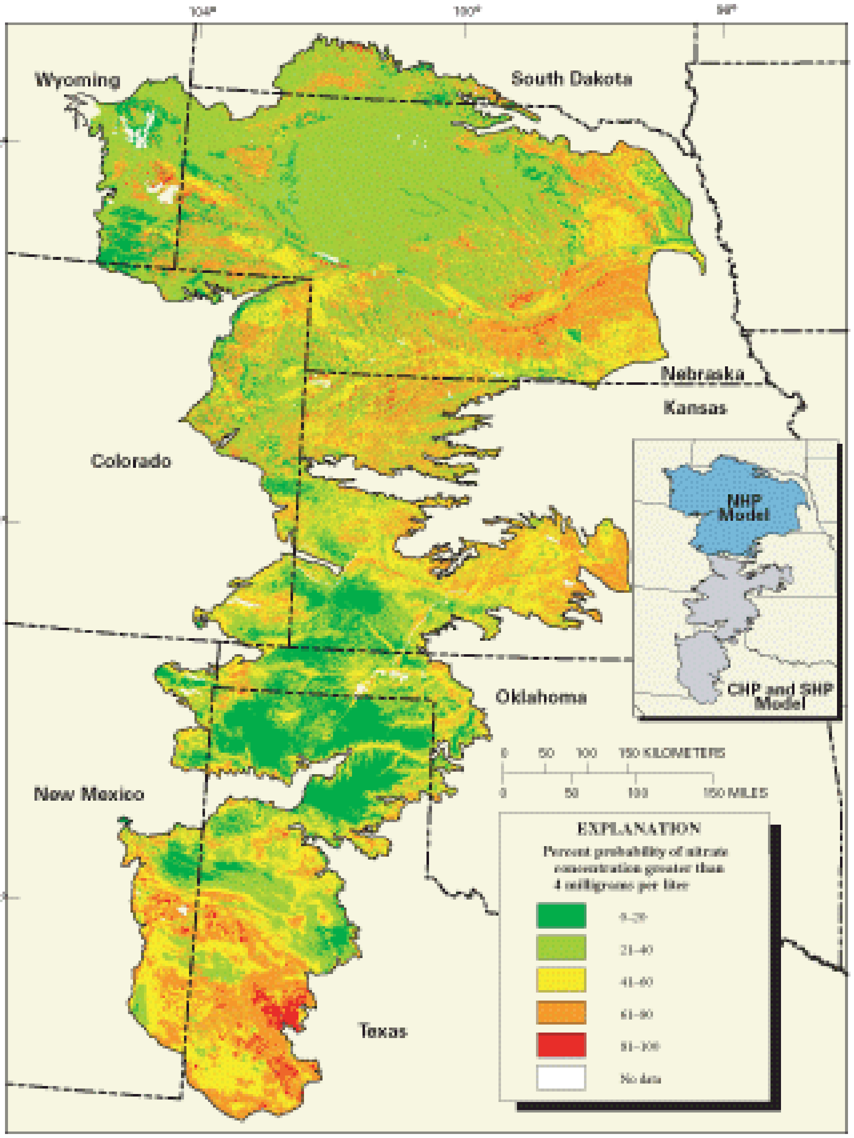
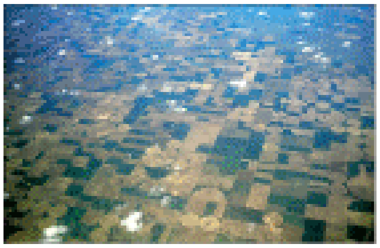
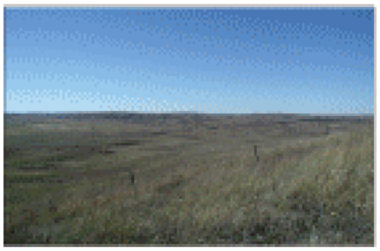
Vulnerability of Recently Recharged Ground Water in the High Plains Aquifer to Nitrate Contamination







Scientific Investigations Report 2006–5050

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

Vulnerability of Recently Recharged Ground Water in the High Plains Aquifer to Nitrate Contamination

By Jason J. Gurdak and Sharon L. Qi

Scientific Investigations Report 2006–5050

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

U.S. Department of the Interior
P. Lynn Scarlett, Acting Secretary

U.S. Geological Survey
P. Patrick Leahy, Acting Director

U.S. Geological Survey, Reston, Virginia: 2006

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

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

Suggested citation:
Gurdak, J.J., and Qi, S.L., 2006, Vulnerability of recently recharged ground water in the High Plains aquifer to nitrate contamination: U.S. Geological Survey Scientific Investigations Report 2006–5050, 39 p.

Contents

Figures

Tables

Conversion Factors, Vertical Datum, Acronyms, and Abbreviations

Multiply	By	To obtain
	Length	
centimeter (cm)	0.3937	inch (in.)
millimeter (mm)	0.03937	inch (in.)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Area	
square meter (m^2)	0.0002471	acre
square kilometer (km^2)	247.1	acre
square kilometer (km^2)	0.3861	square mile (mi^2)
	Volume	
liter (L)	1.057	quart (qt)
cubic meter (m^3)	0.0008107	acre-foot (acre-ft)
cubic meter (m^3)	264.2	gallon (gal)
	Velocity and flow rate	
centimeter per year (cm/yr)	0.3937	inch per year (in/yr)
meter per day (m/d)	3.281	foot per day (ft/d)
meter per year (m/yr)	3.281	foot per year (ft/yr)
millimeter per year (mm/yr)	0.03937	inch per year (in/yr)
cubic meter per day (m^3/d)	264.2	gallon per day (gal/d)
	Mass	
gram (g)	0.03527	ounce, avoirdupois (oz)
kilogram (kg)	2.205	pound avoirdupois (lb)
	Hydraulic conductivity	
meter per day (m/d)	3.281	foot per day (ft/d)

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

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

Acronyms and Abbreviations

^{14}C	carbon-14
AIC	Akaike information criteria
BP	before present
CHP	Central High Plains
d	day
ERS	Economic Research Service
GIS	geographic information system
HL	Hosmer-Lemeshow
LLR	log-likelihood ratio
MCL	Maximum Contaminant Level
mg/L	milligrams per liter
mg/L as N	milligrams per liter as nitrogen
N	nitrogen
N_2	nitrogen gas
NASS	National Agricultural Statistic Service
NAWQA	National Water-Quality Assessment

Acronyms and Abbreviations—Continued

NH_4^+	ammonium
NHP	Northern High Plains
NLCD	National Land Cover Data
NO_2^-	nitrite
NO_3^-	nitrate
NRC	National Research Council
PC	percent correct
PRISM	Parameter-elevation Regression on Independent Slopes Model
ROC	Receiver Operating Characteristic
SC	Schwartz criterion
SHP	Southern High Plains
STATSGO	State soil geographic database
TU	tritium units
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VIF	Variance Inflation Factor

Vulnerability of Recently Recharged Ground Water in the High Plains Aquifer to Nitrate Contamination

By Jason J. Gurdak and Sharon L. Qi

Abstract

Nitrate concentrations greater than background levels have been detected in ground water of the High Plains aquifer. Empirically based models and corresponding maps were developed that predict the vulnerability of the aquifer to nonpoint-source nitrate contamination. The models predict the probability of detecting nitrate concentrations larger than 4 milligrams per liter in ground water of the High Plains aquifer that was recharged during the last 50 years. The models were calibrated by correlating concentrations of nitrate in ground water from 336 wells that intercept recently recharged (less than 50 years) water with anthropogenic and hydrogeologic explanatory variables. Particle-tracking simulations delineated well-contributing areas and determined well-screen depths that intercept recently recharged ground water. The models were developed using multivariate logistic regression, and a map was generated from these models using a geographic information system. Two multivariate logistic regression models of vulnerability were found to have the most statistical significance and the best model fit and predictive ability. The two models represent the Northern High Plains and the combined Central and Southern High Plains, and they indicate that ground-water vulnerability of the entire High Plains aquifer is best explained by the spatial distribution of nonirrigated and irrigated agricultural lands, organic matter of the soil, depth to the regional water table, and clay content of the unsaturated zone. Vulnerability of the Northern High Plains is greater in areas that have more nonirrigated and irrigated agricultural lands and less organic matter in the soil. The vulnerability of the Central and Southern High Plains also is greater in areas that have more nonirrigated and irrigated agricultural lands and also in areas with shallow depths to water table and less clay in the unsaturated zone. The majority (53.3 percent) of the High Plains aquifer has less than a 40-percent predicted probability of nitrate concentrations larger than 4 milligrams per liter. Approximately 21.1 percent of the High Plains aquifer has a relatively high (greater than 60 percent) predicted probability of nitrate concentrations greater than or equal to 4 milligrams per liter. Areas with relatively high predicted probability are located in the southwestern, southern, and eastern areas of the Northern High Plains, in the eastern arm of the Central High Plains, and in southern areas of the Southern High Plains. Areas of the aquifer with relatively low (less than 40 percent) predicted vulnerability to nitrate concentrations greater than or equal to 4 milligrams per liter are located in the northwestern and north-central areas of the Northern High Plains, the central and southern areas of the Central High Plains, and a band across the north-central part of the Southern High Plains. Uncertainty of these vulnerability predictions was estimated by Latin hypercube sampling to address propagation of model and data errors inherently associated with estimates of model coefficients and explanatory variables. Results of the Latin hypercube sampling simulations are presented as uncertainty maps of the lower 5th and upper 95th percentile of the output probability distribution, which represents the 90-percent prediction interval that contains the true probability of detecting nitrate greater than or equal to 4 milligrams per liter. Generally, these uncertainty maps show greater prediction uncertainty in areas with relatively higher predicted vulnerability and lower uncertainty in areas with relatively lower predicted vulnerability.

Introduction

Assessment of the quality of the Nation's water resources and its risk to contamination is important because of the implications to human and aquatic health. Studies that aid this assessment provide valuable information for water and land management, conservation, and remediation. In 1991, the U.S. Geological Survey (USGS) began full implementation of the National Water Quality Assessment (NAWQA) Program. The long-term goals of the NAWQA Program are to describe the status and trends in the quality of the Nation's surface- and ground-water resources and determine the natural and anthropogenic factors affecting the water quality (Gilliom and others, 1995).

The High Plains Regional Ground Water study, which entered the NAWQA Program in 1998, represents a modification of the traditional NAWQA design in that the ground-water resource is the primary focus of study. The High Plains aquifer underlies 450,658 km^2 in parts of eight States (fig. 1) and is an important National resource, encompassing approximately

27 percent of the Nation's irrigated agricultural land and yielding about 30 percent of the ground water used for irrigation in the United States (Dennehy, 2000). Because of the large areal extent of the High Plains aquifer and logistical considerations, the study area has been divided into three geographic subregions of the High Plains aquifer (fig. 1): Northern High Plains (249,934 km^2); Central High Plains (125,614 km^2); and Southern High Plains (75,110 km^2).

Nitrate is the most ubiquitous chemical contaminant in the world's ground water (Spalding and Exner, 1993). Ingestion of nitrate is thought to be one of a number of cofactors that play a sometimes complex role in causing methemoglobinemia, or "blue baby" disorder, a potentially fatal condition caused by low oxygen levels in the blood of infants (Fewtrell, 2004). To address potential health concerns, the U.S. Environmental Protection Agency (USEPA) established a Maximum Contaminant Level (MCL) of 10-mg/L nitrate as N in drinking water (U.S. Environmental Protection Agency, 2004). In addition, ingestion of nitrate also has been linked to spontaneous abortion among some women (Centers for Disease Control and Prevention, 1996).

Similar to other aquifers beneath agricultural areas, the ground water in the unconfined High Plains aquifer is vulnerable to nitrate contamination. Nitrate concentrations above background levels in ground water of the High Plains aquifer are direct evidence of this vulnerability and have been documented by a number of recent studies of ground-water quality (Pope and others, 2000; Becker and others, 2002; Fahlquist, 2003; Bruce and others, 2003), including a retrospective analysis of historical ground-water-quality data of the study area (Litke, 2001). Nitrate contamination of ground water has created a potential health concern for the nearly 1.9 million people that rely on ground water from the High Plains aquifer as the primary source of domestic and public-supply water (Dennehy, 2000). A recent epidemiologic study of drinking water across parts of the Northern High Plains in eastern Nebraska identified a link between increased risk of non-Hodgkins lymphoma and long-term consumption of ground water containing nitrate (as N) greater than 4 mg/L (Ward and others, 1996).

The factors affecting the spatial distribution of nitrate in the High Plains aquifer are complex and poorly understood; however, some general trends have been identified. Results from previous ground-water-quality studies indicate that nitrate concentrations in ground water varies spatially across the High Plains aquifer (Pope and others, 2000; Becker and others, 2002; Fahlquist, 2003; Bruce and others, 2003), and anthropogenic activity during the last 50 years has resulted in nitrate concentrations in recently recharged water that are significantly larger than concentrations in older water in the aquifer (McMahon and others, 2000; McMahon, 2001; Litke, 2001; McMahon, Böhlke, and Christenson, 2004).

To date, no studies have been designed to estimate the potential for nonpoint-source nitrate contamination of ground water recharged during the last 50 years to the High Plains aquifer. It is likely that a predictive model of nitrate vulnerability of recently recharged ground water can be successfully established and validated using empirical relations between ground-water-quality data and independent or "explanatory" variables (Nolan and others, 2002; Rupert 2003).

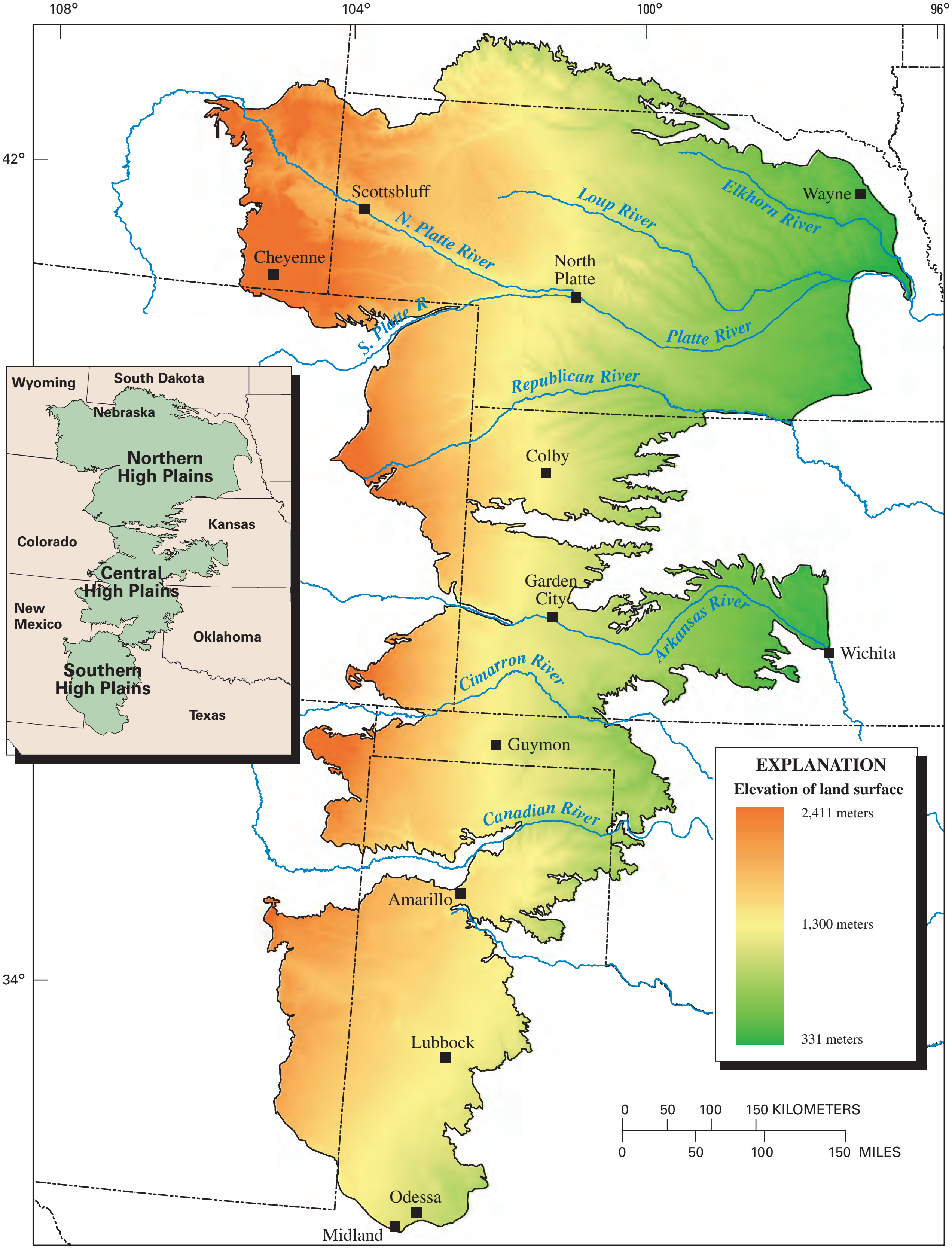
Background

Ground-water vulnerability assessments are designed to estimate and predict the potential for nonpoint-source contamination of a ground-water resource. Federal, State, and local water-management programs have used ground-water vulnerability assessments to identify areas most likely to have contamination and to make informed decisions regarding allocation of resources for ground-water protection and remediation.

The vulnerability of ground water is a function of the properties of the ground-water-flow system (intrinsic susceptibility), as well as the proximity of contamination sources, characteristics of the contaminant, and other factors that cause loading of contaminants to the land surface and delivery to the ground water (Focazio and others, 2002). The intrinsic susceptibility of an aquifer defines the natural and inherent hydrogeologic properties of the ground-water-flow system, including the sources of water and stresses on the system. Therefore, intrinsic susceptibility assessment evaluates how such properties as hydraulic conductivity (horizontal and vertical), porosity, hydraulic gradient, soil taxonomy, depth to water, unsaturated zone lithology, and rates and type of recharge and discharge influence contamination, without specifically addressing properties of that contaminant. A complete ground-water vulnerability assessment to contamination considers the intrinsic susceptibility, and the location and types of sources of naturally occurring and anthropogenic contamination, relative locations of wells, and the fate and transport of the contaminant(s) (Focazio and others, 2002).

Many definitions of ground-water vulnerability have arisen (Aller and others, 1985; Rao and Alley, 1993; U.S. Environmental Protection Agency, 1993; Vowinkel and others, 1996; Focazio and others, 2002), because vulnerability assessments are a complex indicator of contamination. Previous studies have tailored vulnerability assessment definition and approach to fit specific ground-water-resources (intrinsic susceptibility) and contaminant-specific information (National Research Council, 1993). As Rupert (2003) noted, even though multiple definitions have been used for the term "vulnerability," they all attempt to address the same underlying question: "What is the potential for [nonpoint source] ground-water contamination?"

The definition of ground-water vulnerability is important because it guides which approach is most appropriate to scientifically estimate the potential for ground-water vulnerability. Three broad approaches have been used to assess ground-water vulnerability: overlay and index, statistical, and process based. Although no approach is better than the others for all situations, statistical approaches have been used successfully for



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 1. Location of the High Plains aquifer.

subregional to National scale assessments of ground-water vulnerability (Tesoriero and Voss, 1997; Nolan and Stoner, 2000; Rupert, 1998, 2003; Nolan, 2001; Nolan and others, 2002). The National Research Council (NRC) (1993) proposed that ground-water vulnerability to contamination is "the tendency or likelihood for contamination to reach a specified position in the ground-water system after introduction at some location above the uppermost aquifer." This definition has particular relevance for regional-scale assessments that use statistical approaches to estimate ground-water vulnerability. Specific statistical approaches, such as logistic regression analysis, follow the National Research Council (1993) definition because actual probabilities of nitrate detections, or "likelihood for contamination," are quantified. The National Research Council (1993) proposed that a vulnerability assessment could include a selected part, or thickness, of an aquifer, that is, "to reach a specified position in the ground water system," rather than assessing the entire saturated thickness of an aquifer system. Furthermore, the National Research Council (1993) defined contamination as originating from "some location above the uppermost aquifer," implying surface-derived nitrate, possibly from anthropogenic activity.

Two previous studies, Nolan and others (2002) and Rupert (2003), have partially estimated ground-water vulnerability of the High Plains aquifer to nitrate contamination. Nolan and others (2002) developed a logistic regression model to predict the probability of nitrate contamination exceeding 4 mg/L in predominantly shallow, recently recharged ground waters of the conterminous United States. The logistic regression model was based on statistically significant empirical relations between nitrate-concentration data from 1,280 wells across the United States and 6 explanatory variables (nitrogen fertilizer loading, percentage of cropland as pasture, natural log of human population density, percentage of well-drained soils, depth to seasonally high water table, and presence or absence of unconsolidated sand and gravel aquifers) (Nolan and others, 2002). This predictive model was applied to the entire conterminous United States, including the High Plains aquifer, even though few data used to calibrate the model were located within the High Plains aquifer boundary (Nolan and others, 2002). Therefore, those calibration data were not spatially representative of the intrinsic susceptibility or vulnerability of the High Plains aquifer.

Rupert (2003) used logistic regression analyses to develop maps showing the probability of detecting nitrate concentrations greater than 5 mg/L as N in ground water in Colorado. Two nitrate models were constructed: one included fertilizer use and the other did not. The statistically significant explanatory variables identified were available water capacity of soil, clay content of soil, organic matter content of soil (fertilizer-only model), and the following land-cover classifications: shrubland, row crops, and small grains land cover. Similar to Nolan and others (2002), the majority of the calibration data was collected outside the High Plains aquifer boundary. Only 38,829 km^2 of the High Plains aquifer intersect Colorado, which is 8.62 percent of the total area of the aquifer, or stated differently, approximately 1.23×10^{11} m^3 of water in storage, or less than 4 percent of the total volume of water (3.68×10^{12} m^3 in the year 2000) in storage in the High Plains aquifer (McGuire and others, 2003). A ground-water vulnerability assessment of the High Plains aquifer would be greatly improved using ground-water-quality data and explanatory variables collected exclusively from within the aquifer boundary.

Purpose and Scope

The purpose of this report is to present an empirically based model and corresponding maps that identify the vulnerability of recently (less than 50 years) recharged ground water in the High Plains aquifer to nonpoint-source nitrate contamination. The predicted probability of detecting concentrations of nitrate (NO_3^- as N) larger than 4 mg/L is described. The model was developed using multivariate logistic regression and represents the relation between nitrate concentration in recently recharged ground water from samples collected across the High Plains aquifer during 1990 through 2004, and independent, or explanatory, variables of land use and hydrogeologic features of the aquifer. Data of nitrate concentration in ground water were selected from 336 wells with short screen intervals (less than 15 m) that intercept recently recharged ground water. Ground-water flow and particle-tracking simulations were used to select appropriate screen-interval depths and to define contributing areas for each well. The probability map is based on the final multivariate models and corresponding explanatory variables from existing and constructed Geographic Information System (GIS) data. Results of Latin hypercube simulations were used to evaluate the uncertainty of model predictions due to model and data error and are presented as 90-percent prediction interval maps.

Acknowledgments

This study was part of the USGS NAWQA Program, High Plains Regional Ground Water study. The authors would like to recognize Scott C. Christenson, USGS, and Richard R. Luckey, USGS, for their guidance during modeling design of this study, and Curtis Price, USGS, for assistance during GIS spatial processing. Dr. John E. McCray (Colorado School of Mines), Bernard T. Nolan (U.S. Geological Survey), and Michael G. Rupert (U.S. Geological Survey) provided helpful reviews of this report.

Description of Study Area

The High Plains study area covers about 450,660 km^2 (fig. 1), which includes the higher elevations of the Great Plains physiographic province. The study area rises in elevation from the Central lowlands (330 m) in the east to the Rocky Mountains (2,375 m) in the west. The High Plains extend from South Dakota to west-central Texas, and cover parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. The topography of the High Plains is characterized by flat to gently rolling terrain, which is remnant of a vast plain of sediments deposited from streams of the ancestral Rocky Mountains. Erosion has isolated the High Plains and formed escarpments that mark the plains boundaries. More detailed description of the study area can be found in Gutentag and others (1984) and Weeks and others (1988). The geohydrology and land and water use are described in the following sections.

Geohydrologic Setting

The High Plains aquifer generally is unconfined and consists mainly of hydraulically connected geologic units of late Tertiary and Quaternary age. The Ogallala Formation of Pliocene age has the largest areal extent of the geologic units that comprise the High Plains aquifer and underlies approximately 347,060 km^2 (Luckey and others, 1986). The formation was deposited from streams that flowed eastward from the ancestral Rocky Mountains and consists primarily of unconsolidated clay, silt, sand, and gravel with locally cemented zones consisting of calcium carbonate. The carbonate-cemented zones commonly occur near the top of the formation and form escarpments that typically mark the boundary of the High Plains. Locally and sometimes extensive clay layers are present within the unsaturated sediments and below the water table. Sedimentary bedrock units underlie the High Plains aquifer and range in age from Permian to Tertiary (Gutentag and others, 1984).

In 2000, the depth to water in the High Plains aquifer ranged from less than 3 to greater than 90 m below land surface (V.L. McGuire, U.S. Geological Survey, written commun., 2001), and the saturated thickness ranged from less than 1 to greater than 379 m (S.L. Qi, U.S. Geological Survey, written commun., 2001). Depths to water are greatest in the Central and Southern High Plains, whereas the saturated thickness is greatest in the Northern High Plains (fig. 2).

The hydraulic gradients in the High Plains aquifer indicate that regional ground-water movement is generally from west to east. However, local variability in hydraulic gradients indicates ground-water movement from northwest to southeast occurs across parts of the Northern and Southern High Plains. Gutentag and others (1984) established a range of hydraulic gradients across the entire High Plains aquifer from 0.001 to 0.003 m/m; however, local hydraulic gradients can vary depending on the intensity of well pumpage, lithology, recharge and discharge, proximity to streams, and other hydrogeologic factors.

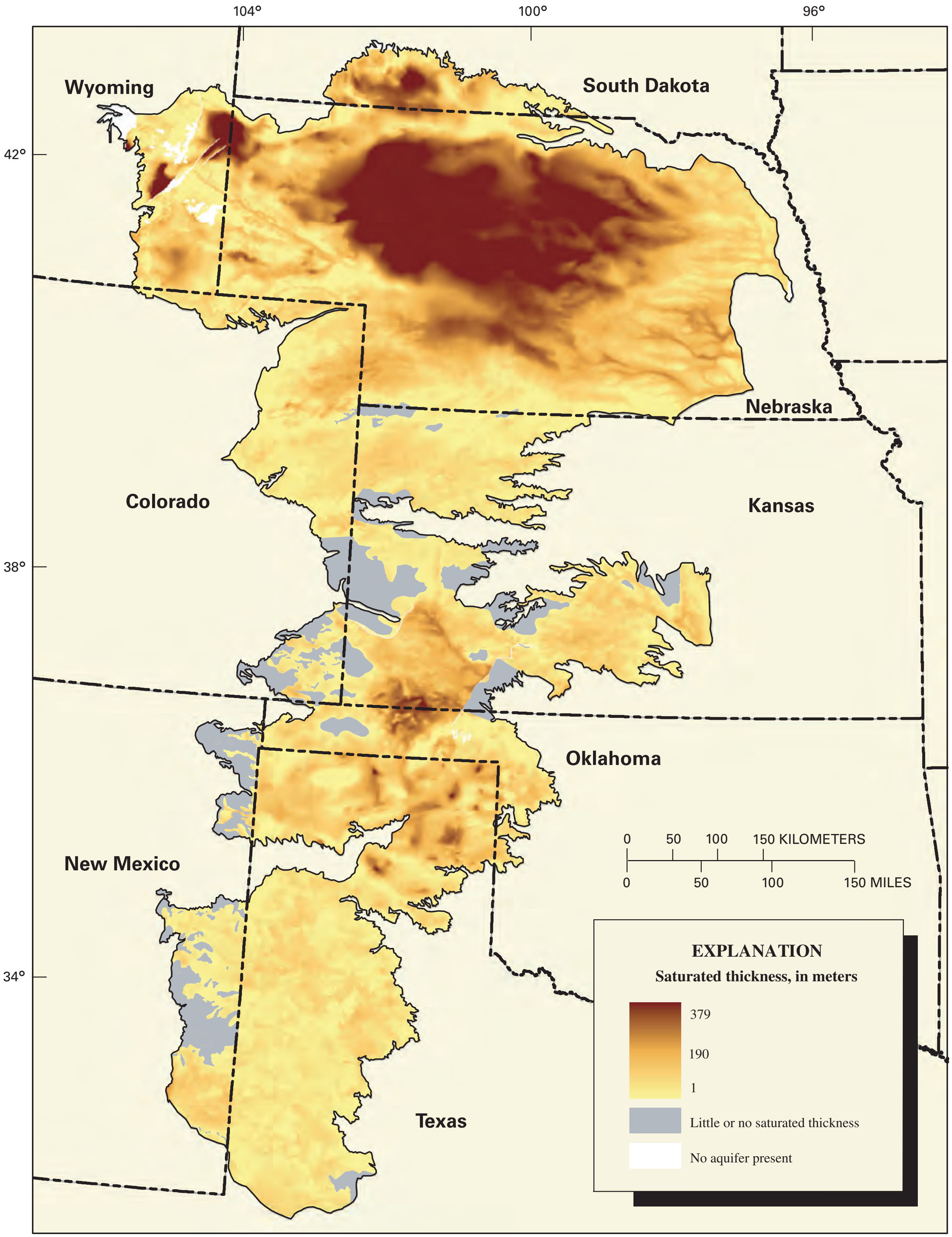
The horizontal hydraulic conductivity of the High Plains aquifer is dependent upon sediment size, shape, sorting, and degree of cementation of the aquifer material. Gutentag and others (1984) concluded that on a regional scale, the sediments that compose the aquifer are distributed randomly in the vertical plane, and they developed a procedure to estimate the hydraulic conductivity based on a thickness-weighted average aquifer characteristic, as determined from drillers' logs. Estimates of hydraulic conductivity range from less than 7.6 to greater than 91 m/d across the High Plains aquifer (fig. 3) (Gutentag and others, 1984).

Recharge to the High Plains aquifer occurs by infiltration of irrigation water, areally diffuse infiltration from precipitation, focused infiltration of stormwater runoff through streambeds and other topographic depressions, and upward movement of water from underlying aquifers (McMahon, 2001). McMahon and others (2006) provide detailed estimates of unsaturated zone water fluxes at discrete locations across the High Plains aquifer. Based on these estimates, approximate average rates of diffuse recharge are 100 mm/yr in the Northern High Plains, 50 mm/yr in the Central High Plains, and 10 mm/yr in the Southern High Plains. Discharge from the High Plains aquifer occurs primarily by well pumping, discharge to streams and underlying aquifers, ground-water flow across the eastern boundary of the aquifer, and evapotranspiration (McMahon, 2001).

Land and Water Use

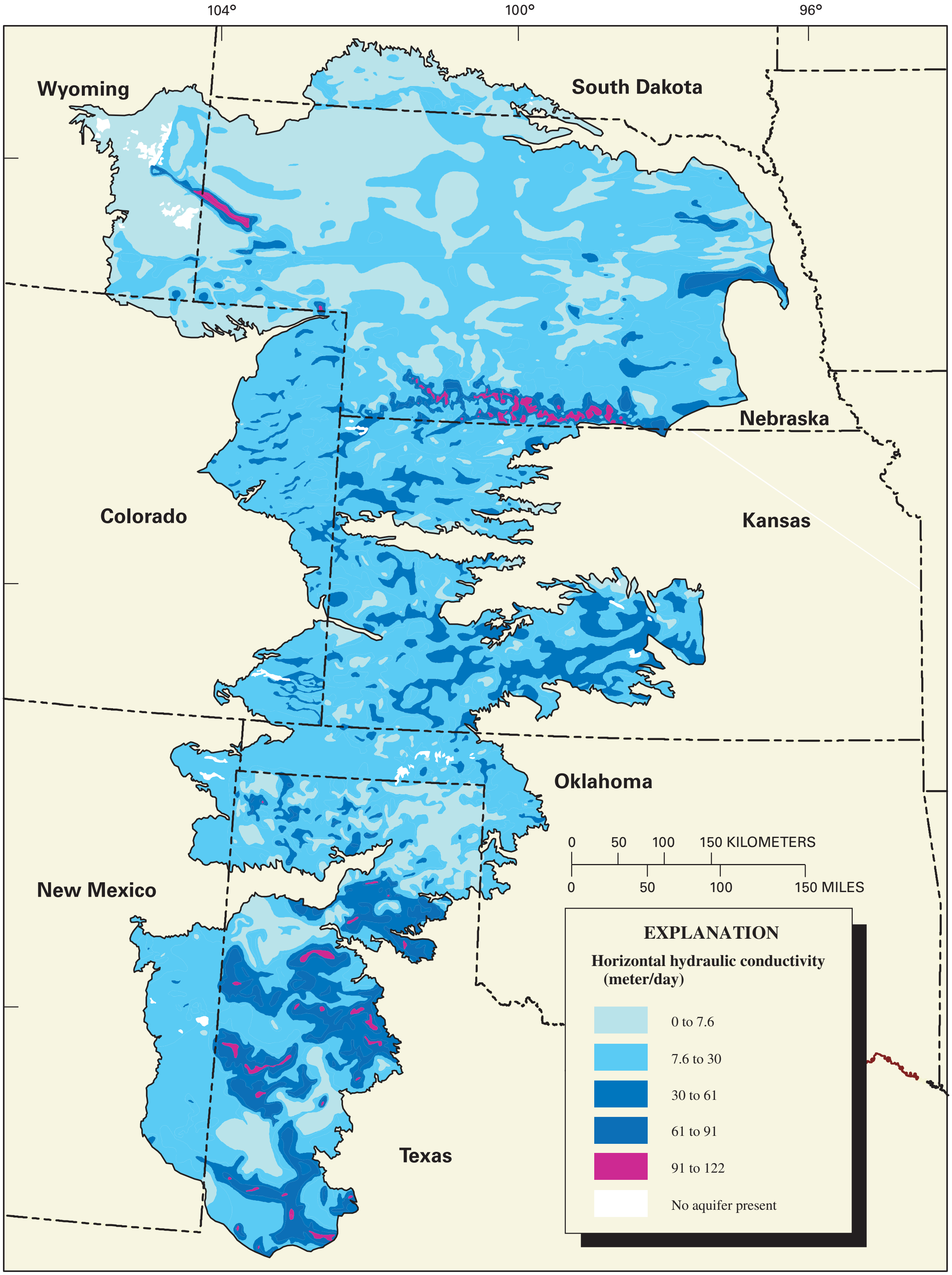
The gently sloping plains and small relief of the High Plains area make it ideal for rangeland and agriculture. The land in the study area is used primarily for grazing and the production of crops, including irrigated and nonirrigated (dryland) cropland. The main crops are alfalfa, corn, cotton, sorghum, and wheat. The High Plains also produces a substantial percentage of the feedlot beef raised in the United States; many of the crops grown there are used for animal feed at confined animal-feeding operations.

To support the agricultural-based economy, farmers have relied on ground water from the High Plains aquifer for irrigation. The majority of the High Plains has a middle latitude dry continental climate, with annual precipitation ranging from 30 cm/yr in the west to 86 cm/yr in the east. About 75 percent of the precipitation falls as rain during the growing season, April through September; much of the rains result from local thunderstorms, creating large variations in rainfall, spatially and temporally (Gutentag and others, 1984). Without irrigation, this semiarid climate supports natural vegetation of shortgrass prairie and only drought-resistant crops such as wheat (Dennehy and others, 2002). In the 1940's, farmers began extensive use of ground water for irrigation. The estimated irrigated acreage increased rapidly from 1940 to 1980 and did not change greatly from 1980 to 1997: 1949—8,498.7 km^2, 1980—55,443.9 km^2, and 1997—56,253.3 km^2 (McGuire and others, 2003). The irrigated land of the High Plains contains 128,720 registered irrigation wells (S.L. Qi, U.S. Geological Survey, written commun., 2004). By 1995, about 94 percent of the water pumped from the aquifer was used for irrigation (McGuire and others, 2003); accounting for more than 5.677×10^7 m^3/d (Qi and others, 2002).



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 2. Saturated thickness of the High Plains aquifer.



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 3. Distribution of hydraulic conductivity in the High Plains aquifer, modified from Cederstrand and Becker (1998).

Methods of Investigation

The methods used to develop models and maps showing the probability of nitrate concentrations exceeding 4 mg/L in recently recharged ground water of the High Plains aquifer are as follows. Ground-water flow and particle-tracking simulations were run using Visual MODFLOW 2.8 (Waterloo Hydrogeologic, 1999) to define depth criteria for selection of wells that intercept recently (less than 50 years) recharged ground water and to delineate the shape of the contributing areas for each well. All relevant and available data in electronic format were compiled for these wells and include ground-water-quality, anthropogenic, and hydrogeologic data. Using GIS and previously defined contributing areas, anthropogenic and hydrogeologic data were extracted for each well. These data were analyzed using SAS 8.02 (SAS Institute, 1999) statistical software. Using multivariate logistic regression analysis, several preliminary multivariate models with various combinations of explanatory variables were constructed and evaluated. The models that best predicted the occurrence of nitrate concentrations greater than or equal to 4 mg/L in recently recharged ground water were selected based on statistical significance, model fit, and predictive ability. An independent set of nitrate-concentration data from recently recharged ground water were used for model validation. Coefficients of the explanatory variables from the final multivariate models were entered into the GIS, and a probability map was created using map-algebra techniques. Corresponding uncertainty estimates were developed using Latin hypercube simulations and are presented as 90-percent prediction interval estimates for the final probability map. The specific details of ground-water flow and particle-tracking simulations, data compilation, statistical methods, and uncertainty estimation methods are discussed in the following sections.

Defining Recently Recharged Ground Water and Delineating Contributing Areas

The methods used to define recently recharged ground water and delineate contributing areas are described in the following section. As used in this report, the contributing area represents an area on land surface unique to each well that likely has provided recent recharge to the ground water and contributes water and dissolved constituents to the well of interest. Prior to statistical modeling, wells that produce predominantly recently recharged ground water were identified. The dependent variable (nitrate concentration) was compiled from water-quality data from these wells, and the independent variables (for example, land use and hydrogeology) were extracted from the corresponding contributing area or point location of the well. Nitrate-concentration data have been collected by various studies from wells with screens at various depths in the High Plains aquifer; therefore, this study needed to develop a method to select only the wells with screen depths that intercept ground water that has recharged during the last 50 years. For each well expected to intercept recently recharged ground water, an approximate contributing area was estimated. The outline of the contributing area or the well location was used in GIS to extract the independent variables used during the statistical modeling.

The hydraulic properties of the High Plains aquifer were used to determine which wells likely received recently recharged water and the size and location of the contributing area and, therefore, constrained the empirically based statistical models within the conceptual framework of the hydrogeologic model of the study area. Other statistically based ground-water vulnerability assessments have delineated the land cover potentially affecting ground water of interest using circular buffers of various radii (Rupert, 2003). A 500-m radius has been the most common radial length for the circular buffer method (Eckhardt and Stackelberg, 1995; Nolan and others, 2002). If the ground-water flow regime is not known, circular buffers provide an adequate first approximation of the contributing area. However, knowing the ground-water flow vector could reveal that circular buffers result in weaker correlations or introduce potential Type I errors (rejecting the null hypothesis when in fact true) or Type II errors (failing to reject the null hypothesis when in fact false) depending on the variability of land use with respect to ground-water flow direction and well location. Lorenz and others (2003) characterized land use near individual wells within an upgradient pie-shaped sector and determined that the sector method presented a more accurate estimate of land use affecting water quality in an individual well than the circular-buffer method. A comparison of the sector and circular-buffer method is illustrated in figure 4.

A modified version of the sector method was developed and used to delineate the contributing area of wells for this ground-water vulnerability assessment. For each well in the study, an upgradient 90-degree sector was established from regional GIS-based ground-water-flow maps. The radial length of each sector was determined based on hypothetical ground-water-flow modeling and particle-tracking simulations.

Simulation of Ground-Water Flow and Particle Tracking

All simulations of ground-water flow and particle tracking were developed in Visual MODFLOW 2.8, a fully integrated, three-dimensional, graphical-modeling environment. Visual MODFLOW incorporates the USGS three-dimensional finite-difference code MODFLOW-88 (McDonald and Harbaugh, 1988) and the particle-tracking program MODPATH (Pollock, 1994). Hypothetical ground-water-flow simulations were conducted across an idealized section of the aquifer. Representative boundary conditions and hydrogeologic parameters of the High Plains aquifer reported in the literature (Gutentag and others, 1984; Luckey and others, 1986; McMahon and others, 2006) were used to constrain the simulations. A sensitivity

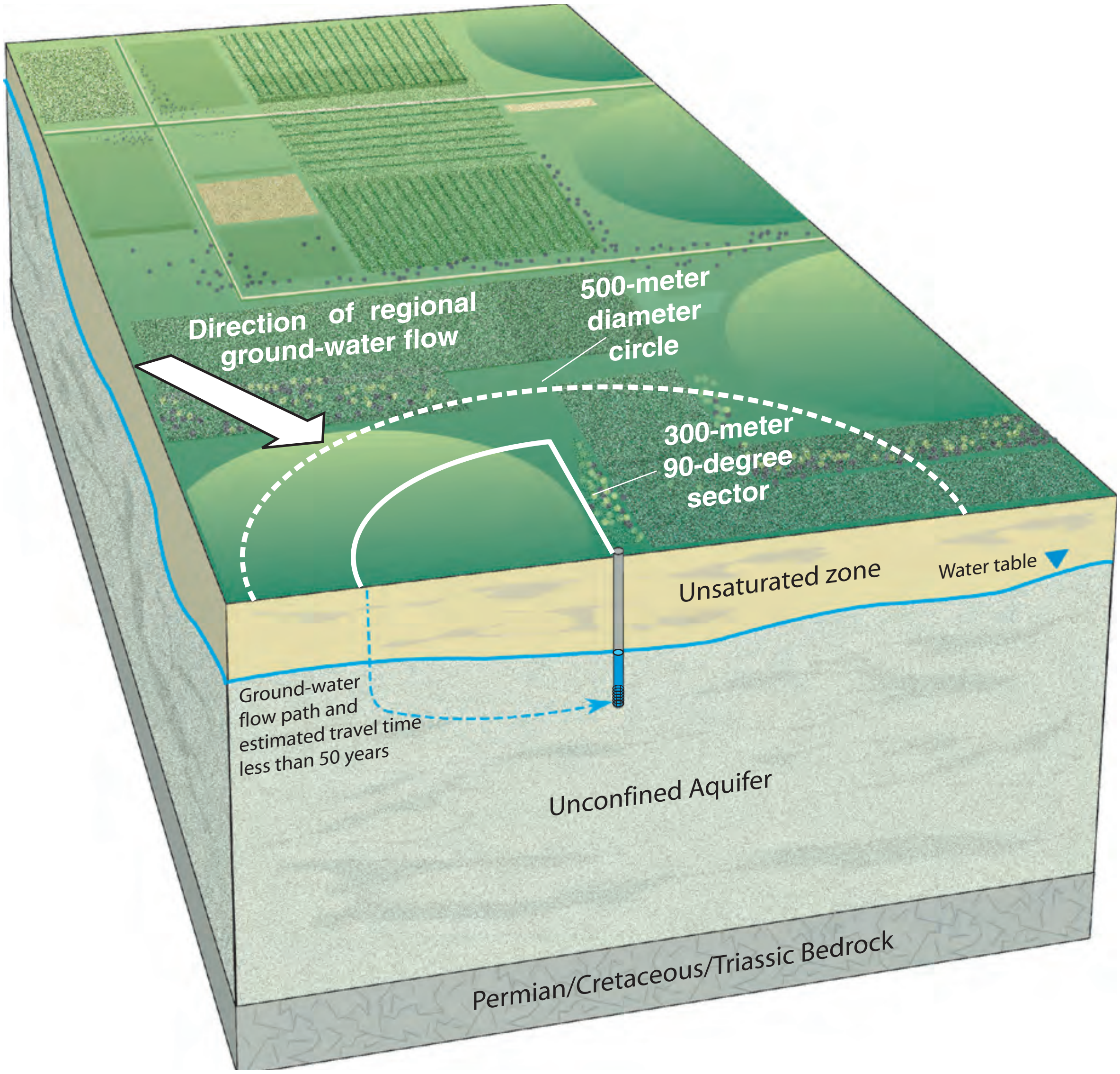


Figure 4. Conceptual diagram of sector method versus circular-buffer method for delineating land use of areas contributing recharge to a well.

analysis was conducted to evaluate model response to variations in reported boundary conditions and hydrogeologic parameters. This approach was favorable because of insufficient data required for model calibration at each monitoring well location considered for this study. Because simulations were run for a hypothetical, idealized system, model calibration was not required (Reilly and Harbaugh, 2004).

Simulating the flow paths of ground water in the aquifer system helps to conceptualize and quantify the sources of water to wells (Reilly and Harbaugh, 2004). Therefore, ground-water-flow simulation results were used as input for the particle-tracking simulations to estimate the advective path of nitrate transport within the saturated zone from the water-table expression of the contributing areas to the well screens. Results of the simulated horizontal-particle travel lengths represent appropriate sector radial lengths of contributing areas during a 50-year travel path. Results from the simulated vertical-particle travel represent appropriate well-screen depths that intercept recently recharged ground water (defined as recharge during the last 50 years).

Model Discretization, Boundary Conditions, and Model Stresses

A hypothetical idealized section of the High Plains aquifer is represented in the numerical model by a two-dimensional grid of cells that consists of 500 columns and 60 layers. The cells are 20 m in the horizontal dimension and 1 m in the vertical dimension. The 60 model layers in the vertical dimension are used to discretize the model in the vertical direction and allow for a refinement of flow paths near the water table surface.

Using the simplified model discretization of a hypothetical idealized section of the High Plains aquifer, hydrogeologic flow properties were systematically changed during each simulation to evaluate (1) likely particle-tracking distances in the horizontal and vertical directions; and (2) to determine which properties had the greatest influence on particle-tracking distances. The hydrogeologic flow properties were selected from the literature, see table 1, to bracket the estimated range of variability in the High Plains aquifer. All model simulations were run under steady-state flow conditions. Specified (constant) head boundary conditions were established at both ends of the grid to represent the flow and head constraints within the hypothetical flow domain. Recharge to the aquifer was modeled by applying the source to the top-grid layer of the model, which was kept active during all simulations. Recharge was specified as a uniform flux across the entire top-grid layer of the model. Estimates of recharge rates were based on measured water fluxes in the unsaturated zone using tritium and chloride mass balance methods at nine unsaturated zone installation sites located across the High Plains aquifer (McMahon and others, 2006). Variability in climate and other factors are assumed to be partially the reason for difference in measured water fluxes among the subregions of the High Plains; approximately averaging 10 mm/yr in the Southern High Plains, 50 mm/yr in the Central High Plains, and 100 mm/yr in the Northern High Plains (McMahon and others, 2006).

Although simulating ground-water flow and particle tracking across an idealized section of the aquifer was favorable because of insufficient data required for model calibration at each monitoring well location considered for this study, assumptions used in this approach can affect the accuracy of the simulation results. Assumptions used in model construction and aquifer properties are a simplified representation of a more complex aquifer system. Therefore, results from these simulations are approximations of particle-tracking distances and could be different if a more complex representation of the aquifer was used.

Model Simulation and Sensitivity Analysis

A forward particle-tracking method was used for all simulations. In this tracking method, one particle was started at the water table and tracked forward over a 50-year simulation. Porosity, horizontal hydraulic conductivity, and recharge rate were systematically varied to evaluate influence on model sensitivity, and the horizontal and vertical movement of the particle was calculated. This method was used to determine the maximum depth and horizontal travel within the saturated zone, which aided selection of the maximum well-screen depth and length of the contributing area that intercept recently recharged ground water. As a check against the results of the forward-tracking simulations, backward particle-tracking simulations also were run. Because of the lack of weak or strong sinks in the simulations, backward tracking produced nearly identical results as in the forward-tracking simulations and therefore are not presented in this report.

The tracking simulations indicate that horizontal particle movement is most sensitive to variability in hydraulic conductivity. For each simulation, uniform horizontal hydraulic conductivity was estimated as 7.6, 30, 61, or 91 m/d, while other hydrogeologic parameters remained constant. Results

Table 1. Summary of parameters in the ground-water particle-tracking model representing hypothetical cross sections of the High Plains aquifer.

[m, meter; mm/yr, millimeter per year; m/d, meter per day]

Model geometry	
Grid	Two-dimensional finite-difference grid (1 row × 500 columns; 1 m × 20 m).
Layers	60 layers (1 m) for vertical refinement of flow paths near water table.
Boundary conditions	
Base boundaries	No-flow cells.
Lateral boundaries	Constant head cells, representing the water table.
Recharge	Uniform flux across the entire top grid layer.
	Southern High Plains = 10 mm/yr; Central High Plains = 50 mm/yr; Northern High Plains = 100 mm/yr (McMahon and others, 2006).
Streams	Not evaluated during simulations.
Well	Not evaluated during simulations.
Hydraulic properties	
Horizontal hydraulic conductivity	Uniform horizontal hydraulic conductivity (7.6 – 91 m/d) (Gutentag and others, 1984).
Vertical hydraulic conductivity	Ratio of horizontal to vertical hydraulic conductivity = 10.
Hydraulic gradient	Average gradient of 0.001 (dimensionless) (Gutentag and others, 1984).
Porosity	Uniform porosity ranged from 0.2 to 0.4 across all layers. Reported by (McMahon and others, 2006).
Specific yield	Specific yield held constant at 10 (percent) for all simulations (Gutentag and others, 1984).

Table 2. Results of (A) horizontal and (B) vertical forward particle movement, in meters.

[Grey box represents particle movement in meters; m, meter; m/d, meter per day; mm/yr, millimeter per year]

(A) Horizontal Movement (m) of Forward Particle Tracking

Hydraulic gradient = 0.001; porosity = 0.2

Horizontal hydraulic conductivity (m/d)	Recharge (mm/yr) 10	50	100
7.6	655	670	625
30	2,900	3,014	1,850
61	5,600	5,500	6,200
91	8,350	8,345	8,338

Hydraulic gradient = 0.001; porosity = 0.4

Horizontal hydraulic conductivity (m/d)	Recharge (mm/yr) 10	50	100
7.6	370	330	275
30	1,375	1,100	1,370
61	2,785	2,980	3,145
91	4,170	3,985	3,840

(B) Vertical Movement (m) of Forward Particle Tracking

Hydraulic gradient = 0.001; porosity = 0.2

Horizontal hydraulic conductivity (m/d)	Recharge (mm/yr) 10	50	100
7.6	2.8	5.5	13.7
30	2.5	8.9	14.2
61	2.4	9.6	17.8
91	2.3	8.5	13.2

Hydraulic gradient = 0.001; porosity = 0.4

Horizontal hydraulic conductivity (m/d)	Recharge (mm/yr) 10	50	100
7.6	1.4	5.5	9.2
30	1.2	5.0	8.0
61	1.4	5.5	9.0
91	1.5	6.0	9.8

of the horizontal particle movement are shown in table 2A. Horizontal particle movement, in meters, was greatest for hydraulic conductivity equal to 91 m/d, porosity equal to 0.2, and recharge equal to 10 mm/yr. The smallest horizontal particle movement was simulated using horizontal hydraulic conductivity equal to 7.6 m/d, porosity equal to 0.4, and recharge equal to 100 mm/yr. To evaluate model sensitivity, the range in differences between the minimum and maximum movement (table 2A) were compared for each variable, while keeping the others constant. This evaluation revealed horizontal particle movement varied the greatest (range was approximately 3,500 to 7,700 m) depending on horizontal hydraulic conductivity. The horizontal particle movement was less sensitive to porosity (range was approximately 285 to 4,500 m) and recharge (range was approximately 12 to 1,160 m).

Modeling results of the horizontal particle movement were used to estimate the radial length of the 90-degree sector (corresponding to traveltime of 50 years) representing the approximate contributing area for each well across the study area. Using the more conservative modeling results, the approximate average horizontal particle movement (table 2A) for each of the four ranges (7.6, 30, 61, 91 m/d) of hydraulic conductivity for hydraulic gradient equal to 0.001 and porosity equal to 0.4 were incorporated into a GIS to select the appropriate radial length based on well location with respect to Cederstrand and Becker's (1998) estimates of hydraulic conductivity (fig. 3). The radial lengths are summarized in table 3. A detailed description of the techniques used in GIS to delineate sector length based on hydraulic conductivity is described below in Explanatory Variable Extraction.

The tracking simulations indicate that vertical particle movement is most sensitive to recharge. For each simulation, a uniform flux of recharge was estimated as 10, 50, or 100 mm/yr, while other hydrogeologic parameters remained constant. Results of the vertical particle movement are shown in table 2B. Vertical particle movement was greatest for recharge equal to 100 mm/yr, hydraulic conductivity equal to 61 m/d, and porosity equal to 0.2. The smallest vertical particle movement distance was simulated using recharge equal to 10 mm/yr, hydraulic conductivity equal to 30 m/d, and porosity equal to 0.4. The sensitivity analysis revealed that vertical particle movement varied the greatest (range was approximately 6.8 to 15.4 m) depending on recharge. The vertical particle movement was less sensitive to horizontal hydraulic conductivity (range was approximately 0.5 to 4.6 m) and porosity (range was approximately 0 to 8.8 m).

The results of the vertical particle movement were used to estimate an appropriate depth in the saturated zone that well screens must intercept to collect recently recharged ground water. This depth was used as selection criteria for wells that contain ground-water nitrate-concentration data to be used during the statistical modeling. For each subregional recharge estimate (10, 50, and 100 m/yr), the vertical particle movement was averaged then rounded to the nearest 1.5-m interval. This average depth is the well-screen depths criteria for selecting wells for statistical modeling and are as follows: Southern High Plains: 3 m, Central High Plains: 10 m, and Northern High Plains: 13.5 m below the estimated water table for 2000 from McGuire and others (2003). Using these depth-selection criteria for wells with screened intervals less than 15 m, 336 low-production monitoring and primarily domestic wells were selected for the study to represent recently recharged ground water across the High Plains aquifer.

Table 3. Summary of sector radius length averaged from particle-tracking simulations.

[m/d, meter per day; m, meter]

Horizontal hydraulic conductivity range[1] (m/d)	Sector radius length (m)
0 to 7.6	300
7.6 to 30	300
30 to 61	1,300
61 to 91	3,000
91 to 122	4,000

[1]Values used from Cederstrand and Becker (1998).

Using the tritium method, the apparent age of ground water was determined for selected wells identified as intercepting recently recharged ground water. The apparent ground-water age provides a check and an independent means to validate the results of the particle-tracking simulations and overall selection process of wells representing recently recharged ground water.

Tritium is a radioactive isotope of hydrogen that is present in some molecules of ground water and is derived from atmospheric concentrations in precipitation that recharged the aquifer. Although tritium is produced naturally by cosmic radiation, production of tritium was raised several orders of magnitude during the atmospheric testing of thermonuclear bombs. Before atmospheric testing began in the early 1950's, the tritium content of precipitation across the High Plains was probably on the order of 8 Tritium Units (TU) (Thatcher, 1962). Tritium is radioactive, with a half-life of 12.43 years. Therefore, ground water that is derived completely from precipitation that fell before the early 1950's atmospheric testing would contain less than 0.5 TU in 2004. However, water samples with original bomb tritium would contain larger tritium concentrations. Tritium values greater than 0.5 TU indicate that those ground-water samples contain at least a portion of water that was recharged during the last 50 years. The tritium value of 0.5 TU provides a convenient threshold to evaluate if wells that were selected by the particle-tracking simulations as intercepting recently recharged ground water actually contain ground water that recharged during the last 50 years.

Ground-water samples were analyzed for tritium at 82 wells (table 4) that were identified by the particle-tracking simulations as intercepting recently recharged ground water. All ground-water samples, except for 11, contained greater than or equal to 0.5 TU of tritium. This indicates that, of the wells selected by the particle tracking simulations and also containing ground water analyzed for tritium, greater than 86 percent were identified correctly as representing recently recharged ground water. The 11 ground-water samples that contained less than 0.5 TU were collected from wells in the Southern High Plains (6), Central High Plains (4), and Northern High Plains (1). All of these wells were screened within 3 m of the water table, except for one well in the Central High Plains that was screened within 5.5 m of the water table, and were located in areas with relatively thick unsaturated zones, ranging from 41 to 59 m. The nitrate concentrations in ground water from these 11 wells were less than 3.10 mg/L. The relatively thick unsaturated zone and lack of elevated nitrate and tritium in the ground water may indicate that recent recharge has not occurred at these 11 locations or that mixing with older ground water has diluted chemical signatures indicative of recently recharged water. Although the tritium-validation data set was available for only 82 wells, it is encouraging that a high percentage (86 percent) of wells were correctly selected by the ground-water flow and particle-tracking methods as intercepting recently recharged ground water in the High Plains aquifer.

Compilation of Ground-Water-Quality, Anthropogenic, and Hydrogeologic Data

Historical and recently collected ground-water-quality, anthropogenic, and hydrogeologic data were compiled and constructed for use by this study. The anthropogenic and hydrogeologic data were available in GIS format from a variety of sources and developed specifically for use as explanatory variables during the statistical modeling (table 5).

Ground-Water-Quality Data

A total of 336 measurements of nitrate concentration were compiled from samples of ground water from the 336 wells identified as producing recently recharged ground water during 1990 to 2004. All ground-water-quality data used for this study were (1) collected and reported by the High Plains NAWQA study (Pope and others, 2000; McMahon, 2001; Becker and others, 2002; Pope and others, 2002; Bruce and others, 2003; Fahlquist, 2003; McMahon, Böhlke, and Christenson, 2004; McMahon, Böhlke, and Lehman, 2004), and (2) compiled and reported by Litke (2001).

Ground-water-quality data from 125 of the 336 wells were collected by the High Plains NAWQA study. These samples were collected following NAWQA sampling procedures, as described by Koterba and others (1995), and were analyzed by the USGS National Water Quality Laboratory in Lakewood, Colorado. Quality-control samples were collected to assess the effectiveness of sample collection, processing, and analysis procedures for acquiring representative environmental data. Approximately 30 percent of the environmental samples collected were quality-control samples and included field blank and replicate environmental samples. The field blanks were used to verify that field and laboratory procedures did not contribute constituents of interest to the water samples and that decontamination protocols were adequate. Replicate samples were collected to assess the effects of field and laboratory procedures on measurement variability.

Nitrate data (number of wells equal to 211) were selected from Litke's (2001) retrospective analysis of historical ground-water quality of the High Plains aquifer to supplement the spatial coverage of data from the High Plains NAWQA study (number of wells equal to 125). Data were screened using Litke's (2001) data-quality-rating method, which was based on criteria related to sampling protocols, field quality-assurance practices, and laboratory analytical methods. The highest quality data were selected and have accepted or strict sampling protocols and field quality-assurance samples (field blanks and replicates) and were analyzed with USEPA-accepted laboratory methods. Additional screening procedures were applied. Wells were selected only if well-construction data, such as well-screen depth, were available. If time-series samples were collected at an individual well, the most recent ground-water sample was used during the vulnerability assessment. This process was to remove temporal variability caused by

Table 4. Summary of ground-water-quality data for ground-water vulnerability assessment.

[No., number; >, greater than; mg/L, milligrams per liter; <, less than; N, nitrogen; --, no available data; NHP, Northern High Plains; CHP, Central High Plains; SHP, Southern High Plains]

Constituent (reporting units)	No. of analyses/ No. of sites	No. of analysis >4-mg/L nitrate	No. of analysis <2-mg/L dissolved oxygen	Minimum value	25th-percentile value	Median value	75th-percentile value	Maximum value
All data								
Nitrate (mg/L as N)	336/336	162	--	0.02	2.10	3.82	7.09	31
Dissolved oxygen (mg/L)	140/336	--	30	0.10	5.03	6.69	7.56	10
NHP								
Nitrate (mg/L as N)	192/192	96	--	0.02	2.10	4.00	6.79	31.0
Dissolved oxygen (mg/L)	44/192	--	11	0.10	3.30	6.84	8.25	9.34
CHP								
Nitrate (mg/L as N)	91/91	39	--	0.02	2.30	3.66	7.56	26.99
Dissolved oxygen (mg/L)	61/91	--	5	0.31	5.44	6.70	7.40	10.00
SHP								
Nitrate (mg/L as N)	53/53	27	--	0.21	1.98	4.03	6.61	21.95
Dissolved oxygen (mg/L)	35/53	--	4	0.44	4.61	6.00	7.35	9.31
Subset for model calibration								
Nitrate (mg/L as N)	232/232	108	23	0.02	2.04	3.63	6.77	29.2
Subset for model validation								
Nitrate (mg/L as N)	104/104	54	7	0.02	2.73	4.23	7.96	31.0
Tritium (tritium units)	82/336	--	--	–0.3	1.0	4.5	14.4	110

Table 5. Summary of anthropogenic and hydrogeologic explanatory variables used for model development.

[No., number; %, percent; °, degree; kg, kilogram; USGS, U.S. Geological Survey; N, nitrogen; gal, gallon; d, day; cm, centimeter; m, meter; hr, hour; mm, millimeter; k, universal soil loss factor]

Explanatory variable	Description	Data source(s)	GIS extraction method	No. of samples	Minimum value	Median value	Maximum value
		Anthropogenic data					
Irrag	Irrigated agricultural land in sector (%)	S.L. Qi and others (2002)	90° sector	336	0.0	8.49	100
Nonirrag	Nonirrigated agricultural land in sector (%)	S.L. Qi and others (2002)	90° sector	336	0.0	31.5	100
Agland	Irrigated and nonirrigated agricultural land in sector (%)	S.L. Qi and others (2002)	90° sector	336	0.0	81.0	100
Nfert	County-based expenditures on commercial fertilizer (kg/county)	D.L. Lorenz, USGS, written commun. (2004)	point	336	7.46E+5	9.84E+6	2.46E+7
Nmanr	County-based expenditures on manure (kg/county)	D.L. Lorenz, USGS, written commun. (2004)	point	336	1.08E+5	3.10E+6	1.93E+7
Nfarmfert	County-based expenditures on fertilizer for agriculture (kg/county)	B.C. Ruddy and others (2006)	point	336	7.39E+5	9.56E+6	2.43E+7
Nnonffert	County-based expenditures on fertilizer for nonagriculture (kg/county)	B.C. Ruddy and others (2006)	point	336	89.0	3,016	1.25E+6
ManCon	County-based expenditures on manure from confined animals (kg/county)	B.C. Ruddy and others (2006)	point	336	5,400	1.47E+6	1.47E+7
ManUncon	County-based expenditures on manure from unconfined animals (kg/county)	B.C. Ruddy and others (2006)	point	336	0.0	1.29E+6	5.95E+6
Atmdep	Estimated atmospheric deposition of nitrogen (kg/county)	B.C. Ruddy and others (2006)	point	336	4,847	9,444	33,125
Nresfert	Estimated residual nitrogen in soil from commercial fertilizer (kg of N/acre)	K.J. Hitt, USGS, written commun. (2004)	point	336	0.0	9.82	30.0
Nresmanr	Estimated residual nitrogen in soil from manure (kg of N/acre)	K.J. Hitt, USGS, written commun. (2004)	point	336	0.0	10.6	34.1
Irrwells	Number of irrigation wells in sector	S.L. Qi, USGS, written commun. (2004)	90º sector	336	0.0	0.0	32.0
Withdrw	Water withdrawal, adjusted for irrigated land (gal/d/acre of irrigated land)	Solley and others (1998)	point	336	61.9	1,209	19,601
		Hydrogeologic data					
Precip	Average annual precipitation (cm)	Taylor and others (1997)	point	336	33.0	48.3	78.7
S_thick	Soil thickness (m)	STATSGO, U.S. Department of Agriculture (1991)	point	336	1.1	1.5	2.0
S_perm	Soil permeability (cm/hr)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.35	3.8	44.2
S_om	Soil organic matter content (% by weight)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.10	0.50	1.7
S_awc	Soil available water capacity (cm/cm)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.13	0.38	0.53
S_clay	Soil clay content (% of material less than 2 mm)	STATSGO, U.S. Department of Agriculture (1991)	point	336	2.8	23.9	43.7
S_k	Universal soil loss factor (k)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.11	0.31	0.42
S_drain	Soil drainage (1 = excessive drainage to 7 = very poorly drained)	STATSGO, U.S. Department of Agriculture (1991)	point	336	1.2	3.0	5.8
S_slope	Soil surface slope (%)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.0	1.6	20.2
S_ll	Soil liquid limit (% moisture by weight)	STATSGO, U.S. Department of Agriculture (1991)	point	336	7.6	31.9	51.5
S_hydric	Occurrence of hydric soils (0 = no hydric soils to 1 = all hydric soils)	STATSGO, U.S. Department of Agriculture (1991)	point	336	0.0	0.0	0.80
S_hydro	Soil hydrologic characteristics (1 = high infiltration, deep soils, excessively drained soils to 4 = very slow infiltration rates, soils are clayey or impervious	STATSGO, U.S. Department of Agriculture (1991)	point	336	1.1	2.1	4.0
S_flood	Annual flood frequency of soil (1 = greater than 50% to 4 = no annual flood)	STATSGO, U.S. Department of Agriculture (1991)	point	336	1.3	3.9	4.0
Uzclay	Average clay content of unsaturated zone (%)	Developed for this study	90º sector	336	0.0	57.4	89.0
Uzsand	Average sand content of unsaturated zone (%)	Developed for this study	90º sector	336	0.0	20.6	70.5
Dtw	Depth to regional water table (m)	V.L. McGuire, USGS, written commun. (2001)	point	336	0.0	25.9	103.6
Sathik	Aquifer saturated thickness (m)	S.L. Qi and others (2002); and V.L. McGuire, USGS, written commun. (2004)	point	336	0.49	23.9	308

changes in nitrate concentration, which Litke (2001) identified as increasing at varying rates in most of the hydrogeologic units of the High Plains aquifer. Nitrate-concentration data also were excluded if the water table dropped below the screened interval of the well after the sample-collection date. The screening procedures used to identify wells that intercept recently recharged ground water were discussed in Defining Recently Recharged Ground Water and Delineating Contributing Areas.

Methods described by Mueller and others (1995) were used to combine equivalent nitrogen species because nitrate data were collected and reported by different agencies that use different laboratory methods and sampling and reporting conventions. Concentrations of nitrate therefore are reported as milligrams per liter of nitrate as nitrogen (mg/L as N).

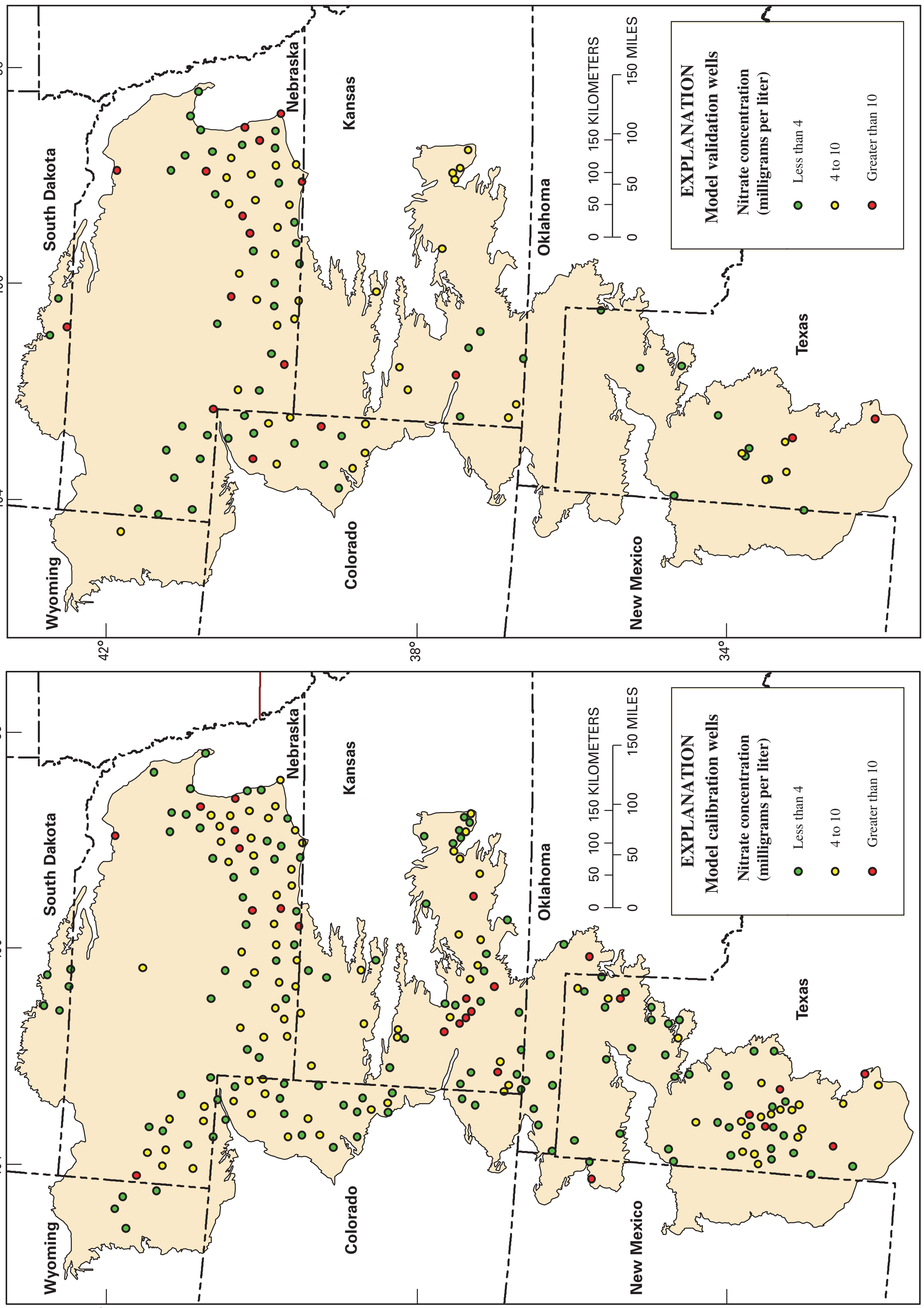
Concentrations of nitrate from the 336 samples ranged from 0.02 to 31.0 mg/L, with a median concentration of 3.82 mg/L (table 4). Approximately 12.7 percent (43 of 336) of the samples have nitrate concentrations that exceed the MCL of 10 mg/L, and 48.2 percent (162 of 336) of the samples have nitrate concentrations that exceed 4 mg/L. There are no apparent patterns in the spatial distribution of nitrate concentrations in ground water from the selected data set (fig. 5).

Establishing Relative Background Concentration of Nitrate

Nitrate in ground water of the High Plains aquifer is derived primarily from sources of nitrogen that originate at land surface and processes that mobilize these sources. Because nitrate is highly soluble in water and stable under oxic conditions, it is the dominant form of nitrogen in ground water of the study area. Minor forms of nitrogen in ground water of the High Plains aquifer include nitrite (NO_2^-) and ammonium (NH_4^+). Nitrogen sources include natural, agricultural, atmospheric, and urban sources. The predominant natural source of nitrogen includes the biological assimilation of atmospheric N_2 by nitrogen-fixing organisms. Natural nitrogen cycling (for example, ammonification and nitrification) can convert this organic nitrogen to nitrate. Natural source of nitrate is generally small and stable and often is referred to as the “background concentration.” Although natural sources tend to accumulate in the soil and subsoil, nitrate reservoirs can be mobilized to the deeper unsaturated zone and ground water when natural rangeland is converted to agricultural land (Walvoord and others, 2003; Bruce and others, 2003). The primary anthropogenic source is from the application of commercial fertilizer and manure to irrigated and nonirrigated agricultural land. Nitrification readily converts the ammonium-based fertilizers to nitrate, although substantial amounts can be lost to volatilization. The estimated average application rate of combined fertilizer and manure across the High Plains study area is 3,805 kg of N per km^2 per year (Litke, 2001). Atmospheric sources of nitrogen occur from precipitation and dry deposition and average about 392.9 kg of N per km^2 per year, equivalent to a concentration in precipitation of about 0.7 mg/L (National Atmospheric Deposition Program, 2000). Other anthropogenic sources of nitrogen include lawn fertilizers and septic systems; however, these sources are expected to be relatively minor in the High Plains because of low population density and predominance of agricultural land.

Because nitrate can occur from natural and anthropogenic sources, establishing a threshold based on nitrate concentration that distinguishes between the two sources is useful when quantifying ground-water vulnerability to nitrate. The term “background” concentration has been used to identify concentrations in ground water resulting from natural processes. However, defining the background based on concentration alone is subject to considerable uncertainty because any sample of ground water may reflect a mixture of waters with nitrate originating from a number of potential sources. Unless isotopic analyses on the nitrate in the sample are done to identify the source, a concentration-based background level may reflect nitrate from mixed natural and anthropogenic sources. Therefore, this report uses “relative background” concentration as a threshold to identify ground water that has been relatively unaffected by anthropogenic influence. The use of relative background also reflects the fact that most ground water in the study area, except for possibly the deepest parts of aquifer, likely has experienced human influence to some degree (McMahon, Böhlke, and Christenson, 2004). As defined here, ground water with nitrate below the relative background concentration has a nitrogen input from mostly natural sources (Nolan and Hitt, 2003). Establishing the relative background concentration can be complicated further by the fact that background nitrate concentrations are aquifer specific, and anthropogenic activity at land surface can mobilize sometimes relatively large concentrations of naturally occurring nitrate.

Previously established relative background and background concentrations of nitrate have a considerable range and indicate that natural inputs of nitrate are aquifer specific. Burkart and Stoner (2001) reported a range of background nitrate (as N) concentrations in ground water from 0.2 mg/L in Ohio (Baker and others, 1989) to as much as 100 mg/L in the Sahel of Africa (Edmunds and Gaye, 1997). Extensive analysis of nitrate-concentration data from across the United States by Madison and Brunett (1985) concluded that 3 mg/L may indicate possible human inputs. A background level of 3 mg/L may be too low from an environmental perspective because one-half of the samples compiled by Madison and Brunett (1985) across the United States do not have detectable levels of nitrate. Spalding and Exner (1993) suggest that from a regulatory standpoint, a more practical background level should coincide with the MCL for drinking water established by the USEPA. However, other investigators suggest background nitrate concentration is lower than 3 mg/L. Mueller and Helsel (1996) analyzed nitrate data from shallow ground water in forest and pasture areas of the United States and suggest that the background nitrate concentration is 2 mg/L. In predominantly rangeland settings of the United States,



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 5. Nitrate concentrations in ground water from wells used during vulnerability analysis of the High Plains aquifer, 1998–2004.

Nolan and Hitt (2003) reported 2.30 mg/L as the 75th percentile of nitrate concentration in shallow ground waters, which may indicate a relative background concentration within that particular land-use classification.

Background nitrate concentrations in the High Plains aquifer have been estimated using ground-water samples collected in the study area before the onset of modern (post-1950's) anthropogenic activity in the study area, and using samples of paleoground water that were recharged thousands of years before modern anthropogenic activity. Nitrate in paleoground water help determine a concentration threshold that represents naturally occurring nitrate in ground water because of natural mobilization. Litke (2001) compiled nitrate data from the 1930's to the 1960's and suggested the median nitrate concentration of 1.7 mg/L may represent background concentrations in the High Plains aquifer. However, McMahon, Böhlke, and Christenson, 2004 reported nitrate concentrations in paleorecharge (range = 2,600 – 12,800 ^{14}C years BP) with a narrow range (1.90 – 3.49 mg/L) and small average (2.7 mg/L) compared to nitrate concentrations in recent recharge under irrigated fields (range = 1.27 – 61.1 mg/L, average = 12.39 mg/L). The maximum range of nitrate concentrations in paleorecharge indicates that natural soil processes and mobilization produce nitrate concentrations in ground water of the High Plains aquifer that can range up to 3.49 mg/L. Therefore, a relative background concentration should reflect this nitrate range in paleorecharge.

For the purposes of this study, 4 mg/L was selected to represent the relative background nitrate concentration of ground water in the High Plains aquifer. A nitrate concentration of 4 mg/L is slightly larger than published estimates of background nitrate concentration and, therefore, reflects a conservative estimate of ground water influenced by anthropogenic activity. In addition, a nitrate concentration of 4 mg/L is of interest from a human-health perspective because of the findings by Ward and others (1996) that indicate long-term consumption of water containing nitrate (as N) greater than 4 mg/L increased the risk of non-Hodgkin's lymphoma. The relative background concentration of 4 mg/L was used as the binary threshold for the statistical modeling that is described in Statistical Methods.

Anthropogenic Data

Land-use classification data were obtained from two sources. The National Land Cover (NLCD) data set was used to determine the location of agricultural land (includes all cropland and pasture) across the High Plains aquifer. Qi and others (2002) used satellite imagery from Landsat Thematic Mapper (nominal date 1992) from the raw NLCD satellite data to classify irrigated and nonirrigated land. Ground-reference information from 2,500 km^2 was compared to the classified irrigated land data to determine an error estimate and percentage correct classification for irrigated land. Qi and others (2002) reported that the overall percentage correct for the irrigated land density map was 77.5 to 79.8 percent, and attribute much of this error to the anomalously wet conditions of eastern Nebraska. Land-use-classification explanatory variables used during this assessment include irrigated agricultural land (explanatory variable, Irrag), nonirrigated agricultural land (Nonirrag), and total agricultural land (Agland) (table 5).

Estimates of nitrogen applied as commercial fertilizer (Nfert) and as manure (Nmanr) were developed by David Lorenz (U.S. Geological Survey, written commun., 2001). Fertilizer was estimated on a county basis; statewide totals of fertilizer sales were prorated to each county based on expenditures by farmers, as reported in the 1997 Census of Agriculture. Similarly, animal manure estimates were developed from the 1997 Census of Agriculture for each county and were based on animal population and projected nitrogen content in amount of animal manure per day. The fertilizer and manure estimates recently were updated by Ruddy and others (2006) and include annual county-based expenditures of fertilizer apportioned by agricultural (Nfarmfert) and nonagricultural (Nnonffert) land use, annual manure estimates weighted by agricultural census reported populations of confined (ManCon) and unconfined (ManUncon) animals, and annual estimates of nitrogen from atmospheric deposition (Atmdep). Manure typically is applied to the land surface within 80 km of where it is generated because of transportation costs (Litke, 2001), and rates of commercial fertilizer application are higher on irrigated agricultural lands.

Estimates of residual nitrogen in the soil from fertilizer (Nresfert) and manure (Nresmanr) were obtained from Kerie Hitt (U.S. Geological Survey, written commun., 2004). These data were developed from the U.S. Department of Agriculture (USDA) Economic Research Service (ERS) compilation of estimates of excess nitrogen (N) in soil from 1990–93 National Agricultural Statistic Service (NASS) survey and study points. The following assumptions were used during compilation: credits were given for N from legumes in crop rotation; crop N uptake was estimated; crop N uptake was constant; no N export from farm in animals; manure application rates disaggregated to county level based on enterprise; manure rates constant within a county; manure N volatilization considered; and only estimates from cropland were considered (Kerie Hitt, U.S. Geological Survey, written commun., 2004).

The location of irrigation wells on the High Plains (Irrwells) was obtained by merging databases of registered wells maintained by the eight States within the High Plains area. Six databases reported that 99 percent of the total irrigation wells in each State were accounted for in the database. The two remaining databases (Texas and Wyoming) stated a known underestimate of the true number of irrigation wells in the respective States.

Water withdrawal from the High Plains aquifer (Withdrw) data were obtained from county estimates of ground-water withdrawal from the USGS water-use program (Solley and others, 1998).

Hydrogeologic Data

Estimates of average annual precipitation (Precip) for 1961–90 were obtained from Oregon State University **P**arameter-elevation **R**egressions on **I**ndependent **S**lopes **M**odel (PRISM) model (Taylor and others, 1997).

The soils data were obtained from Schwarz and Alexander (1995) who compiled variables from the STATSGO soils data (U.S. Department of Agriculture, 1991). The compiled data set contains weighted vertical averages of various soil characteristics that are present in the complete STATSGO database. A summary of the soil characteristics is listed in table 5.

Unsaturated-zone lithology data sets were created by the authors of this study using 56,000 lithologic logs from wells across the High Plains that were drilled deeper than the 2000 water table. The lithologic intervals described in the logs were compiled, simplified, and assigned terms of gravel, sand, clay, silt, or rock based on available descriptions from the logs. Each lithology was calculated as a percentage of the total unsaturated zone thickness. These data were then interpolated using a Kriging method to create GIS surfaces of percentage of each lithology in the unsaturated zone of the study area; percentages of sand and clay in the unsaturated zone are shown in figure 6. Explanatory variables tested during the vulnerability assessment include percentage of clay (Uzclay) and percentage of sand (Uzsand) in the unsaturated zone (table 5).

Estimates of the depth to regional water table (Dtw) were developed by Virginia McGuire (U.S. Geological Survey, written commun., 2001). Using a kriging method, an interpolated surface was created from wells of the High Plains water-level monitoring network (McGuire and others, 2003). The depth to water was calculated by subtracting the 2000 water-table elevation from the land-surface elevation. Saturated thickness (Sathik) then was calculated by subtracting the resulting depth to water surface from a surface representing the base of the aquifer. Virginia McGuire (U.S. Geological Survey, written commun., 2004) developed this surface by modifying an existing digital map of the base of the High Plains aquifer (Becker, 1998).

Explanatory Variable Extraction

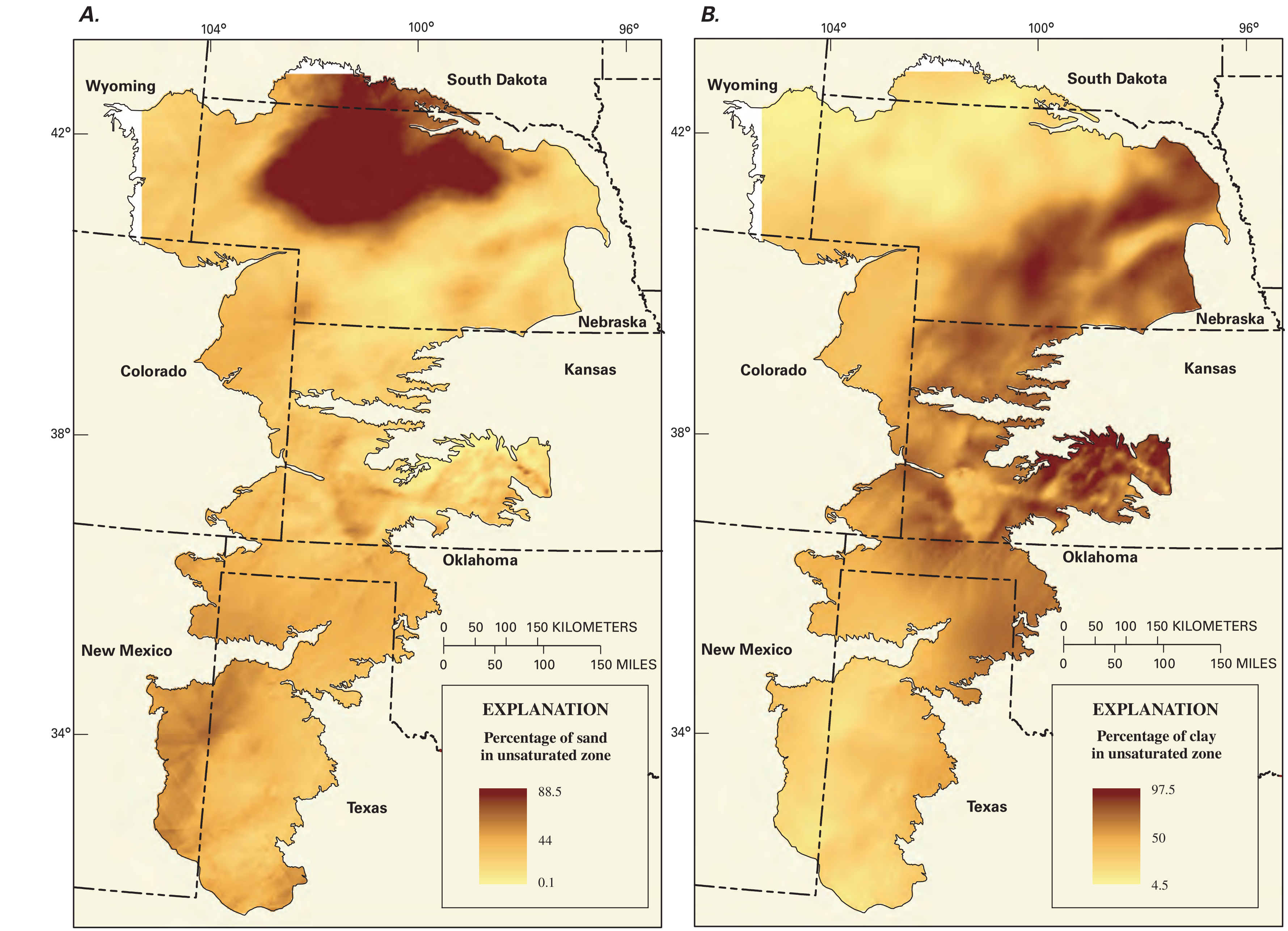
Prior to explanatory variable extraction, the 90-degree sector contributing areas for each of the 336 wells were delineated using a GIS. Programs were written in GIS to automate the delineation, which included orienting and sizing the contributing areas using well location with respect to the ground-water flow regime and estimated horizontal hydraulic conductivity. The orientation of the contributing area at each well location was determined using a 180-degree transformation of the azimuth direction of ground-water flow, calculated using the slope direction of the surface of the water table. The hydraulic-conductivity values were extracted for each well location using Cederstrand and Becker's (1998) estimated hydraulic-conductivity (fig. 3) GIS coverage. The program orients the center of the 90-degree sector upgradient (180-degrees) from the azimuth ground-water flow direction and simultaneously adjusts the size of the sector based on the hydraulic-conductivity point value and the corresponding radial length, summarized in table 3.

After delineation of the contributing areas, explanatory variables were extracted in GIS using either the 90-degree sector contributing area (for example, explanatory variables: Irrag, Nonirrag, Agland, Irrwells, and Uzclay) or the point location of the well (for example, explanatory variables: Withdrw, Precip, S_om, Dtw, and Sathik). Contributing area-based extractions were performed on all explanatory variables, except those with inappropriately large spatial resolutions, such as the soils data set (table 5). Point-based extractions were used on the remaining variables. All extracted explanatory data were exported from GIS into SAS (SAS Institute, 1999), where all statistical modeling was performed.

The 90-degree sector method provided a more representative delineation of the contributing area than approaches that use circular-buffer extraction methods because the sector radius size is based on regional ground-water flow regime and hydrogeologic properties of the aquifer. The Wilcoxon rank-sum test was used to evaluate if the data extracted by the sector method resulted in explanatory data sets with statistically different populations than corresponding explanatory variables extracted by the circular-buffer method. To make this comparison, at each of the calibration wells where explanatory variables were extracted by the sector method, the 500-m circular-buffer method also was used to extract selected corresponding explanatory data. These variables included percentage of agricultural land, percentage of irrigated agricultural land, and number of irrigation wells. Results of the Wilcoxon test indicate that the extracted data for percentage of agricultural land (Agland) (p-value = 0.051), percentage of irrigated agricultural land (Irrag) (p-value = 0.128), and number of irrigation wells (Irrwells) (p-value = less than 0.001) differ between the sector and circular-buffer methods, at the significance level of alpha (α) = 0.15. Lorenz and others (2003) corroborate these findings and suggest that the sector method provides a better correlation to nitrate concentrations in the ground water than the circular-buffer method.

Statistical Methods

Logistic regression analysis is a statistical method that predicts the probability of a binary or categorical response based on explanatory variables. Often, the objective of a ground-water vulnerability assessment is to predict the occurrence of the target water-quality constituent above a certain level or threshold. For this study, univariate and multivariate logistic regression analysis was used to predict the probability of detecting concentrations of nitrate in recently recharged ground water greater than or equal to the relative background concentration of 4 mg/L. Hosmer and Lemeshow (1989) present a thorough review of theory and application of logistic regression.



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 6. Percentage of sand (*A*) and (*B*) clay in the unsaturated zone of the High Plains aquifer.

Logistic regression is well suited for analysis of ground-water vulnerability because the binary response can be established using a threshold meaningful for specific management issues. Some examples of thresholds include the drinking-water standard, laboratory detection level, or relative background concentration. A major assumption of logistic regression is that the natural logarithm of the odds ratio is linearly related to the explanatory variables. The odds ratio (eq. 1) defines the probability of being in a response category

$$\text{Odds ratio} = \frac{p}{1 - p} \tag{1}$$

where

p is the probability of exceeding the established binary threshold value (Helsel and Hirsch, 1992).

The log of the odds ratio, or logit, transforms a variable constrained between 0 and 1 into a continuous variable that is a linear function of the explanatory variables (Helsel and Hirsch, 1992). The logit transformation is

$$\ln\left(\frac{p}{1 - p}\right) = b_o + bx \tag{2}$$

where

b_o is the logistic regression constant,

and

bx is the vector of slope coefficients and explanatory variables.

Predicted values of the response variable are converted back into probability units by using the logistic transformation, with the logistic regression model taking the form of

$$P = \frac{e^{(b_o + bx)}}{1 + e^{(b_o + bx)}} \tag{3}$$

where

P is the probability of the binary response event, defined here as detecting nitrate in ground water at a concentration greater than or equal to 4 mg/L,

and

e is the base of natural logarithm.

A number of statistical parameters are calculated using logistic regression. These parameters aid the modeler in deciding how well the overall model works, how important each of the explanatory variables are in the overall model, and if the form of the model appears to be correct (Menard, 2002). The predictive ability of the overall logistic regression model is of importance and also is evaluated.

The log-likelihood ratio (LLR), commonly called G statistic, measures the success and significance of the logistic regression model as a whole by comparing observed with predicted values (Hosmer and Lemeshow, 1989). The highest LLR indicates the most significant model, taking into account the degrees of freedom (number of explanatory variable) in the model. The p-values of the LLR indicate model significance of the model coefficients (null hypothesis is that slope = 0). Specifically, α of 0.05 indicates a significance level of 5 percent; α of 0.01 indicates a significance level of 1 percent.

Logistic regression model-fitting criteria used in this study include the partial likelihood ratio, percent correct (PC) responses, model sensitivity, and area under the ROC (Receiver Operating Characteristic) curve. Other commonly used model-fit criteria are the Akaike information criterion (AIC) and Schwartz criterion (SC); however, these are not applicable for comparisons between models with different numbers of explanatory variables. The partial likelihood ratio is similar to the LLR but is evaluated to determine the significance of adding one or more new variables to an existing multivariate logistic regression model (Helsel and Hirsch, 1992). A model with the addition of one new variable is more significant than the original model if the partial likelihood ratio is greater than the value of the chi-square distribution with degrees of freedom equal to one. The partial likelihood ratio was used exclusively during the iterative processes of the multivariate logistic regression analysis to select the explanatory variables that produce the best fitting model. Because of the large number of iterations, partial likelihood ratios and corresponding preliminary multivariate models are not listed in this report. The overall rate of correct classification, or percent correct (PC) responses, is the number of observed exceedances predicted by the model as exceedances, plus the number of observed nonexceedances predicted as nonexceedances, divided by the combined number of observed exceedances and nonexceedances (Hosmer and Lemeshow, 1989). Sensitivity is defined as the number of observed exceedances predicted as exceedances divided by the total number of observed exceedances. Higher values of PC and sensitivity indicate better fitting models. The area under the ROC curve, represented by the c statistic, is a measure of the model's ability to discriminate between ground-water samples with nitrate greater than or equal to 4 mg/L and those samples that do not. Hosmer and Lemeshow (1989) suggest that 0.7 less than c statistic less than 0.8 is acceptable discrimination.

Model calibration is evaluated using the degree of correspondence between the predicted probabilities of nitrate exceeding the threshold and the actual nitrate concentrations exceeding the threshold. The Hosmer-Lemeshow (HL) goodness-of-fit test statistic was used to evaluate the model calibration. For the HL test, the data are grouped into typically 10 deciles of risk or bins, with each bin containing approximately 10 percent of the total number of observations. The null hypothesis of the HL test is that the model fits the data; therefore, a higher HL p-value indicates a well-calibrated model (Hosmer and Lemeshow, 1989).

Model diagnostic statistics were evaluated for each individual observation to determine which observations were most poorly fit by the model. The Pearson residual statistic was used, which is the difference between the observed and estimated probabilities divided by the binomial standard deviation of the estimated probability (Menard, 2002). Therefore, a residual value equal to zero indicates that the probability of nitrate concentration exceeding the threshold at that particular well is exactly what would be expected based on observation, a positive residual indicates a higher probability than what is expected, and a negative residual indicates a lower probability than expected (Helsel and Hirsch, 1992).

Because the explanatory variables were reported in different units, coefficients were standardized after final model selection and prior to comparing the strength of the relations between individual explanatory variables and the response (nitrate greater than or equal to 4 mg/L). The advantage to standardized coefficients is that the relative impact and magnitude of effect of the explanatory variables can be directly compared. The standardized logistic regression coefficients are calculated using a standardization technique from Menard (2002)

$$b^* = b s_x r / s_y \tag{4}$$

where

- b^* is the standardized logistic regression coefficient,
- b is the unstandardized logistic regression coefficient,
- s_x is the standard deviation of the explanatory variable of interest,
- r is the square root of coefficient of determination,

and

- s_y is the standard deviation of the estimated logit (eq. 2).

Once standardized, a 1 standard deviation increase in the explanatory variable produces a b* standard deviation change in the probability of detecting nitrate concentration greater than or equal to 4 mg/L.

Multicollinearity is a major concern for multivariate logistic regression models and is the result of strong correlations between two or more explanatory variables. If strongly correlated explanatory variables are included in a multivariate logistic regression model, problems with the model may arise. Multicollinearity may inflate the variance of the parameter estimates, causing a lack of statistical significance of individual explanatory variables, even though the overall model may be strongly significant. Most importantly, incorrect conclusions about relations in the model may be drawn if multicollinearity is present because an unrealistic model coefficient sign or unstable slope coefficients may result. To detect multicollinearity, Pearson correlation coefficients and multicollinearity diagnostic statistics were examined during model development and selection. A Pearson's correlation coefficient greater than 0.7 indicates there is a strong correlation between two explanatory variables. However, if several explanatory variables are interdependent, correlation coefficients may not be sufficient to detect multicollinearity. The Tolerance and Variance Inflation Factor (VIF) were used as multicollinearity diagnostic statistics and are based on linear regression analysis of explanatory variables. The Tolerance is defined as $1 - r^2$, where r^2 is the coefficient of determination for the regression of one independent variable on all remaining independent variables (Allison, 1991; Menard, 2002). The VIF is equal to the reciprocal of the tolerance and describes how inflated the variance of coefficient is compared to what it would be if there were no multicollinearity (Allison, 1991). Although there are no formal thresholds to use for the Tolerance or VIF in detecting the presence of multicollinearity, Allison (1991) suggests that Tolerance values less than 0.4 (VIF greater than 2.5) may indicate the presence of multicollinearity.

Uncertainty Estimation

Uncertainty is inherent during ground-water vulnerability assessment (Loague and others, 1996; Gurdak and McCray, 2005). However, few published ground-water vulnerability assessments have accounted for the uncertainty, and to the knowledge of the authors of this report, no assessments that have used logistic regression and GIS methods have quantified uncertainty. The term *uncertainty* addresses the reliability or confidence surrounding the estimate of vulnerability, expressed as predicted probabilities. The probability values reported in this report are estimates, approximating the true probability of detecting nitrate greater than or equal to 4 mg/L in recently recharged ground water. Therefore, the reported probabilities have prediction *error*, defined as the difference between the true and estimated probability. Because the true probabilities and, therefore, prediction error, are never known exactly, uncertainty represents the magnitude of this difference.

Prediction errors (uncertainty) are a function of data error from GIS-based explanatory variables (expressed as "x" in eq. 3) and model error of estimated logistic regression coefficients (expressed as "b_o" and "b" in eq. 3). The accuracy and precision of GIS databases as representations of the world are never perfect (Longley and others, 2001). Therefore, GIS-based explanatory variables inherently introduce data error into logistic regression models. In addition, logistic regression coefficients are subject to estimation error (van Horssen and others, 2002). Both sources of error propagate through the logistic regression model to produce a combined prediction error (uncertainty) in the model output, expressed in this report as a vulnerability map.

To estimate uncertainty associated with output of the logistic regression based vulnerability model and map, a stochastic modeling approach additionally was used. Latin hypercube sampling (McKay and others, 1979), which is a constrained variation on the standard Monte Carlo stochastic sampling method, was used to illustrate the propagation of input error (explanatory variable data error and logistic regression coefficient model error) through the logistic regression models. During Latin hypercube sampling, values are randomly drawn from predefined input probability distributions, which represent input data and model errors. Sampling is done repetitively, with one sample drawn every iteration from each input probability distribution. Latin hypercube sampling uses a stratified sampling routine that forces samples drawn to correspond more closely with input distribution and, thus, converges faster on the true statistics of the input distribution (Palisade Corporation, 2002).

For this study, the results of the Latin hypercube sampling were expressed as 90-percent prediction intervals for the final probability models. This prediction interval defines the error range from the lower 5th to upper 95th percentile of the output probability distribution, representing a 90-percent prediction interval that contains the true probability of detecting nitrate greater than or equal to 4 mg/L. Because the propagated model output error will likely be spatially variable (Phillips and Marks, 1996), the 90-percent prediction intervals were represented as uncertainty maps to accompany the final vulnerability map of the High Plains aquifer. The uncertainty maps assisted a comparative analysis of spatial uncertainty for the vulnerability map. For example, areas of the aquifer with relatively larger 90-percent prediction interval have greater uncertainty surrounding the estimated vulnerability, compared to areas of the aquifer with smaller 90-percent prediction interval.

Estimating the Probability of Nitrate in Ground Water Exceeding the Relative Background Concentration

Using logistic regression, statistical models were developed that predict the probability of detecting nitrate concentrations exceeding 4 mg/L in recently recharged ground water of the High Plains aquifer. Model development included a univariate and multivariate logistic regression analysis. Evaluation and selection of two final multivariate models was based on statistical significance, model fit, and predictive ability. The two final models represent the Northern High Plains (NHP model) and the combined Central and Southern (CHP and SHP model) High Plains. Nitrate data from a random subset of wells was selected and used for the validation. Using the final models, probabilities were calculated at the location of each validation well, where the percentage of actual nitrate detection was compared to predicted probabilities and evaluated using r-squared values. The two final models were combined in GIS to create a map illustrating the probability of detecting nitrate concentrations exceeding 4 mg/L in recently recharged ground water of the High Plains aquifer.

Development of Nitrate Model

Univariate relations between nitrate concentration greater than or equal to 4 mg/L and explanatory variables were evaluated and are summarized in table 6. The coefficients listed in table 6 indicate the nature of the univariate relation; coefficient values greater than zero indicate positive relations, and coefficient values less than zero indicate inverse relations with nitrate greater than or equal to 4 mg/L. An alpha level of 0.2 was chosen as the inclusion criteria for selecting explanatory variables into the multivariate analysis rather than the more traditional alpha level of 0.10. Hosmer and Lemeshow (1989) suggest that an alpha level of 0.10 has failed to identify variables known to be important during some multiple logistic regression analyses. Twenty-one of the 31 explanatory variables initially were carried forward for multivariate analyses. However, all explanatory variables were evaluated later using the partial likelihood ratio during multivariate analyses. The variable selection for multivariate model development required too many iterative steps to list in this report. Details of the final two multivariate logistic regression models (NHP model and CHP and SHP model) are presented below.

Two final multivariate logistic regression models representing the probability of detecting nitrate exceeding 4 mg/L in recently recharged ground water were constructed. The NHP model is expressed as:

$$P_{NHP} = \frac{e^{(-0.374 + (0.023*\mathrm{Nonirrag}) + (0.017*\mathrm{Irrag}) + (-1.487*\mathrm{S\_om}))}}{1 + e^{(-0.374 + (0.023*\mathrm{Nonirrag}) + (0.017*\mathrm{Irrag}) + (-1.487*\mathrm{S\_om}))}} \quad (5)$$

where

P_{NHP} is the probability of detecting nitrate greater than or equal to 4 mg/L in recently recharged ground water of the NHP aquifer;

Nonirrag is percentage of nonirrigated agricultural land in sector;

Irrag is percentage of irrigated agricultural land in sector;

and

S_om is soil organic matter content.

The probability of detecting nitrate exceeding 4 mg/L in recently recharged ground water of the CHP and SHP model is expressed as:

$$P_{\mathrm{CHP\ and\ SHP}} = \frac{e^{(1.158 + (-0.010*\mathrm{Dtw}) + (0.013*\mathrm{Nonirrag}) + (0.011*\mathrm{Irrag}) + (-0.019*\mathrm{Uzclay}))}}{1 + e^{(1.158 + (-0.010*\mathrm{Dtw}) + (0.013*\mathrm{Nonirrag}) + (0.011*\mathrm{Irrag}) + (-0.019*\mathrm{Uzclay}))}} \quad (6)$$

where

$P_{\mathrm{CHP\ and\ SHP}}$ is the probability of detecting nitrate greater than or equal to 4 mg/L in recently recharged ground water of the CHP and SHP aquifer;

Dtw is depth to regional water table;

Nonirrag is percentage of nonirrigated agricultural land in sector;

Irrag is percentage of irrigated agricultural land in sector;

and

Uzclay is average percentage of clay in the unsaturated zone of the sector.

Table 6. Results of univariate logistic regression analyses, listing logistic regression coefficients, and individual p-values of independent variables related to the detection of nitrate in ground water greater than or equal to 4 milligrams per liter.

[Values not enclosed in parentheses are logistic regression coefficients; values enclosed in parantheses are p-values; bold were initially selected for multivariate analysis; %, percent; kg, kilogram; N, nitrogen; gal, gallon; d, day; cm, centimeter; m, meter; hr, hour; mm, millimeter; k, universal soil loss facter]

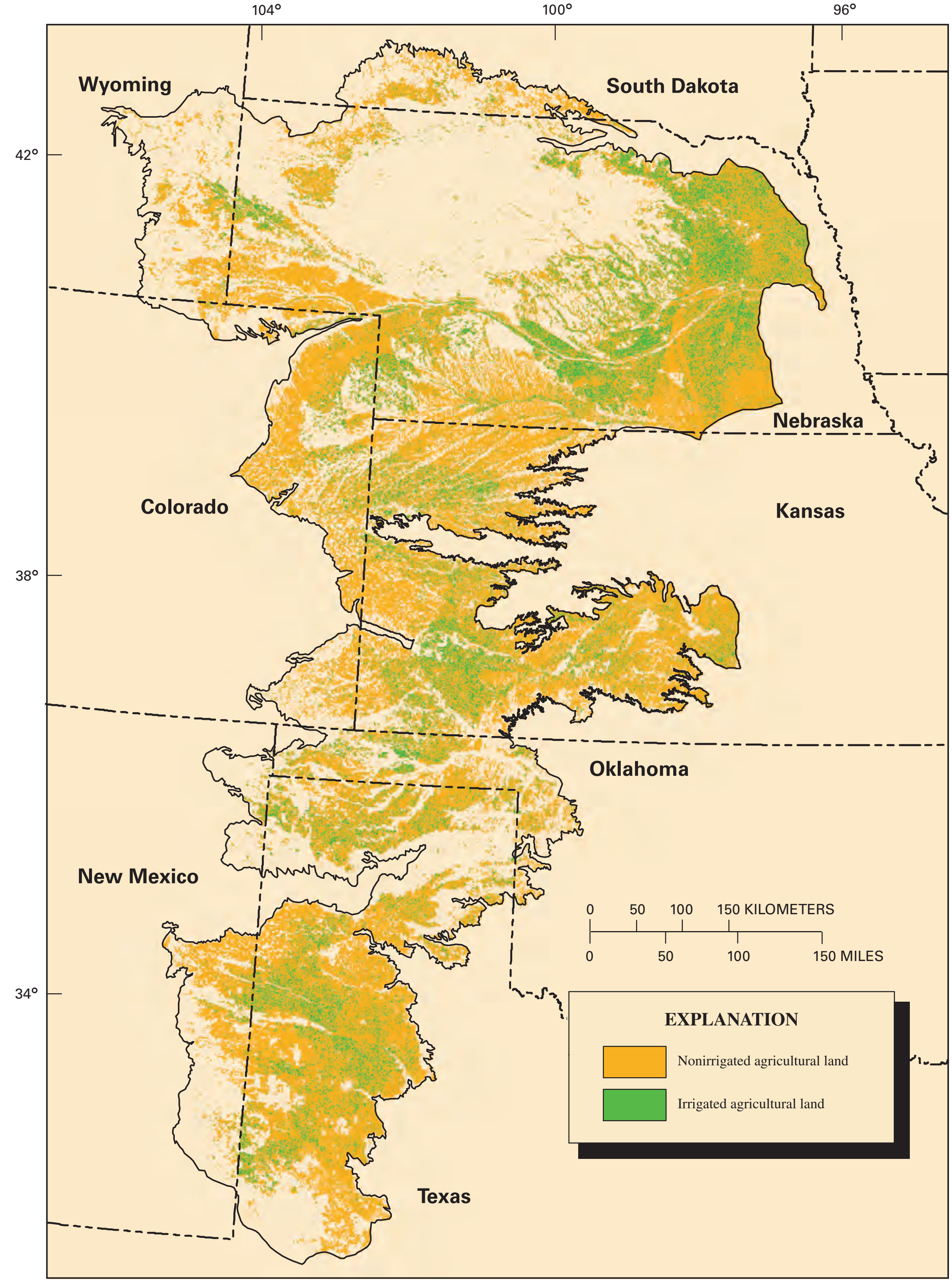
Explanatory variable		Northern High Plains Model	Central and Southern High Plains Model
	Anthropogenic		
Irrag	Irrigated agricultural land in sector (%)	**0.007 (0.200)**	0.001 (0.859)
Nonirrag	Nonirrigated agricultural land in sector (%)	**0.012 (0.025)**	**0.010 (0.065)**
Agland	Irrigated and nonirrigated agricultural land in sector (%)	**0.016 (0.002)**	**0.010 (0.069)**
Nfert	County-based expenditures on commercial fertilizer (kg/county)	–2.55E-08 (0.425)	**5.10E-8 (0.125)**
Nmanr	County-based expenditures on manure (kg/county)	**–1.12E-7 (0.086)**	**–6.27E-8 (0.170)**
Nfarmfert	County-based expenditures on fertilizer for agriculture (kg/county)	–2.43E-08 (0.465)	**5.38E-08 (0.121)**
Nnonffert	County-based expenditures on fertilizer for nonagriculture (kg/county)	**–3.00E-05 (0.122)**	5.06E-07 (0.674)
ManCon	County-based expenditures on manure from confined animals (kg/county)	–6.22E-08 (0.482)	–7.07E-08 (0.215)
ManUncon	County-based expenditures on manure from unconfined animals (kg/county)	**–4.87E-07 (0.004)**	**–2.78E-07 (0.119)**
Atmdep	Estimated atmospheric deposition of nitrogen (kg/county)	**–7.00E-05 (0.026)**	**–7.00E-05 (0.100)**
Nresfert	Estimated residual nitrogen in soil from commercial fertilizer (kg of N/acre)	0.011 (0.616)	–0.009 (0.749)
Nresmanr	Estimated residual nitrogen in soil from manure (kg of N/acre)	0.010 (0.650)	**–0.016 (0.121)**
Irrwells	Number of irrigation wells in sector	–0.017 (0.699)	0.089 (0.548)
Withdrw	Water withdrawal, adjusted for irrigated land (gal/d/acre of irrigated land)	**2.64E-04 (0.170)**	–1.70E-04 (0.363)
	Hydrogeology		
Precip	Average annual precipitation (cm)	0.046 (0.251)	0.029 (0.549)
S_thick	Soil thickness (m)	0.064 (0.457)	0.005 (0.851)
S_perm	Soil permeability (cm/hr)	**–0.078 (0.131)**	0.016 (0.824)
S_om	Soil organic matter content (% by weight)	–0.314 (0.670)	–0.525 (0.440)
S_awc	Soil available water capacity (cm/cm)	**8.968 (0.055)**	1.739 (0.797)
S_clay	Soil clay content (% of material less than 2 mm)	**0.036 (0.109)**	**–0.033 (0.166)**
S_k	Universal soil loss factor (k)	**5.653 (0.017)**	–1.613 (0.605)
S_drain	Soil drainage	–0.024 (0.935)	0.131 (0.832)
S_slope	Soil surface slope (%)	–0.019 (0.687)	–0.054 (0.502)
S_ll	Soil liquid limit (% moisture by weight)	**0.043 (0.057)**	**–0.050 (0.095)**
S_hydric	Occurrence of hydric soils	–1.804 (0.422)	3.418 (0.440)
S_hydro	Soil hydrologic characteristics	–0.134 (0.749)	**–0.455 (0.171)**
S_flood	Annual flood frequency of soil	**0.638 (0.056)**	–0.478 (0.609)
Uzclay	Average clay content of unsaturated zone in sector (%)	**0.015 (0.083)**	–0.005 (0.635)
Uzsand	Average sand content of unsaturated zone in sector (%)	**–0.038 (0.033)**	–0.012 (0.891)
Dtw	Depth to regional water table (m)	3.94E-04 (0.892)	**–0.007 (0.006)**
Sathik	Aquifer saturated thickness (m)	–2.00E-05 (0.986)	0.002 (0.427)

The results of the final multivariate model analysis are summarized in table 7 and reveal statistical significance with good model fit. Explanatory variables included in the NHP model are percentage of nonirrigated agriculture (Nonirrag) (fig. 7) in contributing area, percentage of irrigated agriculture (Irrag) (fig. 7) in contributing area, and organic matter in the soil (S_om) (fig. 8). Explanatory variables included in the CHP and SHP model are depth to regional water table (Dtw) (fig. 9), percentage of nonirrigated agriculture (Nonirrag) (fig. 7) in contributing area, percentage of irrigated agriculture (Irrag) (fig. 7) in contributing area, and average percentage of clay in the unsaturated zone of the contributing area (Uzclay) (fig. 6*B*). The log-likelihood ratio and p-values for the NHP model (LLR = 13.9, p = 0.003) and CHP and SHP model (LLR = 16.2, p = 0.003) indicate high statistical significance (table 7). Model fit was good, indicated by reasonable percent correct (NHP model = 65.8 percent; CHP and SHP model = 70.4 percent) and sensitivity (NHP model = 72.4 percent; CHP and SHP model = 58.0 percent). The final models have acceptable discrimination (both models: c = 0.7). The overall model fit was excellent (HL p-value: NHP model = 0.989; CHP and SHP model = 0.959). Linear regressions were constructed between the percentage of observed detections of nitrate concentrations exceeding 4 mg/L in recently recharge ground water and the average predicted probabilities for deciles calculated with the NHP (eq. 5) and CHP and SHP models

Table 7. Results of multivariate logistic regression analyses and listing of logistic regression coefficients, p-values, and standardized coefficients.

[N, number of observations; LLR, log-likelihood ratio; b*, standardized coefficient; HL, Hosmer Lemeshow goodness-of-fit; r^2, coefficient of determination; PC, percent correct; %, percent; c, area under the Receiver operating characteristic curve; NHP, Northern High Plains, CHP, Central High Plains; SHP, Southern High Plains; m, meter]

Multivariate model	N	Explanatory variables			LLR (p-value)	Logistic regression constant	Explanatory variable coefficient (p-value)	Standardized coefficients (b*)	Model fit				
		Name	Range	Median					HL (p-value)	r^2	PC (%)	Sensitivity (%)	c
NHP	117				13.9 (0.003)	–0.374			0.989	0.910	65.8	72.4	0.7
		Nonirrigated agricultural land in sector (%)											
		(Nonirrag)	0 to 100	25.3			0.023 (0.001)	0.328					
		Irrigated agricultural land in sector (%)											
		(Irrag)	0 to 100	7.78			0.017 (0.009)	0.237					
		Soil organic matter content (% by weight)											
		(S_om)	0.27 to 1.69	0.60			–1.487 (0.079)	–0.155					
CHP and SHP	115				16.2 (0.003)	1.158			0.959	0.891	70.4	58.0	0.7
		Depth to regional water table (m)											
		(Dtw)	2.27 to 96.3	40.8			–0.010 (0.002)	–0.332					
		Nonirrigated agricultural land in sector (%)											
		(Nonirrag)	0 to 100	35.0			0.013 (0.043)	0.186					
		Irrigated agricultural land in sector (%)											
		(Irrag)	0 to 100	6.96			0.011 (0.122)	0.145					
		Average clay content of unsaturated zone (%)											
		(Uzclay)	17.1 to 86.3	50.3			–0.019 (0.122)	–0.138					



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 7. Irrigated and nonirrigated lands of the High Plains aquifer.

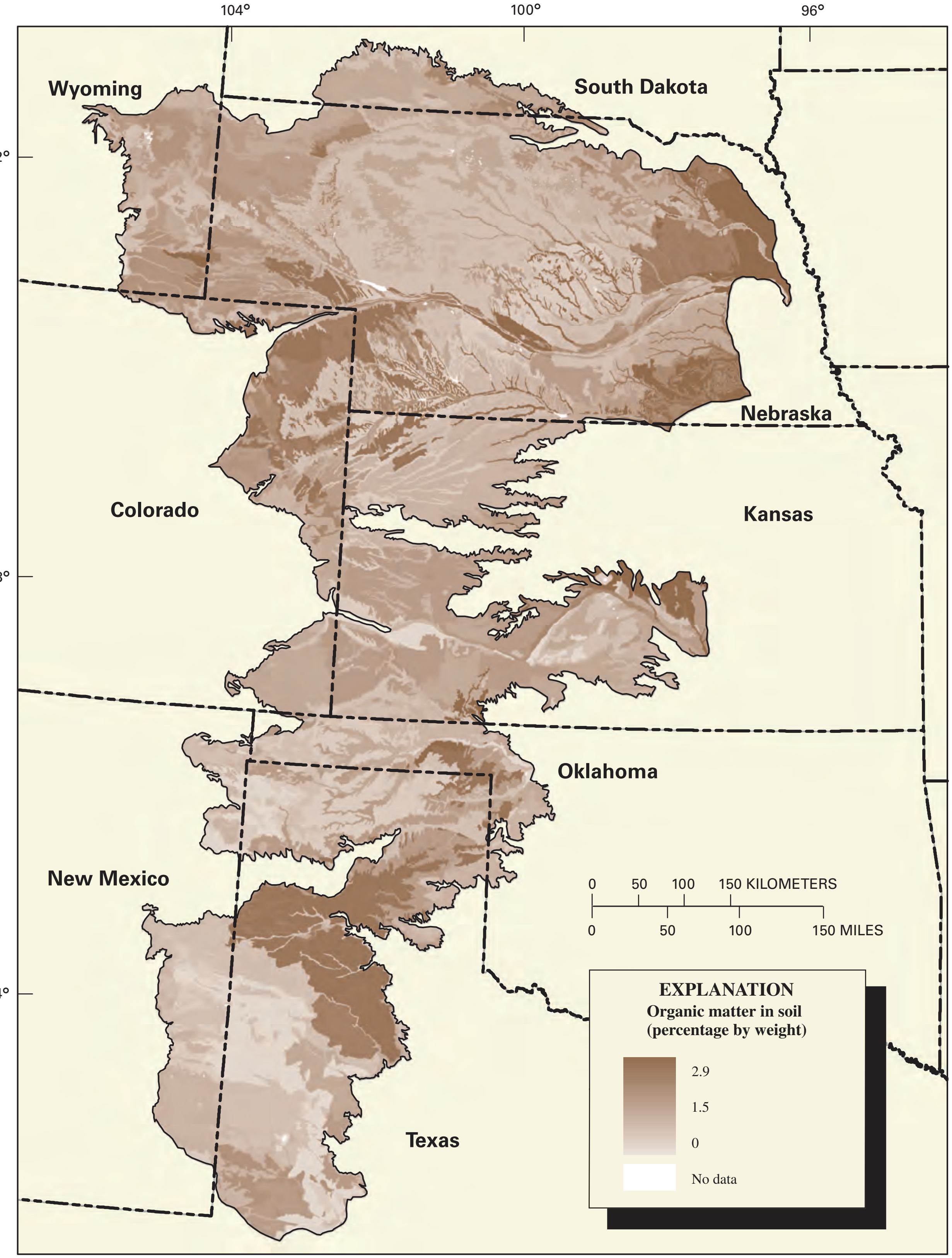
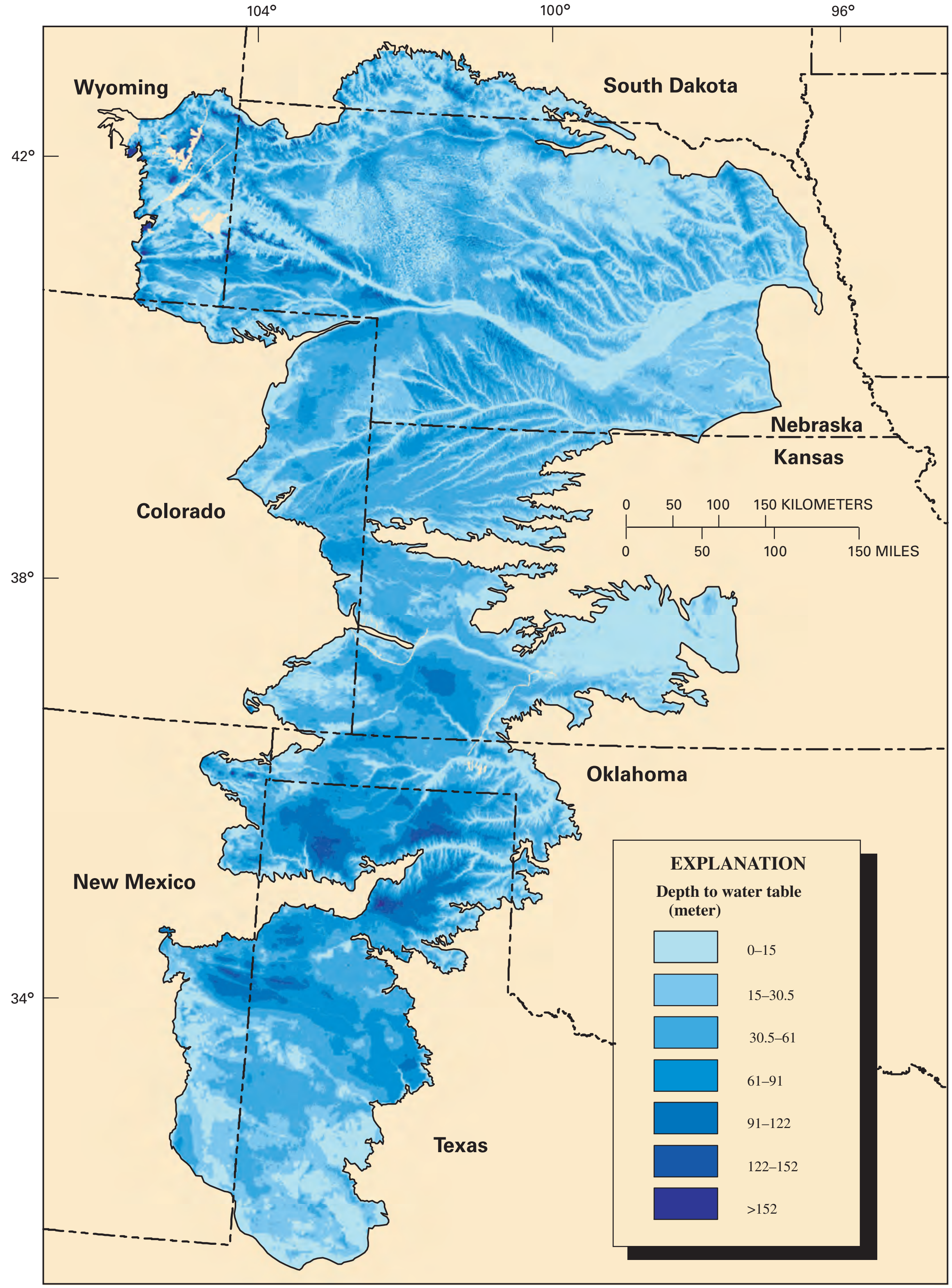


Figure 8. Organic matter in the soil of the High Plains aquifer.



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 9. Depth to regional water table of the High Plains aquifer.

(eq. 6). Actual nitrate concentration data were converted to binary classification of “zero” for nitrate concentrations less than 4 mg/L and “one” for nitrate concentrations equaling or exceeding 4 mg/L. This binary conversion allowed for the percentage of actual detections to be compared to predicted probabilities. The resulting regression (fig. 10) confirms the excellent model calibration, with r^2 value of 0.910 (NHP model) and 0.891 (CHP and SHP model). Diagnostic statistics indicated a lack of multicollinearity for all explanatory variables in both final models (table 8). To identify wells for which the models worked poorly, an analysis of residuals was performed using Pearson residual, which is the difference between the observed and estimated probabilities divided by the standard deviation of the estimated probability (Menard, 2002). The resulting Pearson residuals followed a normal distribution with the mean approximately equal to 0 and standard deviation equal to 1. The lack of large positive (greater than 2) or negative (less than –2) Pearson residual values indicates that the models (eqs. 5 and 6) are correct and fit all calibration data exceptionally well, with no apparent extreme outliers.

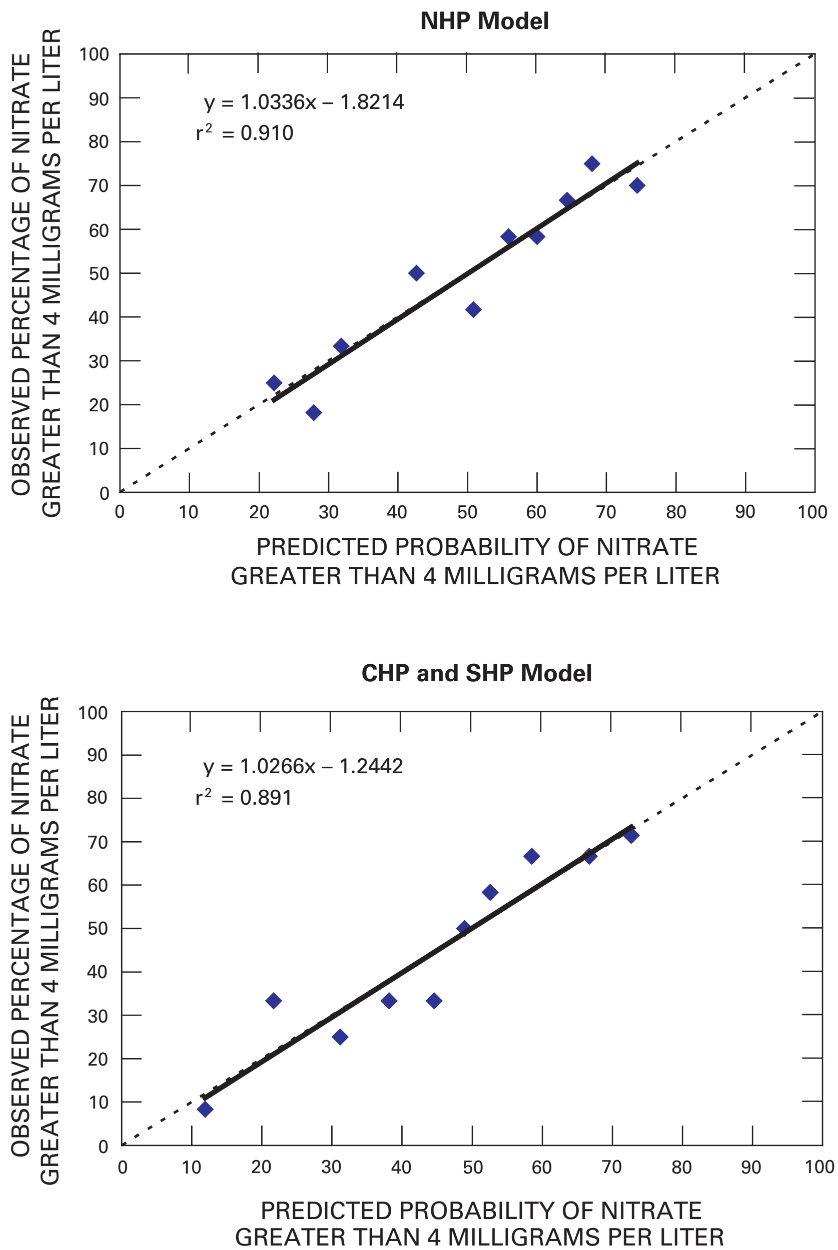


Figure 10. Percentage of actual nitrate detections at or greater than background concentrations and the predicted probability of detecting nitrate at or greater than background (calibration data).

Table 8. Pearson's coefficient matrix and multicollinearity diagnostics for Northern High Plains (NHP) model and Central and Southern High Plains (CHP and SHP) models.

[%, percent; m, meter; Nonirrag, nonirrigated agricultural land in sector (%); Irrag, irrigated agricultural land in sector (%); S_om, soil organic matter content (% by weight); Dtw, depth to regional water table (m); Uzclay; average clay content of unsaturated zone (%)]

NHP model	Pearson's correlation coefficients			Tolerance	Variance inflation factor (VIF)
	Nonirrag	Irrag	S_om		
Nonirrag	1			0.744	1.345
Irrag	–0.374	1		0.822	1.216
S_om	0.316	0.067	1	0.860	1.162

CHP and SHP model	Pearson's correlation coefficients				Tolerance	Variance inflation factor (VIF)
	Dtw	Nonirrag	Irrag	Uzclay		
Dtw	1				0.854	1.171
Nonirrag	0.167	1			0.872	1.146
Irrag	–0.071	–0.347	1		0.858	1.165
Uzclay	–0.314	–0.004	–0.073	1	0.873	1.145

Validation of Nitrate Model

The final multivariate models were validated to evaluate predictive ability. Using the subset of validation wells (number of wells = 104) that sample recently recharged ground water, predicted probabilities were calculated with the final model and compared to observed detections of nitrate exceeding 4 mg/L. Nitrate concentrations from the validation wells have a similar range as the wells used for model development and calibration (number of wells = 232) (table 4) and were converted to binary classification of "zero" for nitrate concentration less than 4 mg/L and "one" for nitrate concentrations equaling or exceeding 4 mg/L. This binary conversion allowed for the percentage of observed detections to be calculated and compared to the average predicted probabilities within each 10 percent decile. The validation shows good predictive ability, with an $r^2 = 0.834$ and negligible systematic bias with respect to the 1:1 ratio line (fig. 11). It is possible the r^2 value from the validation data set is smaller than from the calibration data set because there are less wells in the validation data set than in the calibration data set. Validation of these models may be improved using a greater number of nitrate concentration samples of recently recharged ground water.

Construction of Nitrate Probability and Uncertainty Maps

Maps showing the predicted probability (fig. 12) and uncertainty (fig. 13) of recently recharged ground water of the High Plains aquifer having nitrate concentrations greater than or equal to 4 mg/L were constructed using logistic regression models for NHP (eq. 5) and CHP and SHP (eq. 6). A GIS map-algebra approach was used, rather than interpolation techniques (for example, kriging), because explanatory variables are expressed in GIS; thus probabilities can be calculated at each grid cell in the study area using equations 5 and 6. First, GIS data of explanatory variables in equations 5 and 6 were converted to 80-m grid spacing. Next, contributing area sectors were constructed and properly oriented upgradient for each grid cell across the study area (see Explanatory Variable Extraction). Probabilities were calculated for each grid cell using data extracted from contributing areas and the probability equations (eqs. 5 and 6). The resulting surface was multiplied by 100 to represent the percent probability of nitrate concentrations greater than or equal to 4 mg/L in recently recharged ground water (fig. 12). The percentages of total area of the aquifer represented by each group of percent probability were calculated in GIS as follows: 0 to 20 percent = 12.0 percent; 21 to 40 percent = 41.3 percent; 41 to 60 percent = 25.6 percent; 61 to 80 percent = 19.7 percent; and 81 to 100 percent = 1.4 percent of the aquifer.

The maps showing uncertainty of the probability map (fig. 13) as 90-percent prediction intervals were constructed similar to the probability map (fig. 12). Uncertainty was estimated at each grid cell during calculations of probability, using the risk analysis program @RISK (Palisade Corporation, 2002). Because of the excessively large number of grid cells (80 m) within the study area (80-m grid spacing in 450,658 km^2), additional steps were taken to reduce PC computational run times; GIS data resolution was decreased from 80-m to 500-m grid spacing, and Latin hypercube sampling technique was used rather than traditional Monte Carlo sampling. Results of a Wilcoxon-rank sum test indicated that the increase in resolution from 80-m to 500-m grid spacing did not significantly (p-values greater than 0.050: Nonirrag p-value = 0.84; Irrag p-value = 0.39; S_om p-value = 0.78; Dtw p-value = 0.35; and Uzclay p-value = 0.99) change the

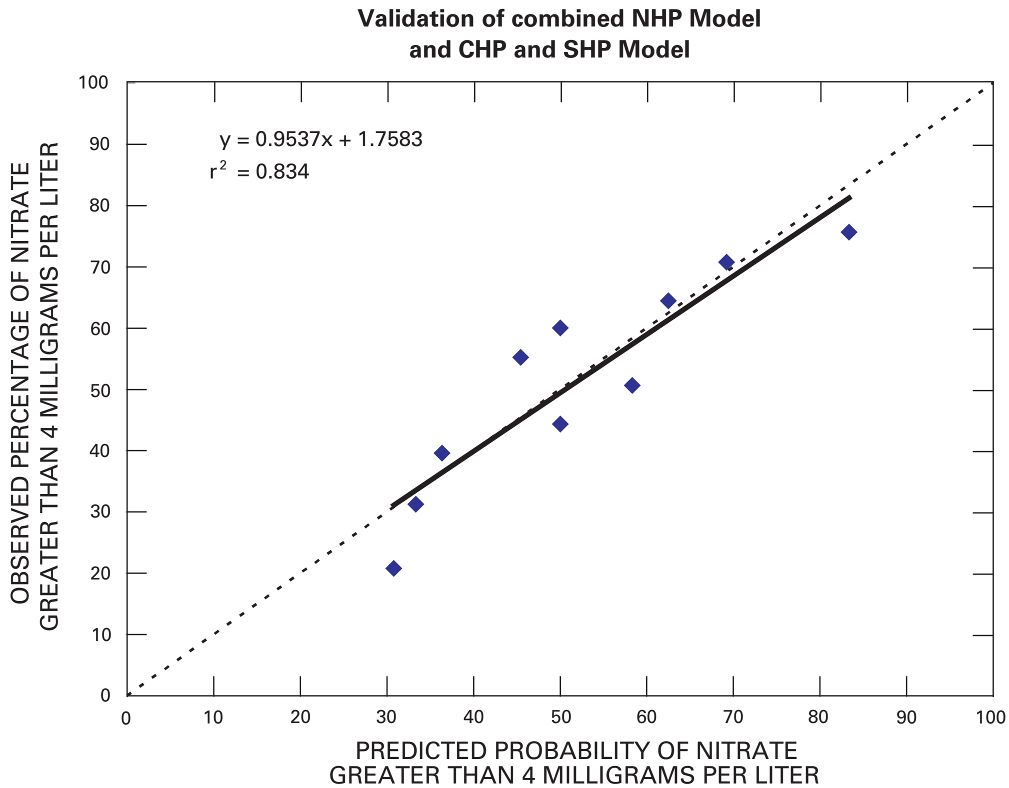


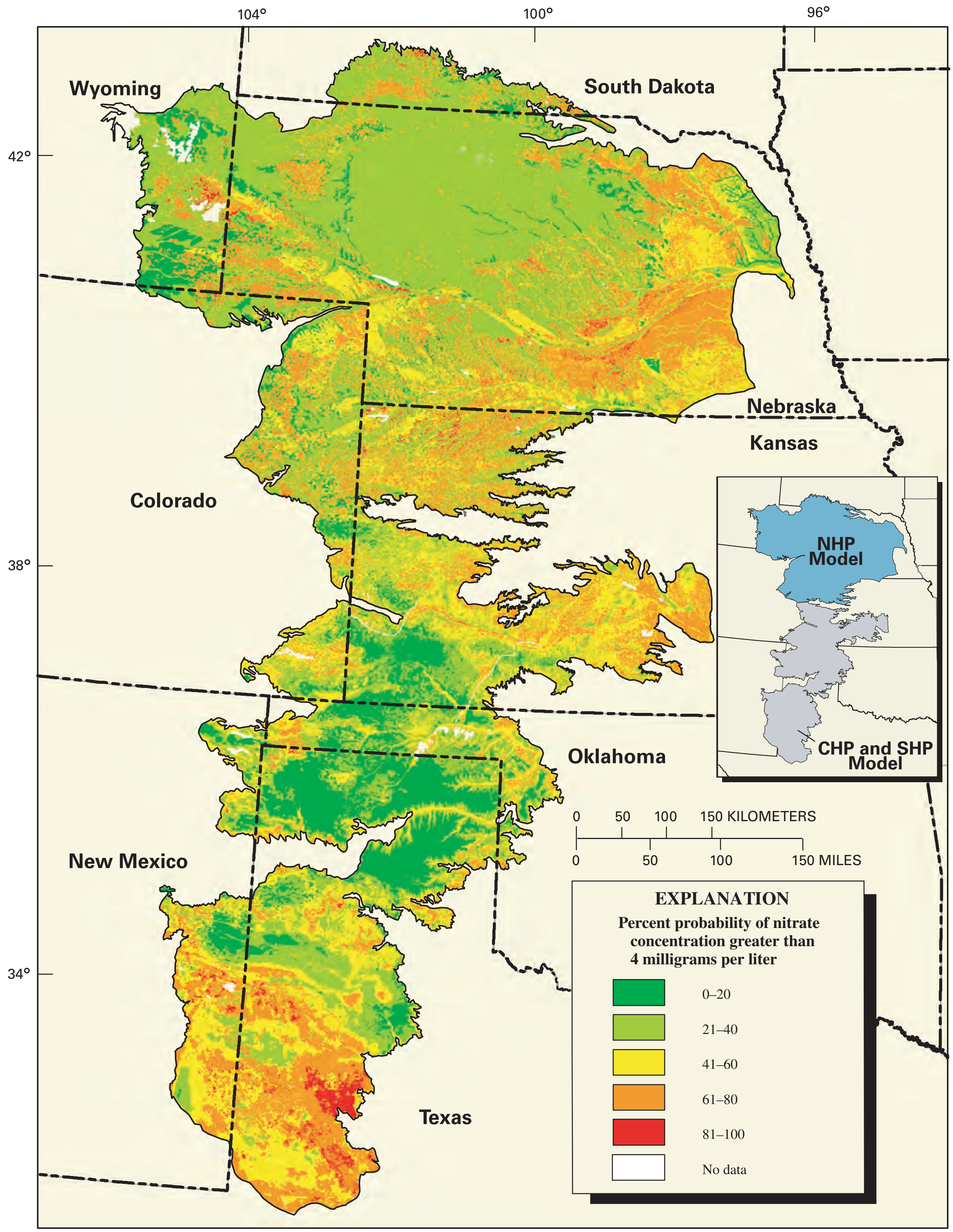
Figure 11. Percentage of actual nitrate detections at or greater than background concentrations and the predicted probability of detecting nitrate at or greater than background (validation data).

data representing each explanatory variable. The stratified sampling approach of Latin hypercube sampling allowed for a faster convergence on the input probability distribution with less number of iterations, as compared to Monte Carlo sampling. As suggested by Phillips and Marks (1996), the input probability distributions were assumed normal; each distribution mean was equal to the estimated value for that grid cell, and the standard deviation was equal to the square root of the estimation error. The estimation errors were defined using a conservative range of errors for each explanatory variable, and Wald 95-percent confidence intervals defined errors for logistic regression model coefficients (summarized in table 9). Measures of errors typically are not available for spatial databases. However, estimates of error for explanatory variables were attributed to range from 10 to 28 percent, based on reported percent correct for the irrigated and nonirrigated lands from Qi and others (2002); root-mean-square prediction error from cross-validation, during kriging of the unsaturated zone lithology; a conservative estimate for depth to water (Dtw) (McGuire and others, 2003); and an arbitrary error estimate for soil organic matter (S_om) because measures of uncertainty are not available for STATSGO data (U.S. Department of Agriculture, 1991). Using the Latin hypercube sampling procedure, 1,000 random realizations were simultaneously drawn from the input probability distributions for each logistic regression coefficients and explanatory variable as the logistic regression probabilities were calculated using equations 5 and 6 for each grid cell of the study area. The coupling of Latin hypercube sampling with logistic regression identified the propagation of input errors through the vulnerability model, which are expressed as the lower 5th and upper 95th percentile of the output probability distribution surrounding the vulnerability prediction using equations 5 and 6. This 90-percent prediction interval was represented for each GIS grid cell as the uncertainty maps illustrated in figure 13.

Vulnerability of Recently Recharged Ground Water to Nitrate Contamination

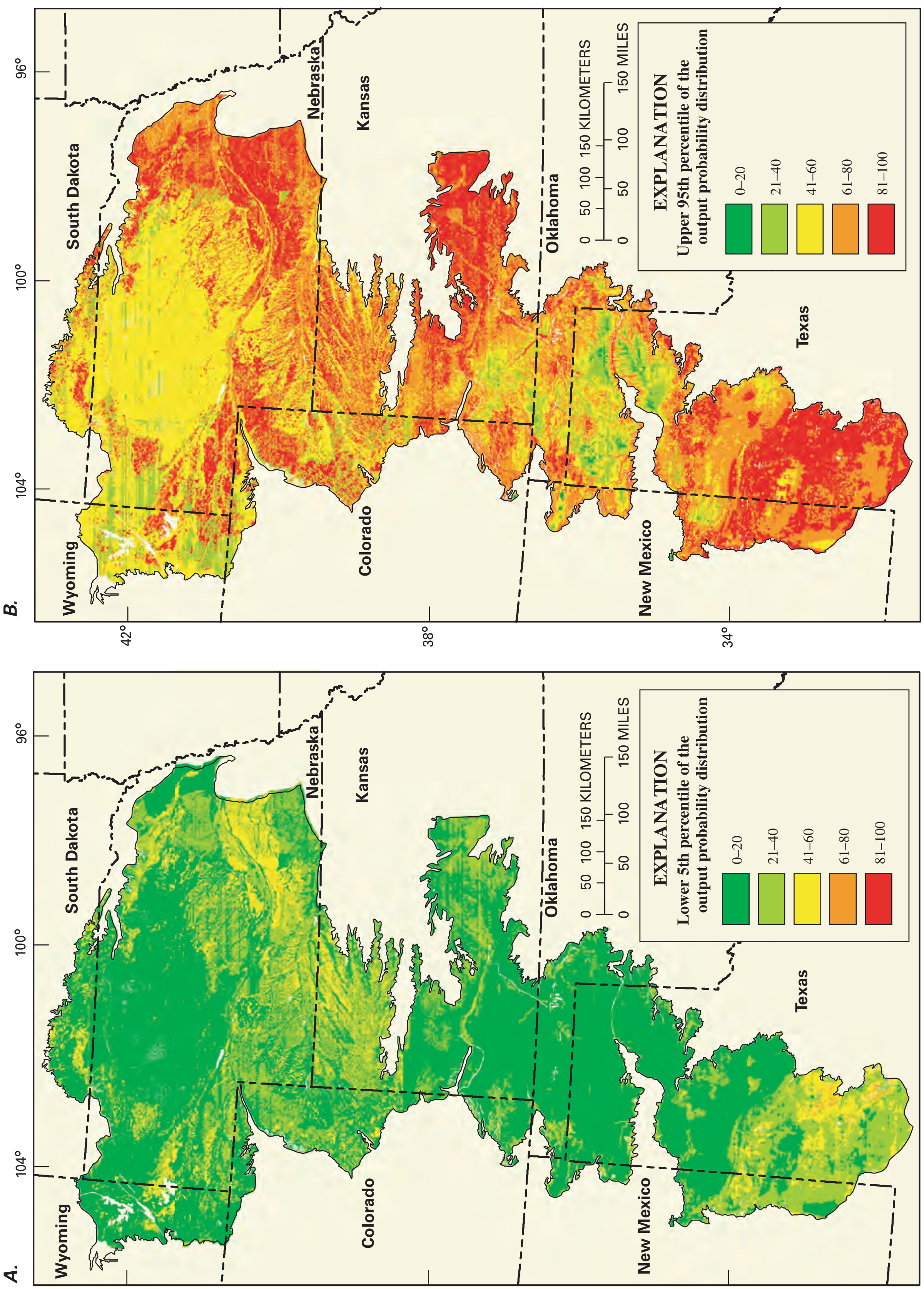
The vulnerability of recently recharged ground water of the High Plains aquifer to nitrate concentrations greater than or equal to 4 mg/L has been expressed as a percent probability (fig. 12) and defined by NHP model (eq. 5) and CHP and SHP model (eq. 6).

The model of vulnerability for the NHP is defined by the percentage of nonirrigated and irrigated agricultural land and organic matter that is present in the soil (eq. 5). NHP vulnerability is directly related to the percentage of nonirrigated and irrigated agricultural land, and inversely related to organic matter in the soil. Therefore, vulnerability of the NHP may



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 12. Probability of detecting nitrate concentrations greater than or equal to background (4 milligrams per liter as nitrogen) concentration in ground water of the High Plains aquifer.



Base information from U.S. Geological Survey digital data, 1:100,000
Albers Equal-Area projection
Standard Parallels 29°30' and 45°30', central meridian –96°

Figure 13. The lower 5th percentile (*A*) and upper 95th percentile (*B*) of the output probability distribution, which represents the 90-percent prediction interval (uncertainty maps) for Northern High Plains and Central and Southern High Plains models of the probability of detecting nitrate concentrations greater than or equal to 4 milligrams per liter in recently recharged ground water of the High Plains aquifer.

Table 9. Wald 95-percent confidence intervals and estimated explanatory variable errors used during Uncertainty Estimation.

[NHP, Northern High Plains; CHP, Central High Plains; SHP, Southern High Plains; %, percent; Nonirrag, nonirrigated agricultural land in sector (%); Irrag, irrigated agricultural land in sector (%); S_om, soil organic matter content (% by weight); Dtw, depth to regional water table, in meters; Uzclay; average clay content of unsaturated zone (%)]

Multivariate model	Logistic regression constant	Explanatory variable	Explanatory variable coefficient	Wald 95-percent confidence limit		Explanatory variable error (%)
				Lower	Upper	
NHP						
	–0.374			–1.477	0.729	
		Nonirrag	0.023	0.009	0.036	20
		Irrag	0.017	0.004	0.030	20
		S_om	–1.487	–3.148	0.173	20
CHP and SHP						
	1.158			–0.632	2.948	
		Dtw	–0.01	–0.017	–0.004	10
		Nonirrag	0.013	0.000	0.026	20
		Irrag	0.011	–0.003	0.026	20
		Uzclay	–0.019	–0.043	0.005	28

be characterized as a function of anthropogenic activity, such as field tillage or irrigation practices that mobilize anthropogenic (commercial fertilizer and manure) and natural sources of nitrate, and the inherent ability of soil to remove nitrate, possibly through denitrification. The relative importance of nonirrigated and irrigated agricultural land within this conceptual framework is corroborated by numerous other studies that have established agriculture as the most extensive influence on nitrate delivered to ground water (Hallberg and Keeney, 1993). Previous logistic regression studies also have identified positive relations between the likelihood of detecting nitrate in ground water and the percentage of agricultural land use near wells (Rupert, 2003; Nolan, 2001; Tesoriero and Voss, 1997). However, a comparison of the standardized coefficients (table 7) reveals the percentage of nonirrigated agriculture (standardized coefficient = 0.328) to have a greater influence than the percentage of irrigated agriculture (standardized coefficient = 0.237) on the vulnerability of the NHP. Few other vulnerability assessments have identified a distinction between nonirrigated and irrigated agricultural land influence, or that nonirrigated agricultural land has a greater influence on aquifer vulnerability than irrigated agricultural land. Although further investigations are needed to understand the underlying processes, a few possible explanations are offered. Both nonirrigated and irrigated agricultural land can receive applications of commercial fertilizer and manure, in addition to the fact there is more nonirrigated land than irrigated land across the study area, may indicate that nonirrigated land is simply a larger source of nitrogen. An alternative explanation is that moisture contents of soils in nonirrigated agricultural lands do not reach saturation, as often happens in irrigated agricultural lands soils, thus anaerobic conditions and accompanying denitrification are not as likely to occur in nonirrigated lands. Plant and root mass in soils may be less in nonirrigated compared to irrigated lands because of cropping patterns and irrigation differences. Less plant and root mass in nonirrigated soils may allow leaching of nitrate below the soil to occur more readily than in irrigated soils. Because of the lack of isotopic data, this vulnerability assessment was not able to distinguish between processes or contributions to nitrate concentrations from mobilization of naturally occurring soils and leaching of applied agricultural nitrogen. However, the application of commercial fertilizer and manure and mobilization of naturally occurring soil nitrate might contribute to aquifer vulnerability. The significance of agricultural land (irrigated and nonirrigated) is consistent with findings by Nolan and others (2002), Rupert (2003), and Ceplecha and others (2004) of ground-water vulnerability within parts of the High Plains aquifer. However, the occurrence of agricultural land alone can be a poor predictor of nitrate concentration in recently recharged ground water (Nolan and Stoner, 2000). The only explanatory variable that represents the inherent susceptibility of the NHP is the amount of organic matter in the soil (S_om). As indicated by the negative coefficient in the NHP model (eq. 6), organic matter in the soil has a significant role in reducing vulnerability. It is likely that organic matter in the soil represents an electron-donating substrate used during microbial respiration by denitrifying bacteria. Denitrification occurs under microaerobic and anaerobic conditions. Therefore, organic matter in the soil may have been identified as a significant explanatory variable in the NHP, rather than in the CHP or SHP, because of generally wetter conditions and more irrigated agriculture that may promote more saturated conditions and denitrification in the soil.

The CHP and SHP model incorporates depth to water table, percentage of nonirrigated and percentage of irrigated agricultural land, and percentage of clay in the unsaturated zone (eq. 6). Vulnerability of these two subregions is directly related to the percentage of nonirrigated and irrigated agricultural land and inversely related to depth to the water table and percentage of clay in the unsaturated zone. Similar to the NHP, the vulnerability of CHP and SHP is characterized as a function of anthropogenic activity (nonirrigated and irrigated agriculture), as discussed above. However, the vulnerability of the CHP and SHP is more influenced by aquifer susceptibility features such as depth to water table (Dtw) and the percentage of clay in the unsaturated zone (Uzclay), which may increase traveltimes and impede nitrate transport to the ground water. The depth-to-water table represents the thickness of the unsaturated zones, or the total distance that surface-derived nitrate must travel to reach the ground water, and thus vulnerability to nitrate is less in areas of the CHP and SHP aquifer with greater depths to water table. However, Burkart and others (1999), Nolan (2001), and Nolan and others (2002) reported a positive relation between depth to water and increased aquifer vulnerability in shallow, unconfined aquifers. A positive relation appears counterintuitive because increasing depth to water generally involves a greater travel distance. However, a positive relation between depth to water table and nitrate concentration can be explained by the presence of anaerobic conditions in soils caused by very shallow ground water, which tends to decrease nitrate concentrations due to denitrification (Nolan, 2001). The depth-to-water table in the combined CHP and SHP aquifer has a median value of 40.8 m, which is substantially greater than in studies that identified positive relations between depth to water and nitrate concentration of shallow ground water. Furthermore, the ground water of the High Plains has generally oxidizing conditions (dissolved oxygen greater than 2 mg/L, table 4), which is not conducive to extensive denitrification (McMahon and others, 2000; McMahon, 2001). Therefore, the inverse relation identified between depth to water and nitrate concentration follows the conceptual model of the CHP and SHP aquifer. This inconsistency is attributed to other studies that incorporated data from outside the study area. The percentage of clay in the unsaturated zone (Uzclay) of the CHP and SHP model also represents an inherent feature of the High Plains aquifer that may help reduce vulnerability. Vulnerability to nitrate greater than or equal to 4 mg/L is inversely related to the percentage of clay in the unsaturated zone (eq. 6). The inverse relation may be explained by the large specific sorption capacity of clay (Bekesi and McConchie, 2000) or its ability to impede water flux, both of which would reduce nitrate concentrations or transport of nitrate to the water table.

Overall model fit was good for the NHP and CHP and SHP models. However, the classification sensitivity of the CHP and SHP model was approximately 14 percent lower than the NHP model (table 7). The lower sensitivity of the CHP and SHP model indicates the model tends to underpredict the occurrence of nitrate greater than or equal to 4 mg/L in recently recharged ground water. The lower sensitivity may result from a loss of statistical correlation due to the greater depths to water in the CHP and SHP. However, the underprediction by the CHP and SHP model indicates that possible preferential flow in the unsaturated zone, downward leakage through the annular space of improperly sealed wells (wellbore leakage), or spatially variable focused recharge zones that occur in the CHP and SHP (McMahon and others, 2003; McMahon and others, 2006) may contribute to rapid nitrate transport that is not accounted for by the current vulnerability model. Further investigations of these mechanisms likely will result in better predictions of nonpoint-source contamination of ground water in the aquifer.

Nolan and others (2002) estimated that some of the highest probabilities of detecting large nitrate concentrations in recently recharged ground water of the conterminous United States are in the High Plains aquifer, because of high nitrogen fertilizer loading and well-drained soils overlying unconsolidated, coarse-grained deposits. Findings from the current vulnerability assessment generally agree with many of the larger spatial patterns that Nolan and others (2002) predicted for the High Plains aquifer. However, the vulnerability predictions shown in figure 12 identify aquifer vulnerability at a finer resolution than is available with a national scale study.

Spatial patterns of predicted nitrate vulnerability generated by this model for the High Plains aquifer are illustrated in figure 12. The majority (53.3 percent) of the study area has less than a 40-percent predicted probability of nitrate greater than or equal to 4 mg/L. Areas of the aquifer with relatively low (less than 40 percent) predicted vulnerability to nitrate greater than 4 mg/L are located in northwestern and north-central areas of the NHP, the central and southern areas of the CHP, and a band across the north-central part of the SHP. Less agricultural land is present in those areas of the NHP, and relatively thick unsaturated zones may help protect the CHP and SHP aquifer in those areas of relatively low predicted vulnerability. Approximately 21.1 percent of the study area has relatively high (greater than 60 percent) predicted probabilities of nitrate greater than or equal to 4 mg/L. These areas generally are located in the southwestern, southern, and eastern areas of the NHP, the eastern arm of the CHP, and the central and southern areas of the SHP. The areas tend to have higher density of agricultural land and shallower depths to water table. The most vulnerable areas (greater than 80 percent) are scattered across the eastern and south areas of the NHP and the southern areas of the SHP. These areas generally have the highest percentages of agricultural land, shallowest depths to water table, and little organic matter or clay in the unsaturated zone. Areas of the aquifer with predicted percent probabilities between 41 and 60 percent may represent areas of moderate vulnerability.

Relatively larger uncertainties are associated with vulnerability predictions in areas with relatively higher predicted vulnerability and, generally, are located across much of the CHP and SHP (fig. 13). The 90-percent prediction intervals also indicate that the lowest uncertainty is generally in the NHP and is located in areas with relatively lower predicted probabilities. Prediction uncertainty can be reduced in future vulnerability assessments of the region by expanding the spatial network of wells that discretely intercept recently recharge ground water, especially across the CHP and SHP, and by developing GIS data with finer spatial resolution.

Appropriate Uses of the Probability and Uncertainty Maps

The probability maps developed and presented in this report illustrate the predicted probability of detecting nitrate greater than or equal to the relative background concentration (4 mg/L) in recently recharged (less than 50 years old) ground water of the High Plains aquifer (fig. 12). This map does not show actual contamination of recently recharged ground water, but rather, depicts areas of the aquifer that have the potential or likelihood of recently recharged ground water with nitrate concentrations that exceed the proposed relative background concentrations of 4 mg/L. Generally, recently recharged ground water is present only near the top of the water table. Although the probability maps show predictions of nitrate detection as a percent probability, there is inherent uncertainty within these predictions that is not shown in the probability map. Estimates of model and data error were used to quantify uncertainty associated with the probability map and are represented as 90-percent prediction intervals by two uncertainty maps representing the lower 5th and upper 95th percentiles of the output probability distribution (fig. 13).

The probability and uncertainty maps provide a tool to help resource decisionmakers to prioritize areas for ground-water-quality monitoring or implement alternative management practices. The maps are intended for regional, subregional, or county-scale use and may have several limitations for use at the site- or field-scale. The probability and uncertainty maps are not appropriate at any scale larger than 1:250,000, as determined by the STATSGO soil data, which has the smallest scale (1:250,000) of the explanatory variables used in the final statistical models. The models and maps do not account for local point sources of nitrate or features and processes that may promote focused recharge, preferential flow, or bypass mechanisms. Therefore, models and maps may not appropriately support local-scale decisions.

The probability and uncertainty maps were created using nitrate data and explanatory variables that were collected from 1990 to 2004 to illustrate spatial predictions of nitrate vulnerability. Because agricultural practices that cause nitrogen loading and mobilization have remained relatively constant during this time period, these maps represent the probability and associated uncertainty of detecting nitrate under current conditions. Temporal validation of these maps using data collected from previous time periods has not been evaluated. Because these maps were based on empirical observations at point locations from a discrete time period, the authors acknowledge that forecasting of future aquifer vulnerability conditions using the presented models or maps may not be appropriate and would require additional validation that is beyond the scope of this study.

Summary

The High Plains aquifer is the principal aquifer system for parts of eight States (Colo., Kans., Nebr., N. Mex., Okla., S. Dak., Tex., and Wyo.) in the Western Great Plains of the United States. Ground water from this aquifer is the primary domestic and public supply for nearly 1.9 million people and supports extensive irrigated agricultural, making it one of the most productive agricultural regions in the United States. However, elevated nitrate concentrations above background levels have been detected in ground water throughout this aquifer. Widespread elevated nitrate concentrations in ground water indicate the aquifer is vulnerable to nonpoint-source contamination. Factors affecting the spatial distribution of nonpoint source contaminants in ground water of regional aquifers systems are complex. This report evaluates the potential vulnerability of ground water recharged during the last 50 years in the High Plains aquifer for nonpoint-source nitrate contamination and identifies the major factors that cause nitrate vulnerability of this regional aquifer system.

Empirically based models and corresponding maps were developed to evaluate the vulnerability of the aquifer to nitrate contamination from nonpoint sources. The models were developed using multivariate logistic regression analysis to predict the probability of detecting nitrate concentrations greater than or equal to 4 mg/L in ground water of the High Plains aquifer that was recharged during the last 50 years. Results from ground-water flow and particle-tracking simulations were used to select wells with screen intervals that intercept recently recharged ground water and to define contributing areas for each of those wells. A total of 336 wells were selected that intercept recently recharged ground water, that have short screened intervals (less than 15 m), and that have ground-water samples that were analyzed for nitrate concentrations during 1990 and 2004. Nitrate-concentration data were used as the response variable, and explanatory variables for logistic regression modeling were extracted from a geographic information system using the contributing area. Nitrate concentration of 4 mg/L was selected to represent the background concentration in ground water of the aquifer and used to establish the binary threshold during logistic regression modeling. Results of Latin hypercube simulations were used to evaluate the uncertainty of model predictions due to model and data error and are represented by 90-percent prediction intervals as uncertainty maps.

Two final multivariate models were selected to represent the vulnerability of ground water to nitrate greater than or equal to 4 mg/L: Northern High Plains (NHP model) and the combined Central and Southern High Plains (CHP and SHP model). These models were selected based on statistical significance (p-values: NHP model = 0.003, CHP and SHP model = 0.003) and model fit (percent correct: NHP = 65.8 percent, CHP and SHP = 70.4 percent). The explanatory variables in the NHP model are nonirrigated

agricultural land, irrigated agricultural land, and organic matter of the soil; and the explanatory variables in the CHP and SHP model are depth to regional water table, nonirrigated agricultural land, irrigated agricultural land, and the percentage of clay in the unsaturated zone. Vulnerability of the NHP is greater in areas that have more nonirrigated and irrigated agricultural lands and less organic matter in the soil. The vulnerability of the CHP and SHP also is greater in areas that have more nonirrigated and irrigated agricultural lands and also in areas with shallow depths to water table and less clay in the unsaturated zone. The NHP and CHP and SHP models were validated using an independent set of nitrate-concentration data. A regression between predicted probabilities from the NHP and CHP and SHP models and actual detections of nitrate greater than or equal to 4 mg/L in the validation data set had an r^2 value of 0.834.

The probability map produced by this assessment illustrates predicted vulnerability to nitrate greater than 4 mg/L, rather than actual contamination of recently recharged ground water. The majority (53.3 percent) of the High Plains aquifer has less than a 40-percent predicted probability of detecting nitrate greater than or equal to 4 mg/L in recently recharged ground water. Approximately 21.1 percent of the study area has relatively high (greater than 60 percent) predicted probability of detecting nitrate greater than or equal to 4 mg/L. These areas with high predicted vulnerability are located in the southwestern, southern, and eastern areas of the NHP, the eastern arm of the CHP, and the southern areas of the SHP. Larger uncertainties associated with vulnerability predictions are located in areas with relatively higher predicted vulnerability and across much of the CHP and SHP. The 90-percent prediction intervals represented by uncertainty maps represent inherent uncertainty of the vulnerability prediction due to model and data errors. It is likely that such uncertainty can be reduced in future vulnerability assessments of the High Plains aquifer by expanding the spatial network of wells that discretely intercept recently recharge ground water and improving GIS data that represent explanatory variables.

References Cited

Aller, L., Bennett, T., Lehr, J.H., and Petty, R.J., 1985, DRASTIC—A standardized system for evaluating ground water pollution potential using hydrogeologic settings: U.S. Environmental Protection Agency Report EPA/600/2–85/018, 163 p.

Allison, P.D., 1991, Logistic regression using the SAS system theory and application: Cary, N.C., SAS Institute Inc., 288 p.

Baker, D.B., Wallrabenstein, L.K., Richards, R.P., and Creamer, N.L., 1989, Nitrate and pesticides in private wells of Ohio A state atlas: The Water Quality Laboratory, Heidelberg College, Tiffin, Ohio.

Becker, M.F., Bruce, B.W., Pope, L.M., and Andrews, W.J., 2002, Ground-water quality in the central High Plains aquifer, Colorado, Kansas, New Mexico, Oklahoma, and Texas, 1999: U.S. Geological Survey Water-Resources Investigations Report 02–4112, 63 p.

Becker, M.F., 1998, Digital map of base of aquifer for High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming: U.S. Geological Survey Open-File Report 98–393, accessed June 2004 at *http://water.usgs.gov/GIS/metadata/usgswrd/XML/ofr98-393_aqbase.xml*

Bekesi, G., and McConchie, J., 2000, Empirical assessment of the influence of the unsaturated zone on aquifer vulnerability, Manawatu Region, New Zealand: Ground Water, v. 38, no. 2, p. 193–199.

Bruce, B.W., Becker, M.F., Pope, L.M., and Gurdak, J.J., 2003, Ground-water quality beneath irrigated agriculture in the central High Plains aquifer, 1999–2000: U.S. Geological Survey Water-Resources Investigations Report 03–4219, 39 p.

Burkart, M.R., and Stoner, J.D., 2001, Nitrogen in groundwater associated with agricultural systems, *in* Follett, R.F., and Hatfield, J.L., eds., Nitrogen in the environment—Sources, problems, and management: N.Y., Elsevier, p. 123–145.

Burkart, M.R., Kolpin, D.W., Jaquis, R.J., and Cole, K.J., 1999, Agrichemicals in ground water of the Midwestern USA—Relations to soil characteristics: Journal of Environmental Quality, v. 28, p. 1908–1915.

Cederstrand, J.R., and Becker, M.F., 1998, Digital map of hydraulic conductivity for the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming: U.S. Geological Survey Open-File Report 98–548, accessed June 2004 at *http://water.usgs.gov/GIS/metadata/usgswrd/ofr98-548.html*

Centers for Disease Control and Prevention, 1996, Spontaneous abortions possibly related to ingestion of nitrate-contaminated well water, La Grange County, Indiana, 1991–1994: Morbidity and Mortality Weekly Report, v. 45, p. 569–572.

Ceplecha, Z.L., Waskom, R.M., Bauder, T.A., Sharkoff, J.L., and Khosla, R., 2004, Vulnerability assessments of Colorado ground water to nitrate contamination: Water, Air, and Soil Pollution, v. 59, p. 373–394.

Dennehy, K.F., 2000, High Plains regional ground-water study: U.S. Geological Survey Fact Sheet FS–091–00, 6 p.

Dennehy, K.F., Litke, D.W., and McMahon, P.B., 2002, The High Plains aquifer; USA—Ground-water development and sustainability, *in* Hiscock, K.M., Rivett, M.O., and Davison, R.M., eds., Sustainable ground-water development: London Geological Society, Special Publication 193, p. 99–119.

Edmunds, W.M., and Gaye, C.B., 1997, Naturally high nitrate concentrations in groundwaters from the Sahel: Journal of Environmental Quality, v. 26, no. 5, p. 1231–1239.

Eckhardt, D.A.V., and Stackelberg, P.E., 1995, Relation of ground-water quality to land use on Long Island, New York: Ground Water, v. 33, no. 6, p. 1019–1033.

Fahlquist, L., 2003, Ground-water quality of the southern High Plains aquifer, Texas and New Mexico, 2001: U.S. Geological Survey Open-File Report 03–345, 59 p.

Fewtrell, L., 2004, Drinking-water nitrate, methemoglobinemia, and global burden of disease—A discussion: Environmental Health Perspectives, v. 112, p. 1371–1374.

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

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.

Gurdak, J.J., and McCray, J.E., 2005, Groundwater vulnerability to pesticides—An overview of approaches and methods of evaluation: Encyclopedia of Water, paper #GW–451.

Gutentag, E.D., Heimes, F.J., Krothe, N.C., Luckey, R.L., and Weeks, J.B., 1984, Geohydrology of the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming: U.S. Geological Survey Professional Paper 1400–B, 63 p.

Hallberg, G.R., and Keeney, D.R., 1993, Nitrate, *in* Alley, W.M., ed., Regional Ground-Water Quality: N.Y., Van Nostrand Reinhold, p. 297–322.

Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: N.Y., Elsevier Science Publishing Company, Inc., 522 p.

Hosmer, D.W., and Lemeshow, S., 1989, Applied logistic regression: N.Y., John Wiley & Sons, Inc., 307 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.

Litke, D.W., 2001, Historical water-quality data for the High Plains regional ground-water study area in Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming, 1930–98: U.S. Geological Survey Water-Resources Investigations Report 00–4254, 65 p.

Loague, K., Bernknopf, R.L., Green, R.E., and Giambelluca, T.W., 1996, Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii—Review: Journal of Environmental Quality, v. 25, p. 475–490.

Longley, P.A., Goodchild, M.F., Maguire, D.J., and Rhind, D.W., 2001, Geographic information systems and science: West Sussex, England, John Wiley & Sons, 454 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.

Luckey, R.L., Gutentag, E.D., Heimes, F.J., and Weeks, J.B., 1986, Digital simulation of ground-water flow in the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming: U.S. Geological Survey Professional Paper 1400–D, 57 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.

McDonald, M.G., and Harbaugh, A.W., 1988, A modular three-dimensional finite-difference ground-water flow model: U.S. Geological Survey Techniques of Water-Resources Investigations, book 6, chap. A1, 586 p.

McGuire, V.L., Johnson, M.R., Schieffer, R.L., Stanton, J.S., Sebree, S.K., and Verstraeten, I.M., 2003, Water in storages and approaches to ground-water management, High Plains aquifer, 2000: U.S. Geological Survey Circular 1243, 51 p.

McKay, M.D., Conover, W.J., and Beckman, R.J., 1979, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code: Technometrics, v. 221, p. 239–245.

McMahon, P.B., 2001, Vertical gradients in water chemistry in the Central High Plains aquifer, southwestern Kansas and Oklahoma panhandle, 1999: U.S. Geological Survey Water-Resources Investigations Report 01–4028, 47 p.

McMahon, P.B., Dennehy, K.F., Bruce, B.W., Böhlke, J.K., Michel, R.L., Gurdak, J.J., and Hurlbut, D.B., 2006, Storage and transit time of chemicals in thick unsaturated zones under rangeland and irrigated cropland, High Plains, United States: Water Resources Research, v. 42, doi:10.1029/2005WR004417.

McMahon, P.B., Böhlke, J.K., and Christenson, S.C., 2004, Geochemistry, radiocarbon ages, and paleorecharge conditions along a transect in the central High Plains aquifer, southwestern Kansas, USA: Applied Geochemistry, v. 19, p. 1655–1686.

McMahon, P.B., Böhlke, J.K., and Lehman, T.M., 2004, Vertical gradients in water chemistry and age in the Southern High Plains aquifer, Texas, 2002: U.S. Geological Survey Scientific Investigations Report 2004–5053, 53 p.

McMahon, P.B., Dennehy, K.F., Michel, R.L., Sophocleous, M.A., Ellett, K.M., and Hurlbut, D.B., 2003, Water movement through thick unsaturated zones overlying the Central High Plains aquifer, southwestern Kansas, 2000–2001: U.S. Geological Survey Water-Resources Investigations Report 03–4171, 32 p.

McMahon, P.B., Bruce, B.W., Becker, M.F., Pope, L.M., and Dennehy, K.F., 2000, Occurrence of nitrous oxide in the central High Plains aquifer, 1999: Environmental Science & Technology, v. 34, p. 4873–4877.

Menard, S., 2002, Applied logistic regression analysis: Thousand Oaks, Calif., Sage Publications, Inc., 111 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., Hamilton, P.A., Helsel, D.R., Hitt, K.J., and Ruddy, B.C., 1995, Nutrients in ground water and surface water of the United States—An analysis of data through 1992: U.S. Geological Survey Water-Resources Investigations Report 95–4031, 74 p.

National Atmospheric Deposition Program, 2000, Nitrogen in the Nation's rain: National Atmospheric Deposition Program Brochure 2000–01a, 13 p.

National Research Council, 1993, Ground water vulnerability assessment, contamination potential under conditions of uncertainty: Washington, D.C., National Academy Press, 210 p.

Nolan, B.T., and Hitt, K.J., 2003, Nutrients in shallow ground waters beneath relatively undeveloped areas of the conterminous United States: U.S. Geological Survey Water-Resources Investigations Report 02–4289, 17 p.

Nolan, B.T., Hitt, K.J., Ruddy, B.C., 2002, Probability of nitrate contamination of recently recharged groundwaters in the conterminous United States: Environmental Science & Technology, v. 36, no. 10, p. 2138–2145.

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 Stoner, J.D., 2000, Nutrients in groundwater of the conterminous United States, 1992–1995: Environmental Science & Technology, v. 34, no. 7, p. 1156–1165.

Palisade Corporation, 2002, Guide to using @RISK, risk analysis and simulation add-in for Microsoft Excel: Newfield, N.Y., Palisade Corporation, 499 p.

Phillips, D.L., and Marks, D.G., 1996, Spatial uncertainty analysis—Propagation of interpolation errors in spatially distributed models: Ecological Modeling, v. 91, p. 213–229.

Pollock, D.W., 1994, User's guide to MODPATH/MODPATH_PLOT, version 3—A particle tracking post-processing package for MODFLOW, the U.S. Geological Survey modular three-dimensional finite-difference ground-water-flow model: U.S. Geological Survey Open-File Report 94–464 (variously paged).

Pope, L.M., Bruce, B.W., Rasmussen, P.P., and Milligan, C.R., 2002, Quality of shallow ground water in areas of recent residential and commercial development, Wichita, Kansas, 2000: U.S. Geological Survey Water-Resources Investigations Report 02–4228, 67 p.

Pope, L.M., Bruce, B.W., Hansen, C.V., 2000, Ground-water quality in Quaternary deposits of the central High Plains aquifer, south-central Kansas, 1999: U.S. Geological Survey Water-Resources Investigations Report 00–4259, 44 p.

Qi, S.L., and Gurdak, J.J., 2006, Percentage of probability of nonpoint-source nitrate contamination of recently recharged ground water in the High Plains aquifer: U.S. Geological Survey Data Series 192, on the World Wide Web at *http://water.usgs.gov/lookup/getspatial?ds192_hp_npctprob* (accessed May 15, 2006).

Qi, S.L., Konduris, A., Litke, D.W., and Dupree, J., 2002, Classification of irrigated land using satellite imagery, the High Plains aquifer, nominal data 1992: U.S. Geological Survey Water-Resources Investigations Report 02–4236, 31 p.

Rao, P.S.C., and Alley, W.M., 1993, Pesticides, *in* Regional ground-water quality, Alley, W.M., ed.: New York, Van Nostrand Reinhold, p. 345–377.

Reilly, T.E., and Harbaugh, A.W., 2004, Guidelines for evaluating ground-water flow models: U.S. Geological Survey Scientific Investigations Report 2004–5038, 30 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.

Rupert, M.G., 2003, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate in ground water in Colorado: U.S. Geological Survey Water-Resources Investigations Report 02–4269, 35 p.

Rupert, M.G., 1998, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate (NO_2+NO_3–N) in ground water of the Idaho part of the upper Snake River Basin: U.S. Geological Survey Water-Resources Investigations Report 98–4203, 32 p.

SAS Institute, 1999, SAS User's Manual–The Logistic Procedure: Cary, N.C., p. 1901–2043.

Schwarz, G.E., and Alexander, R.B., 1995, State Soil Geographic (STATSGO) data base for the conterminous United States: U.S. Geological Survey Open-File Report 95–449, on the World Wide Web at *http://water.usgs.gov/GIS/metadata/usgswrd/XML/ussoils.xml* (accessed 02/21/06).

Solley, W.B., Piercem, R.R., and Perlman, H.A., 1998, Estimated use of water in the United States in 1995: U.S. Geological Survey Circular 1200, 71 p.

Spalding, R.F., and Exner, M.E., 1993, Occurrence of nitrate in groundwater—A Review: Journal of Environmental Quality, v. 22, p. 392–402.

Taylor, G., Daly, C., Gibson, W., Sibul-Weisburg, J., 1997, Digital and map products produced using PRISM: 10th Conference on Applied Climatology, Reno, Nev., American Meteorological Society, p. 217–218.

Tesoriero, A.J., and Voss, F.D., 1997, Predicting the probability of elevated nitrate concentrations in the Puget Sound Basin—Implications for aquifer susceptibility and vulnerability: Ground Water, v. 35, no. 6, p. 1029–1039.

Thatcher, L.L., 1962, The distribution of tritium fallout in precipitation over North America: Bulletin of the International Association of Scientific Hydrology, v. 7, p. 48–58.

U.S. Department of Agriculture, 1991, State Soil Geographic (STATSGO) database: U.S. Department of Agriculture, Soil Conservation Service, Miscellaneous Publication 1492, 88 p.

U.S. Environmental Protection Agency, 2004, Current drinking water standards: Office of Water, on the World Wide Web at *http://www.epa.gov/safewater/mcl.html* (accessed 06/01/04)

U.S. Environmental Protection Agency, 1993, A review of methods for assessing aquifer sensitivity and ground water vulnerability to pesticide contamination: U.S. Environmental Protection Agency, EPA/813/R–93/002, 147 p.

van Horssen, P.W., Pebesma, E.J., and Schot, P.P., 2002, Uncertainties in spatially aggregated predictions from a logistic regression model: Ecological Modelling. v. 154, p. 93–101.

Vowinkel, E.F., Clawges, R.M., Buxton, D.E., Stedfast, D.A., and Louis, J.B., 1996, Vulnerability of public drinking water supplies in New Jersey to pesticides: U.S. Geological Survey Fact Sheet FS–165–96, 3 p.

Walvoord, M.A., Phillips, F.M., Stonestrom, D.A., Evans, D.R., Hartsough, P.C., Newman, B.D., and Striegl, R.G., 2003, A reservoir of nitrate beneath desert soils: Science, v. 302, p. 1021–1024.

Ward, M.H., Mark, S.D., Cantor, K.P., Weisenburger, D.D., Correa-Villasenor, A., and Zahm, S.H., 1996, Drinking water nitrate and the risk of non-Hodgkins lymphoma: Epidemiology, v. 7, p. 465–471.

Waterloo Hydrogeologic, 1999, Visual MODFLOW User's Manual: Waterloo, Ontario, Canada, 280 p.

Weeks, J.B., Gutentag, E.D., Heimes, F.J., and Luckey, R.B., 1988, Summary of the High Plains Regional Aquifer-System Analysis in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming: U.S. Geological Survey Professional Paper 1400–A, 30 p.