Probability of Detecting Atrazine/ Desethyl-atrazine and Elevated Concentrations of Nitrate in Ground Water in Colorado

By Michael G. Rupert

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 02–4269

Prepared in cooperation with the
COLORADO DEPARTMENT OF AGRICULTURE,
COLORADO DEPARTMENT OF PUBLIC HEALTH AND ENVIRONMENT, and
COLORADO STATE UNIVERSITY COOPERATIVE EXTENSION

Denver, Colorado
2003

U.S. DEPARTMENT OF THE INTERIOR
GALE A. NORTON, Secretary

U.S. GEOLOGICAL SURVEY
Charles G. Groat, Director

For additional information write to:

District Chief
U.S. Geological Survey
Box 25046, Mail Stop 415
Denver Federal Center
Denver, CO 80225–0046

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Box 25286
Denver Federal Center
Denver, CO 80225

CONTENTS

FIGURES

TABLES

CONVERSION FACTORS, VERTICAL DATUM, AND ABBREVIATED WATER-QUALITY UNITS

Multiply	By	To obtain
acre	4,047	square meter
acre-foot (acre-ft)	1,233	cubic meter
cubic foot per second (ft^3/s)	0.02832	cubic meter per second
foot (ft)	0.3048	meter
foot per mile (ft/mi)	0.1894	meter per kilometer
[1]foot squared per day (ft^2/d)	0.0929	meter squared per day
gallon per minute (gal/min)	0.06309	liter per second
mile (mi)	1.609	kilometer
meter	3.281	feet
square mile (mi^2)	2.590	square kilometer

Sea level: In this report, “sea level” refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)—a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

Abbreviated Water-Quality Units

µg/L micrograms per liter
mg/L milligrams per liter

[1]The standard unit for transmissivity is cubic foot per day per square foot [(ft^3/d)/ft^2] times foot of aquifer thickness . In this report, the mathematically reduced form, foot squared per day (ft^2/d), is used for convenience.

Probability of Detecting Atrazine/Desethyl-atrazine and Elevated Concentrations of Nitrate in Ground Water in Colorado

By Michael G. Rupert

Abstract

Draft Federal regulations may require that each State develop a State Pesticide Management Plan for the herbicides atrazine, alachlor, metolachlor, and simazine. Maps were developed that the State of Colorado could use to predict the probability of detecting atrazine and desethyl-atrazine (a breakdown product of atrazine) in ground water in Colorado. These maps can be incorporated into the State Pesticide Management Plan and can help provide a sound hydrogeologic basis for atrazine management in Colorado. Maps showing the probability of detecting elevated nitrite plus nitrate as nitrogen (nitrate) concentrations in ground water in Colorado also were developed because nitrate is a contaminant of concern in many areas of Colorado.

Maps showing the probability of detecting atrazine and(or) desethyl-atrazine (atrazine/DEA) at or greater than concentrations of 0.1 microgram per liter and nitrate concentrations in ground water greater than 5 milligrams per liter were developed as follows: (1) Ground-water quality data were overlaid with anthropogenic and hydrogeologic data using a geographic information system to produce a data set in which each well had corresponding data on atrazine use, fertilizer use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well construction. These data then were downloaded to a statistical software package for analysis by logistic regression. (2) Relations were observed between ground-water quality and the percentage of land-cover categories within circular regions (buffers) around wells. Several buffer sizes were evaluated; the buffer size that provided the strongest relation was selected for use in the logistic regression models. (3) Relations between concentrations of atrazine/DEA and nitrate in ground water and atrazine use, fertilizer use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well-construction data were evaluated, and several preliminary multivariate models with various combinations of independent variables were constructed. (4) The multivariate models that best predicted the presence of atrazine/DEA and elevated concentrations of nitrate in ground water were selected. (5) The accuracy of the multivariate models was confirmed by validating the models with an independent set of ground-water quality data. (6) The multivariate models were entered into a geographic information system and the probability maps were constructed.

INTRODUCTION

Ground-water quality is a water-resource management concern in Colorado. Pesticides (a generic term for herbicides, insecticides, and rodenticides) have been detected in greater than 90 percent of the wells sampled in urban and agricultural areas of the South Platte River Basin of northeast Colorado (Dennehy and others, 1998). Atrazine and desethyl-atrazine (DEA), a breakdown product of atrazine, were detected in water from 61 and 54 percent, respectively, of the wells sampled in the South Platte River Basin; alachlor, metolachlor, and simazine also were detected (Dennehy and others, 1998). Pesticide data from the South Platte River Basin were compared with pesticide data collected in many different basins of the United States. The South Platte River Basin ranked in the highest 25 percent of the basins for frequency of pesticide detections (U.S. Geological Survey, 1999), and types of pesticides detected in the South Platte River Basin were similar to those found across the United States. Frequency of pesticide detections from 1,034 wells sampled in agricultural and urban settings across the United States were atrazine (38.2 percent), DEA (34.2 percent), simazine (18.0 percent), metolachlor (14.6 percent), and prometon (13.9 percent) (Kolpin and others, 1998). Not all areas in Colorado have high frequencies of pesticide detections; data collected from the San Luis Valley in south-central Colorado showed few pesticide detections (Anderholm, 1996) in spite of coarse-textured soils and a water table close to the land surface. Ground-water samples collected in the Upper Colorado River Basin also showed few pesticide detections (Spahr and others, 2000).

Draft Federal regulations (U.S. Environmental Protection Agency, 1996) may require that each State develop a State Pesticide Management Plan (PMP) for the herbicides atrazine, alachlor, metolachlor, and simazine. The Colorado Agricultural Chemicals and Groundwater Protection Program, a cooperative effort of the Colorado Department of Agriculture (CDA), the Colorado Department of Public Health and Environment (CDPHE), and the Colorado State University Cooperative Extension (CSUCE), is developing a PMP for each of the herbicides and would benefit from a map that could be used to predict the probability of detecting atrazine and DEA in ground water. The map could be incorporated into the PMP and provide a sound hydrogeologic basis for atrazine management in Colorado. Other organizations and programs that could benefit from maps that predict the probability of detecting atrazine, DEA, and elevated concentrations of nitrate in ground water include the agri-chemical industry, county and city governments, farmers, planning and zoning commissions, education programs for applicators, and State programs related to Wellhead Protection, Drinking Water, Home-A-Syst, and Best Management Plans (BMP's). To address these needs, the U.S. Geological Survey (USGS), in cooperation with the CDA, CDPHE, and CSUCE, conducted a study to develop maps to predict the probability of detecting atrazine and(or) DEA and elevated concentrations of nitrate in ground water in Colorado.

Background

Ground-water vulnerability maps are designed to portray the potential for contamination of ground water in an area on the basis of anthropogenic (related to human activities) and hydrogeologic factors. Several definitions have been used for the term "ground-water vulnerability." The National Research Council (1993) defines ground-water vulnerability to contamination as "the tendency or likelihood for contaminants to reach a specified position in the ground-water system after introduction at some location above the uppermost aquifer." The National Research Council (1993) refined the definition on the basis of whether the assessment was contaminant specific, defined as "specific vulnerability," or for any contamination in general, "intrinsic vulnerability." Rao and Alley (1993) defined "intrinsic vulnerability" to be the time of travel of water from the point of contaminant entry to the reference location in the ground-water system. The U.S. Environmental Protection Agency (1993) defined aquifer vulnerability to pesticides as "the relative ease with which a contaminant (such as a pesticide) applied on or near the land surface can migrate to the aquifer of interest under a given set of agronomic management practices, pesticide characteristics, and hydrogeologic sensitivity conditions." The U.S. Environmental Protection Agency (1993) groups factors such as depth to ground water, geology, and soils into hydrogeologic sensitivity; sensitivity plus contaminant input and contaminant characteristics constitutes ground-water vulnerability. Vowinkel and others (1996) defined vulnerability as sensitivity plus intensity, where intensity is a measure of the source of contamination. 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 ground-water contamination?

The most widely known ground-water vulnerability mapping procedure is the DRASTIC model (Aller and others, 1985). DRASTIC was designed to evaluate the potential for ground-water contamination in a given area on the basis of hydrogeologic factors. The DRASTIC acronym refers to the seven factors considered in the model: **d**epth to water, net **r**echarge, **a**quifer media, **s**oil media, **t**opography, **i**mpact of vadose zone media, and hydraulic **c**onductivity of the aquifer (Aller and others, 1985, p. iv). The DRASTIC model has been used to develop ground-water vulnerability maps in many parts of the Nation, but the validity and accuracy of the model has met with mixed success (Koterba and others, 1993, p. 513; Barbash and Resek, 1996; Rupert, 2001). DRASTIC maps usually use a vulnerability point rating system that is based upon best professional judgment instead of calibration to actual contaminant concentrations.

One of the first published ground-water vulnerability maps in Colorado was developed for the greater Denver area by Hearne and others (1995) using the DRASTIC model. Hall (1998) developed a ground-water sensitivity map for the entire State of Colorado using a modified form of the DRASTIC model. Four factors were incorporated into Hall's map: (1) Location of principal aquifers, (2) depth to water table, (3) soil hydrologic group, and (4) the presence of irrigated agriculture. Hall's map (1998) was constructed at a relatively small scale (1-km grid cells), and the point rating scheme was developed on the basis of best professional judgment.

Nolan and others (1997) developed a ground-water vulnerability map for the contiguous 48 States that shows the risk of nitrate contamination. Nolan and others (1998) used estimates of nitrogen input (such as fertilizers and animal manure), soil drainage, and woodland-to-cropland ratios. Nolan and others (1998) observed a good fit when they compared their maps to actual ground-water quality data.

The validity and accuracy of ground-water vulnerability maps have been improved by calibrating the vulnerability rating system to actual ground-water quality. Rupert (2001) improved the validity and accuracy of a modified DRASTIC vulnerability map by calibrating (adjusting) the vulnerability point ratings to measured nitrite plus nitrate as nitrogen (nitrate) concentrations in ground water using nonparametric statistical tests.

The validity and accuracy of ground-water vulnerability maps also have been improved by using logistic regression to relate water-quality data to anthropogenic and hydrogeologic factors (Koterba and others, 1993; Druliner and others, 1996; Nolan and Clark, 1997; Tesoriero and Voss, 1997; Rupert, 1998). Logistic regression (Hosmer and Lemeshow, 1989; Kleinbaum, 1994) is a statistical method that can be used to predict the probability of occurrence of an event of interest (such as detection of a contaminant in ground water) as a function of a set of independent variables (such as land cover and soils). Logistic regression is an improvement over the nonparametric statistical methods used by Rupert (2001) because the actual probabilities of a detection are quantified and the weighting of the independent variables is quantified.

Purpose and Scope

This report presents maps developed to show the probability of detecting concentrations of atrazine and(or) DEA (atrazine/DEA) at or greater than 0.1 µg/L and elevated concentrations of nitrate greater than 5 mg/L in ground water in Colorado. The maps are based on predictions of the probability of detecting contamination calculated using logistic regression and are termed “probability” maps. The probability maps are analogous to the U.S. Environmental Protection Agency's (1993) vulnerability maps because hydrogeologic, land cover, and chemical-use variables are combined. These probability maps are intended to be used by the Colorado PMP to identify areas in greatest need of ground-water protection. The study used a broad-brush approach, and additional site-specific data would be needed for site-specific evaluations.

The maps were developed from existing Geographic Information System (GIS) data and existing ground-water quality data; resources to develop or collect additional data were not available. Relations between atrazine/DEA and nitrate in ground water and anthropogenic and hydrogeologic factors were examined using a GIS and statistical analysis software. Statistical models that predict the probability of detecting atrazine/DEA and elevated concentrations of nitrate in ground water were developed using logistic regression statistical methods. The only herbicide probability map developed by the study was for atrazine/DEA because only atrazine/DEA had a sufficient number of detections in the available data sets for analysis by logistic regression. Ground-water quality data used in the study were collected from 1992 to 2000 by the State of Colorado and the USGS. In this report, nitrate refers to nitrite plus nitrate measured as nitrogen.

Acknowledgments

The author acknowledges the assistance and information provided by Bradford Austin of the Colorado Department of Public Health and Environment, Troy Bauder and Reagan Waskom of the Colorado State University Cooperative Extension, Robert Wawrzynski of the Colorado Department of Agriculture, and Dana Kolpin and Bernard T. Nolan of the U.S. Geological Survey.

CLIMATE, GEOLOGY, AND WATER USE

Colorado (fig. 1) has been called “The Rooftop of the Nation” because it contains more peaks higher than 14,000 ft than all the other States combined (Litke, 1990). The headwaters of four major rivers—the Arkansas, Colorado, Rio Grande, and South Platte—are in Colorado. The lowest elevation (about 3,300 ft) is in southeastern Colorado, where the Arkansas River flows into Kansas.

Colorado’s midlatitude interior-continental location and high-altitude, mountainous terrain combine to produce a complex and diverse climate. Annual precipitation ranges from about 7 inches in the San Luis Valley of south-central Colorado to about 60 inches in the mountains east of Steamboat Springs (fig. 1) (Collins, 1991). Seasonal large-scale atmospheric circulation interacts with the mountainous topography to produce three major precipitation patterns in the State (Collins, 1991). Throughout the winter, frontal systems from the Pacific Ocean travel generally eastward over the Colorado Rockies, causing relatively large amounts of precipitation west of the Continental Divide and small amounts of precipitation east of the divide. This important precipitation pattern provides the snowpack for snowmelt runoff that produces much of Colorado’s water supply. The second precipitation pattern affects eastern Colorado during the spring and summer, where moisture from the Gulf of Mexico produces periodic, widespread rainfall and occasionally severe thunderstorms. The third precipitation pattern affects the southern Rocky

Figure 1. Location of counties, towns and cities, major rivers, and the San Luis Valley in Colorado.

Mountains during the summer, where subtropical moisture from the Pacific and Atlantic oceans drifts northward into the Southwestern United States, resulting in frequent summer thunderstorms (Collins, 1991).

The geology of Colorado can be generalized by subdividing the State into four regions. Eastern Colorado contains the High Plains (fig. 2), a gently rolling upland composed mostly of shales, limestones, and siltstones of Cretaceous age, overlain by gravels of Tertiary age (Chronic, 1980). The central portion of Colorado contains the Rocky Mountains, a complex area of folded and faulted sedimentary, igneous, and metamorphic rocks that have been uplifted to elevations greater than 14,000 ft. West of the Rocky Mountains is a region of sedimentary rocks of Paleozoic, Cretaceous, and Tertiary age that are part of the Colorado Plateaus province, a vast area that extends into Utah, Arizona, and New Mexico. In many areas, these relatively flat-lying sedimentary rocks have been incised with deep canyons. The southwest portion of Colorado contains the San Juan Mountains, which are composed of a complex assortment of volcanic rocks of Tertiary age at elevations up to 14,000 ft.

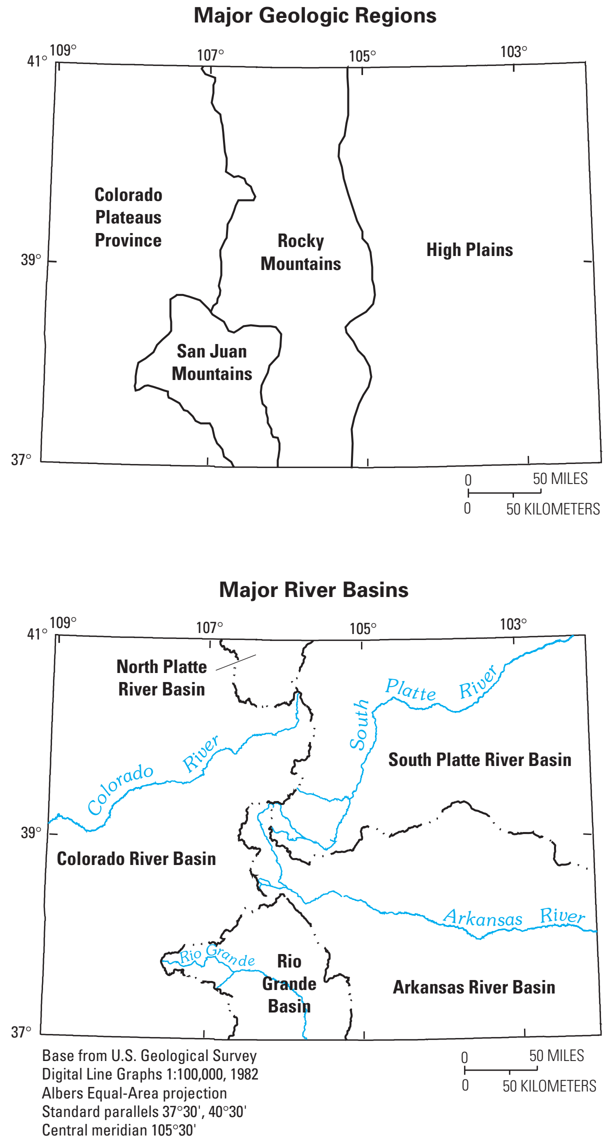
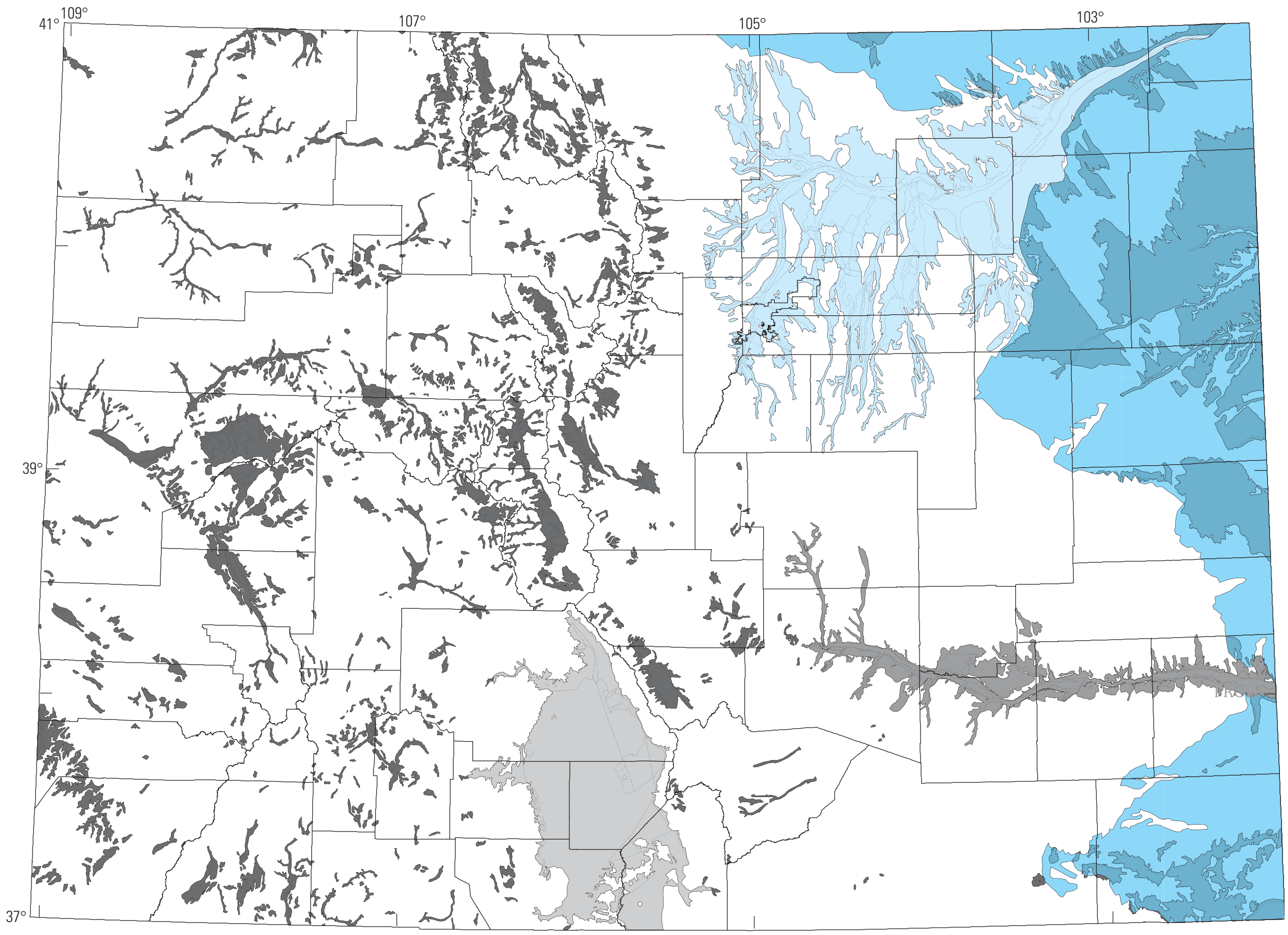


Figure 2. Major geologic regions and river basins in Colorado.

The most productive and easily developed aquifers in Colorado (and commonly some of the most vulnerable to contamination) are composed of unconsolidated deposits of Tertiary or Quaternary age (Hurr and Hearne, 1985). The most significant unconsolidated aquifers are composed of alluvial and eolian deposits near the South Platte River and its tributaries (fig. 3), alluvial and eolian deposits near the Arkansas River and its tributaries, alluvial and eolian deposits in the High Plains aquifer in eastern Colorado, and sand, gravel, and eolian deposits in the San Luis Valley of south-central Colorado (Hurr and Hearne, 1985). An important regional aquifer in eastern Colorado is the High Plains aquifer of Tertiary and Quaternary age (fig. 3), also referred to as the "Ogallala aquifer." The High Plains aquifer extends into Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming. The High Plains aquifer is a water-table aquifer that is composed mainly of sand and gravel deposits with some silt and clay.

The western one-half of Colorado holds three-fourths of the State's surface-water resource but contains only 10 percent of the State's population (Litke, 1990). The eastern one-half of the State, which includes the populous Front-Range urban corridor surrounding Denver, must rely on scarce rainfall (8–16 inches per year), ground-water resources, and water diverted from the western part of the State to fulfill its water demands (Litke, 1990).

Ground water constitutes 18 percent of the total water used in Colorado (Hurr and Hearne, 1985). Of the total population, 15 percent obtains drinking water from ground water. In many rural areas of Colorado, ground water is the main source for domestic and irrigation supply. Of the total 2.7 million acres irrigated in Colorado, 2.1 million acres are irrigated with ground water or a combination of ground water and surface water (Hurr and Hearne, 1985).



Base from U.S. Geological Survey Digital Line Graphs
1:100,000, 1982
Albers Equal-Area projection
Standard parallels 37°30', 40°30'
Central meridian 105°30'

0 10 20 30 40 50 MILES
0 10 20 30 40 50 KILOMETERS

EXPLANATION

Hydrogeomorphic region

- Alluvial and eolian deposits overlying the High Plains aquifer
- Sand and gravel deposits of the High Plains aquifer
- Interbasin sand, gravel, eolian, and glacial deposits
- Alluvial and eolian deposits near the Arkansas River and tributaries
- Sand, gravel, and eolian deposits in the San Luis Valley
- Alluvial and eolian deposits near the South Platte River and tributaries
- Not mapped

Figure 3. Hydrogeomorphic regions in Colorado.

METHODS OF INVESTIGATION

Maps showing the probability of atrazine/DEA detections and elevated concentrations of nitrate in ground water in Colorado were developed in several steps.

1. All suitable anthropogenic, hydrogeologic, and ground-water quality data were compiled.
2. Ground-water quality data were overlaid with anthropogenic and hydrogeologic data using a GIS to produce a data set in which each well had corresponding data on atrazine use, fertilizer use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well construction. These data then were downloaded to a statistical software package (SYSTAT) for analysis (SPSS, Inc., 2000).
3. Relations were observed between ground-water quality and the percentage of land-cover categories within circular regions (buffers) around wells. Several buffer sizes were evaluated; the buffer size that provided the strongest relation was selected for use in the logistic regression models.
4. Several preliminary multivariate models with various combinations of independent variables were constructed.
5. The multivariate models that best predict the presence of atrazine/DEA in ground water were selected.
6. The models were validated with a second, independent set of ground-water quality data to verify their accuracy.
7. The multivariate models were entered into the GIS, and the probability maps were constructed.

The specific details of data compilation, statistical methods, model development, model validation, and construction of the probability maps are discussed in the following sections.

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

Anthropogenic, hydrogeologic, and ground-water quality data were compiled for use by this study. Anthropogenic and hydrogeologic data include herbicide use, fertilizer use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well construction. These data were available in GIS format from a variety of sources.

Two independent sets of ground-water quality data were used in this study. One set of data was collected by the State of Colorado; the other was collected by the USGS. The data collected by the State of Colorado were used to calibrate the models because the data were collected at the greatest number of wells and the wells had a wider distribution across the State. The data collected by the USGS were used to validate the models.

Anthropogenic and Hydrogeologic Data

Herbicide-use data were obtained from two sources. The first source was the CDA, which surveyed all commercial pesticide applicators during 1997 (Robert Wawrzynski, Colorado Department of Agriculture, written commun., 2001). Information on compounds applied, the amounts of active ingredients, and the number of acres being treated in each county was compiled. This survey included commercial chemical applicators, residential landscaping applicators, and structural/termiticide applicators, but retail sales by sources such as hardware stores and home gardening stores were not included. Farmers were included in the survey, but those data could not be used in this study because the survey did not include the county where the farmer applied the pesticide.

The second source of herbicide-use data was Battaglin and Goolsby (1994), who prepared GIS coverages of herbicide-use estimates for the 20 most-used herbicides in the conterminous United States based on data compiled by Gianessi and Puffer (1991). Gianessi and Puffer's (1991) estimates for each county included the number of acres treated, pounds of active ingredient used, and pounds used per square mile. Herbicide-use estimates by county were generated by Gianessi and Puffer (1991) for 1987 by the following procedure: (1) Collect statistics by State, by crop, on percentage of acres treated with a given herbicide, and average annual application rates of the herbicide from surveys sent to Extension Service weed scientists in 1987 and 1989; (2) augment survey data with published information available from some States; (3) establish herbicide-use profiles, by State and by crop, containing the percentage of acres treated and average annual application rates; (4) apply herbicide-use profiles to county-level crop-acreage estimates from the 1987 Census of Agriculture (U.S. Department of Commerce, 1989); and (5) tabulate pounds of active ingredient of herbicides used by crop and by county. Although crop-acreage data represent the 1987 growing year, the herbicide-use estimates generally reflect 1989 usage amounts (Gianessi and Puffer, 1991). In this study, usage amounts in Battaglin and Goolsby (1994) were converted to ounces to allow comparison with the usage data compiled by the CDA. In order to produce the most accurate results with logistic regression, usage amounts also were divided by 100,000 to transform the data to the same relative magnitude as all other independent variables. When the final probability maps were created for this study, atrazine use was assigned only to areas in GIS coverages that delineate irrigated agricultural land cover because these are the areas where most, if not all, agricultural atrazine use occurs.

Estimates of nitrogen fertilizer use for 1997 were developed by David Lorenz (U.S. Geological Survey, written commun., 2001). First, estimates of the total amount of nitrogen fertilizer product that was sold in Colorado during 1997 were obtained from the Association of American Plant Food Control Officials (AAPFCO) at the University of Kentucky. The statewide total was prorated to each county based upon amounts of fertilizer expenditures by farmers that were reported in the 1997 Census of Agriculture. In order to produce the most accurate results with logistic regression, nitrogen fertilizer estimates were divided by 1,000,000 to transform the data to the same relative magnitude as all other independent variables prior to analysis by logistic regression. When the final probability maps were created for this study, nitrogen fertilizer use was assigned only to areas in GIS coverages that delineate agricultural land cover.

Estimates of average annual precipitation for 1961–90 were developed by Daly and others (1994), with the **P**arameter-elevation **R**egressions on **I**ndependent **S**lopes **M**odel (PRISM). PRISM uses climatic point data and a digital elevation model (DEM) to generate gridded estimates of climatic parameters. PRISM has been used extensively to map precipitation and minimum and maximum temperature over the United States, Canada, and other countries.

Green (1992) digitized geology from the original scribe sheets used to prepare the published Geologic Map of Colorado (Tweto, 1979). For this study, geologic units were grouped to allow comparison of ground-water quality among similar geologic materials. For instance, sedimentary deposits of Quaternary age were grouped into one category and metamorphic and intrusive rocks of Precambrian age were grouped into a different category.

A GIS data set delineating hydrogeomorphic regions in Colorado was developed for this study (fig. 3). Hydrogeomorphic regions are similar in concept to regional aquifers, but are distinguished from regional aquifers in that hydrogeomorphic regions are delineated on the basis of general geographic locations of geologic materials and not on actual aquifer locations. A comprehensive coverage of regional aquifers for all of Colorado was not available, but hydrogeomorphic regions satisfied the needs of this study. Hydrogeomorphic regions were delineated by combining information from three sources. The first source was the geologic map of Colorado (Tweto, 1979), which was digitized by Green (1992). Unconsolidated alluvial regions were delineated by using the alluvial, gravel, eolian, and glacial deposits of Quaternary age that are delineated on the geologic map of Colorado. The second source was a GIS coverage delineating the boundary of the High Plains aquifer (Cederstrand and Becker, 1999) that was compiled from a digital coverage created for publication of paper maps in McGrath and Dugan (1993). The third source was a GIS coverage digitized from reports that delineate the boundary of the valley-fill aquifer in the lower Arkansas River Basin (Hurr and Moore, 1972; Nelson and others, 1989a, 1989b, 1989c).

Land-cover data for Colorado were obtained from National Land Cover Data (NLCD) developed by the U.S. Geological Survey (2000). NLCD were produced as part of a cooperative project between the U.S. Environmental Protection Agency and the USGS to produce a consistent land-cover GIS data layer for the conterminous United States based on 30-meter-resolution Landsat thematic mapper (TM) data. The NLCD contains 21 categories of land-cover information—open water, perennial ice and snow, low-intensity residential, high-intensity residential, commercial/industrial/transportation, bare rock/sand/clay, quarries/strip mines/gravel pits, transitional, deciduous forest, evergreen forest, mixed forest, shrubland, orchards/vineyards, grasslands, pasture/hay, row crops, small grains, fallow, urban/recreational grasses, woody wetlands, and emergent herbaceous wetlands. The NLCD files were too large for the computers available to this study to manipulate in raw form, so the files were generalized to 60-meter resolution. This generalization may have improved the consistency of the data by reducing the number of isolated single-cell occurrences of a particular land-cover classification, which were probably an artifact of spectral processing and not true differences in land cover.

Soils data were obtained from the State Soil Geographic (STATSGO) database (U.S. Department of Agriculture, 1991). The finer scale Soil Survey Geographic (SSURGO) database (U.S. Department of Agriculture, 1995) was not available for all regions in Colorado. The STATSGO data were not suitable for use by this study in raw form, so STATSGO data compiled by Schwarz and Alexander (1995) were used. These later data included weighted averaging of many of the soil characteristics contained in the database (table 1). Available water-capacity data were multiplied by 100 to transform the data to the same relative magnitude as all other independent variables prior to analysis by logistic regression. The U.S. Department of Agriculture (1993) provides more information on these soil characteristics.

Ground-Water Quality Data for Model Calibration

Herbicide and nitrate ground-water quality data collected by the State of Colorado for the years 1992 through 1999 were used as the model calibration data set for this study. The data were collected through a cooperative program with the CDPHE, CDA, and CSUCE to meet the requirements of the Colorado Agriculture Chemicals and Groundwater Protection Act. This monitoring program collects samples from different regions of the state each year, and collects samples from selected wells in Weld County annually. The water samples were analyzed by the CDA laboratory in Denver. The minimum laboratory reporting limits for alachlor, atrazine, DEA, and metolachlor were lowered during the study period as the laboratory improved its analytical capabilities. All ground-water quality monitoring data were censored to the highest minimum laboratory reporting limit (0.1 μg/L for atrazine) for analysis by logistic regression. The ground-water quality data collected by the State of Colorado were collected from shallow wells completed in the uppermost portion of unconfined aquifers. The median depth of the wells is 64 ft, the median water level is 14 ft below land surface, and the median saturated thickness penetrated by the wells is 40 ft.

Table 1. Summary of soil characteristics described in the STATSGO soil database (Schwarz and Alexander, 1995)

[mm, millimeters]

Soil characteristic	Description
Available water capacity	The volume of water that should be available to plants if the soil, exclusive of rock fragments, was at field capacity.
Clay content	Clay content of the soil or horizon, expressed as a percentage of material less than 2 mm in size.
Liquid limit	The water content at the change between the liquid and plastic state of the soil. It is measured on thoroughly puddled soil material that has passed a number 40 sieve (0.43 mm) and is expressed on a dry-weight basis.
Occurrence of hydric soils	Hydric soils are soils that are saturated, flooded, or ponded long enough during the growing season to develop anaerobic conditions in the upper part. They make up part of the criteria for the identification of wetlands.
Organic matter content	The amount of organic material in the soil, in percent by weight.
Soil permeability	The amount of water that will move downward through a unit area of saturated soil in unit time, under unit hydraulic gradient, measured in inches per hour. Based upon laboratory measurements.
Soil drainage	The natural drainage condition of the soil, based on the frequency and duration of periods when the soil is free from saturation. Ratings are composed of seven categories ranging from very poorly drained to excessively drained.
Soil erodibility	A relative index of the susceptibility of bare, cultivated soil to particle detachment and transport by rainfall. Measured by applying simulated rainfall on freshly tilled plots.
Soil hydrologic group	The minimum steady-ponded infiltration rate for a bare ground. Ratings are composed of four categories, A through D, with A having the highest saturated hydraulic conductivity.
Soil surface slope	Slope of soil surface, measured in percent.
Soil thickness	The weighted average thickness of all soil layers, in inches.

Ground-water quality data collected by the State of Colorado were screened to use only one analysis from each well. For wells in which the compound was detected in more than one sample, the largest concentration was used. For wells in which a compound was detected in one sample, but not in another sample, the sample with the detection was used in the logistic regression calculations. The reason for this approach is because the well has a likelihood for future contamination if the compound was detected at any time in water from that well, and the most valid logistic regression model will be developed if all detections are used.

Herbicides and elevated concentrations of nitrate primarily were detected in water from wells in the eastern half of Colorado and in the San Luis Valley. Atrazine/DEA was detected mostly in the Arkansas and the South Platte River Basins (fig. 4). Concentrations of atrazine ranged from below the minimum laboratory reporting level of 0.1 μg/L to a maximum concentration of 4.2 μg/L (table 2), which exceeds the maximum contaminant level of 3 μg/L established for drinking water by the U.S. Environmental Protection Agency (2000). Metolachlor and alachlor were detected in the South Platte River valley, but at a much lower frequency than atrazine (table 2). Nitrate concentrations are greater than the maximum contaminant level of 10 mg/L (U.S. Environmental Protection Agency, 2000) in the Arkansas River Basin, the South Platte River Basin, and the San Luis Valley in the Rio Grande Basin (fig. 5).

Ground-Water Quality Data for Model Validation

Herbicide and nitrate ground-water quality data analyzed by the USGS National Water Quality Laboratory for the years 1993 through 2000 were used as the model validation data set for this study. Most of the data were collected by the USGS National Water-Quality Assessment Program (NAWQA), which is studying water quality in three basins in Colorado in addition to many other basins in the United States. Data collected in northeast Colorado were collected by the South Platte River Basin Study Unit (Bruce and McMahon, 1998). Data collected in central and western Colorado were collected by the Upper Colorado River Basin Study Unit (Apodaca, 1997), and the data collected in the San Luis Valley were collected by the Rio Grande Basin Study Unit (Anderholm, 1996). Atrazine data from the lower Arkansas River valley (southeastern Colorado) were collected by investigators of the State of Colorado as part of the cooperative program with CDA, CDPHE, and CSUCE but were included in the USGS data set because the samples were analyzed by the USGS National Water Quality Laboratory using the same USGS laboratory methods used for the NAWQA data. The reporting level for atrazine in the USGS data set (0.001 μg/L) (Zaugg and others, 1995) is lower than the reporting levels in the Colorado data set (0.1 μg/L). The ground-water quality data in the USGS data set were collected from shallow wells completed in the uppermost portion of unconfined aquifers. The median depth of the wells is 36 ft, the median water level is 13 ft below land surface, and the median saturated thickness penetrated by the wells is 12 ft.

Herbicides and elevated concentrations of nitrate in the USGS data set were detected in areas similar to detections in the State of Colorado data set, but the USGS data set had a larger frequency of herbicide detection because of the lower laboratory reporting limit. Atrazine and DEA were commonly detected in the Arkansas and South Platte River Basins (fig. 6). Atrazine and DEA were not detected in the San Luis Valley, probably because atrazine is not used by local weed control districts (Eric Lane, Colorado Department of Agriculture, oral commun., 2002) or farmers in the area (Sandra McDonald, Colorado State University Cooperative Extension, oral commun., 2002). The USGS data set had a higher frequency of detection of atrazine, metolachlor, and simazine than the State of Colorado data set probably because of the lower laboratory reporting limits of the USGS data (table 2). USGS atrazine concentrations ranged from below the minimum laboratory reporting limit of 0.001 to 1.6 μg/L. For the most part, nitrate concentrations were above the maximum contaminant level of 10 mg/L in the South Platte River Valley and the San Luis Valley (fig. 5).

Statistical Methods and Regression Models

Logistic regression (Hosmer and Lemeshow, 1989; Kleinbaum, 1994) is conceptually similar to multiple linear regression because relations between one dependent variable and several independent variables are evaluated. In logistic regression, the dependent variable (for this study, atrazine/DEA detection or elevated nitrate concentration) was transformed

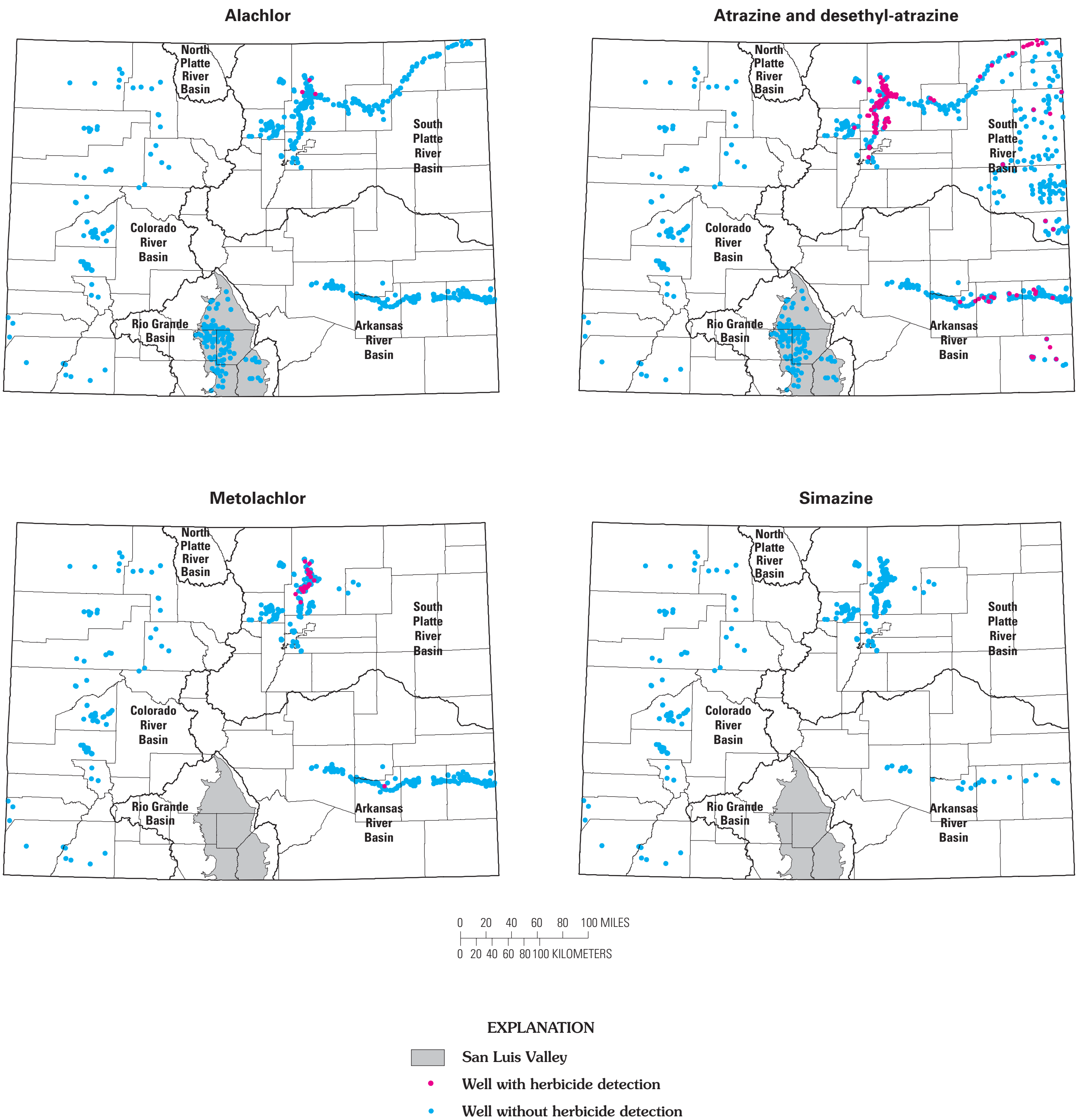


Figure 4. Wells with and without detections of alachlor, atrazine and desethyl-atrazine, metolachlor, and simazine in Colorado sampled by the State of Colorado, 1992–99.

Table 2. Nitrate and selected pesticides in ground water in Colorado sampled by the State of Colorado and the U.S. Geological Survey, 1992–2000

[Concentrations of nitrate are in milligrams per liter; concentrations of alachlor, atrazine, desethyl-atrazine, metolachlor, and simazine are in micrograms per liter; e, estimated concentration; MCL, maximum contaminant level; HA, health advisory level; RSD, risk-specific dose; ne, not established; <, less than]

Constituent	Number of wells sampled	Number of samples with detection	Percentage of detections	Minimum laboratory reporting level	Minimum concentration measured	Median (50th-percentile) concentration	Maximum concentration measured	MCL, HA, or RSD
State of Colorado—calibration data								
Nitrate	655	602	92	<0.5	<0.5	3.8	83.6	10
Alachlor	554	3	<1	<.1	<.1	<.5	3	2
Atrazine	683	110	16	<.1	<.1	<.1	4.2	3
Desethyl-atrazine	305	30	10	<.1	<.08e	<.2	1.3	ne
Metolachlor	383	21	5	<.1	<.1	<.1	11.75	100
Simazine	266	0	0	<.2	<.2	<.2	<.2	4
U.S. Geological Survey—validation data								
Nitrate	330	302	92	<.05	<.05	3.1	61	10
Alachlor	228	1	<1	<.002	<.002	<.002	.120	2
Atrazine	228	83	36	<.001	<.001	<.001	1.6	3
Desethyl-atrazine	228	78	34	<.002	<.002	<.002	.51	ne
Metolachlor	228	36	16	<.002	<.002	<.002	2.7	100
Simazine	228	57	25	<.005	<.005	<.005	.220	4

to a binary variable (detection or nondetection). A major advantage of logistic regression over multiple regression is that the former is well suited for analysis of data sets with a large number of nondetections.

Logistic regression calculates several statistical parameters that determine the predictive success of the model. The log-likelihood ratio measures the success of the model as a whole by comparing observed with predicted values (Hosmer and Lemeshow, 1989, p. 13); specifically, it tests whether model coefficients of the entire model are significantly different from zero. The most significant model is the one with the highest log-likelihood ratio, taking into account the number of independent variables (degrees of freedom) used in the model. The log-likelihood ratio follows a chi-squared distribution, and the computed p-value indicates whether model coefficients are significantly different from zero. In other words, the computed p-value is the significance level attained by the data; the smallest p-value indicates the best model. A p-value of 0.05 indicates a significance level of 95 percent; a p-value of 0.01 indicates a significance level of 99 percent. McFadden's rho-squared (SPSS, Inc., 2000, p. I–571) is a transformation of the log-likelihood statistic and is intended to mimic the r-squared of linear regression. Rho-squared is always between zero and one; a rho-squared approaching 1 corresponds to more significant results. Rho-squared tends to be smaller than r-squared, so a small number does not necessarily imply a poor fit. Values between 0.20 and 0.40 indicate good results (SPSS, Inc., 2000, p. I–571). The percentage of correct predictions is a measure of how many actual atrazine/DEA detections and nondetections are present, compared with what was predicted by the model; the largest number denotes the best model. The percentage of correct responses is calculated as the number of observed detections predicted by the model as detections, plus the number of observed nondetections predicted as nondetections, divided by the combined number of observed detections and nondetections (Nolan and Clark, 1997, p. 855). The partial-likelihood ratio was used to compare nested models to determine the significance of adding one or more new variables (Helsel and Hirsch, 1992; Nolan and Clark, 1997). A nested model contains all of the independent variables in the original model, plus one or more additional independent variables. To determine whether the model is improved by adding the independent variable, the logistic regression model is calculated without that new variable. Logistic regression calculates a partial-likelihood ratio. The logistic regression model then is rerun, this time with the additional new independent variable; the second model also calculates a partial-likelihood ratio. The difference in partial-likelihood ratios between the two models is calculated, and a chi-squared approximation is calculated with degrees of freedom equal to the number of additional variables in the new model. If the p-value from the chi-squared distribution is less than 0.10, the model has been significantly improved at the 90-percent significance level.

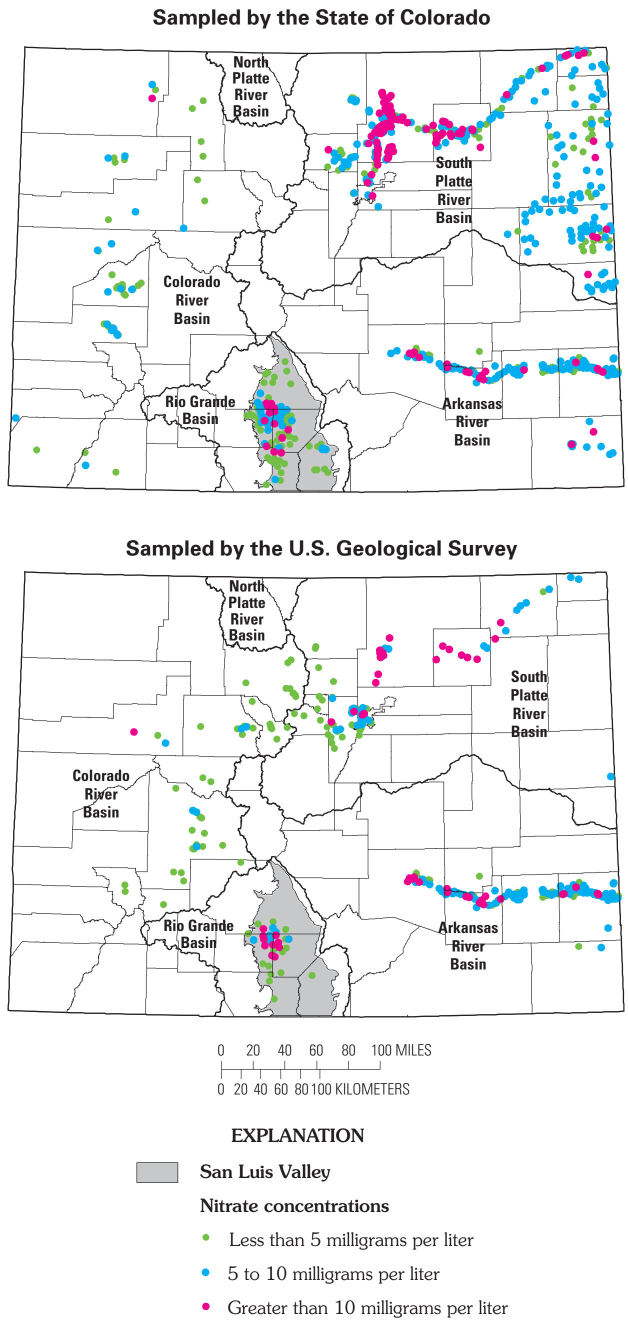


Figure 5. Nitrate concentrations in water from wells sampled by the State of Colorado (1992–99) and the U.S. Geological Survey (1993–2000), Colorado.

In logistic regression, a model is generated that predicts the probability (P) of detecting atrazine/DEA or elevated concentrations of nitrate in ground water similar to equation 1:

$$P = \frac{e^{a + b_1(H) + b_2(S) + b_3(L) + b_4(CU)}}{1 + e^{a + b_1(H) + b_2(S) + b_3(L) + b_4(CU)}} \qquad (1)$$

where

- P is the probability of detecting atrazine/DEA or elevated nitrate in ground water;
- a is logistic regression constant;
- b_1 is coefficient for hydrogeomorphic region;
- H is hydrogeomorphic region (table 3; fig. 3);
- b_2 is coefficient for soil characteristic;
- S is soil characteristic (available water capacity, clay content, organic matter content, occurrence of hydric soils) (tables 2 and 3);
- b_3 is coefficient for land cover;
- L is land cover (percent row crops, percent small grains, percent low-intensity residential, and(or) percent pasture/hay) (table 3);
- b_4 is coefficient for chemical use; and
- CU is chemical use (atrazine or nitrogen fertilizer) (table 3).

Ground-water quality data collected by the State of Colorado were used to calibrate the logistic regression models. Ground-water quality data collected by the USGS were used to validate the performance of the models. All data on atrazine/DEA and nitrate concentrations in ground water were converted to binary coding of “zero” for wells with no detection and “one” for wells with detections to satisfy the input data requirements of logistic regression. To produce the largest possible sample set, and because DEA is a breakdown product of atrazine, atrazine and DEA data were combined into one dependent variable.

Atrazine use, fertilizer use, percent land cover, precipitation, soils, and well construction were modeled as continuous variables. Land cover was modeled as percentage of a certain land-cover classification within

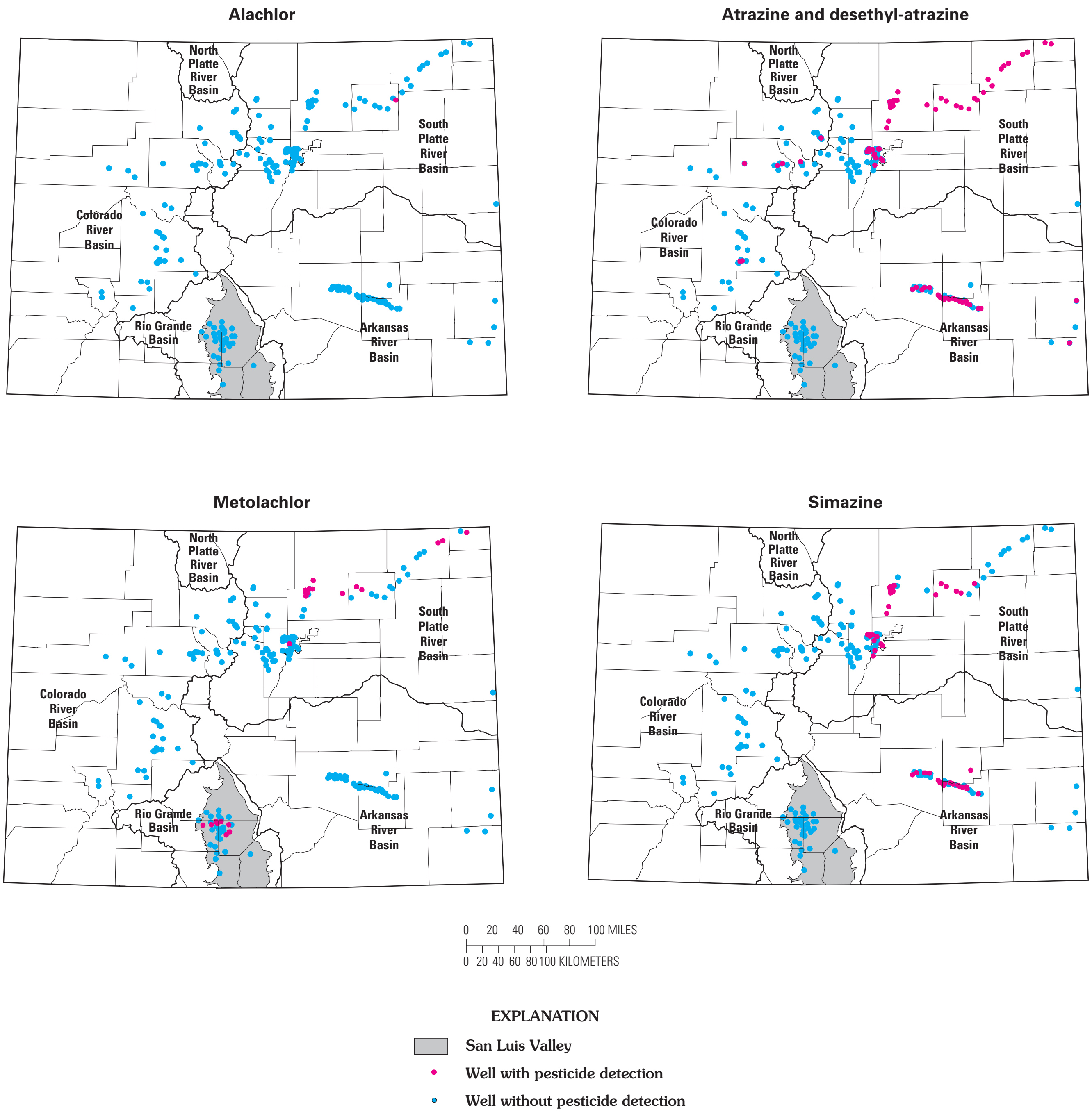


Figure 6. Wells with and without detections of alachlor, atrazine and desethyl-atrazine, metolachlor, and simazine in Colorado sampled by the U.S. Geological Survey, 1993–2000.

Table 3. Logistic regression coefficients and individual p-values of independent variables significantly related with the detection of atrazine/desethyl-atrazine concentrations at or greater than 0.1 microgram per liter and nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado

[--, no relation observed; values not enclosed in parentheses are logistic regression coefficients; values enclosed in parentheses are individual p-values; independent variables in bold are used in equation 1 for the final models; <, less than]

Independent variable	Atrazine/ desethyl-atrazine model, with atrazine use	Atrazine/ desethyl-atrazine model, without atrazine use	Nitrate model, with fertilizer use	Nitrate model, without fertilizer use
Logistic regression constant	**–5.098 (<0.001)**	**–5.250 (<0.001)**	**–1.989 (<0.001)**	**–1.425 (0.007)**
Depth to water in well	--	--	--	--
Total well depth	--	--	--	--
Surface elevation	--	--	--	--
Hydrogeomorphic region—alluvial and eolian deposits overlying the High Plains aquifer	**.723 (.316)**	**1.240 (.106)**	**–.933 (.066)**	**–.535 (.267)**
Hydrogeomorphic region—sand and gravel deposits of the High Plains aquifer	**.950 (.230)**	**1.568 (.062)**	**–.854 (.105)**	**–0.486 (.339)**
Hydrogeomorphic region—interbasin sand, gravel, eolian, and glacial deposits	**–13.264 (.964)**	**–13.658 (.962)**	**–2.508 (.022)**	**–2.544 (.020)**
Hydrogeomorphic region—alluvial and eolian deposits near the Arkansas River and tributaries	**.261 (.678)**	**.239 (.704)**	**.842 (.030)**	**.712 (.061)**
Hydrogeomorphic region—sand, gravel, and eolian deposits in the San Luis Valley	**–14.187 (.945)**	**–14.990 (.940)**	**–.416 (.405)**	**–0.677 (.165)**
Hydrogeomorphic region—alluvial and eolian deposits near the South Platte River and tributaries	**1.329 (.027)**	**1.549 (.010)**	**.893 (.028)**	**1.231 (.001)**
Available water capacity of soil	**.402 (< .001)**	**.394 (<.001)**	**.129 (.010)**	**.112 (.023)**
Clay content of soil	**–.126 (<.001)**	**–.143 (<.001)**	**–.054 (.013)**	**–.067 (.002)**
Soil liquid limit	--	--	--	--
Occurrence of hydric soils	**1.912 (.025)**	**1.745 (.040)**	**–1.171 (.103)**	**–1.184 (.084)**
Organic matter content of soil	**–1.619 (.028)**	**–1.522 (.038)**	**–1.235 (.030)**	**–.934 (.085)**
Soil permeability	--	--	--	--
Soil drainage	--	--	--	--
Soil erodibility	--	--	--	--
Soil hydrologic group	--	--	--	--
Soil surface slope	--	--	--	--
Soil thickness	--	--	--	--
Precipitation	--	--	--	--
Percentage of low-intensity residential land cover within a 500-meter buffer	--	**.015 (.081)**	--	--
Percentage of high-intensity residential land cover within a 500-meter buffer	--	--	--	--
Percentage of commercial/industrial/transportation land cover within a 500-meter buffer	--	--	--	--
Percentage of urban/recreational grasses land cover within a 500-meter buffer	--	--	--	--
Percentage of shrubland within a 2,000-meter buffer	--	--	--	--
Percentage of pasture/hay land cover within a 2,000-meter buffer	--	**.020 (.092)**	--	--
Percentage of row crops land cover within a 2,000-meter buffer	**.027 (<.001)**	**.039 (<.001)**	**.036 (<.001)**	**.042 (<.001)**
Percentage of small grains land cover within a 2,000-meter buffer	**–.078 (<.001)**	**–.076 (.001)**	--	--
Percentage of fallow land cover within a 2,000-meter buffer	--	--	--	--
Atrazine use	**0.430 (.059)**	--	--	--
Fertilizer use	--	--	**.063 (.001)**	--

a specific buffer area. Precipitation values were determined from contour maps; values were set at one-half of the contour interval. For example, wells lying between precipitation contour intervals of 10 and 20 inches were assigned precipitation values of 15 inches.

Because of their categorical nature, geology and hydrogeomorphic regions were modeled as discrete (design) variables. Discrete variables were coded as “one” if a well was located in a particular geologic or hydrogeologic unit and coded “zero” if the well was not located in a unit. An example is the hypothetical case where two geologic units exist, Unit A and Unit B. If a well is located in Unit A, then the database would be coded “Unit A = 1,” “Unit B = 0.” Hosmer and Lemeshow (1989) provide more information on the use of continuous and discrete variables in logistic regression.

Selecting the Most Significant Land-Cover Buffer Size

Land cover in many areas in Colorado consists of a patchwork of many different types. For instance, fields of irrigated row crops can be commingled with irrigated pasture and nonirrigated rangeland (fig. 7). Ground-water quality in areas with multiple land-cover classifications probably reflects the combined effects of all the land-cover classifications in those areas. To account for mixtures of land-cover classifications in the vicinity of wells, this study incorporated the percentage of land-cover classifications within circular buffers around wells into the logistic regression models.

The significance of relations between the percentage of land-cover classifications and ground-water quality is affected by the size of the buffers around wells. To determine which buffer size had the most significant relation with ground-water quality, univariate logistic regression relations between ground-water quality data collected by the State of Colorado and percentage of land-cover classifications were evaluated. The following buffer radii were evaluated: zero, 100, 250, 500, 1,000, 2,000, 4,000, 8,000, 16,000, 20,000, 30,000, 40,000, and 50,000 meters. A wide range of buffers was used to bracket the optimum buffer size. McFadden’s rho-squared (SPSS, Inc., 2000, p. I–571) was calculated for each buffer size, and the results were plotted to highlight the most significant relation (figs. 8 and 9).

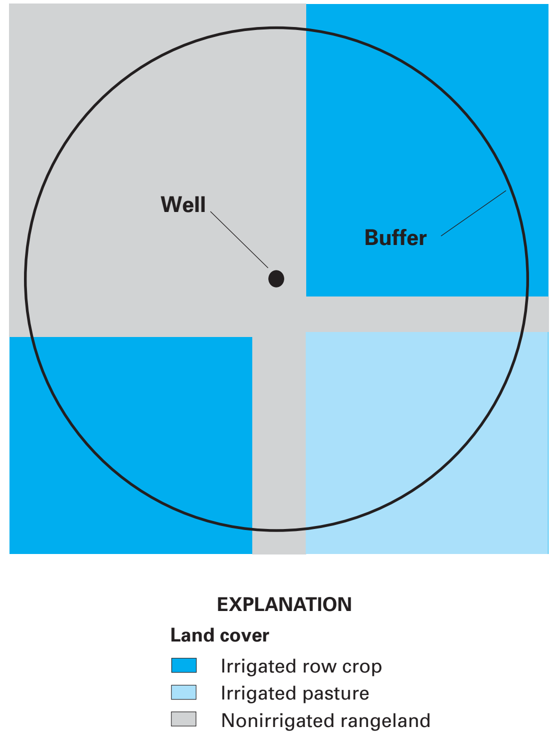


Figure 7. Example land-cover classifications within a circular buffer around a well.

For agricultural lands, an optimum buffer size of 2,000 meters was determined for atrazine/DEA detections and elevated concentrations of nitrate (fig. 8). Pasture/hay, row crops, small grains, and fallow land-cover classifications were regressed with atrazine/DEA detections greater than 0.1 µg/L and nitrate detections greater than 5 mg/L; row crops had the most significant relation, followed by small grains (fig. 8). Pasture/hay and fallow did not have a significant relation. Logistic regression calculated p-values for each independent variable in the univariate relations; p-values for row crops and small grains were less than 0.004, denoting these relations are statistically significant.

For urban lands, an optimum buffer size of 500 meters was determined for atrazine/DEA detections and elevated concentrations of nitrate (fig. 9). Low-intensity residential, high-intensity residential, commercial/industrial/transportation, and urban/recreational grasses land-cover classifications were regressed with atrazine/DEA detections greater than 0.1 µg/L and nitrate detections greater than 5 mg/L; commercial/industrial/transportation had the most significant relations, followed by urban/recreational grasses and high-intensity residential (fig. 9). Low-intensity residential did not have a significant relation. Individual p-values for commercial/industrial/transportation were less than 0.005, which indicates these relations are statistically significant. Individual p-values for urban/recreational grasses and high-intensity residential ranged from 0.100 to 0.254, which indicates much weaker relations.

Agricultural land cover had a larger significant buffer size than urban land cover, possibly because wells in agricultural lands are deeper than wells in urban land cover. Out of 439 wells sampled by the State of Colorado with well depth information, 418 wells have row crop land cover within a 2,000-meter buffer. The median depth of these wells is 65 ft, and the mean well depth is 117 ft. From this same data set, 95 wells have commercial/industrial/transportation land cover within a 500-meter buffer; the median well depth is 50 ft and the mean well depth is 86 ft. Agricultural land cover also may have a larger buffer size than urban land cover because ground-water quality in agricultural land is typically influenced by nonpoint pollution such as wide-spread fertilization of fields, whereas urban land may be more influenced by point-source pollution from local sources such as spills and leaks.

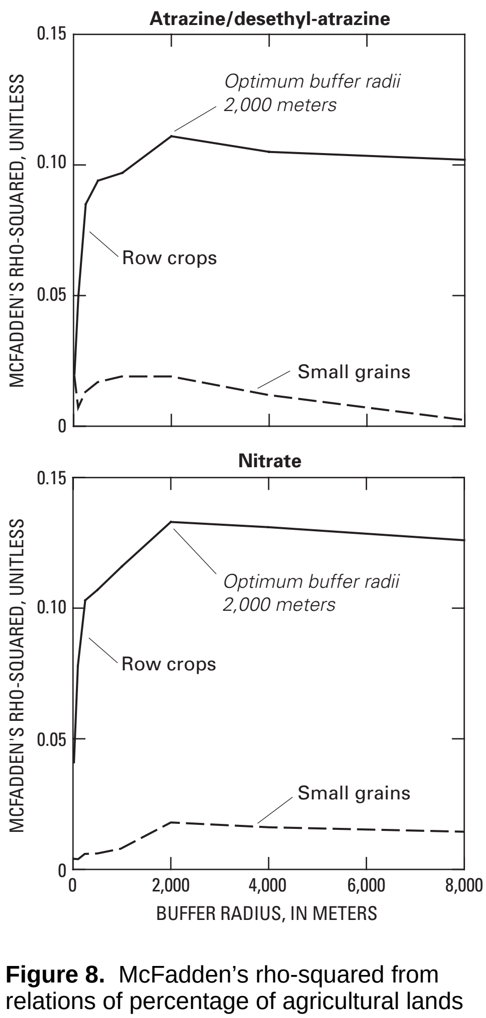


Figure 8. McFadden's rho-squared from relations of percentage of agricultural lands within circular buffers around wells with atrazine/desethyl-atrazine detections at or greater than 0.1 microgram per liter and nitrate detections greater than 5 milligrams per liter in ground water in Colorado.

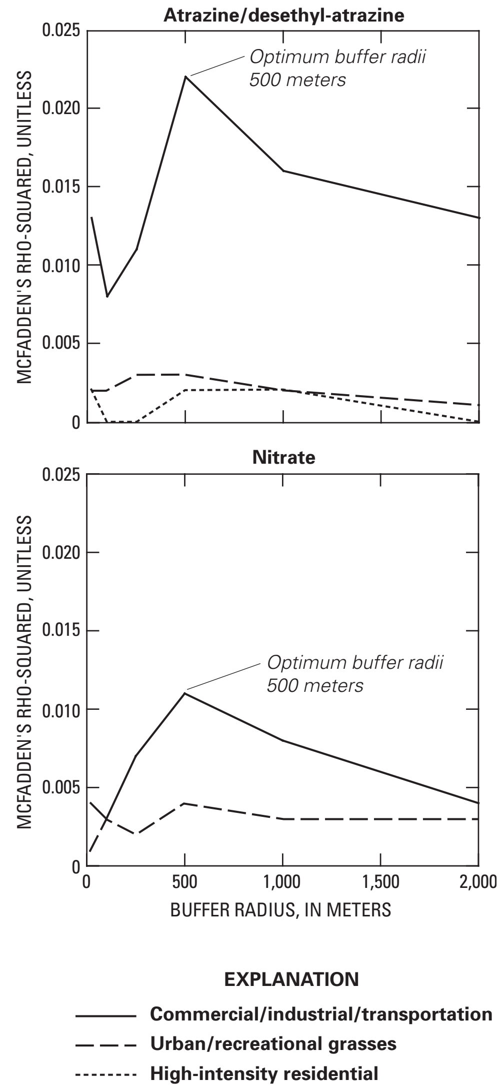


Figure 9. McFadden's rho-squared from relations of percentage of urban lands within circular buffers around wells with atrazine/desethyl-atrazine detections at or greater than 0.1 microgram per liter and nitrate detections greater than 5 milligrams per liter in ground water in Colorado.

ESTIMATING THE PROBABILITY OF DETECTING ATRAZINE/ DESETHYL-ATRAZINE AND ELEVATED CONCENTRATIONS OF NITRATE IN GROUND WATER

Statistical models predicting the probability of detecting atrazine/DEA at or greater than concentrations of 0.1 µg/L in ground water were developed using logistic regression techniques. All possible combinations of independent variables were evaluated to develop the most accurate models. Overall model validity and accuracy were determined by evaluating the log-likelihood ratio, McFadden's rho-squared, p-values calculated for each independent variable, and the percentage of correct responses. The models were built by including each individual variable in the model and evaluating the resulting test statistics. One of the most useful methods to determine whether addition of a particular independent variable made a significant improvement to the model was to compare the partial-likelihood ratios calculated before and after addition of that variable. As described in the "Statistical Methods and Regression Models" section, the independent variable was determined to significantly improve the model if the chi-squared p-value calculated from the difference of the partial-likelihood ratios was less than 0.1.

Development of Atrazine/ Desethyl-atrazine Model

Two models were developed that predict the probability of detecting atrazine/DEA in ground water: one with and one without estimates of atrazine use. Both models were developed because atrazine use can improve the statistical strength of the model, but the use estimates are inexact estimates and are available only on a county-by-county basis. As a result, in some cases, probability ratings change substantially at county boundaries.

Hydrogeomorphic regions, available water capacity of soils, clay content of soils, occurrence of hydric soils, organic matter content of soils, percentage of land in row crops within a 2,000-meter buffer, and percentage of land in small grains within a 2,000-meter buffer were significant variables in both atrazine/DEA models (table 3). Percentage of low-intensity residential land within a 500-meter buffer and percentage of pasture/hay land within a 2,000-meter buffer were significant variables in the model without atrazine use but were not significant in the model with atrazine use, probably because atrazine use contains similar information but has higher statistical significance than low-intensity residential and pasture/hay land cover. This higher statistical significance is confirmed by a smaller p-value for atrazine use (0.059) than low-intensity residential (0.081) and pasture/hay (0.092) land covers without atrazine use (table 3). Overall performance of both models was good, with the chi-squared p-value calculated from the log-likelihood ratio of the entire model less than 0.001, McFadden's rho-squared values greater than 0.338, and the percentage of correct predictions greater than 0.8 (table 4).

The p-values corresponding with each significant land cover and soil variable incorporated in the models were all less than 0.10, and several were less than 0.001 (table 3), confirming that the land cover and soil variables used in the models were significantly related to atrazine/DEA detections in ground water. The significance of the hydrogeomorphic regions is more difficult to interpret than the percent land cover and soils variables because hydrogeomorphic regions were modeled as discrete variables. Because they were discrete variables, the

Table 4. Statistical results from models that predict the probability of atrazine/desethyl-atrazine and elevated concentrations of nitrate in ground water in Colorado

[LLR, log-likelihood ratio of logistic regression model; DF, degrees of freedom of log-likelihood ratio; LLR-P, chi-squared p-value calculated from log-likelihood ratio; McF, McFadden's rho-squared calculated with logistic regression; PC, percentage of correct predictions; <, less than]

Model	LLR	DF	LLR-P	McF	PC
Atrazine/desethyl-atrazine, with atrazine use	214	13	<0.001	0.338	0.808
Atrazine/desethyl-atrazine, without atrazine use	215	14	<.001	.339	.808
Nitrate, with fertilizer use	243	12	<.001	.259	.677
Nitrate, without fertilizer use	232	11	<.001	.248	.670

hydrogeomorphic region variables were modeled as a group, and the overall effect on the model was determined by observing the partial-likelihood ratios calculated before and after their addition to the model. The chi-squared p-value calculated from the difference of the partial-likelihood ratios was less than 0.001, denoting that the inclusion of hydrogeomorphic regions significantly improved the atrazine/DEA model overall. Individual p-values of the hydrogeomorphic region variables ranged from 0.010 to 0.964 (table 3); those p-values are related to the percentage of detections of atrazine measured in each hydrogeomorphic region. Hydrogeomorphic regions with a large percentage of atrazine/DEA detections (fig. 10) tend to have lower p-values and larger coefficients in the atrazine/DEA logistic regression models (table 3), because stronger relations occur in regions with detections.

The positive and negative signs of the model coefficients are consistent with expectations. For instance, the relation between the percentage of row crops and atrazine/DEA detections was positive (table 3), which is to be expected because the greatest atrazine use occurs in irrigated row crops. On the other hand, the relation between the percentage of small grains and atrazine/DEA detections was negative. The small grains land-cover classification comprises mostly dryland grains on which little or no atrazine is used; as the percentage of small grains land cover increases, the occurrence of atrazine/DEA detections decreases. The relation between clay content and organic-matter content of soils and atrazine/DEA detections was negative. Increased clay and organic-matter content in soil decreases the likelihood that atrazine/DEA will reach the water table.

To help confirm that the atrazine/DEA models are calibrated to the ground-water quality data, regressions were made between the percentage of actual atrazine/DEA detections and the predicted probability of atrazine/DEA detections (fig. 11). The percentage of actual atrazine/DEA detections was determined by dividing the predicted probabilities for the entire study area into groupings of 10 percent (0 to 10 percent, greater than 10 to 20 percent, greater than 20 to 30 percent, and so on). The percentage of atrazine detections within each group then was calculated and included in the regressions shown in figure 11. Both atrazine/DEA models showed good calibration, with r-squared values larger than 0.95.

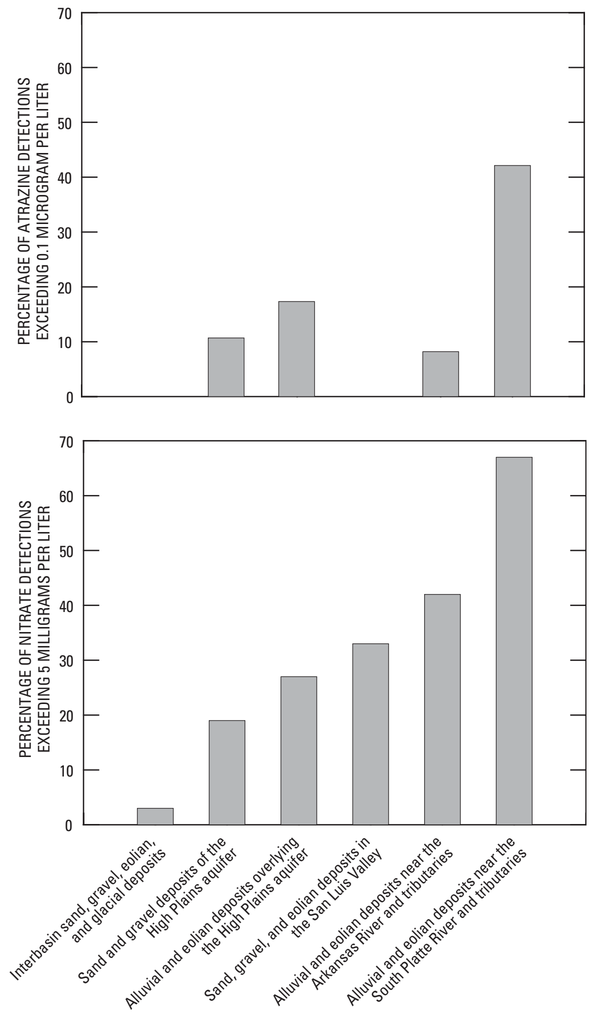


Figure 10. Percentage of atrazine and nitrate detections in ground-water samples collected in hydrogeomorphic regions in Colorado.

Development of Nitrate Model

Logistic regression models were developed to predict the probability of detecting elevated concentrations of nitrate by using the same methods as those used to develop the atrazine/DEA models. Six models were developed. Three different nitrate concentrations were tested to determine which concentration produced the best models—greater than 2 mg/L, greater than 10 mg/L, and greater than 5 mg/L. Background nitrate concentrations are generally less than 2 mg/L (Nolan and others, 1998) and 10 mg/L is the maximum contaminant level for nitrate (U.S. Environmental Protection Agency, 2000). Because 5 mg/L is one-half the maximum contaminant level, the presence of nitrate at this concentration may be a useful indicator for identifying problem areas before concentrations exceed the maximum contaminant level. Two models were developed for each of these concentrations: one with and one without nitrogen fertilizer use. Nitrate data were converted to binary classifications of “zero” for nondetection and “one” for detection at each of the three nitrate concentrations, and multivariate models were constructed using logistic regression.

The most significant logistic regression models were developed at a nitrate concentration greater than 5 mg/L (fig. 12); the r-squared values of the regressions were the largest. The log-likelihood was much larger for the models developed at greater than 5 mg/L (greater than 230) than the models developed at greater than 2 mg/L (less than 163) or greater than 10 mg/L (less than 184). Soil organic matter and occurrence of hydric soils became insignificant in the models developed at greater than 2 mg/L and 10 mg/L, which severely limits the accuracy of the resulting maps. The numbers of detections and nondetections of nitrate were evenly distributed at greater than 5 mg/L, which probably aided the development of the most significant models at this concentration.

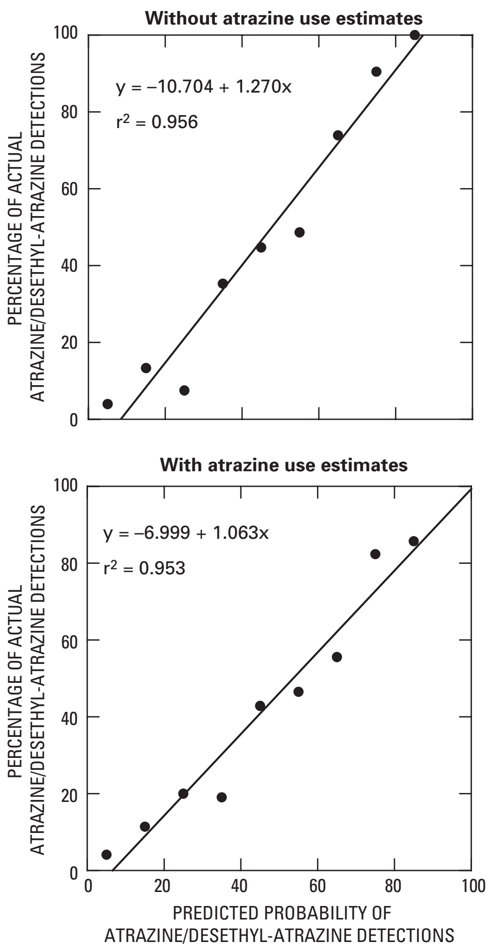


Figure 11. Percentage of actual atrazine/desethyl-atrazine detections at or greater than 0.1 microgram per liter and the predicted probability of atrazine/desethyl-atrazine detections at or greater than 0.1 microgram per liter using ground-water quality data collected by the State of Colorado (calibration data).

Validation of the Atrazine/Desethyl-atrazine and Nitrate Models

To validate the models, comparisons were made with an independent set of ground-water quality data collected by the USGS. The atrazine/DEA ground-water quality data collected by the USGS had a lower laboratory reporting level (0.001 µg/L) than the data collected by the State of Colorado (0.1 µg/L). To make the data comparable, the USGS data were converted to binary classifications of “zero” for nondetection and “one” for detection at or greater than a concentration of 0.1 µg/L before validating the models. Relations of the atrazine/DEA models with ground-water quality data collected by the USGS (validation data) had an r-squared of 0.849 for the model with atrazine use and an r-squared of 0.792 for the model without atrazine use (fig. 13). The r-squared values from the validation data set are smaller than from the calibration data set probably because the validation data set had a lower overall number of wells and a lower percentage of atrazine/DEA detections; out of 228 wells, water from only 32 wells had atrazine/DEA detections at or greater than 0.1 µg/L. Relations with the validation data set would probably have larger r-squared values if additional ground-water quality data were available.

Validations of the nitrate models with an independent set of ground-water quality monitoring data collected by the USGS were good; an r-squared of 0.966 was calculated for the model without fertilizer use, and an r-squared of 0.910 was calculated for the model with fertilizer use (fig. 14). USGS data were converted to binary classifications of “zero” for no detection and “one” for detection greater than a concentration of 5 mg/L before validating the models.

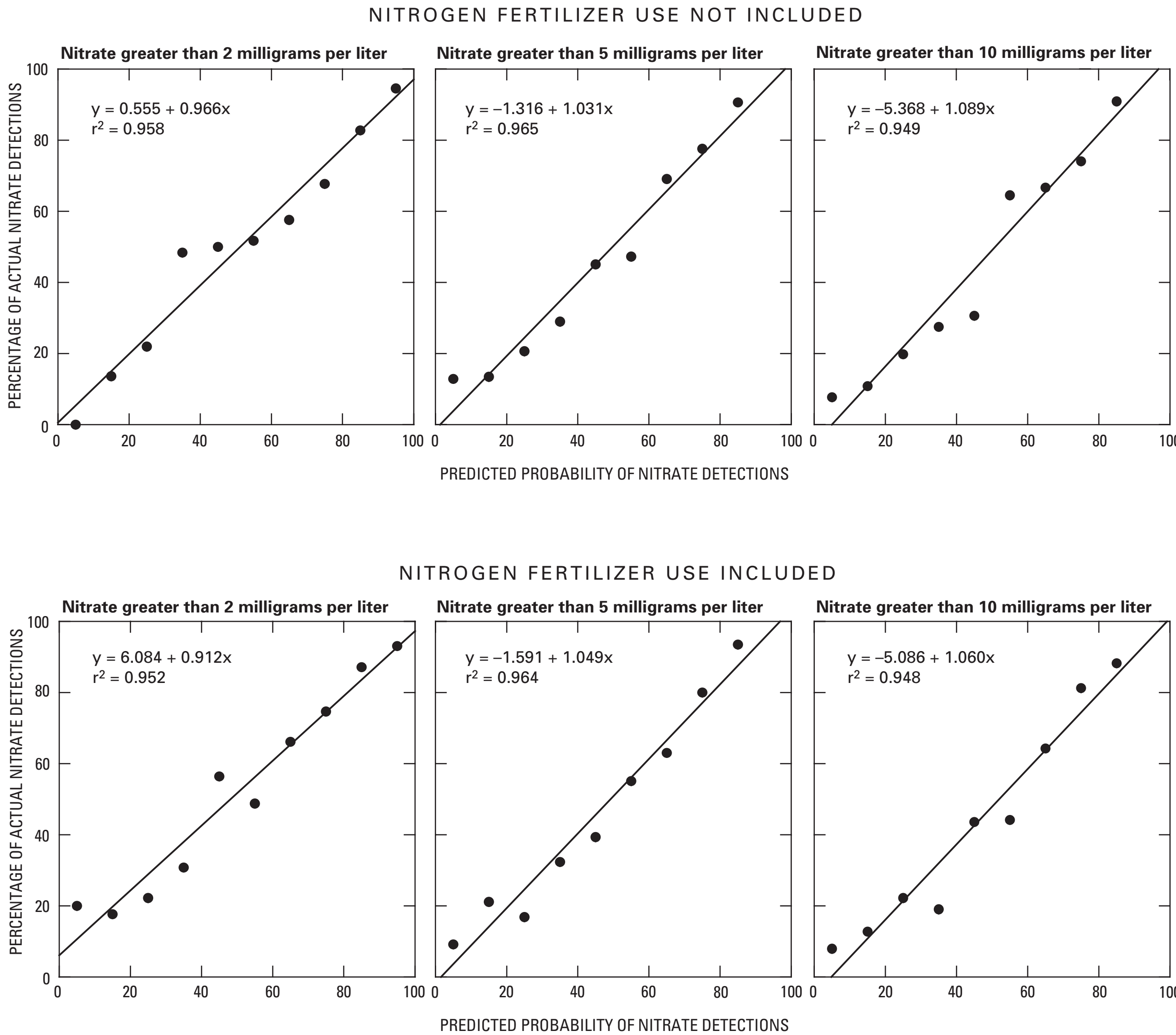


Figure 12. Percentage of actual detections of nitrate greater than concentrations of 2, 5, and 10 milligrams per liter and the predicted probability of detecting elevated concentrations of nitrate from models that did and did not include fertilizer use, using ground-water quality data collected by the State of Colorado (calibration data).

Construction of Atrazine/Desethyl-atrazine and Nitrate Probability Maps

Maps showing the probability of ground water in Colorado of detecting atrazine/DEA at or greater than concentrations of 0.1 µg/L (fig. 15) and nitrate concentrations greater than 5 mg/L (fig. 16) were constructed using the logistic regression models. Before constructing the maps, all GIS data were converted to grids with 60-meter spacing. Then, the logistic regression models similar to equation 1 were entered into a GIS and a probability rating was calculated for each of the approximately 78,000,000 grid nodes across Colorado. The atrazine/DEA and nitrate models that did not incorporate chemical (atrazine and nitrogen fertilizer) use are shown in figures 15 and 16. The models that incorporate chemical use are not shown in this publication because the differences between the maps with and without chemical use are not visible at the scale the maps are presented.

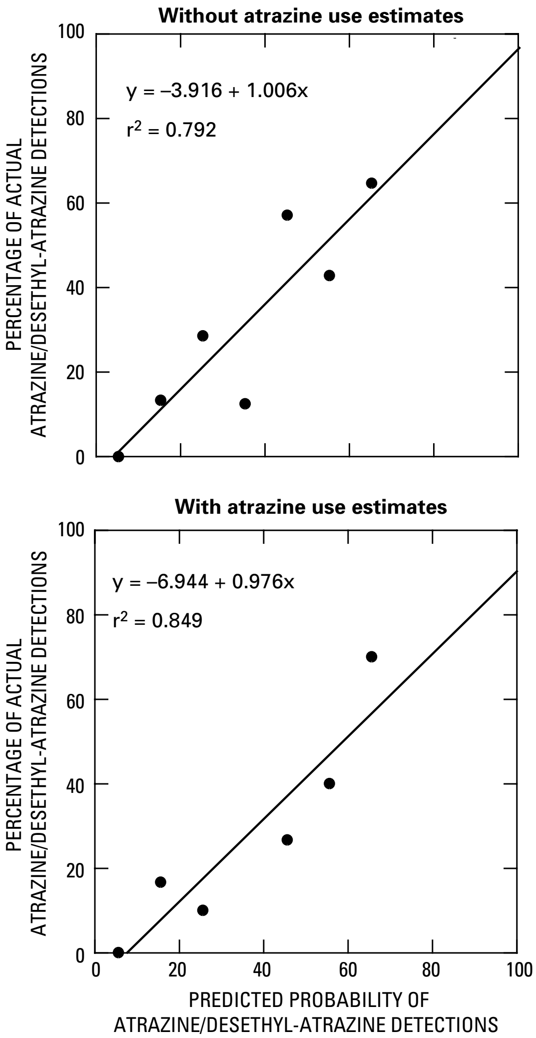


Figure 13. Percentage of actual atrazine/desethyl-atrazine detections at or greater than 0.1 microgram per liter and the predicted probability of detecting atrazine/desethyl-atrazine at or greater than 0.1 microgram per liter, using ground-water quality data collected by the U.S. Geological Survey (validation data).

MODEL RESULTS

Atrazine-use estimates by the CDA (Robert Wawrzynski, Colorado Department of Agriculture, written commun., 2001) and Battaglin and Goolsby (1994) were evaluated to determine which estimates were most closely related with atrazine/DEA detections in ground water. The estimates by Battaglin and Goolsby (1994) had more significant relations with atrazine/DEA detections in ground water than the CDA data, probably because the CDA data were based on use estimates by commercial applicators but did not include estimates of atrazine use by farmers.

The probabilities of detecting atrazine/DEA and nitrate in urban areas may be anomalously low. In the State of Colorado data set, a limited number of wells were near urban areas, which may have caused the logistic regression model to give urban areas an anomalously low probability rating. The validity and accuracy of the probability maps in urban areas may be improved in the future if additional ground-water quality data become available.

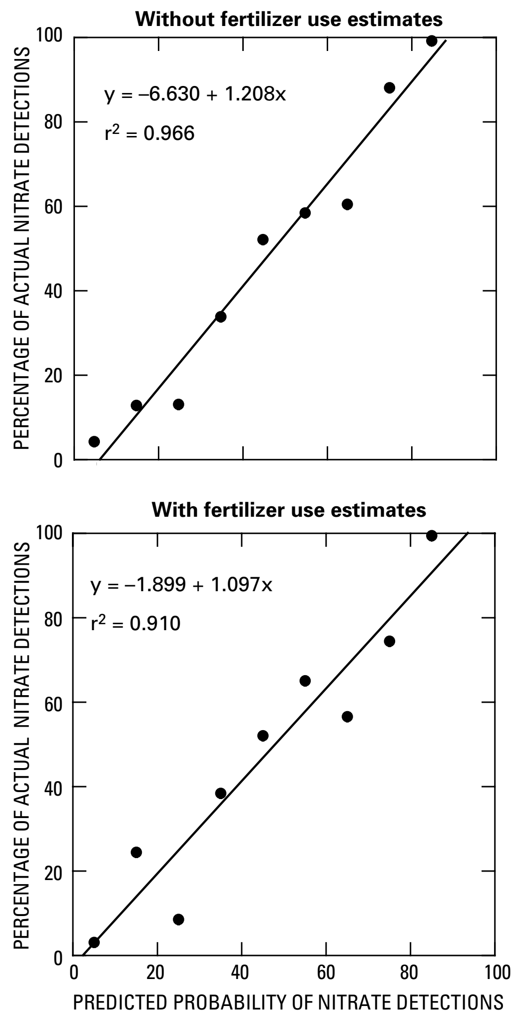


Figure 14. Percentage of actual detections of nitrate greater than 5 milligrams per liter and the predicted probability of nitrate detections greater than 5 milligrams per liter, using ground-water quality data collected by the U.S. Geological Survey (validation data).

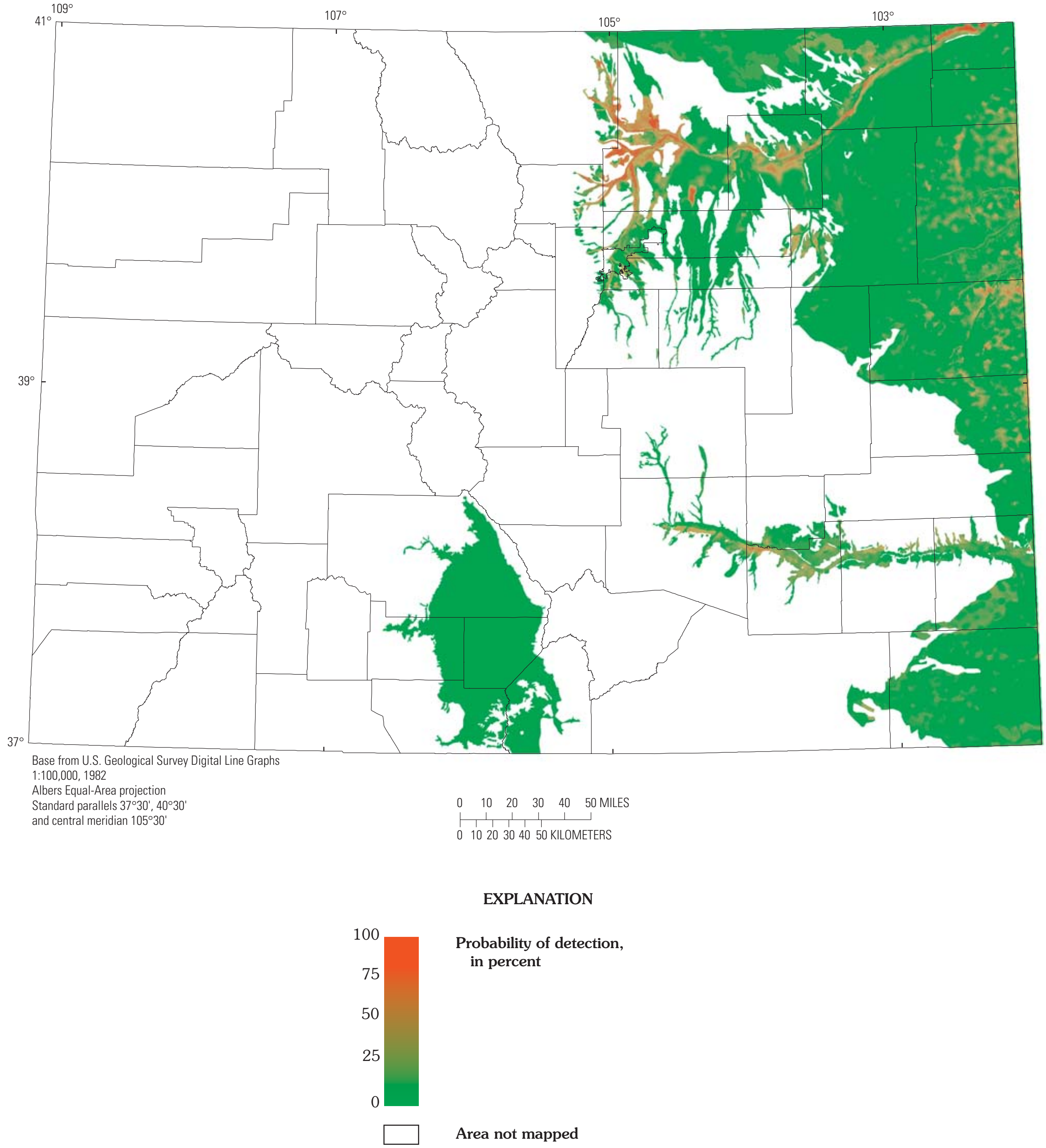


Figure 15. Probability of detecting atrazine/desethyl-atrazine concentrations at or greater than 0.1 microgram per liter in ground water in Colorado, atrazine use not included.

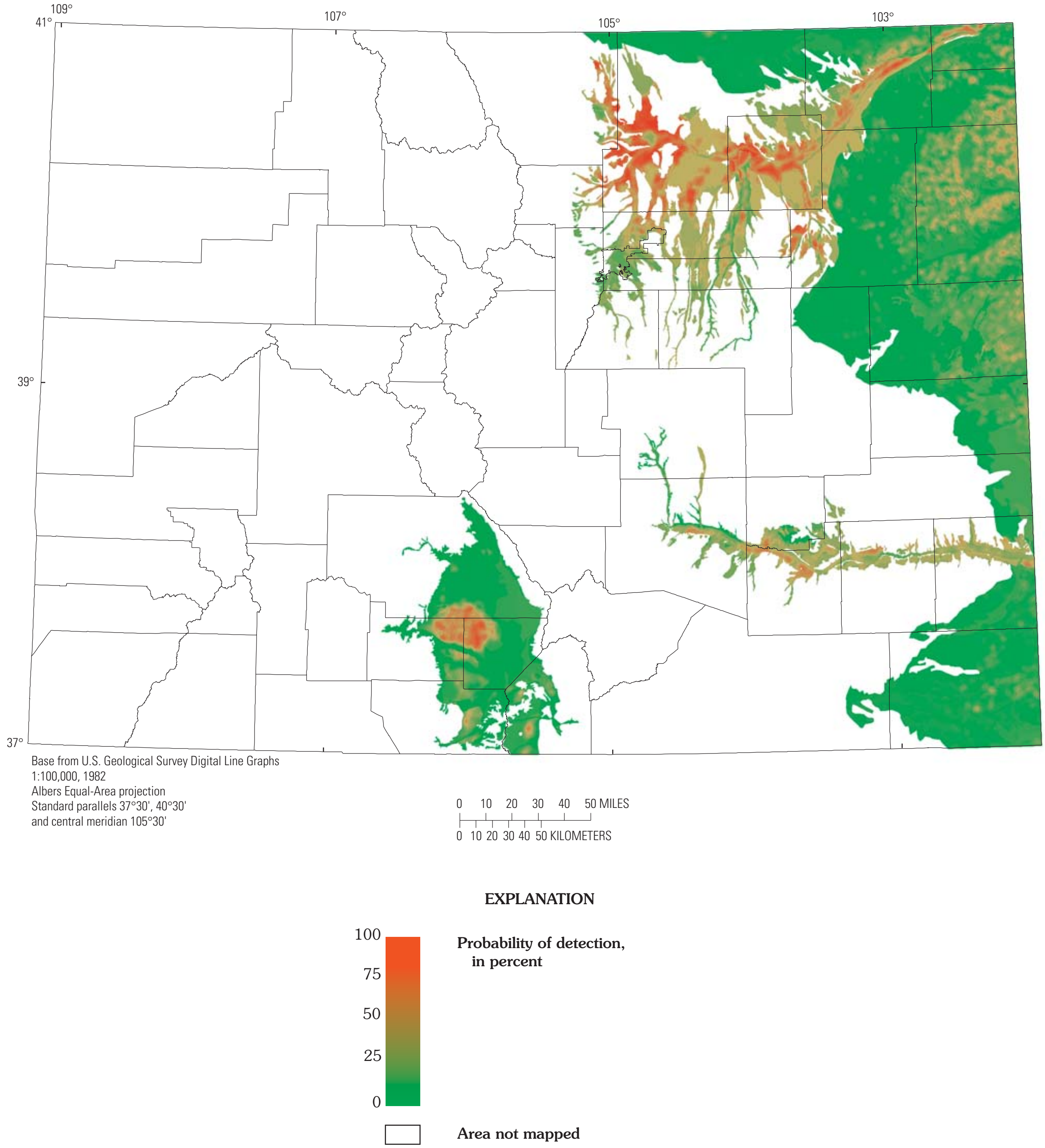


Figure 16. Probability of detecting nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado, fertilizer use not included.

The logistic regression models were significantly improved when more than one soil characteristic was incorporated. Acceptable models were developed using only one soil characteristic, but the models were much more significant when several soil characteristics were used. Contaminant migration through the soil is affected by complex interactions with the soil; using more than one soil characteristic can constrain these interactions and reduce the statistical variability in the model.

The same ground-water monitoring, hydrogeologic, and anthropogenic data used to create the atrazine/DEA models were used to create the nitrate models, thus allowing a direct comparison between atrazine/DEA and nitrate models. For the most part, the same independent variables that were significant in the atrazine/DEA models were significant in the nitrate models (table 3); hydrogeomorphic regions, available water capacity of soils, clay content of soils, occurrence of hydric soils, organic matter content of soils, and percentage of row crops within a 2,000-meter buffer were important variables in the atrazine/DEA and nitrate models. Percentage of low-intensity residential land within a 500-meter buffer, percentage of pasture/hay land cover within a 2,000-meter buffer, and percentage of small grains land cover within a 2,000-meter buffer were significant variables in the atrazine/DEA model without atrazine use but were not significant in either of the nitrate models (table 3). The differences between atrazine/DEA and nitrate models were attributed to differences in the extent and rates of application of these compounds, and possibly to differences in the chemical behavior of these compounds in the environment.

Hydrogeomorphic regions are significant in the atrazine/DEA and nitrate models because a large variation in percentage of atrazine/DEA detections and elevated concentrations of nitrate occurs between the various hydrogeomorphic regions of Colorado (fig. 10). Hydrogeomorphic regions primarily are a surrogate for aquifer lithology in the logistic regression models, but chemical use, climate, and cropping practices also vary regionally among the hydrogeomorphic regions. Hydrogeomorphic regions are an important variable because they allow logistic regression to compartmentalize the variability in ground-water quality conditions observed in an area as large as Colorado, and they allow logistic regression to build statistical models that are significant over a wide variety of environmental conditions. One advantage to using hydrogeomorphic regions is that the probability maps are calibrated to current ground-water quality conditions within each region, creating models that are more accurate under relatively local conditions. One disadvantage is that using hydrogeomorphic regions may give some areas anomalously low probability ratings in regions where a particular chemical compound or contaminant has not yet been used, or where chemical use increased but an insufficient amount of time has passed for those compounds to reach the ground water. Anomalously low probability ratings may occur because the compound has not yet been detected in ground water, so logistic regression would assign low probability to that hydrogeomorphic region.

Models that excluded hydrogeomorphic regions (figs. 17 and 18) were constructed to allow a comparison with models that did include such regions (figs. 15 and 16). In the atrazine/DEA and nitrate models that excluded hydrogeomorphic regions, the probability ratings tend to decrease in the South Platte River Basin and increase in the High Plains and the San Luis Valley (figs. 17 and 18) because the land cover and soils variables are modeled equally across the entire State. Model validity and accuracy were significantly reduced in all models that excluded hydrogeomorphic regions (table 5). For instance, McFadden’s rho-squared and percentage of correct predictions for the atrazine/DEA model without atrazine use and including hydrogeomorphic regions were 0.339 and 0.808, respectively (table 4). When hydrogeomorphic regions and atrazine use were excluded from the model, McFadden’s rho-squared and percentage of correct predictions were reduced to 0.288 and 0.792, respectively (table 5). A negative relation with percentage of shrubland land cover within a 2,000-meter buffer and a negative relation with percentage of small grains land cover within a 2,000-meter buffer became significant when hydrogeomorphic regions were excluded from the atrazine/DEA and nitrate models (table 6).

Models incorporating hydrogeomorphic regions may be useful for predicting probability of detecting atrazine/DEA and elevated concentrations of nitrate in regions that have a relatively long history of atrazine use or nitrogen input from sources such as fertilizers and manure; models excluding hydrogeomorphic

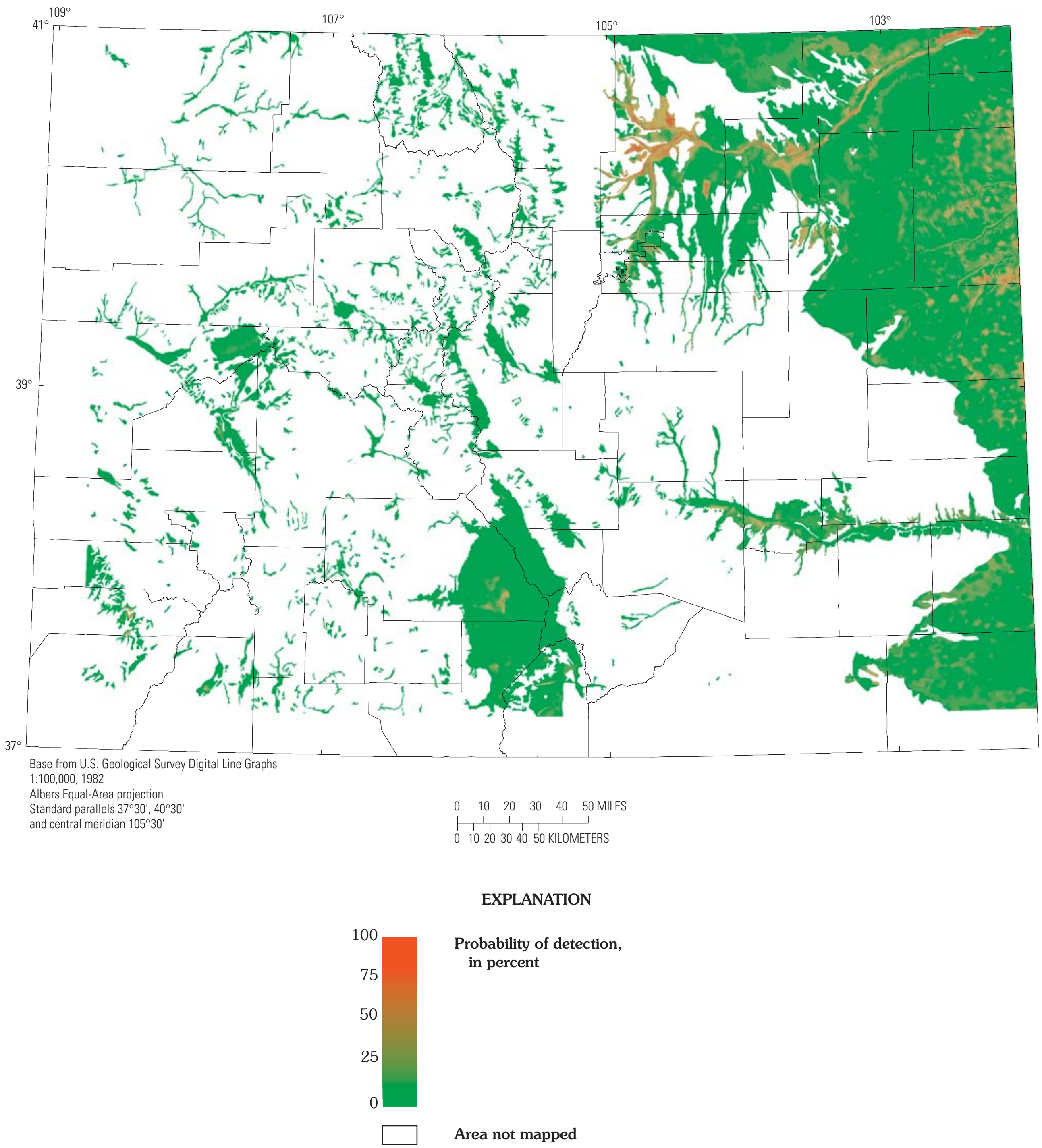


Figure 17. Probability of detecting atrazine/desethyl-atrazine concentrations at or greater than 0.1 microgram per liter in ground water in Colorado, atrazine use and hydrogeomorphic regions not included.

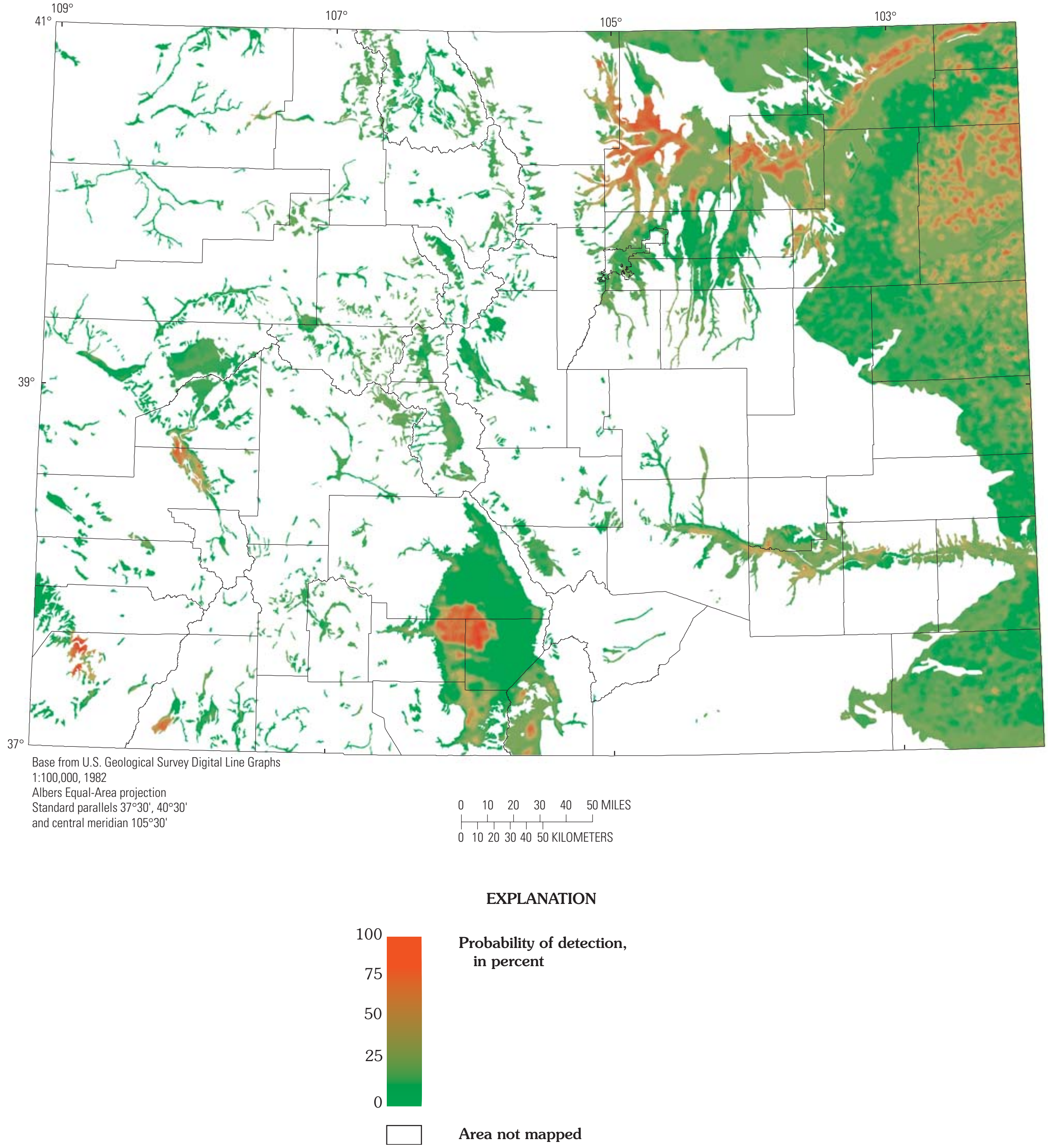


Figure 18. Probability of detecting nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado, fertilizer use and hydrogeomorphic regions not included.

Table 5. Statistical results from models that predict the probability of atrazine/desethyl-atrazine and elevated concentrations of nitrate in ground water in Colorado, hydrogeomorphic regions not included

[LLR, log-likelihood ratio of logistic regression model; DF, degrees of freedom of log-likelihood ratio; LLR-P, chi-squared p-value calculated from log-likelihood ratio; McF, McFadden's rho-squared calculated with logistic regression; PC, percentage of correct predictions; <, less than]

Model	LLR	DF	LLR-P	McF	PC
Atrazine/desethyl-atrazine, with atrazine use, without hydrogeomorphic regions	200	8	< 0.001	0.315	0.802
Atrazine/desethyl-atrazine, without atrazine use and without hydrogeomorphic regions	183	6	<.001	.288	.792
Nitrate, with fertilizer use, without hydrogeomorphic regions	188	7	<.001	.201	.648
Nitrate, without fertilizer use and without hydrogeomorphic regions	172	5	<.001	.184	.638

Table 6. Logistic regression coefficients and individual p-values of independent variables significantly related with the detection of atrazine/desethyl-atrazine concentrations at or greater than 0.1 microgram per liter and nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado, hydrogeomorphic regions not included

[--, no relation observed; values not enclosed in parentheses are logistic regression coefficients; value enclosed in parentheses are individual p-values; independent variables in bold are used in equation 1 for the final models; <, less than]

Independent variable	Atrazine/desethyl-atrazine model, with atrazine use, without hydrogeomorphic regions	Atrazine/desethyl-atrazine model, without atrazine use, without hydrogeomorphic regions	Nitrate model, with fertilizer use, without hydrogeomorphic regions	Nitrate model, without fertilizer use, without hydrogeomorphic regions
Logistic regression constant	**–4.955 (<0.001)**	**–4.042 (<0.001)**	**–1.680 (<0.001)**	**–1.139 (0.005)**
Depth to water in well	--	--	--	--
Total well depth	--	--	--	--
Surface elevation	--	--	--	--
Available water capacity of soil	**.404 (<.001)**	**.343 (<.001)**	**.110 (.025)**	**.074 (.105)**
Clay content of soil	**–.100 (.002)**	**–.110 (<.001)**	**–.037 (.046)**	**–.048 (.010)**
Soil liquid limit	--	--	--	--
Occurrence of hydric soils	**2.388 (.002)**	**2.483 (.001)**	--	--
Organic matter content of soil	**–1.161 (.080)**	--	**–.872 (.054)**	--
Soil permeability	--	--	--	--
Soil drainage	--	--	--	--
Soil erodibility	--	--	--	--
Soil hydrologic group	--	--	--	--
Soil surface slope	--	--	--	--
Soil thickness	--	--	--	--
Precipitation	--	--	--	--

Table 6. Logistic regression coefficients and individual p-values of independent variables significantly related with the detection of atrazine/desethyl-atrazine concentrations at or greater than 0.1 microgram per liter and nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado, hydrogeomorphic regions not included—Continued

[--, no relation observed; values not enclosed in parentheses are logistic regression coefficients; value enclosed in parentheses are individual p-values; independent variables in bold are used in equation 1 for the final models; <, less than]

Independent variable	Atrazine/ desethyl-atrazine model, with atrazine use, without hydrogeomorphic regions	Atrazine/ desethyl-atrazine model, without atrazine use, without hydrogeomorphic regions	Nitrate model, with fertilizer use, without hydrogeomorphic regions	Nitrate model, without fertilizer use, without hydrogeomorphic regions
Percentage of low-intensity residential land cover within a 500-meter buffer	--	--	--	--
Percentage of high-intensity residential land cover within a 500-meter buffer	--	--	--	--
Percentage of commercial/industrial/transportation land cover within a 500-meter buffer	--	--	--	--
Percentage of urban/recreational grasses land cover within a 500-meter buffer	--	--	--	--
Percentage of shrubland within a 2,000-meter buffer	**–.273 (.005)**	**–.409 (<.001)**	**–.029 (.004)**	**–.035 (.001)**
Percentage of pasture/hay land cover within a 2,000-meter buffer	--	--	--	--
Percentage of row crops land cover within a 2,000-meter buffer	**.020 (.002)**	**.028 (<.001)**	**.034 (<.001)**	**.040 (<.001)**
Percentage of small grains land cover within a 2,000-meter buffer	**–.094 (<.001)**	**–.093 (<.001)**	**–.031 (.003)**	**–.027 (.009)**
Percentage of fallow land cover within a 2,000-meter buffer	--	--	--	--
Atrazine use	**.800 (<.001)**	--	--	--
Fertilizer use	--	--	**.056 (<.001)**	--

regions may be more useful for extending probability ratings into regions where atrazine has not been used or sources of nitrogen have not existed. For instance, the San Luis Valley has a low probability of atrazine/DEA detections in the model that incorporates hydrogeomorphic regions (fig. 15). Very little, if any, atrazine is used on crops in the San Luis Valley, and atrazine/DEA have not been detected in ground water in the San Luis Valley (figs. 4, 6, and 10). The lack of atrazine/DEA detections in the San Luis Valley caused logistic regression to give the San Luis Valley low probability ratings. Atrazine/DEA probability ratings in the San Luis Valley are larger in the model that does not include hydrogeomorphic regions (fig. 17), indicating that portions of the valley may have a potential for atrazine/DEA detections in ground water if atrazine is used there in the future.

Only 40 wells were sampled in the interbasin sand, gravel, eolian, and glacial deposits hydrogeomorphic region, and in many of the basins no wells were sampled (figs. 4, 5, and 6). The interbasin sand, gravel, eolian, and glacial deposits hydrogeomorphic region was excluded from the final atrazine/DEA and nitrate probability maps that included hydrogeomorphic regions (figs. 15 and 16) because insufficient ground-water quality data were available to calibrate the atrazine/DEA and nitrate models for this hydrogeomorphic region. The interbasin sand, gravel, eolian, and glacial deposits hydrogeomorphic region was included in the atrazine/DEA and nitrate models that exclude hydrogeomorphic regions, however, because these models are based upon correlations with ground-water quality across the entire State (figs. 17 and 18) and not in particular hydrogeomorphic regions.

USE OF THE ATRAZINE/ DESETHYL-ATRAZINE AND NITRATE MODELS FOR ALACHLOR, METOLACHLOR, AND SIMAZINE

The State of Colorado may have to develop chemical-specific PMP's for alachlor, metolachlor, and simazine. Relations between atrazine/DEA detections and nitrate concentrations greater than 5 mg/L with alachlor, metolachlor, and simazine were evaluated to determine if the atrazine/DEA or nitrate models could be used as a surrogate for alachlor, metolachlor, or simazine detections.

Logistic regression models for alachlor, metolachlor, and simazine could not be constructed because an insufficient number of detections were available in the data set.

The co-occurrence of atrazine/DEA detections and nitrate concentrations greater than 5 mg/L with alachlor, metolachlor, and simazine was observed. Ground-water quality data collected by the USGS were used because of the higher frequency of detections due to the lower laboratory reporting limits (table 2). Out of 57 samples with detections of simazine, atrazine was detected in 52 samples (tables 2 and 7), indicating a relatively large co-occurrence. Metolachlor had a less frequent co-occurrence with atrazine; out of 36 samples with detections of metolachlor, atrazine was detected in 28 samples. Co-occurrence of metolachlor and simazine with nitrate concentrations greater than 5 mg/L was less frequent (table 7). There were insufficient alachlor detections to determine if there is co-occurrence with atrazine/DEA and nitrate concentrations greater than 5 mg/L.

A Spearman correlation matrix was calculated to quantify correlations between alachlor, atrazine, DEA, atrazine/DEA combined, metolachlor, simazine, and nitrate concentrations greater than 5 mg/L (table 8). USGS data were used for the correlation; data were converted to binary coding of "zero" for nondetection and "one" for detection to reduce the influence of concentrations in the correlations. Two factors are strongly correlated when the result from a Spearman correlation approaches plus or minus one; no correlation exists as the Spearman correlation approaches zero (Helsel and Hirsch, 1992). A strong correlation between atrazine, DEA, and atrazine/DEA combined (table 8) verifies co-occurrence; atrazine and DEA should be strongly correlated because DEA is a breakdown product of atrazine. Simazine had Spearman correlations of about 0.6 with atrazine, DEA, and atrazine/DEA combined, indicating a relatively strong correlation. All other correlations were insignificant. The ground-water quality data also were subdivided into individual hydrogeomorphic regions, but the Spearman correlations were not significantly different than those presented in table 8.

The Spearman correlation matrix indicates that the atrazine/DEA and nitrate models probably are not valid for metolachlor and simazine. Additional comparisons between the occurrence of atrazine/DEA and metolachlor and simazine are needed before reaching a final conclusion on the use of the atrazine/DEA model as a surrogate for metolachlor and simazine. The differences in chemical use should be taken into account, as well as the differences or similarities in chemical characteristics of the compounds, such as solubility and sorption coefficient.

Table 7. Number of wells with co-occurrence of alachlor, metolachlor, and simazine with atrazine detections and nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado using ground-water quality data analyzed by the U.S. Geological Survey

[mg/L, milligrams per liter]

Pesticide	Co-occurrence of atrazine and desethyl-atrazine detections	No co-occurrence of atrazine and desethyl-atrazine detections	Co-occurrence of nitrate concentrations greater than 5 mg/L	No co-occurrence of nitrate concentrations greater than 5 mg/L
Alachlor	1	0	1	0
Metolachlor	28	8	24	11
Simazine	52	5	36	21

Table 8. Spearman correlation matrix showing the co-occurrence of alachlor, atrazine, desethyl-atrazine, atrazine and desethyl-atrazine combined, metolachlor, simazine, and nitrate concentrations greater than 5 milligrams per liter in ground water in Colorado using ground-water quality data analyzed by the U.S. Geological Survey

[<, less than]

Compound	Alachlor	Atrazine	Desethyl-atrazine	Atrazine and desethyl-atrazine combined	Metolachlor	Simazine	Nitrate
Alachlor	1.000						
Atrazine	.088	1.000					
Desethyl-atrazine	.093	**.797**	1.000				
Atrazine and desethyl-atrazine combined	.082	**.927**	**.883**	1.000			
Metolachlor	–.029	.363	.362	.325	1.000		
Simazine	–.039	**.621**	**.549**	**.609**	.342	1.000	
Nitrate	.093	.388	.368	.348	.311	.355	1.000
Number of detections/ number of analyses (percent detections)	1/228 (<1)	83/228 (36)	78/228 (34)	91/228 (40)	36/228 (16)	57/228 (25)	302/333 (91)

APPROPRIATE USES OF THE PROBABILITY MAPS, AND SUGGESTIONS FOR IMPROVEMENTS

The probability maps developed by the methods described in this report are designed to portray the potential for contamination of ground water in Colorado. These maps do not show areas that are actually (currently) contaminated, but rather, the areas that have a potential (or likelihood) for being contaminated if a contaminant was released to the environment. More specifically, each map shows the probability of detection (in terms of a percent) of a particular chemical compound (the contaminant) in ground water. Probability is not the same as certainty; a well in a high probability area is not necessarily contaminated because contamination also can depend on well depth and other local factors not taken into account by the models described in this report.

The atrazine/DEA and nitrate probability maps were specifically developed for use by the State of Colorado in their PMP to help provide a sound hydrogeologic basis for the management of atrazine and nitrogen in Colorado. The maps are intended to be a first approximation at developing a consistent rating method for the entire State. Additional site-specific data are needed before site-specific decisions are made, such as pesticide-use restrictions. The most appropriate uses of these maps are to focus prevention programs in areas of greatest concern, to focus ground-water sampling programs in areas of greatest potential for contamination, and to assist educational programs for ground-water quality protection.

The probability maps should not be used at a scale any larger than 1:250,000. The soils data had the smallest scale (1:250,000) of all the independent variables used in the models; therefore, the resulting probability maps should not be used at a larger scale. Soils data at a larger scale were not available for all regions of Colorado.

The accuracy of the probability maps would be improved if (1) larger scale soils data were available in digital form, (2) more complete and detailed chemical-use data were available, and (3) a larger number and wider distribution of ground-water quality data, particularly in rangeland areas, were available. Data differentiating between sprinkler and flood irrigation methods would probably improve the accuracy in agricultural areas. Some site-specific variables, such as improper well construction and local spills of contaminants, were not accounted for in the models. Accounting for ground-water flow direction was beyond the scope of this study.

Depth to ground water was not evaluated by this study because a large-scale statewide coverage was not available. Hall (1998) incorporated a depth to water table coverage in a sensitivity map, but the scale of the data was too small for use in this study (1-kilometer by 1-kilometer grid cells). Depth to ground water may be an important coverage to develop for future studies. Rupert (1998) observed a significant relation between depth to ground water and elevated concentrations of nitrate in the eastern Snake River Plain of Idaho, but significant relations with atrazine/DEA were not observed.

SUMMARY

Draft Federal regulations may require that each State develop a State Pesticide Management Plan (PMP) for the herbicides atrazine, alachlor, metolachlor, and simazine. The Colorado Agricultural Chemicals and Groundwater Protection Program, a cooperative effort of the Colorado Department of Agriculture (CDA), the Colorado Department of Public Health and Environment (CDPHE), and the Colorado State University Cooperative Extension (CSUCE), is developing a PMP for each of the herbicides and would benefit from a map that could be used to predict the probability of detecting atrazine and desethyl-atrazine (DEA) in ground water. The map could be incorporated into the PMP and provide a sound hydrogeologic basis for atrazine management in Colorado. Nitrate also has been identified as a contaminant of concern in ground water in Colorado. To address these needs, the U.S. Geological Survey (USGS), in cooperation with the CDA, CDPHE, and CSUCE, conducted a study to develop maps to predict the probability of detecting atrazine and(or) DEA and elevated concentrations of nitrate in ground water in Colorado.

Maps showing the probability of detecting atrazine and(or) DEA (atrazine/DEA) concentrations at or greater than 0.1 μg/L and nitrite plus nitrate as nitrogen (nitrate) concentrations greater than 5 mg/L in ground water in Colorado were developed using logistic regression statistical methods as follows: (1) Ground-water quality data were overlaid with anthropogenic and hydrogeologic data by using a geographic information system to produce a data set in which each well had corresponding data on atrazine use, fertilizer use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well construction. These data then were downloaded to a statistical software package for analysis by logistic regression techniques; (2) relations were observed between ground-water quality and the percentage of land-cover categories within circular regions (buffers) around wells. Several buffer sizes were evaluated; the buffer size that provided the strongest relation was selected for use in the logistic regression models; (3) relations between atrazine/DEA and nitrate in ground water and atrazine use, geology, hydrogeomorphic regions, land cover, precipitation, soils, and well-construction data were evaluated and several preliminary multivariate models with various combinations of independent variables were constructed; (4) the multivariate models that best predicted the presence of atrazine/DEA and elevated concentrations of nitrate in ground water were selected; (5) the accuracy of the multivariate models was confirmed by validating the models with an independent set of ground-water quality data; and (6) the multivariate models were entered into a geographic information system and probability maps were constructed.

Two models that best predicted the probability of detecting atrazine/DEA in ground water were selected: one with and one without atrazine use. Relations of the predicted probability of atrazine/DEA in ground water with the percentage of actual detections were good; r-squared values were 0.953 and 0.956. Models were validated using a second set of ground-water quality data (validation data). Relations of the predicted probability of atrazine/DEA in ground water with the percentage of actual detections in the validation data set had an r-squared of 0.849 for the model with atrazine use and an r-squared of 0.792 for the model without atrazine use. Smaller r-squared values from the validation data set than from the calibration data set probably result from a smaller number of wells and a lower percentage of atrazine/DEA detections in the validation data set; out of 228 wells, water from only 32 wells had atrazine/DEA detections at or greater than 0.1 μg/L. Relations with the validation data set would probably have larger r-squared values if additional ground-water quality data were available.

Logistic regression also was used to develop models that predict the probability of elevated concentrations of nitrate in ground water in Colorado. Three concentration ranges were tested: 2, 5, and 10 mg/L. The models predicting the probability of nitrate detections at concentrations greater than 5 mg/L were the most significant models, probably because the numbers of detections and nondetections were evenly distributed at that concentration. Two models that best predicted the probability of nitrate

concentrations greater than 5 mg/L in ground water were selected: one with and one without fertilizer use. Relations of the predicted probability of nitrate concentrations greater than 5 mg/L in ground water with the percentage of actual detections were good; r-squared values were 0.964 and 0.965. Models were validated using a second set of ground-water quality data (validation data). Relations of the predicted probability of nitrate concentrations greater than 5 mg/L in ground water with the percentage of actual detections in the validation data set were good; r-squared values were 0.910 and 0.966.

Models that excluded hydrogeomorphic regions were constructed to allow a comparison with models that included these regions. In the atrazine/DEA and nitrate models that excluded hydrogeomorphic regions, the probability ratings tend to decrease in the South Platte River Basin and increase in the High Plains and the San Luis Valley because the land cover and soils variables are modeled equally across the entire State. Model validity and accuracy was significantly reduced in all models that excluded hydrogeomorphic regions, confirming that hydrogeomorphic regions are a significant variable. Models incorporating hydrogeomorphic regions may be useful for predicting probability in regions that have a relatively long history of atrazine use or nitrogen input from sources such as fertilizers and manure; models excluding hydrogeomorphic regions may be more useful for extending probability ratings into regions where atrazine has not been used or sources of nitrogen have not existed. For instance, the San Luis Valley has a low probability of atrazine/DEA detections in the model that incorporates hydrogeomorphic regions. Very little, if any, atrazine is used on crops from the San Luis Valley, and atrazine/DEA have not been detected in ground water in the San Luis Valley. The lack of atrazine/DEA detections in the San Luis Valley caused logistic regression to give the San Luis Valley low probability ratings. Atrazine/DEA probability ratings in the San Luis Valley are larger in the model that does not include hydrogeomorphic regions, indicating that portions of the valley may have a potential for atrazine/DEA detections in ground water if atrazine is used there in the future.

The probability maps developed by the methods described in this report are designed to portray the potential for contamination of ground water in Colorado. These maps do not show areas that are actually (currently) contaminated, but rather, the areas that have a potential (or likelihood) for being contaminated if a contaminant were released to the environment. More specifically, each map shows the probability of detection (in terms of a percent) of a particular chemical compound (the contaminant) in ground water. Probability is not the same as certainty; a well in a high probability area is not necessarily contaminated because contamination can also depend on well depth and other local factors not taken into account by the models described in this report.

The maps produced by this project were developed using the best available data, and can probably be improved as additional data become available. Additional ground-water quality data in areas that are sparsely sampled would probably improve the calibration of the maps. Larger scale soils and chemical input data would probably improve the accuracy of the probability maps. Data differentiating between sprinkler and flood irrigation methods would probably improve the accuracy in agricultural areas. Chemical-use data in urban areas would improve the accuracy of the probability maps in urban areas.

REFERENCES CITED

Aller, Linda, Bennett, Truman, 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, Robert S. Kerr Environmental Research Laboratory, Office of Research and Development, EPA/600/2–85/018, 163 p.

Anderholm, S.K., 1996, Water-quality assessment of the Rio Grande Valley, Colorado, New Mexico, and Texas—Shallow ground-water quality of a land-use area in the San Luis Valley, south-central Colorado, 1993: U.S. Geological Survey Water-Resources Investigations Report 96–4144, 94 p.

Apodaca, L.E., 1997, Analysis of ground-water-quality data of the Upper Colorado River Basin, water years 1972–1992: U.S. Geological Survey Water-Resources Investigations Report 97–4240, 31 p.

Barbash, J.E., and Resek, E.A., 1996, Pesticides in ground water—Distribution, trends, and governing factors: Chelsea, Mich., Ann Arbor Press, Inc., 588 p.

Battaglin, W.A., and Goolsby, D.A., 1994, Spatial data in GIS format on agricultural chemical use, land use, and cropping practices in the United States: U.S. Geological Survey Water-Resources Investigations Report 94–4176, 87 p., accessed July 3, 2001, from URL http://water.usgs.gov/lookup/getspatial?herbicide1

Bruce, B.W., and McMahon, P.B., 1998, Shallow ground-water quality of selected land-use/aquifer settings in the South Platte River Basin, Colorado and Nebraska 1993–95: U.S. Geological Survey Water-Resources Investigations Report 97–4229, 48 p.

Cederstrand, J.R., and Becker, M.F., 1999, Digital map of aquifer boundary 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 99–267, accessed July 3, 2001, from URL http://water.usgs.gov/lookup/getspatial?ofr99-267_aqbound

Chronic, Halka, 1980, Roadside geology of Colorado: Missoula, Mont., Mountain Press Publishing Co., 334 p.

Collins, D.L., 1991, Colorado *in* Paulson, R.W., Chase, E.B., Roberts, R.S., and Moody, D.W., eds., National water summary 1988–89—Hydrologic events and floods and droughts: U.S. Geological Survey Water-Supply Paper 2375, p. 207–214.

Daly, Christopher, Neilson, R.P., and Phillips, D.L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain: Journal of Applied Meteorology, 33, 140–158, accessed January 29, 2002, from URL http://www.ocs.orst.edu/prism/gen_toc.html Geospatial data accessed January 31, 2001, from URL http://www.ftw.nrcs.usda.gov/prism/prismdata_state.html

Dennehy, K.F., Litke, D.W., Tate, C.M., Qi, S.L., McMahon, P.B., Bruce, B.W., Limbrough, R.A., and Heiny, J.S., 1998, Water quality in the South Platte River Basin, Colorado, Nebraska, and Wyoming, 1992–95: U.S. Geological Survey Circular 1167, 38 p.

Druliner, A.D., Chen, H.H., and McGrath, T.S., 1996, Relations of nonpoint-source nitrate and atrazine concentrations in the High Plains aquifer to selected explanatory variables in six Nebraska study areas: U.S. Geological Survey Water-Resources Investigations Report 95–4202, 51 p.

Gianessi, Leonard P., and Puffer, Cynthia, 1991, Herbicide use in the United States: Washington, D.C., Resources for the Future, December 1990 (revised April 1991), 128 p.

Green, G.N., 1992, The digital geologic map of Colorado in ARC/INFO format: U.S. Geological Survey Open-File Report 92–507, metadata accessed January 31, 2001, from URL http://geo-nsdi.er.usgs.gov/metadata/open-file/92-507/metadata.faq.html

Hall, M.D., 1998, Relative sensitivity of Colorado ground-water to pesticide impact: Colorado Department of Agriculture, Agricultural Chemicals in Ground-Water Protection Program, Fact Sheet 16, 7 p.

Hearne, G.A., Wireman, M., Campbell, A., Turner, S., and Ingersoll, G.P., 1995, Vulnerability of the uppermost ground water to contamination in the greater Denver area, Colorado: U.S. Geological Survey Water-Resources Investigations Report 92–4143, 244 p., 1 pl.

Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: New York, Elsevier Science Publishing Company, Inc., 522 p., 1 diskette.

Hosmer, D.W., and Lemeshow, Stanley, 1989, Applied logistic regression: New York, John Wiley & Sons, Inc., 307 p.

Hurr, R.T., and Hearne, G.A., 1985, Colorado ground-water resources, *in* National water summary 1984—Hydrologic events, selected water-quality trends, and ground-water resources: U.S. Geological Survey Water-Supply Paper 2275, p. 153–160.

Hurr, R.T., and Moore, J.E., 1972, Hydrogeologic characteristics of the valley-fill aquifer in the Arkansas River valley, Bent County, Colorado: U.S. Geological Survey Hydrologic Investigations Atlas HA–0461, scale 1:62,500, 2 sheets.

Kleinbaum, D.G., 1994, Logistic regression, a self-learning text: New York, Springer-Verlag, 282 p.

Kolpin, D.W., Barbash, J.E., and Gilliom, R.J., 1998, Occurrence of pesticides in shallow ground water of the United States—Initial results from the National Water-Quality Assessment Program: Environmental Science and Technology, v. 32, no. 5, p. 558–566.

Koterba, M.T., Banks, W.S.L., and Shedlock, R.J., 1993, Pesticides in shallow groundwater in the Delmarva Peninsula: Journal of Environmental Quality, v. 22, no. 3, p. 500–518.

Litke, D.W., 1990, Colorado *in* Carr, J.E., Chase, E.B., Paulson, R.W., and Moody, D.W., eds., National water summary 1987—Hydrologic events and water supply and use: U.S. Geological Survey Water-Supply Paper 2350, p. 183–192.

McGrath, Timothy, and Dugan, J.T., 1993, Water-level changes in the High Plains aquifer—Predevelopment to 1991: U.S. Geological Survey Water-Resources Investigations Report 93–4088, 53 p.

National Research Council, 1993, Ground water vulnerability assessment, contamination potential under conditions of uncertainty: National Academy Press, Washington, D.C., 210 p., accessed December 15, 2000, from URL http://books.nap.edu/books/0309047994/html

Nelson, G.A., Hurr, R.T., and Moore, J.E., 1989a, Hydrogeologic characteristics of the valley-fill aquifer in the Arkansas River Valley, Prowers County, Colorado: U.S. Geological Survey Open-File Report 89–254, scale 1:62,500, 3 sheets.

———1989b, Hydrogeologic characteristics of the valley-fill aquifer in the Arkansas River Valley, Crowley and Otero Counties, Colorado: U.S. Geological Survey Open-File Report 89–255, scale 1:62,500, 3 sheets.

———1989c, Hydrogeologic characteristics of the valley-fill aquifer in the Arkansas River Valley, Pueblo County, Colorado: U.S. Geological Survey Open-File Report 89–256, scale 1:62,500, 3 sheets.

Nolan, B.T., and Clark, M.L., 1997, Selenium in irrigated agricultural areas of the western United States: Journal of Environmental Quality, v. 26, no. 3, p. 849–857.

Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Helsel, D.R., 1997, Risk of nitrate in ground waters of the United States—A national perspective: Environmental Science and Technology, v. 31, no. 8, p. 2229–2236.

Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Helsel, D.R., 1998, A national look at nitrate contamination in ground water: Water Conditioning and Purification, v. 39, no. 12, p. 76–79.

Rao, P.S.C., and Alley, W.M., 1993, Pesticides: New York, Van Nostrand Reinhold, 112 p.

Rupert, M.G., 1998, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate (NO_2+NO_3–N) in ground water in the Idaho part of the upper Snake River Basin: U.S. Geological Survey Water-Resources Investigations Report 98–4203, 32 p., 1 pl., accessed March 10, 2001, from URL http://idaho.usgs.gov/PDF/wri984203/index.html

———2001, Calibration of the DRASTIC ground water vulnerability mapping method: Ground Water, v. 39, no. 4, p. 625–630.

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, accessed March 10, 2001, from URL http://water.usgs.gov/lookup/getspatial?ussoils

Spahr, N.E., Apodaca, L.E., Deacon, J.R., Bails, J.B., Bauch, N.J., Smith, C. Michelle, and Driver, N.E., 2000, Water quality in the Upper Colorado River Basin, 1996–98: U.S. Geological Survey Circular 1214, 41 p.

SPSS, Inc., 2000, SYSTAT 10, Statistics I—Software documentation: Chicago, Ill., SPSS, Inc., 663 p.

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.

Tweto, Ogden, comp., 1979, Geologic map of Colorado: U.S. Geological Survey, scale 1:500,000, 1 sheet.

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

———1993, Soil survey manual: U.S. Department of Agriculture, Handbook 18, 437 p.

———1995, Soil Survey Geographic (SSURGO) data base: U.S. Department of Agriculture, Natural Resources Conservation Service, Miscellaneous Publication 1527, 31 p.

U.S. Department of Commerce, 1989, Census of agriculture, 1987—Final county file: U.S. Department of Commerce, Bureau of the Census [machine-readable data file].

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.

———1996, Pesticides and Ground Water State Management Plan Regulation; Proposed Rule: U.S. Federal Register, v. 61, no. 124, June 26, 1996, p. 33260–33301, accessed March 10, 2001, from URL http://www.epa.gov/docs/fedrgstr/EPA-PEST/1996/June/Day-26/pr-768DIR/pr-768.pdf

———2000, Drinking water standards and health advisories: U.S. Environmental Protection Agency, Office of Water, EPA 822–B–00–001, accessed December 27, 2001, from URL http://www.epa.gov/waterscience/drinking/standards/index.html

U.S. Geological Survey, 1999, The quality of our Nation's waters—Nutrients and pesticides: U.S. Geological Survey Circular 1225, 82 p., accessed December 26, 2001, from URL http://water.usgs.gov/pubs/circ/circ1225/index.html

———2000, National land cover dataset: U.S. Geological Survey Fact Sheet 108–00, 3 p., accessed November 14, 2000, from URL http://mapping.usgs.gov/mac/isb/pubs/factsheets/fs10800.html

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.

Zaugg, S.D., Sandstrom, M.W., Smith, S.G., and Fehlberg, K.M., 1995, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—Determination of pesticides in water by C–18 solid-phase extraction and capillary-column gas chromatography/mass spectrometry with selected-ion monitoring: U.S. Geological Survey Open-File Report 95–181, 49 p.