sulfate	MAR	N	P	P	ns	P	N	ns	N	N
	UIR	P	ns	P	P	N	ns	ns	N	ns
	NGLFR	ns	P	P	P	ns	ns	P	ns	ns
	WAGR	ns	ns	P	P	ns	ns	P	ns	ns
	MCFR	ns	ns	ns	ns	P	ns	P	P	N
Bicarbonate	MAR	P	ns	ns	P	N	P	ns	N	P
	UIR	ns	ns	ns	ns	N	P	P	N	P
	NGLFR	P	P	P	ns	ns	P	ns	ns	ns
	WAGR	P	P	P	ns	ns	ns	P	N	ns
	MCFR	ns	ns	ns	P	N	P	P	P	ns

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Recharge[1]	Average annual precipitation[1]	Average annual temperature[1]	Coarse-grained surficial material[1]	Fine-grained surficial material[1]	Clay in soil[1]	Vertical permeability[1]	Agricultural land use[1]	Forest land use[1]	Urban land use[1]	Land surface slope[1]	Subsurface contact time[1]	Estimated year of ground-water recharge[1]	Arsenic	Barium
Recharge[1]	MAR	--	P	ns	P	N	ns	N	N	P	ns	ns	P	P	N	ns
	UIR	--	P	ns	ns	ns	P	N	P	P	N	P	ns	ns	ns	ns
	NGLFR	--	P	P	N	P	P	N	N	P	P	P	N	P	ns	N
	WAGR	--	P	N	P	P	P	N	N	ns	P	ns	P	ns	N	P
	MCFR	--	P	P	ns	ns	N	P	N	ns	P	N	ns	ns	N	ns
Average annual precipitation[1]	MAR	--	--	P	ns	ns	P	N	P	ns	ns	N	P	N	ns	P
	UIR	--	--	P	ns	ns	P	N	P	ns	N	P	P	P	ns	ns
	NGLFR	--	--	P	N	P	P	N	N	P	P	P	N	P	ns	ns
	WAGR	--	--	N	P	P	P	N	N	ns	ns	N	P	N	N	P
	MCFR	--	--	P	ns	ns	N	ns	N	ns	ns	N	ns	ns	ns	ns
Average annual temperature[1]	MAR	--	--	--	N	P	P	N	P	N	ns	N	P	N	ns	P
	UIR	--	--	--	ns	ns	P	N	P	ns	N	ns	P	P	ns	P
	NGLFR	--	--	--	ns	ns	P	N	N	ns	P	ns	ns	P	ns	ns
	WAGR	--	--	--	N	N	N	ns	ns	ns	P	P	N	ns	ns	ns
	MCFR	--	--	--	N	ns	P	ns	N	N	P	N	P	ns	ns	P
Coarse-grained surficial material[1]	MAR	--	--	--	--	N	N	P	N	P	P	ns	N	P	N	ns
	UIR	--	--	--	--	N	N	P	N	N	P	N	ns	ns	ns	ns
	NGLFR	--	--	--	--	N	N	P	P	N	N	N	ns	ns	ns	ns
	WAGR	--	--	--	--	N	P	ns	N	ns	ns	N	ns	ns	ns	P
	MCFR	--	--	--	--	N	N	P	ns	N	P	N	ns	ns	ns	P
Fine-grained surficial material[1]	MAR	--	--	--	--	--	P	N	P	N	N	ns	P	N	P	ns
	UIR	--	--	--	--	--	P	N	P	P	N	P	ns	N	ns	ns
	NGLFR	--	--	--	--	--	P	N	N	P	ns	P	ns	ns	ns	ns
	WAGR	--	--	--	--	--	P	N	P	N	N	N	P	ns	N	ns
	MCFR	--	--	--	--	--	P	N	ns	P	ns	P	ns	ns	ns	N
Clay in soil[1]	MAR	--	--	--	--	--	--	N	P	N	N	ns	P	N	P	P
	UIR	--	--	--	--	--	--	N	P	ns	N	P	P	P	ns	P
	NGLFR	--	--	--	--	--	--	N	ns	ns	ns	P	P	P	ns	ns
	WAGR	--	--	--	--	--	--	N	N	N	ns	N	P	P	N	P
	MCFR	--	--	--	--	--	--	N	P	ns	N	P	P	ns	ns	ns
Vertical permeability[1]	MAR	--	--	--	--	--	--	--	N	P	P	ns	N	P	ns	N
	UIR	--	--	--	--	--	--	--	N	ns	P	N	N	N	ns	N
	NGLFR	--	--	--	--	--	--	--	P	N	ns	N	N	N	ns	ns
	WAGR	--	--	--	--	--	--	--	ns	P	ns	ns	N	N	P	N
	MCFR	--	--	--	--	--	--	--	N	ns	P	N	ns	ns	ns	ns
Agricultural land use[1]	MAR	--	--	--	--	--	--	--	--	N	N	N	P	N	P	ns
	UIR	--	--	--	--	--	--	--	--	P	N	ns	P	ns	P	ns
	NGLFR	--	--	--	--	--	--	--	--	N	N	N	P	P	ns	ns
	WAGR	--	--	--	--	--	--	--	--	N	N	ns	ns	ns	ns	N
	MCFR	--	--	--	--	--	--	--	--	ns	N	P	ns	ns	ns	N

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Boron	Bromide	Dissolved Oxygen	Iron	Lithium	Manganese	Molybdenum	Nitrite plus nitrate as nitrogen	Silica	Strontium	pH	Well depth
Recharge[1]	MAR	N	ns	ns	P	N	ns	ns	N	N	ns	ns	N
	UIR	N	ns	ns	ns	ns	N	ns	ns	N	ns	ns	P
	NGLFR	N	ns	P	N	ns	N	N	ns	N	ns	N	ns
	WAGR	P	N	N	P	P	P	ns	N	N	ns	N	N
	MCFR	ns	ns	P	N	ns	N	ns	ns	P	ns	N	ns
Average annual precipitation[1]	MAR	N	ns	ns	P	N	P	P	N	N	ns	N	ns
	UIR	P	ns	ns	N	N	N	ns	ns	N	ns	ns	P
	NGLFR	ns	ns	P	N	N	N	ns	ns	N	ns	N	ns
	WAGR	P	N	N	P	P	P	ns	N	N	ns	N	N
	MCFR	ns	ns	ns	ns	ns	ns	ns	ns	P	ns	ns	ns
Average annual temperature[1]	MAR	ns	ns	ns	ns	N	ns	ns	ns	N	ns	N	P
	UIR	P	P	ns	N	N	N	P	P	N	P	ns	P
	NGLFR	P	ns	P	N	N	N	ns	ns	N	ns	N	ns
	WAGR	N	P	P	N	N	N	ns	P	P	ns	P	P
	MCFR	P	ns	ns	ns	ns	ns	ns	P	P	ns	ns	P
Coarse-grained surficial material[1]	MAR	N	ns	ns	ns	N	ns	ns	P	N	ns	ns	ns
	UIR	ns	N	ns	ns	ns	ns	ns	ns	ns	ns	ns	N
	NGLFR	P	ns	N	P	ns	P	P	ns	P	ns	P	ns
	WAGR	ns	N	N	P	P	P	ns	N	N	ns	N	N
	MCFR	P	P	ns	ns	ns	ns	ns	ns	ns	P	ns	ns
Fine-grained surficial material[1]	MAR	P	ns	ns	ns	P	ns	P	N	P	ns	ns	ns
	UIR	ns	P	ns	ns	ns	ns	ns	ns	ns	ns	ns	P
	NGLFR	N	ns	P	N	ns	N	N	ns	N	ns	N	ns
	WAGR	P	P	N	P	P	P	ns	N	N	ns	ns	P
	MCFR	ns	N	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Clay in soil[1]	MAR	P	ns	ns	ns	P	ns	ns	ns	ns	P	N	ns
	UIR	P	ns	ns	N	ns	N	ns	ns	N	P	ns	P
	NGLFR	ns	ns	N	ns	ns	ns	ns	N	N	ns	N	ns
	WAGR	P	ns	N	P	P	P	N	N	N	ns	N	ns
	MCFR	N	N	ns	ns	ns	ns	ns	ns	ns	ns	P	P
Vertical permeability[1]	MAR	ns	ns	ns	ns	ns	ns	N	ns	P	N	ns	ns
	UIR	ns	ns	ns	P	ns	P	ns	ns	P	N	ns	N
	NGLFR	P	ns	ns	P	ns	ns	P	P	P	P	P	ns
	WAGR	ns	ns	ns	ns	N	ns	ns	ns	ns	ns	P	N
	MCFR	P	ns	ns	ns	ns	ns	ns	ns	ns	ns	N	ns
Agricultural land use[1]	MAR	ns	ns	ns	ns	P	ns	P	ns	ns	P	ns	ns
	UIR	ns	N	ns	ns	ns	N	P	ns	N	ns	P	P
	NGLFR	P	ns	ns	ns	ns	ns	P	P	ns	P	ns	ns
	WAGR	ns	ns	P	ns	ns	ns	ns	ns	P	ns	P	ns
	MCFR	N	ns	ns	ns	ns	ns	ns	ns	N	ns	P	ns

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Recharge[1]	Average annual precipitation[1]	Average annual temperature[1]	Coarse-grained surficial material[1]	Fine-grained surficial material[1]	Clay in soil[1]	Vertical permeability[1]	Agricultural land use[1]	Forest land use[1]	Urban land use[1]	Land surface slope[1]	Subsurface contact time[1]	Estimated year of ground-water recharge[1]	Arsenic	Barium
Forest land use[1]	MAR	--	--	--	--	--	--	--	--	--	P	P	N	P	N	ns
	UIR	--	--	--	--	--	--	--	--	--	N	P	ns	ns	ns	N
	NGLFR	--	--	--	--	--	--	--	--	--	ns	P	N	ns	ns	N
	WAGR	--	--	--	--	--	--	--	--	--	ns	ns	N	ns	ns	ns
	MCFR	--	--	--	--	--	--	--	--	--	N	P	N	ns	ns	N
Urban land use[1]	MAR	--	--	--	--	--	--	--	--	--	--	ns	N	P	N	ns
	UIR	--	--	--	--	--	--	--	--	--	--	ns	ns	ns	N	ns
	NGLFR	--	--	--	--	--	--	--	--	--	--	P	N	N	ns	P
	WAGR	--	--	--	--	--	--	--	--	--	--	P	ns	ns	P	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	N	ns	ns	ns	P
Land surface slope[1]	MAR	--	--	--	--	--	--	--	--	--	--	--	N	P	ns	ns
	UIR	--	--	--	--	--	--	--	--	--	--	--	ns	P	N	ns
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	N	ns	N	N
	WAGR	--	--	--	--	--	--	--	--	--	--	--	N	ns	ns	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	--	N	ns	ns	N
Subsurface contact time[1]	MAR	--	--	--	--	--	--	--	--	--	--	--	--	N	P	P
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	P	ns	P
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	ns	P	ns
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	P	ns	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns	ns
Estimated year of ground-water recharge[1]	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	N	ns
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	N	ns
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	N	N
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns
Arsenic	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	ns
Barium	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Boron	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Boron	Bromide	Dissolved Oxygen	Iron	Lithium	Manganese	Molybdenum	Nitrite plus nitrate as nitrogen	Silica	Strontium	pH	Well depth
Forest land use[1]	MAR	N	ns	ns	ns	N	ns	N	N	ns	ns	ns	ns
	UIR	ns	ns	ns	ns	ns	ns	ns	ns	N	ns	ns	P
	NGLFR	N	N	ns	N	N	N	N	N	N	N	P	ns
	WAGR	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
	MCFR	N	N	ns	ns	ns	ns	ns	N	N	N	ns	ns
Urban land use[1]	MAR	ns	P	ns	ns	ns	ns	N	ns	ns	N	N	P
	UIR	ns	P	ns	ns	ns	P	N	ns	P	ns	N	N
	NGLFR	ns	P	ns	ns	P	ns	ns	ns	ns	ns	N	ns
	WAGR	ns	ns	ns	ns	N	N	ns	P	ns	ns	ns	ns
	MCFR	P	P	ns	ns	ns	ns	ns	P	P	P	ns	ns
Land surface slope[1]	MAR	P	ns	P	N	P	ns	ns	ns	ns	P	N	N
	UIR	ns	ns	P	N	ns	N	N	ns	N	ns	N	P
	NGLFR	N	ns	P	N	ns	N	N	ns	N	ns	N	P
	WAGR	N	P	P	N	N	N	ns	P	P	ns	P	ns
	MCFR	N	N	ns	ns	ns	ns	ns	N	N	ns	P	ns
Subsurface contact time[1]	MAR	ns	ns	N	ns	ns	ns	P	ns	N	ns	ns	ns
	UIR	P	ns	ns	ns	N	ns	P	ns	N	ns	ns	ns
	NGLFR	P	ns	N	P	N	P	P	N	P	ns	ns	N
	WAGR	ns	ns	N	P	P	ns	ns	N	N	ns	N	P
	MCFR	P	P	ns	ns	ns	ns	ns	P	P	ns	ns	ns
Estimated year of ground-water recharge[1]	MAR	ns	ns	P	N	ns	N	ns	P	N	ns	ns	N
	UIR	ns	ns	P	N	N	N	N	P	N	ns	N	ns
	NGLFR	ns	N	P	N	--	N	N	P	ns	ns	ns	ns
	WAGR	ns	ns	ns	ns	ns	ns	--	ns	ns	ns	ns	ns
	MCFR	ns	N	P	ns	ns	ns	N	P	ns	ns	ns	ns
Arsenic	MAR	P	ns	N	P	P	ns	P	N	P	P	P	P
	UIR	ns	ns	N	P	P	P	P	N	P	ns	P	ns
	NGLFR	ns	ns	N	P	ns	P	P	N	ns	P	P	ns
	WAGR	ns	ns	ns	ns	ns	ns	P	ns	P	ns	P	ns
	MCFR	ns	ns	N	P	ns	P	P	N	ns	ns	ns	ns
Barium	MAR	ns	ns	N	P	ns	ns	P	N	P	P	ns	P
	UIR	ns	P	N	P	ns	P	P	N	ns	P	ns	P
	NGLFR	P	P	N	P	P	P	P	ns	P	P	ns	ns
	WAGR	P	ns	ns	ns	P	ns	N	ns	ns	ns	N	P
	MCFR	P	P	ns	ns	ns	ns	ns	P	P	P	ns	N
Boron	MAR	--	P	ns	ns	P	ns	P	ns	P	P	ns	ns
	UIR	--	P	N	ns	P	ns	P	ns	ns	P	ns	ns
	NGLFR	--	ns	N	P	ns	P	P	ns	P	P	ns	N
	WAGR	--	ns	N	P	P	P	ns	ns	N	ns	N	P
	MCFR	--	P	N	ns	ns	ns	ns	P	P	P	ns	ns

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Arsenic	Barium	Boron	Bromide	Dissolved Oxygen	Iron	Lithium	Manganese	Molybdenum	Nitrite plus nitrate as nitrogen	Silica	Strontium	pH	Well depth
Bromide	MAR	--	--	--	--	ns	ns	P	ns	N	ns	ns	ns	N	ns
	UIR	--	--	--	--	N	P	ns	ns	ns	N	ns	ns	ns	P
	NGLFR	--	--	--	--	N	ns	P	P	ns	ns	P	ns	ns	P
	WAGR	--	--	--	--	P	ns	ns	N	ns	P	P	P	P	P
	MCFR	--	--	--	--	N	P	P	ns	ns	ns	ns	P	ns	ns
Dissolved Oxygen	MAR	--	--	--	--	--	N	ns	N	N	P	ns	N	N	N
	UIR	--	--	--	--	--	N	ns	N	N	P	ns	N	ns	N
	NGLFR	--	--	--	--	--	N	ns	N	N	P	N	N	ns	ns
	WAGR	--	--	--	--	--	N	N	N	ns	P	P	ns	P	ns
	MCFR	--	--	--	--	--	N	N	N	N	P	ns	N	ns	ns
Iron	MAR	--	--	--	--	--	--	ns	P	P	N	P	ns	P	P
	UIR	--	--	--	--	--	--	ns	P	P	N	P	ns	ns	ns
	NGLFR	--	--	--	--	--	--	ns	P	P	N	P	P	ns	P
	WAGR	--	--	--	--	--	--	P	P	P	N	N	ns	N	ns
	MCFR	--	--	--	--	--	--	ns	P	ns	N	ns	ns	ns	ns
Lithium	MAR	--	--	--	--	--	--	--	N	ns	ns	P	P	ns	ns
	UIR	--	--	--	--	--	--	--	ns	P	N	P	ns	ns	ns
	NGLFR	--	--	--	--	--	--	--	P	P	ns	P	P	ns	ns
	WAGR	--	--	--	--	--	--	--	P	ns	ns	N	P	N	P
	MCFR	--	--	--	--	--	--	--	ns	ns	ns	ns	P	ns	ns
Manganese	MAR	--	--	--	--	--	--	--	--	P	N	ns	ns	N	N
	UIR	--	--	--	--	--	--	--	--	P	N	P	ns	ns	N
	NGLFR	--	--	--	--	--	--	--	--	P	N	ns	P	N	N
	WAGR	--	--	--	--	--	--	--	--	P	N	N	ns	N	N
	MCFR	--	--	--	--	--	--	--	--	P	N	N	P	ns	N
Molybdenum	MAR	--	--	--	--	--	--	--	--	--	N	ns	P	P	ns
	UIR	--	--	--	--	--	--	--	--	--	N	ns	P	P	ns
	NGLFR	--	--	--	--	--	--	--	--	--	N	ns	P	P	ns
	WAGR	--	--	--	--	--	--	--	--	--	ns	ns	ns	P	ns
	MCFR	--	--	--	--	--	--	--	--	--	N	ns	P	P	P
Nitrite plus nitrate as nitrogen	MAR	--	--	--	--	--	--	--	--	--	--	ns	ns	N	N
	UIR	--	--	--	--	--	--	--	--	--	--	N	N	N	N
	NGLFR	--	--	--	--	--	--	--	--	--	--	ns	ns	ns	ns
	WAGR	--	--	--	--	--	--	--	--	--	--	P	P	P	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	P	P	N	ns
Silica	MAR	--	--	--	--	--	--	--	--	--	--	--	P	ns	P
	UIR	--	--	--	--	--	--	--	--	--	--	--	ns	ns	ns
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	P	ns	ns
	WAGR	--	--	--	--	--	--	--	--	--	--	--	P	P	P
	MCFR	--	--	--	--	--	--	--	--	--	--	--	P	P	P

Table A18. Significant (p-value<0.05) Spearman rank correlations for pair-wise comparison of select constituents by hydrochemical region.—Continued

[--, indicates correlation not shown or shown elsewhere in the table; ns, indicates an insignificant correlation; P, indicates positive significant correlation (p-value < 0.05); N, indicates negative significant correlation (p-value < 0.05); MAR, Midwestern Agricultural Region; UIR, Urban-Influenced Region; NGLFR, Northern and Great Lakes Forested Region; WAGR, Western Agriculture and Grassland Region; MCFR, Mountain and Coastal Forested Region]

Correlation variable	Hydro-chemical region	Arsenic	Barium	Boron	Bromide	Dissolved Oxygen	Iron	Lithium	Manganese	Molybdenum	Nitrite plus nitrate as nitrogen	Silica	Strontium	pH	Well depth
Strontium	MAR	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	ns	P
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	ns	ns
pH	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	P
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	P
Well Depth	MAR	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	UIR	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	NGLFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	WAGR	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	MCFR	--	--	--	--	--	--	--	--	--	--	--	--	--	--

[1] Variable is from point overlay on spatial data.

Table A19. Comparison of select water-quality characteristics as described by regional-scale hydrochemical region analysis, and a local-scale study in Michigan.

[mg/L, milligrams per liter]

	Hydrochemical region						Wahrer and others, 1996	
	Midwestern Agricultural Region (most of the Saginaw Lowland is in this region)			Urban-Influenced Region			Glaciofluvial[1] aquifer in Michigan	
Constituent	Mean	Median	Maximum	Mean	Median	Maximum	Maximum	Typical values
Dissolved Solids, mg/L	495	424	4,091	499	460	1,728	12,000	500–12,000
Chloride, mg/L	37	14	2,300	75	43	800	6,700	100–1,000
Iron, mg/L	1	0	10	1	0	28	6	0.1–1
Sulfate, mg/L	67	37	1,386	55	42	520	18,000	10–100

[1] The glaciofluvial aquifer is part of the glacial aquifer system.

Table A20. Comparison of select water-quality characteristics as described by regional-scale hydrochemical region analysis, and a local-scale study in Connecticut.

[mg/L, milligrams per liter; µg/L, micrograms per liter; <, less than].

	Hydrochemical region		Grady, 1988	
	Northern and Great Lakes Forested Region	Mountain and Coastal Forested Region	Minimum and maximum medians reported for four land-use areas of stratified-drift aquifers[1] in Connecticut (land-use category is in parentheses)	
Constituent	Median	Median	Minimum	Maximum
Calcium, in mg/L	61	19	9.2 (undeveloped)	25 (industrial)
Magnesium, in mg/L	17	4.5	1.5 (undeveloped)	4.7 (industrial)
Sulfate, in mg/L	15	13	10 (undeveloped)	18 (industrial)
Boron, in µg/L	15	15.2	<20 (undeveloped)	40 (industrial)
Lithium, in µg/L	6	6	3.2 (residential)	4 (agricultural and industrial)
Strontium, in µg/L	62	86	37 (undeveloped)	91 (urban)

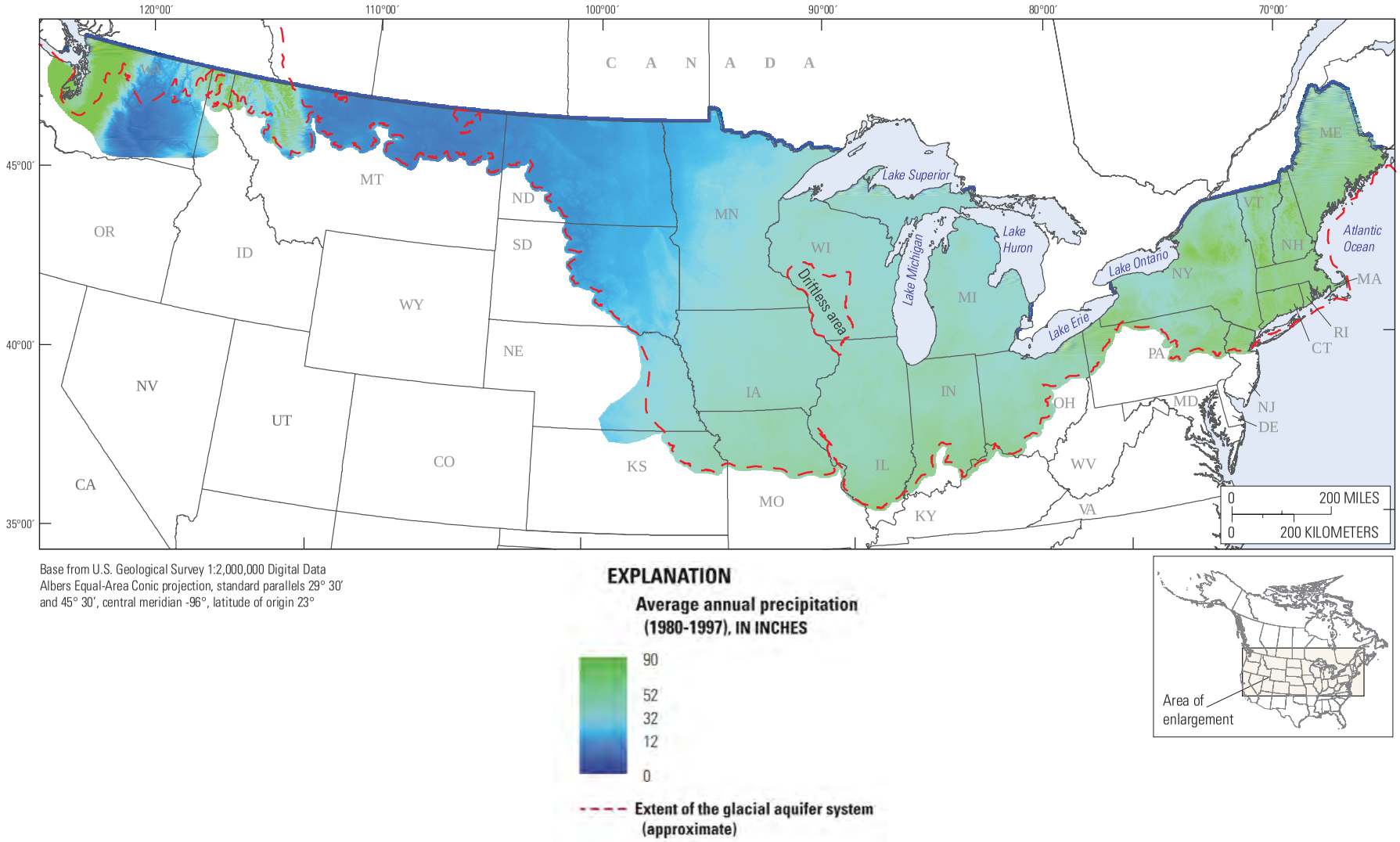
[1] The stratified drift aquifers are part of the glacial aquifer system.

Appendix B—Spatial Data Layers Used to Map Hydrochemical Regions

Table B1. Descriptions and sources of environmental spatial data layers used to map hydrochemical regions.

[GIS, Geographic Information System]

Figure in this report	Environmental characteristic (spatial layer used in GIS)	Data scale	Reference	Description of data	Conceptual framework representation
B1	Average annual precipitation (1980–1997), in centimeters	Continuous	Daymet, 2005	This spatial layer was derived from 1-kilometer resolution Daymet model data, which represented 18 year (1980–1997) precipitation averages obtained from terrain-adjusted daily climatological observations. In this report, precipitation data is in inches.	Vulnerability to natural sources
B2	Average annual temperature (1980–1997), in degrees Celsius	Continuous	Daymet, 2005	This spatial layer was derived from 1-kilometer resolution Daymet model data, which represented 18 year (1980–1997) temperature averages obtained from terrain-adjusted daily climatological observations. In this report, temperature data is in degrees Fahrenheit	Vulnerability to natural sources
B3	Land surface slope, in percent	Continuous	U.S. Geological Survey, 2001	This spatial layer was derived from a digital elevation model, resampled to 500-meter cell size, and calculated as a percent slope (100 times the quotient of change in elevation divided by change in ground distance). Percent slope approaches infinity as the measured slope angle approaches 90 degrees.	Vulnerability to natural sources
B4	1990s land use	Discrete	Vogelmann and others, 2001	Original land cover categories were simplified and grouped for analysis: urban (land cover codes 21 through 23 and 85); bare rock, quarry, or transitional (land cover codes 31 through 33); forest or shrub (land cover codes 41 through 43, and 51); orchards and vineyards (land cover code 61); grasslands or herbaceous (land cover 71); agriculture (land cover codes 81 through 84); wetlands (land cover codes 91 and 92); open water (land cover code 11); or perennial ice or snow (land cover code 12).	Vulnerability to anthropogenic sources
B5	Vertical soil permeability, in inches per hour	Continuous	U.S. Department of Agriculture, 1994; and Hitt, U.S. Geological Survey, unpublished data	This spatial layer was derived from the State Soil Geographic Database (STATSGO) (U.S. Department of Agriculture, 1994). Vertical soil permeability was calculated as the total thickness of soil in a mapping unit divided by the sum, for all soil layers, of the ratio of thickness of the soil layer to the permeability of the soil layer (Wolock, 1997; Hitt, U.S. Geological Survey, unpublished data). A soil layer is defined as a soil horizon in STATSGO (U.S. Department of Agriculture, 1994). STATSGO has data for multiple soil layers in each mapping unit.	Intrinsic susceptibility
B6	Soil clay content, in percent of material less than 2 millimeter in size	Continuous	U.S. Department of Agriculture, 1994; and Wolock, 1997	This spatial layer was derived from the State Soil Geographic Database (STATSGO) (U.S. Department of Agriculture, 1994). Percent soil clay content was calculated as layer-thickness and component-area weighted averages for the mapping unit (Wolock, 1997).	Intrinsic susceptibility
B7	Texture of surficial deposits	Discrete	Fullerton and others, 2004; Clawges and Price, 1999; National Park Service, 1999	This spatial layer is a composite of several different data sets. The texture of surficial deposits was specifically interpreted for use in this analysis based on generalized geologic descriptions of surficial deposits. Surficial texture in the Eastern and Central United States (East of 102 Degrees West Longitude) is modified from Fullerton and others, 2004. Surficial texture West of 102 Degrees West Longitude is modified from Clawges and Price, 1999. Interpretations of surficial texture in Alaska were not made for this analysis (National Park Service, 1999).	Intrinsic susceptibility
B8	Type of surficial deposit	Discrete	Soller and Packard, 1998; Clawges and Price, 1999; National Park Service, 1999	This spatial layer is a composite of several different data sets. The type of surficial deposits in the glaciated area east of the Rocky Mountains is based on Soller and Packard, 1998 (1:1,000,000 scale). The type of surficial deposits shown in the glaciated area west of the Rocky Mountains is based on Clawges and Price, 1999 (1:7,500,00 scale). Types of surficial deposits in Alaska are from National Park Service, 1999 (1:1,584,000).	Intrinsic susceptibility, physical setting
B9	Estimated annual natural ground-water recharge, in inches per year	Continuous	Wolock, 2003b	This spatial layer was created by multiplying a raster grid of base-flow index (BFI) values (Wolock, 2003b) by a grid of mean annual runoff values (Gebert and others, 1987).	Intrinsic susceptibility



Base from U.S. Geological Survey 1:2,000,000 Digital Data
Albers Equal-Area Conic projection, standard parallels 29° 30′ and 45° 30′, central meridian -96°, latitude of origin 23°

Figure B1. Average annual precipitation (1980–1997) across the glacial aquifer system in the northern United States (modified from Daymet, 2005). No data were available for Alaska.

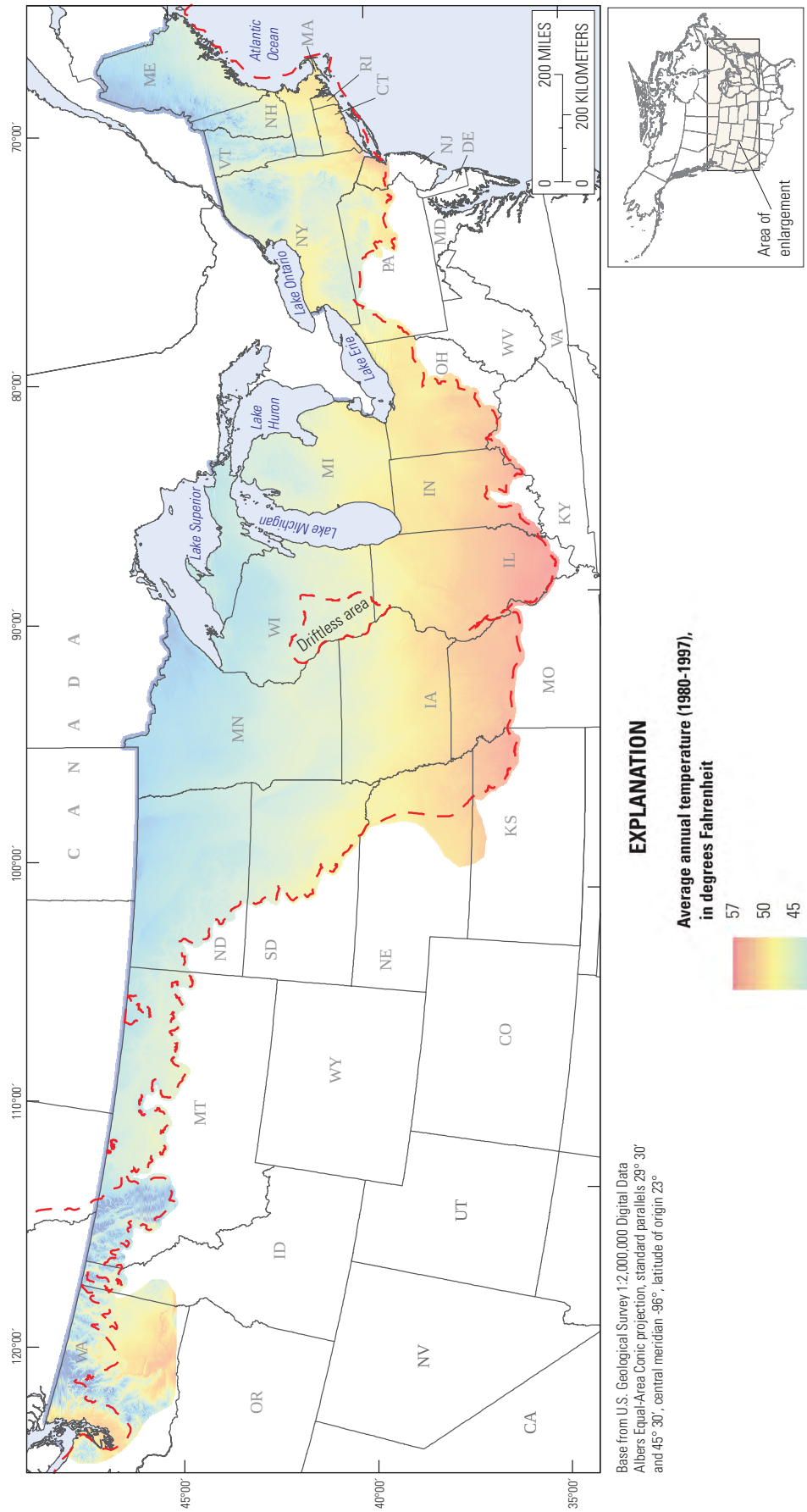
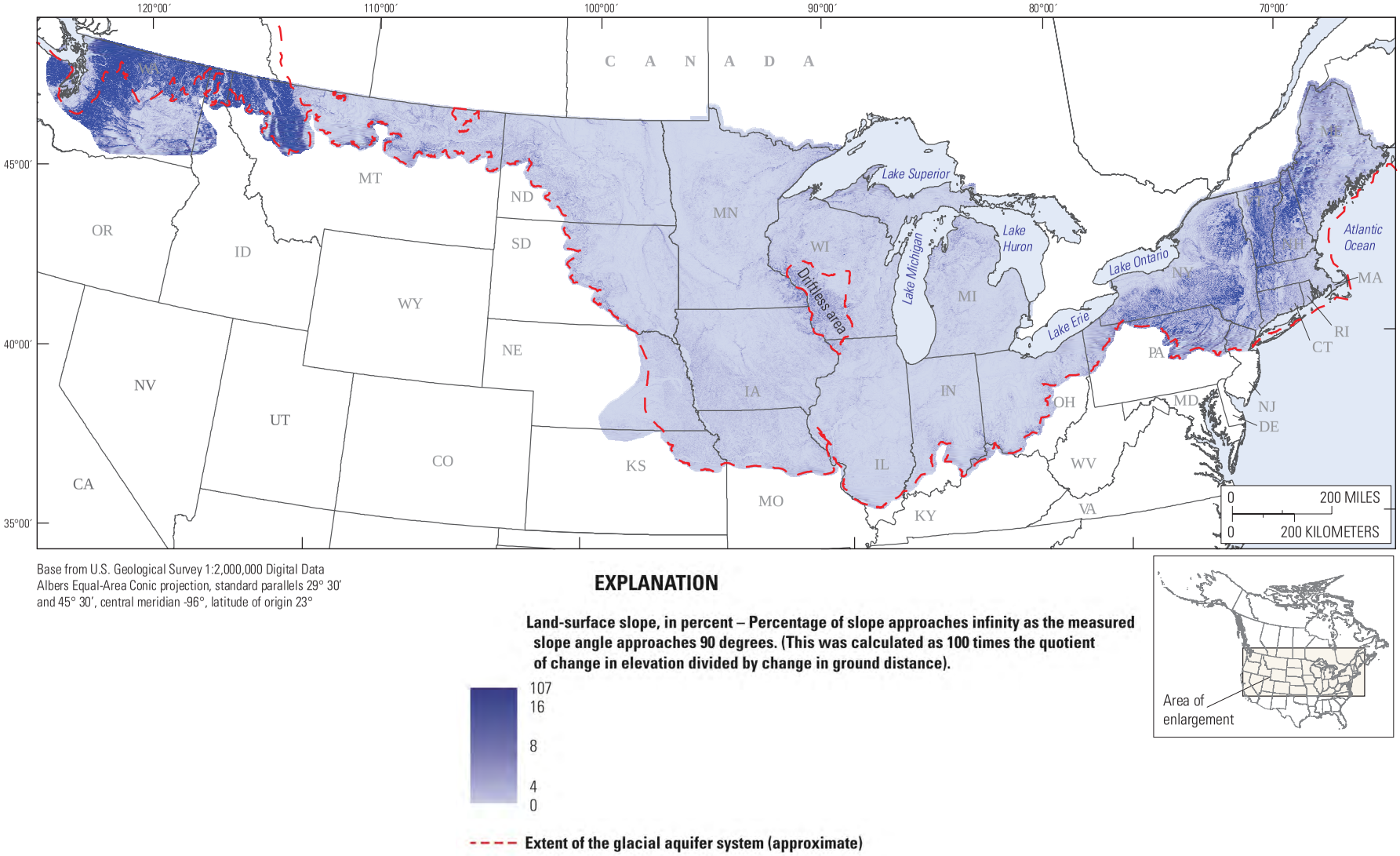


Figure B2. Average annual temperature (1980–1997) across the glacial aquifer system in the northern United States (modified from Daymet, 2005). No data were available for Alaska.



Base from U.S. Geological Survey 1:2,000,000 Digital Data
Albers Equal-Area Conic projection, standard parallels 29° 30′ and 45° 30′, central meridian -96°, latitude of origin 23°

Figure B3. Land-surface slope across the glacial aquifer system in the northern United States (modified from U.S. Geological Survey, 2001). No data were available for Alaska.

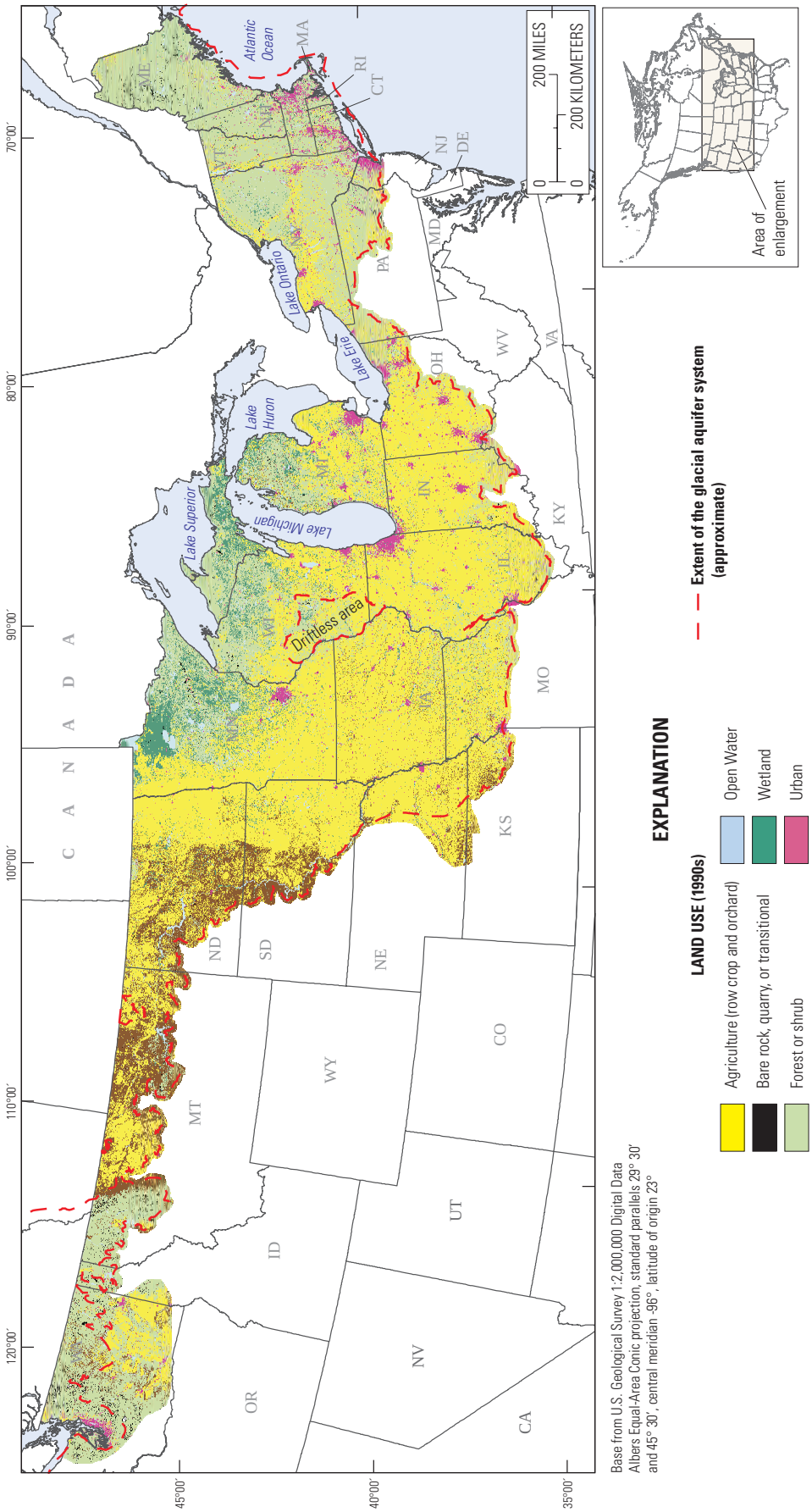
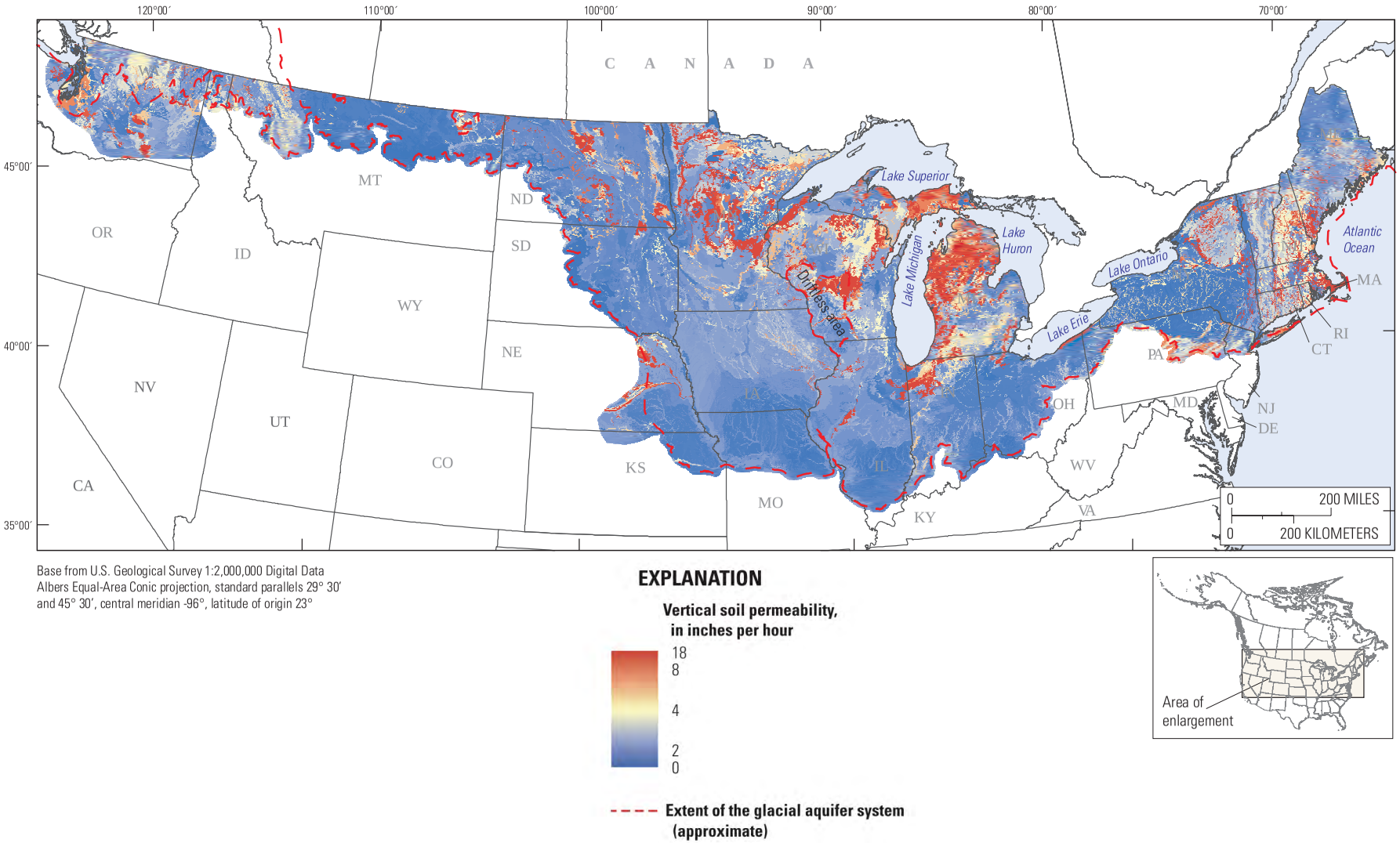


Figure B4. Land use (1990s) across the glacial aquifer system in the northern United States (modified from Vogelmann and others, 2001). No data were available for Alaska.



Base from U.S. Geological Survey 1:2,000,000 Digital Data
Albers Equal-Area Conic projection, standard parallels 29° 30′ and 45° 30′, central meridian -96°, latitude of origin 23°

Figure B5. Vertical soil permeability in the glacial aquifer system in the northern United States (modified from U.S. Department of Agriculture, 1994). No data were available for Alaska.

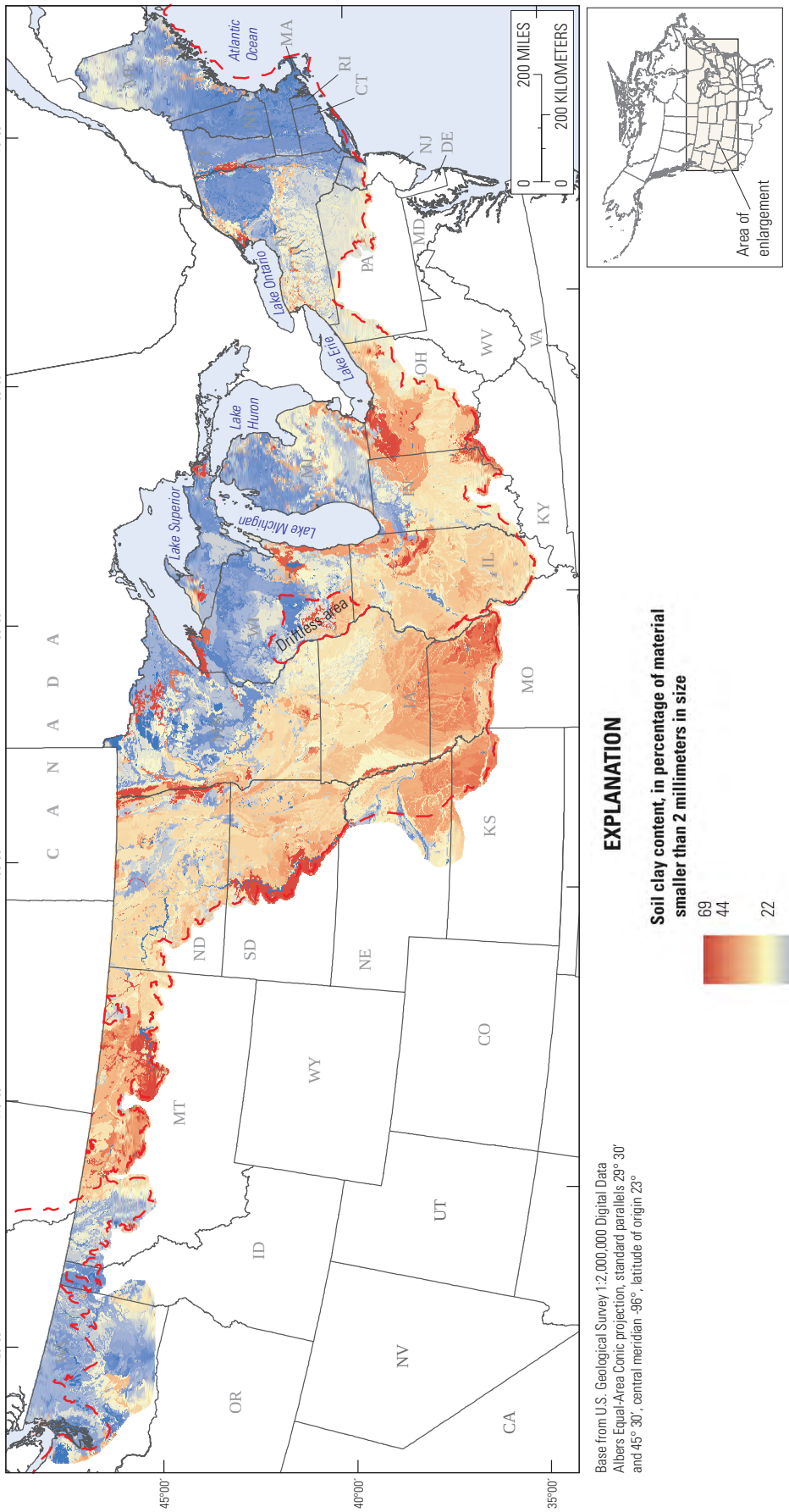


Figure B6. Soil clay content in the glacial aquifer system in the northern United States (modified from Wolock, 1997). No data were available for Alaska.

EXPLANATION

Texture of surficial deposits

- Coarse-grained sediment
- Fine-grained sediment (includes till and patchy Quaternary or organic-rich sediment)
- Undifferentiated or mixed fine- and coarse-grained sediment
- Unclassified

Extent of the glacial aquifer system (approximate)

Figure B7. Texture of surficial deposits in the glacial aquifer system in the northern United States (modified from Fullerton and others, 2004; Clawges and Price, 1999; and National Park Service, 1999). No data were available for Alaska.

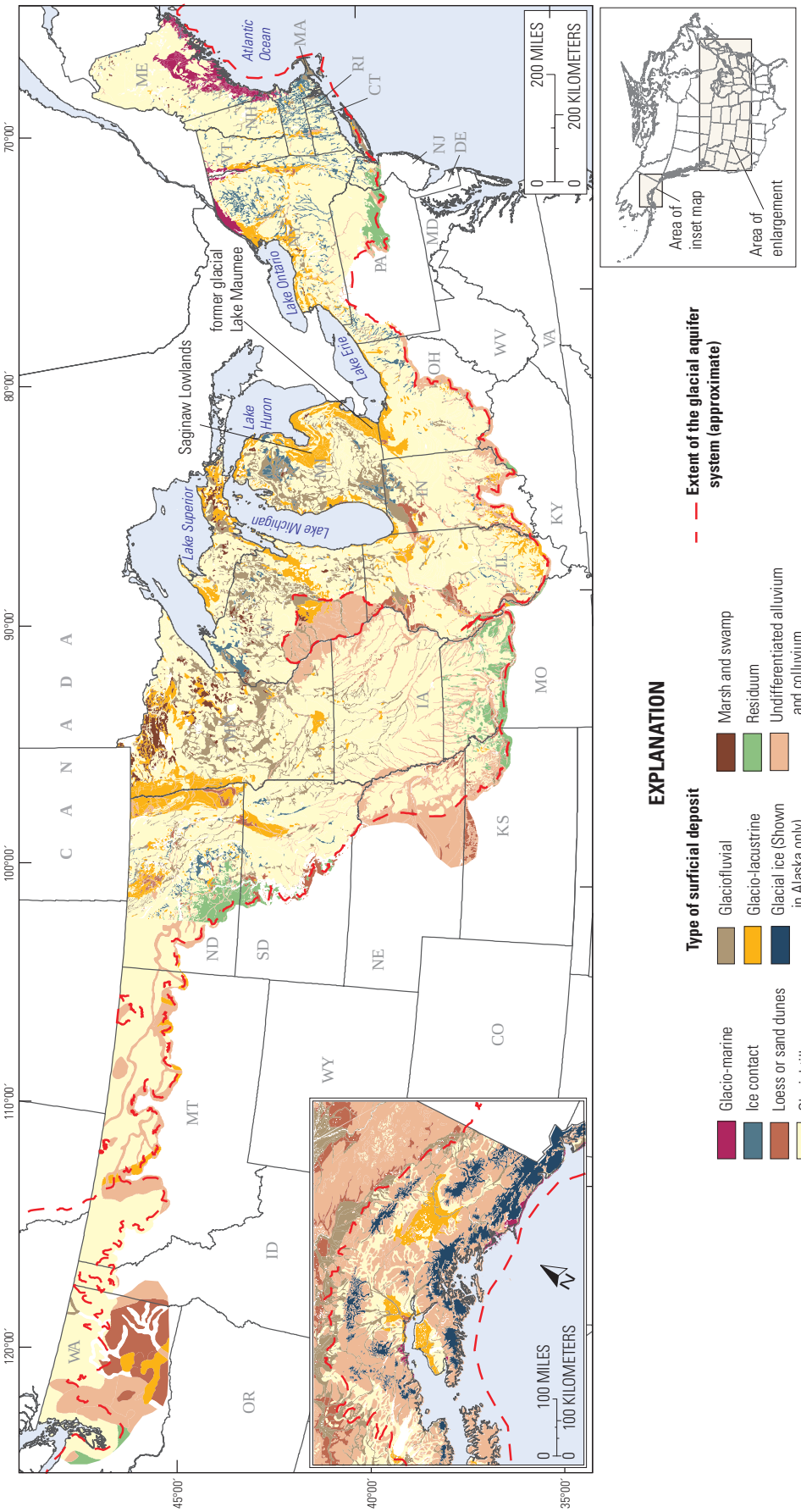
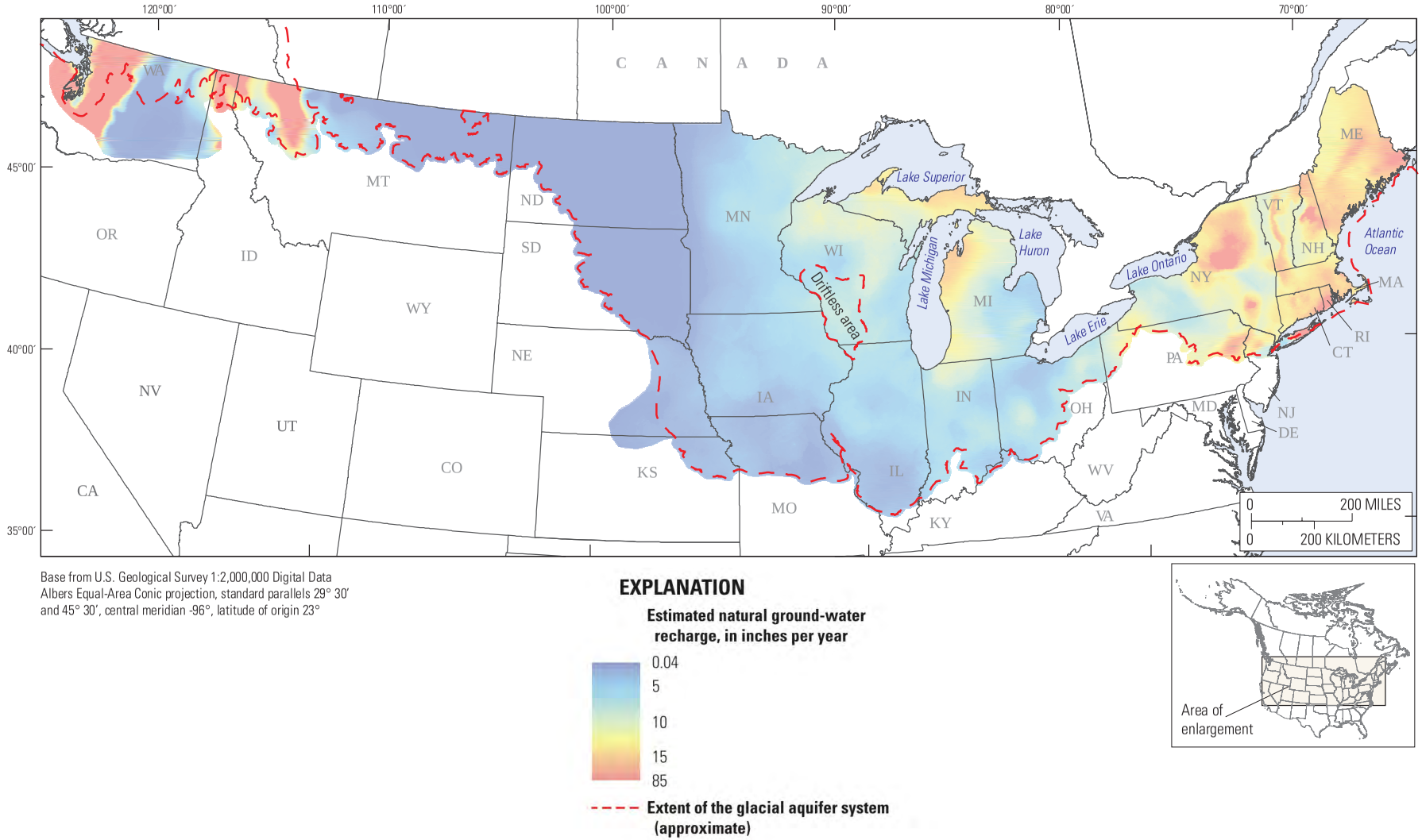


Figure B8. Type of surficial deposits in the glacial aquifer system in the northern United States (modified from Fullerton and others, 2004; Clawges and Price, 1999; and National Park Service, 1999). Saginaw lowlands location is from Westjohn and others (1994) and Wahrer and others (1996). Former glacial Lake Maumee location is from Geological Survey of Canada (1989).



Base from U.S. Geological Survey 1:2,000,000 Digital Data
Albers Equal-Area Conic projection, standard parallels 29° 30′ and 45° 30′, central meridian -96°, latitude of origin 23°

EXPLANATION

Estimated natural ground-water recharge, in inches per year

0.04
5
10
15
85

Extent of the glacial aquifer system (approximate)

Figure B9. Estimated annual natural ground-water recharge in the glacial aquifer system in the northern United States (modified from Wolock, 2003b). No data were available for Alaska.

Prepared by:
Rolla Publishing Service Center

For more information concerning this publication, contact:
Director, USGS, Illinois Water Science Center
1201 West University Avenue, Suite 100
Urbana, IL 61801–2347
(217) 344–0037

Or visit the Illinois Water Science Center Web site at:
http://il.water.usgs.gov

Science Center Director
USGS Illinois Water Science Center
1201 W. University Ave., Suite 100
Urbana, IL 61801

http://il.water.usgs.gov

Printed on recycled paper

Arnold and others —Hydrochemical Regions of the Glacial Aquifer System, Northern United States—Scientific Investigations Report 2008–5015