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

Prepared in cooperation with
Idaho State Department of Agriculture

Probability of Detecting Atrazine/Desethyl-Atrazine and Elevated Concentrations of Nitrite Plus Nitrate as Nitrogen in Ground Water in the Idaho Part of the Western Snake River Plain

Water-Resources Investigations Report 00–4163

USGS
science for a changing world

Probability of Detecting Atrazine/Desethyl-Atrazine and Elevated Concentrations of Nitrite Plus Nitrate as Nitrogen in Ground Water in the Idaho Part of the Western Snake River Plain

By Mary M. Donato

Water-Resources Investigations Report 00–4163

Prepared in cooperation with
Idaho State Department of Agriculture

Boise, Idaho
2000

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

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

Additional information can be obtained from:

District Chief
U.S. Geological Survey
230 Collins Road
Boise, ID 83702-4520
http://idaho.usgs.gov

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Box 25286
Federal Center
Denver, CO 80225
e-mail: infoservices@usgs.gov

CONTENTS

PLATE

FIGURES

FIGURES—Continued

TABLES

CONVERSION FACTORS AND OTHER ABBREVIATED UNITS

Multiply	By	To obtain
inch (in.)	2.54	centimeter
inch per year (in/yr)	2.54	centimeter per year
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
square mile (mi^2)	2.590	square kilometer

Temperature in degrees Celsius (∞C) can be converted to degrees Fahrenheit (∞F) as follows:

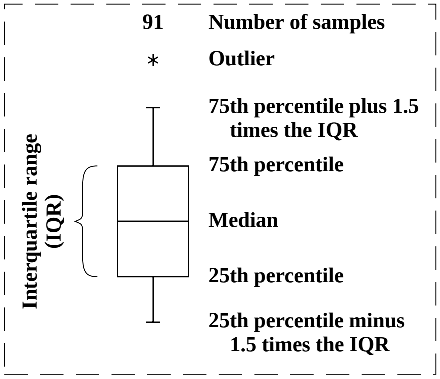
∞F=(1.8) (∞C)+32

Other abbreviated units:

mg/L	microgram per liter
mg/L	milligram per liter

EXPLANATION FOR BOXPLOTS

(Figures 6, 8, and 9)



Probability of Detecting Atrazine/Desethyl-Atrazine and Elevated Concentrations of Nitrite Plus Nitrate as Nitrogen in Ground Water in the Idaho Part of the Western Snake River Plain

By Mary M. Donato

ABSTRACT

As ground water continues to provide an ever-growing proportion of Idaho's drinking water, concerns about the quality of that resource are increasing. Pesticides (most commonly, atrazine/desethyl-atrazine, hereafter referred to as atrazine) and nitrite plus nitrate as nitrogen (hereafter referred to as nitrate) have been detected in many aquifers in the State. To provide a sound hydrogeologic basis for atrazine and nitrate management in southern Idaho—the largest region of land and water use in the State—the U.S. Geological Survey produced maps showing the probability of detecting these contaminants in ground water in the upper Snake River Basin (published in a 1998 report) and the western Snake River Plain (published in this report).

The atrazine probability map for the western Snake River Plain was constructed by overlaying ground-water quality data with hydrogeologic and anthropogenic data in a geographic information system (GIS). A data set was produced in which each well had corresponding information on land use, geology, precipitation, soil characteristics, regional depth to ground water, well depth, water level, and atrazine use. These data were analyzed by logistic regression using a statistical software package. Several preliminary multivariate models were developed and those that best predicted the detection of atrazine were selected. The multivariate models then were entered into a GIS and the probability maps were produced.

Land use, precipitation, soil hydrologic group, and well depth were significantly correlated with atrazine detections in the western Snake River Plain. These variables also were important in the 1998 probability study of the upper Snake River Basin. The effectiveness of the probability models for atrazine might be improved if more detailed data were available for atrazine application.

A preliminary atrazine probability map for the entire Snake River Plain in Idaho, based on a data set representing that region, also was produced. In areas where this map overlaps the 1998 map of the upper Snake River Basin, the two maps show broadly similar probabilities of detecting atrazine.

Logistic regression also was used to develop a preliminary statistical model that predicts the probability of detecting elevated nitrate in the western Snake River Plain. A nitrate probability map was produced from this model. Results showed that elevated nitrate concentrations were correlated with land use, soil organic content, well depth, and water level. Detailed information on nitrate input, specifically fertilizer application, might have improved the effectiveness of this model.

INTRODUCTION

Ground-water quality is an ongoing concern in Idaho because ground water provides a larger proportion of the State's drinking water than ever before. In 1990, ground water supplied approximately 85 percent of the State's drinking water; in 1995, ground water supplied nearly 95 percent of drinking water (Solley and others, 1993, 1998). Pesticides and nitrate (nitrite plus nitrate as nitrogen) have been detected in many aquifers in the State (Crockett, 1995; Rupert, 1994; Rupert and others, 1996). Atrazine was the most commonly detected pesticide in ground water sampled statewide (Crockett, 1995). Atrazine or its breakdown product, desethyl-atrazine (hereafter referred to as atrazine) was detected in water from almost 60 percent of the wells sampled in the Weiser, Idaho, area (Gary Bahr, Idaho State Department of Agriculture, oral commun., 1999). The percentage of wells in this area containing nitrate in concentrations higher than 10 mg/L also was among the largest in Idaho (Crockett, 1995, p. 27).

Maps showing the vulnerability of areas to groundwater contamination are important tools used by resource protection and regulatory agencies to help protect ground-water quality. Maps that delineate areas of high vulnerability to atrazine contamination could be incorporated into the State Pesticide Management Plan to help prevent potential ground-water quality degradation. Other organizations and programs, such as the agri-chemical industry, agricultural producers, the Shoshone-Bannock Tribes, county and city governments, planning and zoning commissions, education programs for applicators, and State programs related to Wellhead Protection, Drinking Water, Home-A-Syst, and Best Management Plans (BMPs), also can benefit from atrazine and nitrate ground-water vulnerability maps.

Background

There are various methods for producing groundwater vulnerability maps. Most maps are produced by using geographic information systems (GIS) to combine data on hydrogeologic and anthropogenic factors such as land use, soil characteristics, and depth to ground water. Some methods emphasize the processes by which contaminants move through the environment to delineate areas that are more or less vulnerable to contamination (Holtschlag and Luukkonen, 1996; Snyder and others, 1998). Others use overlay and index methods that take into account physical characteristics that affect vulnerability. One of the most widely known of these mapping methods is the DRASTIC model, which was developed 15 years ago (Aller and others, 1985). In this method, point ratings are assigned to seven factors: **D**epth to ground 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. The ratings are added together in data layers in GIS to make a map. Maps produced by this method are usually not calibrated (adjusted) to measured contaminant concentrations.

Another type of vulnerability mapping uses statistical methods to correlate various environmental factors with contamination. A statistically valid method to calibrate nitrate ground-water vulnerability maps recently was developed by the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program in the upper Snake River Basin, southeastern Idaho (Rupert, 1997). Unlike the DRASTIC method, this method calculates the probability of contamination on the basis of statistical correlations between measured nitrate concentrations and land use, soils, and depth to ground water.

The USGS recently used an outgrowth of that method to produce an atrazine-specific ground-water vulnerability map for the upper Snake River Basin in a cooperative project with the Idaho State Department of Agriculture (ISDA) (Rupert, 1998). This map was produced by using a statistical method called logistic regression to relate water-quality data to hydrogeologic and anthropogenic factors. This type of map is more appropriately called a probability map, because it delineates areas according to the probability of detecting a contaminant in that area. Calibration of contaminant probability maps with measured water-quality data makes them a superior predictive tool over maps produced by the modified DRASTIC method because the actual probabilities of contaminant detection are quantified.

Experience gained by mapping the upper Snake River Basin now has been applied to producing probability maps for the western Snake River Plain (WSRP). Together, the two maps will provide information on probability of ground-water contamination by atrazine and nitrate for the southern half of Idaho and will serve most of the State's population.

Purpose and Scope

The purpose of this report is to describe how maps were produced that show the probability of detecting atrazine and elevated concentrations of nitrate in ground water in the WSRP in Idaho. The maps are intended to supplement a previous report and maps by the USGS that show the probability of detecting atrazine and elevated concentrations of nitrate in ground water in the Idaho part of the upper Snake River Basin (Rupert, 1998). The areas covered by the two studies are shown in figure 1. Although it is intended that the maps presented in this report and the maps by Rupert (1998) be used side by side, this report addresses only the area encompassed by the WSRP aquifer in Idaho and does not include the entire western Snake River Basin. The study area was restricted because data needed to define depth to ground water in the western Snake River Basin, especially for the tributary valleys, were not available.

Acknowledgments

The author wishes to thank the following people and agencies: Michael G. Rupert, for invaluable advice and encouragement throughout the duration of this project; Gary Bahr and Rick Carlson (ISDA) and Janet Crockett (Idaho Department of Water Resources), for providing data and helpful guidance; and Patrick Lambert, Greg Clark, and Dave Clark, for valuable discussions and substantial improvements to the manuscript.

STUDY AREA DESCRIPTION AND GEOHYDROLOGY

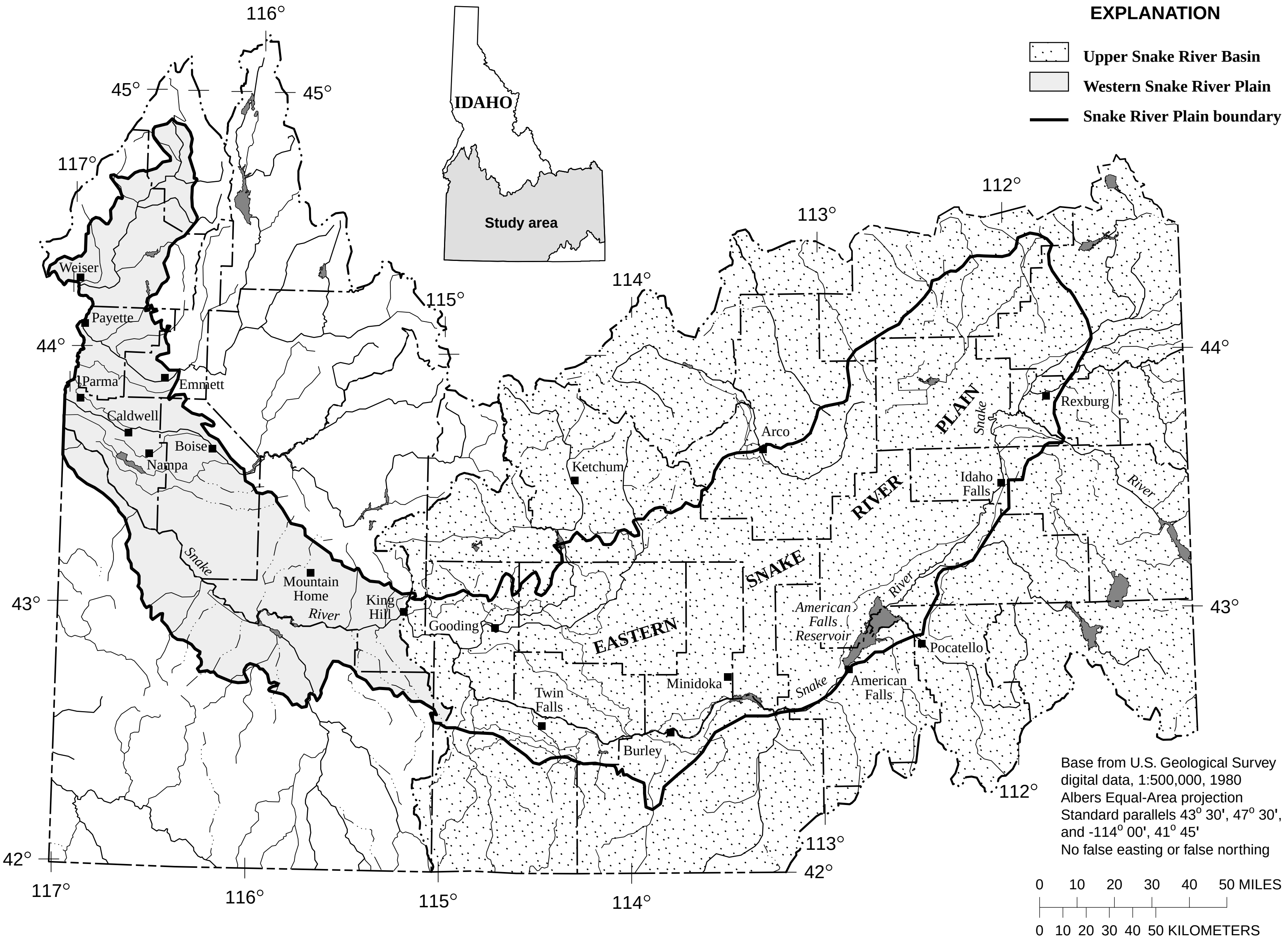
The Snake River Plain is an arcuate topographic and structural depression that extends across southern Idaho. Distinct changes in geology and hydrology near King Hill make feasible a geohydrologic division of the Snake River Plain into eastern and western parts. The 4,800-mi^2 WSRP within Idaho is a graben that extends approximately 130 mi from Weiser on the west to King Hill on the east (fig. 1). The WSRP is as much as 50 mi wide and is bounded on the north and south by high-angle normal faults with at least 9,000 ft of aggregate vertical displacement (Whitehead, 1992). Although the WSRP extends into Oregon, this study refers only to the Idaho part.

The regional aquifer system in the WSRP is composed of three major rock units: upper and middle units of Tertiary and Quaternary sedimentary and volcanic rocks and a lower unit of Tertiary volcanic rocks (Newton, 1991). The upper unit is about 500 ft thick and is composed primarily of sand and gravel. In the Boise River Valley, sand and gravel aquifers also contain many discontinuous clay layers that are locally confining. The middle unit is about 4,000 ft thick and consists of fine-grained sedimentary and volcanic rocks. The volcanic lower unit, estimated to be about 7,000 ft thick, supplies geothermal water as warm as 77°C to numerous springs and wells, largely through faults and fractures. Overlying sedimentary rocks of low hydraulic conductivity in the middle unit confine water in the lower unit and locally separate geothermal water in the lower unit from cold water in the upper unit (Newton, 1991). Water samples used in this study were collected from wells completed in the upper unit.

Regional water movement is to the Snake River; locally, water flows to the Boise River and other tributaries. Depth to first-encountered ground water ranges from 0 to 600 ft (Maupin, 1991) in this region. Water-quality data used in this study were from wells representing the entire range of depths.

MODEL DEVELOPMENT APPROACH

The statistical methods used to produce the maps in this report closely follow those used by Rupert (1998) for the upper Snake River Basin. That study was based on ground-water quality data from approximately 1,600 wells in the upper Snake River Basin included in the Idaho Statewide Ground-Water Monitoring Program (ISGWMP) and on data from 104 additional wells in Cassia, Jerome, Gooding, Lincoln, and Minidoka Counties. This study is based on data collected as part of the ISGWMP from 246 wells in the WSRP. The data represent water samples collected and analyzed for atrazine during January 1, 1995, through December 31, 1998. Because at the outset it was uncertain whether these data would be sufficient to build a statistically rigorous model for producing the probability maps, a larger data set of 541 wells, representing these 246 wells plus an additional 295 wells from the eastern Snake River Plain, was analyzed. This approach had two advantages. First, it assured that if the WSRP data were determined to be inadequate to produce a map, the larger data set could be used to produce a map of the entire Snake River Plain. Although such a map would be more general, it still would provide information for the WSRP that previously had not been available. Furthermore, a new map featuring the entire Snake River Plain would overlap partially with Rupert’s 1998 map of the upper Snake River Basin and



Base from U.S. Geological Survey digital data, 1:500,000, 1980
Albers Equal-Area projection
Standard parallels 43° 30', 47° 30', and -114° 00', 41° 45'
No false easting or false northing

Figure 1. Locations of the western and eastern Snake River Plain and the upper Snake River Basin in Idaho.

would provide a basis for comparison and evaluation of both maps. During the course of the study, it was determined that data for the WSRP alone were indeed adequate to produce an atrazine probability map for that region (pl. 1). However, a preliminary atrazine probability map for the entire Snake River Plain is presented later in this report.

The hydrogeologic and anthropogenic data used to produce the nitrate map for the WSRP were the same as those used to produce the atrazine map, but a larger ground-water quality data set, representing data from 574 wells sampled between 1995 and 1998, was obtained by a separate retrieval from the USGS National Water Information System (NWIS) data base.

The probability maps were produced in a series of steps. First, all ground-water quality data on atrazine and nitrate in ground water of the study area were retrieved from the NWIS data base. Results of analyses of 19 other pesticides were retrieved simultaneously. Information on well depth, water levels at the time of sampling, and use of water also was retrieved and compiled for each well. Next, relations between the atrazine data and the hydrogeologic and anthropogenic data were evaluated using a GIS and statistical methods, including logistic regression analysis. Univariate relations between atrazine and land use, geology, precipitation, soils, depth to ground water, well depth, water level, and atrazine use were investigated, and the significant variables were selected for further analysis. Multivariate logistic regression techniques were used to develop several models representing the combinations of factors that most effectively predicted the presence of atrazine. The final models were entered into a GIS where the maps were produced.

GROUND-WATER QUALITY DATA

Ground-water quality data used in this study were collected as part of the ISGWMP. The ISGWMP is a cooperative program between the USGS and the Idaho Department of Water Resources (IDWR). The primary objectives of the program are to (1) characterize the quality of water in Idaho's aquifers, (2) identify temporal trends in water quality in individual aquifers, and (3) identify aquifers or geographic areas where water-quality problems might exist or be emerging (Idaho Department of Water Resources, 1995). Water-quality data generated by this program are stored in the USGS NWIS data base and the IDWR Environmental Data Management System data base. Samples for the ISGWMP were collected by USGS personnel and analyzed at the USGS National Water Quality Laboratory (NWQL) in Arvada, Colorado. Wells were purged prior to sampling until specific conductance, pH, and temperature stabilized (at least 15 minutes). Details of the ISGWMP sampling design and methods are given in a report by Rupert (1998) and in references contained therein.

All ground-water samples used to develop the atrazine probability model were analyzed using the gas chromatography/mass spectrometry (GC/MS) laboratory method. A separate water-quality data set, analyzed for atrazine by the enzyme-linked immunosorbent assay (ELISA) method and provided by the IDWR, was used solely to assess the completed probability model.

Pesticide Analytical Methods

Solid-phase extraction of analytes and subsequent GC/MS analysis were carried out at the NWQL. Details of the analytical procedures are given in a report by Zaugg and others (1995). The reporting limit for atrazine using the GC/MS method is 0.001 µg/L.

The ELISA method (Vanderlaan and others, 1990) was performed on unfiltered samples using a spectrophotometer at a bench laboratory at the IDWR. This method is designed to measure atrazine concentrations but also has cross reactivity to many different triazine compounds, including desethyl-atrazine, cyanazine, propazine, prometon, simazine, and others. If more than one of these compounds is present, the ELISA method will measure the additive concentrations of all the compounds. A further disadvantage of the ELISA method is that the reporting limit for atrazine is 0.046 µg/L, much higher than the GC/MS reporting limit of 0.001 µg/L.

Duplicates of 210 of the 246 samples in the WSRP analyzed by GC/MS and used in probability modeling also were analyzed by the ELISA method at the IDWR. The duplicate samples were collected concurrently with the samples for GC/MS as part of the ISGWMP. Because of the low minimum detection limit, atrazine was detected by ELISA in only 40 of the 210 samples, whereas atrazine was detected by GC/MS in 128 of the samples. Statistical tests used to compare the results of the two methods indicated a poor linear correlation. These results corroborate those obtained by Rupert

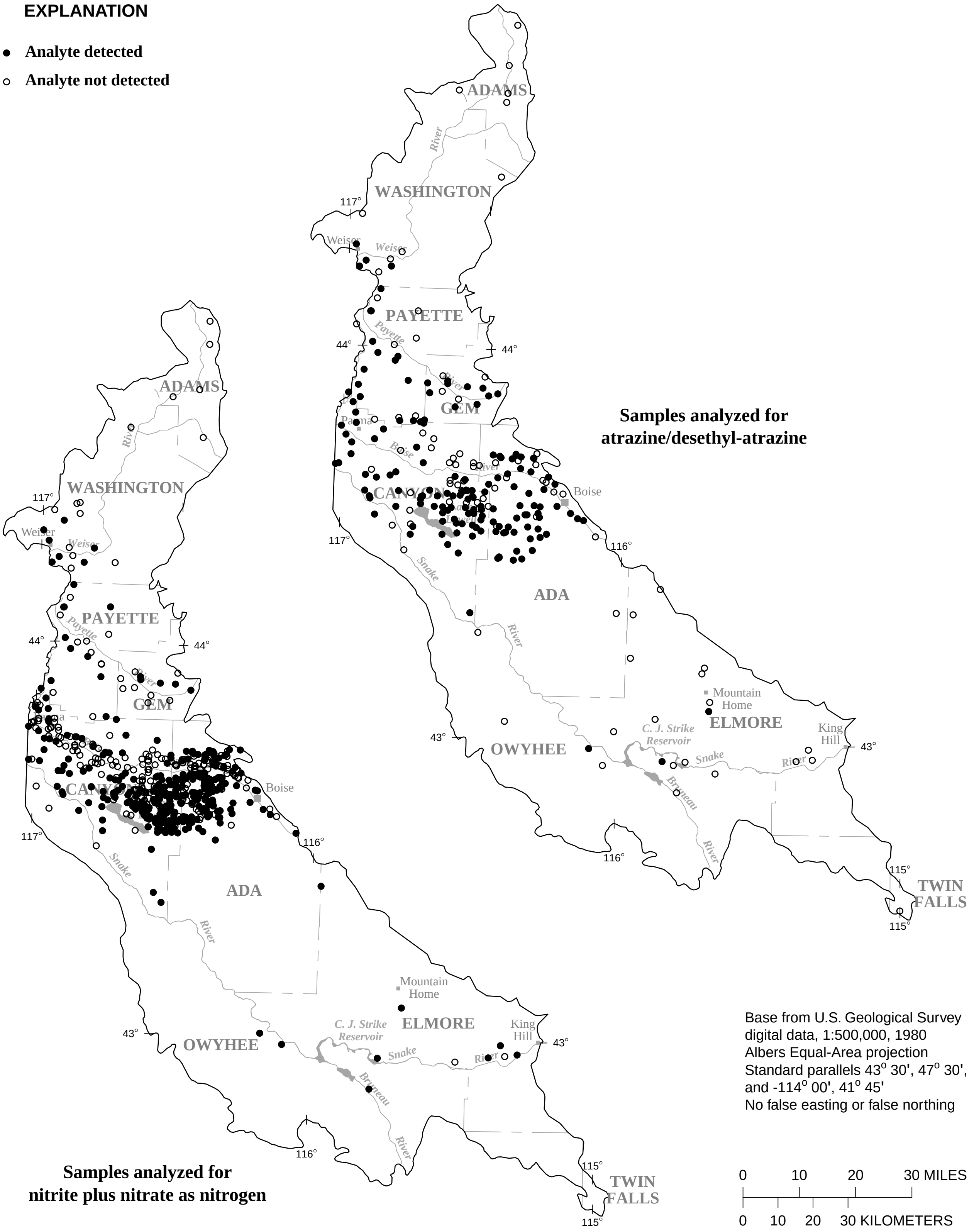


Figure 2. Ground-water samples with and without atrazine/desethyl-atrazine detections and ground-water samples with and without nitrite plus nitrate as nitrogen detections, Idaho part of the western Snake River Plain.

(1998). Therefore, only data obtained by the GC/MS method were used to develop the probability model.

Quality Assurance

The IDWR utilizes a quality assurance/quality control (QA/QC) plan for the ISGWMP to assure that data are collected in a manner that will achieve the objectives of the ISGWMP, that quality control procedures will be sufficient to assure that data are of known and adequate quality, and that data collected will be technically defensible. The plan, which addresses the number and frequency of analyses, sampling procedures, and analytical and calibration methods, is summarized in a report by the IDWR (1995).

Pesticide and Nitrate Detections in Ground Water

Of the 246 wells in the WSRP sampled for atrazine analysis, 80 percent were for domestic use, 12 percent for irrigation, 5 percent for public supply, 2 percent for stock, and 1 percent for other uses. Pesticides were detected in 162 wells sampled for this study (fig. 2). Atrazine was considered to be detected if either atrazine or desethyl-atrazine was detected. Atrazine and desethyl-atrazine, found in 155 wells, were the most commonly detected pesticides. Simazine was detected in 59 wells; prometon, metribuzin, and metolachlor were among the other pesticides detected.

Atrazine was detected in 121 of the 197 domestic wells and in 10 of the 12 public supply wells; however, none of the wells contained concentrations higher than the maximum contaminant level of 0.003 mg/L (3 µg/L) for atrazine established by the U.S. Environmental Protection Agency (EPA).

Of the 574 wells analyzed for nitrate in the WSRP, 85 percent were for domestic use, 11 percent for irrigation, 2 percent for public supply, and the remainder for other uses. Nitrate was detected in 520 wells; concentrations ranged from 0.05 mg/L to more than 45 mg/L; concentrations exceeded 2 mg/L in 351 of the 574 wells and exceeded the EPA minimum contaminant level of 10 mg/L in 42 samples.

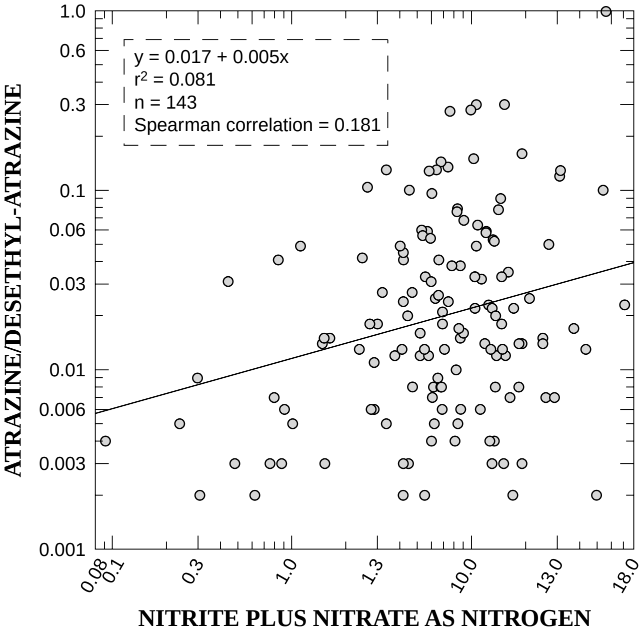


Figure 3. Correlation between atrazine/desethyl-atrazine concentrations determined using gas chromatography/mass spectrometry and nitrite plus nitrate as nitrogen concentrations, Idaho part of the western Snake River Plain.

Correlations Between Atrazine and Nitrate Concentrations

Rupert (1998) looked for relations between atrazine and nitrate concentrations to determine whether nitrate could be used as a surrogate for atrazine data in correlations with hydrogeologic and anthropogenic factors. He found that atrazine and nitrate were poorly correlated and concluded that the use of nitrate as a surrogate for atrazine was not appropriate. In this study, a similar relation between atrazine and nitrate concentrations was sought for samples from the WSRP. Nitrate concentrations were significantly higher in samples in which atrazine was detected, and atrazine concentrations were significantly higher in samples in which nitrate was detected ($p<0.001$). A linear regression of atrazine and nitrate yielded an r-squared value of 0.081 (fig. 3). The Spearman rank-order test produced a correlation coefficient of 0.181. These results suggest poor correlation between the two and verify that using nitrate as a surrogate for atrazine is not feasible.

HYDROGEOLOGIC AND ANTHROPOGENIC DATA

Hydrogeologic and anthropogenic data used in this study were land use, geology, precipitation, soil characteristics, depth to ground water, well depth, water level, and atrazine use. The data were available in GIS format from a variety of sources.

Land-use data were obtained from the IDWR and the Bureau of Reclamation (BOR). GIS coverages obtained from IDWR were derived from three maps: one showing vegetation types, one differentiating between sprinkler and gravity-fed irrigation, and one differentiating between dryland and irrigated agriculture (Rupert and others, 1991). The BOR mapped land cover at 1:40,000 scale from high-altitude aerial photographs taken in 1987 and field checked in 1992. Each land-use data set has its own unique advantages. The IDWR data were mapped at a scale of 1:100,000 and included classifications for lava flows, dryland agriculture, rangeland, and forest land. The BOR data were mapped at a larger scale but combined forest, lava flows, and rangeland into one classification, native lands. Both sets of land-use data were evaluated to determine which produced the best correlation with atrazine detections in ground water.

Geology of the WSRP was obtained from a GIS coverage based on a geologic map by Whitehead (1986). Precipitation data were obtained from a GIS coverage based on an isohyetal map of Idaho (Molnau, 1995), which represents mean annual precipitation in inches during 1961 through 1990.

Soil characteristics were derived from the State Soil Geographic Data Base (STATSGO) developed by the U.S. Natural Resources Conservation Service (U.S. Department of Agriculture, 1991). Six soil characteristics were evaluated for this study: permeability, clay content, drainage, hardpan occurrence, hydrologic group (infiltration rate), and percentage of organic matter. Although STATSGO is considered appropriate for regional interpretations such as this study, the nature of the data base necessitates calculating a series of weighted averages based on the soil map unit identifier (MUID) to obtain a single value for each soil characteristic for each map unit. These averages took into account both the various thicknesses of soil layers and the relative areal proportions of characteristics in each MUID. Because of the necessity to treat the data in this way, soil characteristics are highly generalized. Therefore, the importance of soil characteristics was considered conservatively when the final probability models were developed.

Depth to first-encountered ground water in the WSRP was mapped by Maupin (1991). The map displays depths, in feet below land surface, in five depth intervals: 0 to 100, 101 to 300, 301 to 600, 601 to 900, and >900. The data selection process used to produce the map excluded the ability to distinguish a shallow ground-water system from a deep ground-water system. The GIS coverage used in this study is a revised version of Maupin's 1991 map, in which the depth intervals, in feet, for the WSRP are: 1 to 25, 26 to 50, 51 to 100, 101 to 300, and 301 to 600. No depths greater than 600 ft are indicated.

Accompanying data for well depth were available for all but 10 of the wells used in the statistical analysis. Well depths range from 20 to 471 ft.

Detailed atrazine sales and use information was not available for the study area. Atrazine application data used in this study were obtained from the USGS Pesticide National Synthesis Project (PNSP) and reflect the best county-based estimates of atrazine use in Idaho. The PNSP applied State-based pesticide use coefficients compiled by the National Center for Food and Agricultural Policy to county-level crop averages obtained from the 1992 Census of Agriculture to produce maps showing the distribution of average annual pesticide use. The data obtained from the PNSP indicate the average number of pounds of atrazine applied annually in a county. The county-level estimates are not precise because they are based on average application and treatment rates by State. Furthermore, the data do not reflect pesticide applications to noncropland (for example, home use, greenhouse use). Further discussion of the limitations of the data can be found on the PNSP's Web page (http://water.wr.usgs.gov/pnsp/use92/mapex.html).

County-level atrazine use data were adjusted in this study to reflect the number of pounds used per acre of agricultural land in each county, and that amount was "applied" only to agricultural land; nonagricultural land was assumed to receive no atrazine application. Because the two land-use GIS coverages used in this study classify some lands differently, atrazine use data were adjusted accordingly for each coverage.

DEVELOPMENT OF PROBABILITY MODELS

Maps showing the probability of detecting atrazine in ground water in the WSRP were produced using the methods of Rupert (1998) as outlined previously in this report. A data set containing pesticide concentration and hydrogeologic and anthropogenic data was produced. The data then were analyzed using a statistical software package. Individual (univariate) relations between the occurrence of atrazine and each of the hydrogeologic and anthropogenic factors (independent variables) were evaluated to identify those that were significantly related to atrazine detections. The statistical analysis proceeded with multivariate analysis of independent variables to develop several preliminary models with various combinations of variables. The preliminary models were tested, and those that best predicted the presence of atrazine in ground water were selected. Algorithms representing those multivariate models were entered into the GIS and the probability maps were produced.

Statistical Methods

The independent variables used in this study are of two types: continuous and categorical. A continuous independent variable is one that can assume any one of the infinite number of values on a line interval (Ott, 1993). Examples of variables that were treated as continuous in this study are percent soil clay content, well depth, and atrazine use. In contrast, categorical (sometimes called discrete) variables can assume only a limited number of values. Many of the variables in this study, including land use, geology, precipitation, and depth to water, were treated as categorical variables.

A variety of statistical methods were used to evaluate correlations between atrazine concentrations or detections and the independent variables. One approach was to use the Kruskall-Wallis and Wilcoxon rank-sum tests to evaluate whether atrazine concentrations in different groups (for example, different land-use categories) were significantly different. Another approach was to evaluate whether the value of a continuous variable (for example, well depth) was significantly different for samples in which atrazine was detected compared with those in which it was not. Spearman tests also were used to evaluate correlations between independent variables and atrazine detection.

Logistic regression (Hosmer and Lemeshow, 1989) was used to develop the equations (algorithms) that predict the probability of detecting atrazine in ground water. Logistic regression is a statistical method similar to linear regression, but the dependent variable (in this case, atrazine detection) is transformed to a binary variable (detection or no detection). Therefore, the resulting model predicts the probability that atrazine will be detected, rather than how much atrazine will be present. As with linear regression, logistic regression can be performed on one variable (univariate logistic regression) or on several variables at one time (multiple logistic regression). An advantage of logistic regression as a statistical tool is that normally distributed data are not required.

Multiple logistic regression models were evaluated using parameters calculated by the software, including the likelihood ratio statistic (LR), the rho-squared value (similar to an r-squared value in linear regression), and the standard error, t-ratio, p-value, and odds ratio of the coefficients. The model prediction-success table, which summarizes the classificatory power of the model, also was examined. A brief synopsis of the method and the statistical parameters used to evaluate the predictive success of a model are given in a report by Rupert (1998). For a complete discussion of logistic regression and its applications, refer to the book by Hosmer and Lemeshow (1989). Multiple logistic regression yielded the following model that predicts the probability (p) of detecting atrazine in ground water:

$$p = \frac{e^{[a + b_1(LU) + b_2(P) + b_3(S) + b_4(WL) + b_5(WD)]}}{1 + e^{[a + b_1(LU) + b_2(P) + b_3(S) + b_4(WL) + b_5(WD)]}} \qquad (1)$$

where

p = the probability of detecting atrazine in ground water;
a = intercept;
b_1 = slope coefficient for land use;
LU = land use;
b_2 = slope coefficient for precipitation;
P = precipitation, in inches per year;
b_3 = slope coefficient for soils;
S = soils;
b_4 = slope coefficient for water level;
WL = water level, in feet;
b_5 = slope coefficient for well depth; and
WD = well depth, in feet.

Univariate Analysis

Univariate relations between atrazine detections and land use, geology, precipitation, soils, depth to ground water, well depth, and atrazine use first were evaluated to identify key variables that were likely to be important in the final model. Kruskall-Wallis, Wilcoxon rank-sum, Spearman rank-order, and univariate logistic regression tests were performed on all data by using the commercial statistical software SYSTAT (SPSS, 1999).

During this process, some adjustments in data groupings were made to optimize the effectiveness of the tests. For example, the BOR land-use data originally fell into 10 categories, including 3 types of residential land and 2 types of commercial land. A logistic regression of atrazine with land use was not successful when using the 10 categories (model did not converge). After the number of categories was reduced to 5 (commercial, gravity-irrigated land, sprinkler-irrigated land, residential, and native land) by generalizing some categories, a meaningful result was obtained.

There were statistical differences in atrazine concentrations among samples with atrazine detections from BOR-classified land. Atrazine concentrations in samples from residential land were significantly higher than in samples from irrigated land (p=0.038). Concentrations in samples from native land, however, were not statistically distinguishable from those in samples from residential (p=0.402) and irrigated land (p=0.581). There were no statistical differences in atrazine concentrations among samples from IDWR-classified irrigated, residential, or rangeland (fig. 4).

Correlations among the six soil variables were sought to help eliminate redundancy and achieve a more efficient model. The Spearman rank-order test is a nonparametric test that yields a correlation coefficient between -1 and +1. A coefficient of -1 or +1 indicates that two variables have a perfect linear relation. A coefficient of 0 means that neither of the variables can be predicted from the other by using a linear relation. The test indicated that there were correlations between some pairs of soil variables. The strongest correlations were between drainage and organic content (Spearman correlation coefficient = -0.833), clay content and permeability (-0.70), and permeability and hydrologic group (0.58). This information helped guide final variable selection.

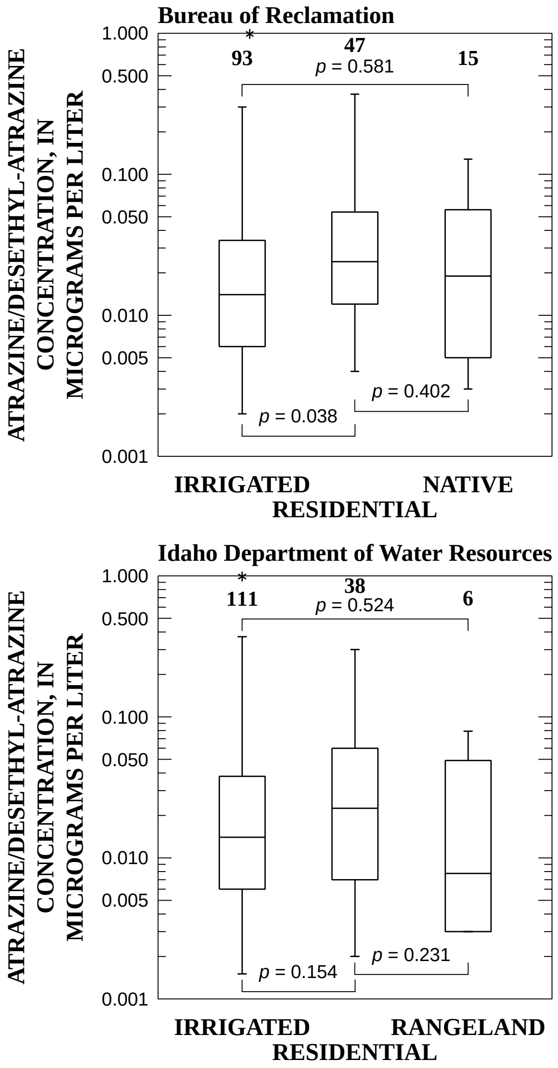


Figure 4. Correlations between atrazine/desethyl-atrazine concentrations and land use classified by the Bureau of Reclamation and the Idaho Department of Water Resources, Idaho part of the western Snake River Plain.

Among the variables that were effective in predicting atrazine detection using univariate analysis were land use, precipitation, soil hydrologic group (one of the six soil characteristics), well depth, and water level. BOR land-use data and IDWR land-use data performed equally well in univariate statistical tests (tables 1 and 2).

Although some variables, including geology and soil clay content, were identified at this stage as probable weak contributors to a model, none were excluded on the basis of univariate analysis. The following variables were carried forward for multivariate logistic regression analysis: BOR-classified land use, IDWR-classified land use, geology, precipitation, the six soil

Table 1. Results from Spearman rank-order tests between independent variables and atrazine/desethyl-atrazine detection, Idaho part of the western Snake River Plain

[IDWR, Idaho Department of Water Resources; BOR, Bureau of Reclamation; Q, Quaternary; T, Tertiary; in., inches; DEA, desethyl-atrazine]

Independent variable	Spearman correlation coefficient
Flood irrigation (IDWR)	0.14
Rangeland (IDWR)	–.23
Urban land (IDWR)	.14
Sprinkler-irrigated land (IDWR)	–.20
Commercial land (BOR)	–.01
Gravity-irrigated land (BOR)	.04
Residential land (BOR)	.17
Sprinkler-irrigated land (BOR)	.08
Native land (BOR)	–.30
Geology (Q-T basalt)	.12
Geology (Q-T sediments)	.01
Geology (other)	–.17
Geology (Q basalt)	–.03
Geology (Q sediments)	–.04
Precipitation (less than 10 in.)	.01
Precipitation (between 10 and 15 in.)	.17
Precipitation (greater than 15 in.)	–.27
Soil permeability	–.10
Soil clay content	–.02
Soil drainage	–.02
Soil hardpan content	.12
Soil organic content	–.03
Soil hydrologic group	–.09
Well depth	–.20
Water level	.05
Atrazine/DEA application (IDWR)	.16
Atrazine/DEA application (BOR)	.14

characteristics, depth to ground water, well depth, water level, and atrazine use.

Multivariate Analysis

Developing a multivariate logistic regression model is a stepwise procedure whereby variables are added to or deleted from the model one at a time until the model that best fits the data is achieved. The model is tested at each step to determine whether it is significantly (statistically) improved by the addition or subtraction of a variable. Stepwise modeling, using a forward-selection, backward-elimination method, begins with one variable and adds variables one at a time until the best model is achieved. Each variable included in the present model is tested for removal each time a new variable is added, because a variable found to be important at an early step can become insignificant at a later step. Because the number of variables is high, and the number of possible combinations is too large to perform manually in a practical timeframe, commercial statistical software commonly is used to perform automatic stepwise modeling. A disadvantage of automatic modeling is that the user has limited control over the order of inclusion or exclusion of variables. This type of modeling can, however, provide an efficient overview of the variables that are likely to be important in the final model.

For this study, forward and backward stepwise modeling routines in commercial statistical software were used to screen variables but not to develop the final models. The multivariate models were developed in part using commercial software to perform a modified version of forward stepwise multiple logistic regression. Starting with a single land-use variable, each of the remaining variables was added, one at a time, to create several two-variable models. Each of the new two-variable models was statistically compared to the one-variable model to see which, if any, significantly improved the one-variable model's predictive capabilities. The likelihood ratio test (Hosmer and Lemeshow, 1989, p. 14-16), which commonly is used to compare nested models (models in which the variables of one model are a subset of the other), was used to make this comparison. The model was considered to be significantly improved at the 95-percent confidence

Table 2. Results from univariate logistic regression of atrazine/desethyl-atrazine detections with independent variables, Idaho part of the western Snake River Plain

[LLR, log-likelihood ratio of logistic regression model; df, degrees of freedom of the log-likelihood ratio; LLR-p, chi-square p-value calculated from log-likelihood ratio; Rho-squared, McFadden's rho-squared calculated with logistic regression; —, logistic regression model did not converge; IDWR, Idaho Department of Water Resources; BOR, Bureau of Reclamation; <, less than; DEA, desethyl-atrazine]

Independent variable	LLR	df	LLR-p	Rho-squared
Land use (IDWR)	27.4	3	<0.001	0.085
Land use (BOR)	28.4	4	<.001	.088
Geology	—	—	—	—
Precipitation	15.4	2	<.001	.048
Soil permeability	4.32	2	.115	.013
Soil clay content	5.05	3	.168	.016
Soil drainage	5.0	2	.082	.015
Soil hardpan content	6.03	4	.197	.019
Soil hydrologic group	7.33	1	.007	.023
Soil organic content	3.91	1	.048	.012
Depth to water	1.18	1	.278	.004
Well depth	21.2	1	<.001	.067
Water level	3.0	1	.083	.01
Atrazine/DEA use (IDWR)	5.32	1	.021	.016
Atrazine/DEA use (BOR)	4.41	1	.036	.014

Table 3. Independent variables significantly correlated in multivariate regressions with the detection of atrazine/desethyl-atrazine and elevated nitrite plus nitrate as nitrogen concentrations, Idaho part of the western Snake River Plain

[IDWR, land-use data developed by the Idaho Department of Water Resources; BOR, land-use data developed by the Bureau of Reclamation; x, significant relation with atrazine detections in ground water; —, no relation observed; DEA, desethyl-atrazine]

	Western Snake River Plain (this study)					Rupert (1998)
Independent variable	Nitrite plus nitrate as nitrogen	Model 4	Model 5	Model 6	Model 8	Model 1
Land use (IDWR)	—	—	—	X	X	—
Land use (BOR)	X	X	X	—	—	X
Precipitation	—	X	X	X	X	X
Soil hydrologic group	—	X	X	X	X	X
Soil organic content	X	—	—	—	—	—
Well depth	X	X	X	X	X	X
Water level	X	X	X	X	X	—
Atrazine/DEA application	—	X	—	—	X	—

level if the chi-square p-value calculated from the likelihood ratio test was <0.05. The most effective two-variable model was selected, and the procedure was repeated with the remaining variables for a three-variable model. This process continued until the model was not significantly improved by the addition of any variables. This procedure differs from the automated procedure in that once a variable was incorporated in the model, it remained there; no tests were performed for removal of variables (no backward elimination).

Model development involved not only interpretation of various statistical indicators, but also purposeful selection of variables based on scientific judgment. For example, one statistically acceptable model that included two soil variables, drainage and hydrologic group, was eliminated because these two variables are weakly correlated (Spearman correlation coefficient = -0.50) and considered somewhat redundant. Furthermore, considering the limitations of the soil data, a model containing two soil variables is believed to overemphasize the importance of soils. This variable selection process yielded preliminary models that were statistically justifiable and scientifically reasonable.

Because IDWR and BOR land-use data both performed well in the univariate analysis phase, models were developed for each land-use type. A version of each model with and without atrazine use also was developed. Early versions of the models included precipitation, well depth, water level at the time of sampling, and one soil variable, either drainage or hydrologic group. Addition of other variables—geology, regional depth to ground water, soil clay content, permeability, organic content, and percent hardpan—either did not significantly improve or actually weakened the model, and these variables were eliminated from further consideration.

Evaluation of Preliminary Logistic Regression Models

Multivariate analysis culminated in selection of four preliminary models, two that use BOR land-use data and two that use IDWR land-use data (table 3). Of the many preliminary models, versions that used soil hydrologic group were as effective as versions that used soil drainage. Models that used hydrologic group were emphasized because they tended to produce smoother distributions of probability values. Models 6 and 8 are based on IDWR-classified land use; models 4 and 5 are based on BOR-classified land use. Models 4 and 8 include atrazine use, whereas models 5 and 6 do not. Statistical results from the four preliminary models are summarized in table 4.

All four models were statistically robust and similarly well fit. Rho-squared values for the multiple logistic regressions ranged from 0.194 to 0.226 (table 4). Total correctly predicted values ranged from 0.65 to 0.67. Log-likelihood p-values for all models were <0.001, indicating that the models were significantly more effective than constant-only models at the 99.9-percent confidence level.

The effectiveness of each of the four models was evaluated further by applying the model to the data upon which it had been built. The probability of atrazine detection was calculated for each sample in the

Table 4. Results from four preliminary models used to predict atrazine/desethyl-atrazine detections, Idaho part of the western Snake River Plain

[Rho-squared, McFadden's rho-squared calculated with logistic regression; LL, log-likelihood of logistic regression model; LLR, log-likelihood ratio of logistic regression model; df, degrees of freedom of the log-likelihood ratio; LLR-p, chi-square p-value calculated from the log-likelihood ratio; total correct, percentage of correct predictions from logistic regression model prediction-success table; Spearman, Spearman correlation coefficient of predicted versus actual detections; <, less than]

	Logistic regression results						Linear regression results			
Model	**Rho-squared**	**LL**	**LLR**	**df**	**LLR-p**	**Total correct**	**r-squared**	**y-intercept**	**Slope**	**Spearman**
4	0.197	–123.8	60.8	12	<0.001	0.65	0.92	–5.8	1.06	0.954
5	.194	–124.3	59.9	11	<.001	.65	.92	4.1	.93	.964
6	.199	–123.5	61.5	10	<.001	.65	.88	–1.7	.98	.952
8	.226	–119.4	69.7	11	<.001	.67	.86	–6.3	1.05	.891

data set. The data set then was sorted by ascending probability rating and was divided into intervals of 10 percent (0 to 10, 10 to 20, 20 to 30, and so on). The percentage of atrazine detections in each interval then was calculated. Linear regression was used to compare the percentage of actual atrazine detections with the predicted probability of a detection (fig. 5). A perfectly fit model is one whose linear regression produces a slope = 1 and y-intercept = 0; models approaching these values were considered well fit. A negative y-intercept for the linear regression indicates that the model tends to predict higher probabilities of detection than are actually detected (high bias), whereas a positive y-intercept indicates that the model tends to underestimate the probability of detection (low bias). Model-fit linear regression slopes for the preliminary models ranged from 0.93 to 1.06; y-intercepts ranged from -6.3 to +4.1. Spearman correlations of predicted versus actual percent detections were strong: 0.89 to 0.96. On the basis of the Wilcoxon test, the differences in probability ratings between water from wells with and without atrazine detections were significant at the 99.9-percent confidence level in all cases (fig. 6).

Model 5 (BOR-classified land use, no atrazine use) is the only model with a positive y-intercept and is thus the only model that has a low bias. The likelihood ratio test indicates that addition of the atrazine use variable to model 5 does not significantly improve its predictive capabilities (model 4; chi-square p-value = 0.331). On the other hand, model 8 (IDWR-classified land use, with atrazine use) is a significant improvement over the equivalent model without atrazine use (model 6; chi-square p-value = 0.004).

Models that include atrazine use (models 4 and 8) demonstrate more strongly negative y-intercepts compared with those of corresponding models without atrazine use, suggesting that models 4 and 8 tend to predict higher probabilities of detection. The aforementioned limitations of the atrazine use data, however, impart some ambiguity to this observation.

Of the two models that include atrazine use, model 8 has the highest rho-squared value and the highest log-likelihood ratio and is considered the statistically most effective model. The two models that do not include atrazine use, models 5 and 6, are nearly as effective in defining the probability of detecting atrazine in ground water as models that include atrazine use.

These results are different from Rupert's (1998) results for the upper Snake River Basin in several ways. In Rupert's study, BOR-classified land-use data produced a more effective model than IDWR-classified land-use data, whether or not atrazine use was included. Adding atrazine use to the models improved them regardless of which land-use data were used. Soil drainage was an effective variable in the vulnerability models developed in this study, whereas it was not useful in the eastern Snake River Plain models. Overall, however, the models for the WSRP are similar to those for the eastern Snake River Plain, especially considering the differences in the basic geohydrologic framework of the two regions. In both studies, the most effective variables were land use, precipitation, soil hydrologic group, and well depth. Geology, depth to ground water, soil organic content, permeability, and clay content were not important contributors to the models.

Production of Atrazine Probability Maps

Results of the preliminary regression models were used to produce the probability maps. Probability ratings then were calculated, using eq. 1, for each of the 20,000 polygons to define the maps. The maps were produced for a well depth of 99 ft and a water level of

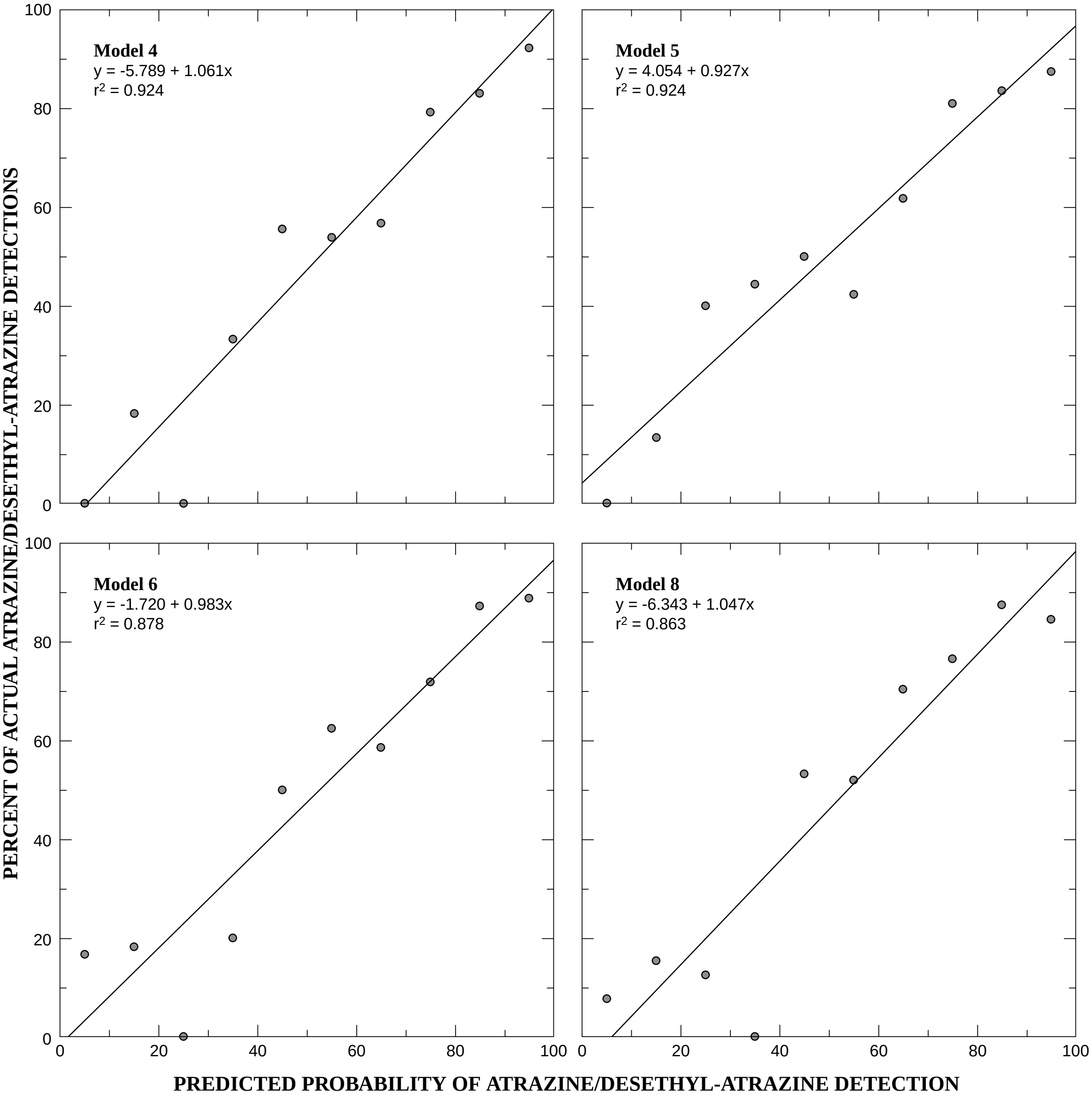


Figure 5. Correlations between predicted probabilities of detecting atrazine/desethyl-atrazine and actual detections, Idaho part of the western Snake River Plain.

30 ft, the median values for wells evaluated in this study. Rupert (1998) and Tesoriero and Voss (1997) used the same procedure. Maps that represent models 4 and 5, both of which use the BOR land-use data, were emphasized to facilitate use of these maps alongside those made by Rupert (1998) for the upper Snake River Basin. The probability map derived using coefficients from model 5 (BOR-classified land use, no atrazine use) is presented in plate 1. For comparison, the same map with a coefficient for atrazine use included in the model (model 4) is shown in figure 7.

EVALUATION AND TESTING OF PROBABILITY MODELS

The effectiveness of regression models 4 and 5 was tested by comparing model results with an independent set of ground-water monitoring data. The independent data set consisted of analyses from 40 wells included in ISDA's ground-water quality monitoring program and sampled for selected pesticides, including atrazine and desethyl-atrazine, during 1996-99. The wells were sampled during ISDA projects 710, 770, and 860, and are clustered in Washington, Payette, Gem, and Owyhee Counties. The data set is biased, because ISDA selected many of the sampled wells on the basis of previous reports of pesticide detections in the region. None of the wells in this data set are included in the data set used to develop the model. The analytical methods varied for the three projects, and detection limits ranged from 0.013 to 0.033 µg/L, well above the 0.001 µg/L limit for the data used to develop the model. Because the regression model is concerned only

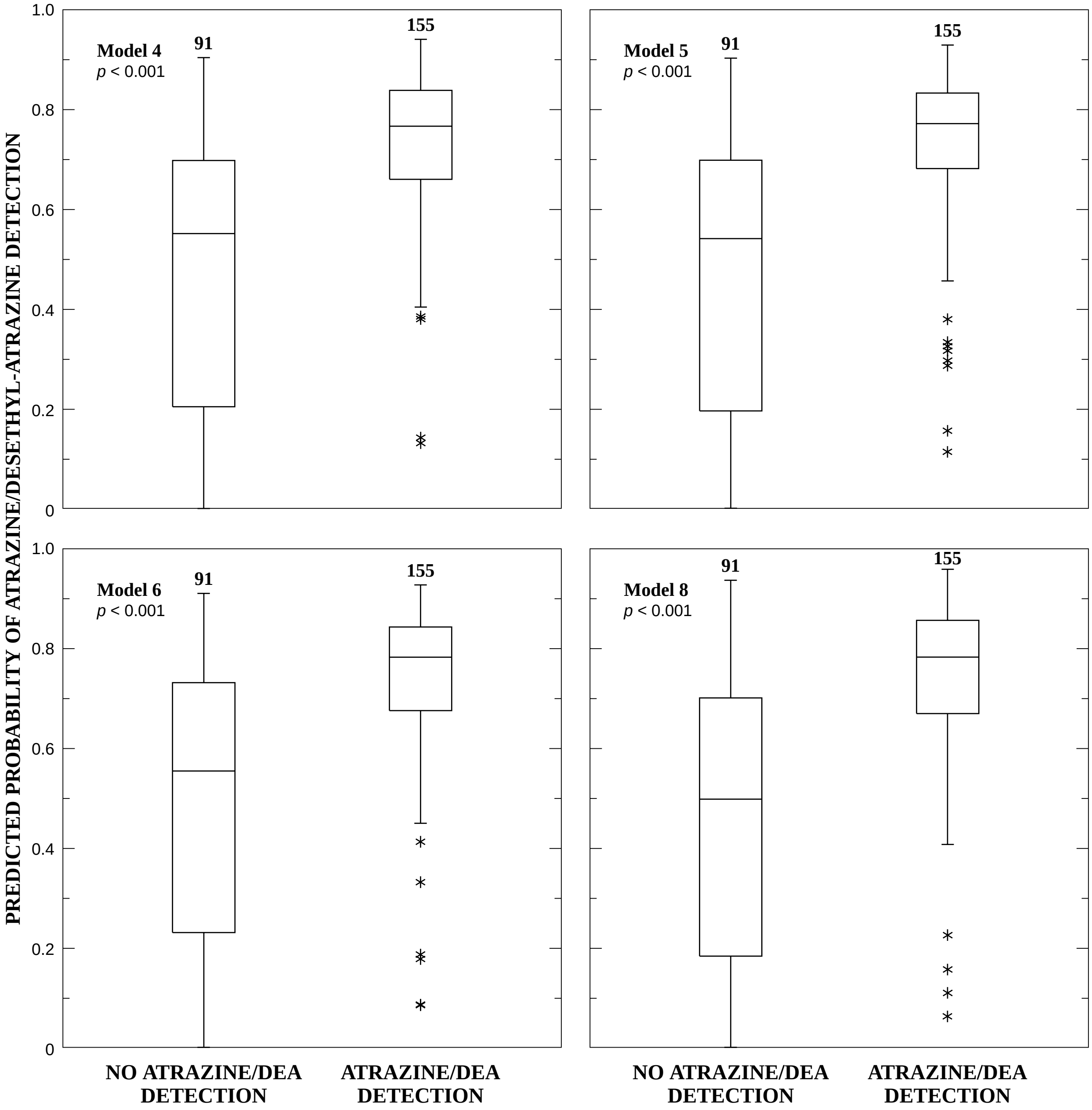


Figure 6. Predicted probabilities of detecting atrazine/desethyl-atrazine for samples with and without actual detections, Idaho part of the western Snake River Plain. (DEA, desethyl-atrazine)

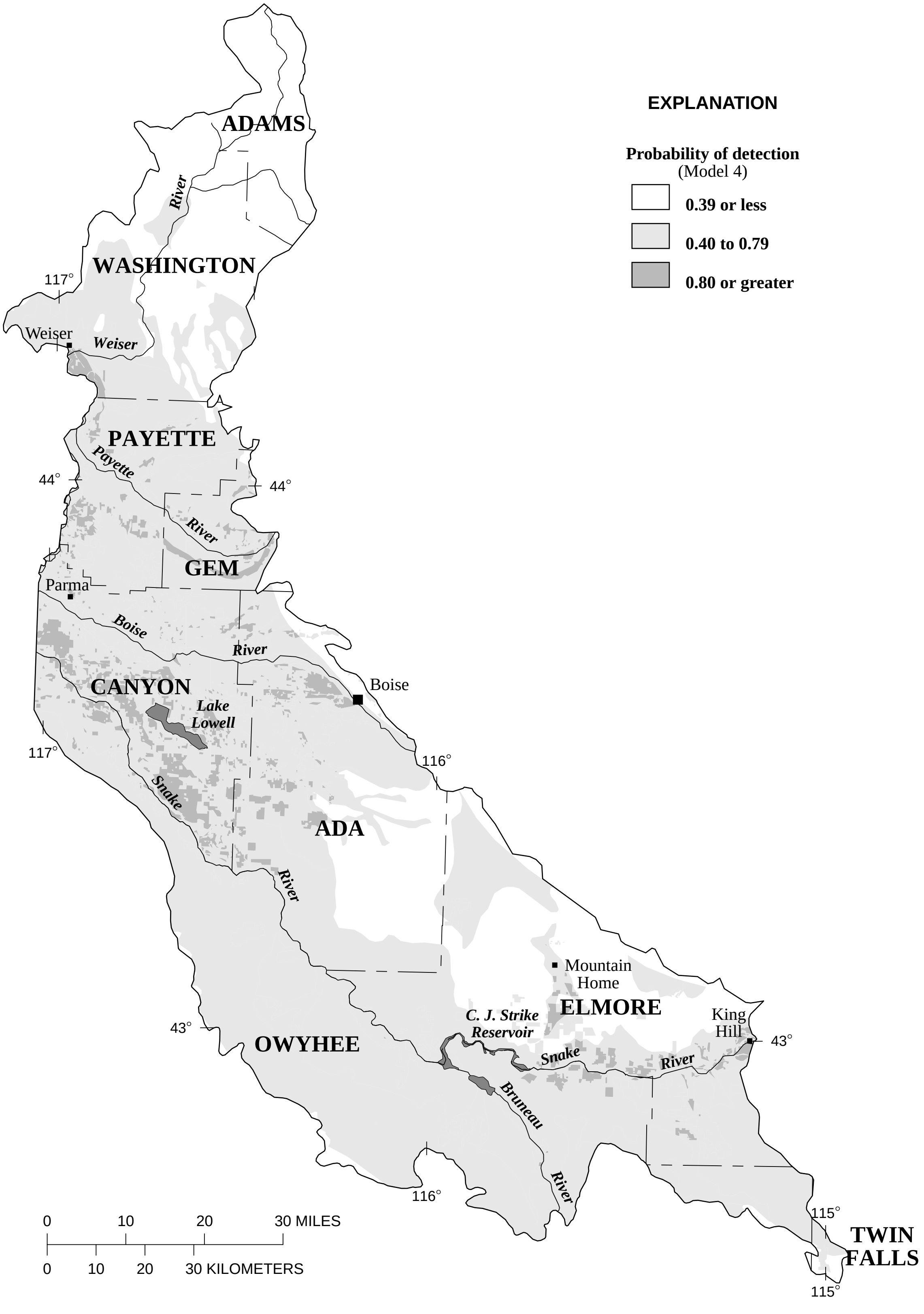


Figure 7. Probability of detecting atrazine/desethyl-atrazine concentrations greater than 0.001 microgram per liter in water from wells 99 feet deep using model 4, atrazine use included, Idaho part of the western Snake River Plain.

with detection of atrazine, not with the exact concentration, data from all projects were combined. Atrazine was detected in 30 of the 40 wells examined. Water-level and well-depth information was available for only 31 of the 40 wells. Static water levels reported in drillers' logs at the time the wells were drilled were used for the calculations; the recency of these water levels is not known.

The ISDA well locations and associated analytical data were converted to GIS format and probabilities of detecting atrazine were calculated for each well. The data were sorted by increasing probability of detection, and the percentage of actual atrazine detections in each interval was calculated. The predicted probabilities ranged from 0.02 to 0.96 for model 4 and from 0.02 to 0.93 for model 5. As before, the predicted probability of detection in each interval was compared with the actual percentage of detections by linear regression. The correlation between predicted and detected values for both models was poor (fig. 8). Model 4 predicted

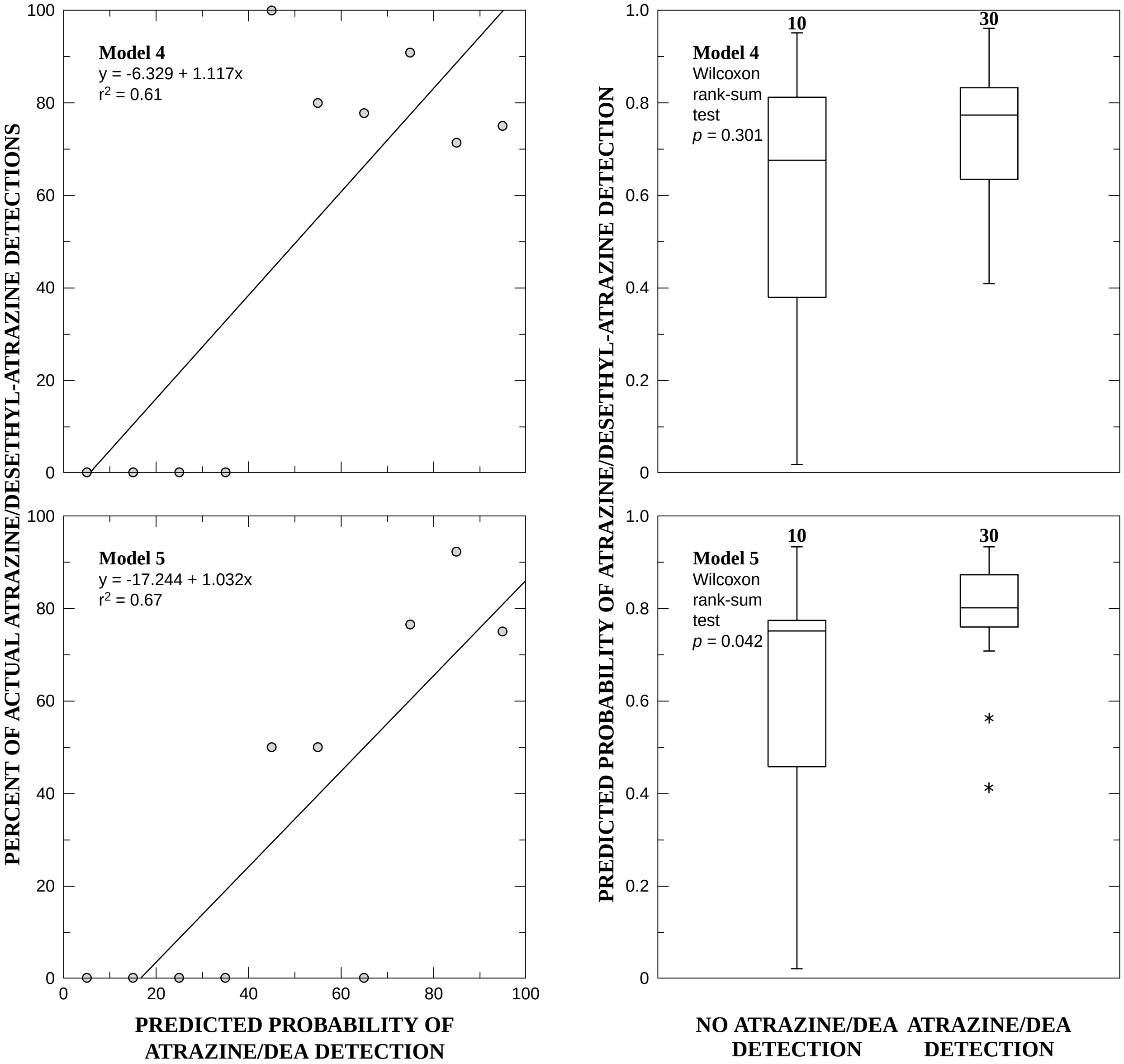


Figure 8. Correlations between predicted probabilities of detecting atrazine/desethyl-atrazine and actual detections in ground-water samples collected by the Idaho Department of Agriculture, Idaho part of the western Snake River Plain. (DEA, desethyl-atrazine)

somewhat higher probabilities of detection for the population of wells with atrazine detections than for those without atrazine detections, but the difference was not statistically significant ($p = 0.301$). This poor correlation is probably a result of the high minimum detection limit associated with the verification data set. Model 5, however, predicted significantly higher probabilities for wells with atrazine detections than for those without atrazine detections ($p = 0.042$).

Even though the test data set contained only 40 wells, and data for only 31 wells were complete, the results suggest that both models do a fair job of predicting the probability of atrazine detection for an independent data set. A more useful test of the models would involve a larger independent data set that represents the entire map area and that was analyzed with a method having a detection limit of 0.001 µg/L. Until such a verification of the models is performed, the probability map, which is based on regional-scale data, should be used cautiously, particularly when applied to localized areas.

Another test of models 4 and 5 was performed using atrazine data provided by IDWR. The IDWR analyses were performed with the ELISA method described previously and have a higher detection limit and a lower percentage of atrazine detections. Nevertheless, they provide some degree of insight as to the effectiveness of the models.

The IDWR test data consisted of 386 samples from wells in the WSRP. This number includes 185 samples from wells that were not part of the USGS data base. As previously described, the IDWR well locations and associated analytical data were converted to GIS format, and probabilities of detecting atrazine were calculated for each well by using models 4 and 5. The data for all 386 wells were sorted and the percentage of detections in each interval was counted. Linear regression was used to compare the percentage of actual atrazine detections with the predicted probability of a detection (fig. 9). Although the higher detection limit of the ELISA analytical method decreased the percentage of detections, the linear correlation between percentages predicted and detected for both models was reasonably good. The Wilcoxon rank-sum test indicated that the probabilities for samples with atrazine detections were significantly higher at the 95-percent confidence level than probabilities for those with no detections ($p < 0.001$).

The same tests were applied only to the 185 wells not included in the USGS data set and, thus, represent a more independent test of the models. This data set did not produce a meaningful linear regression because of the low number of actual detections, but the Wilcoxon test again indicated that probabilities for samples with atrazine detections were significantly higher at the 95-percent confidence level than probabilities for those without atrazine detections ($p = 0.029$ for model 4 and $p = 0.032$ for model 5).

Because the geohydrologic framework of the WSRP is different from that of the upper Snake River Basin, and because the models for these two areas were based on different data sets, there is no reason to expect that the respective models would produce similar results. Nevertheless, several of the variables that were important in Rupert's (1998) model also were important in this study: land use, precipitation, soil hydrologic group, and well depth. Soil drainage also was an effective variable for models in this study, whereas soil drainage was not effective for the eastern Snake River Plain atrazine model. It is not known whether this difference reflects the contrast in hydrologic conditions in the WSRP, or whether it indicates differences in the way the soil data were manipulated in the two studies.

At the eastern margin of the map presented in plate 1 and adjacent to Rupert's (1998) map, probabilities are generally higher than 0.40, whereas on the western boundary of Rupert's map, probabilities generally are between 0.20 and 0.40. As mentioned earlier, the maps are intended to give a regional view of probabilities of detection, and they should not be interpreted in too literal a sense.

Three general areas on plate 1 display high probabilities of detection (>60 percent): (1) an east-west-trending belt adjacent to the Snake River in the southern WSRP, (2) an area generally coinciding with Canyon and northern Ada Counties, and (3) a west-northwest-trending zone that roughly coincides with the Payette River Valley. In all these areas, high probabilities most likely are related to the predominance of agricultural land use, which is weighted heavily by the probability model. In addition, some areas in Canyon and Ada Counties are classified as residential land, which also influences the probability calculation. In area 3, the curved southern boundary of the 40- to 60-percent probability zone coincides with the rainfall contour representing 10 in. of annual precipitation. Areas of low probability generally coincide with areas of greater rainfall (greater than 15 in/yr has a strong negative effect on the probability) or areas of less agricultural land use. The effect of soil characteristics on

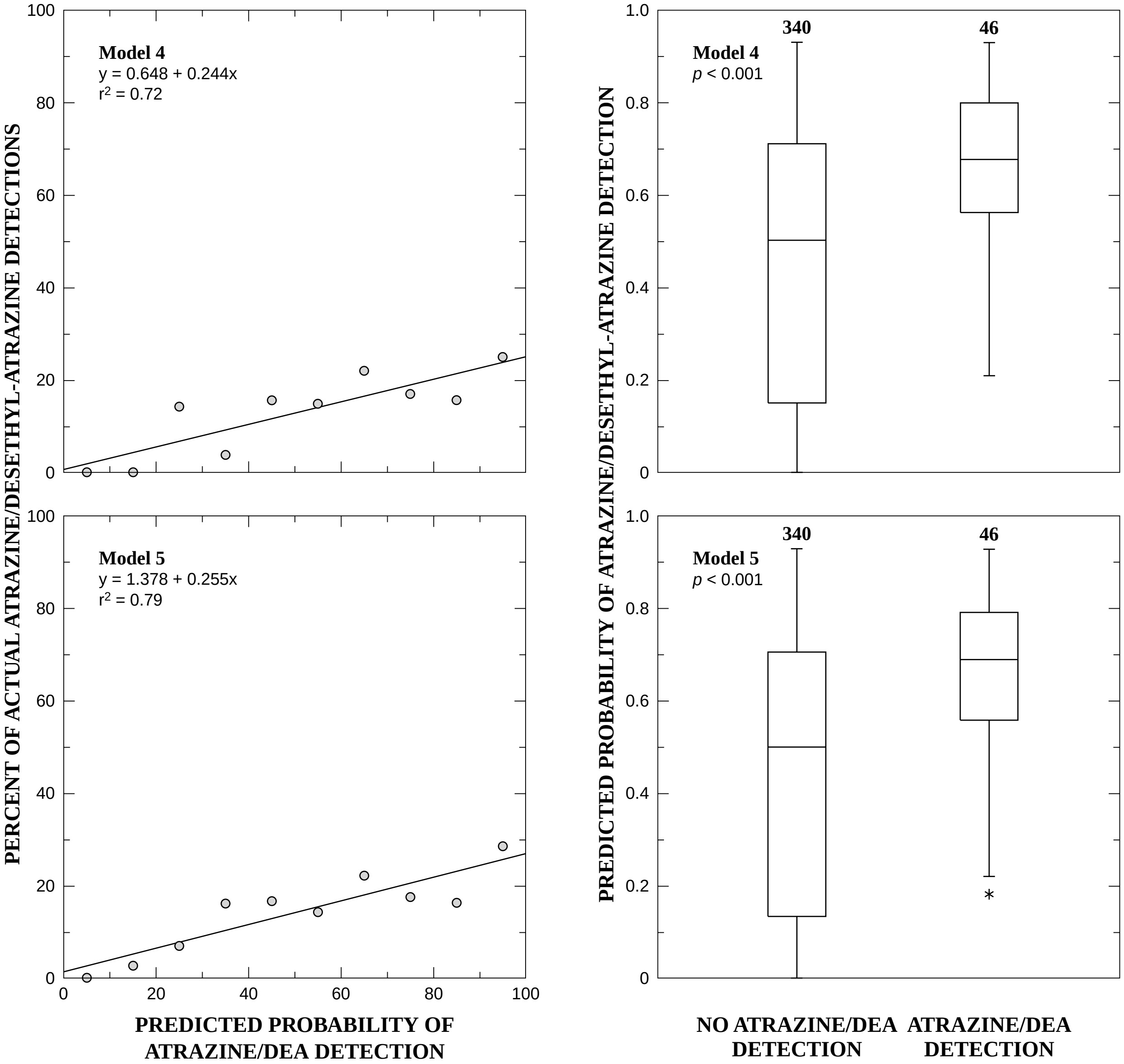


Figure 9. Correlations between predicted probabilities of detecting atrazine/desethyl-atrazine and actual detections determined by enzyme-linked immunosorbent assay, Idaho part of the western Snake River Plain. (DEA, desethyl-atrazine)

probability can be seen in several places on the map where probability boundaries mimic the soil map (for example, the irregular shapes in central Ada County and in Washington County).

The probability maps produced by using models 4 (atrazine use included) and 5 (atrazine use not included) are very similar in overall appearance. Most differences are evident within the areas where probabilities are greater than or equal to 0.40 (brown, pink, and red on plate 1). Comparison of total acreage in each of the five probability intervals shows that model 5 includes more land in the intervals 0.40 to 0.59 and greater than 0.79, and less land in the interval 0.60 to 0.79. Differences between the two maps are most evident in areas of agricultural land use. Atrazine use was assigned only to irrigated land, using the assumption that the primary agricultural use for atrazine is weed control. The coefficients for sprinkler- and gravity-irrigated land reflect this difference: These land types are weighted more heavily in the probability calculation in model 5 than in model 4. One area where this difference is evident is southeast of Weiser. The map for model 5 shows this area to be in the 0.40- to 0.59-probability range, whereas in the map for model 4 (fig. 7), this area lies in the 0.40- to 0.79-probability range. Slight redistribution of the two highest probability intervals in areas southwest of Mountain Home is another notable difference between the two maps.

PRELIMINARY ATRAZINE PROBABILITY MAP FOR THE ENTIRE SNAKE RIVER PLAIN

A larger data set representing the entire Snake River Plain was analyzed concurrently with the data for the WSRP in case a rigorous atrazine probability model could not be developed for the WSRP alone. Although the WSRP data were sufficient to develop a model, a preliminary probability model and map were produced for the entire Snake River Plain so that a comparison could be made with Rupert's (1998) map. The most effective model for the entire Snake River Plain consists of the following variables: BOR-classified land use, precipitation, soil hydrologic group, and depth to water. The preliminary atrazine probability map for the entire Snake River Plain is shown in figure 10. The map is produced for a median well depth of 150 ft.

In the areas where the map in figure 10 overlaps Rupert's (1998) map of the upper Snake River Basin, the probabilities of detection are broadly similar, but Rupert's map indicates slightly lower probabilities of detection in several areas. Rupert's map is probably a more appropriate representation of atrazine detection probabilities in the eastern Snake River Plain, because the data upon which his map is based are more specific to that region.

PROBABILITY OF DETECTING ELEVATED NITRATE IN GROUND WATER OF THE WESTERN SNAKE RIVER PLAIN

Logistic regression also was used to develop a preliminary model that predicts the probability of nitrate concentrations greater than 2 mg/L in ground water of the WSRP. The threshold value of 2 mg/L was chosen because the background concentrations of nitrate in ground water are typically about 2 mg/L (Rupert, 1996).

Nitrogen input and loss data were estimated using Rupert's (1996) methods to compute an overall nitrogen balance for the WSRP. Inputs comprised fertilizer and manure application, septic tank effluent, legume crops, and atmospheric nitrogen deposition. Losses comprised crop uptake, denitrification, and crop decomposition. Data were estimated at the county level because crop, livestock, population, fertilizer, and other data were available only at the county level, but were adjusted to account for land use (for example, nitrogen from fertilizer, manure application, and legume crops was considered to be applied only to agricultural land). No clear statistical correlations were identified between concentration of nitrate in ground water and net nitrogen input, total nitrogen input, or any of the simple components of the nitrogen balance. Adding nitrate input to the nitrogen probability model as an independent variable did not improve its effectiveness. The poor statistical correlations could be due to the inexact nature of the county-level data regarding fertilizer application and crop input and uptake.

The statistically most effective model for predicting excessive nitrate concentrations in ground water included BOR-classified land use, soil organic content, well depth, and water level (model B, table 5). It is recognized that the organic content of soil enhances the

Table 5. Results from two preliminary models used to predict nitrite plus nitrate as nitrogen concentrations greater than background levels (2 milligrams per liter), Idaho part of the western Snake River Plain

[Rho-squared, McFadden's rho-squared calculated with logistic regression; LL, log-likelihood of logistic regression model; LLR, log-likelihood ratio of logistic regression model; df, degrees of freedom of the log-likelihood ratio; LLR-p, chi-square p-value calculated from the log-likelihood ratio; total correct, percentage of correct predictions from logistic regression model prediction-success table; Spearman, Spearman correlation coefficient of predicted versus actual detections; A, model including soil hardpan only; B, model including soil organic content only (preferred model); C, model including both soil hardpan and soil organic content; <, less than; —, test not performed]

	Logistic regression results						Linear regression results			
Model	**Rho-squared**	**LL**	**LLR**	**df**	**LLR-p**	**Total correct**	**r-squared**	**y-intercept**	**Slope**	**Spearman**
A	0.199	–307.09	152.76	11	<0.001	0.65	0.95	3.79	0.93	0.964
B	.195	–308.62	149.71	9	<.001	.64	.95	–.35	1.00	.998
C	.207	–303.98	159.00	13	<.001	.65	—	—	—	—

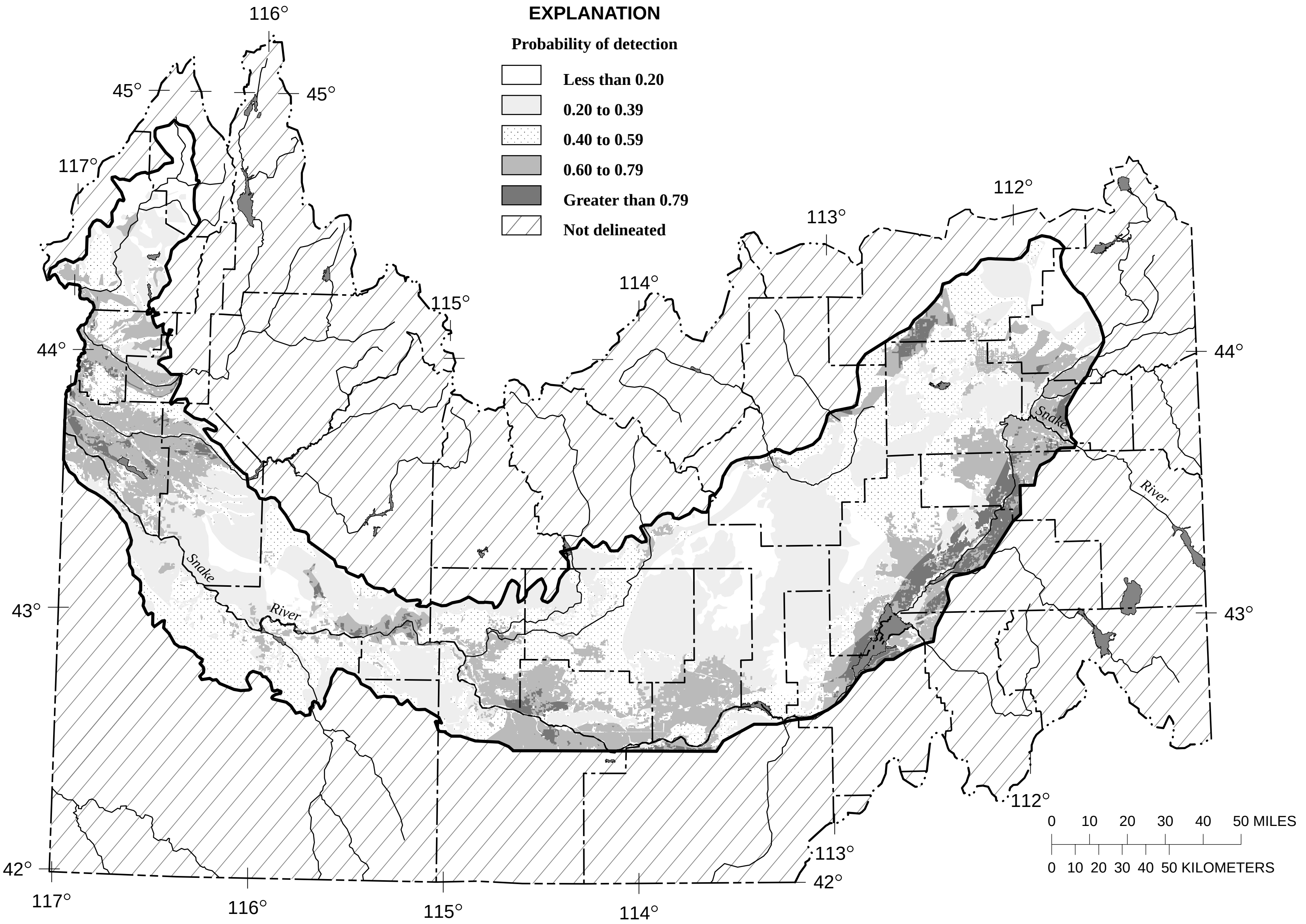


Figure 10. Preliminary map showing the probability of detecting atrazine/desethyl-atrazine concentrations greater than 0.001 microgram per liter in water from wells 150 feet deep, Idaho part of the Snake River Plain.

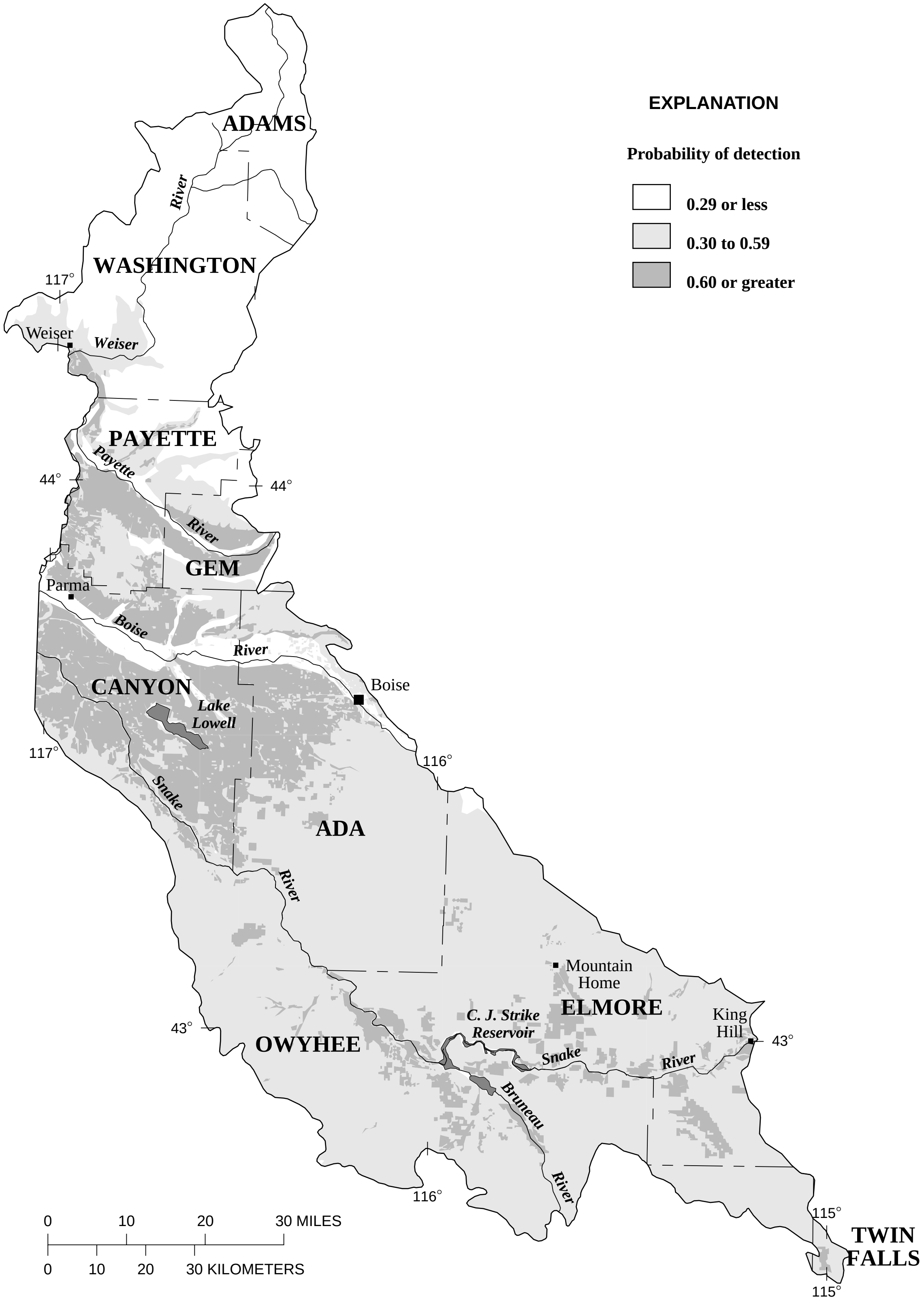


Figure 11. Probability of detecting nitrite plus nitrate as nitrogen concentrations greater than background levels (2 milligrams per liter), Idaho part of the western Snake River Plain.

process of denitrification, the chemical process by which nitrates are converted to nitrogen gas. Denitrification can lead to a net loss of nitrate from the soil. Thus, wells associated with soils high in organic content might be less likely to be contaminated by excess nitrate. The coefficients for organic content in the model exactly reflect this relation. The model is well fit, according to linear regression of predicted versus actual detections. The probability of excess nitrate based on the preferred model is shown in figure 11.

The nitrate probability map and the atrazine probability map for the WSRP have many features in common because both models are influenced by land use and soil characteristics. Therefore, areas that have high probabilities of atrazine detection also tend to have high probabilities of nitrate detection (for example, Canyon and northern Ada Counties). The effect of soil organic content is especially evident in the nitrate probability map. For example, boundaries between the narrow, low-probability areas adjacent to the Boise and Payette Rivers and contiguous higher probability areas coincide with soil map boundaries and reflect differences in the amount of organic material in the soil.

SUMMARY

Ground-water quality is an ongoing concern in Idaho because ground water provides an ever-growing proportion of the State's drinking water. Pesticides and nitrate have been detected in many aquifers in the State. Atrazine was the most commonly detected pesticide in ground water sampled statewide.

Maps showing the vulnerability of areas to ground-water contamination are important tools used by resource protection and regulatory agencies to help protect ground-water quality. In an earlier study (1998), the U.S. Geological Survey, in cooperation with the Idaho Department of Agriculture, produced maps showing the probability of detecting atrazine and elevated concentrations of nitrate in ground water of the upper Snake River Basin. This study adopted methods used in the 1998 study to produce maps showing the probability of detecting atrazine and nitrate in ground water of the western Snake River Plain (WSRP). The maps presented here, together with the previously published maps for the upper Snake River Basin, provide a sound hydrogeologic basis for atrazine and nitrate management in all of southern Idaho.

The atrazine probability map for the WSRP was produced by overlaying ground-water quality data with hydrogeologic and anthropogenic data in a geographic information system (GIS). A data set was produced in which each well had corresponding information on land use, geology, precipitation, soil characteristics, depth to ground water, well depth, water level, and atrazine use. A variety of statistical methods were used to evaluate correlations between atrazine concentrations or detections and the independent variables. The data then were further analyzed by logistic regression using a statistical software package. Several preliminary multivariate models were constructed, and those that best predicted the detection of atrazine were selected. The multivariate models then were entered into a GIS and the probability maps were produced.

Land use, precipitation, soil hydrologic group, and well depth were significantly correlated with atrazine detections in the WSRP. These variables also were important in the 1998 probability study of the upper Snake River Basin.

A preliminary atrazine probability map for the entire Snake River Plain in Idaho, based on a data set representing that region, also was produced. In areas where this map overlaps the 1998 map of the upper Snake River Basin, the two maps show broadly similar probabilities of detecting atrazine.

Logistic regression also was used to produce a preliminary map showing the probability of detecting elevated nitrate in ground water of the WSRP. Nitrogen input and loss data were estimated and an overall nitrogen balance for the WSRP was prepared. Inputs comprised fertilizer and manure application, septic tank effluent, legume crops, and atmospheric nitrogen deposition. Losses comprised crop uptake, denitrification, and crop decomposition. No clear statistical correlations were identified between concentration of nitrate in ground water and net nitrogen input, total nitrogen input, or any of the simple components of the nitrogen balance. Elevated nitrate concentrations were correlated with land use, soil organic content, well depth, and water level. The effectiveness of the probability models for atrazine and nitrate might be improved if more detailed data were available for atrazine and fertilizer application, respectively.

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.

Crockett, J.K., 1995, Idaho Statewide Ground Water Quality Monitoring Program—summary of results, 1991 through 1993: Idaho Department of Water Resources Water Information Bulletin no. 50, pt. 2, 67 p.

Holtschlag, D.J., and Luukkonen, C.L., 1996, Vulnerability of ground water to atrazine leaching in Kent County, Michigan: U.S. Geological Survey Water-Resources Investigations Report 96–4198, 49 p.

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

Idaho Department of Water Resources, 1995, Quality assurance plan for Idaho's Statewide Ground Water Quality Monitoring Program: Idaho Department of Water Resources, 15 p.

Maupin, M.A., 1991, Depth to water in the western Snake River Plain and surrounding tributary valleys, southwestern Idaho and eastern Oregon, calculated using water levels from 1980 to 1988: U.S. Geological Survey Water-Resources Investigations Report 91–4020, 1 sheet, scale 1:750,000.

Molnau, Myron, 1995, Mean annual precipitation, 1961–1990, Idaho: Moscow, University of Idaho, Agricultural Engineering Department, State Climate Program, scale 1:1,000,000.

Newton, G.D., 1991, Geohydrology of the regional aquifer system, western Snake River Plain, southwestern Idaho: U.S. Geological Survey Professional Paper 1408–G, 52 p., 1 pl.

Ott, R.L., 1993, An introduction to statistical methods and data analysis: Belmont, Calif., Duxbury Press, Wadsworth Publishing Company, 1051 p.

Rupert, M.G., 1994, Analysis of data on nutrients and organic compounds in ground water in the upper Snake River Basin, Idaho and western Wyoming, 1980–91: U.S. Geological Survey Water-Resources Investigations Report 94–4135, 40 p.

——— 1996, Major sources of nitrogen input and loss in the upper Snake River Basin, Idaho and western Wyoming, 1990: U.S. Geological Survey Water-Resources Investigations Report 96–4008, 15 p.

——— 1997, Nitrate (NO2+NO3–N) in ground water of the upper Snake River Basin, Idaho and western Wyoming, 1991–95: U.S. Geological Survey Water-Resources Investigations Report 97–4174, 47 p.

——— 1998, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate (NO2+NO3–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.

Rupert, M.G., Dace, Tana, Maupin, M.A., and Wicherski, Bruce, 1991, Ground-water vulnerability assessment, Snake River Plain, southern Idaho: Boise, Idaho Department of Health and Welfare, Division of Environmental Quality, 25 p.

Rupert, M.G., Stone, M.A.J., and Ott, D.S., 1996, National Water-Quality Assessment Program—nitrate and pesticides in ground water, Blaine, Cassia, Lincoln, and Minidoka Counties, south-central Idaho: U.S. Geological Survey Fact Sheet FS–246–95, 4 p.

Snyder, D.T., Wilkinson, J.M., and Orzol, L.L., 1998, Use of a ground-water flow model with particle tracking to evaluate ground-water vulnerability, Clark County, Washington: U.S. Geological Survey Water-Supply Paper 2488, 63 p.

Solley, W.B., Pierce, R.R., and Perlman, H.A., 1993, Estimated use of water in the United States in 1990: U.S. Geological Survey Circular 1081, 76 p.

——— 1998, Estimated use of water in the United States in 1995: U.S. Geological Survey Circular 1200, 71 p.

SPSS, Inc., 1999, SYSTAT 9.0, Statistics—software documentation: Chicago, Ill., SPSS, Inc., 2 vols., 1212 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 Magazine, v. 35, no. 6, p. 1029–1039.

U.S. Department of Agriculture, 1991, State soil geographic data base (STATSGO): U.S. Department of Agriculture, Soil Conservation Service, Miscellaneous Publication no. 1492, 88 p.

Vanderlaan, M., Stanker, L.H., and Watkins, B.E., 1990, Immunoassays for trace chemical analysis, in Vanderlaan, M., Stanker, L.H., Watkins, B.E., and Roberts, D.W., Immunoassays for trace chemical analysis: Washington, D.C., American Chemical Society, ACS Symposium Series 451, p. 2–13.

Whitehead, R.L., 1986, Geohydrologic framework of the Snake River Plain, Idaho and eastern Oregon: U.S. Geological Survey Hydrologic Investigations Atlas HA–681, 3 sheets, scale 1:100,000.

——— 1992, Geohydrologic framework of the Snake River Plain regional aquifer system, Idaho and eastern Oregon: U.S. Geological Survey Professional Paper 1408–B, 32 p., 6 pls.

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.