

National Water-Quality Assessment Program

Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells in the Glacial Aquifer System in the United States



Scientific Investigations Report 2010–5100

Front Cover. Most private wells are located in rural areas with surrounding agricultural land use. The photograph was taken by Kelly Warner and Troy Daugherty on June 28, 2010, near Eureka, IL.

Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells in the Glacial Aquifer System in the United States

By Kelly L. Warner and Terri L. Arnold

National Water-Quality Assessment Program

Scientific Investigations Report 2010–5100

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

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2010

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

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

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:

Warner, K.L., and Arnold, T.L., 2010, Relations that affect the probability and prediction of nitrate concentration in private wells in the glacial aquifer system in the United States: U.S. Geological Survey Scientific Investigations Report 2010–5100, 55 p. plus appendixes.

ISBN 978–1–4113–2951–5

Foreword

The U.S. Geological Survey (USGS) is committed to providing the Nation with reliable 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, 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 quality of our Nation's streams and groundwater? How are conditions changing over time? How do natural features and human activities affect the quality of streams and groundwater, 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 to 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/studies/study_units.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 groundwater, and by determining water-quality 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, 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
Associate Director for Water

Contents

Foreword	iii
Abstract	1
1.0 Introduction and Hydrostratigraphy of the Glacial Aquifer System	2
2.0 Methods	10
3.0 Background Nitrate Concentration	14
4.0 Comparison of Nitrate Concentration in the Glacial Aquifer System and Other Principal Aquifers	16
5.0 Nitrate Sources	20
6.0 Nitrate Transport	22
6.1 Nitrate and Groundwater Age	22
6.2 Nitrate Concentrations and Depth	26
7.0 Nitrate Concentrations in Private, Public, and Monitoring Wells	28
7.1 Nitrate Concentrations in Water from Private Supplies	28
7.2 Comparison of Nitrate Concentrations in Water from Private, Public, and Monitoring Wells	32
8.0 Statistical Models of Nitrate Concentration	36
9.0 Probability of Nitrate Occurrence Above Threshold Concentrations	40
10.0 Prediction of Mean Nitrate Concentrations in Networks of Wells	44
11.0 Conclusions	48
References Cited	51
Glossary	55

Appendixes

1. Variable names, descriptions, units, and sources used in the models described in this report	57
2. Results of a comparison of nitrate measured in private and monitoring wells in the glacial aquifer system to the nationwide model of probability of nitrate concentration greater than 4 milligrams per liter as nitrogen based on all monitoring wells (Nolan and others, 2002). (Data are from a spatial overlay of wells in the glacial aquifer system with a model of probability of nitrate contamination of groundwater (Nolan and others, 2002))	58
3. Results of using a logistic regression model with six categories to calculate the probability of nitrate concentration greater than 4 milligrams per liter as nitrogen	59
4. Results of recalibrating a logistic regression model (Nolan and others, 2002): (A) Coefficients resulting from recalibrating the regression equation from Nolan and others (2002) and results of test to determine whether estimated coefficients of new variables are significant predictors; and (B) observed and predicted probabilities using the 6 categories of predicted probabilities in Nolan and others (2002) and the standard 10 categories	60–61
5. Logistic regression parameter estimates and odds ratio estimates for each of 10 logistic regression models	62–65

6.	Scatterplots showing predicted and observed probabilities of nitrate concentration greater than the specified concentration in private wells in the glacial aquifer system for all 10 logistic regression models developed for this study.....	66
7.	Spearman rank correlations of nitrate concentration with selected variables, probability significance (p), and number of samples.....	67–70
8.	Variables selected in each of three iterations of the development of the 10 logistic regression models developed for this study	71–72
9.	Linear regression to predict network mean nitrate concentration using the same variables used in the logistic regression models. The variables were used because they were used in the logistic regression models. The significance level of variables was not a consideration for this model.....	73

Figures

1–1.	Agriculture is the major land use overlying the glacial aquifer system.....	2
1–2.	The highest rate (1992–2001) of nitrogen fertilizer application is over the Midwestern states which are part of the glacial aquifer system	3
1–3.	Despite the gaps in the spatial distribution of wells across the glacial aquifer system in the United States, all four glacial framework areas are represented in the wells sampled as part of the U.S. Geological Survey National Water-Quality Assessment program. Framework areas are based on similarity of physiography and source of aquifer material. The wells in Nebraska are included in this study because they are located in areas of reworked glacial deposits.	5
1–4.	Wells in buried bedrock valleys are less susceptible to nitrate contamination because of the thick overlying layer of till which is mostly fine-grained material, whereas wells in valley-fill deposits and in sheets, layers, and lenses of sand are more likely to have a connection to the land surface.....	6
1–5.	Fine-grained sediment is present at land surface across most of the glacial aquifer system, but the texture of sediment at land surface is not indicative of the presence or absence of an underlying aquifer.....	8
1–6.	There are numerous soil types overlying the glacial aquifer system. Alfisols are soils that overlie approximately 25 percent of the glacial aquifer system, mostly in the Midwest. Alfisols are typically formed in humid areas under deciduous forests and mixed vegetation.....	9
3–1.	The graph shows the deviation of the mean and median nitrate concentrations from background in groundwater samples from reference wells selected to represent areas minimally affected by human activities. Mean and median nitrate concentrations in water from reference wells in Nebraska were greater than the background nitrate concentration used in this report (1 milligram per liter as nitrogen); however, mean and median nitrate concentrations in water from reference wells in other parts of the glacial aquifer system generally were less than the background concentration.....	15
4–1.	Locations of U.S. Geological Survey National Water-Quality Assessment (NAWQA) program regional assessments in principal aquifers in the United States. The glacial aquifer system covers about 953,000 square miles and parts of 26 States. The two highest concentrations of nitrate in private wells measured during the NAWQA program (1995–2005) were in water from the glacial aquifer system	16

4-2.	Although the glacial aquifer system had a larger range of nitrate plus nitrite as nitrogen measured in water from private and public-supply wells in selected principal aquifers and other aquifers of different lithologies sampled as part of the U.S. Geological Survey National Water-Quality Assessment (NAWQA) Program (1995–2005) are shown.....	17
4-3.	The glacial aquifer system receives a higher nitrogen load than most other aquifers, but the capacity for nitrogen attenuation also is greater	19
5-1.	Median nitrate concentration was lowest in water from wells in areas where agriculture is the predominant land use within a 1,640 foot (500 meter) radius of the well, but nitrate concentrations in these areas also were the most variable. Median nitrate concentration was highest in areas where urban land use is predominant within a 1,640 foot radius of the well	21
6.1-1.	Groundwater flow to large- and small-diameter wells may represent recharge that was a mixture of groundwater of different ages (Sandra Eberts, U.S. Geological Survey, written commun., 2007). The large-diameter wells generally are shallow and contain “young” (recharged after 1955) water.....	22
6.1-2.	All samples in which nitrate concentrations were greater than 4 milligrams per liter as nitrogen were from private wells containing only groundwater that was recharged after 1955 (“young”).....	24
6.1-3.	Groundwater recharged after 1955 (“young”) consisted of a higher percentage of oxic water than groundwater recharged before 1955 (“old”). Nitrate is mobile under oxic conditions.....	25
6.1-4.	(A) The glacial aquifer system in the United States has been divided into four framework areas on the basis of physiography and source of aquifer material. (B) Private wells in these areas contained a mixture of old and young groundwater. The wells in the West area contained mostly “young” (recently recharged) groundwater.....	25
6.2-1.	The difference in mean and median well depth shows the large variability of well depths associated with nitrate concentrations less than 6 milligrams per liter. The samples with concentrations greater than 6 milligrams per liter were from relatively shallow wells	27
6.2-2.	Mean and median nitrate concentrations were highest in water from private wells less than 20 feet deep, but most private wells were more than 60 feet deep.....	27
7.1-1.	Agriculture was the predominant land use within 1,640 feet of the private wells sampled	28
7.1-2.	Nitrate concentrations in most of the private wells sampled in this study were less than the background nitrate concentration of 1 milligram per liter as nitrogen (<1 mg/L as N). The background nitrate concentration of <1 mg/L as N is less than one-tenth of the U.S Environmental Protection Agency Maximum Contaminant Level (MCL) of 10 mg/L as N; such concentrations are not likely a human-health concern. In this study, samples with nitrate concentrations between 1 and 10 mg/L as N are considered to approach concentrations of potential human-health concern, and nitrate concentrations greater than the MCL of 10 mg/L as N are of potential human-health concern.....	29
7.1-3.	The Central framework area of the glacial aquifer system had a high nitrate application rate but a low median concentration of nitrate in groundwater from private wells. The Central and West-central areas had a large variability in nitrate concentrations.....	29

7.1-4.	Most private wells sampled in this study were constructed during 1990-99 and the water from these wells sampled by NAWQA had the highest mean nitrate concentration compared to water from other private wells.....	30
7.2-1.	Most private and monitoring wells sampled were located in agricultural settings, whereas most public-supply wells sampled were in forested or mixed-land-use areas. The public-supply wells sampled by NAWQA were mostly in the northeastern United States, where forest and mixed land use are common	32
7.2-2.	Private and monitoring wells in the glacial aquifer system sampled as part of the U.S. Geological Survey National Water-Quality Assessment program had similar geographic distribution. More public-supply wells were sampled in the East framework area than in other framework areas.....	34
7.2-3.	Survival curves were plotted to examine the difference in measured nitrate concentrations among private, public-supply, and monitoring wells in the glacial aquifer system. Nitrate concentrations exceeded 4 milligrams per liter as nitrogen in about 12 percent of private wells, 0 percent of public-supply wells, and 36 percent of monitoring wells	35
8-1.	Scatterplots showing the accuracy of the nationwide nitrate model (Nolan and others, 2002) for predicting the probability of nitrate concentration greater than 4 milligrams per liter as nitrogen in private wells in the glacial aquifer system	38
9-1.	The interaction between open-interval diameter and depth to top of open interval indicates that the relation between nitrate concentration and open-interval diameter in private wells in the glacial aquifer system changes at depths of about 71 feet below land surface	41
9-2.	Fewer variables are needed to predict higher nitrate concentrations (7-10 milligrams per liter as nitrogen) than lower nitrate concentrations (1-6 milligrams per liter as nitrogen); at lower concentrations, most variables represent natural conditions and well-construction characteristics	42
9-3.	Probability of occurrence of nitrate concentrations greater than 1 milligram per liter in private wells in the glacial aquifer system. This map was developed by setting the depth to top of open interval and open-interval length to the mean of the values for these variables in all private wells in the glacial aquifer system (91 and 8 feet, respectively)	43
10-1.	The relation between mean observed nitrate concentrations and mean nitrate concentrations predicted with the linear regression model indicates that the model is a good fit ($R^2=0.82$)	45
10-2.	Mean observed and predicted nitrate concentrations show the goodness of fit for the linear regression model, which is based on data from these 17 well networks. The networks for which the residuals are closer to zero have the best fit to the model. Data needed for the model were unavailable for Alaska.....	46
10-3.	The area of highest predicted mean nitrate concentration in water from private wells in the glacial aquifer system is in the glacial deposits in west-central Illinois.	47
11-1.	There was no distinct geographic pattern in nitrate concentration in water from private wells in the glacial aquifer system, although most high concentrations (greater than 10 milligrams per liter as nitrogen) were in the central part of the glaciated United States.....	49
11-2.	Changes in nitrate concentrations in water from private wells in the highly heterogeneous glacial aquifer system were most variable with depth.....	50

Tables

1-1.	The glacial aquifer system has many local and regional names, as shown by the nomenclature for Quaternary aquifers of glacial origin used in the U.S. Geological Survey's Regional Aquifer System Assessment (RASA) program. In this report, the term "glacial aquifer system" is used to represent all of these aquifers as one system	6
4-1.	Mean nitrate concentrations are higher in sand and gravel aquifers than most other lithologies	18
5-1.	Comparison of nitrate concentration in groundwater from private wells to nitrogen sources indicates that atmospheric deposition of nitrogen, septic-system density, and sewer density are significantly related to nitrate concentration with 95 percent confidence, but the relations are weak.....	20
6.1-1.	Most detections of nitrate in private wells were in "young" (recently recharged) groundwater that is oxic. In this study, only samples of young, oxic groundwater contained nitrate at concentrations greater than 10 milligrams per liter as nitrogen.....	23
6.2-1.	Shallow wells (75 feet deep or less) generally were associated with higher nitrate concentration, younger groundwater, and more oxic conditions than wells deeper than 75 feet.....	26
7.2-1.	Private, public-supply, and monitoring wells in the glacial aquifer system differ with respect to groundwater nitrate concentration and well-construction characteristics.....	33
7.2-2.	Samples from all well types in the glacial aquifer system represent a mixture of groundwater of different ages.....	35
8-1.	Variables used in selected models to predict probability of nitrate concentration exceeding a given level are compared. The logistic regression models included some or all of the variables in the linear regression model highlighted in this report, plus an interaction term for models with nitrate thresholds of 5 mg/L or greater.....	37
9-1.	Variables used in logistic regression models for 10 thresholds of nitrate concentrations include source, transport, and attenuation variables. The intercept is a statistical variable that was important in all models for thresholds of 1 to 10 mg/L.....	40
10-1.	Coefficient values, standard error and p-values of variables used in the linear regression model used to predict mean nitrate concentration in private-well networks in the glacial aquifer system	45

Conversion Factors

SI to Inch/Pound

Multiply	By	To obtain
Length		
centimeter (cm)	0.3937	inch (in.)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
Area		
square kilometer (km ²)	0.3861	square mile (mi ²)
Volume		
liter (L)	0.2642	gallon (gal)
Mass		
kilogram (kg)	2.205	pound avoirdupois (lb)
cubic meter per second	22.82	million gallons per day (mgd)
megaliters per day	0.2642	million gallons per day (mgd)
Application rate		
kilograms per hectare per year [(kg/ha)/yr]	0.8921	pounds per acre per year [(lb/acre)/yr]

Abbreviated water-quality units:

µg/L micrograms per liter

mg/L milligrams per liter

Abbreviations

CFC	Chlorofluorocarbon
DO	Dissolved oxygen
MCL	Maximum Contaminant Level
NAWQA	National Water-Quality Assessment Program
RASA	Regional Aquifer System Assessment
Redox	Oxidation and reduction potential
ROC	Receiver Operating Characteristic
STATSGO	U.S. Department of Agriculture State Soil Geographic database
TU	Tritium Units
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey.
>	Greater than
<	Less than
%	Percent

Terms in text that are shown in **boldface** are defined in the glossary at the back of the report.

Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells in the Glacial Aquifer System in the United States

By Kelly L. Warner and Terri L. Arnold

Abstract

Nitrate in private wells in the glacial aquifer system is a concern for an estimated 17 million people using private wells because of the proximity of many private wells to nitrogen sources. Yet, less than 5 percent of private wells sampled in this study contained nitrate in concentrations that exceeded the U.S. Environmental Protection Agency (USEPA) Maximum Contaminant Level (MCL) of 10 mg/L (milligrams per liter) as N (nitrogen). However, this small group with nitrate concentrations above the USEPA MCL includes some of the highest nitrate concentrations detected in groundwater from private wells (77 mg/L). Median nitrate concentration measured in groundwater from private wells in the glacial aquifer system (0.11 mg/L as N) is lower than that in water from other unconsolidated aquifers and is not strongly related to surface sources of nitrate. Background concentration of nitrate is less than 1 mg/L as N.

Although overall nitrate concentration in private wells was low relative to the MCL, concentrations were highly variable over short distances and at various depths below land surface. Groundwater from wells in the glacial aquifer system at all depths was a mixture of old and young water. Oxidation and reduction potential changes with depth and groundwater age were important influences on nitrate concentrations in private wells.

A series of 10 logistic regression models was developed to estimate the probability of nitrate concentration above various thresholds. The threshold concentration (1 to 10 mg/L) affected the number of variables in the model. Fewer explanatory variables are needed to predict nitrate at higher threshold concentrations. The variables that were identified as significant predictors for nitrate concentration above 4 mg/L as N included well characteristics such as open-interval diameter, open-interval length, and depth to top of open interval. Environmental variables in the models were mean percent silt in soil, soil type, and mean depth to saturated soil. The 10-year mean (1992–2001) application rate of nitrogen fertilizer applied to farms was included as the potential source variable. A linear regression model also was developed to predict mean nitrate concentrations in well **networks**. The model is based on network averages because nitrate concentrations are highly variable over short distances. Using values for each of the predictor variables averaged by network (network mean value) from the logistic regression models, the linear regression model developed in this study predicted the mean nitrate concentration in well networks with a 95 percent confidence in predictions.

1.0 Introduction and Hydrostratigraphy of the Glacial Aquifer System

The glacial aquifer system is characterized by the heterogeneity of the aquifer material. A framework for the system groups areas of similar intrinsic susceptibility and natural vulnerability.

The glacial aquifer system is the principal aquifer for most of the northern United States and is present in parts of 26 States. A principal aquifer is a regionally extensive aquifer system that is used and has the potential to be used as a source of potable water (Lapham and others, 2005). The area overlying the glacial aquifer system receives some of the largest surface applications of nitrate as a fertilizer in the Nation, and the population using private water supplies from the glacial aquifer system is large. Approximately 41 million people relied on the glacial aquifer system for public supply and private use in 2000 (Warner and Arnold, 2006). The population served by private supply from the glacial aquifer system was estimated at 17 million people in 2000. Total private withdrawals were estimated at 1,200 million gallons per day, which is about 40 percent of the 2000 groundwater withdrawals for private supply nationwide.

The glacial aquifer system is defined as all unconsolidated aquifers above bedrock north of the line of continental glaciation in the United States. This aquifer system is an important resource for the highly agricultural part of the United States including Illinois, Indiana, Iowa, Nebraska, and elsewhere because it is the uppermost aquifer. The aquifer extends from land surface to depths of more than 1,000 ft in some areas, so it is easily accessible for private well installation.

Nitrogen is a concern for private well owners because it is often applied as a fertilizer in areas near private wells. Nitrogen in groundwater is in the forms of nitrite and nitrate. Measurements of nitrite plus nitrate as nitrogen (N) in this report are hereafter referred to as “nitrate.” Nitrate is a concern for private well owners if the water is consumed because nitrate can restrict oxygen transport in the bloodstream, which is especially harmful to infants and the elderly. The U.S. Environmental Protection Agency (USEPA) has

established a drinking-water standard of 10 mg/L as N for nitrate, which is the Maximum Contaminant Level (MCL) (U.S. Environmental Protection Agency, 2009). Nitrate can originate from natural sources or from human activities such as fertilizer application and wastewater disposal. The concentration and transformation of nitrate are controlled by oxidation and reduction potential (redox) conditions and mediated by microorganisms. Nitrate is stable and mobile under oxidizing conditions. Additionally, it is soluble in water and can easily pass through soil to the groundwater table, where it can persist for decades. The amount of nitrate may increase because of continuing yearly inputs (Nolan and Stoner, 2000).

In a study of nitrate in private wells in the United States, it was reported to be the only contaminant derived from primarily manmade sources that was found at concentrations greater than a human-health benchmark in more than 1 percent of the wells (DeSimone and others, 2009). Concentrations of nitrate were greater than the MCL in nearly 25 percent of the private wells in areas of intense agricultural land use (DeSimone and others, 2009). The dominant land use overlying the glacial aquifer system is agriculture (fig. 1–1). Nitrate is applied at high rates to land (fig. 1–2) overlying the glacial aquifer system and its occurrence is a known human-health issue in groundwater (Follett and Hatfield, 2001). Private wells are commonly located in rural parts of the country where large areas are fertilized. Nitrate is a concern in many of these rural areas because the residents rely on private wells in the glacial aquifer system for drinking water and these wells are not routinely monitored for nitrate. States typically recommend that water from new wells or wells on transferred property be analyzed for nitrate and bacteria, but this analysis is at the discretion of the well owner.

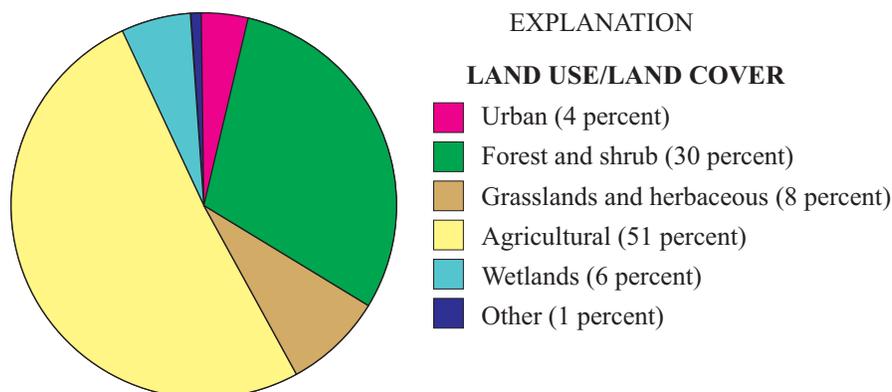
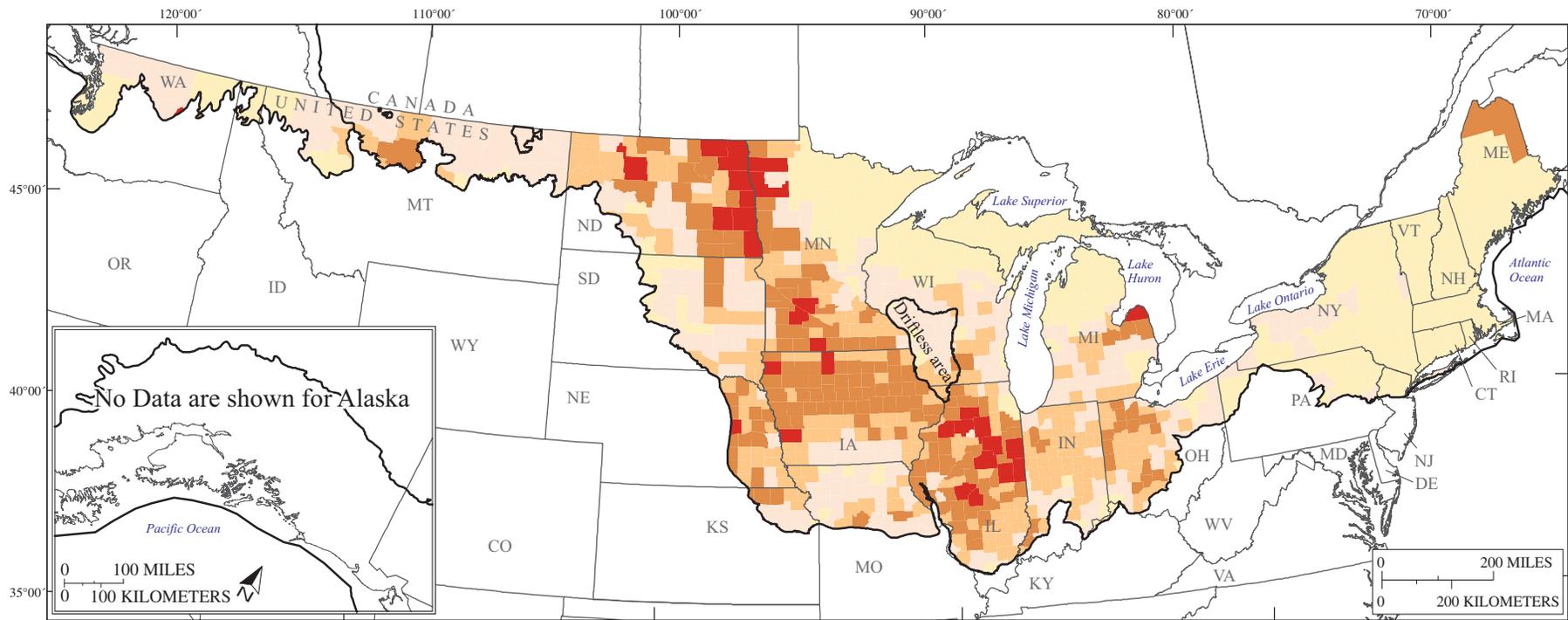


Figure 1–1. Agriculture is the major land use overlying the glacial aquifer system.



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

**Mean nitrogen fertilizer
 applied to farms (1992-2001),
 in million kilograms per year**

- 0 to 1.75
- Greater than 1.75 to 5.02
- Greater than 5.02 to 8.77
- Greater than 8.77 to 14.60
- Greater than 14.60 to 36.00

— Glacial aquifer system extent
 (approximate)

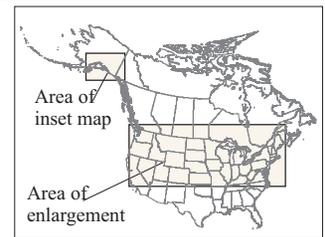


Figure 1-2. The highest rate (1992–2001) of nitrogen fertilizer application is over the Midwestern states which are part of the glacial aquifer system. (Modified from Ruddy and others, 2006.)

4 Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells, Glacial Aquifer System

The purpose of this study was to determine the vulnerability of private supplies in the glacial aquifer system to nitrate contamination. A well is susceptible if water can flow easily from the land surface to the well, but the well becomes vulnerable when there is a known source of nitrate that can be transported to the susceptible well. This report determines the vulnerability of private wells in three ways:

1. Determine the relation of nitrate in water from private wells to potential sources;
2. Determine the relation of nitrate in water from private wells to environmental variables that may affect the transport of nitrate; and
3. Determine which variables are related to the probability and prediction of nitrate contamination in private wells in the glacial aquifer system.

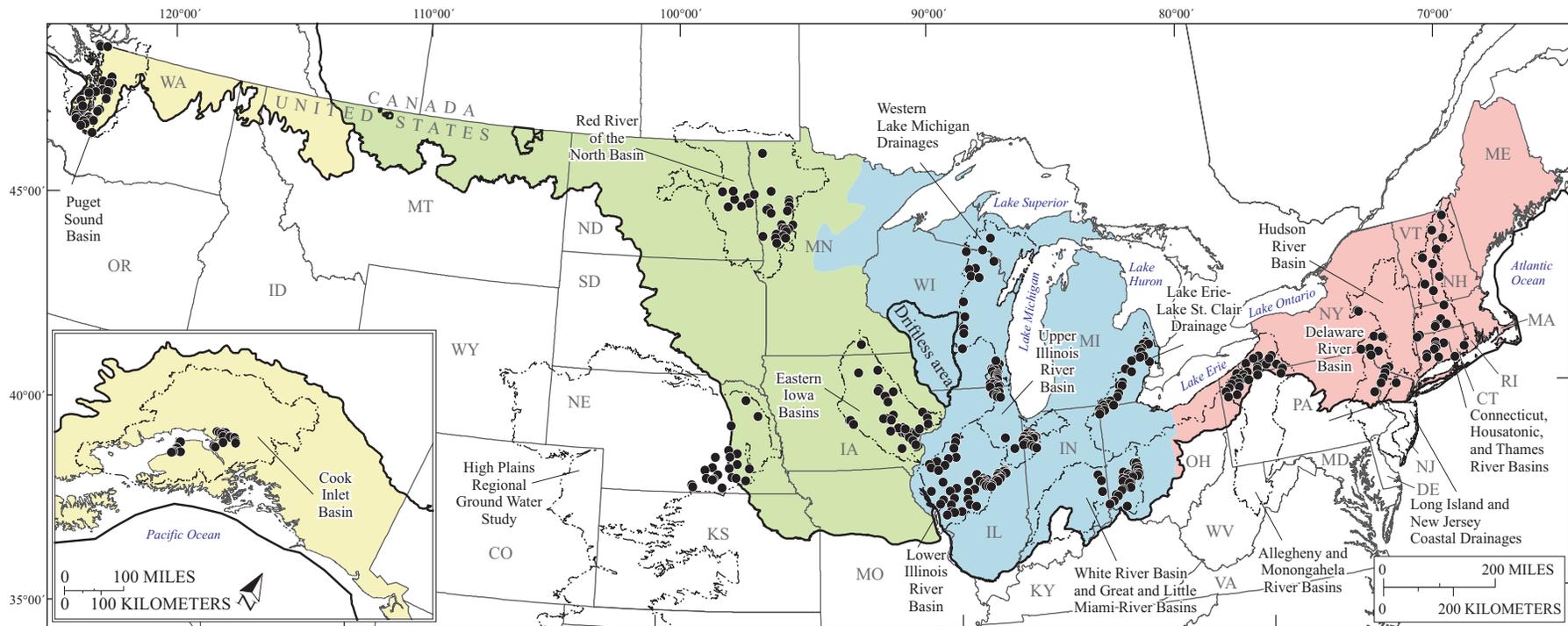
This report is based on the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) program's groundwater sampling of the glacial aquifer system from 1991 to 2005. The data set includes data collected and analyzed in a consistent manner as defined by the NAWQA program (Koterba and others, 1995). The data analyzed in this study are combined from 15 geographically separated study areas (NAWQA study units). Only the wells originally selected to address the overall water quality of the glacial aquifer system—379 private wells—were used in this study (fig. 1–3).

Extensive quality-assurance data have been collected to measure the accuracy of the data (Mueller and Titus, 2005), and concentrations of nitrite plus nitrate were determined to be essentially unaffected by contamination. Laboratory nitrite concentration commonly is less than the reporting level of 0.01 mg/L as N in groundwater sampled as part of the NAWQA program, resulting in negligible contribution to the nitrite plus nitrate concentration (Nolan and Stoner, 2000).

The nomenclature for the glacial aquifer system is varied. Other names for the glacial aquifer system have included stratified drift, surficial aquifer, unconsolidated deposits, and sand and gravel aquifers, as well as many local aquifer names. The term “stratified-drift aquifer” was formerly used in the eastern States where proglacial lakes or marine embayments developed in major valleys or lowlands during deglaciation and were filled with a sequence of stratified sediment (Randall, 2001; Kontis and others, 2004). The exceptions are Cape Cod and Long Island, which are covered in sheet-type sands. The term “surficial” or “glacial” aquifer has been used in the mid-continent area for aquifers in sands and tills (Bugliosi, 1999;

Eberts and George, 2000). In this part of the glacial aquifer system, buried bedrock valleys are filled with sediment in which the water-producing units are sand lenses within clayey deposits or coarse-grained deposits at the bottom of buried bedrock valleys. In the western glaciated area, the unconsolidated deposits of glacial origin are part of the Puget Sound aquifer system (Jones, 1999). The USGS Regional Aquifer System Assessment (RASA) program delineated hydrogeologic units in the glacial aquifer system in different geographic areas (table 1–1). Additionally, the USGS Ground Water Atlas of the United States refers to all of these aquifers as the surficial aquifer system of sand and gravel aquifers (Miller and Appel, 1997; Trapp and Horn, 1997; Lloyd and Lyke, 1995) or unconsolidated-deposit aquifers (Whitehead, 1996). All of these aquifers contain groundwater in the unconsolidated material that ranges from land surface to bedrock (including materials of glacial and alluvial origin). This aquifer system is referred to as the **glacial aquifer system** in this report.

The three general environmental settings for private wells in the glacial aquifer system are buried bedrock valleys; sheets, layers, or discontinuous lenses of sand; and valley-fill deposits (fig. 1–4). Buried bedrock valleys are known in some areas and the glacial aquifer system in this setting is often sought for public and private supplies. Aquifers in buried bedrock valleys commonly provide large water supplies for cities and private homeowners. Because wells that withdraw water from sand and gravel in a buried bedrock valley typically are very deep, it is not likely they will contain water with measurable nitrate concentrations. The valley-fill environmental setting is similar to the buried bedrock valley except there is frequently hydraulic connection to land surface and surface-water bodies (fig. 1–4). Groundwater from the glacial aquifer system in this setting is more likely to contain measurable nitrate because of the hydraulic connection of sand and gravel to the land surface. If wells are drilled deep in a valley-fill setting, then nitrate reduction is likely. The environmental setting of sheets, layers, or lenses of discontinuous sand is the most difficult to assess for vulnerability to nitrate. If wells are drilled in shallow sands and gravels, nitrate applied at land surface may enter the groundwater. If wells are drilled deeper, below some clays or clayey tills, then nitrate is less likely to be in the groundwater because water does not travel easily through clay. Understanding the geology and environmental setting of a private well helps in assessing the vulnerability of the aquifer to contamination from nitrate applied at the land surface.



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

- | | |
|-------------------------------|---|
| Glacial framework area | — Glacial aquifer system extent (approximate) |
| ■ East | --- Study-unit boundary |
| ■ Central | ● Private well |
| ■ West-central | |
| ■ West | |

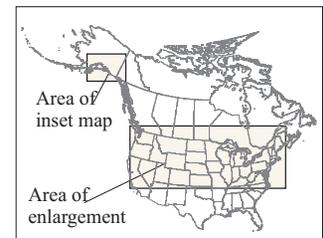


Figure 1-3. Despite the gaps in the spatial distribution of wells across the glacial aquifer system in the United States, all four glacial framework areas are represented in the wells sampled as part of the U.S. Geological Survey National Water-Quality Assessment program. Framework areas are based on similarity of physiography and source of aquifer material. The wells in Nebraska are included in this study because they are located in areas of reworked glacial deposits.

6 Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells, Glacial Aquifer System

Table 1–1. The glacial aquifer system has many local and regional names, as shown by the nomenclature for Quaternary aquifers of glacial origin used in the U.S. Geological Survey’s Regional Aquifer System Assessment (RASA) program. In this report, the term “glacial aquifer system” is used to represent all of these aquifers as one system.

Glacial framework area	Reference	Regional Aquifer System Assessment (RASA) study area	Name of Quaternary aquifers of glacial origin
East	Randall, 2001	Stratified-drift aquifers	Stratified-drift aquifers.
Central	Bugliosi, 1999	Midwestern Basins and Arches	Glacial deposits and confining unit.
Central and West-Central	Young, 1992	Cambrian-Ordovician aquifer	Drift aquifer.
West	Jones, 1999	Puget Sound Aquifer System	Various river aquifers, Fraser aquifer, or Puget aquifer.

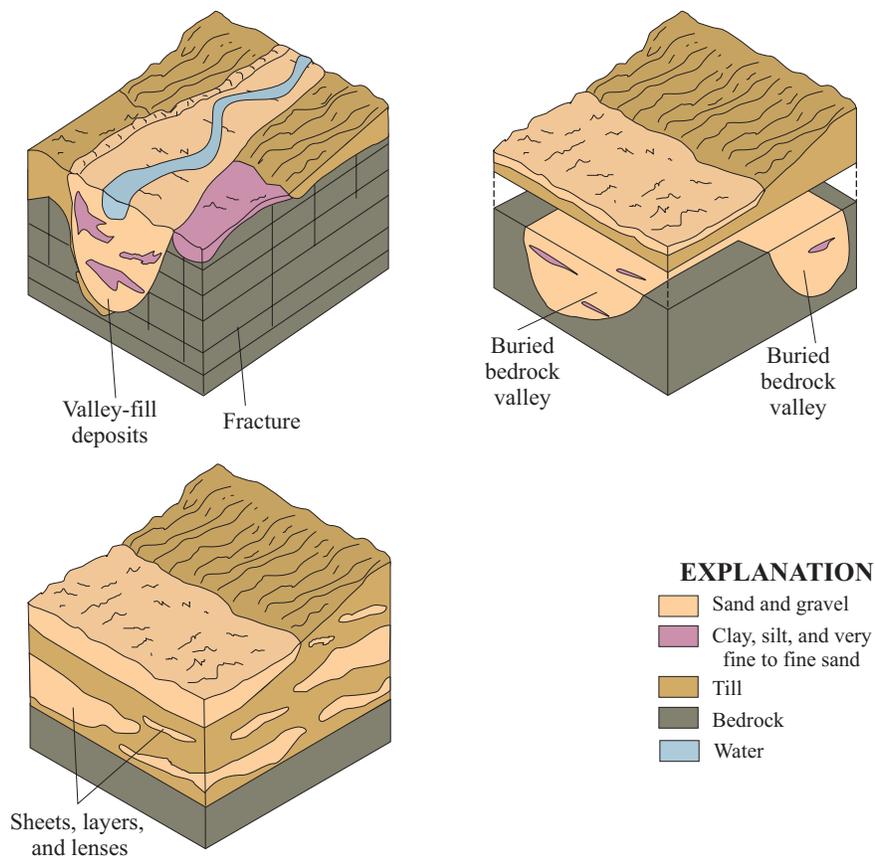


Figure 1–4. Wells in buried bedrock valleys are less susceptible to nitrate contamination because of the thick overlying layer of till which is mostly fine-grained material, whereas wells in valley-fill deposits and in sheets, layers, and lenses of sand are more likely to have a connection to the land surface. (Modified from Whitehead, 1996.)

Heterogeneity characterizes the geology of the glacial aquifer system. Because of its clay content, especially in the Midwest, the glacial aquifer system differs from other unconsolidated aquifer systems, such as the High Plains of Colorado and Nebraska. The distribution of the sand and gravel units that make up the glacial aquifer system and the clay and silt confining units that are interbedded can be extremely complex (Miller, 1999). Because the lobes of continental ice associated with these glacial deposits advanced repeatedly and from different directions, different materials were eroded, transported, and deposited, depending on the predominant rock types underlying the ice. The mixtures of clay and sand are widely variable both spatially and vertically, and this variability affects the transport of nitrate through the aquifer system. Randall (2001) notes that the susceptibility of a given aquifer to contamination, if determined from aquifer material near the surface (fig. 1–5), can be misleading because the distribution of aquifer material with depth does not match the surficial distribution. For example, some of the coarse-grained deposits that make up the productive part of the glacial aquifer system may be buried beneath fine-grained deposits and generally are not depicted on maps of surficial geology.

Another key difference between the glacial aquifer system and other unconsolidated aquifers is the mixture of clay and organic material it contains. The clay and organic-matter content is an important aquifer characteristic because these materials can slow down the transport of contaminants to deeper groundwater. Subsurface drains, which are installed in agricultural areas with clayey soils such as **alfisols**, can impede nitrate in recharge water from reaching the water table. McMahon and others (2008) found that the extent of nitrate attenuation is related to the depositional environment of the

aquifer materials and generally increases from fluvial sand to glacial sand to glacial/marine clay, and, finally, to marine shale. Susceptibility of wells in the glacial aquifer system to contamination also is affected by organic material, which commonly originates as peat and soils buried beneath glacial deposits and in forested areas. Organic material can affect the redox of the groundwater system by promoting consumption of dissolved oxygen and denitrification. Alfisols, naturally productive soils typically formed under deciduous forest or mixed vegetative cover in humid areas, overlie 25 percent of the glacial aquifer system—mostly in the Midwestern United States (fig. 1–6). This soil type results from weathering processes that leach clay and minerals from the surface layer into lower soil layers. Alfisols are characterized by discolored features that indicate particular redox conditions, such as recurring periods of soil saturation (U.S. Department of Agriculture, 2006).

Grouping areas of similar hydrogeology may help to assess the regional water quality of the glacial aquifer system. The glacial aquifer system was divided into four areas by Warner and Arnold (2006) based on factors affecting the intrinsic susceptibility (related to the ease of groundwater flow) and the natural vulnerability of wells. These four areas of the glacial framework are used in this report to delineate areas of similar susceptibility, but local information on geochemical conditions and land use would be needed to refine the determination of vulnerability to nitrate contamination. Bedrock is not considered as part of this framework because nitrate is not likely to be derived from a bedrock source. The 379 private wells sampled as part of this study include 71 wells in the East **framework area**, 174 in the Central framework area, 83 in the West-central framework area, and 51 in the West framework area.

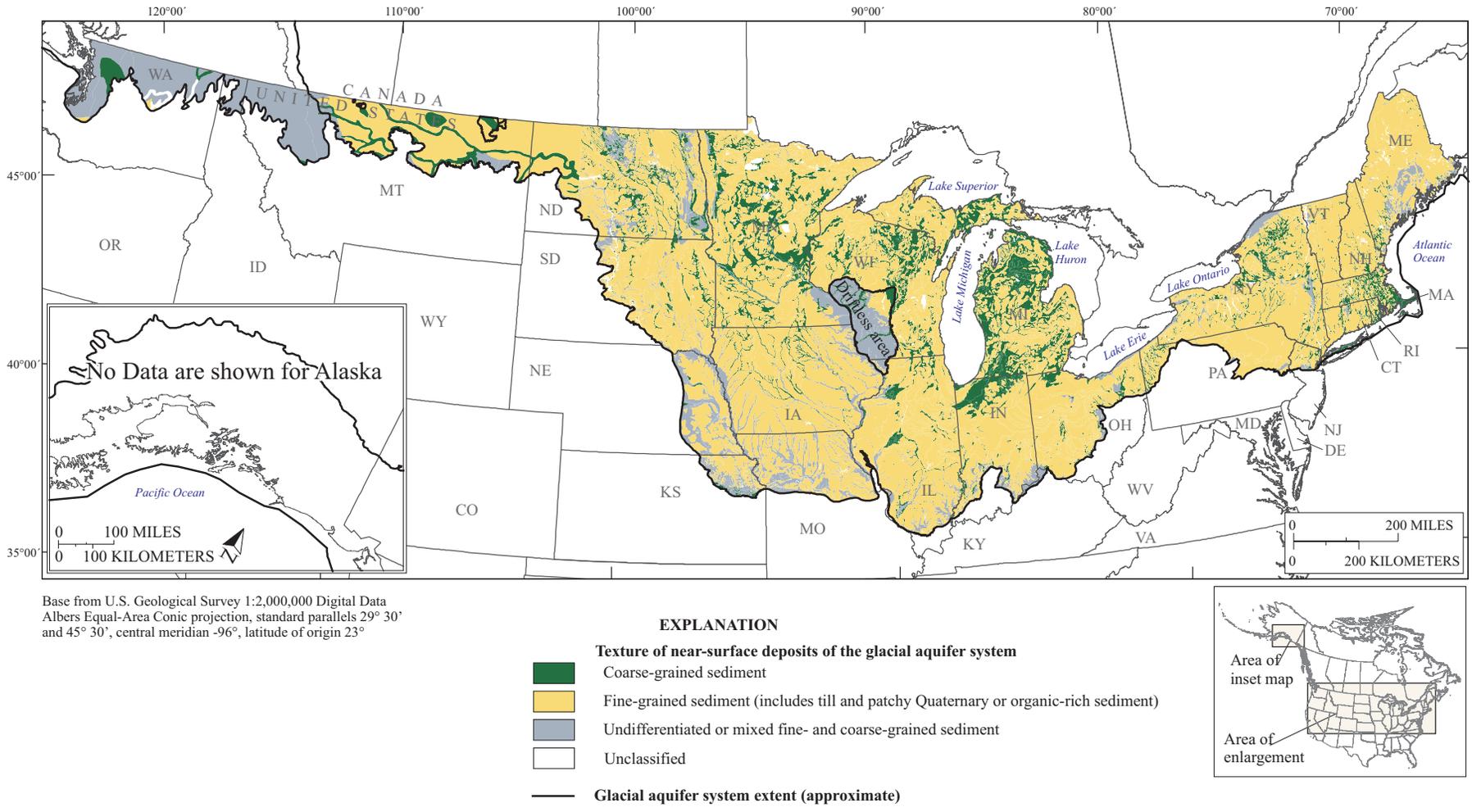
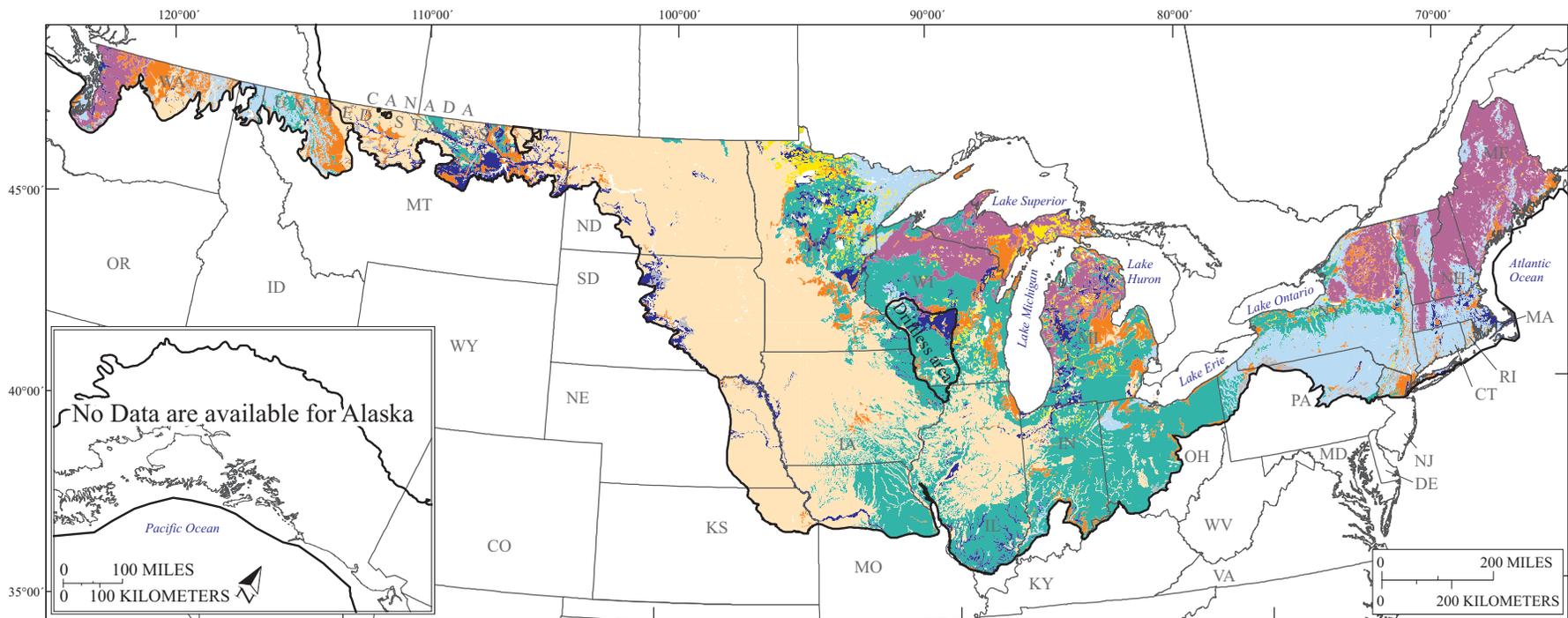


Figure 1-5. Fine-grained sediment is present at land surface across most of the glacial aquifer system, but the texture of sediment at land surface is not indicative of the presence or absence of an underlying aquifer. (From Arnold and others, 2008.)



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

Predominant soil type (greater than 50% of soil)

	Alfisols		Mollisols
	Entisols		Spodosols
	Histosols		Other (andisols, aridisols, ultisols, or vertisols)
	Inceptisols		Mixed—none predominant

— Glacial aquifer system extent (approximate)

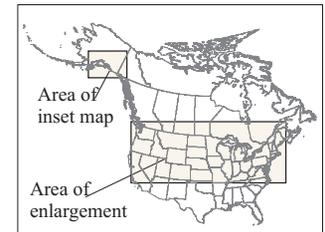


Figure 1-6. There are numerous soil types overlying the glacial aquifer system. Alfisols are soils that overlie approximately 25 percent of the glacial aquifer system, mostly in the Midwest. Alfisols are typically formed in humid areas under deciduous forests and mixed vegetation. (Modified from Wolock, 1997.)

2.0 Methods

Standardized sample collection and analysis aids in regional aggregation of data. Nonparametric statistics and models are used to interpret the analytical results.

The USGS NAWQA program is the primary source of long-term, nationwide water-quality data from groundwater and surface water (<http://water.usgs.gov/nawqa/>). Groundwater data collected as part of the NAWQA program are based on networks of 20 to 30 wells. Nitrate data were collected at 1,290 groundwater-sampling sites in 15 studies across the glacial aquifer system during 1991–2005. Within these 15 studies, there are 17 well networks that were designed to collect data targeted for a specific purpose. Most of the 430 private wells in the glacial aquifer system that were sampled in the 15 studies were selected to assess the overall water quality of the major aquifer system, but 51 of the private wells were initially selected to address the impact of land use on water quality or for other, special studies (Lapham and others, 1995). Only the private or public supply wells in networks selected to assess the overall water quality of the major aquifer system are included in this report. This study analyzed data from 379 private wells, 41 public-supply wells, and 870 shallow monitoring wells. Public supplies sampled as part of the NAWQA assessment of major aquifers were selected if the wells were pumped at a low capacity (Lapham and others, 1995). All of the private and public wells sampled for this study were wells installed by the owners; all of the monitoring wells were installed by the USGS specifically to examine the effects of land use on shallow, recently recharged groundwater (generally less than 10 years old) (Gilliom and others, 1995). Additionally, 20 reference wells that are a subset of the monitoring wells were selected in areas identified as minimally impacted by humans. Data have been collected at these wells since 1994 and include 78 analyses (completed by January 2007). These reference-well data were used to determine the background nitrate concentration.

Data were collected following NAWQA protocols (Lapham and others, 1995; Koterba and others, 1995; Koterba, 1998; and Squillace and Price, 1996). Analyses for groundwater-age indicators were completed at several laboratories. Tritium was measured at the USGS Menlo Park Isotope Research Laboratory in Menlo Park, CA, or at the University of Miami Tritium Laboratory in Miami, FL. Chlorofluorocarbon (CFC) and tritium-helium were measured at the USGS Chlorofluorocarbon Laboratory in Reston, VA, or at the Lamont-Doherty Laboratory in Palisades, NY. Nitrogen compounds were measured at the USGS National Water-Quality Laboratory using methods documented in Fishman (1993) and Patton and Kryskalla (2003).

Some nitrate data used in this analysis were censored (the concentration was less than the reporting limit) at 0.05 mg/L as N. For most of this analysis, nitrate concentrations that were censored were treated as if they had been measured at

0.05 mg/L as N. These treated values were used in the calculations of general descriptive statistics. Spearman correlation coefficients were used to examine relations among variables related to nitrate concentrations in private wells and in monitoring wells. A confidence level of 90 percent (where the probability of the relation being one of chance is less than or equal to 10 percent, or $\alpha = 0.10$) was considered a “significant correlation.” Low Spearman correlation coefficients generally indicate a weak correlation; however, in the case of the data set that was used in this analysis, the sample size was sufficiently large ($n > 300$) that even small coefficients could be significant. Nonparametric statistics, such as the Kruskal-Wallis rank sum test, were used to compare groups of data. The Kruskal-Wallis rank sum test is a robust nonparametric test for differences among groups; this test is largely unaffected by the presence of outliers. **Survival curves** were plotted to examine the difference among nitrate concentrations measured in private, public-supply and monitoring wells in the glacial aquifer system. Survival curves resemble upside-down **cumulative frequency distribution curves** that can take into account **censored data** (Helsel, 2005). The nonparametric rank of the nitrate concentrations was used to draw the survival curves.

Data used to model probabilities and predictions of nitrate concentrations were derived mostly from spatial data; however, some data were from well-construction records. Well-construction data used in the models include open-interval diameter, open-interval length, and depth to top of open interval. The term “open interval,” also called the screened interval, describes the length of the well that is open to the aquifer materials, where the water can enter the well. The open-interval-diameter variable was used to represent the diameter of the well. If the open-interval diameter was not available, then casing diameter was used. Spatial data include depth to saturated soil, mean silt in soil, and percent of alfisols, all of which were derived from the U.S. Department of Agriculture’s State Soil Geographic (STATSGO) database (Wolock, 1997). Ten-year mean (1992–2001) nitrogen fertilizer applied to farmland was derived from Ruddy and others (2006). The 10-year mean fertilizer application was calculated for this report by summing the county nitrogen-fertilizer applications from 1992 through 2001 (Ruddy and others, 2006) and dividing by 10. All spatial data were calculated as area-weighted averages for the area within a 1,640-ft radius (500 m) around the well. Soils data were area-weighted by the fraction of a soil unit defined in STATSGO that overlapped the area within a 1,640-ft radius around a well. The farm-fertilizer data were area-weighted by county and agricultural land-use area within a 1,640-ft radius around a well. Descriptions of variables, variable names, and data sources used for models

described in this report are shown in appendix 1. Alaska private-well data were not included in the models because spatial data were unavailable.

Nitrate concentrations in private wells in the glacial aquifer system were compared to results of a nationwide logistic regression model that predicts the probability of occurrence of nitrate concentrations greater than 4 mg/L as N in recently recharged (shallow) groundwater in all aquifers across the United States (Nolan, 2001; Nolan and others, 2002). The nationwide logistic regression model was applied to the glacial aquifer system in three ways: spatial overlay of points representing wells over the spatial model, using the exact regression equation of the nationwide model, and recalibrating the model specifically for the data from the glacial aquifer system.

Private wells in the glacial aquifer system were first overlain with the nationwide spatial model of probability of groundwater nitrate contamination (Nolan and others, 2002) to identify the probability category for each well. The observed and average predicted probabilities were calculated (app. 2) and plotted on a scatter plot for the same six categories used in Nolan and others (2002).

Nitrate concentrations in private wells in the glacial aquifer system also were compared to results of the nationwide logistic regression model by calculating the probability of occurrence of a nitrate concentration greater than 4 mg/L as N for each private well in the glacial aquifer system using the logistic regression equation and coefficients from Nolan and others (2002). The following equation was used:

$$P = e^{F(x)} / (1 + e^{F(x)}) ,$$

where

$$P = \text{probability of being in a category and}$$

$$F(x) = (-5.541 + (0.004 * 1992 \text{ fertilizer nitrogen, in kilograms per hectare}) + (0.016 * \text{percent agricultural land use}) + (0.229 * \text{natural log of 1990 population density}) + (0.025 * \text{percent of well-drained soils}) + (1.088 * \text{depth to seasonally high water table, in meters}) + (0.424 * \text{presence or absence of unconsolidated sand and gravel aquifers})).$$

The sources of variables used in the equation were calculated for a 1,640-ft radius around a well and provided more accurate representation of well attributes than information at the well point, except for the sand and gravel presence/absence indicator variable, which was still available only as a point on a polygon overlay. The results were divided into the same six categories as in Nolan and others (2002). The observed and the mean predicted probabilities were calculated for each of the six categories (app. 3) and plotted on a scatter plot. The number of wells with a measured nitrate concentration greater than 4 mg/L as N was tabulated for each of the six categories and the observed probability was calculated by dividing the number of wells with nitrate concentration greater than

4 mg/L as N by the total number of wells in the category. The mean predicted probability was calculated as the mean of all probabilities resulting from the logistic regression equation for each of the six categories.

The nationwide logistic regression model was recalibrated for private wells in the glacial aquifer system to obtain new coefficients (app. 4). The following recalibrated model equation was developed:

$$P = e^{F(x)} / (1 + e^{F(x)}) ,$$

where

$$P = \text{probability of being in a category and}$$

$$F(x) = (-4.274 + (0.0000938 * 1992 \text{ fertilizer nitrogen, in kilograms per hectare}) + (0.0083 * \text{percent agricultural land use}) + (0.0759 * \text{natural log of 1990 population density}) + (0.0058 * \text{percent of well-drained soils}) + (0.256 * \text{depth to seasonally high water table, in meters}) + (-0.0127 * \text{presence or absence of unconsolidated sand and gravel aquifers})).$$

The model variables were examined for significance of predicting probability of nitrate concentration greater than 4 mg/L as N using partial t-tests and an analysis of deviance chi-square test (app. 4). The predicted and observed probabilities of a nitrate concentration greater than 4 mg/L as N were calculated and plotted on a scatter plot using the same six probability categories as in Nolan and others (2002); however, only the three lower probability categories were represented by observed data. To better evaluate the recalibrated model fit, the predicted and observed probabilities were calculated and plotted on a scatter plot using 10 probability categories (app. 4).

Ten regional logistic regression models for assessing the probability of nitrate occurrence in groundwater in private wells were developed for thresholds from 1 to 10 mg/L as N (app. 5, app. 6). Logistic regression modeling was used to more closely examine relations between significantly correlated variables and nitrate concentrations in private wells. The logistic regression models used a binary variable to indicate (flag) whether the nitrate concentration in a sample was above or below a given threshold concentration and resulted in a predicted probability of encountering a nitrate value greater than the various thresholds, based on a set of independent predictor variables. The nitrate thresholds that were modeled overlapped each successive threshold. For example, using the threshold value “nitrate concentration greater than 1 mg/L as N” flagged all nitrate concentrations that were greater than 1 mg/L as N and overlapped the next threshold, “nitrate concentration greater than 2 mg/L as N.” In other words, every nitrate concentration that was flagged as a concentration greater than 2 mg/L as N was also flagged as a nitrate concentration greater than 1 mg/L as N.

A series of more than 100 variables in relation to measured nitrate concentration was visually examined on scatter plots. Forty-six of these variables appeared visually to be linearly related to nitrate concentration in private wells in the glacial aquifer system. The relations of the 46 variables to nitrate concentration were examined with Spearman correlations (app. 7). Eighteen variables were correlated at the 95-percent confidence level ($p < 0.05$).

All 10 logistic regression models for thresholds from 1 to 10 mg/L nitrate as N were developed using a semi-automated method in SAS (SAS Institute Inc., 2003) where the program selected a subset of independent variables using stepwise forward selection to predict the binary dependent variable (above or below the threshold). For a variable to be added and retained in a model, the addition of the variable to the model had to contribute to explaining the variability in the nitrate concentrations at a significance level of at least 0.05. In general, it was theorized that variables important for a model of one threshold would most likely also be important for the model of the next lower threshold because the concentrations in the samples in each successive threshold overlapped. An example of overlapping thresholds is the set of wells with nitrate concentrations greater than 2 mg/L as N, which is included in the set of wells with nitrate concentrations greater than 1 mg/L as N.

Each of the 10 logistic regression models was run three times. The first iteration for each model allowed the potential inclusion of 18 variables that were correlated to nitrate concentrations at the 95-percent confidence level (app. 8) plus the 10-year mean (1992–2001) nitrate fertilizer applied to farms, which was included as an additional potential predictor variable. Nitrogen fertilizer is known to be a potential source of nitrate in groundwater and has been used successfully as a predictor variable in other nitrate models (Nolan and Hitt, 2003; Nolan and Hitt, 2006). Some variables were used in all 10 models, whereas others were used in only 1 or 2 models. The variables that were used in only a few of the models were evaluated as to whether the inclusion was scientifically justified and whether there was a similar variable in the model for the next higher or lower threshold. If the variable was not used in the model for the next higher or lower threshold, it was excluded from further iterations of the logistic regression models. For example, in the first iteration, soil **bulk density** was one of the variables selected for the logistic regression model in which the threshold level was nitrate concentration greater 5 mg/L as N; however, soil bulk density was not selected to predict the probability of nitrate concentration greater than 4 mg/L as N or greater than 6 mg/L as N. Therefore, soil bulk density was not included as a possible predictor variable for subsequent iterations for the final models. The third iteration resulted in the final 10 logistic regression models, in which the variables were specified for each model (app. 8).

Several of the models include an interaction term. Interaction occurs when the relation between an independent variable and a dependent variable depends on the value of a

different independent variable (Kleinbaum and others, 1988; Ahlbom and Alfredsson, 2005).

The significance of all variables in each model was tested using the Wald chi-square test statistic, which is the squared ratio of the estimate to the standard error of the respective predictor. The variables used in the final logistic regression models also were evaluated for tolerances and variance inflation factors to ensure they were not collinear (SAS Institute Inc., 2003). Collinearity means that the variables might be representing similar variability in the models; therefore, collinear variables should not be used in the same model. The quality of the models was evaluated using two goodness-of-fit statistics: area under the Receiver Operating Characteristic (ROC) curve, and the Hosmer-Lemeshow goodness-of-fit statistic (Hosmer and Lemeshow, 2000). Each of these goodness-of-fit statistics evaluates the association of predicted probabilities and observed responses and is used to compare the results of the initial models to those of the final, modified models in which only eight of the possible variables were used.

The ROC curve statistic describes the model's ability to discriminate between the two cases of the binary dependent variable—in this case, nitrate concentration above a threshold (1 to 10 mg/L as N). The value of the statistic is the area under the ROC curve. The value ranges from zero to one and is a measure of the model's ability to discriminate between samples that have a nitrate concentration greater than a threshold (Hosmer and Lemeshow, 2000). A value of the ROC statistic that is above 0.8 is considered "excellent discrimination."

The Hosmer-Lemeshow statistic uses a Pearson chi-square test to evaluate the relation between observed and predicted probability. Predicted probabilities are divided into a number of percentile groups, usually deciles, and compared to the probability that an observed value will be in that percentile category (Hosmer and Lemeshow, 2000). Probabilities should be greater than an alpha level of 0.05 to indicate a good fit, which means that the observed and predicted probabilities in the percentile groups are not significantly different. The model for nitrate concentration greater than 4 mg/L as N was validated using an interactive variable selection method based on the Mallows' Cp statistic (Mallows, 1973).

A linear regression model also was developed to evaluate how accurately actual nitrate concentrations could be predicted based on the same variables used in the logistic regressions (regardless of p-value). The linear regression equation used the same variables (averaged by network) as the logistic regressions, cumulatively; however, the interaction term was omitted from the linear regression analysis (app. 9). The predictor variables were not selected specifically for this linear regression; therefore, some of the p-values are insignificant. The linear regression model included network means of the following variables: depth to saturated soil, percent silt in soil, percent alfisols soil type, 10-year mean nitrogen applied to farms, diameter of open interval, depth to top of open interval, and length of open interval. Each variable was averaged for wells within each network. Data for a network were used in

the linear regression model only if the network consisted of five or more wells. **Residuals** resulting from the linear regression model were plotted against quantiles of the standard normal distribution to check for a normal distribution. The approach of using network averages is based on the methods

of the Groundwater Vulnerability Assessment for Drinking Water (GWAVA-DW) model developed by Nolan and Hitt (2006). Unlike the GWAVA-DW model, the prediction model in this report was developed with a linear relation.

3.0 Background Nitrate Concentration

Background concentration of nitrate in the glacial aquifer system is low (less than 1 mg/L as N).

The background concentration of nitrate in groundwater is used as a threshold value above which influence by anthropogenic sources is indicated. This value is important because it identifies areas where the nitrate concentration may be increasing and provides the lowest threshold for modeling the vulnerability of private wells to nitrate contamination used in this study. Relative background concentration can vary depending, in part, on land use, hydrogeology, and climate. Several nationwide assessments of nitrate concentration compared to background concentration have been done, but without regard to the type of aquifer. In one of the first assessments of background concentration of nitrate, Madison and Brunett (1985) chose 3 mg/L as N to represent the level above which human input is likely. A concentration less than 0.2 mg/L as N represented natural background concentrations and a concentration between 0.21 and 3.0 mg/L as N was transitional, meaning the water may or may not be affected by human activity. Mueller and Helsel (1996), in their nationwide assessment of nitrate concentrations in groundwater from more than 10,000 wells, concluded that concentrations less than 2 mg/L as N were unaffected by human activity. More recently, Nolan and Hitt (2003) assessed background concentrations of nitrate in shallow, recently recharged groundwater beneath areas in the United States that are relatively unaffected by agricultural and urban land uses. Their assessment was based on 81 wells sampled as part of NAWQA in addition to 320 of the wells used by Mueller and Helsel (1996). Nolan and Hitt (2003) concluded that 1 mg/L as N is a reasonable upper-bound estimate of background concentration of nitrate. Data from 232 spring-water and 200 karst well-water samples in the Midwest were examined statistically by Panno and others (2006) to identify threshold nitrate concentrations by separating background from anomalous nitrate concentrations using cumulative probability plots. The thresholds of 2.5 mg/L as N (springs) and 2.1 mg/L as N (karst wells) were interpreted to represent the present-day background concentration of nitrate.

Reference wells in the glacial aquifer system that are minimally impacted by human activities, including two to three monitoring wells in each of the 15 NAWQA studies, were installed to assess background groundwater quality.

Data have been collected from these reference wells as part of NAWQA since 1994. The mean nitrate concentration in these wells was 1.16 mg/L as N (when censored data were set to the censored value, which may have slightly overestimated the mean), but the median was only 0.13 mg/L as N, indicating substantial variability. Groundwater from the glacial aquifer system reference wells in Nebraska had a much higher mean concentration of 6.2 mg/L as N that may be caused by the limited area unaffected by agricultural land use. The nitrate concentration in the Nebraska wells influenced the mean nitrate concentration for all reference wells in the glacial aquifer system, but the median, which is not affected by data extremes, was similar regardless of whether or not the estimate of the median included the data from Nebraska (0.13 mg/L as N or 0.10 mg/L as N, respectively). Nitrate concentrations in 75 percent of the wells were less than 0.49 mg/L as N. Therefore, the background nitrate concentration in the glacial aquifer system may vary geographically, but the nitrate concentration in more than half of the reference wells was less than 0.13 mg/L as N and in 75 percent of the reference wells was less than 0.49 mg/L as N. These values are slightly lower than the background concentration for all aquifers of 1.0 mg/L as N as determined by Nolan and Hitt (2003) based on the 75th percentile of nitrate concentrations and slightly higher than the background concentration of 0.14 mg/L as N used in a study of the glacial aquifer system in Connecticut (Garabedian and others, 1998). Water in the reference wells was mostly (72 percent) oxic (dissolved-oxygen concentration greater than or equal to 0.5 mg/L as N), so the concentration of nitrate in reference wells was not unusually low for the glacial aquifer system. The concentration of 1.0 mg/L as N used by Nolan and Hitt (2003) is a conservative estimate of background concentration of nitrate that is applicable to the glacial aquifer system because nitrate concentrations in water from 97 percent of glacial aquifer system reference wells were below this value. Therefore, for the purposes of this report, water containing nitrate concentrations greater than 1.0 mg/L as N is considered to be impacted by human activities, although local differences in background concentration may exist, such as in Nebraska (fig. 3–1).

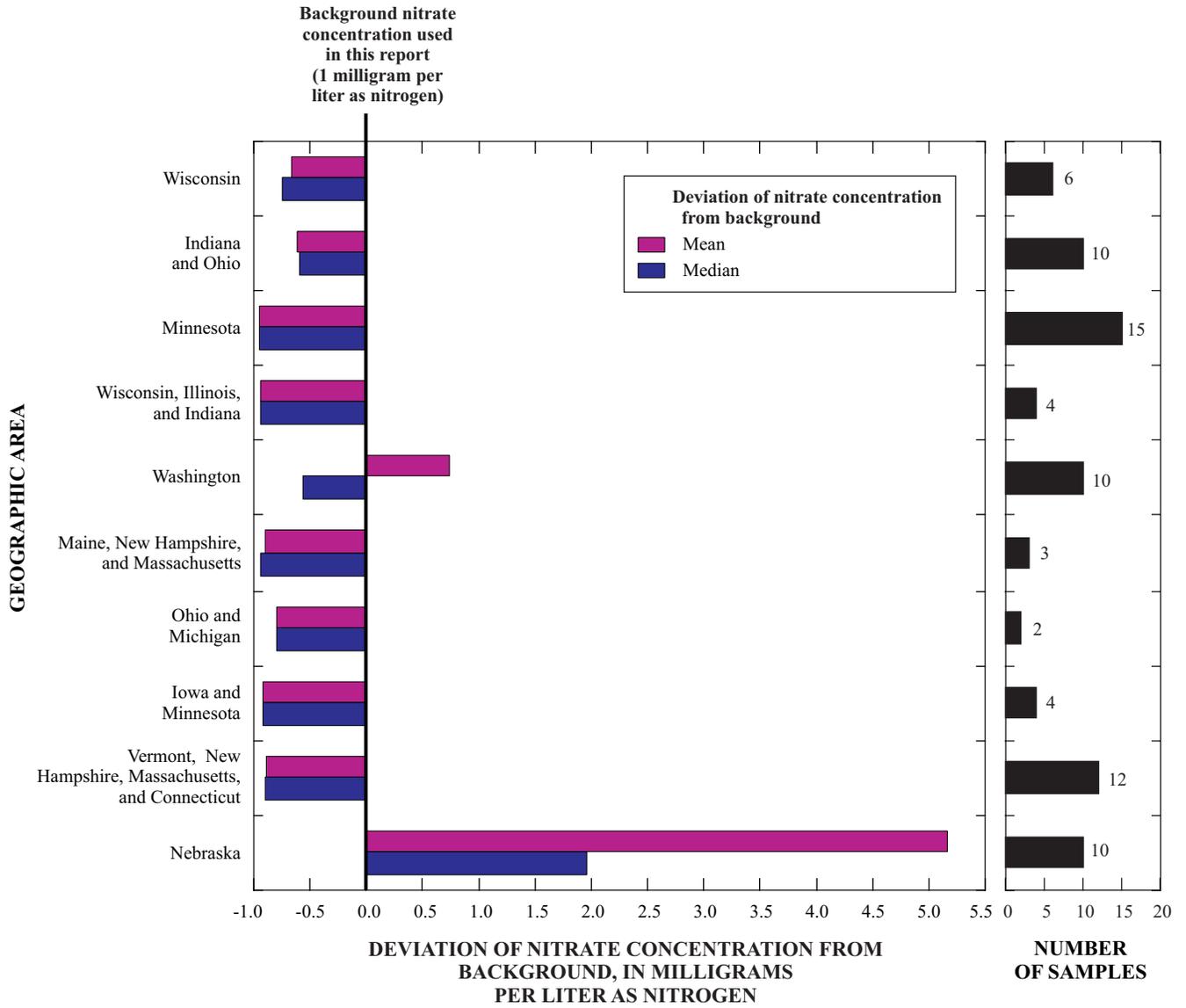


Figure 3-1. The graph shows the deviation of the mean and median nitrate concentrations from background in groundwater samples from reference wells selected to represent areas minimally affected by human activities. Mean and median nitrate concentrations in water from reference wells in Nebraska were greater than the background nitrate concentration used in this report (1 milligram per liter as nitrogen); however, mean and median nitrate concentrations in water from reference wells in other parts of the glacial aquifer system generally were less than the background concentration.

4.0 Comparison of Nitrate Concentration in the Glacial Aquifer System and Other Principal Aquifers

Median nitrate concentration measured in water from private wells in the glacial aquifer system was lower than that in water from other sand and gravel aquifers.

The USGS, as part of its NAWQA program, currently (2010) is studying the quality of water in principal aquifers of the United States (fig. 4–1). Sand and gravel aquifers generally are susceptible to nitrate contamination because of ease of transport of contaminants through the permeable aquifer materials. Among the principal aquifers currently being studied, the glacial aquifer system had one of the lowest median nitrate concentrations measured in private well water and the lowest median concentration (0.11 mg/L as N) among sand and gravel aquifers (fig. 4–2).

In contrast, the two highest nitrate concentrations (77 and 76 mg/L as N) in NAWQA wells in principal aquifers were measured in water from private wells in the glacial aquifer system (fig. 4–1). These concentrations were measured in

groundwater from older, weathered glacial deposits in Illinois and in thick glacial deposits in North Dakota. The next highest nitrate concentration measured in water from private wells in all the principal aquifers studied was 75 mg/L as N in water from the Central Valley aquifer system, a principal aquifer in central California that consists of sand and gravel (fig. 4–2).

Overall, the median nitrate concentration in water from private wells in the glacial aquifer system was low—even when compared to concentrations in water from other sand and gravel aquifers or bedrock aquifers (fig. 4–2, table 4–1). The low nitrate concentrations in water from the glacial aquifer system may be related to the extreme heterogeneity of the deposits comprising the aquifer system and associated soil processes (McMahon and Bohlke, 2005). Heterogeneity

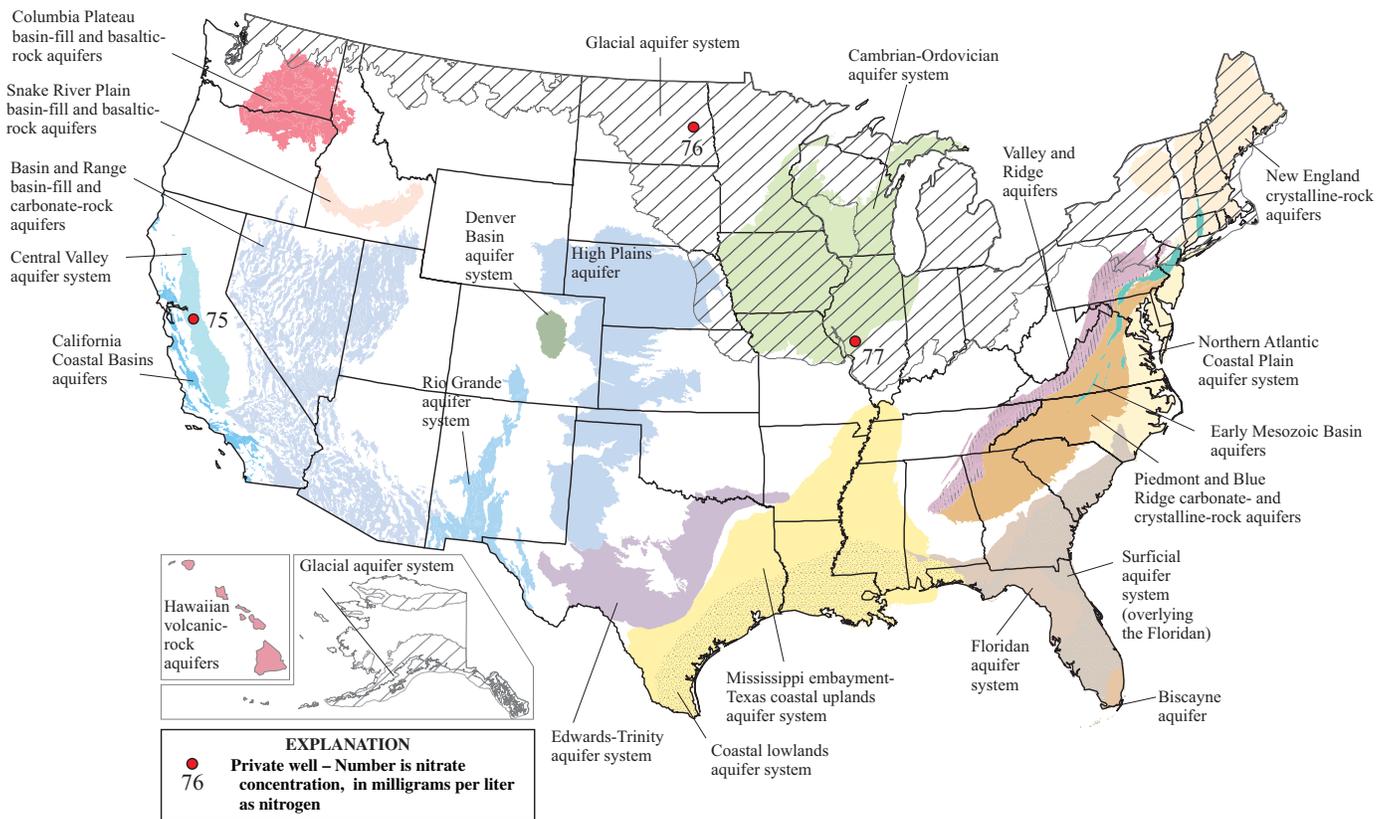


Figure 4–1. Locations of U.S. Geological Survey National Water-Quality Assessment (NAWQA) program regional assessments in principal aquifers in the United States. The glacial aquifer system covers about 953,000 square miles and parts of 26 States. The two highest concentrations of nitrate in private wells measured during the NAWQA program (1995–2005) were in water from the glacial aquifer system

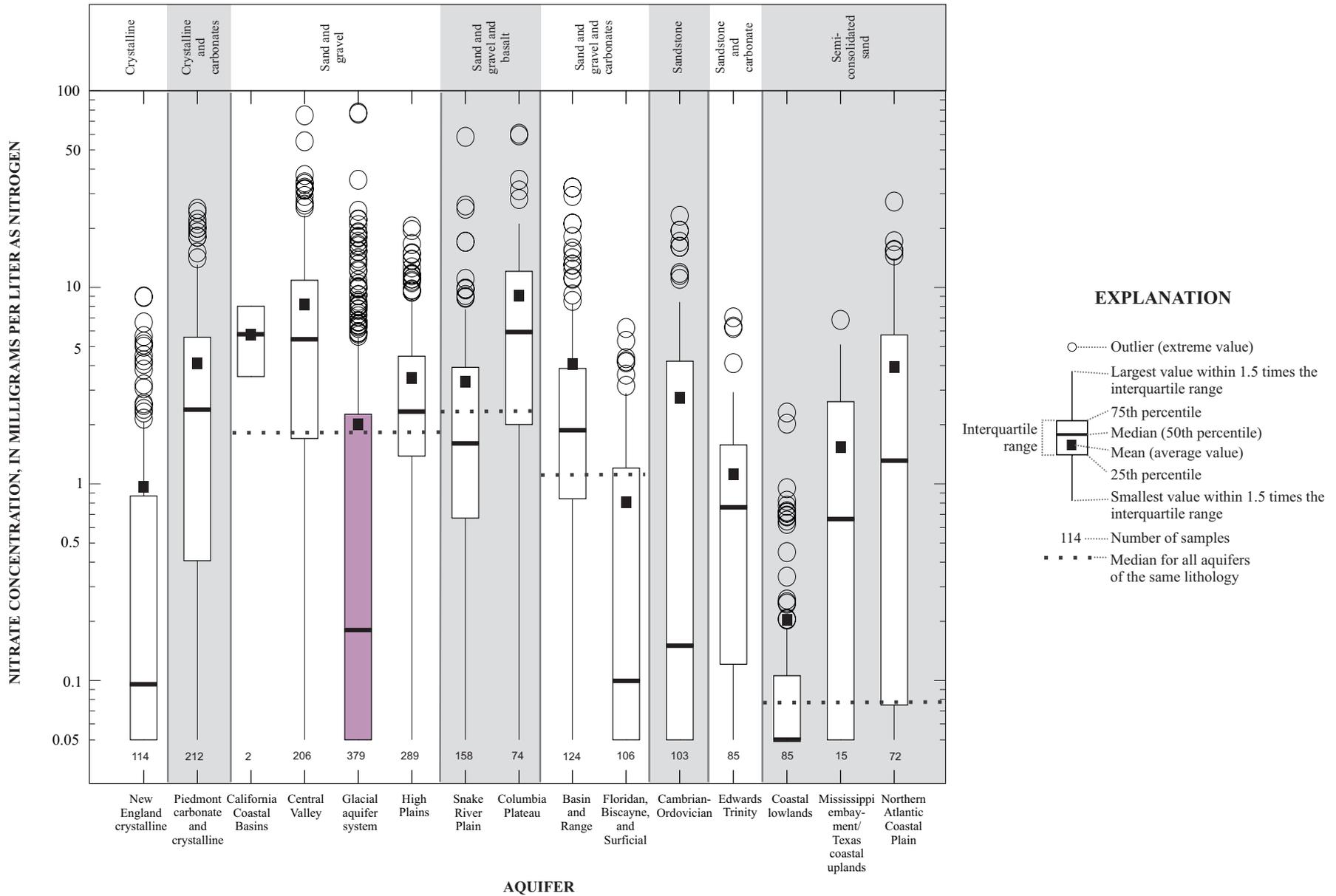


Figure 4-2. Although the glacial aquifer system had a larger range of nitrate plus nitrite concentration compared to other principal aquifers, the median concentration of nitrate was relatively low. Concentration of nitrate plus nitrite as nitrogen measured in water from private and public-supply wells in selected principal aquifers and other aquifers of different lithologies sampled as part of the U.S. Geological Survey National Water-Quality Assessment (NAWQA) Program (1995–2005) are shown.

18 Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells, Glacial Aquifer System

Table 4-1. Mean nitrate concentrations are higher in sand and gravel aquifers than most other lithologies.

Primary lithology	Aquifer	Nitrate concentration, in milligrams per liter as nitrogen							Number of samples
		Minimum	25th percentile	Mean	50th percentile (median)	75th percentile	Maximum	Standard deviation	
Crystalline	New England crystalline-rock aquifer	0.05	0.05	0.96	0.10	0.86	8.96	1.77	114
Crystalline and carbonates	Piedmont and Blue Ridge carbonate- and crystalline-rock aquifers	0.05	0.41	4.09	2.38	5.53	25	5.08	212
Sand and gravel	California Coastal Basins aquifers	3.51	4.62	5.74	5.74	6.86	7.97	3.16	2
	Central Valley aquifer system	0.05	1.73	8.11	5.42	10.79	74.65	9.52	206
	Glacial aquifer system	0.05	0.05	2.00	0.11	1.93	77	6.29	379
	High Plains aquifer	0.05	1.38	3.46	2.32	4.44	20.34	3.30	289
Sand and gravel and basalt	Snake River Plain basin-fill and basaltic-rock aquifers	0.05	0.67	3.31	1.6	3.88	58	5.88	158
	Columbia Plateau basin-fill and basaltic rock aquifers	0.05	2.05	9.08	5.9	11.75	60	11.19	74
Sand and gravel and carbonate	Basin and Range basin-fill and carbonate-rock aquifers	0.05	0.84	4.08	1.87	3.85	32	6.03	124
	Floridan	0.05	0.05	0.81	0.10	1.20	6.19	1.23	106
Sandstone	Cambrian-Ordovician	0.05	0.05	2.73	0.15	4.05	23	4.80	103
Sandstone and carbonate	Edwards-Trinity aquifer system	0.05	0.12	1.11	0.76	1.57	7	1.34	85
Semi-consolidated sand	Coastal lowlands aquifer system	0.05	0.05	0.20	0.05	0.11	2.30	0.38	85
	Mississippi embayment/Texas coastal uplands aquifer systems	0.05	0.05	1.54	0.66	2.09	6.8	2.19	15
	Northern Atlantic Coastal Plain aquifer system	0.05	0.08	3.92	1.31	5.58	27.21	5.48	72

of glacial sediments also is the main reason that redox varies more in the glacial aquifer system than in the other principal aquifers. The glacial aquifer system and the semi-consolidated sandstone aquifer in the southern coastal area of the United States had the largest percent of samples with mixed redox processes (McMahon and Chapelle, 2008).

In recent studies, decadal-scale changes in nitrate concentration (Rupert, 2008) and redox characteristics (McMahon and Chapelle, 2008) were compared among aquifer lithologies. Although nitrogen fertilizer and manure application overlying the glacial aquifer system is higher than on most other aquifer systems (Ruddy and others, 2006), nitrate concentrations are low, possibly as a result of the large attenuation capacity of the glacial aquifer system (fig. 4–3). In some places, such as the

unconsolidated deposits in alluvium that are part of the glacial aquifer system in Wisconsin (Saad, 2008), the aquifer system is less heterogeneous and contains minimal **organic matter**. In this part of Wisconsin, decadal changes in nitrate concentration in the glacial aquifer system generally were greater than those in lithologies of basalt, sandstone, and carbonate (Rupert, 2008). Additionally, decadal-scale changes in nitrate concentration were significantly smaller in water from wells with reduced conditions than in oxidized and mixed waters. The glacial aquifer system includes both environments: oxic conditions in deposits such as alluvium with gravel where decadal-scale changes in nitrate concentration are large and reducing conditions in deposits such as clay and silt deposits where decadal-scale changes in nitrate concentration are small.

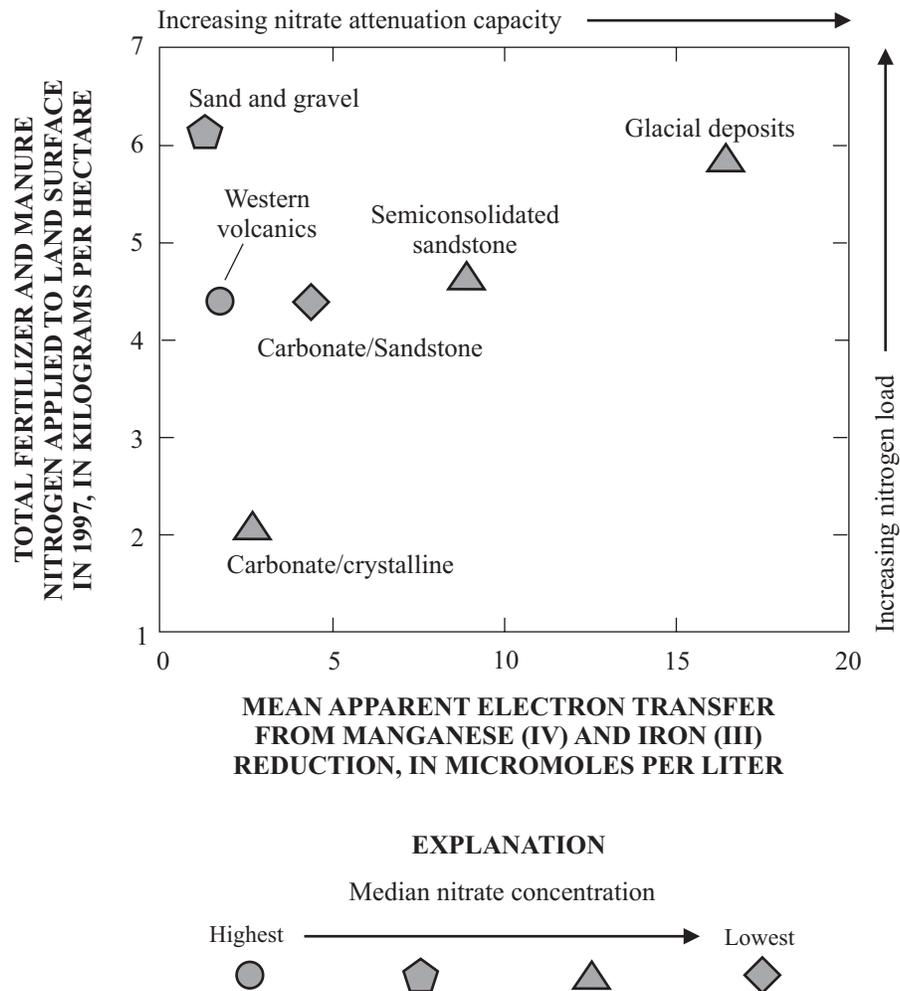


Figure 4–3. The glacial aquifer system receives a higher nitrogen load than most other aquifers, but the capacity for nitrogen attenuation also is greater. (From McMahon and Chapelle, 2008.)

5.0 Nitrate Sources

Nitrate detections in private wells were not strongly related to surface sources of nitrate.

Nitrate in the environment has diverse sources and is affected by a suite of biological and chemical transformation processes (Follett and Hatfield, 2001). Nitrate can be released from the soils when organic matter is broken down or mineralized. Previous studies assessing near-surface nitrate concentrations in groundwater have found relations between nitrate concentration and point sources such as waste lagoons, septic systems, and feedlots, or nonpoint sources such as precipitation, mineralization in soils, and fertilizer application. Atmospheric sources of nitrogen in the northeastern United States are associated with vehicle emissions or industry (Elliott and others, 2007). A study of private-well water supplies in the unconsolidated surficial aquifer on the Delmarva Peninsula showed that agriculture was the most important influence on groundwater quality on a regional scale (Debrewer and others, 2007). In England, it is estimated that more than 70 percent of the nitrate in natural waters is derived from agricultural land use (Lake and others, 2003). In a study of private wells in the glacial aquifer system, DeSimone and others (2009) state that nitrate concentration greater than 10 mg/L as N was primarily the result of human activities. Increases in agricultural use of fertilizer has been identified as a major source of nitrate to the glacial aquifer system (Puckett and Cowdery, 2002; Saad, 2008). Thus, nitrate is a known contaminant in many aquifers.

Correlation of sources to nitrate in water from private wells in the glacial aquifer system was weak on a regional scale. Nitrogen sources assessed in this study include land use (Nakagaki and others, 2007), fertilizer use (Ruddy and others, 2006), atmospheric deposition (Ruddy and others, 2006), density of septic tanks or cesspools (U.S. Census Bureau, 1992), density of public sewers (U.S. Census Bureau, 1992), and population density (Hitt, 2003), all of which were compared to nitrate concentrations in private water supplies (table 5–1).

Land use was evaluated as an indicator for nitrate in private wells, but it was not significant ($p > 0.05$) (table 5–1). Examination of the 1,640-ft- radius area around a well can provide a broad characterization of local land use affecting the well (fig. 5–1), but it may not adequately characterize the land use in the entire capture area of the well, which may also affect the quality of water in the well (Lorenz and others, 2003). DeSimone and others (2009) found that nitrate concentrations greater than 10 mg/L as N were most frequently detected in agricultural areas. Another potential indicator for nitrate is the 10-year mean (1992–2001) of nitrogen applied to farms (fig. 1–2), which was a significant relation for nitrate concentrations greater than background. The 10-year mean (1992–2001) of nitrogen applied to farms was not related to nitrate concentrations for all samples. The time period

Table 5–1. Comparison of nitrate concentration in groundwater from private wells to nitrogen sources indicates that atmospheric deposition of nitrogen, septic-system density, and sewer density are significantly related to nitrate concentration with 95 percent confidence, but the relations are weak.

[p, probability; <, less than; ft, feet; %, percent; values in boldface indicate significance at the 95-percent confidence level]

Variable	Spearman correlation	Significance
10-year-mean nitrogen deposition	-0.25	<0.0001
Septic-system density	-0.13	0.01
Public-sewer density	0.13	0.01
Percent forested land use within 1,640 ft of the well	0.10	0.06
Percent agricultural land use within 1,640 ft of the well	-0.10	0.06
Percent urban land use within 1,640 ft of the well	0.09	0.09
Forested land use greater than 50% within 1,640 ft of the well (binary)	-0.08	0.10
10-year mean nitrogen fertilizer application to farms	-0.07	0.16
Urban land use greater than 50% within 1,640 ft of the well (binary)	0.06	0.24
Agricultural land use greater than 50% within 1,640 ft of the well (binary)	-0.04	0.43
Population density (for the year 2000)	0.02	0.72

1992–2001 was selected because it coincides with the range of sampling dates in this study.

Other sources may be related to nitrate concentration in private wells, but the variability of the data obscures these individual relations. Three sources were weakly but significantly ($p < 0.05$) correlated to nitrate concentration in private wells—10-year mean nitrogen atmospheric deposition (-0.25 rho), septic-system density (-0.13 rho), and public-sewer density (0.13 rho). The relation of nitrate to wastewater was represented by the percentage of housing units served by sewers or using septic systems (table 5–1). The statistical result based on available data is counterintuitive if the nitrate source is sewers. The correlation results indicated that when the percentage of housing units served by public sewers increased, the nitrate concentration in private wells increased, but when the percentage of housing units using septic tanks or cesspools increased, the nitrate concentration in private wells decreased. The percentage of housing units served by sewers or septic systems was interpreted from 1990 census data (U.S. Census Bureau, 1992). There are several possible

reasons for these counterintuitive relations. Public sewers generally are associated with areas of high population density ($p < 0.05$ significance). Other sources of nitrate, including lawn fertilizer applications, could be present in these high population areas. In addition, new sewer or septic-system subdivisions are commonly constructed on land previously used for agriculture, and these areas may have subsurface drains that impede the transport of nitrate to groundwater. Although septic systems may not be a nitrate source on a regional scale, septic systems could affect groundwater quality locally.

Another gauge of sewer and septic-system impact on groundwater is the presence of fecal indicators. Analyses for fecal indicator bacteria in samples of groundwater from 18 principal aquifers indicated that counts of coliform bacteria were significantly lower in the glacial aquifer system than in other aquifer systems sampled as part of NAWQA (Embrey and Runkle, 2006). This may be another line of evidence indicating that sewers and septic systems are not major sources of nitrate to the glacial aquifer system on a regional scale.

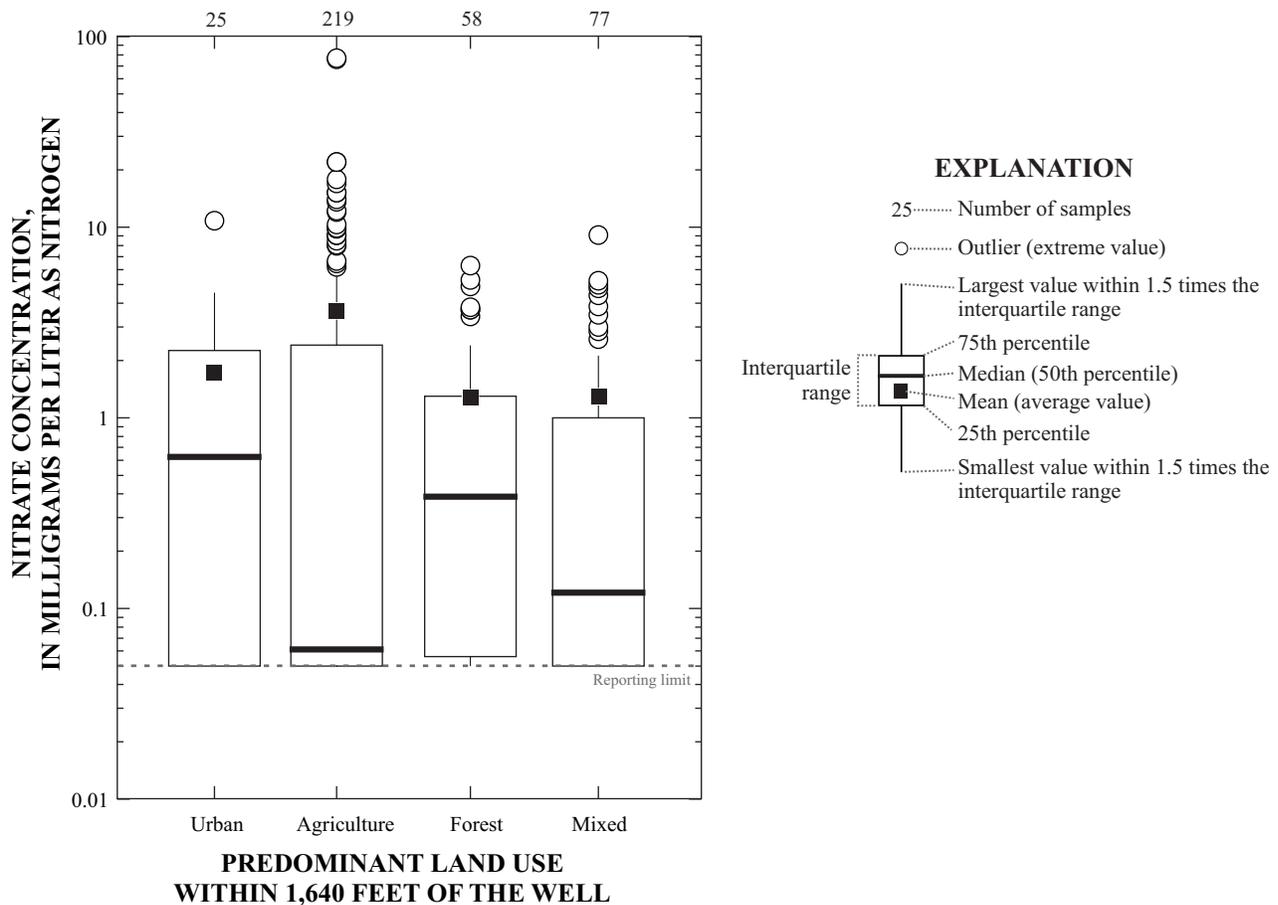


Figure 5–1. Median nitrate concentration was lowest in water from wells in areas where agriculture is the predominant land use within a 1,640 foot (500 meter) radius of the well, but nitrate concentrations in these areas also were the most variable. Median nitrate concentration was highest in areas where urban land use is predominant within a 1,640 foot radius of the well.

6.0 Nitrate Transport

6.1 Nitrate and Groundwater Age

The flow of groundwater in a regional, heterogeneous system like the glacial aquifer system is complex. Groundwater from wells in the glacial aquifer system commonly is a mixture of old and young water, regardless of depth, because “short-circuiting” allows mixing of older and younger groundwater over short distances and at various depths.

Groundwater generally moves slowly from its point of recharge to its point of use or discharge. Groundwater age, defined as the time elapsed since recharge of water, has important implications for understanding the vulnerability and sustainability of the aquifer. Aquifers that are more vulnerable to contamination originating at the land surface generally contain more recently recharged water than aquifers that are less vulnerable (Clark and Fritz, 1997). In the Loess Hills, southwestern Iowa, time-of-travel calculations and tritium-concentration data indicated that groundwater resides in these aquifers for decades (Tomer and Burkart, 2003). A study of nitrate and groundwater age of the glacial aquifer system in Canada showed that nitrate concentrations in groundwater recharged after 1960 contained nitrate in concentrations near 10 mg/L as N, but older groundwater contained nitrate in concentrations less than 2.5 mg/L as N (Johnston and others, 1998). The older groundwater was present under anaerobic conditions suitable for denitrification (the conversion of nitrate to nitrogen gas (N_2)), indicating that the removal of dissolved oxygen (DO) may be a slow process taking decades or that the older groundwater may simply predate the period of intensive fertilizer application.

Groundwater age can be determined by examining constituents that were introduced to the water at the time of recharge. For example, tritium is a radioactive tracer whose natural presence in groundwater is negligible but was found at high levels after nuclear-weapons testing occurred in the 1950s. Tritium decays with a half-life of 12.32 years, so groundwater that does not contain detectable tritium is inferred to have recharged prior to 1955 (the beginning of atmospheric nuclear-weapons testing). In contrast, groundwater that contains tritium (above 4 tritium units (TU)) is inferred to include at least some water that recharged after 1955 (Clark and Fritz, 1997). Groundwater-age indicators used in this report include tritium, tritium-helium, chlorofluorocarbons (CFCs), and CFCs with tritium. CFCs, which were introduced to the atmosphere with refrigeration and air conditioning, are useful tracers for groundwater age because these compounds are mostly resistant to degradation. Denitrification, a microbial-mediated process that transforms nitrate to nitrogen gas in oxygen-deficient environments, can affect groundwater-age estimates by degrading CFCs, thereby resulting in decreased concentrations of CFCs. A study in Canada found the denitrifying zone to be

6 to 7 m deep in a glacial aquifer, a zone below which use of CFC data resulted in an overestimation of the groundwater age (Sebol and others, 2007). It is becoming more difficult to accurately date groundwater with tritium because the half-life is relatively short and the concentrations are small and diminishing. In this report, estimated dates of recharge in or prior to 1955 are grouped together as “old” and estimated dates of recharge after 1955 are “young”. Wells for which only tritium data are available are considered to contain recently recharged water if the tritium concentration was greater than 4 TU.

Most groundwater in private wells is of mixed age. A NAWQA study of public-supply wells has shown that mixing of groundwater of different ages occurs even in samples from wells with short screens (Eberts and others, 2006). Water in wells in close proximity to each other may have very different groundwater ages that depend on factors related to nitrate

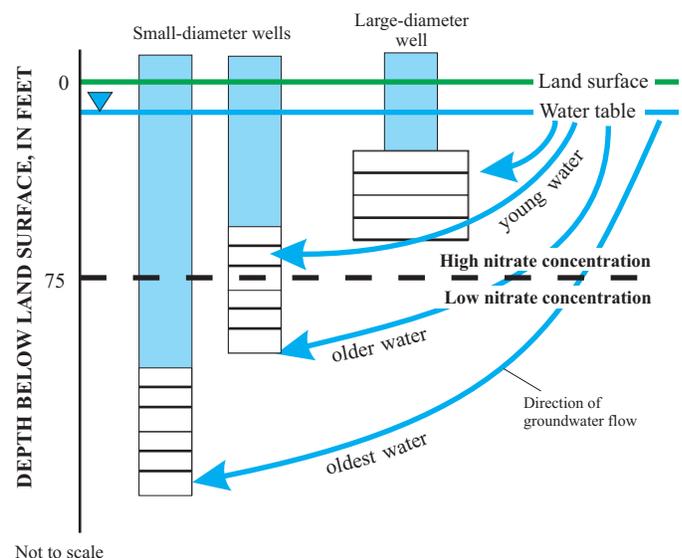


Figure 6.1–1. Groundwater flow to large- and small-diameter wells may represent recharge that was a mixture of groundwater of different ages (Sandra Eberts, U.S. Geological Survey, written commun., 2007). The large-diameter wells generally are shallow and contain “young” (recharged after 1955) water.

transport, such as depth of the well within the aquifer system and changes in geologic material (fig. 6.1–1). Water from some private wells in which the groundwater was mostly old contained recently applied pesticides, indicating mixing of old and young groundwater (Warner and Morrow, 2007). In many instances, the older water contained relatively low concentrations of nitrate because of denitrification or recharge of the aquifer prior to land application of nitrogen-containing fertilizer. As a result, groundwater that is now used for private supply may contain only a small fraction of water that recharged during periods of intensive nitrate application. The contamination of recently recharged groundwater near the water table can be an additional concern because this water will eventually recharge the deeper aquifer.

Approximately 210 samples from private wells were analyzed for groundwater age and nitrate concentration (table 6.1–1). Nitrate concentrations in young and old groundwater from private wells differed by an order of magnitude. The mean nitrate concentration in mostly young groundwater was 3.0 mg/L as N (median of 0.41 mg/L as N), whereas the mean nitrate concentration in old groundwater was 0.27 mg/L as N (median of less than 0.05 mg/L as N). All private wells in the glacial aquifer system in which nitrate concentrations exceeded 4 mg/L as N contained predominantly young groundwater (fig. 6.1–2). Generally, the private-well water samples that are old groundwater were collected from deeper

wells than the samples that are young groundwater, but mixing and “short-circuiting” (movement of groundwater along preferential flow paths) can occur. The mean depth of private wells in the glacial aquifer system with old water was 145 ft, and the mean depth of private wells with young water was 63 ft. Water from private wells that was young and oxic (DO greater than or equal to 0.5 mg/L) was more likely to have nitrate contamination than older and more reducing water (table 6.1–1). The nine samples in which nitrate concentrations were greater than or equal to the MCL of 10 mg/L as N were collected from recently recharged water (young) under oxic conditions. Old groundwater in private wells was mostly **anoxic** (fig. 6.1–3). Approximately 36 percent of the private wells sampled as part of this study contained older groundwater. Most of the private wells in which nitrate concentrations are greater than 1 mg/L as N contained young, oxic groundwater (table 6.1–1). The wells that contain young groundwater under anoxic conditions (DO less than 0.5 mg/L) may be subject to denitrification that could decrease the nitrate concentration. The necessary conditions for denitrification include (1) presence of specific bacteria (microbial denitrifiers); (2) availability of electron donors such as organic carbon; (3) low DO concentration; and (4) availability of nitrate as nitrogen (Follett, 2001). In the glacial aquifer system, bacteria are present. The availability of organic carbon is greater in areas with buried soils (paleosols), less sand, and more clay or peat or other organic

Table 6.1–1. Most detections of nitrate in private wells were in “young” (recently recharged) groundwater that is oxic. In this study, only samples of young, oxic groundwater contained nitrate at concentrations greater than 10 milligrams per liter as nitrogen.

[<=, less than or equal to; >, greater than; >=, greater than or equal to; mg/L, milligrams per liter; N, nitrogen; <, less than]

Oxidation-reduction state	Number of samples with specified nitrate concentration											
	Number of samples in each groundwater-age category		Percent of samples in each groundwater-age category		Nitrate concentration <= 1 mg/L as N	Nitrate concentration > 1 mg/L as N ¹	Nitrate concentration >= 10 mg/L as N	Nitrate concentration <= 1 mg/L as N	Nitrate concentration > 1 mg/L as N ¹	Nitrate concentration >= 10 mg/L as N	Median nitrate concentration (mg/L as N)	
	Young	Old	Young	Old	Young			Old			Young	Old
Oxic (dissolved oxygen concentration >= 0.5 mg/L)	82	22	60	30	34	48	9	16	6	0	1.95	0.05
Anoxic (dissolved oxygen concentration < 0.5 mg/L)	55	51	40	70	46	9	0	51	0	0	0.05	0.05

¹. The set of samples in which nitrate concentrations are greater than 1 mg/L as N contains the set of samples in which nitrate concentrations are greater than 10 mg/L as N.

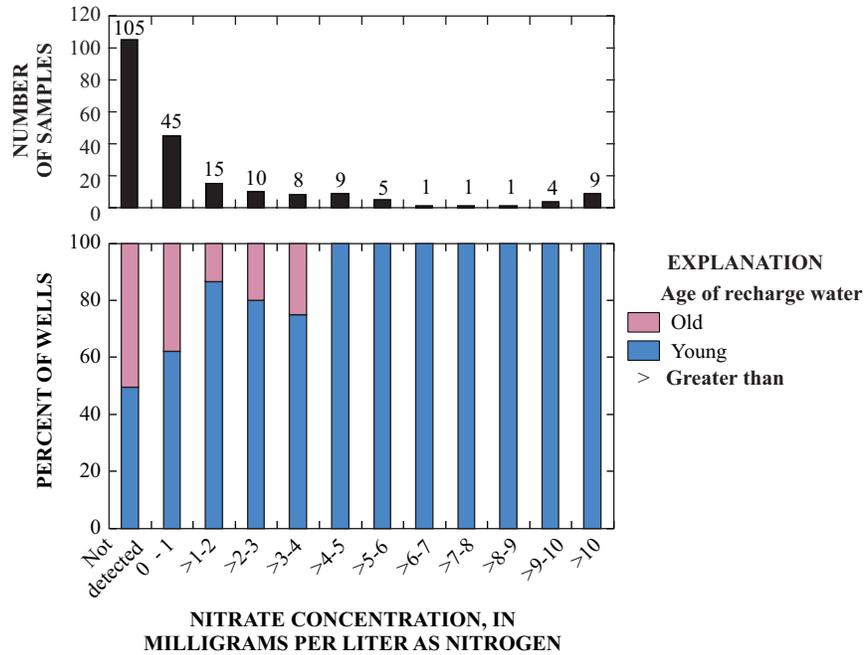


Figure 6.1-2. All samples in which nitrate concentrations were greater than 4 milligrams per liter as nitrogen were from private wells containing only groundwater that was recharged after 1955 (“young”).

matter. Dissolved oxygen concentrations tend to decrease with increasing residence time and depth (Puckett and others, 2002). Additionally, Puckett and others (2002) found that riparian-zone groundwater older than 8 years in the glacial aquifer system was extremely oxygen deficient (O_2 less than 0.51 mg/L).

There were differences in groundwater age for wells of the same depth in different framework areas that likely result from the heterogeneity of the aquifer across the northern United States. No samples from private wells in the Eastern area of the glacial aquifer system were analyzed for age indicators. Samples from the West-central area of the glacial

aquifer system were a mixture of nearly equal amounts of old and young groundwater (fig. 6.1-4). The groundwater from private wells in the West area was mostly young water, whereas samples from the West-Central area contained more young than old groundwater. Private-well networks in two local areas of the glacial aquifer system, central Illinois and northeastern Nebraska, contained mostly old water. In Illinois, this water is in a deep, buried bedrock valley aquifer; in northeastern Nebraska, it is in poorly permeable glacial till (Stanton and others, 2007). Finally, groundwater age indicates the time it takes recharge to travel to a well and this varies with location and is related to redox conditions.

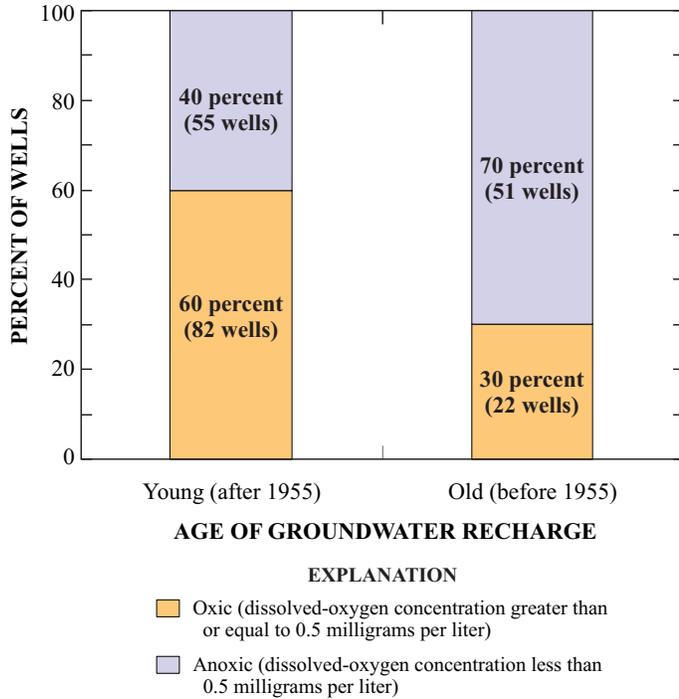


Figure 6.1-3. Groundwater recharged after 1955 (“young”) consisted of a higher percentage of oxidic water than groundwater recharged before 1955 (“old”). Nitrate is mobile under oxidic conditions.

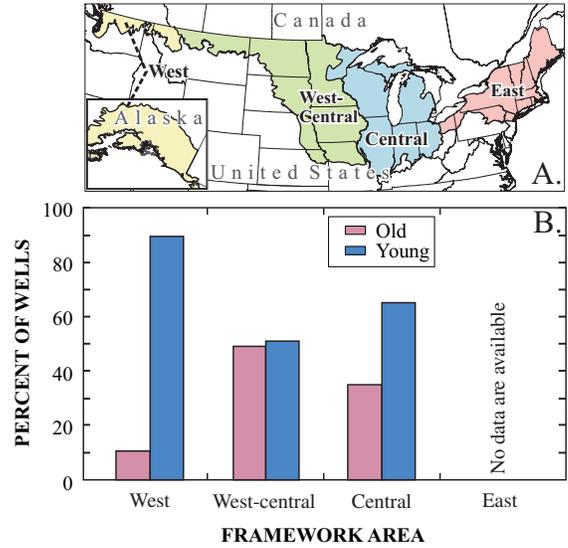


Figure 6.1-4. (A) The glacial aquifer system in the United States has been divided into four framework areas on the basis of physiography and source of aquifer material. (B) Private wells in these areas contained a mixture of old and young groundwater. The wells in the West area contained mostly “young” (recently recharged) groundwater.

6.0 Nitrate Transport—Continued

6.2 Nitrate Concentrations and Depth

Nitrate concentrations in water from private wells varied vertically more than areally.

The glacial aquifer system is the shallowest aquifer in an area where multiple stacked aquifers are used for supply. For example, the Cambrian-Ordovician aquifer is a principal aquifer underlying the glacial aquifer system in the Midwest, and the New England crystalline-rock aquifers underlie the glacial aquifer system in the Northeast. Generally, because private-well owners have neither the need nor the resources to drill deep wells (greater than 500 ft) into the underlying aquifers, their wells are screened in the shallow glacial material, except in New England, where private wells are commonly drilled into bedrock.

In the glacial aquifer system, nitrate concentrations increased as well depth decreased (fig. 6.2–1). Approximately half of the private wells in this study are less than 75 ft deep. The private wells within 20 ft of the land surface are most susceptible to high nitrate concentrations. The mean nitrate concentration in water from the private wells that are 20 ft deep or less (9.6 mg/L) was close to the MCL, but these wells account for only 1 percent of the wells in the study (fig. 6.2–2).

Nitrate concentrations were higher and more variable in water from private wells that are 75 ft deep or less (standard deviation, 8.5 mg/L; maximum concentration, 77 mg/L as N) than in wells that are deeper than 75 ft (standard deviation, 1.5 mg/L; maximum concentration, 9.1 mg/L as N at 78 ft). Several factors contribute to this difference. Most private wells less than 75 ft deep draw young groundwater that likely was recharged during times of greater fertilizer application rather than older water that is typically tapped by deeper wells

(table 6.2–1). Additionally, the shallower groundwater was mostly oxic, whereas the deeper groundwater was commonly anoxic and conditions are conducive to denitrification. Therefore, owners of shallow private wells need to be aware of the potential for nitrate contamination.

Regardless of the source of nitrogen, geochemical processes such as redox can limit nitrate concentration at depth. Redox is a major factor in denitrification, which has been shown to decrease the nitrate concentration in the glacial aquifer system (McMahon and Chapelle, 2008; Tesoriero and others, 2007; Puckett and others, 2002). Therefore, although fertilizer application above the glacial aquifer system may be large, redox changes with depth in the aquifer can facilitate denitrification and consequently, decrease the concentration of nitrate. DO concentration can be easily measured in the field to help understand geochemical conditions in the aquifer (Lewis, 2006). McMahon and Chapelle (2008) found that nitrate concentrations were significantly higher in water samples in which DO was greater than or equal to 0.5 mg/L than in samples containing DO less than 0.5 mg/L. Deeper wells tend to contain older water with less DO; therefore, nitrate concentrations also tend to be low. Seventy percent of the oxic groundwater samples from private wells in the glacial aquifer system were from wells that are 75 ft deep or less (table 6.2–1). All concentrations of nitrate greater than the MCL of 10 mg/L as N were in samples from private wells that are 75 ft deep or less.

Table 6.2–1. Shallow wells (75 feet deep or less) generally were associated with higher nitrate concentration, younger groundwater, and more oxic conditions than wells deeper than 75 feet.

[mg/L, milligrams per liter; N, nitrogen; DO, dissolved-oxygen concentration; ≥, greater than or equal to; <, less than; >, greater than]

Depth range (feet below land surface)	Number of wells	Mean nitrate concentration (mg/L as N)	Maximum nitrate concentration (mg/L as N)	Percent of wells with young groundwater in depth range	Percent of wells with old groundwater in depth range	Percent of wells from which all samples are oxic (DO ≥ 0.5 mg/L)	Percent of wells from which all samples are anoxic (DO < 0.5 mg/L)
0–75	194	3.17	77	74	18	70	39
>75	184	0.77	9.11	26	82	30	61

Locally, heterogeneity of the glacial deposits makes it difficult to discern patterns of nitrate concentration. McMahon and Chapelle (2008) found that private wells in the glacial aquifer system, even those with short well-screen lengths, had a mixed redox signature resulting from closely spaced zones of differing redox conditions. These redox zones are often related to the amount of fine or coarse material overlying the aquifer. The mixed redox signature may be caused by localized zones of organic carbon and (or) slow groundwater-flow velocity. If the redox signature is mixed, then the likelihood of nitrate detection is mixed. Although groundwater generally becomes more reducing with depth, depth is not always a good indication of groundwater age because the characteristics of the aquifer materials affect the velocity of groundwater flow. In the Midwest, till is commonly clayey and poorly permeable, whereas the coarser-grained outwash and alluvial deposits have a wider range of permeability (Tomer and Burkart, 2003). Even within a short well screen or between

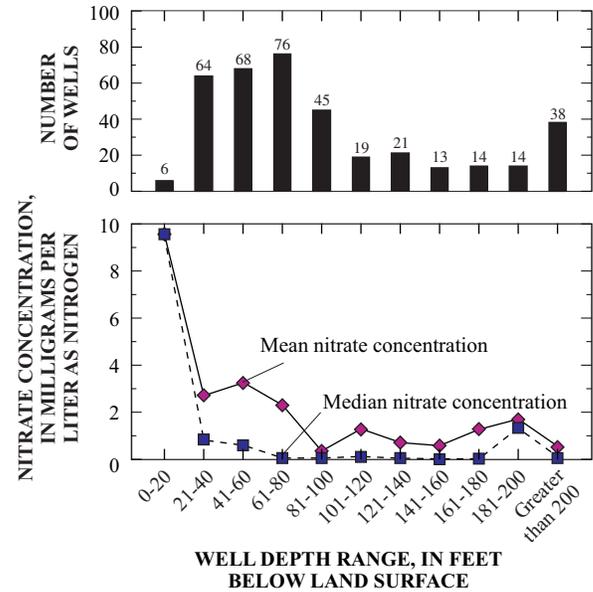


Figure 6.2-2. Mean and median nitrate concentrations were highest in water from private wells less than 20 feet deep, but most private wells were more than 60 feet deep.

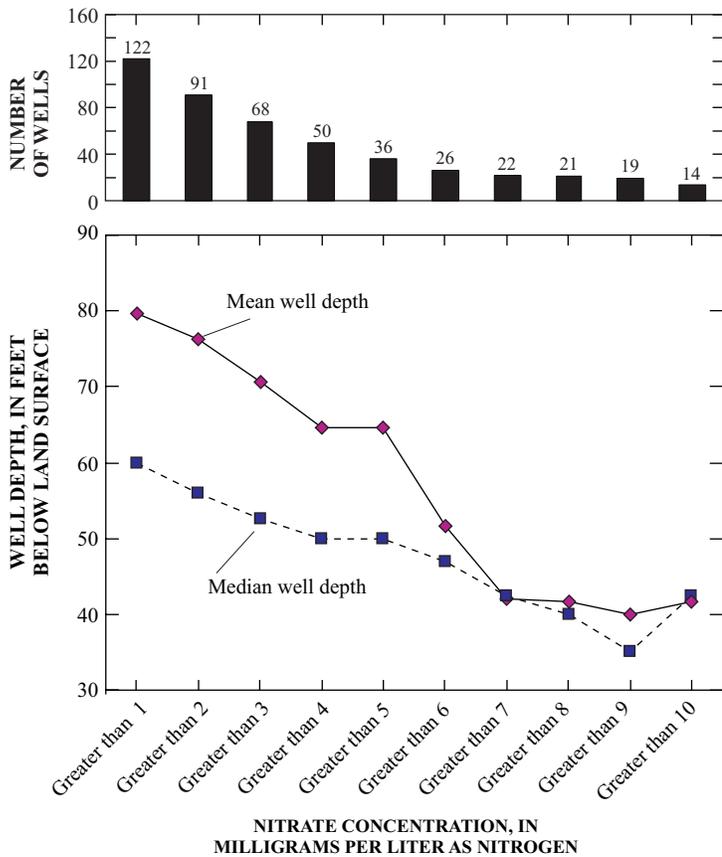


Figure 6.2-1. The difference in mean and median well depth shows the large variability of well depths associated with nitrate concentrations less than 6 milligrams per liter. The samples with concentrations greater than 6 milligrams per liter were from relatively shallow wells.

wells in close proximity to each other, there can be large differences in the amount of clay and sand. This heterogeneity of glacial deposits is helpful to private well owners concerned about nitrate contamination because it implies that redox may change quickly over short distances and, consequently, nitrate may not persist for long distances along flow paths in the glacial aquifer system. Regional patterns of decreasing nitrate concentration, increasing groundwater age, and changing redox with depth have been observed, but local heterogeneity of the aquifer materials makes it difficult to predict the probability of high nitrate concentration at a given well.

7.0 Nitrate Concentrations in Private, Public, and Monitoring Wells

7.1 Nitrate Concentrations in Water from Private Supplies

Although overall nitrate concentrations in water from private wells in the glacial aquifer system were low relative to the MCL, concentrations were highly variable over short distances and with depth.

The advantage of assessing only private wells is that it implicitly recognizes that the type of well will affect the water-quality results because well construction and well siting tend to differ among well types (Alley, 1993). An additional sampling-design consideration is the selection of which private wells to include in the regional assessment. The 379 private wells included in this study were selected because the wells represent local groundwater use. Differences in data-collection objectives and analytical methods can bias a regional assessment, but the consistency of the NAWQA sampling design and protocols assures similarity in random selection and analysis (Bartholomay and others, 2007). The mean and median nitrate concentrations (7.8 and 5.5 mg/L as N, respectively) were higher in water from private wells in networks designed to assess the effects of land use on water quality than in water from private wells in networks designed to assess the general water quality in the aquifer (2.0 and 0.1 mg/L as N, respectively). Nitrate concentrations in water from private wells of networks designed to assess effects of agricultural land use were strongly correlated with land-use practices near the well (agricultural land as a percentage of land use within 1,640 ft of the well). In water from private wells selected to assess land-use impacts, high concentrations of nitrate (greater than 7 mg/L as N) were found at all depths, but not in water from private wells selected to assess the overall quality of water in the aquifer system. In contrast, nitrate concentration was not significantly ($p < 0.05$) correlated with percent agricultural land use surrounding private wells selected to assess the quality of water in major aquifers, even though more than 50 percent of the area surrounding those private wells was agricultural (78 percent of the 379 wells in this study) (fig. 7.1–1). Therefore, the land use that affects the concentration of nitrate in private wells may be from farther than 1,640 ft from the well, or the well may draw older groundwater unaffected by recent land use. In regional water-quality analyses, well type and the purpose of sampling should be considered before selecting the wells. Thus, only data from the 379 private wells sampled for the purpose of regional assessment of major aquifers are used in this report. Nitrate was not a potential health concern at most of the 379 private wells sampled (fig. 7.1–2). The mean and median nitrate concentrations in water from these private wells (2.0 mg/L and 0.11 mg/L as N, respectively) were below the USEPA MCL for nitrate (10 mg/L as N). Fewer than 5 percent of the wells contained groundwater with nitrate at concentrations above 10 mg/L as N, which is a known human

health concern (fig. 7.1–2). This is unexpected because the nitrogen load at land surface above the glacial aquifer system is higher than that above most other regional aquifers (McMahon and Chapelle, 2008).

The distribution of nitrate in private wells of the glacial aquifer system differed by framework area (fig. 1–3, 7.1–3). The wells in the West-Central area had longer open intervals (almost 13 ft on average) than the wells in the other framework areas. The West-central area differs from the other framework areas in that the region consists predominantly of irrigated agricultural land underlain by clayey till (Arnold and others, 2008). The cyclic pumping and recharging of irrigation water in parts of this area may concentrate nitrate in the groundwater over time. Mean well depths and casing diameters of the private wells were similar among the framework areas. Large-diameter (10- to 45-in.) wells were sampled in all framework areas except the West, where the largest casing diameter was 6 in.

Nitrate concentrations varied throughout the extent of the glacial aquifer system but there was not a distinct areal pattern of elevated nitrate in water from private wells (fig. 11–1).

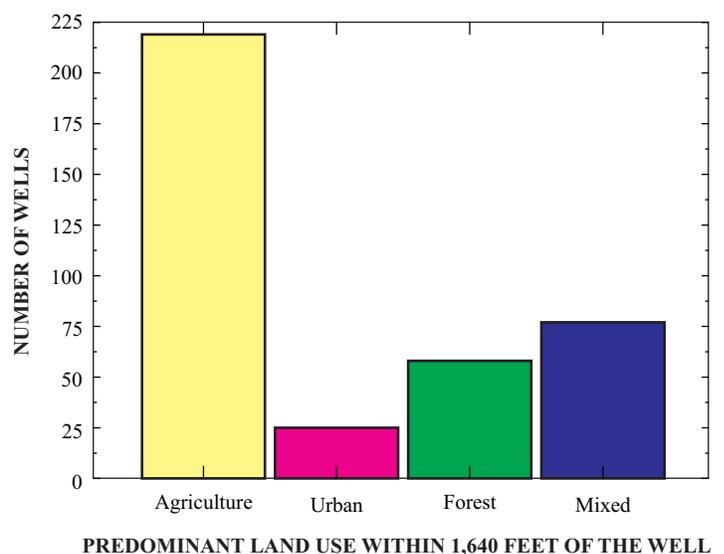


Figure 7.1–1. Agriculture was the predominant land use within 1,640 feet of the private wells sampled.

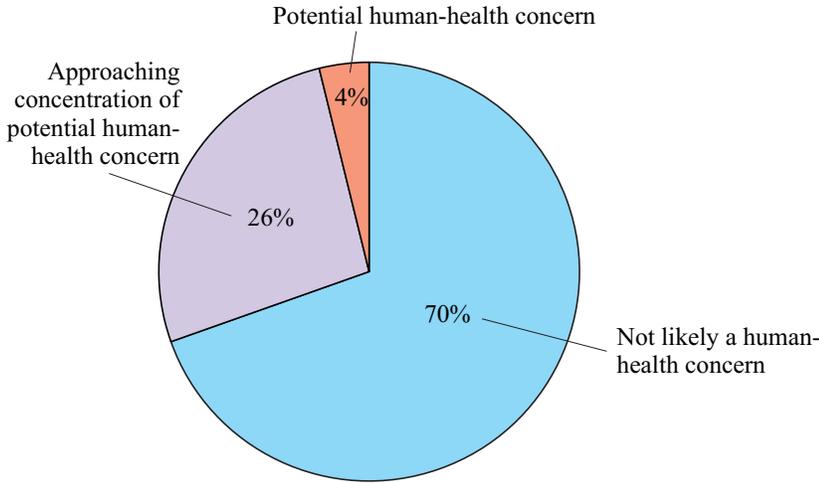


Figure 7.1-2. Nitrate concentrations in most of the private wells sampled in this study were less than the background nitrate concentration of 1 milligram per liter as nitrogen (<1 mg/L as N). The background nitrate concentration of <1 mg/L as N is less than one-tenth of the U.S Environmental Protection Agency Maximum Contaminant Level (MCL) of 10 mg/L as N; such concentrations are not likely a human-health concern. In this study, samples with nitrate concentrations between 1 and 10 mg/L as N are considered to approach concentrations of potential human-health concern, and nitrate concentrations greater than the MCL of 10 mg/L as N are of potential human-health concern.

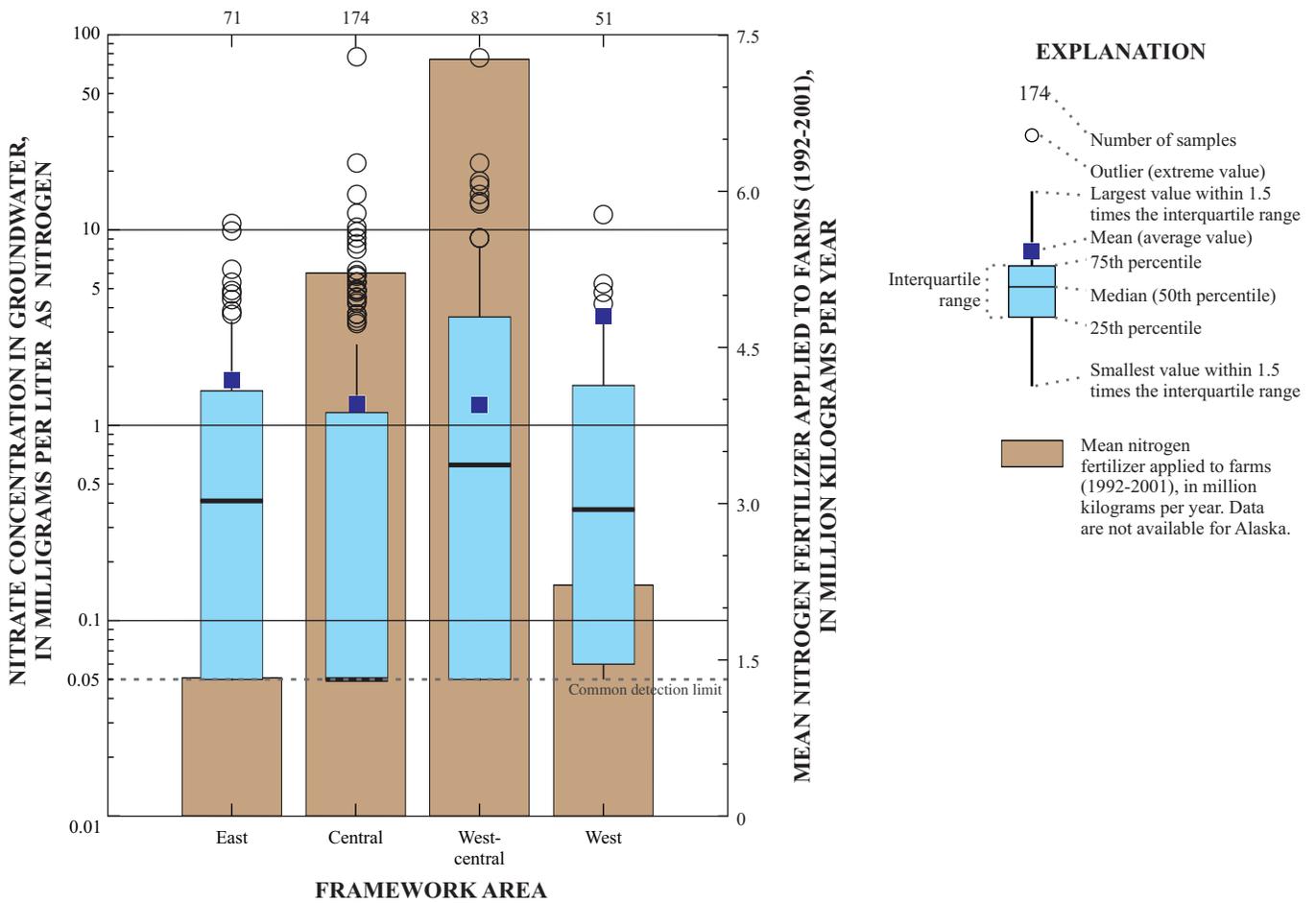


Figure 7.1-3. The Central framework area of the glacial aquifer system had a high nitrate application rate but a low median concentration of nitrate in groundwater from private wells. The Central and West-central areas had a large variability in nitrate concentrations. (Modified from Ruddy and others, 2006.)

In an exploratory data analysis, “hot-spot differentiation” (Environmental Systems Research Institute, 2008) was used initially to indicate geographic areas where nitrate concentrations were more than one or two standard deviations higher or lower than surrounding concentrations. Results of this analysis showed that west-central Illinois and east-central North Dakota areas are geostatistically different from other areas. High nitrate concentrations tend to be clustered in these areas compared to other parts of the glacial aquifer system. These same areas have relatively high 10-year-mean nitrate fertilizer application rates. In other areas of the glacial aquifer system, such as in the East area, nitrate concentration in water from private wells was lower than in the rest of the glacial aquifer system. Although there are areal differences in nitrate concentrations in water from private wells, these differences do not occur in a distinct pattern.

The glacial aquifer system is different from other unconsolidated aquifers because of the abundance of clay and organic layers in the aquifer. The heterogeneity of glacial

sediments results in highly variable rates of water recharge and, consequently, nitrate movement through the aquifer. The heterogeneity is one of the most important factors affecting variability in flow of water to the wells in the glacial aquifer system. The highest mean and median nitrate concentrations (3.6 and 0.63 mg/L as N, respectively) were in groundwater in the West-central framework area, which includes Iowa, Minnesota, Nebraska, and North Dakota.

Well construction also is important in determining the vulnerability of private wells in the glacial aquifer system to contamination with nitrate. Most of the private wells sampled for this study were constructed during 1990–2000, and it is water from these wells in which the mean concentration of nitrate was highest (fig. 7.1–4). Open-interval diameters in sampled private wells ranged from 1.25 to 45 in., but most were 6 in. Twenty-four wells had very large diameters (24–45 in.); most of these wells were constructed between 1990 and 2000. The larger-diameter wells are built to maximize groundwater storage within the well bore. Networks with

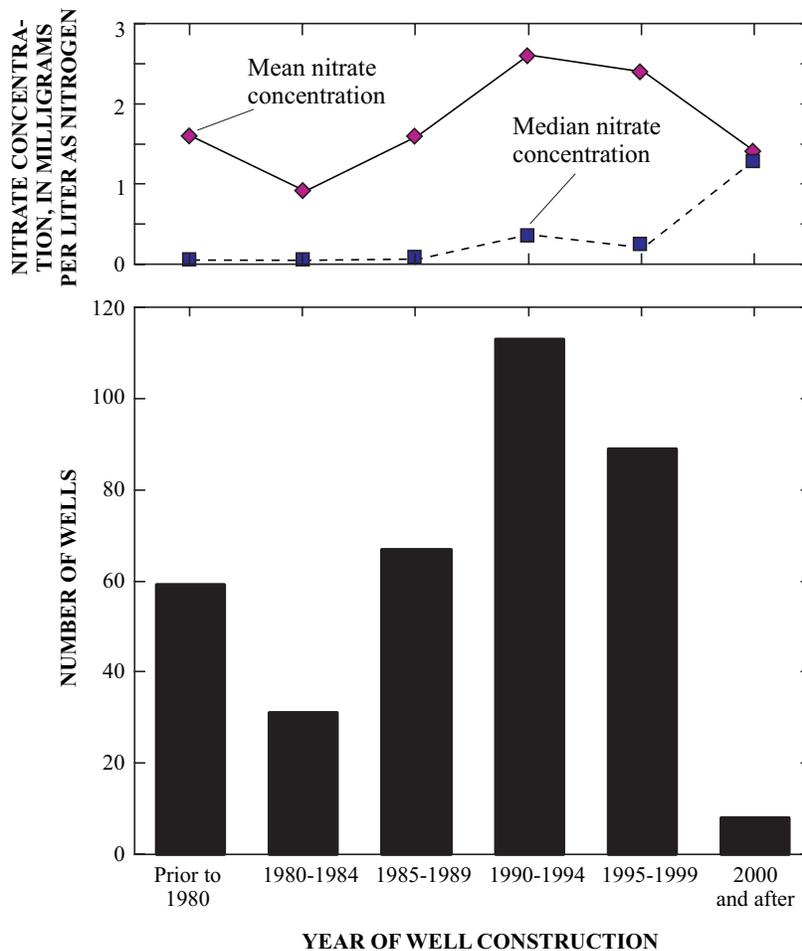


Figure 7.1-4. Most private wells sampled in this study were constructed during 1990-99 and the water from these wells sampled by NAWQA had the highest mean nitrate concentration compared to water from other private wells.

large-diameter wells (greater than 24 in.) are in Connecticut, Illinois, Iowa, Minnesota, and North Dakota.

Even in areas with poorly permeable glacial sediments and low transmissivity, the glacial aquifer system is used for drinking-water supply, as indicated by the large-diameter wells in these areas. In Illinois, the southern half of the State is an area of thin glacial sediments, and more than 50 percent of the private wells in this area are large-diameter wells (McKenna and Keefer, 1991). Although studies have shown that rural, private well construction and placement are factors that facilitate nitrate contamination of wells, it is important to note that well construction and placement are not sources of contamination.

The depth to water in sampled private wells varied. The deepest water level was 202 ft below land surface, but the median water level was 27 ft below land surface. Water levels were deepest in wells in the West and West-central framework areas (Nebraska, Kansas, Washington, and Alaska) or in the deep glacial aquifer system in the Central area (Illinois).

Although the glacial deposits are unconsolidated, the depth to water is not always equivalent to the depth of the water table because large amounts of overlying clay (mostly in the Midwest) can confine the water and cause differences in the hydraulic head or **perched groundwater**.

The relation of nitrate concentration to potential explanatory variables was complex because of the heterogeneity across the aquifer. The variables with the strongest correlations to nitrate concentration in private wells were soils (bulk density of soils, -0.29; percent alfisols, -0.21; percent organic matter, -0.14; percent **silt content**, 0.16; mean depth to soil saturation, 0.25), land-use infrastructure (percent of houses with drilled wells, -0.26; percent of houses with dug wells, 0.12; percent of houses served by septic systems, -0.13; percent of houses served by public sewers, 0.13), depth (well depth, -0.24; depth to top of open interval, -0.24), well construction (diameter of open interval, 0.23; open interval length, 0.13) and 10-year-mean atmospheric nitrate deposition (-0.25).

7.0 Nitrate Concentrations in Private, Public, and Monitoring Wells

7.2 Comparison of Nitrate Concentrations in Water from Private, Public, and Monitoring Wells

Nitrate concentration was significantly lower in groundwater from private wells than in groundwater from monitoring wells.

Public, private, and monitoring wells sampled as part of NAWQA differ with respect to well-construction characteristics (table 7.2–1). The median well-casing diameter and screen length were greater for public-supply wells (18 in. and 10 ft) than for private wells (6 in. and 5 ft) or monitoring wells (2 in. and 5 ft). Public-supply wells generally are open to more of the aquifer than private or monitoring wells so there is the potential to draw water from a large zone. The large diameter of public-supply wells is typically intended to increase production capacity, whereas in private wells the large diameter is commonly used in poorly producing areas to enhance water storage. The depths of public and private wells that were sampled as part of this study were similar (mean 91 and 96 ft below land surface, respectively; median 62 and 75 ft below land surface, respectively) (table 7.2–1). However, the private wells were cased at greater depths than the public wells, as shown by the median depth to the top of the open interval (71 ft and 53 ft, respectively) and generally had a shorter open interval than the public wells (5 ft and 10 ft, respectively). Because the monitoring wells were shallower (mean, 29 ft; median, 25 ft below land surface) than the other two well types, the open interval may be near to surface sources of nitrate.

More than 50 percent of the sampled private wells were surrounded by agricultural land use. The nature of a private well implies individual ownership; therefore, these wells are typically in isolated areas such as rural areas. Most of the sampled public-supply wells were in mixed-land-use areas (fig. 7.2–1). A majority of the sampled monitoring wells, like the private wells, were in agricultural areas. The sampled monitoring and private wells were similar with respect to surrounding land use (within 1,640 ft of the well), but differ with respect to depth and well-construction characteristics.

The differences in nitrate concentrations among private, public, and monitoring wells reflect the purpose of the well installation and well-construction characteristics. The nitrate concentration was higher in recently recharged groundwater from monitoring wells than in groundwater from private wells (table 7.2–1). The lower nitrate concentrations in the water from private wells may result from denitrification, dilution, or constraints on nitrate transport from land surface to the deeper private wells. The mean and median nitrate concentrations in water from public supplies were similar, indicating little variability; in contrast, the mean and median nitrate concentra-

tions in water from private supplies were different, indicating greater variability.

Groundwater used for private supply contained a larger percentage of old water than groundwater from public or monitoring wells (table 7.2–2). This means that public and monitoring wells are pulling in recharge water that was more recently in contact with the land surface and potential nitrate sources. Public-supply wells are known to withdraw groundwater of multiple ages even over short screen lengths (Eberts and others, 2006) because of the large volume of water withdrawn. The NAWQA studies on susceptibility of public-water supplies found that from 20 to nearly 100 percent of water produced from these wells was young groundwater, and drawdown by short circuiting may be the primary pathway by which contaminants enter the wells (Eberts and others, 2006). Similar mixing is likely to occur in private wells, but on a smaller scale.

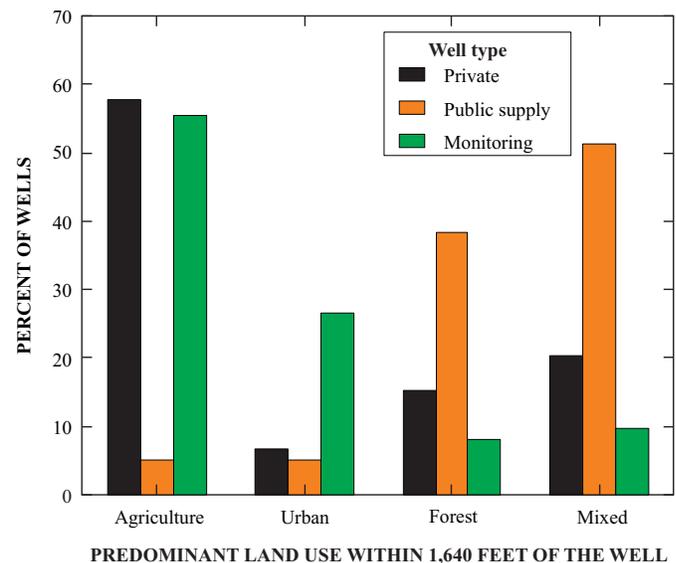


Figure 7.2–1 Most private and monitoring wells sampled were located in agricultural settings, whereas most public-supply wells sampled were in forested or mixed-land-use areas. The public-supply wells sampled by NAWQA were mostly in the northeastern United States, where forest and mixed land use are common.

Table 7.2-1. Private, public-supply, and monitoring wells in the glacial aquifer system differ with respect to groundwater nitrate concentration and well-construction characteristics.

Statistic	Well type			
	Private	Public supply	Monitoring	All well types
Nitrate concentration, in milligrams per liter as nitrogen				
Minimum	0.05	0.05	0.05	0.05
Mean	2.00	0.69	4.7	3.8
Median	0.11	0.56	2.0	0.89
Maximum	77	2.5	70	77
Number of wells	379	41	870	1,290
Well depth, in feet below land surface				
Minimum	12	27	7	7
Mean	96.23	91.09	28.69	50.48
Median	75	62.4	24.61	32
Maximum	365	420	137	420
Number of wells	378	41	870	1,289
Depth to top of open interval, in feet below land surface				
Minimum	3	21	2	2
Mean	91.06	77.61	23.05	43.55
Median	71	53.2	19	25
Maximum	357	348	113	357
Number of wells	347	38	867	1,252
Length of open interval, in feet				
Minimum	0	3	2	0
Mean	7.87	14.11	5.12	6.11
Median	8	15	5	5
Maximum	78	72	36.6	78
Number of wells	314	38	867	1,219
Open-interval diameter, in inches				
Minimum	1.25	4	0.17	0.17
Mean	7.05	15.56	2.08	3.95
Median	5.63	16	2	2
Maximum	45	24	4	45
Number of wells	371	40	867	1,278

Table 7.2-2. Samples from all well types in the glacial aquifer system represent a mixture of groundwater of different ages.

Well type	Number of wells for which age data are available	Percent of wells with old water	Percent of wells with young water	Mean depth, in feet below land surface
Private	212	35	65	92
Public	28	18	82	124
Monitoring	632	16	84	29

Kruskal-Wallis rank sum tests showed a statistically significant (95-percent confidence) difference in median nitrate concentrations between monitoring (1.97 mg/L as N) and private wells (0.11 mg/L as N). There was no statistical difference in median nitrate concentrations between private and public-supply wells although most public-supply wells sampled by NAWQA are in the east (fig. 7.2-2). Additionally, survival curves were plotted to examine the difference between nitrate concentrations measured in water from

private, public-supply, and monitoring wells in the glacial aquifer system. The survival curves indicate that, in the glacial aquifer system, nitrate concentrations exceeded 4 mg/L as N in about 12 percent of private wells and 36 percent of monitoring wells (fig. 7.2-3). Therefore, the results of this study indicate some differences in nitrate concentration with well type, especially between shallow monitoring wells and deeper public or private wells.

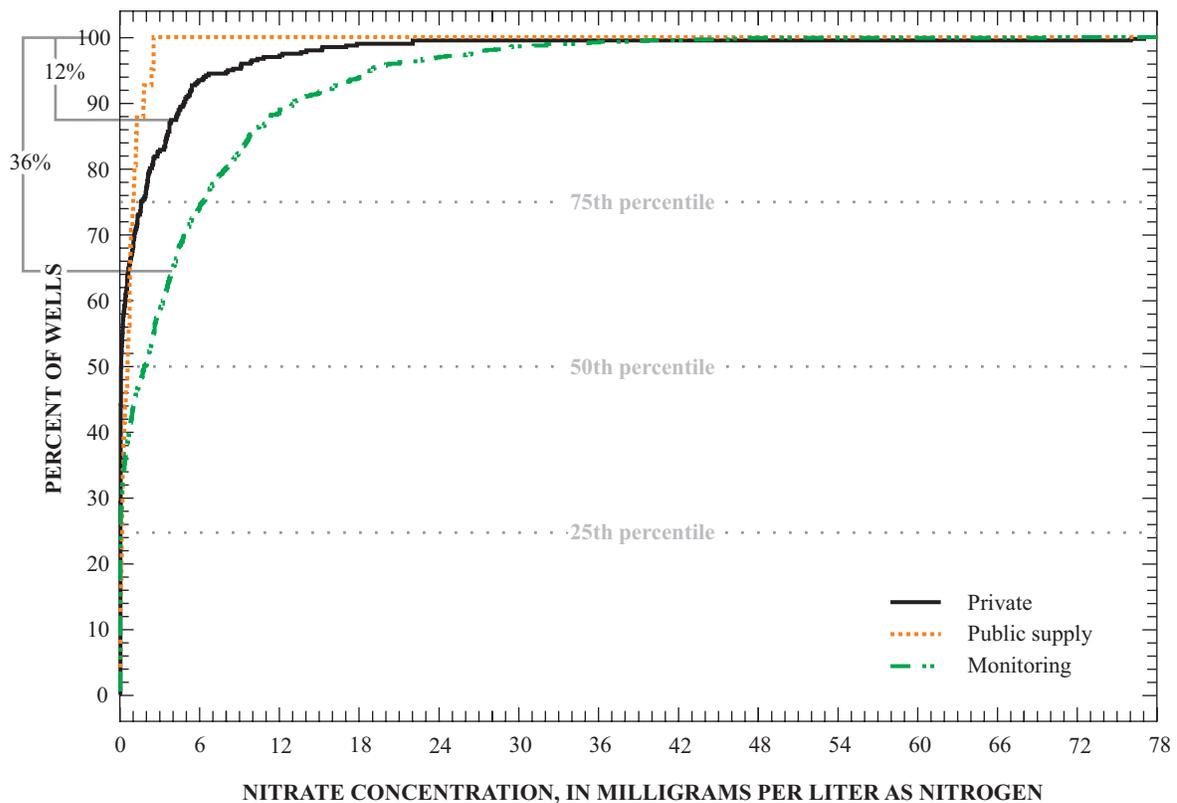


Figure 7.2-3 Survival curves were plotted to examine the difference in measured nitrate concentrations among private, public-supply, and monitoring wells in the glacial aquifer system. Nitrate concentrations exceeded 4 milligrams per liter as nitrogen in about 12 percent of private wells, 0 percent of public-supply wells, and 36 percent of monitoring wells.

8.0 Statistical Models of Nitrate Concentration

Model predictions of nitrate concentration based on nationwide water-quality data are higher than nitrate concentrations measured in private wells in the glacial aquifer system

Nitrate concentrations in water from private wells in the glacial aquifer system were compared to results of a nationwide logistic regression model that predicts the probability of occurrence of nitrate concentrations greater than 4 mg/L as N in recently recharged (shallow) groundwater in all aquifers across the United States (Nolan, 2001; Nolan and others, 2002). The nationwide model (Nolan and others, 2002) was based on 1992–95 nitrate-concentration data from shallow monitoring wells in all principal aquifers of the United States. The variables used in the nationwide logistic regression models (Nolan and Hitt, 2003; Nolan and Hitt, 2006) represent source, transport, and attenuation of nitrate (table 8–1). The nationwide models used data from all aquifers and the predictions do not fit data for private wells in the glacial aquifer system.

The fit of the nationwide model was assessed in four ways (fig. 8–1). First, the model was assessed by determining the probability of nitrate concentration greater than 4 mg/L as N for each private well in the glacial aquifer system, which was determined by spatially overlaying the private wells on the probability map (Nolan and others, 2002) and comparing to observed results. The results of this spatial overlay showed that the nationwide logistic regression model over-predicted the probability of nitrate contamination in the private wells of the glacial aquifer system (fig. 8–1A). It is important to note that the nationwide model was developed using data from monitoring wells that are shallower than the private wells assessed in this report (median depth to water in private wells in the glacial aquifer system is 39 ft compared to 15 ft for monitoring wells used in the nationwide model of all aquifers). Scatter plots of predicted and observed nitrate concentrations greater than 4 mg/L as N show that the nationwide model predicts a higher probability of finding nitrate concentrations greater than 4 mg/L as N than actually was found for private wells in the glacial aquifer system. The nationwide nitrate model fit monitoring wells in the glacial aquifer system more closely; however, the probability of a nitrate concentration

greater than 4 mg/L as N was still slightly over-predicted. Secondly, the logistic regression equation used to create the nationwide probability map (Nolan and others, 2002) was applied to the private wells of the glacial aquifer system to evaluate whether the fit achieved with the equation was superior to the fit achieved using the map. The equation with the coefficients used in the nationwide model did not fit the glacial aquifer system data (fig. 8–1B). Next, the nationwide model was recalibrated with the private wells in the glacial aquifer system to test the fit of the model with new coefficients. The recalibrated nationwide model was an excellent fit for the private wells when using the same categories of probability, but there were only three categories with observed data from the glacial aquifer system (fig. 8–1C). Hosmer and Lemeshow (2000) found that evaluations of fewer than six categories of probability almost always indicated a model fit. Finally, the fit of the recalibrated nationwide model was evaluated with the commonly used 10 categories of probability. The observed and predicted probabilities of nitrate concentration greater than 4 mg/L as N were moderately correlated ($R^2 = 0.56$) for the 10 categories (fig. 8–1D).

The probability predictions of the nationwide model generally are higher than the observed probability of nitrate occurrence in private wells; therefore, some refinement of the model would be needed to accurately predict the probability of nitrate occurrence in private wells. To examine why the nationwide nitrate model did not accurately predict the probability of nitrate occurrence in private wells in the glacial aquifer system, it was hypothesized that private wells are generally deeper than monitoring wells, have different positions within the aquifer flow system, and are constructed differently, affecting the ease with which nitrate from the surface can reach the groundwater. The deeper private wells are more likely to be screened beneath clay layers that partially confine the aquifer, increasing the likelihood of reducing redox conditions and decreasing the susceptibility of the aquifer to nitrate contamination.

Table 8–1. Variables used in selected models to predict probability of nitrate concentration exceeding a given level are compared. The logistic regression models included some or all of the variables in the linear regression model highlighted in this report, plus an interaction term for models with nitrate thresholds of 5 mg/L or greater.

[GWAVA-DW, model titled Ground WATER Vulnerability Assessment of Drinking Water; GWAVA-S, model titled Ground WATER Vulnerability Assessment of Shallow, recently recharged water; kg/ha, kilograms per hectare; %, percent; km², square kilometers; km²/cm, square kilometers per centimeter; m, meters; ML/d, Megaliters per day]

Method, variable type, and variables used in reference	Reference		
	Nolan and Hitt, 2006		
	Nolan and others, 2002	GWAVA-DW	GWAVA-S
Regression method	Logistic	Nonlinear	Nonlinear.
Variable values used in regression	For each well	Average of each well network	Average of each well network.
Types of wells used in regression	Monitoring wells in all aquifers	Private and public- supply wells in all aquifers	Monitoring wells and selected private wells in all aquifers.
Source variables	1992 nitrogen fertilizer applied to farms, in kg/ha	Farm fertilizer, in kg/ha	Farm fertilizer, in kg/ha.
	Agricultural land, in %	Manure from confined animal feeding operations, in kg/ha	Manure from confined animal feeding operations, in kg/ha.
	1990 population density, in people/km ²	Orchards/vineyards, in %	Orchards/vineyards, in %.
		Population density, in people/km ²	Population density in people/km ² . Crop/pasture/fallow, in %
Transport variables	Well-drained soils, in %	Water input (irrigated land/precipitation), in km ² /cm	Water input (irrigated land/precipitation), in km ² /cm.
	Depth to seasonally high water table, in m ¹	Semiconsolidated sand aquifer, presence/absence	Carbonate rock, presence/absence.
	Unconsolidated sand and gravel aquifers, presence/absence	Sandstone/carbonate aquifer, presence/absence	Basaltic/volcanic rock aquifer, presence/absence.
		Glacial till, presence/absence	Glacial till, presence/absence.
		Drainage ditch, in km ²	Drainage ditch, in km ² .
	Hortonian overland flow, in % of streamflow	Slope, in %.	
	Clay sediment, in %.		
Attenuation variables		Fresh surface water withdrawals, in ML/d	Fresh surface water withdrawals, in ML/d.
		Irrigation tailwater recovery, in km ²	Irrigation tailwater recovery, in km ² .
		Dunne overland flow, in % of streamflow	Histosol soil type, in %.
		Well depth, in m	Wetlands, in %.

¹. This variable is equivalent to the depth to saturated soil variable that was used in the logistic and linear regression models for this report

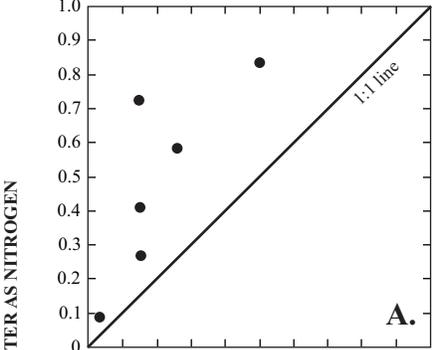
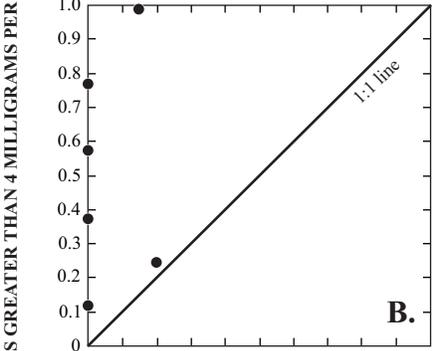
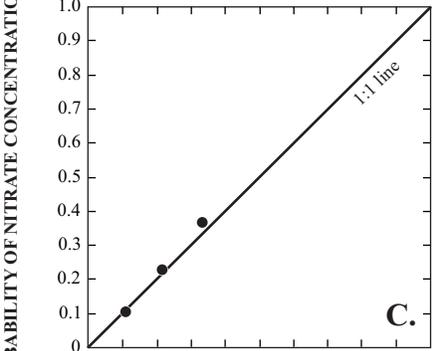
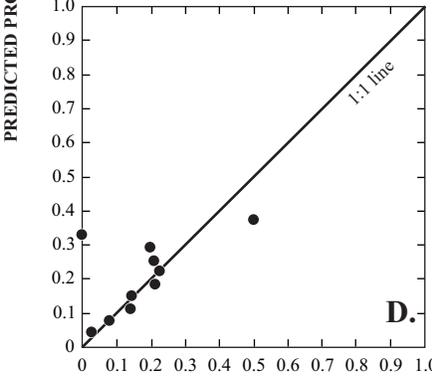
Observed probabilities plotted against model predicted probabilities	R-squared value	Method	Interpretation	Related table
 <p>A.</p>	0.79	Spatial overlay with national model shown with six probability categories	Scatter plots of predicted and observed nitrate concentrations greater than 4 milligrams per liter as nitrogen showed that the nationwide model was predicting a higher probability of finding nitrate concentrations greater than 4 milligrams per liter as nitrogen than actually were measured in private wells of the glacial aquifer system.	Table A2
 <p>B.</p>	0.12	National model equation shown with six probability categories	The logistic regression equation with the coefficients used in the nationwide model did not fit the glacial aquifer system data.	Table A3
 <p>C.</p>	1.00	Recalibrated national model shown with six probability categories. Three categories were not represented by data and did not plot on the graph.	The recalibrated nationwide model was an excellent fit for the private wells when using the same categories of probability, but there were only three categories with observed data. Evaluations of fewer than six categories of probability almost always indicate a good model fit (Hosmer and Lemeshow, 2000).	Table A4
 <p>D.</p>	0.56	Recalibrated national model shown with 10 standard probability categories	The fit of the recalibrated nationwide model was evaluated with the more commonly used 10 categories of probability. The observed and predicted probabilities of nitrate greater than 4 milligrams per liter as nitrogen were moderately correlated for the 10 categories.	Table A4

Figure 8-1. Scatterplots showing the accuracy of the nationwide nitrate model (Nolan and others, 2002) for predicting the probability of nitrate concentration greater than 4 milligrams per liter as nitrogen in private wells in the glacial aquifer system.

9.0 Probability of Nitrate Occurrence Above Threshold Concentrations

The variables that are important predictors of the probability of nitrate concentration above a selected threshold depend on the actual nitrate-concentration threshold chosen.

A series of 10 logistic regression models was developed to estimate the probability of nitrate concentration above various thresholds in groundwater from private wells in the glacial aquifer system across the United States (except Alaska, for which no spatial data were available for use in models). Thresholds of nitrate concentration ranged from 1 to 10 mg/L as N. Forty-six variables were assessed for correlation with nitrate concentration. Eighteen variables that were significantly correlated ($p < 0.05$) with nitrate concentrations in water from private wells in the glacial aquifer system were used in an initial forward-selection process to identify those variables that were important for determining the probability that the nitrate concentration would be above a given threshold value. The variables used in the final 10 models (table 9–1) include various combinations of variables describing soil properties (depth to soil saturation, percent silt in soil, and percent alfisols soil type), variables describing well-construction characteristics (diameter of the open interval, depth to top of

open interval, and length of open interval), and a source variable (10-year mean nitrogen applied to farms). Several of the models include an interaction term between depth to top of open interval and casing diameter. Not all variables were used to model all thresholds.

The 10-year mean (1992–2001) of farm-applied nitrogen fertilizer was included as a potential predictor variable for an indication of nitrate source. This variable was significantly correlated (0.37 rho and $p < 0.0001$) to nitrate concentrations in groundwater above background, although the correlation was not significant for all samples. It was included as a potential predictor variable because farm-applied nitrogen fertilizer is known to be a potential source of nitrate in water and has been used successfully as a predictor variable (4-year averages) in other nitrate models (Nolan and others, 2002; Nolan and Hitt, 2006). There are undoubtedly other important source variables for which no data are available.

Table 9–1. Variables used in logistic regression models for 10 thresholds of nitrate concentrations include source, transport, and attenuation variables. The intercept is a statistical variable that was important in all models for thresholds of 1 to 10 mg/L.

[mg/L, milligrams per liter; N, nitrogen; >, greater than; kg/yr, kilograms per year; — variable not used in model; S, significant variable for logistic regression model at a particular threshold nitrate concentration; logistic regression models used variable values for each private well in the glacial aquifer system]

Variable type	Variable description	Variable abbreviation (see appendix A, table A1)	Nitrate concentration, in mg/L as N										
			> 10	> 9	> 8	> 7	> 6	> 5	> 4	> 3	> 2	> 1	
Source variables	10-year mean (1992–2001) nitrogen applied to farms, in kg/yr	NFARM_10YR	S	S	S	S	S	S	S	S	—	—	—
Transport variables	Depth to top of open interval, in inches	DEPTH_TOI_	S	S	S	S	S	S	S	S	S	S	S
	Open interval Diameter, in inches	OI_CASE_DI	S	S	S	S	S	S	S	—	—	—	—
	Length of open interval, in inches	OILENGTH_I	—	—	—	—	—	—	—	—	S	S	S
	Silt in soil, in percent	SILTAVE	—	—	—	—	—	S	S	S	S	S	S
	Depth to saturated soil, in feet	WTDEPAVE	—	—	—	—	S	S	S	S	S	S	S
Attenuation variables	Alfisols, in percent	ORDA	—	—	—	—	S	S	S	S	S	S	S
	Interaction term between depth to top and diameter of open interval	OI_CASW_D*DEPTH_TOI	S	S	S	S	S	S	—	—	—	—	—

An interaction between open-interval diameter and depth to top of open interval was identified as a possible predictor variable for nitrate. Nitrate concentration increased as well diameter increased to a depth of approximately 71 ft; at greater depths, nitrate concentration decreased as well diameter increased (fig. 9–1). The relation of nitrate concentration in water from private wells to open-interval diameter depends on the depth to the top of the open interval in the well. Specifically, trends in nitrate concentration are different in wells shallower than 71 ft than in deeper wells (fig. 9–1). This interaction is reasonable because nitrate concentration was greater in large open-interval diameter wells which were not deeper than 71 ft. The open-interval diameter and depth to top of open interval interact to partly predict nitrate concentrations.

The final 10 logistic regression models were evaluated for goodness of fit using the ROC and the Hosmer-Lemeshow statistics (app. 8). All 10 of the final logistic regression models developed for this study had ROC statistics between 0.81 and 0.95, which indicates the models fit the data well. The probabilities corresponding to the Hosmer-Lemeshow statistics for all of the final logistic regression models for nitrate in private wells developed in this study were above the alpha level of 0.05, indicating that each of the models fit the independent variables adequately. The best fit was for the probability of nitrate concentration above 10 mg/L as N. Evaluations of collinearity showed that the variables used in the final logistic regression models were not collinear and therefore were satisfactory variables to use in the models.

The 10 models were developed for different nitrate thresholds (from 1 to 10 mg/L as N) using the most relevant variables for predicting occurrence of nitrate above selected concentrations in private wells in the glacial aquifer system. The set of significant variables selected for predicting the probability of nitrate concentration exceeding a threshold depends on the threshold that is modeled, as stated earlier (fig. 9–2, table 9–1). At the lower thresholds, more variables are needed in the model to describe natural soil-texture characteristics and depth to water. The 1 mg/L as N nitrate model is an indicator of background nitrate concentration and requires mostly variables that are related to natural characteristics of the aquifer and soil. No nitrate-source term is included in this model because there should be no specific source of “unnaturally high” nitrate under “background conditions.” As the thresholds of nitrate concentration increase, generally fewer variables are included and, for the larger thresholds, the variables mostly represent well-construction characteristics and a specific nitrate source. High nitrate concentrations in the glacial aquifer system are not natural in most areas; therefore, the probability of nitrate concentrations near the MCL can be predicted by a few selected variables. No natural variables are included in the models with thresholds greater than 7 mg/L nitrate as N.

A map that displays the probability of nitrate concentration greater than a given threshold value requires generalizations because the data displayed have a spatial and continuous aspect. However, the glacial aquifer system logistic regression

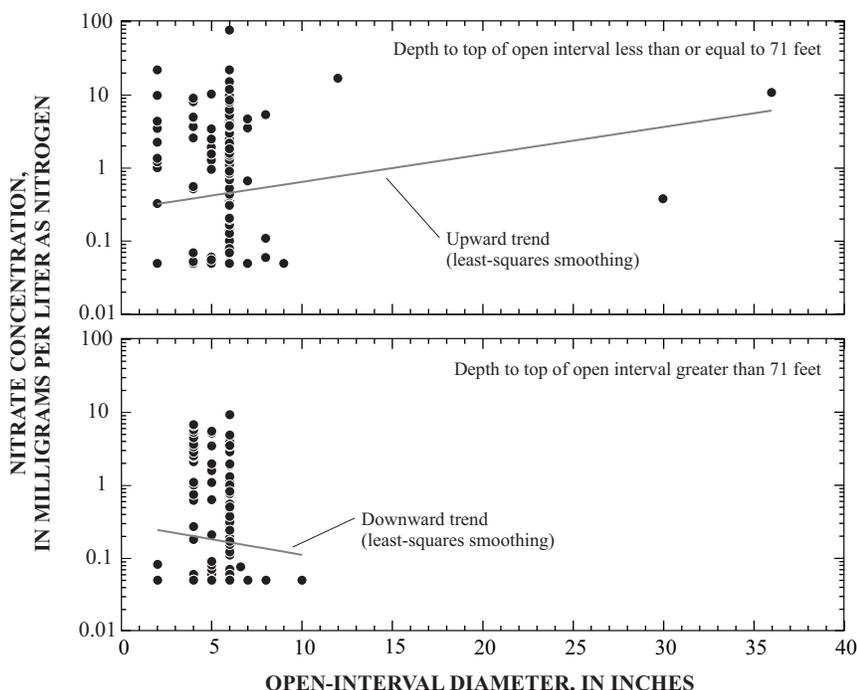


Figure 9–1. The interaction between open-interval diameter and depth to top of open interval indicates that the relation between nitrate concentration and open-interval diameter in private wells in the glacial aquifer system changes at depths of about 71 feet below land surface.

models for predicting probability of occurrence of a given nitrate concentration depend on well-construction data, which are not continuous variables (such as depth to top of open interval, open-interval length, and open-interval diameter). A generalized map displaying the probability of occurrence of nitrate concentrations greater than 1 mg/L as N was developed by setting the discrete variables (depth to top of open interval and open-interval length) to the mean of those variables in all private wells in the glacial aquifer system (91 ft and 8 ft, respectively). This map can give a regional indication of where nitrate concentrations are more or less likely to be above a background level of 1 mg/L as N (fig. 9–3), but

specific well-construction information is needed to refine the map for an individual private well.

The accuracy of the maps displaying the predicted probability of occurrence of a nitrate concentration above a given value in water from private wells decreases rapidly as the probability becomes more dependent on the well-construction variables because the maps do not include individual well-construction variables. The accuracy of the maps for thresholds above 4 mg/L as N was very poor when the appropriate well-construction variables were set to a mean for all private wells in the glacial aquifer system.

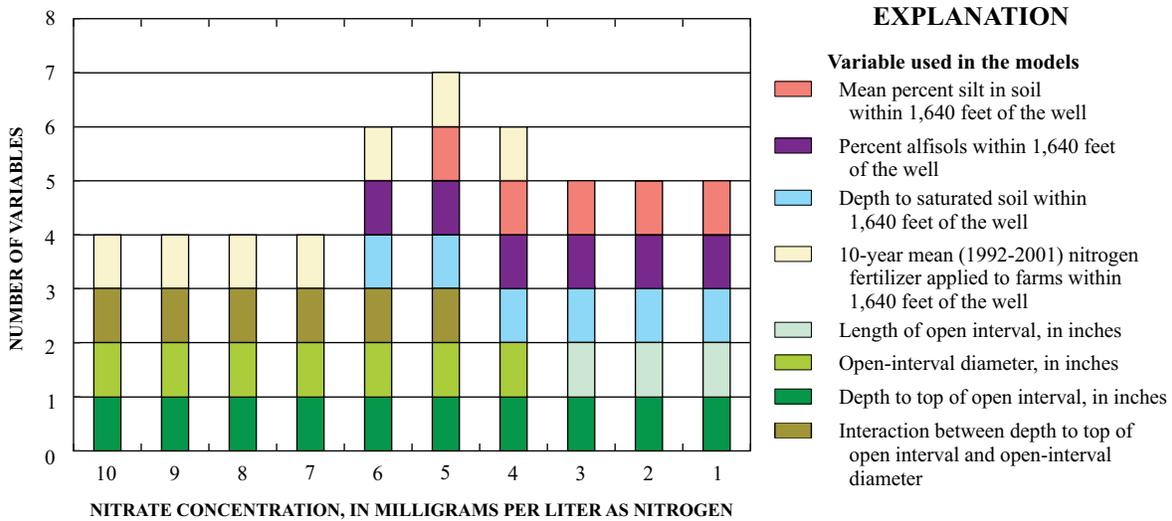
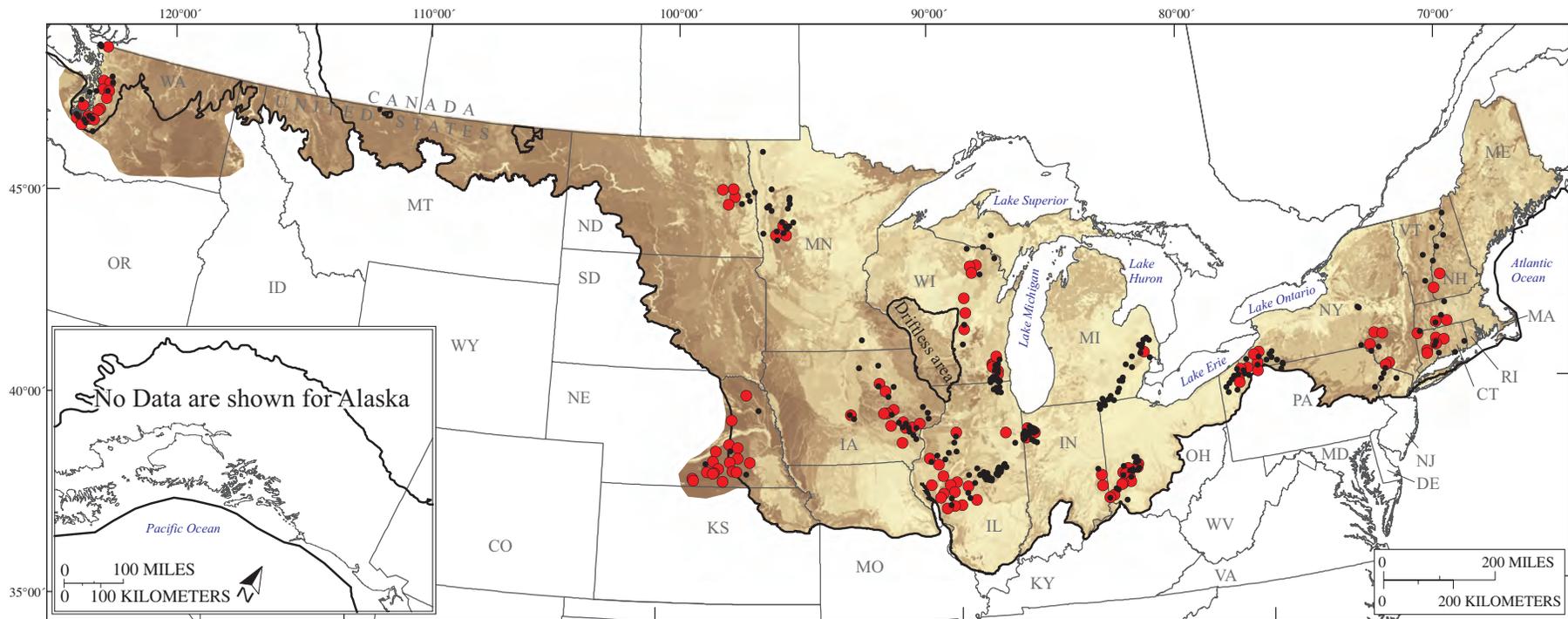


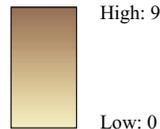
Figure 9–2. Fewer variables are needed to predict higher nitrate concentrations (7–10 milligrams per liter as nitrogen) than lower nitrate concentrations (1–6 milligrams per liter as nitrogen); at lower concentrations, most variables represent natural conditions and well-construction characteristics.



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

Predicted probability that nitrate concentration is greater than 1 milligram per liter as nitrogen, in percent



— Glacial aquifer system extent (approximate)

Measured nitrate concentration, in milligrams per liter as nitrogen

- Less than or equal to 1 (background concentration)
- Greater than 1 (some human influence)

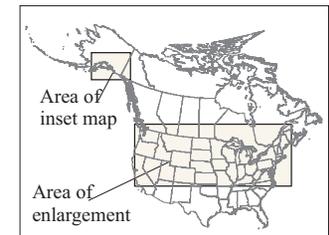


Figure 9-3. Probability of occurrence of nitrate concentrations greater than 1 milligram per liter in private wells in the glacial aquifer system. This map was developed by setting the depth to top of open interval and open-interval length to the mean of the values for these variables in all private wells in the glacial aquifer system (91 and 8 feet, respectively).

10.0 Prediction of Mean Nitrate Concentrations in Networks of Wells

Use of network averages of variables affecting nitrate concentration in linear regression models helps minimize the well-to-well differences in nitrate concentration and aids in predicting nitrate concentration in water from private wells in the glacial aquifer system at the regional scale.

A linear regression model also was developed to predict nitrate concentrations at the regional scale. In order to remove some of the variability, nitrate concentrations were averaged by well network. The model is based on network averages because nitrate concentrations were highly variable over short distances and it was otherwise not possible for a linear regression model to fit the individual data points for the glacial aquifer system. The nitrate concentrations were averaged by network under the assumption that wells in the same network would likely be related and be subject to similar influences on nitrate concentration. The linear regression was calculated using the same predictor variables used in the logistic regressions (cumulatively); however, the interaction term was omitted from the linear regression analysis. The linear regression model includes network averages of the following variables: depth to saturated soil, percent silt in soil, percent alfisols soil type, 10-year mean (1992–2001) nitrogen applied to farms, diameter of the open interval, depth to top of open interval, and length of open interval (table 10–1). Using values for each of the predictor variables averaged by network (network mean value) from the logistic regression models, the linear regression model significantly predicted the mean nitrate concentration in networks ($R^2 = 0.82$ and $p\text{-value} < 0.0001$) (fig. 10–1). Residuals resulting from the linear regression model were plotted against quantiles of the standard normal distribution and were found to be normally distributed.

The nitrate prediction model is a better fit in some geographic areas than in others. To determine the goodness of fit for the linear prediction model, the predicted mean

concentration for a network was compared to the observed mean concentration. Predicted nitrate concentrations for networks with residuals greater than zero are too low, whereas, concentrations for networks with residuals less than zero are over predicted by the model. In figure 10–2, the network results are ordered by the model residual obtained for each network. The closer the residual is to zero, the better the fit of the model. The networks that fit this model best are those in Connecticut, New Hampshire, and Vermont; in Washington; and in Illinois (fig. 10–2). The overall fit of the model is good ($R^2=0.82$).

The area of highest predicted mean nitrate concentration in private wells was in west-central Illinois (fig. 10–3). This area is characterized by a relatively high application of nitrate, older aquifer materials with fractured till, thinner glacial materials than in areas covered by Wisconsin glacial deposits, and predominantly large-diameter wells. The mean measured nitrate concentration in this west-central Illinois network was 5.9 mg/L as N. This network also had the largest mean well diameter (23.7 in.).

The linear regression model can be used to predict and compare regional nitrate concentrations in water from private wells in the glacial aquifer system. This model is designed for large regional interpretations. The map showing nitrate concentrations predicted by using this model (fig. 10–3) can be used to indicate where nitrate concentrations in water from private wells in the glacial aquifer system are likely to be elevated.

Table 10-1. Coefficient values, standard error, and p-values of variables used in the linear regression model used to predict mean nitrate concentration in private-well networks in the glacial aquifer system.

[ft, feet; kg/yr, kilograms per year]

Coefficient (see appendix 1, table A1 for a description of coefficient variables)	Coefficient value	Standard error	T-value	Probability (p-value)
Intercept	-2.184	1.420	-1.538	0.163
Open-interval diameter, in inches (mDiamOlin)	0.158	0.081	1.963	0.085
Depth to saturated soil, in ft (mwtdepave)	0.479	0.320	1.495	0.173
Silt in soil, in percent (msiltave)	0.054	0.029	1.878	0.097
10-year-mean (1992-2001) nitrogen applied to farms, in kg/yr (mnfarm.10y)	0.000	0.000	1.448	0.186
Length of open interval, in inches (moilength)	-0.003	0.006	-0.483	0.642
Depth to top of open interval, in inches (mdepthtoi)	-0.001	0.001	-2.083	0.071
Alfisols, in percent (morda)	-0.017	0.011	-1.519	0.167

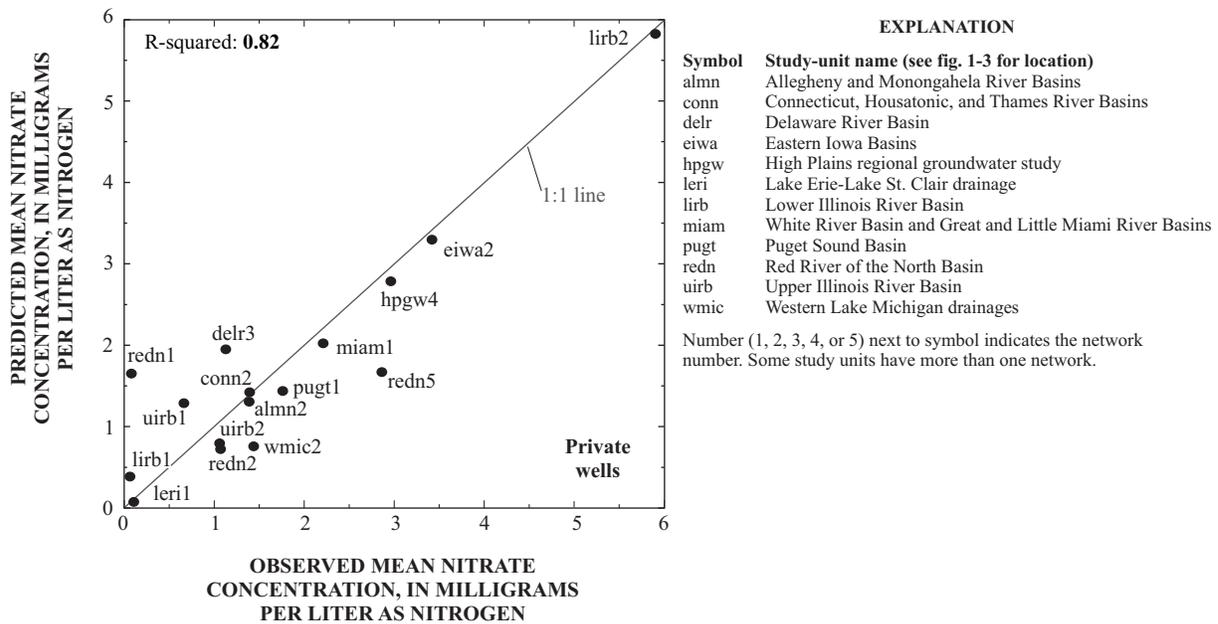


Figure 10-1. The relation between mean observed nitrate concentrations and mean nitrate concentrations predicted with the linear regression model indicates that the model is a good fit ($R^2=0.82$).

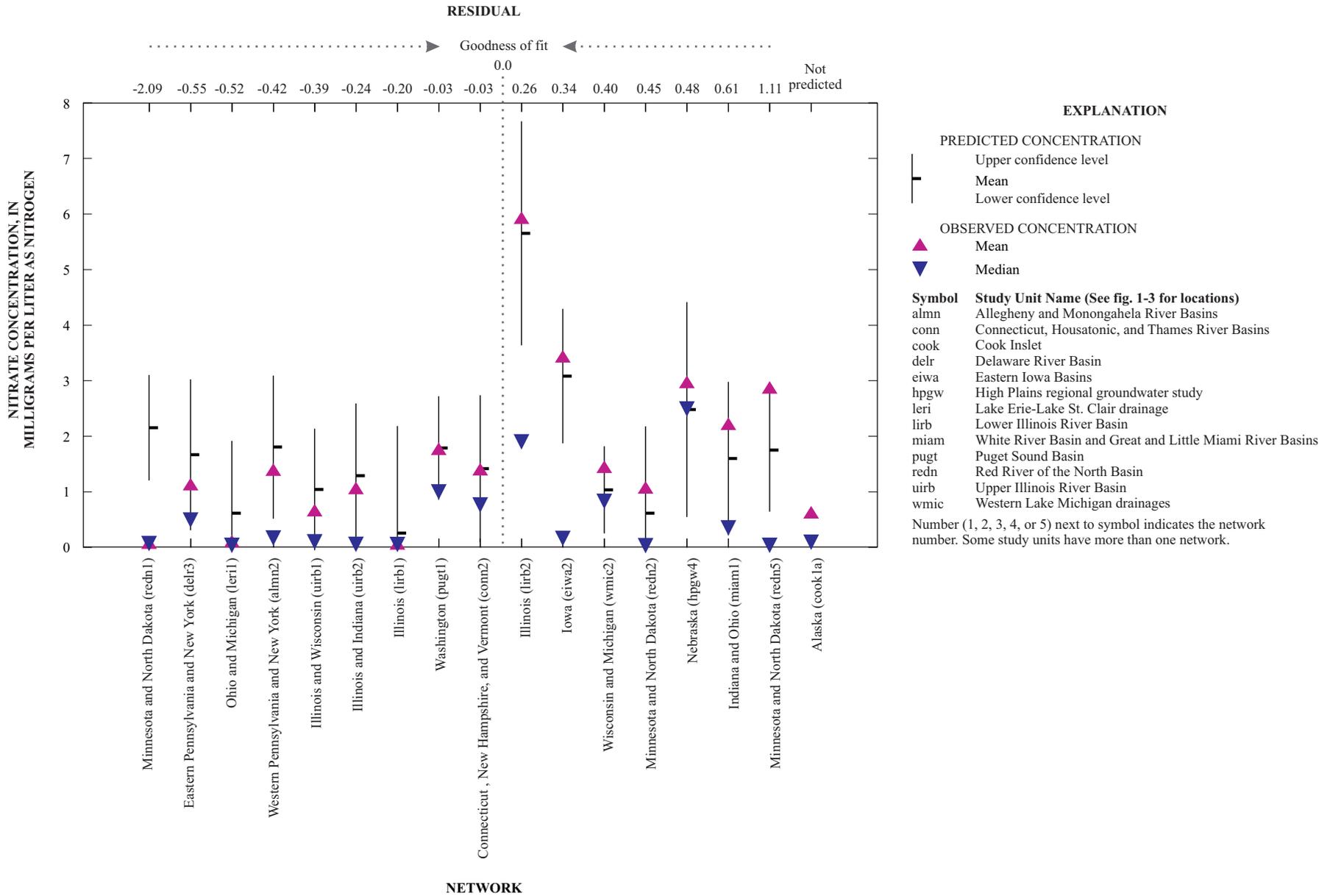
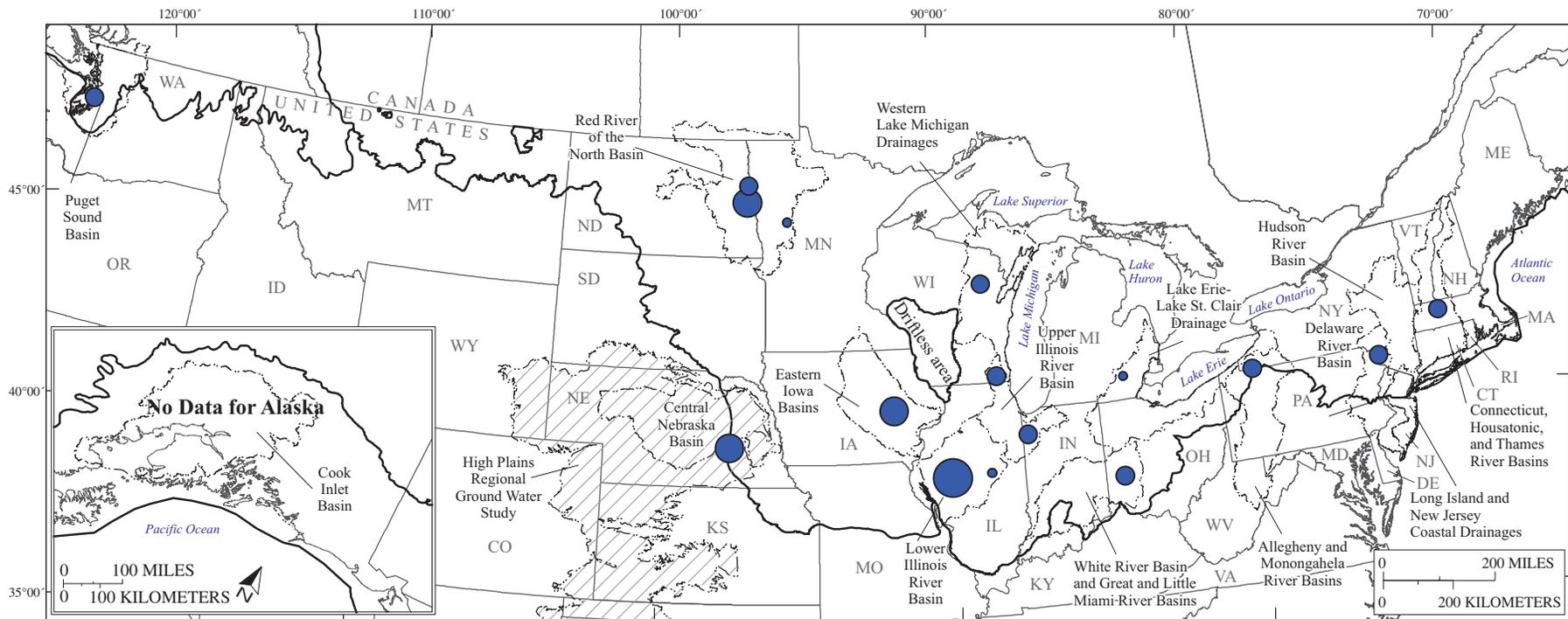


Figure 10–2. Mean observed and predicted nitrate concentrations show the goodness of fit for the linear regression model, which is based on data from these 17 well networks. The networks for which the residuals are closer to zero have the best fit to the model. Data needed for the model were unavailable 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°

EXPLANATION

- Glacial aquifer system extent (approximate)
- - - Study-unit boundary
- Predicted mean nitrate concentration for a network of wells, in milligrams per liter as nitrogen**
- 0 to 1
- Greater than 1 to 2
- Greater than 2 to 4
- Greater than 4

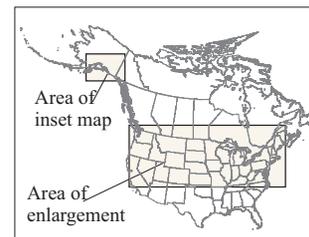


Figure 10-3. The area of highest predicted mean nitrate concentration in water from private wells in the glacial aquifer system is in the glacial deposits in west-central Illinois.

11.0 Conclusions

Most people in rural America obtain their drinking water from private wells, and the potential for nitrate contamination is one of the most pressing water-quality issues for these private-well users. Forty-one million people rely on the glacial aquifer system for drinking water. Although only 4 percent of private wells contain groundwater with nitrate at concentrations greater than 10 mg/L as N, another 26 percent of private wells are approaching a concentration of potential human-health concern—that is, they contain nitrate at concentrations greater than 1 mg/L as N. A summary of data from reference wells indicates that most occurrences of nitrate above 1 mg/L as N represent nitrate concentrations above background and likely result from anthropogenic sources. If anthropogenic nitrate is found in water from a private well, the well owner may want to check for other anthropogenic compounds, as other compounds may be transported from land surface to the well. DeSimone (2009) found that 73 percent of private wells had mixtures of two or more contaminants. Nitrate alone should not be used as the screening tool for other contaminants, however. Squillace and others (2002) showed that volatile organic compounds or pesticides were present in 43 percent of samples even though “anthropogenic” nitrate was not present. Testing private well water for nitrate and other contaminants would be needed to ensure safe drinking-water quality.

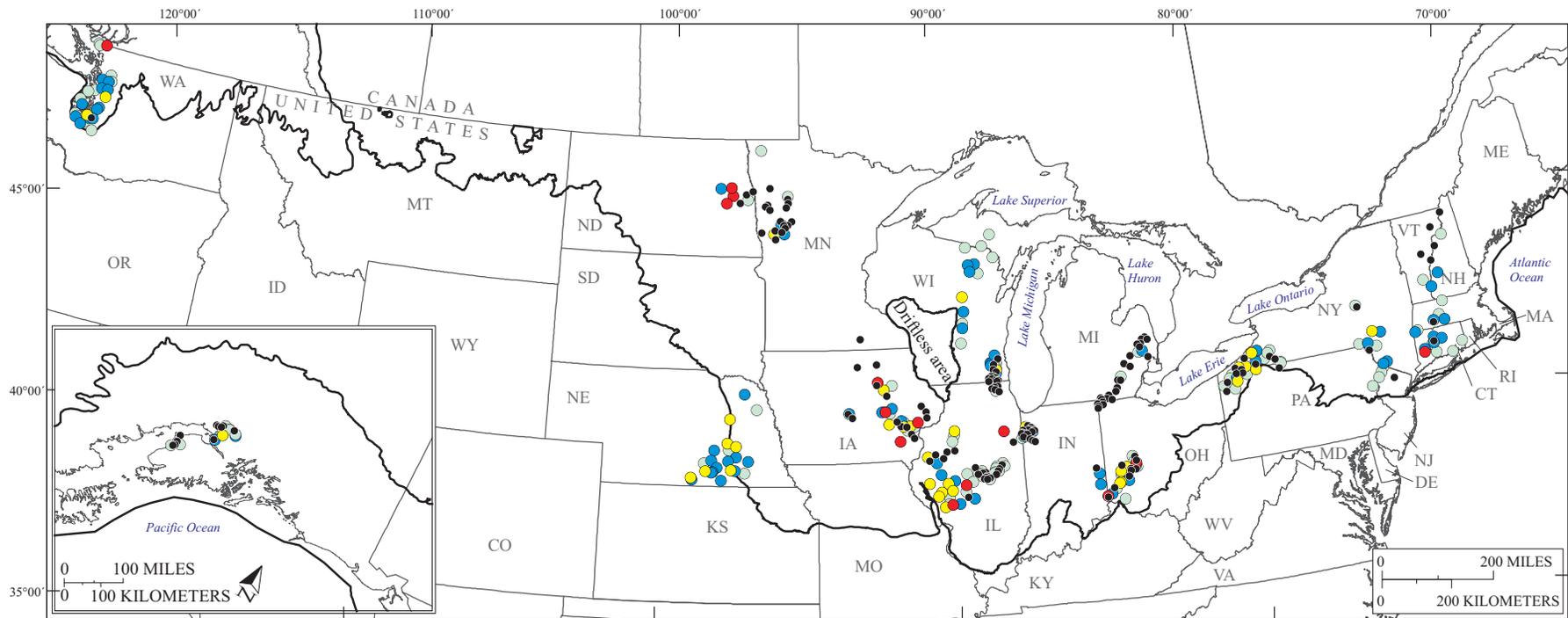
One reason that nitrate alone may not be a good predictor of other anthropogenic constituents is that nitrate is redox sensitive, but not all anthropogenic constituents are redox sensitive. When the dissolved-oxygen concentration in groundwater is high, nitrate is relatively stable and may persist. Nitrate is less likely to be present in water with low dissolved-oxygen content where it may change to nitrogen gas through denitrification. Oxygen content can decrease as well depth increases or if recharge is constrained by a confining layer, such as clay or clayey till. In areas of high clay content, wells commonly are drilled to greater depths to improve the reliability of the supply of groundwater. Understanding the geologic framework, including the heterogeneity of the aquifer material with depth, is the key to understanding the distribution of nitrate and other redox-sensitive constituents.

The concentrations of nitrate in water from private wells in the glacial aquifer system assessed in this study varied with time, location, and depth. Comparisons of age of recharge water indicates that the mean nitrate concentration was significantly greater in young groundwater than in old groundwater. Results of a recent study (Saad, 2008) that assessed trends in the shallow glacial aquifer system in Wisconsin indicated that nitrate concentration in the shallow, recently recharged groundwater increased substantially in near-decadal sampling (1994 to 2002) as fertilizer application increased. This study was conducted in a highly permeable part of the glacial aquifer system; it is not known whether the nitrate concentration also increases significantly in private wells in deeper or less permeable parts of the aquifer system.

Nitrate concentrations in water from private wells in the glacial aquifer system generally decreased with depth. The glacial aquifer system may differ from other unconsolidated aquifers as a result of its heterogeneity and its content of fine-grained material, such as clay and buried soils overlying the highly agricultural part of the aquifer system. This heterogeneity with depth obscures a distinct spatial pattern of nitrate concentration (fig. 11–1). Fine-grained materials confine the deeper aquifer material, causing the geochemical environment to be reducing and, therefore, low in nitrate. The deeper the well, the more likely the water is to be reducing, and under confined or semi-confined conditions. The decrease in nitrate concentration with increasing well depth is more distinct than any spatial or temporal patterns (fig. 11–2) because depth can be an indicator of geochemical conditions. Local differences in sediment thickness and texture also can affect nitrate concentrations; the decreased risk of nitrate contamination with depth is strongly dependent on the local hydrogeologic conditions.

Because collection and analysis of groundwater samples everywhere in the glacial aquifer system is impractical, methods are needed for extrapolating the vulnerability of private wells to nitrate contamination from sampled to unsampled areas. The methods and data used in the probability and prediction models presented in this report may be useful for extrapolation of vulnerability for a number of anthropogenic constituents with properties similar to nitrate. The models of nitrate in private wells indicate that only a small amount of information is needed to predict the regional vulnerability to nitrate contamination; the same may be true for other surface-derived contaminants as well. Although variables such as density of septic systems (as percent of housing units on septic systems within 1,640 ft of the well) are correlated with nitrate concentration in water from individual wells, they cannot be used to predict nitrate concentration in the glacial aquifer system on a regional scale, but may be useful to a local well owner. Additional information, such as soil type and aquifer material, also is important for predicting nitrate concentration. Similarly, well characteristics such as depth to the top of the well screen, the length of the screen, and the diameter of the open interval can be used to predict nitrate concentration because these variables indicate what part of the flow system is utilized. A source variable such as the rate of nitrogen applied to farms was useful in predicting regional nitrate concentration.

It is apparent in the various model trials that altering the set of variables alters the outcome of the models; however, more or less equivalent models can be developed that fit the dependent variable using entirely different sets of independent predictor variables. If, in future modeling efforts, values of the exact variables that were used in these models are not available, alternate variables that describe similar characteristics may be substituted to obtain a result that is quantitatively similar to the results presented in this report.



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

- Glacial aquifer system extent (approximate)
- Nitrate concentration, in milligrams per liter as nitrogen**
 - Less than or equal to 0.05 (not detected)
 - Greater than 0.05 to 1.0 (background level)
 - Greater than 1 to 4 (some impact from humans)
 - Greater than 4 to 10 (high concentration)
 - Greater than 10 (above the U.S. Environmental Protection Agency Maximum Contaminant Level)

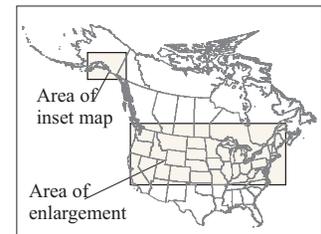


Figure 11-1. There was no distinct geographic pattern in nitrate concentration in water from private wells in the glacial aquifer system, although most high concentrations (greater than 10 milligrams per liter as nitrogen) were in the central part of the glaciated United States.

Because the glacial aquifer system covers a large area, aggregating network information helps in predicting nitrate concentrations. The prediction model based on network means indicates that nitrate in water from private wells is less likely to be a problem in the two coastal framework areas than in parts of the Central and West-central glacial framework areas. In these areas, continued monitoring of nitrate concentrations and study of the distribution of glacial materials would help in understanding the variations in nitrate concentration in the glacial aquifer system.

Some characteristics of the models of nitrate probability and prediction must be taken into consideration when applying the model results for making management decisions and water-quality assessments. Important among these considerations are that the models (1) focus on spatial rather than temporal detail; (2) integrate data over large networks to calculate spatial averages for model input; (3) include only the water-quality variables that are electronically available and statistically correlated with nitrate concentration; and (4) favor

water-quality comparisons across broad regions as opposed to local systems. Some regional characteristics that may be critical in understanding water-quality conditions are not well documented for the entire multi-state region. Examples include tile-drain location and distribution, land-management practices, and septic-system locations. The effects of such characteristics were evaluated by extension through available variables or explored indirectly by evaluating model fit. Maps developed by applying these models may be improved over time as new geographic information becomes available or as spatial resolution improves. Specifically, the models of nitrate concentration in private wells can be used to (1) establish links between water quality and constituent sources; (2) assess the natural processes that attenuate nitrate; and (3) predict where changes in water quality may result from changes in environmental management or well construction. Finally, the more information available about well construction and nearby nitrogen sources, the more likely it is that the vulnerability of a given private well to nitrate contamination can be determined.

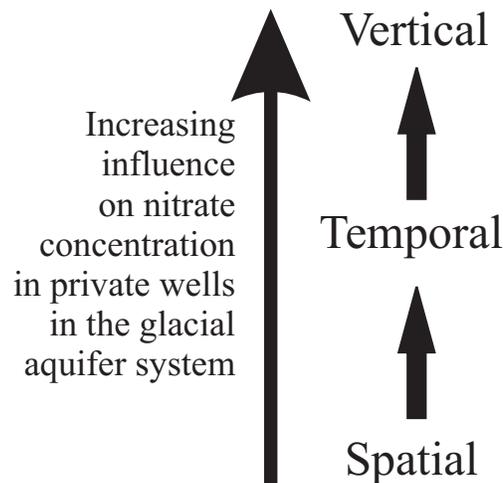


Figure 11–2. Changes in nitrate concentrations in water from private wells in the highly heterogeneous glacial aquifer system were most variable with depth.

References Cited

- Ahlbom, Anders, and Alfredsson, Lars, 2005, Interaction—A word with two meanings creates confusion: *European Journal of Epidemiology*, v. 20, p. 563–564.
- Alley, W.M., ed., 1993, *Regional ground-water quality*: New York, Van Nostrand Reinhold, 634 p.
- American Society of Civil Engineers, Committee on Ground Water, 1961, *Ground water basin management*: ASCE Manuals of Engineering, no. 40, 160 p.
- 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. plus app.
- Bartholomay, R.C., Carter, J.M., Qi, S.L., Squillace, P.J., and Rowe, G.L., 2007, Summary of selected U.S. Geological Survey data on domestic well water quality for the Centers for Disease Control’s National Environmental Public Health Tracking Program: U.S. Geological Survey Scientific Investigations Report 2007–5213, 57 p.
- Bugliosi, E.F., 1999, The Midwestern Basins and Arches Regional Aquifer System in parts of Indiana, Ohio, Michigan, and Illinois—Summary: U.S. Geological Survey Professional Paper 1423–A, 46 p.
- Clark, Ian, and Fritz, Peter, 1997, *Environmental isotopes in hydrogeology*: Boca Raton, Fla., Lewis Publishers, 328 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, at http://water.usgs.gov/lookup/getspatial?ofr99-77_geol75m.
- Debrewer, L.M., Ator, S.W., and Denver, J.M., 2007, Factors affecting spatial and temporal variability in nutrient and pesticide concentrations in the surficial aquifer on the Delmarva Peninsula: U.S. Geological Survey Scientific Investigations Report 2005–5257, 44 p.
- DeSimone, L.A., Hamilton, P.A., and Gilliom, R.J., 2009, Quality of water from domestic wells in principal aquifers of the United States, 1991–2004—Overview of major findings: U.S. Geological Survey Circular 1332, 48 p.
- Eberts, S.M., and George, L.L., 2000, Regional ground-water flow and geochemistry in the Midwestern Basins and Arches Aquifer System in parts of Indiana, Ohio, Michigan, and Illinois: U.S. Geological Survey Professional Paper 1423–C, p. C1–C103.
- Eberts, S., Bohlke, J.K., and Kauffman, L., 2006, Application of ground-water dating techniques for evaluating the susceptibility of aquifers and public-supply wells to contamination (abs.): Proceedings of the 2006 National Monitoring Conference, San Jose, CA, May 7–11, 2006, http://oh.water.usgs.gov/tanc/pubs/Eberts_NWQMC_Abstract.pdf.
- Elliott, E.M., Kendall, C., Wankel, S.D., Burns, S.A., Boyer, E.W., Harlin, K., Bain, S.J., and Butler, T.J., 2007, Nitrogen isotopes as indicators of NO_x source contributions to atmospheric nitrate deposition across the Midwestern and Northeastern United States: *Environmental Science and Technology*, v. 41, p. 7661–7667.
- Embrey, S.S., and Runkle, D.L., 2006, Microbial quality of the Nation’s ground-water resources, 1993–2004: U.S. Geological Survey Scientific Investigations Report 2006–5290, 34 p.
- Environmental Systems Research Institute, Inc., 2008, ArcGIS v. 9.3 desktop help: Environmental Systems Research Institute, Inc., Redlands, Calif., digital media.
- Fishman, M.J., ed., 1993, *Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determination of inorganic and organic constituents in water and fluvial sediments*: U.S. Geological Survey Open-File Report 93–125, 217 p.
- Follett, R.F., 2001, Nitrogen transformation and transport processes, in Follett, R.F., and Hatfield, J.L., eds., *Nitrogen in the environment—Sources, problems and management*: Amsterdam, The Netherlands, Elsevier Science B.V., p. 17–44.
- Follett, R.F., and Hatfield, J.L., eds., 2001, *Nitrogen in the environment—Sources, problems, and management*: Amsterdam, Netherlands, Elsevier B.V., 520 p.
- Freeze, R.A. and Cherry, J.A., 1979, *Groundwater*: Englewood Cliffs, New Jersey, Prentice-hall, Inc., 604 p.
- Fullerton, D.S., Bush, C.A., and 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, at <http://pubs.usgs.gov/imap/i-2789/>.

- Garabedian, S.P., Coles, J.F., Grady, S.J., Trench, E.C.T., and Zimmerman, M.J., 1998, Water quality in the Connecticut, Housatonic, and Thames River Basins, Connecticut, Massachusetts, New Hampshire, New York, and Vermont, 1992–95: U.S. Geological Survey Circular 1155, 32 p.
- Gebert, W.A., Graczyk, D.J., and Krug, W.R., 1987, Average annual runoff in the United States, 1951–80, 1:5,000,000 scale digital data from <http://water.usgs.gov/GIS/metadata/usgswrd/XML/runoff.xml>, last accessed December 21, 2009.
- Gilliom, R. J., Alley, W. M., and Gurtz, M. E., 1995, Design of the National Water-Quality Assessment Program; Occurrence and distribution of water-quality conditions: U.S. Geological Survey Circular 1112, 33 p.
- Helsel, D.R., 2005, Nondetects and data analysis: Hoboken, N.J., John Wiley & Sons, Inc., 250 p.
- Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: New York, Elsevier Science Publishing Company, 522 p.
- Hitt, K.J., 2003, 2000 population density by block group for the conterminous United States: U.S. Geological Survey digital data, accessed July 9, 2009, at <http://water.usgs.gov/lookup/getspatial?uspopd00x10g>.
- Hosmer, D.W., and Lemeshow, Stanley, 2000, Applied logistic regression (2d ed.): Hoboken, N.J., John Wiley and Sons, Inc., 375 p.
- Johnston, C.T., Cook, P.G., Frape, S.K., Plummer, L.N., Busenberg, E., and Blackport, R.J., 1998, Ground water age and nitrate distribution within a glacial aquifer beneath a thick unsaturated zone: *Ground Water*, v. 36, no. 1, p. 171–180.
- Jones, M.S., 1999, Geologic framework for the Puget Sound Aquifer System, Washington and British Columbia: U.S. Geological Survey Professional Paper 1424-C, p. C1–31.
- Kleinbaum, D.G., Kupper, L.L., and Muller, K.E., 1988, Applied regression analysis and other multivariate methods: Boston, Mass., PWS-Kent Publishing Company, 718 p.
- Kontis, A.L., Randall, A.D., and Mazzaferro, D.L., 2004, Regional hydrology and simulation of flow in stratified-drift aquifers in the glaciated Northeastern United States: U.S. Geological Survey Professional Paper 1415-C, 156 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., and 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.
- Lake, L.R., Lovett, A.A., Hiscock, K.M., Betson, Mark, Foley, Aidan, Sunnenberg, Gisela, Evers, Sarah, and Fletcher, Steve, 2003, Evaluating factors influencing groundwater vulnerability to nitrate pollution—Developing the potential of GIS: *Journal of Environmental Management*, v. 66, no. 3, p. 315–328.
- Lapham, W.W., 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.
- Lapham, W.W., 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 of wells, and collection of related data: U. S. Geological Survey Open-File Report 95–398, 69 p.
- Lewis, M.E., 2006, Dissolved oxygen: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A6., section 6.2, accessed July 13, 2009, at <http://pubs.water.usgs.gov/twri9A6/>.
- Lloyd, O.B., Jr., and Lyke, W.L., 1995, Ground water atlas of the United States—Segment 10, Illinois, Indiana, Kentucky, Ohio, Tennessee: U.S. Geological Survey Hydrologic Investigations Atlas 730-K, 30 p.
- Lorenz, D.L., Goldstein, R.M., Cowdery, T.K., and Stoner, J.D., 2003, Comparison of two methods for delineating land use near monitoring wells used for assessing quality of shallow ground water: U.S. Geological Survey Water-Resources Investigations Report 03–4067, 13 p.
- Madison, R.J., and Brunett, J.O., 1985, Overview of the occurrence of nitrate in ground water of the United States, *in* National Water Summary 1984: U.S. Geological Survey Water-Supply Paper 2275, p. 93–105.
- Mallows, C.L., 1973, Some comments on Cp: *Technometrics*, v. 15, p. 661.

- McKenna, D.P., and Keefer, D.A., 1991, Potential for agricultural chemical contamination of aquifers in Illinois: Illinois State Geological Survey Open-File Series 1991-7R, 16 p.
- McMahon, P.B., and Bohlke, J.K., 2005, Variability of nitrate sources in a regional aquifer—Role of soil processes and land use: *Geochimica et Cosmochimica Acta*, v. 69, no. 10, A604 Suppl. S.
- McMahon, P.B., and Chapelle, F.H., 2008, Redox processes and water quality of selected principal aquifer systems: *Ground Water*, v. 46, no. 2, p. 259–271.
- McMahon, P.B., Bohlke, J.K., Kauffman, L.J., Kipp, K.L., Landon, M.K., Crandall, C.A., Burow, K.R., and Brown, C.J., 2008, Source and transport controls on the movement of nitrate to public supply wells in selected principal aquifers of the United States: *Water Resources Research*, v. 44, W04401.
- Miller, J.A., 1999, Ground water atlas of the United States—Introduction and National summary: U.S. Geological Survey Hydrologic Atlas 730-A, 36 p.
- Miller, J.A., and Appel, C.L., 1997, Ground water atlas of the United States—Segment 3, Kansas, Missouri, Nebraska: U.S. Geological Survey Hydrologic Atlas 730-D, 24 p.
- Mueller, D.K., and Helsel, D.R., 1996, Nutrients in the Nation's waters—Too much of a good thing? U.S. Geological Survey Circular 1136, 24 p.
- Mueller, D.K., and Titus, C.J., 2005, Quality of nutrient data from streams and ground water sampled during water years 1992–2001: U.S. Geological Survey Scientific Investigations Report 2005-5106, 27 p.
- Nakagaki, N., Price, C.V., Falcone, J.A., Hitt, K.J., and Ruddy, B.C., 2007, Enhanced National Land Cover Data 1992 (NLCDe 92): U.S. Geological Survey digital data accessed July 9, 2009, at <http://water.usgs.gov/GIS/metadata/usgswrd/XML/nlcde92.xml>.
- National Park Service, 1999, State surficial geology map of Alaska: National Park Service Alaska Support Office, digital data, 1:1,584,000 scale.
- Natural Resources Conservation Service, 2008, Celebrating the next generation: Illinois cooperative soil survey—Soils of Illinois: accessed December 29, 2008, at <http://www.il.nrcs.usda.gov/technical/soils/soilsofil.html>.
- Nolan, B.T., 2001, Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States: *Ground Water*, v. 39, no. 2, p. 290–299.
- Nolan, B.T., and Hitt, K.J., 2003, Nutrients in shallow ground waters beneath relatively undeveloped areas in the conterminous United States: U.S. Geological Survey Water-Resources Investigations Report 02-4289, 17 p.
- Nolan, B.T., and Hitt, K.J., 2006, Vulnerability of shallow groundwater and drinking-water wells to nitrate in the United States: *Environmental Science and Technology*, v. 40, no. 24, p. 7834–7840.
- Nolan, B.T., Hitt, K.J., and Ruddy, B.C., 2002, Probability of nitrate contamination of recently recharged ground waters in the conterminous United States: *Environmental Science and Technology*, v. 36, no. 10, p. 2138–2145.
- Nolan, B.T., and Stoner, J.D., 2000, Nutrients in ground waters of the conterminous United States, 1992–95: *Environmental Science and Technology*, v. 34, no. 7, p. 1156–1165.
- Panno, S.V., Kelly, W.R., Martinsek, A.T., and Hackley, K.C., 2006, Estimating background and threshold nitrate concentration using probability graphs: *Groundwater*, v. 44, no. 5, p. 697–709.
- Patton, C.J., and Kryskalla, J.R., 2003, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Evaluation of alkaline persulfate digestion as an alternative to Kjeldahl digestion for determination of total and dissolved nitrogen and phosphorus in water: U.S. Geological Survey Water-Resources Investigations Report 03-4174, 33 p.
- Puckett, L.J., and Cowdery, T.K., 2002, Transport and fate of nitrate in a glacial outwash aquifer in relation to ground water age, land use practices, and redox processes: *Journal of Environmental Quality*, v. 31, p. 782–796.
- Puckett, L.J., Cowdery, T.K., McMahon, P.B., Tornes, L.H., and Stoner, J.D., 2002, Using chemical, hydrologic, and age dating analysis to delineate redox processes and flow paths in the riparian zone of a glacial outwash aquifer-stream system: *Water Resources Research*, v. 38, no. 8, p. 9–1 to 9–20.
- Randall, A.D., 2001, Hydrogeologic framework of stratified-drift aquifers in the glaciated northeastern United States: U.S. Geological Survey Professional Paper 1415-B, 179 p.
- Ruddy, B.C., Lorenz, D.L., and Mueller, D.K., 2006, County-level estimates of nutrient inputs to the land surface of the conterminous United States, 1982–2001: U.S. Geological Survey Scientific Investigations Report 2006-5012, 17 p., also online at <http://pubs.usgs.gov/sir/2006/5012/>.
- Rupert, M.G., 2008, Decadal-scale changes of nitrate in ground water of the United States, 1998–2004: *Journal of Environmental Quality*, v. 37, p. S-240 to S-248.

- Saad, D.A., 2008, Agriculture-related trends in groundwater quality of the glacial deposits aquifer, central Wisconsin: *Journal of Environmental Quality*, v. 37, p. S-209-S-225.
- SAS Institute Inc., 2003, SAS version 9.1.3 for Windows: Cary, N.C., SAS Institute Inc., computer software accessed May 20, 2008 at <http://www.sas.com/>.
- Sebol, L.A., Roberston, W.D., Busenberg, E., Plummer, L.N., Ryan, M.C., and Schiff, S.L., 2007, Evidence of CFC degradation in groundwater under pyrite-oxidizing conditions: *Journal of Hydrology*, v. 347, no. 1-2, p. 1-12.
- 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.
- Squillace, Paul J., Scott, Jonathon, Moran, Michael J., Nolan, B.T., and Kolpin, Dana W., 2002, VOCs, pesticides, nitrate, and their mixtures in groundwater used for drinking water in the United States: *Environmental Science & Technology*, v. 36, no. 9, p. 1923-1930.
- Stanton, J.S., Steele, G.V., and Vogel, J.R., 2007, Occurrence of agricultural chemicals in shallow ground water and the unsaturated zone, northeast Nebraska glacial till, 2002-2004: U.S. Geological Survey Scientific Investigations Report 2007-5228, 51 p.
- Tesoriero, A.J., Saad, D.A., Burow, K.R., Frick, E.A., Puckett, L.J., and Barbash, J.E., 2007, Linking ground-water age and chemistry data along flow paths—Implications for trends and transformations of nitrate and pesticides: *Journal of Contaminant Hydrology*, v. 94, p. 139-155.
- Tomer, M.D., and Burkart, M.R., 2003, Long-term effects of nitrogen fertilizer use on groundwater nitrate in two small watersheds: *Journal of Environmental Quality*, v. 32, p. 2158-2171.
- Trapp, Henry, Jr., and Horn, M.A., 1997, Ground water atlas of the United States—Segment 11, Delaware, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia: U.S. Geological Survey Hydrologic Atlas 730-L, 24 p.
- U.S. Census Bureau, 2004, U.S. Census Database, 2000: U.S. Geological Survey National Atlas, Reston, Va., digital data from <http://edcftp.cr.usgs.gov/pub/data/nationalatlas/ce2000t.tar.gz>, last accessed May 1, 2009.
- U.S. Census Bureau, 1992, Census of population and housing, 1990—Summary tape file 3A on CD-ROM (machine-readable data file): Washington, D.C., The Bureau (producer and distributor).
- U.S. Department of Agriculture, 2006, Keys to soil taxonomy, tenth edition: USDA Natural Resources Conservation Service, chap. 5, p. 35.
- U.S. Environmental Protection Agency, 2009, Technical Factsheet on nitrate/nitrate: U.S. Environmental Protection Agency, online technical fact sheet, accessed on Dec. 16, 2009, at <http://www.epa.gov/OGWDW/pdfs/factsheets/ioc/tech/nitrates.pdf>.
- U.S. Geological Survey, 2010, National Water-Quality Assessment (NAWQA) Program—About NAWQA study units: accessed June 30, 2010, at http://water.usgs.gov/nawqa/studies/study_units.html.
- Warner, K.L., and Arnold, T.L., 2006, 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., and Morrow, W.S., 2007, Pesticide and transformation product detections and age-dating relations from till and sand deposits: *Journal of the American Water Resources Association*, v. 43, no. 4, p. 1-12.
- Whitehead, R.L., 1996, Ground water atlas of the United States--Segment 8, Montana, North Dakota, South Dakota, Wyoming: U.S. Geological Survey Hydrologic Atlas 730-I, 24 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, at <http://water.usgs.gov/lookup/getspatial?muid>.
- Wolock, D.M., 2003, Estimated mean annual natural ground-water 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. 1273-1286.
- Young, H.L., 1992, Summary of ground-water hydrology of the Cambrian-Ordovician aquifer system in the northern Midwest—United States: U.S. Geological Survey Professional Paper 1405-A, 55 p.

Glossary

Alfisols — generally light-colored soils formed under forest environment (Natural Resources Conservation Service, 2008).

Anoxic — water samples, containing less than 0.5 mg/L of dissolved oxygen.

Bulk density — for soils, mass or weight of moist soil per unit volume (in this report, in grams per cubic centimeter).

Censored data — results of laboratory analysis of water samples in which the analyte concentration is too low to be reliably detected and quantified by the laboratory equipment.

Cumulative frequency distribution curve — visually shows how many observations in a data set are above or below a given value (Helsel and Hirsch, 1992).

Framework area — in this report, a preliminary classification scheme that identifies the most general and important regional-scale characteristics of the glacial aquifer system (Warner and Arnold, 2006). Four framework areas are defined: East, Central, West-Central, and West.

Glacial aquifer system — as defined in this report, it is the unconsolidated geologic material above bedrock that lies on or north of the line of maximum glacial advance within the United States.

Hot-spot differentiation — local clustering of high or low values that are significantly (statistically) different from the mean for the entire study area (Environmental Systems Research Institute, Inc., 2008).

Network — group of wells from which water samples were collected for a specific purpose.

Organic matter — in this report is the non-living component of soils which are a heterogeneous mixture composed largely of products resulting from microbial and chemical transformations of organic debris. Organic matter content of soils used in this report is expressed in percent by weight.

Perched groundwater — a discontinuous saturated lense, with unsaturated conditions above and below (Freeze and Cherry, 1979).

Residual — the difference between an observed value and a predicted value derived from a statistical model. A residual of 0 indicates that the predicted and observed values are identical. A negative residual indicates that the observed value is less than the predicted value; a positive residual indicates that the observed value is greater than the predicted value.

Silt content — in this report, the percentage by weight of a sample of unconsolidated material that consists of silt. Silt consists of small-diameter mineral particles that are larger than clay but smaller than sand.

Survival curve — see cumulative frequency distribution curve.

Study Unit — one of 51 geographically distinct study areas, based on the major river basins and aquifers of the United States that is an organizational unit of the USGS NAWQA program (U.S. Geological Survey, 2010). Data from 15 NAWQA study units were analyzed for the study described in this report.

Appendixes 1–9

Appendix 1. Variable names, descriptions, units, and sources used in the models described in this report.

[ft, feet; %, percent; NAWQA, National Water-Quality Assessment Program; STATSGO, State Soil Geographic Database]

Variable name in logistic regression models	Variable name in linear regression model	Description of variable	Units	Values used in logistic regressions	Values used in linear regression	Source of data
OI_CASE_DI	mDiamOlin	Open-interval diameter	inches	Values for each well	Mean for well network	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HOME:1931200546639998).
WTDEPAVE	mwtdepave	Depth to saturated soil within 1,640 ft of well	feet	Values for each well	Mean for well network	STATSGO soil data (Wolock, 1997). This variable is named WTDEPAVE, depth to seasonally high water table, in Wolock (1997).
SILTAVE	msiltave	Mean % silt in soil within 1,640 ft of the well	area weighted percent	Values for each well	Mean for well network	STATSGO soil data (Wolock, 1997).
NFARM_10YR	mnfarm.10y	10-year mean (1992-2001) nitrogen applied to farms within 1,640 ft of the well	kilograms per year	Values for each well	Mean for well network	10-year mean was calculated by summing nitrogen applications for the years 1992 through 2001 (Ruddy and others, 2006) and dividing by 10 years.
OILENGTH_I	moilength	Length of open interval	inches	Values for each well	Mean for well network	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HOME:1931200546639998).
DEPTH_TOI_	mdepthtoi	Depth to top of open interval	inches	Values for each well	Mean for well network	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HOME:1931200546639998).
ORDA	morda	% alfisols within 1,640 ft of well	area weighted percent	Values for each well	Mean for well network	STATSGO soil data (Wolock, 1997).
OI_CASE_D*DEPTH_TOI_	Not used in linear regression	Interaction of casing diameter and depth to top of open interval	dimensionless	Values for each well	Mean for well network	Calculated by model.

Appendix 2. Results of a comparison of nitrate measured in private and monitoring wells in the glacial aquifer system to the nationwide model of probability of nitrate concentration greater than 4 milligrams per liter as nitrogen based on all monitoring wells (Nolan and others, 2002). Data are from a spatial overlay of wells in the glacial aquifer system with a model of probability of nitrate contamination of groundwater (Nolan and others, 2002).

[mg/L, milligrams per liter; N, nitrogen; NA, not applicable; see figure 8-1A for scatter plot of the relation between observed and predicted probabilities]

Category of predicted probabilities	Number of wells in which measured nitrate concentration is less than or equal to 4 mg/L as N in each category	Number of wells in which measured nitrate concentration is greater than 4 mg/L as N in each category	Total number of wells in each category	Mean observed probability	Mean predicted probability	C-test statistic (Hosmer and Lemeshow, 2000, p. 148)	Multiple R-squared regression statistic for observed and predicted probabilities
Private wells							
0 to 0.17	88	3	91	0.03	0.09	3.56	NA
0.17 to 0.33	78	14	92	0.15	0.27	6.25	NA
0.33 to 0.50	73	13	86	0.15	0.41	23.97	NA
0.50 to 0.67	40	14	54	0.26	0.58	23.09	NA
0.67 to 0.83	23	4	27	0.15	0.73	45.20	NA
0.83 to 1.0	1	1	2	0.50	0.83	1.61	NA
All categories	303	49	352	NA	NA	103.68	0.79
Monitoring wells							
0 to 0.17	120	35	155	0.23	0.10	28.29	NA
0.17 to 0.33	152	45	197	0.23	0.25	0.66	NA
0.33 to 0.50	153	74	227	0.33	0.42	7.67	NA
0.50 to 0.67	89	78	167	0.47	0.57	7.62	NA
0.67 to 0.83	46	56	102	0.55	0.74	20.56	NA
0.83 to 1.0	3	8	11	0.73	0.85	1.41	NA
All categories	563	296	859	NA	NA	66.22	0.96

Appendix 3. Results of using a logistic regression model (Nolan and others, 2002) with six categories to calculate the probability of nitrate concentration greater than 4 milligrams per liter as nitrogen.

[mg/L, milligrams per liter; N, nitrogen; NA, not applicable; see figure 8–1A for scatter plot of relation between observed and predicted probabilities]

Category of predicted probabilities¹	Number of wells in which measured nitrate concentration is less than or equal to 4 mg/L as N in each category	Number of wells in which measured nitrate concentration is greater than 4 mg/L as N in each category	Total number of wells in each category	Mean observed probability	Mean predicted probability	Multiple R-squared regression statistic for observed and predicted probabilities
0 to 0.17	8	0	8	0	0.12	NA
0.17 to 0.33	4	1	5	0.2	0.25	NA
0.33 to 0.50	4	0	4	0	0.37	NA
0.50 to 0.67	7	0	7	0	0.57	NA
0.67 to 0.83	9	0	9	0	0.77	NA
0.83 to 1.0	276	48	324	0.15	0.99	NA
All categories	308	49	357	NA	NA	0.12

¹ Six probability Categories as in Nolan and others (2002)

60 Relations That Affect the Probability and Prediction of Nitrate Concentration in Private Wells, Glacial Aquifer System

Appendix 4.A Results of recalibrating a logistic regression model (Nolan and others, 2002): coefficients resulting from recalibrating the regression equation from Nolan and others (2002) and results of test to determine whether estimated coefficients of new variables are significant predictors (significantly different from 0) at $\alpha = 0.1$.

[NA, not applicable; %, percent; m, meters; km², square kilometers; kg/ha, kilograms per hectare; mg/L milligrams per liter; N, nitrogen; see figures 8-1C and 8-1D for scatter plots of relation between observed and predicted probabilities]

Coefficient variable	Value	Standard error	t-value	Wald statistic	Deviance	Residual degrees of freedom	Residual deviance	Probability of Chi-square value if coefficient variable is added sequentially, first to last
Intercept (NULL model)	-4.274	0.947	-4.51	-4.51	NA	356	285.56	NA
1992 nitrogen fertilizer applied to farms, in kg/ha	0.000	0.000	0.76	0.76	6.20	355	279.36	0.01
Agricultural land, in %	0.008	0.009	0.89	0.89	0.08	354	279.28	0.77
1990 population density, in people/ km ² (natural log)	0.076	0.122	0.62	0.62	0.09	353	279.19	0.77
Well-drained soils, in %	0.006	0.008	0.71	0.71	5.51	352	273.68	0.02
Depth to seasonally high water table, in m ¹	0.259	0.175	1.48	1.48	2.17	351	271.51	0.14
Unconsolidated sand and gravel aquifers, presence/absence	-0.013	0.335	-0.04	-0.04	0.00	350	271.51	0.97

¹ This variable is equivalent to the variable called 'depth to saturated soil' that was used in the logistic and linear regression models for the glacial aquifer system as described in this report.

Appendix 4B. Results of recalibrating a logistic regression model (Nolan and others, 2002): observed and predicted probabilities using the 6 categories of predicted probabilities in Nolan and others (2002) and the standard 10 categories.

[NA, not applicable; %, percent; m, meters; km², square kilometers; kg/ha, kilograms per hectare; mg/L milligrams per liter; N, nitrogen; see figures 8–1C and 8–1D for scatter plots of relation between observed and predicted probabilities]

Category of predicted probabilities	Number of wells in which measured nitrate concentration is less than or equal to 4 mg/L as N in each category	Number of wells in which measured nitrate concentration is greater than 4 mg/L as N in each category	Total number of wells in each category	Mean observed probability	Mean predicted probability	Multiple R-squared regression statistic for observed and predicted probabilities
Observed and predicted probabilities using the 6 categories as in Nolan and others (2002)						
0 to 0.17	233	28	261	0.11	0.10	NA
0.17 to 0.33	73	20	93	0.22	0.22	NA
0.33 to 0.50	2	1	3	0.33	0.36	NA
0.50 to 0.67	0	0	0	NA	NA	NA
0.67 to 0.83	0	0	0	NA	NA	NA
0.83 to 1.0	0	0	0	NA	NA	NA
All categories	308	49	357	NA	NA	1.00
Observed and predicted probabilities using the standard 10 categories						
0.02 to 0.058	35	1	36	0.03	0.04	NA
0.058 to 0.09	71	6	77	0.08	0.08	NA
0.09 to 0.13	61	10	71	0.14	0.11	NA
0.13 to 0.17	65	11	76	0.14	0.15	NA
0.17 to 0.20	26	7	33	0.21	0.19	NA
0.20 to 0.24	24	7	31	0.23	0.22	NA
0.24 to 0.28	15	4	19	0.21	0.25	NA
0.28 to 0.31	8	2	10	0.2	0.29	NA
0.31 to 0.35	2	0	2	0	0.33	NA
0.35 to 0.39	1	1	2	0.5	0.38	NA
All categories	308	49	357	NA	NA	0.56

Appendix 5. Logistic regression parameter estimates and odds ratio estimates for each of 10 logistic regression models.

[<, less than; >, greater than; mg/L, milligrams per liter; N, nitrogen; —, not applicable; see Appendix 6 for scatter plots of the relation between predicted and observed probabilities for each logistic regression model]

Coefficient ²	Value	Standard error	Chi-square value	Probability the coefficient is significant	Odds ratio estimate ¹			
					Effect	Point estimate	95% Wald confidence limits	
Nitrate concentration > 1 mg/L as N								
Intercept	-3.9552	0.762	26.931	<.0001	—	—	—	—
WTDEPAVE	0.7219	0.140	26.788	<.0001	WTDEPAVE	2.058	1.566	2.706
SILTAVE	0.0356	0.011	10.728	0.001	SILTAVE	1.036	1.014	1.059
ORDA	-0.0164	0.004	13.631	0.0002	ORDA	0.984	0.975	0.992
DEPTH_TOI_	-0.0012	0.000	23.397	<.0001	DEPTH_TOI_	0.999	0.998	0.999
OILENGTH_I	0.00309	0.001	4.554	0.033	OILENGTH_I	1.003	1	1.006
Nitrate concentration > 2 mg/L as N								
Intercept	-5.1108	0.908	31.697	<.0001	—	—	—	—
WTDEPAVE	0.8013	0.163	24.285	<.0001	WTDEPAVE	2.228	1.62	3.065
SILTAVE	0.0535	0.013	17.970	<.0001	SILTAVE	1.055	1.029	1.081
ORDA	-0.0188	0.005	14.563	0.0001	ORDA	0.981	0.972	0.991
DEPTH_TOI_	-0.00152	0.000	26.045	<.0001	DEPTH_TOI_	0.998	0.998	0.999
OILENGTH_I	0.000985	0.001	0.615	0.433	OILENGTH_I	1.001	0.999	1.003
Nitrate concentration > 3 mg/L as N								
Intercept	-4.6698	0.944	24.479	<.0001	—	—	—	—
WTDEPAVE	0.6468	0.168	14.765	0.0001	WTDEPAVE	1.909	1.373	2.656
SILTAVE	0.0481	0.013	12.955	0.0003	SILTAVE	1.049	1.022	1.077
ORDA	-0.0156	0.005	8.845	0.003	ORDA	0.985	0.974	0.995
DEPTH_TOI_	-0.00165	0.000	23.763	<.0001	DEPTH_TOI_	0.998	0.998	0.999
OILENGTH_I	0.00111	0.001	0.790	0.374	OILENGTH_I	1.001	0.999	1.004

¹. For continuous explanatory variables, the odds ratios correspond to a unit increase in the risk factors. For continuous explanatory variables, the point estimate corresponds to the odds ratio for a unit increase of the corresponding effect variable. Odds ratios are computed only for variables not involved in an interaction term.

². See Appendix 1 for a description of the coefficient variables.

Appendix 5. Logistic regression parameter estimates and odds ratio estimates for each of 10 logistic regression models.—Continued

[<, less than; >, greater than; mg/L, milligrams per liter; N, nitrogen; —, not applicable; see Appendix 6 for scatter plots of the relation between predicted and observed probabilities for each logistic regression model]

Coefficient ²	Value	Standard error	Chi-square value	Probability the coefficient is significant	Odds ratio estimate ¹			
					Effect	Point estimate	95% Wald confidence limits	
Nitrate concentration > 4 mg/L as N								
Intercept	-5.3277	1.014	27.600	<.0001	—	—	—	—
NFARM_10YR	0.000211	0.000	6.100	0.014	NFARM_10YR	1	1	1
OI_CASE_DI	0.0151	0.022	0.465	0.496	OI_CASE_DI	1.015	0.972	1.06
DEPTH_TOI_	-0.00187	0.000	16.159	<.0001	DEPTH_TOI_	0.998	0.997	0.999
ORDA	-0.0127	0.006	4.905	0.027	ORDA	0.987	0.976	0.999
SILTAVE	0.0326	0.014	5.137	0.023	SILTAVE	1.033	1.004	1.063
WTDEPAVE	0.7355	0.192	14.668	0.0001	WTDEPAVE	2.086	1.432	3.04
Nitrate concentration > 5 mg/L as N								
Intercept	-6.8249	1.296	27.746	<.0001	—	—	—	—
DEPTH_TOI_	-0.004	0.001	11.821	0.0006	—	—	—	—
NFARM_10YR	0.00042	0.000	13.843	0.0002	NFARM_10YR	1	1	1.001
OI_CASE_DI	-0.05	0.042	1.437	0.231	—	—	—	—
ORDA	-0.019	0.007	6.943	0.008	ORDA	0.981	0.967	0.995
SILTAVE	0.039	0.018	4.927	0.026	SILTAVE	1.04	1.005	1.076
WTDEPAVE	0.8839	0.251	12.409	0.0004	WTDEPAVE	2.42	1.48	3.958
DEPTH_TOI*OI_CASE_DI	0.000367	0.000	3.839	0.050	—	—	—	—

¹ For continuous explanatory variables, the odds ratios correspond to a unit increase in the risk factors. For continuous explanatory variables, the point estimate corresponds to the odds ratio for a unit increase of the corresponding effect variable. Odds ratios are computed only for variables not involved in an interaction term.

² See Appendix 1 for a description of the coefficient variables.

Appendix 5. Logistic regression parameter estimates and odds ratio estimates for each of 10 logistic regression models.—Continued

[<, less than; >, greater than; mg/L, milligrams per liter; N, nitrogen; —, not applicable; see Appendix 6 for scatter plots of the relation between predicted and observed probabilities for each logistic regression model]

Coefficient ²	Value	Standard error	Chi-square value	Probability the coefficient is significant	Odds ratio estimate ¹			
					Effect	Point estimate	95% Wald confidence limits	
Nitrate concentration > 6 mg/L as N								
Intercept	-4.7459	1.240	14.648	0.0001	—	—	—	—
NFARM_10YR	0.000555	0.000	15.837	<.0001	NFARM_10YR	1.001	1	1.001
OI_CASE_DI	-0.0909	0.052	3.024	0.082	—	—	—	—
WTDEPAVE	0.7818	0.277	7.982	0.005	WTDEPAVE	2.185	1.271	3.759
ORDA	-0.0211	0.009	5.567	0.018	ORDA	0.979	0.962	0.996
DEPTH_TOI_	-0.00563	0.002	11.507	0.001	—	—	—	—
OI_CASE_D*DEPTH_TOI_	0.000541	0.000	4.003	0.045	—	—	—	—
Nitrate concentration > 7 mg/L as N								
Intercept	-1.187	1.012	1.375	0.241	—	—	—	—
NFARM_10YR	0.000422	0.000	8.746	0.003	NFARM_10YR	1	1	1.001
OI_CASE_DI	-0.1599	0.067	5.694	0.017	—	—	—	—
DEPTH_TOI_	-0.00793	0.003	9.660	0.002	—	—	—	—
OI_CASE_D*DEPTH_TOI_	0.000772	0.000	5.072	0.024	—	—	—	—
Nitrate concentration > 8 mg/L as N								
Intercept	-1.396	1.076	1.684	0.194	—	—	—	—
NFARM_10YR	0.00047	0.000	9.257	0.002	NFARM_10YR	1	1	1.001
OI_CASE_DI	-0.1614	0.068	5.608	0.018	—	—	—	—
DEPTH_TOI_	-0.00812	0.003	9.385	0.002	—	—	—	—
OI_CASE_D*DEPTH_TOI_	0.000776	0.000	4.992	0.026	—	—	—	—

¹ For continuous explanatory variables, the odds ratios correspond to a unit increase in the risk factors. For continuous explanatory variables, the point estimate corresponds to the odds ratio for a unit increase of the corresponding effect variable. Odds ratios are computed only for variables not involved in an interaction term.

² See Appendix 1 for a description of the coefficient variables.

Appendix 5. Logistic regression parameter estimates and odds ratio estimates for each of 10 logistic regression models.—Continued

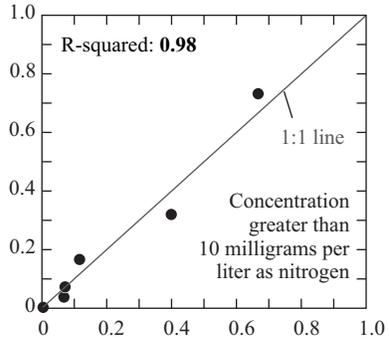
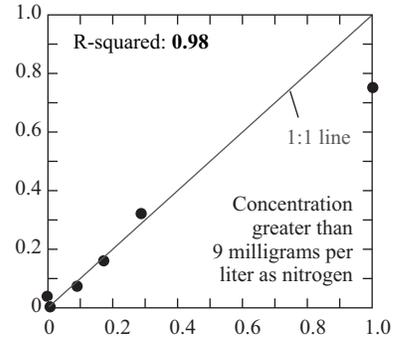
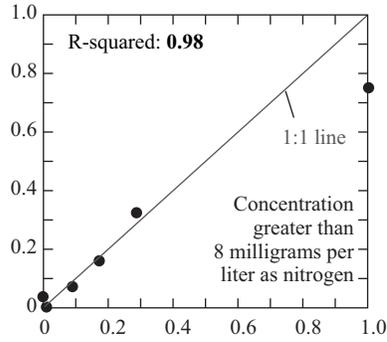
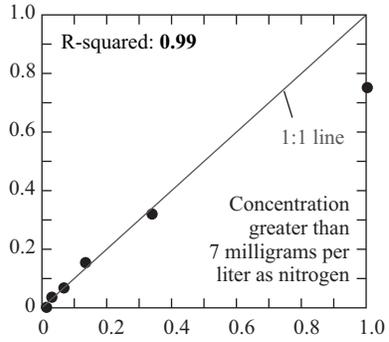
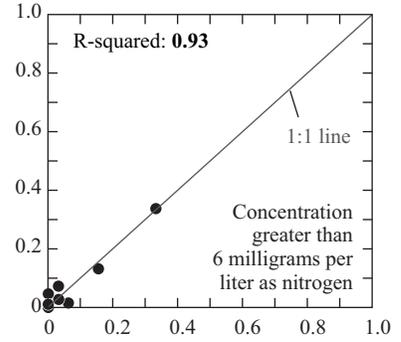
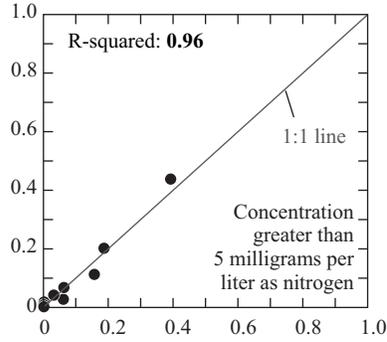
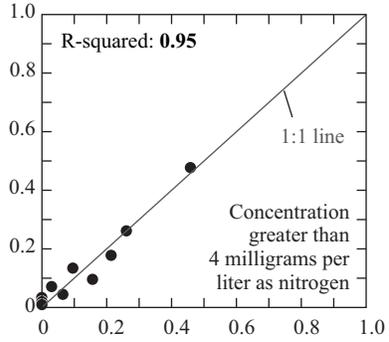
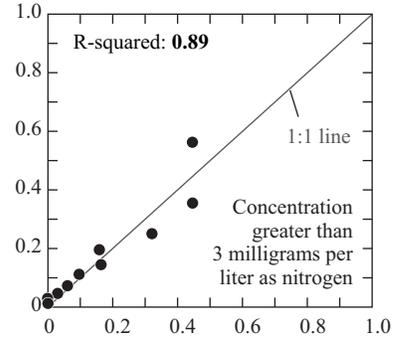
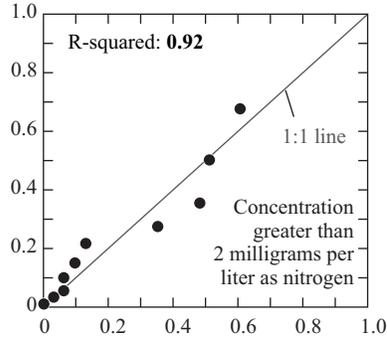
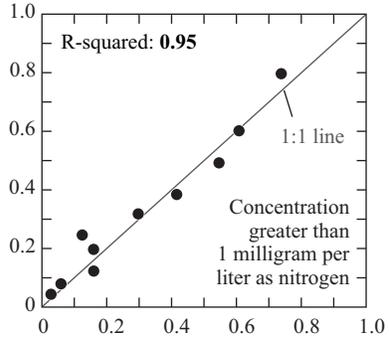
[<, less than; >, greater than; mg/L, milligrams per liter; N, nitrogen; —, not applicable; see Appendix 6 for scatter plots of the relation between predicted and observed probabilities for each logistic regression model]

Coefficient ²	Value	Standard error	Chi-square value	Probability the coefficient is significant	Odds ratio estimate ¹			
					Effect	Point estimate	95% Wald confidence limits	
Nitrate concentration > 9 mg/L as N								
Intercept	-1.0447	1.223	0.729	0.393	—	—	—	—
NFARM_10YR	0.000474	0.000	7.496	0.006	NFARM_10YR	1	1	1.001
OI_CASE_DI	-0.2021	0.078	6.721	0.010	—	—	—	—
DEPTH_TOI_	-0.0104	0.003	10.107	0.002	—	—	—	—
OI_CASE_D*DEPTH_TOI_	0.00103	0.000	6.656	0.010	—	—	—	—
Nitrate concentration > 10 mg/L as N								
Intercept	-2.1454	1.724	1.549	0.213	—	—	—	—
NFARM_10YR	0.000658	0.000	5.970	0.015	NFARM_10YR	1.001	1	1.001
OI_CASE_DI	-0.26	0.095	7.482	0.006	—	—	—	—
DEPTH_TOI_	-0.0123	0.004	8.785	0.003	—	—	—	—
OI_CASE_D*DEPTH_TOI_	0.00135	0.000	8.276	0.004	—	—	—	—

¹ For continuous explanatory variables, the odds ratios correspond to a unit increase in the risk factors. For continuous explanatory variables, the point estimate corresponds to the odds ratio for a unit increase of the corresponding effect variable. Odds ratios are computed only for variables not involved in an interaction term.

² See Appendix 1 for a description of the coefficient variables.

PREDICTED PROBABILITY THAT THE NITRATE CONCENTRATION IS GREATER THAN THE SPECIFIED CONCENTRATION



OBSERVED PROBABILITY THAT THE NITRATE CONCENTRATION IS GREATER THAN THE SPECIFIED CONCENTRATION

OBSERVED PROBABILITY THAT THE NITRATE CONCENTRATION IS GREATER THAN THE SPECIFIED CONCENTRATION

OBSERVED PROBABILITY THAT THE NITRATE CONCENTRATION IS GREATER THAN THE SPECIFIED CONCENTRATION

Appendix 6. Scatterplots showing predicted and observed probabilities of nitrate concentration greater than the specified concentration in private wells in the glacial aquifer system for all 10 logistic regression models developed for this study.

Appendix 7. Spearman rank correlations of nitrate concentration with selected variables, probability significance (p), and number of samples:

[mg/L as N, milligrams per liter as nitrogen; %, percent; >, greater than; in., inches; ft, feet; ln(m), natural log of value in meters; km, kilometers; kg/km²/yr, kilograms per square kilometers per year; kg/km², kilograms per square kilometer; km², square kilometers; in./hr, inches per hour; NAWQA, National Water-Quality Assessment Program; DEM, digital elevation model; gray shading indicates insignificant correlations; bold indicates correlations significant at the 95-percent confidence level; plain text indicates correlations significant at the 90-percent confidence level]

Variable	Spearman correlation coefficient	Probability of a significant correlation	Number of samples	Variable description	Original data source or variable	Description of variable values
A00631	1	NA	379	Nitrate (mg/L as N)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO:ME:1931200546639998)	Value at well.
PCTAG	-0.096	0.063	379	Agricultural land use (%)	Nakagaki and others (2007)	Area-weighted average for area within 1,640 ft of the well.
PCTFOR	0.098	0.056	379	Forest land use (%)	Nakagaki and others (2007)	Area-weighted average for area within 1,640 ft of the well.
PCTURBAN	0.088	0.089	379	Urban land use (%)	Nakagaki and others (2007)	Area-weighted average for area within 1,640 ft of the well.
AGGT50	-0.040	0.433	379	Agricultural land use > 50 % of area within 1,640 ft of a well, true/false binary indicator	Calculated for this report from Nakagaki and others (2007)	Determined from area-weighted average for area within 1,640 ft of the well.
FORGT50	0.083	0.105	379	Forest land use > 50 % of area within 1,640 ft of a well, true/false binary indicator	Calculated for this report from Nakagaki and others (2007)	Determined from area-weighted average for area within 1,640 ft of the well.
URBGT50	0.061	0.237	379	Urban land use > 50 % of area within 1,640 ft of a well, true/false binary indicator	Calculated for this report from Nakagaki and others (2007)	Determined from area-weighted average for area within 1,640 ft of the well.
MIXEDLU	-0.063	0.224	379	Mixed land use, no land use greater than 50 % of within 1,640 ft of a well, true/false binary indicator	Calculated for this report from Nakagaki and others (2007)	Determined from area-weighted average for area within 1,640 ft of the well.
AWCAVE	0.054	0.305	357	Mean available water capacity (fraction)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
BDAVE	-0.291	<0.0001	357	Mean bulk density (grams per cubic centimeter)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
CLAYAVE	-0.007	0.902	357	Mean clay content (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
SANDAVE	-0.103	0.051	357	Mean sand content (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.

Appendix 7. Spearman rank correlations of nitrate concentration with selected variables, probability significance (p), and number of samples.—Continued

[mg/L as N, milligrams per liter as nitrogen; %, percent; >, greater than; in., inches; ft, feet; ln(m), natural log of value in meters; km, kilometers; kg/km²/yr, kilograms per square kilometers per year; kg/km², kilograms per square kilometer; km², square kilometers; in./hr, inches per hour; NAWQA, National Water-Quality Assessment Program; DEM, digital elevation model; gray shading indicates insignificant correlations; bold indicates correlations significant at the 95-percent confidence level; plain text indicates correlations significant at the 90-percent confidence level]

Variable	Spearman correlation coefficient	Probability of a significant correlation	Number of samples	Variable description	Original data source or variable	Description of variable values
SILTAVE	0.159	0.003	357	Mean silt content (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
RECHARGE _{BF}	-0.004	0.945	356	Mean annual recharge from base-flow analysis and Gebert runoff (in.)	Wolock (2003)	Area-weighted average for area within 1,640 ft of the well.
RUNOFF	0.049	0.353	356	Mean annual runoff (1951–80) from Gebert (in.)	Gebert and others (1987)	Area-weighted average for area within 1,640 ft of the well.
CONTACT	-0.067	0.212	350	Subsurface flow contact time (days)	Wolock and others (1997)	Area-weighted average for area within 1,640 ft of the well.
DEPMIN	-0.103	0.051	357	Depth to minimally permeable soil layer (in.)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
HGA	0.044	0.404	357	Soil hydrologic group A (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
WTDEPAVE	0.246	<0.0001	357	Mean depth to saturated soil (ft)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
OMAVE	-0.135	0.011	357	Soil organic-matter content (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
ORDA	-0.213	<0.0001	357	Soil order A (alfisols) (%)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
PCTCOARSE	0.059	0.252	379	Coarse-grained surficial material (%)	Fullerton and others (2004); and Clawges and Price (1999); and National Park Service (1999).	Average for area within 1,640 ft of the well.
PCTFINE	-0.087	0.090	379	Percent fine-grained surficial material (%)	Fullerton and others (2004); and Clawges and Price (1999); and National Park Service (1999).	Average for area within 1,640 ft of the well.
SLOPE	0.159	0.003	356	Slope from 1-km-resolution DEM (%)	K.J. Hitt, U.S. Geological Survey, written commun., 2006	Averaged over all soil layers and for area within 1,640 ft of the well.
TWI	-0.068	0.199	356	Topographic wetness index from 1-km-resolution DEM (ln(m))	K.J. Hitt, U.S. Geological Survey, written commun., 2006	Averaged over all soil layers and for area within 1,640 ft of the well.

Appendix 7. Spearman rank correlations of nitrate concentration with selected variables, probability significance (p), and number of samples.—Continued

[mg/L as N, milligrams per liter as nitrogen; %, percent; >, greater than; in., inches; ft, feet; ln(m), natural log of value in meters; km, kilometers; kg/km²/yr, kilograms per square kilometers per year; kg/km², kilograms per square kilometer; km², square kilometers; in./hr, inches per hour; NAWQA, National Water-Quality Assessment Program; DEM, digital elevation model; gray shading indicates insignificant correlations; bold indicates correlations significant at the 95-percent confidence level; plain text indicates correlations significant at the 90-percent confidence level]

Variable	Spearman correlation coefficient	Probability of a significant correlation	Number of samples	Variable description	Original data source or variable	Description of variable values
PERDUN	-0.022	0.673	356	Percent Dunne overland flow (%)	K.J. Hitt, U.S. Geological Survey, written commun., 2006	Averaged over all soil layers and for area within 1,640 ft of the well.
PERHOR	0.077	0.154	349	Percent Hortonian overland flow (%)	K.J. Hitt, U.S. Geological Survey, written commun., 2006	Averaged over all soil layers and for area within 1,640 ft of the well.
PERMAVE	0.027	0.618	357	Average soil permeability (in./hr)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
ROCKDEPAVE	-0.192	0.000	357	Average total soil thickness examined (in.)	Wolock (1997)	Averaged over all soil layers and for area within 1,640 ft of the well.
DEPTH_TOI_	-0.240	<0.0001	347	Depth to top of open interval (ft)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO ME:1931200546639998)	Value at well.
DEPTHBOI_F	-0.198	0.000	314	Depth to bottom of open interval (ft)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO ME:1931200546639998)	Value at well.
OILENGTH_I	0.125	0.027	314	Open interval length (in.)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO ME:1931200546639998)	Value at well.
OI_CASE_DI	0.228	<0.0001	371	Diameter of open interval (in.)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO ME:1931200546639998)	Value at well.
WELLDPTH_I	-0.240	<0.0001	378	Well depth (ft)	NAWQA data warehouse (http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:HO ME:1931200546639998)	Value at well.
NDEP_10YRA	-0.253	<0.0001	357	10-year mean nitrogen atmospheric deposition (kg/km ² /yr)	Calculated for this report from Ruddy and others (2006)	Average for area within 1,640 ft of the well.

Appendix 7. Spearman rank correlations of nitrate concentration with selected variables, probability significance (p), and number of samples.—Continued

[mg/L as N, milligrams per liter as nitrogen; %, percent; >, greater than; in., inches; ft, feet; ln(m), natural log of value in meters; km, kilometers; kg/km²/yr, kilograms per square kilometers per year; kg/km², kilograms per square kilometer; km², square kilometers; in./hr, inches per hour; NAWQA, National Water-Quality Assessment Program; DEM, digital elevation model; gray shading indicates insignificant correlations; bold indicates correlations significant at the 95-percent confidence level; plain text indicates correlations significant at the 90-percent confidence level]

Variable	Spearman correlation coefficient	Probability of a significant correlation	Number of samples	Variable description	Original data source or variable	Description of variable values
NFARM_10YR	-0.074	0.162	357	10-year mean nitrogen fertilizer applied to farms (kg/yr)	Calculated for this report from Ruddy and others (2006)	Average for area within 1,640 ft of the well.
NNONF_10YR	0.015	0.778	357	10-year mean nitrogen fertilizer applied to other than farms (kg/yr)	Calculated for this report from Ruddy and others (2006)	Average for area within 1,640 ft of the well.
PDENS00	0.019	0.718	379	2000 population density (people/km ²)	U.S. Census Bureau (2004)	Average for area within 1,640 ft of the well.
PDENS90	0.003	0.949	379	1990 population density (people/km ²)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
SEWPUB	0.127	0.014	379	Public sewer (% of housing units)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
SEWSEP	-0.131	0.010	379	Septic tank or cesspool (% of housing units)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
MEDYRBLT	-0.054	0.297	379	Median year housing built	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
WATDRWEL	-0.257	<0.0001	379	Individual well, drilled (% of housing units)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
WATDUWEL	0.124	0.016	379	Individual well, dug (% of housing units)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.
WATPUBPV	0.199	<0.0001	379	Public system or private company (% of housing units)	U.S. Census Bureau (1992)	Average for area within 1,640 ft of the well.

Appendix 8. Variables selected in each of three iterations of the development of the 10 logistic regression models developed for this study.

[>, greater than; mg/L as N, milligrams per liter as nitrogen; grey shaded indicates selected; grey shaded and crossed out indicates selected but later omitted from subsequent iterations of the model; no shading indicates not selected; ROC, Receiver Operating Characteristic Curve statistic; H-L, Hosmer-Lemeshow chi-square goodness-of-fit statistic; see Appendix 1 for description of variables]

Logistic regression models										
Variable, goodness-of-fit statistics, and number of wells	Nitrate concentration > 10 mg/L as N	Nitrate concentration > 9 mg/L as N	Nitrate concentration > 8 mg/L as N	Nitrate concentration > 7 mg/L as N	Nitrate concentration > 6 mg/L as N	Nitrate concentration > 5 mg/L as N	Nitrate concentration > 4 mg/L as N	Nitrate concentration > 3 mg/L as N	Nitrate concentration > 2 mg/L as N	Nitrate concentration > 1 mg/L as N
First iteration of variable selection										
intercept										
nfarm_10yr										
welldpth_i										
oi_case_di										
wtdepave										
orda										
depthtoi_i										
rockdepave										
siltave										
oilength_i										
bdave										
depthtoi*oi_case_di										
ROC	0.86	0.929	0.901	0.892	0.863	0.851	0.797	0.825	0.832	0.808
Second iteration of variable selection										
intercept										
depthtoi_i										
oi_case_di										
nfarm_10yr										
wtdepave										
orda										
siltave										
oilength_i										
depthtoi*oi_case_di										
ROC	0.952	0.918	0.901	0.89	0.877	0.87	0.797	0.819	0.826	0.806

Appendix 8. Variables selected in each of three iterations of the development of the 10 logistic regression models developed for this study.—Continued

[>, greater than; mg/L as N, milligrams per liter as nitrogen; grey shaded indicates selected; grey shaded and crossed out indicates selected but later omitted from subsequent iterations of the model; no shading indicates not selected; ROC, Receiver Operating Characteristic Curve statistic; H-L, Hosmer-Lemeshow chi-square goodness-of-fit statistic; see Appendix 1 for description of variables]

Logistic regression models										
Variable, goodness-of-fit statistics, and number of wells	Nitrate concentration > 10 mg/L as N	Nitrate concentration > 9 mg/L as N	Nitrate concentration > 8 mg/L as N	Nitrate concentration > 7 mg/L as N	Nitrate concentration > 6 mg/L as N	Nitrate concentration > 5 mg/L as N	Nitrate concentration > 4 mg/L as N	Nitrate concentration > 3 mg/L as N	Nitrate concentration > 2 mg/L as N	Nitrate concentration > 1 mg/L as N
Third iteration of variable selection and final models										
intercept										
depthtoi_i										
oi_case_di										
nfarm_10yr										
wtdepave										
orda										
siltave										
oilength_i										
depthtoi*oi_case_di										
ROC	0.948	0.918	0.901	0.890	0.877	0.870	0.820	0.815	0.826	0.806
H-L	2.0	5.3	6.5	4.0	9.6	4.4	4.9	6.0	8.0	4.6
H-L probability	0.9	0.6	0.6	0.9	0.3	0.8	0.8	0.6	0.4	0.8
H-L degrees of freedom	6	7	8	8	8	8	8	8	8	8
Number of wells used in each model	14	19	21	22	26	36	50	68	91	123

Appendix 9. Linear regression to predict network mean nitrate concentration using the same variables used in the logistic regression models. The variables were used because they were used in the logistic regression models. The significance level of variables was not a consideration for this model.

[NA, not applicable; see figure 10.1 for scatter plot of relation between observed and predicted probabilities; see appendix 1 for description of variables; multiple R-squared = 0.83; F-statistic = 5.45 on 7 and 8 degrees of freedom and p-value = 0.02; the variables were used because they were used in the logistic regression models and the significance level of variables was not a consideration for this linear regression model]

Coefficient variables	Value	Standard error	Probability the coefficient is significant	Probability of significance if coefficient variable is added sequentially, first to last
Intercept (NULL Model)	-2.184	1.420	0.16	NA
mDiamOlin	0.158	0.081	0.09	0.00
mwtdepave	0.479	0.320	0.17	0.12
morda	-0.017	0.011	0.17	0.61
msiltave	0.054	0.029	0.10	0.40
mnfarm.10y	0.000	0.000	0.19	0.41
moilength	-0.003	0.006	0.64	0.82
mdepthtoi	-0.001	0.001	0.07	0.07

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

<http://il.water.usgs.gov>

