GIS-BASED STATISTICAL, GEOSTATISTICAL, AND STOCHASTIC ANALYSES OF NITRATE CONTAMINATION IN THE CIMARRON TERRACE AQUIFER IN OKLAHOMA

By

MOTI LAL K.C.

Bachelor's Degree in Civil Engineering

Tribhuvan University

Institute of Engineering

Lalitpur, Nepal

2002

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2007

GIS-BASED STATISTICAL, GEOSTATISTICAL, AND STOCHASTIC

ANALYSES OF NITRATE CONTAMINATION IN THE

CIMARRON TERRACE AQUIFER IN

OKLAHOMA

Thesis Approved:

Dr. William F. McTernan

Thesis Co-Advisor

Dr. Avdhesh K. Tyagi

Thesis Co-Advisor

Dr. Mahesh N. Rao

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGEMENTS

My sincere gratitude is warmly extended to my co-advisors, Dr. William F. McTernan and Dr. Avdhesh K. Tyagi, who provided continual technical support, encouragement and a never ending sense of humor throughout my academic endeavors. I am very much grateful to Dr. Tyagi for providing the financial assistance in the form of research assistantship. This thesis would have not been completed without the support of many people who have been very kind to me. Their advise, guidance, and patience have enriched my whole graduate experience. I would like to take this opportunity to thank my thesis co-advisor, Dr. William F. McTernan, who has been very supportive and has provided exemplary guidance and supervision. I appreciate his willingness to help me think through issues in different levels. I really learned a lot from him. I would also like to thank Dr. Mahesh N. Rao for his time, valuable comments and support as a committee member.

I am really grateful to my late father, Jay Kisan Khatri, who always supported and showed me the right path. Thank you mom for your love. Especial thanks also go to my brother, Bali Ram Khatri, who had really worked hard for my abroad study plan. I would especially like to thank my fiancée, Ojaswi for her sacrifice, patience and moral support during my graduate studies.

Lastly, I would like to thank the City of Enid in Oklahoma for providing water quality data for this research.

TABLE OF CONTENTS

Chapter **Page**

Chapter	Page

LIST OF FIGURES

Figure **Page**

Figure	Page

Figure	**Page**

Figure **Page**

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Nitrate is the most common groundwater contaminant in the United States (Burkart and Stoner, 2002) because it is highly leachable in soils. Nitrate can accumulate in groundwater to high levels as more nitrogen is applied to the land surface every year. Agricultural activities are possibly the most significant anthropogenic source of nitrate contamination in groundwater (Livingston and Corey, 1998). This contaminant in groundwater is an indicator of overall water quality that has been used in agricultural research to assess the effectiveness of nitrogen management strategies (Hong et al., 2006). Numerous studies throughout the United States have shown that Midwest agricultural areas including that of north-west Oklahoma tend to have among the highest nitrate levels in groundwater in the nation (Spalding and Exner, 1993; Nolan et al., 1997; Bukart and Stoner, 2002). Highly permeable soils, shallow well depths and intensive agriculture are the key factors associated with high nitrate levels in those areas.

Potential sources of nitrate in groundwater include inorganic fertilizers, animal manure, septic systems and atmospheric deposition. Fertilizer nitrogen that is not taken up by plants, are either volatilized or carried away by surface runoff. Nitrogen from surface runoff leaches to the groundwater in the form of nitrate and can persist in shallow groundwater for years. Natural sources of nitrate include organic nitrogen in plant matter and fixed ammonium in till and loess deposits (Boyce et al., 1976; Hendry et al., 1984).

Nitrate is highly soluble in water and is not prone to ion exchange (Stumm and Morgan, 1996). Nitrate itself is not volatile and, thus, can not be lost through volatilization but it can be lost through denitrification which is a microbial process that transforms nitrate into nitrogen, a harmless gas that constitutes approximately 80% of the atmosphere. The entire nitrogen cycle consists of ammonia is oxidized to nitrites in the presence of water and then again to nitrates.

Due to high solubility and mobility, nitrate leaches through the soil zone to underlying aquifers. Nitrate is also not affected by chlorination, the most common method of treating most public water. Reverse Osmosis is one method to remove nitrate from water but this is an expensive process. Additional treatment technologies include ion exchange and denitrification (Kapoor and Viraraghavan, 1997).

Groundwater vulnerability was defined by the National Research Council (1993) as "the tendency or likelihood for contaminants to reach a specified position in the ground-water system after introduction at some locations above the uppermost aquifer" (Rupert, 2003).The study of groundwater vulnerability has been conducted in many areas using the DRASTIC method (Aller et al., 1985). The DRASTIC method has been extensively used to develop maps at a variety of scales such as national, (Lynch et al., 1994), statewide (Hamerlinck and Ameson, 1998), and local (Shukla et al., 2000). This index method is a popular approach to groundwater vulnerability assessments because it is less expensive, straightforward, and uses data that are readily available, and produces a visual map that can easily be interpreted and incorporated into the decision-making process. This model includes: Depth to water, Net Recharge, Aquifer media, Soil media, Topography, Impact of vadose zone media, and Hydraulic Conductivity of the aquifer. The subjective point rating system that is based on best professional judgment and the lack of calibration to actual groundwater quality data has been the major deficiency of DRASTIC in predicting ground water vulnerability (Koterba et al., 1993; Rupert, 2001). Rupert (2001) calibrated the vulnerability point ratings to measured nitrate value in ground water using non parametric statistical test to overcome some of the problems of traditional DRASTIC vulnerability mapping. DRASTIC is deficient however in defining actual impact to an aquifer from a pollution event. The subjective score only relates to aquifer vulnerability not to actual conditions where pollution may occur. As a supplement to DRASTIC, other modeling approaches have been attempted. Chief among them is logistic regression which overcomes some of the deficiencies of traditional vulnerability mapping also by calibrating to actual ground water quality data (Scanlon et al., 2003)

Determining where and to what extent the groundwater is at risk of nitrate contamination can help managers build aquifer protection strategies. Evaluation of nitrate contamination and its relationship to explanatory variables has been addressed in national and regional scale in many previous studies (Nolan et al., 2002; Squillace et al., 2002; Greene et al., 2004). Nolan et al. (1997), and Tesoriero and Voss (1997) used Geographic Information System (GIS) buffer and overlay analysis, and statistical analysis to determine risk of nitrate contamination in shallow aquifers. Use of geostatistics and process-based simulation models are other approaches to assessing aquifer vulnerability to contamination.

Statistical methods are commonly used because they are inherently flexible, can readily accommodate differences in spatial scale and can effectively describe uncertainty (NRC, 1993). Stochastic models can also identify and attempt to address the inherent variability of natural phenomena and ultimately can address uncertainties. Geostatistics provide tools to describe and predict spatial variation, carry out spatial interpolation and obtain a probabilistic assessment of groundwater contaminants. Use of map algebra in Geographic Information System (GIS) can readily enhance those processes to analyze probabilistic study of contaminant occurrences and hence to create risk maps and their use in decision-making for risk management.

1.2 Statement of the problem

Total ground-water withdrawals in United States were 77,500 Mgal/d in 1995 which provided drinking water for more than one-half of the people in the United States (Solley et al., 1995). The City of Enid and its surrounding area in Oklahoma are solely dependent on groundwater for its drinking water supply. More than 3 billion gallons of groundwater annually are pumped from the Cimarron terrace aquifer. Pumpage data provided by the City of Enid reveals that this amount accounts for approximately 90 percent of drinking water supply for the City. Contaminants in groundwater that may cause health problems, such as nitrate, are of great concern. Infants under six months of age are most vulnerable to elevated levels of nitrates in drinking water. Methemoglobinemia in infants is a potentially fatal disease and results from low oxygen levels in the blood caused by injestion of high nitrate water (Spalding and Exner, 1993). Mathemoglobinemia “A blue baby syndrome” results from the oxidation of reduced iron,

Fe^{2+}, in hemoglobin to its oxidized form Fe^{3+}. The resulting meth-moglobin (MeHb) is unable to release oxygen to body tissue (Comly, 1945; Bosch et al., 1950). The U.S Environmental Protection Agency (US EPA) has established a Maximum Contaminant Level (MCL) of 10 mg/l nitrate as nitrogen as drinking water criteria. Increased risk of Non-Hodgkin's lymphoma has also been related to nitrate concentration $\geq$ 4 mg/l nitrate as nitrogen in community water supply wells (Ward et al., 1996).

An understanding of relative importance of various sources of nitrate in groundwater is important for agricultural management practices. Some techniques are available to extend monitoring data over space and time as well as identifying the most critical, contributory variables associated with nitrate contamination of groundwater. GIS and geostatistics can quantify the distribution of spatial pattern of monitoring nitrate data over space and time. Logistic regression has been used to determine most significant variables in several national, regional and local assessments of nitrate and pesticide contamination (Teso et al., 1996; Tesoriero and Voss, 1997; Nolan et al., 1998; 2002; Nolan and Stoner, 2000; Nolan, 2001; Scanlon et al., 2003; Greene et al., 2004). In addition to logistic regression, this research focuses on stochastic models that are used to address uncertainty associated with each significant variable. Once sensitivities of input variables are identified, proper management can be applied to stem further deterioration of ground water quality.

1.3 Investigative approach

Deterioration of groundwater quality from nitrate contamination throughout the world has grown significantly in recent years. To overcome this challenge, management objectives require scientific assessments of the potential for groundwater resources to become contaminated from anthropogenic, as well as natural source of contamination. The difficulty and high cost of remediating contaminated groundwater (McKay and Cherry, 1989) have increased the attention of regulatory agencies. It is important to identify possible sources of nitrate contamination and areas that are susceptible to contamination for land managers to build aquifer protection strategies. Following are overall solution approaches of this thesis:

a) Use geostatistics and GIS to prepare broad database of nitrate from point values to area-wide values throughout the aquifer, to quantify the distribution of spatial pattern of nitrate and to perform probabilistic assessment of nitrate to create risk maps and their use in decision-making for risk management

b) Use GIS to extract land cover variables and use these data to establish the best areas of well influence when analyzing the relationship between land cover and groundwater nitrate for sampled wells.

c) Determine explanatory variables that significantly influence nitrate concentration in Cimarron terrace aquifer using logistic regression models.

d) Address uncertainty associated with significant variables of final logit models in predicting probability of nitrate concentration exceeding 4 mg/l and 10 mg/l of NO_3-N. While there are many methods available to address uncertainty but those based upon the Monte Carlo Algorithm are frequently used (Kaplan and McTernan, 1993). In addition to determining most significant variables, this research will focus on stochastic modeling to address uncertainty.

CHAPTER 2

DESCRIPTIONS OF STUDY SITE

2.1 Location of study area

The Cimarron terrace aquifer is located in northwestern Oklahoma. The location map of study site is shown in figure 2-1. This aquifer consists of 1242.5 square miles of area and includes Quaternary-age terrace deposits. The deposits are unconformably overlying Permian red-bed formations (Reely, 1992). Geographically the study area extends from the 98° 36' W to 97°44' W on the horizontal and 36°10' N to 36°34' N on the vertical. The study area is a part of Cimarron River watershed and is connected by a dense network of smaller streams.

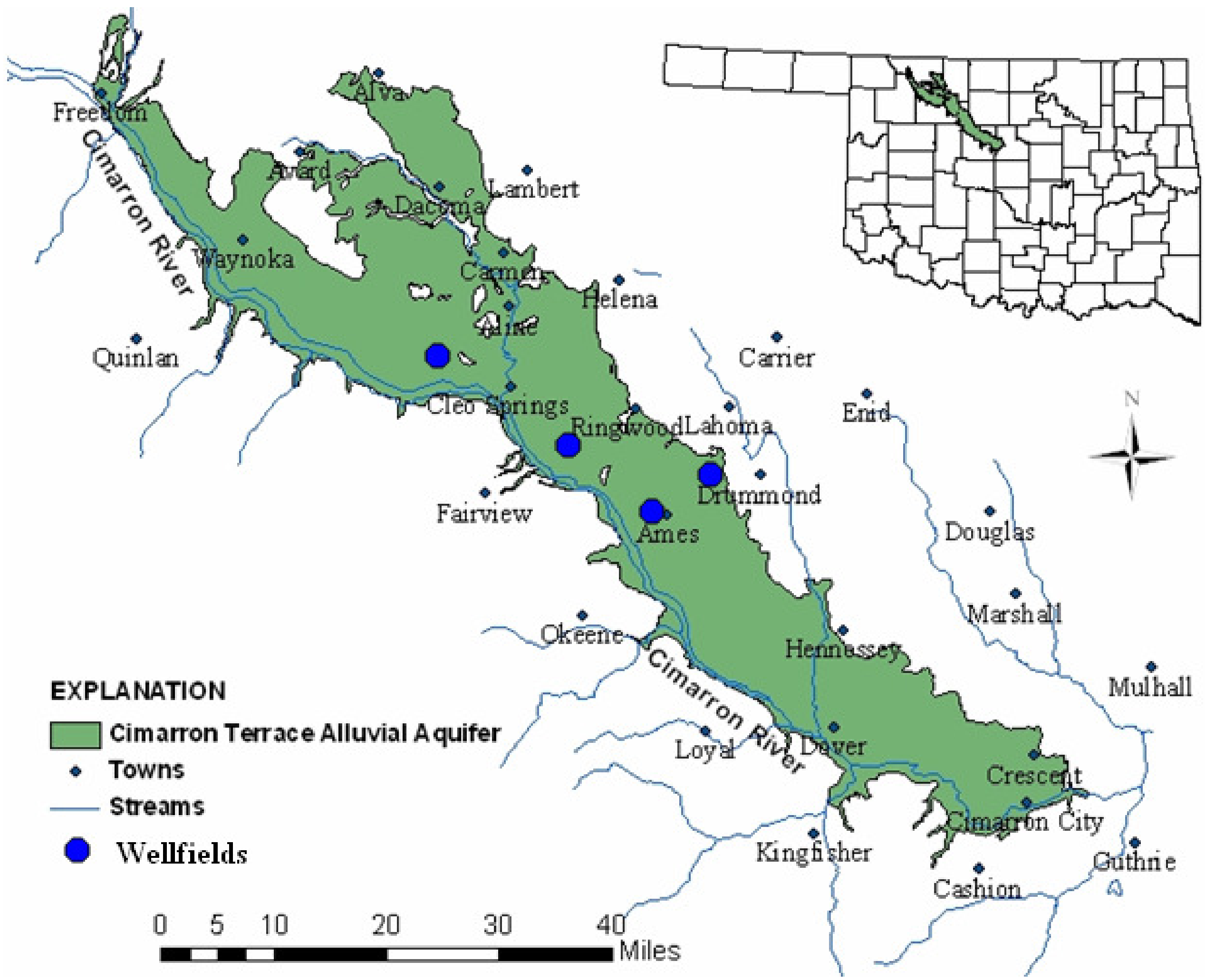


Figure 2-1. Location map of Cimarron terrace aquifer

2.2 Water use

Groundwater in the Cimarron terrace aquifer has been an important economic resource for northwest Oklahoma. In 2000, approximately 63 % of total water use were withdrawn from the Cimarron terrace aquifer by five counties; Alfalfa, Garfield, Kingfisher, Major, and Woods (Masoner and Mashburn, 2004). The aquifer produced 4.27 billion gallons of water for public supply in 2000 and more than 4.40 billion gallons of water were used for irrigation and livestock purposes (Masoner and Mashburn, 2004). More than 3 billion gallons of groundwater annually are pumped by the City of Enid from the Cimarron terrace aquifer. Pumpage data provided by City of Enid reveals that this amount accounts for 90 percent of drinking water supply for the City. Public water supply withdrawn by City of Enid from 2004 to 2006 indicates that the water demand seems to be generally increasing (Fig. 2-2), making protection of the existing groundwater resource even more important.

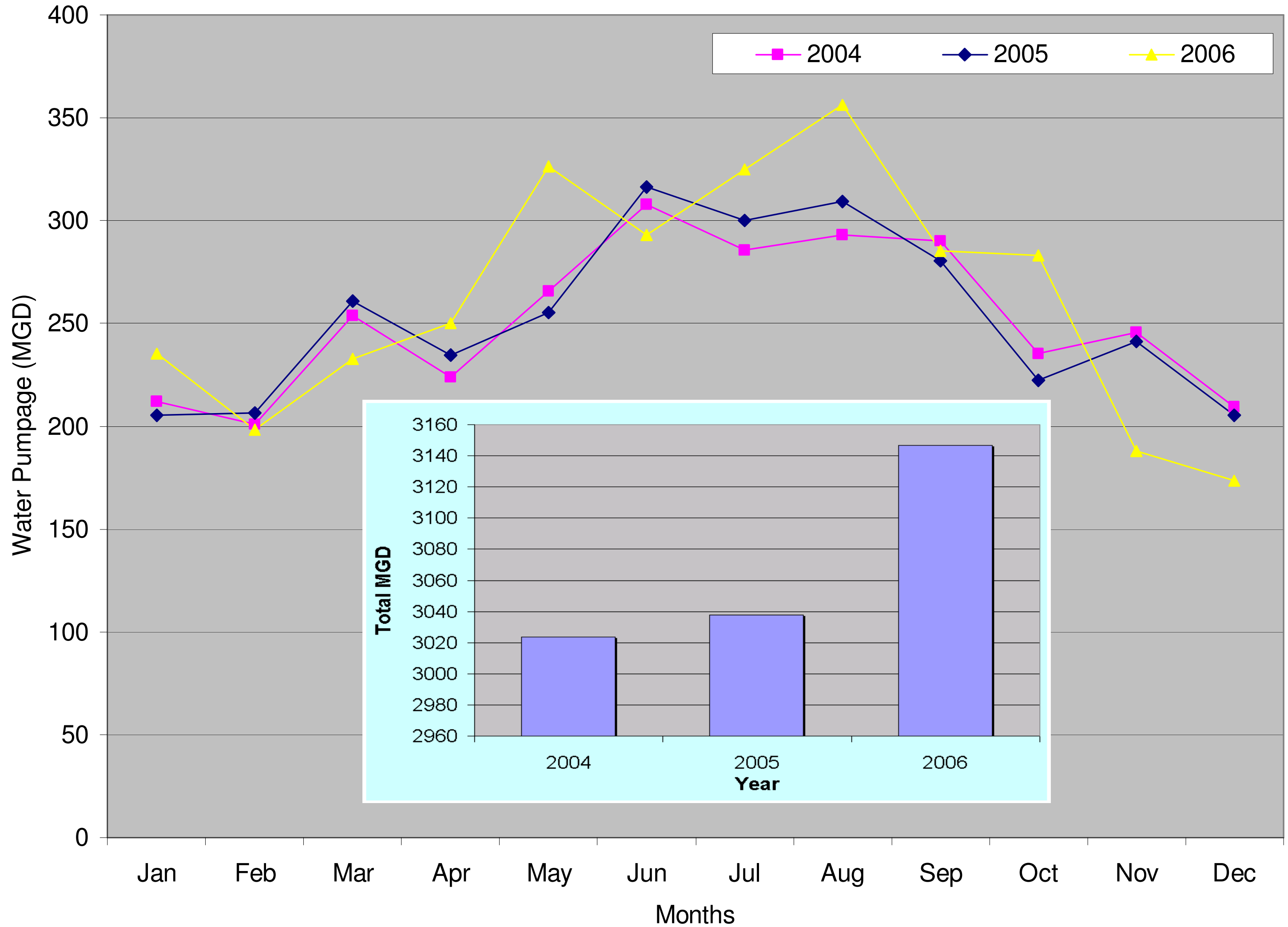


Figure 2-2. Water withdrawn from Cimarron terrace aquifer by the City of Enid

2.3 Land covers

Land use land covers in 1992 (Fig. 2-3) and 2001 (Fig. 2-4) for the entire study area were taken from National Land Cover Database (NLCD). Agriculture was found to be the most predominant land cover in the Cimarron terrace aquifer. Agricultural land in the study area refers to areas that have been planted or are intensively managed for the production of livestock for food (Masoner and Mashburn, 2004). The agricultural land use overlying the aquifer in 1992 consisted of 41.93 percent small grains, 7.04 percent cultivated crops, and 6.24 percent pasture and, while in 2001, 46.54 percent were in cultivated crops and 0.32 percent pasture and hay. Additional land cover areas in 2001 in the aquifer consisted of grassland (41.09 percent), developed area (5.11 percent), and forested upland (4.25 percent). Table of land cover distributions are shown in table 2-1 below. Marred

Table 2-1. Distribution of land covers in 1992 and 2001

Land cover types	1992	2001
Open Water	1.89	1.32
Developed Area	0.35	5.11
Barren	0.55	0.34
Forested Upland	4.39	4.25
Shrubland	10.44	0.08
Grasslands/Herbaceous	26.47	41.09
Pasture/Hay	6.24	0.32
Cultivated Crops	7.04	46.54
Small Grains	41.93	0.00
Wetlands	0.70	0.95

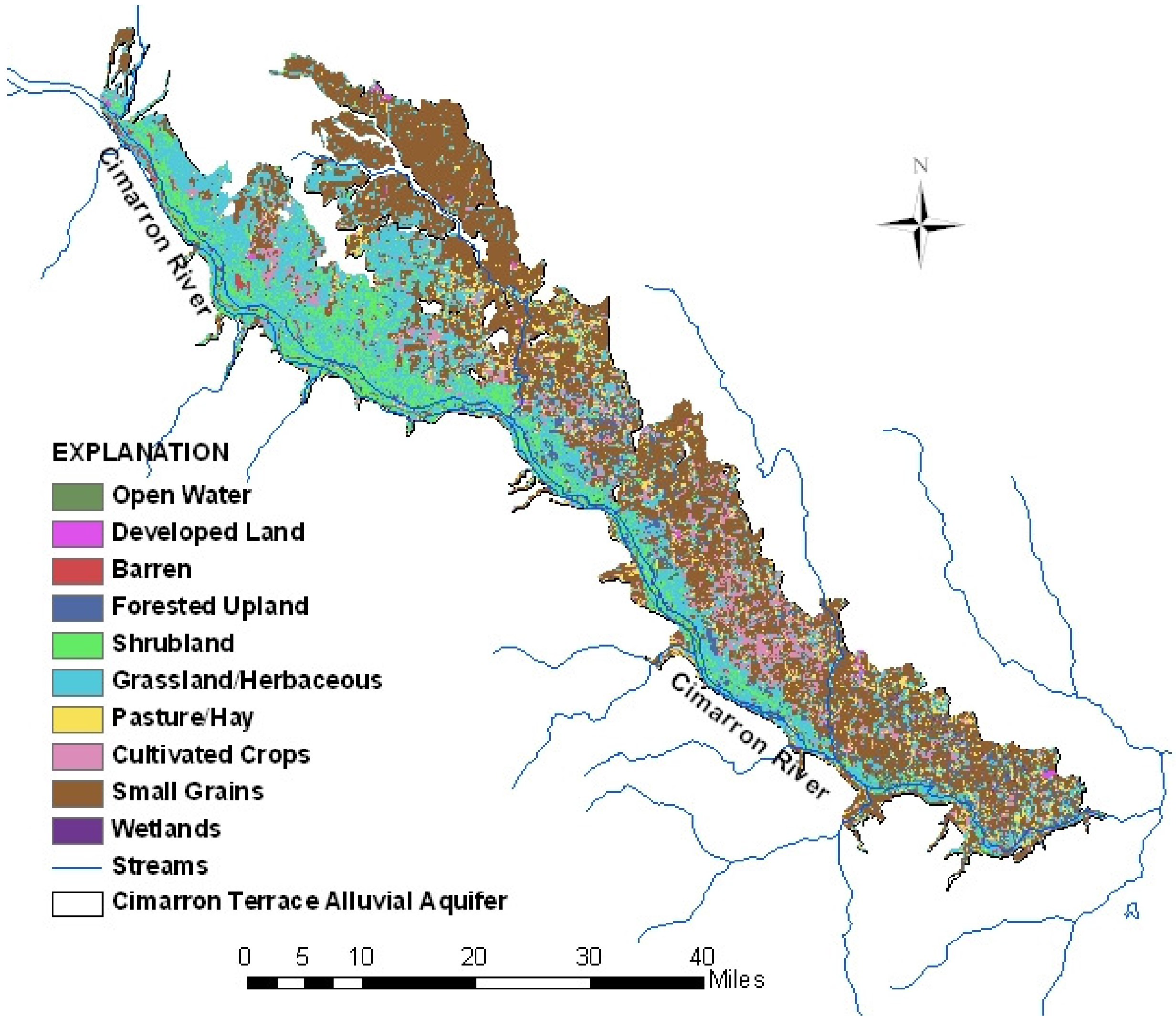


Figure 2-3. Major land cover distribution taken from 1992 National Land Cover Database

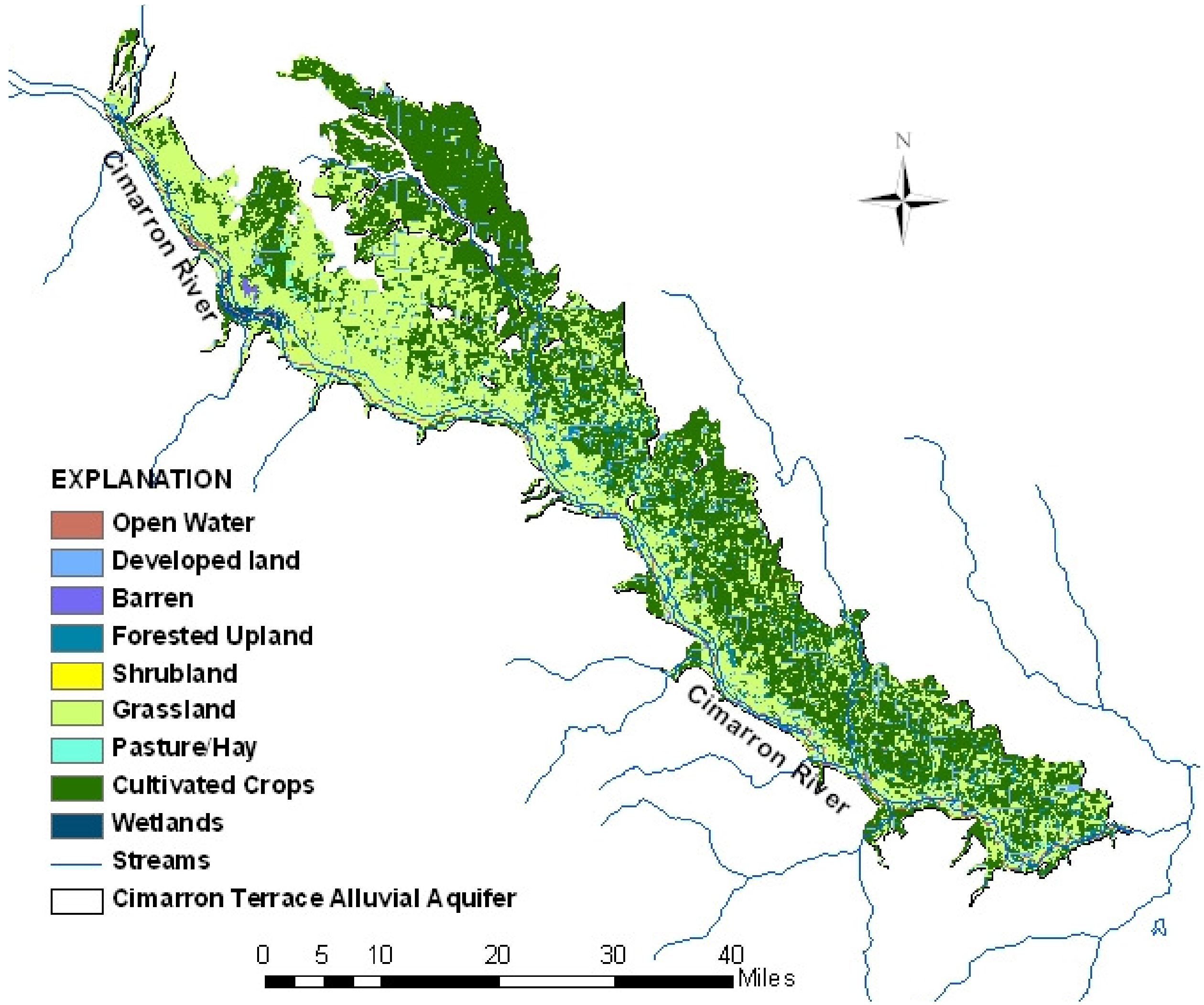


Figure 2-4. Major land cover distributions taken from 2001 National Land Cover Database

To study land cover changes between 1992 and 2001 before starting to analyze their effects to groundwater contamination, land cover areas were compared (Fig. 2-5). Total agriculture area in 1992 was 55.21 percent while in 2001 it was 46.86 percent. The figure 2-5 explains that the cultivation of small grains has been totally changed to the production of cultivated crops from 1992 to 2001. Grassland and developed areas have also been increased from 26.47 percent to 41.09 percent and 0.35 percent to 5.11 percent respectively. The reason of such high agriculture cultivation in the Cimarron terrace aquifer is the modern irrigation systems that have facilitated an increase in the cultivation of wheat, corn, and oats (USDA, 1996).

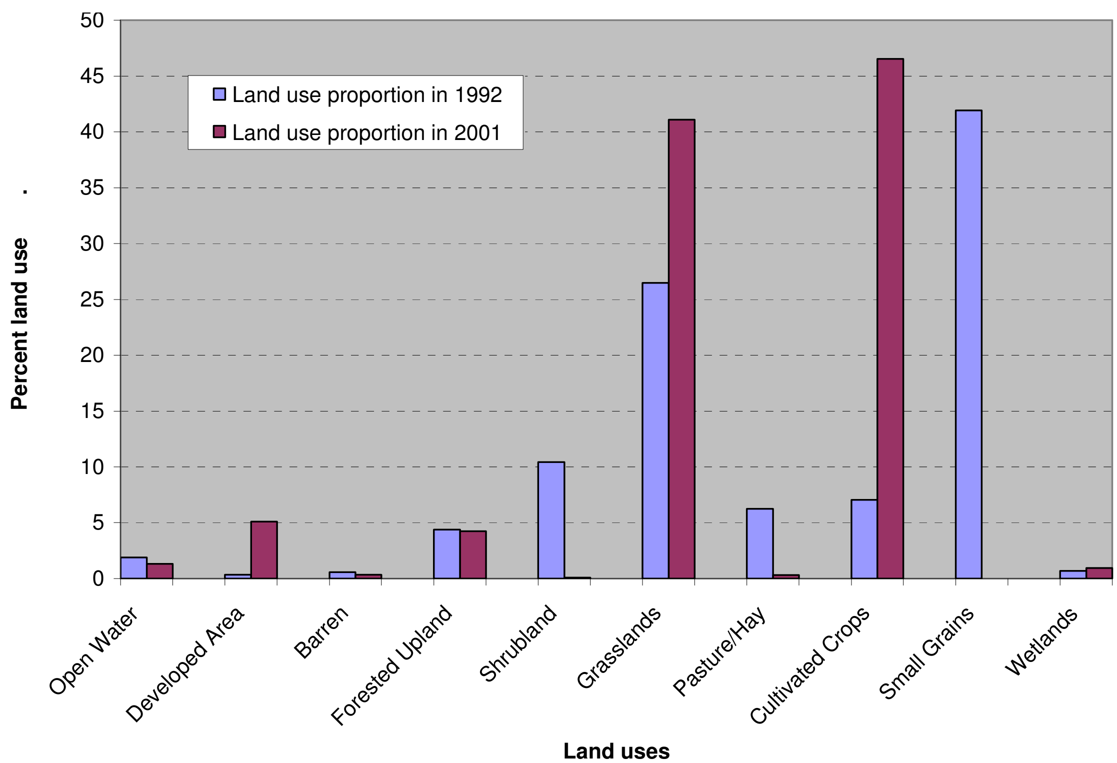


Figure 2-5. Comparison of land cover proportions between 1992 to 2001 in the study area overlying the Cimarron terrace aquifer

2.4 Hydrogeology

The general direction of groundwater flow within the Cimarron terrace deposits is from northeast to southwest, flowing towards the Cimarron River, except where flow direction is influenced by perennial tributaries to the Cimarron River (Adams and Bergman, 1996). Cimarron terrace dunes were originally deposited by the southward migration of the ancestral Cimarron River (Adams and Bergman, 1996). Because of spatially varied lithologies, ground water surface geometry is undulating through out the aquifer (Reely, 1992). Surface water is not a major source of recharge to the aquifer. The two major sources of recharge in the aquifer are infiltration of precipitation and irrigation return flow (Adams and Bergman, 1996). The regional groundwater gradient is 0.0035 feet/feet and the saturated thickness ranges from 0 to over 80 feet in several locations within the study area (Reely, 1992). Based on pumping tests, transmissivity in the Cimarron terrace aquifer ranges from 800 ft^2/day to 10, 200 ft^2 /day with an average value of 2,670 ft^2/day, while the specific yield ranges from 0.018 to 0.131, with an average value of 0.065 (Reed et al., 1952).

CHAPTER 3

METHODOLOGY

3.1 Background

This chapter explains methods used in this research to address problems mentioned in Chapter 1. Investigative approaches utilized two types of statistical models-logit and geostatistic, which were used for determining the most significant explanatory variables that best explained the occurrence and distribution of elevated levels of nitrate in shallow groundwater Cimarron terrace aquifer in northwestern Oklahoma. Stochastic modeling, a Monte Carlo simulation, was used to evaluate the impact of variation in previously determined significant variables of logistic models. An overview of water quality parameter selected and database development for logistic regression models are also being covered by this chapter.

3.2 An overview of water quality parameter selected

The City of Enid performed sampling of wells from 1997 to 2005 in four wellfields of Cimarron terrace aquifer, Oklahoma. A total of 821 samples were collected in four wellfields located at central part of the aquifer and these sampling wells in each wellfield are clustered in much closed space. Each well was given a co-ordinate so that it could be easily mapped out for further analysis. The City of Enid, Oklahoma, measured nitrate as nitrogen (NO_3-N) as water quality parameter along with other chemical characteristics such as chloride, TDS, manganese etc. Nitrates were frequently observed to be exceeded Maximum Contaminant Level (MCL) set by EPA. Such elevated levels of nitrate can cause low oxygen levels in the blood of infants ‘known as blue-baby syndrome’, a potentially fatal condition (Bosch et al., 1950). These data were analyzed in Minitab (Minitab, 2003) to understand the general overview of selected water quality parameter before proceeding to advanced methodologies such as GIS, geostatistics, logistic regression, and stochastic modeling. Geostatistic and GIS were used to expand these point values to represent area-wide values. These tools analyzed spatial-temporal trend analysis of nitrate monitoring data in the aquifer and then probabilistic assessment for specified thresholds of nitrate as nitrogen. These tools provided an overview of nitrate point values to area-wide values in the aquifer.

It is important for land managers to identify major sources of nitrate contamination in the groundwater because once these sources are identified; proper management can be applied to stem further deterioration of groundwater quality. To address this problem, logistic regression models were executed to determine most significant sources of nitrate in the aquifer. Furthermore, stochastic modeling was implemented to magnify inherent variability of previously determined significant explanatory variables of those logistic regression models.

3.3 Geostatistics

Geostatistical estimation methods (David, 1977; Journel and Huijbregts, 1978) were developed to create mathematical models of spatial correlation structures (Isaaks and Srivastava, 1989; Goovaerts, 1997). The fitted function to the experimental variogram provides the input parameters for spatial prediction by kriging (Krige,1951). The application of estimation methodology to problems of environmental pollution has been addressed in many studies (Moore and McLaughlin, 1980; Cooper and Istok, 1988a and 1988b), Istok et al., 1993; Cinnirella et al., 2005).

Monitoring of groundwater quality involves building strategies and methodologies of field surveys for choosing the most reliable possible data at a closely spaced network of observational points. The criteria of maximizing information and minimizing costs are always the top priority of water managers or decision makers for planning and evaluating groundwater resources. There are always uncertainties associated with data and manager's priority of maximizing information and minimizing costs which arise the following questions:

1. How much information is required to design efficient monitoring network? Or what is the optimal sample size to achieve this goal?
2. How to identify the optimum locations for further sampling?

One way of approaching at solutions of above questions is to quantify uncertainties associated with the prediction of field values by spatial arrangement of monitoring well data. Kriging geostatistical method, an optimal estimator, always seeks to minimize the estimation uncertainty, represented as estimation variance or kriging variance.

Data collected by the City of Enid, Oklahoma from the year 1997 to 2005 were used in geostatistical analysis. Geostatistical models were fitted to the data using ArcGIS Geostatistical Analyst extension (ESRI, 2006). Ordinary kriging was used to quantify the distribution of spatial pattern of nitrate. Spatial maps were integrated to visualize and quantify areas of temporal difference between 1997 and 2005 using ArcGIS and statistical hypothesis tests in ArcGIS Spatial Analyst. While indicator kriging, that provides a methodology for risk evaluation (Smith and Williams, 1996), was applied to determine indicator variography for probabilistic assessment of NO_3-N.

3.3.1 Ordinary kriging

The core of geostatistical techniques is the analysis of the spatial structure of the variable of interest through variogram analysis. A variogram is a plot of the average squared differences between values as a function of the separation distance. For groundwater nitrate as nitrogen, empirical semi-variogram is defined as (Hendry et al., 1984):

$$\gamma(h,\alpha) = \frac{1}{2N(h,\alpha)} \sum_{i=1}^{N(h)} \left[z(x_i) - z(x_i + h) \right]^2 \qquad (3\text{-}1)$$

where

$\gamma(h,\alpha)$ = semi-variance, which is a function of both the magnitude of the lag distance (h) and its direction α.

N = number of pair values

$Z(x_i)$ = random variable at location x_i

Many kinds of variogram models such as Linear, Spherical, Power, Exponential, Gaussian etc. can be used to transfer $\gamma(h)$ values from the practical model to theoretical model. The figure 3-1 below is drawn to better explain the theoretical semi-variogram. The semi-variance increases with lag distance (h) between sample locations, rising up to a constant value called "Sill" at a given separation distance known as "Range" of spatial dependence. Beyond this separation distance (Range), data do not have significant statistical dependence because variation in the amount tends to be null. At given range, the 'Sill' seeks to estimate the sample variance (σ^2_{krig}) for stationary data. The intercept at y-axis is termed as the nugget effect and is due to measurement errors at microstructure levels. The vertical distance where the variation in amount in the process occurs is called "Partial Sill".

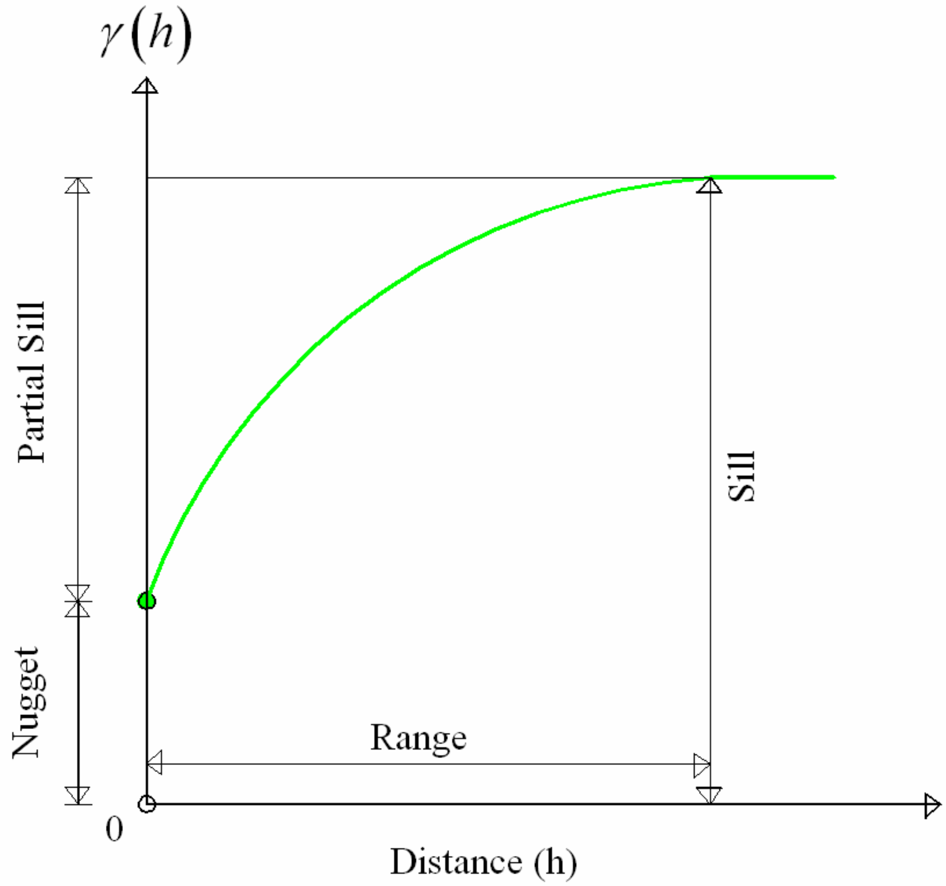


Figure 3-1. Theoritical semi-variogram

Models were fitted to the variogram by components of semi-variogram as explained earlier. The trial and error method was used to fit those models which minimized the square differences between the empirical semi-variogram values and the theoretical model. The directional tool called "Anisotropy" provided by the ArcGIS's Geostatistical Analyst was also used to statistically quantify directional influences while fitting those models. Searching neighborhood option in Geostatistical Analyst was also utilized by defining a circle to enclose the points that were used to predict values at unmeasured locations. The enclosed data points indicated the weights that were associated with each location in the prediction of unknown values.

To evaluate the cross-validation results, statistical criteria (Isaaks and Srivastava, 1989; and Kitanidis, 1997) such as Mean Error (ME) and Root Mean Square-standardized Error (RMSE), were computed. The kriged Mean Error (ME) was used as a criterion for examining the degree of systematic error present and was calculated as:

$$ME = \frac{1}{N}\sum_{i=1}^{N}\left(z_i - z_i^*\right) \tag{3-2}$$

The kriged Root Mean Square-standardized Error (RMSE) was used to test the consistency between the estimation errors and the standard deviation of the actual values and that was calculated as:

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{z_i - z_i^*}{SD}\right\}^2\right]^{\frac{1}{2}} \tag{3-3}$$

Where z_i is the observed value, z_i^* represents the expected value at location i, and SD represents standard deviation of observed values. For a model to predict accurately, the Mean Error (ME) should be close to 0 and the Root-Mean Square-standardized Error (RMSE) should be close to 1 (LaMotte and Greene, 2007).

3.3.2 Spatial-temporal trend analysis using GIS

GIS is a computer based tool that can be used for managing, compiling, and analyzing spatial data. GIS can be used to identify areas affected by groundwater contamination. GIS based groundwater quality maps are important for decision makers because these maps can be used for groundwater planning strategies. GIS Spatial Analyst was used to transform kriging estimates to kriged blocks (size =100m x 100m and points in cell = 10 horizontal x 10 vertical) in order to best represent average nitrate values over discrete blocks.

Ordinary kriging block estimates of groundwater quality monitoring data were integrated into GIS to provide a quantitative, statistical, and weighted means of nitrates defining the statistical significance of geographic apparent change between 1997 and 2005. As mentioned earlier, kriging always seeks to minimize estimation uncertainty which is represented as kriging variance, or standard error in ArcGIS Geostatistical Analyst. This quantity (σ^2_{Krig}) defines the likelihood that a kriged estimate lies within a specified confidence interval under a normal probability distribution. Normal distributions of nitrates in two sampling periods were generated with estimated likelihood (Fig. 3-2). With two kriged estimates at same location in two different sampling periods and their corresponding kriging variances, maps of statistical significance changes in areas were determined at 5% and 30% confidence levels in ArcGIS's Spatial Analyst.

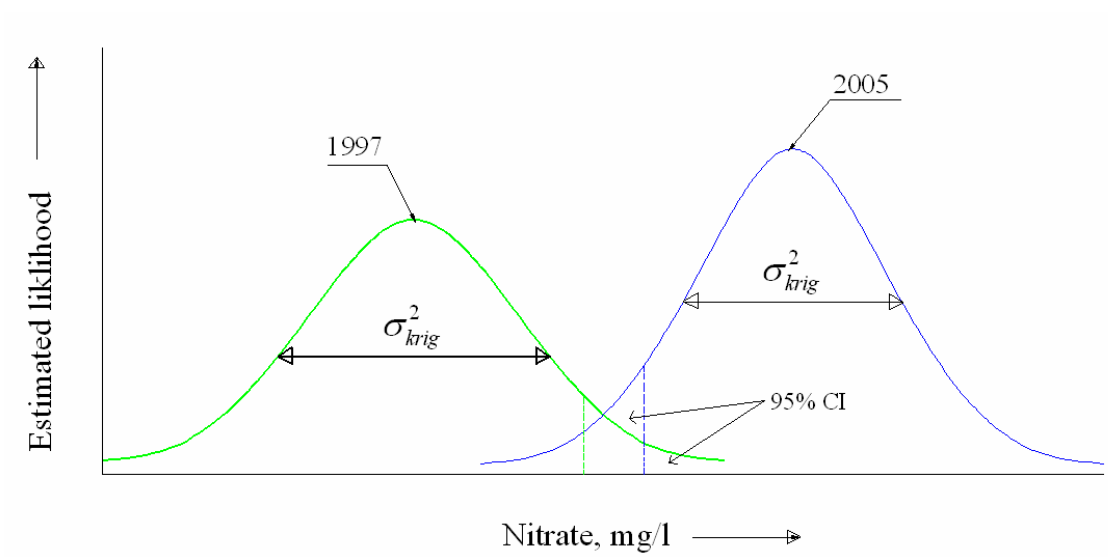


Figure 3-2. Kriging estimates at the same location at two different sampling periods

The applied global null hypothesis of a one-tailed comparison was that one kriging estimate is indistinguishable from the other. GIS Spatial Analyst was extensively used to identify areas affected by nitrates. From statistical significance changes of nitrate maps, areas of nitrate concentrations increased by at least 4 mg/l and 10 mg/l from 1997 to 2005 were determined in specified confidence intervals. In addition to defining temporal trends, areas of decreasing trends and increasing trends in the aquifer were mapped at aforementioned confidence levels.

To quantify areas of spatial-temporal trends, the null hypothesis, $H_0 : \mu_A = \mu_B, or\, H_0 : \mu_A - \mu_B = 0$ (two sided test) was evaluated to test the significance difference at given confidence levels between two kriged blocks of two different sampling periods. The equation used in ArcGIS's Spatial Analyst was

$$z = \frac{D}{s_D} = \frac{\left(X_{krig}\right)_A - \left(X_{krig}\right)_B}{\sqrt{\left(\left(\sigma^2{}_{krig}\right)_A + \left(\sigma^2{}_{krig}\right)_B\right)}} \qquad (3\text{-}4)$$

where X= mean nitrate estimate of kriged blocks,

$\sigma^2{}_{krig}$ = kriging variance, and

A and B are two sampling periods.

To better understand the classification of Z-scores of kriged blocks of average nitrates, a figure of two sided hypothesis test was drawn (Fig. 3-3). The z-score values, amount of distributions falling in normal distribution, were classified based on abovementioned confidence level values (z_α) as explained in figure 3-3 below. Areas rejecting null hypothesis of "means are equal" were displaced in maps as areas defining statistically significant changes between 1997 and 2005.

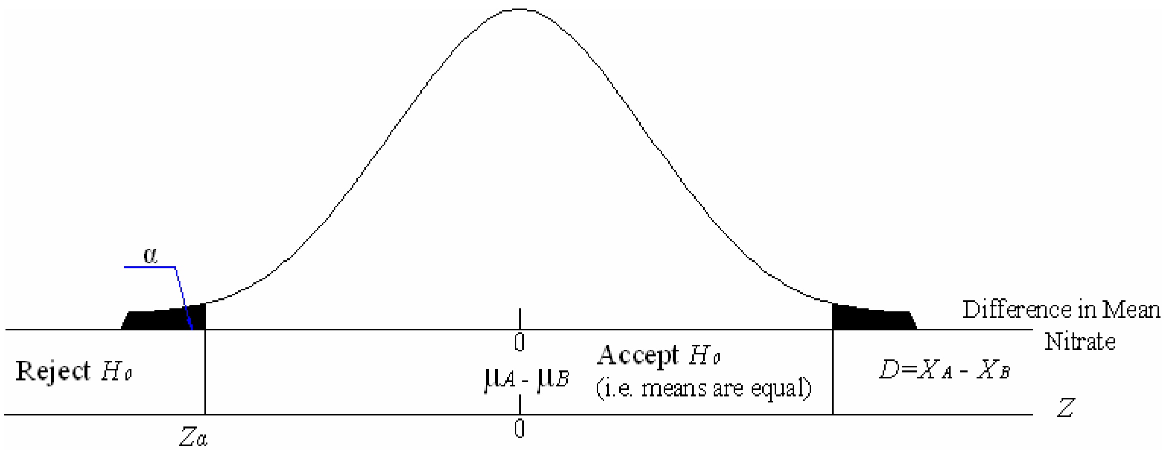


Figure 3-3. Two sided null hypothesis test used in ArcGIS's Spatial Analyst

3.3.3 Indicator kriging

Geostatistics and GIS are essential partners for spatial analysis. The GIS user only needs to interpolate point data so that they can be displayed or visualized, or combined simply with other data. For the decision-making process related to estimates of groundwater contamination at any spatial location, uncertainties associated with these estimates should be recognized. If these estimates are more than a management threshold (e.g. NO_3-N as MCL), safety measures may be applied. Such estimates are usually affected by large uncertainty, occurring from data sampling, modeling and interpolation, which must be quantified to allow an evaluation of the risk involved in any assessments (Buttafuoco et al., 2000). The use of geostatistics allows the user to assess such uncertainty through the determination of a Conditional Cumulative Distribution Function (CCDF) of an unknown attribute value. A non-parametric geostatistical approach is mostly appropriate for downscaling processes (Lanz et al., 2001) and point kriging can be applied to interpolate environmental indicators (i.e. NO_3-N). Numerous studies in various disciplines have used non-parametric geostatistics to define areas with high and low certainty of exceeding a threshold value such as in soil sciences (Halvorson et al.,1996; Castrignano et al., 1999), hydrology (Allard, 1998; Atkinson and Lloyd, 1998), geology (Smith and Williams, 1996), and in environmental science (Goovaerts, 1994; Istok and Rautman, 1996; Krivoruchko, 2001; Cinnirella et al., 2005).

The non-linear Kriging, so called Indicator Kriging technique, defines the probability of a contaminant level exceeding a given threshold value at a given location. Indicator variography is the assignment of a binary transform value, either 0 or 1. Hence, binary indicator function is defined as:

$$I(x) = \begin{cases} 1 & \text{If } Z(x) \geq M_t \\ 0 & \text{If } Z(x) < M_t \end{cases} \tag{3-5}$$

$I(x)$ is the binary variable determined by whether the variable of interest $Z(x)$ is exceeding the threshold M_t at location x = [X,Y].

This indicator variable provides an estimate of the conditional Cumulative Distribution Function (CDF) at a threshold (Smith and Williams, 1996) as:

$$\begin{aligned} M_t : i\left(x; M_t\right) &= E\left\{I\left(x; M_t \mid (n)\right)\right\} \\ &= \Pr ob\left\{Z\left(x\right) \geq \mid \left(n\right)\right\} \\ &= F\left(x; M_t \mid \left(n\right)\right) \\ &= CDF \end{aligned} \tag{3-6}$$

The least square estimate of the indicator [$i(x; M_t)$] is also the least square estimate of its conditional expectation. Thus the ccdf F(($x; M_t$)|(n)] can be estimated by kriging the indicator [$i(x; M_t)$].

The symbol (n) means conditional to n sample data taken in the neighborhood x. Once the indicator values are generated and the variogram is fitted, the models are applied on those values using an ordinary kriging method, an optimal estimator, as follows (Smith and Williams, 1996):

$$F\left(x; M_t \mid (n)\right) \cong \left[i\left(x; M_t\right)\right]^* = \sum_{j=1}^{n} a_j \left(x; M_t\right) i\left(x; M_t\right) \qquad (3\text{-}7)$$

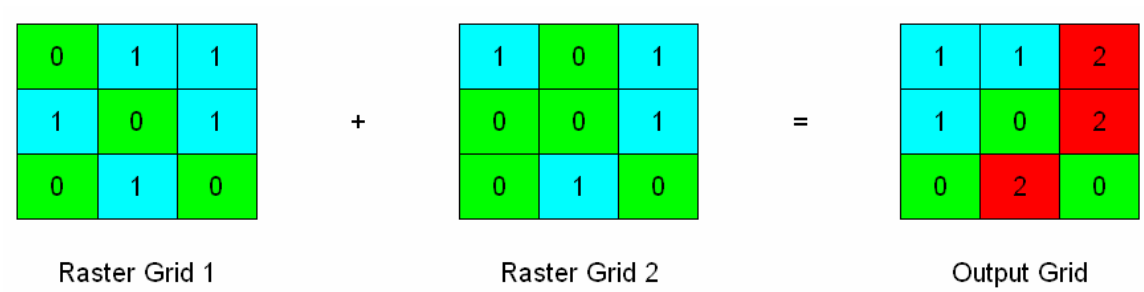
Where $F\left(x; M_t \mid (n)\right)$ estimated value at location x., based on threshold M_k, and $a_j(x;M_t)$ for the j=1,2,…,n kriging weights.

Symbol (*) indicates that the estimated indicator values [$i(x; M_t)$] will take on continuous (0, 1) values rather than the discrete (0, 1) values of the transformed sample data.

3.3.4 Mapping chronic exceedance of NO_3-N using GIS

Another way of addressing uncertainties associated with the prediction of field values by spatial arrangement of monitoring well data is to perform a probabilistic assessment of monitoring nitrate data and extend them to whole aquifer. The use of indicator kriging addressed such uncertainties through the determination of conditional cumulative distribution function of an unknown attribute value. The integration of probability maps of monitoring nitrate data using GIS can help decision makers to build aquifer protection strategies. Probability maps of nitrate data generated from indicator kriging at thresholds 4 mg/l and 10 mg/l have been manipulated for the enhancement of proper decision making to protect ground water quality.

Probability maps of indicator kriging were transformed into block kriging (size =100 m x 100 m and points in cell = 10 horizontal x 10 vertical) using GIS in order to best represent average probability values over discrete blocks. Probability maps of exceeding 4 mg/l and 10 mg/l of nitrate in the aquifer were classified as “1” if the chance of occurring nitrate was 70% or more, otherwise “0”. Same procedure was repeated for 95% or more chance of nitrate occurring in the aquifer. Simple map algebric function available in GIS was utilized to find areas probably affected by nitrates for chronic exceedance of contamination at 0.70 or more and 0.95 or more probabilities. Classified kriged blocks of probability maps for nine years (1997-2005) of nitrate data were added to the ArcGIS’s Spatial Analyst to determine the chronic exceedance maps of probabilities exceeding specified thresholds. A simple example has been demonstrated below on how ArcGIS’s spatial analyst can be used to determine chronic exceedance maps of monitoring nitrate data.



Grid 1 and 2 are the reclassified probability matrices for two sampling years where 1 represents occurrence of nitrate exceeding specified probabilities (0.70 and 0.95) and 0 represents none occurrence. The resulting output GIS grid matrix is the cumulative nitrate occurrence for specified probabilities - 0 for none occurring at all, 1 for occurring once in two sampling years at a location (i.e. 50% - chronic exceedance) and 2 for occurring twice in two sampling years at a location (i.e. 100% - chronic exceedance).

3.4 Logistic regression

Logistic regression has been used extensively in medical science since late 1960s to predict a dichotomous response from possible explanatory variables (Lemeshow et al., 1988) and is becoming more powerful statistical tool to solve environmental problems these days. Potential explanatory variables are important in predicting probability of groundwater nitrate concentrations greater than specified management threshold. In this research, threshold nitrate concentrations of 4 mg/l and 10 mg/l were chosen because the threshold of 4 mg/l of NO_3-N has been related to increased risk of Non-Hodgkins lymphoma (Ward et al., 1996) and has also been used for national assessment of nitrate in groundwater (Nolan et al., 2002). On the other hand, EPA has established 10 mg/l of NO_3-N as the maximum contaminant level (U.S. EPA, 1996) because elevated concentrations of nitrate in drinking water can cause low oxygen levels in the blood of infants, known as methemoglobinemia, a potentially fatal condition (Bosch et al., 1950; Comly, 1945).

Previous studies (Cain et al., 1989; Hay and Battaglin, 1990; Tesoriero and Voss, 1997; Greene et al., 2004; Scanlon et al., 2003; Gardner and Vogel, 2005) have shown that there is a significant relation between shallow groundwater nitrate concentration and the types of land cover around a sampled well. The logistic regression analysis precedes a hypothesis test (p-value) for each explanatory variable that determines whether the variable explains a significance amount of contamination probability for specified thresholds. The 2003 nitrate database was selected for this analysis because it was the most complete of all the years available.

3.4.1 Explanatory variables

As mentioned earlier, there is a significant relation between shallow groundwater nitrate concentration and the types of land cover around a sampled well. Land cover data for Cimarron terrace aquifer were downloaded from the National Land Cover Database (NLCD) surveyed in 2001. A total of 15 classes (Open water, Developed open space, Developed low intensity, Developed medium intensity, Developed high intensity, Barren land, Deciduous forest, Evergreen forest, Mixed forest, Shrubland, Grassland, Pasture/hay, Cultivated crop land, Palustrine forested wetland, and Estuarine forested wetland) that represent the Cimarron terrace aquifer were obtained. Open water and wetlands were added together as wetland, all types of developed lands were aggregated as **developed land**, all types of forest were combined as forest and then to wetland **(forest-wetland)**, barren and shrubland were added to **grassland,** and pasture/hay and cultivated cropland were combined together as **cropland-pasture**. These 15 classes were first aggregated to similar 4 classes (developed, forested-wetland, grassland, and cropland-pasture) because major 4 original classes (developed, forest, grassland, and cropland) cover more than 90% of the aquifer area. These final 4 classes were used as explanatory variables of logistic regression models to establish a relationship between land covers and nitrates in the aquifer. Aggregated land cover classes were developed as continuous variables with unit equal to percentage land covers.

Nitrogen loading and aquifer susceptibility to contamination have been previously studied to occurrence of elevated nitrate concentrations (Nolan, 2001, Tesoriero and Voss, 1997; Greene et al., 2004; Scanlon et al., 2003; Gardner and Vogel, 2005). Some aquifer susceptibility terms such as hydrological soil groups, percent well drained soils or combination hydrological soils A and B, and permeability of soils have been directly associated with the textures of surficial and sub-surficial geology. These variables are related to travel medium for agriculture nitrogen and hence should be addressed to understand the process of nitrogen leaching from agricultural areas to the aquifer. Potential variables such as atmospheric nitrogen deposition, population density, and rainfall were not included as explanatory variables because wellfields of Cimarron terrace aquifer are located in three counties and the countywide or larger spatial data of these variables were not helpful to explain the occurrence of elevated nitrate concentration in regression analysis.

Depth to seasonally high water table represents the unsaturated zone thickness and percent organic matter that represents denitrification potential in aquifer were also included as explanatory variables of logit models. Fertilizer N was apportioned equally to agricultural and developed land (urban) to account for residential fertilizer use. Average annual animal waste nitrogen in counties over Cimarron terrace aquifer ranges from 0.33 kg/ha to 1.85 kg/ha (Storm et al., 2000). This load was not applied as separate explanatory variable. The exact amount of fertilizer load in study area was not obtained. However, the fertilizer load of 110 kg/ha applied to wheat in Oklahoma (Storm et al., 2003) was used as the total nitrogen loading.

3.4.2 Database development

Land covers were derived from NLCD 2001 and aggregated to final four classes as explained earlier were included as explanatory variables of logistic regression models. The surficial and sub-surficial geology or soil fractions data set created by using the texture class information to estimate the percents of sand, silt, and clay in the fine (less than 2 mm) fraction of each layer of each component for each STATSGO map unit, and interpolating the results to a set of 11 standard layers (Fig. 3-4), and computing a weighted average of the values for all components of the map unit was derived from spatial database maintained by the Earth System Science Center (ESSC) at Pennsylvania State University (http://dbwww.essc.psu.edu/). The GIS was used to determine average percentages of sand, silt, and clay for each component for each STATSGO map unit from percentages at 11 standard layers. Similarly, an organic matter by weight was also derived from the same database.

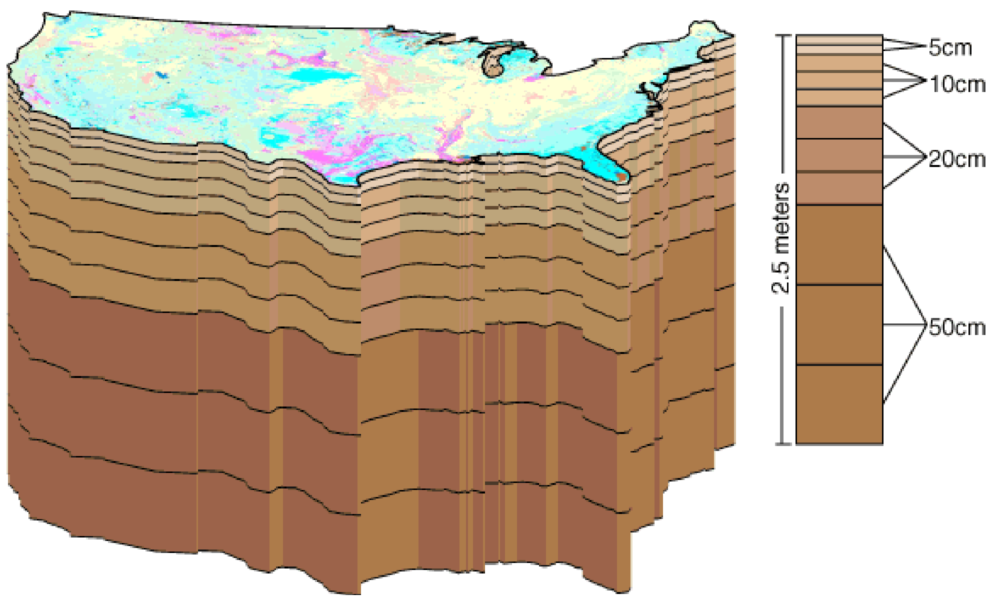


Figure 3-4. Soil fractions sampled at 11 layers (source: http://www.soilinfo.psu.edu)

After the aquifer level spatial data were added to the GIS software ArcMap (ESRI, 2006), GIS buffer analysis was created around each well to build a database for use in determining the best area of well influence. If the area of well influence was set too small, land cover characteristics were not reflected properly in the groundwater quality, and if the area of well influenced was set too large, unrelated land covers may appear to influence the groundwater quality. Various well buffers (100m, 250m, 500m, 750m, 1,000m, 1,500m, and 1,750m) were set to analyze the relationship between land cover and groundwater nitrate sampled at wells in 2003.

Percentage by land covers at different radial distances, as mentioned earlier, around each well location were extracted and examined to determine the best radius of well influence. The best fit logistic models were obtained for threshold 4 mg/l and 10 mg/l by finding the radius that maximized the likelihood ratio test (G-statistic) and Wald statistic. These statistics are well explained in subsections later in this chapter. The same radius of well influence that was determined for land covers was applied to all explanatory variables (soil fractions, organic matter, fertilizer N) that were the function of area. Explanatory variables such as depth to groundwater table and extracted values of spatial data for best radius of well influence and the response variable 'nitrate' were compiled together and uploaded to SAS (SAS, 2006) database for use in logistic regression analysis.

3.4.3 Logit model

To develop a logistic regression model for aforementioned thresholds, various variables were examined individually to determine if they were significant predictors. Univariate logistic regression models were applied to each of the explanatory variable to check whether the variable should be added to multivariate model. A total of 10 potential explanatory variables (land cover variables; percent of developed land, cropland-pasture, grassland, and forested-wetland, other variables; percent sand, percent silt , percent clay, organic matter percent by weight, nitrogen-fertilizer application rate and depth to water) were initially considered in univariate logistic regression models. At p-value of 0.10, significance level of 10%, Likelihood ratio test (G-statistic) and Wald statistic were examined to determine if the variable had significant effect on occurrence of nitrate at specified thresholds. The coefficient of determination (R^2) value along with Hosmer-Lemeshow p-value is also important criteria that describes how well the plot of observed versus predicted value of deciles of risk fit the line 1:1 (Nolan, 2002). This criterion was also examined to upgrade variable from univariate to multivariate models. Then statistically significant explanatory variables from univariate logistic regression models were used in stepwise logistic regression models to build final multivariate logistic regression models.

The logit of the multiple logistic regression model is given by the equation (Hosmer and Lemeshow, 2000):

$$G\left(x\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots\dots\dots\dots\dots\dots + \beta_n x_n \qquad (3\text{-}8)$$

Logistic regression is used when the dependent variable is dichotomous (presence/absence, above/below, yes/no).

The odds ratio is the probability of the event exceeding a threshold value, divided by the probability of the event not exceeding the threshold value (Helsel and Hirch, 2002)

$$Odds \;\; ratio = \frac{p}{1-p} \qquad (3\text{-}9)$$

Where, p= probability of exceeding the threshold value. The log of the odds ratio, logit, transforms a variable constrained between 0 and 1.

The logit can then be modeled as a linear function of one or more explanatory variables to produce logistic regression. Now, transforming equation (3-9) to logit and combining to equation (3-8)

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots\ldots\ldots + \beta_n x_n \tag{3-10}$$

Thus, the odds ratio is modeled as:

$$\left(\frac{p}{1-p}\right) = e^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots\ldots\ldots\ldots\ldots+\beta_n x_n)} \tag{3-11}$$

To convert the estimated values of the response variable back to original units, the logistic transformation, the inverse of logit transformation was used:

$$p = \frac{e^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots\ldots\ldots\ldots\ldots+\beta_n x_n)}}{1+e^{(\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots\ldots\ldots\ldots\ldots+\beta_n x_n)}} \tag{3-12}$$

Where β_0 is intercept and $X = x_1, x_2, \ldots\ldots\ldots, x_n$ explains variables with corresponding slopes $\beta = \beta_1, \beta_2, \ldots\ldots\ldots\ldots, \beta_n$, or $\beta_i = \beta_{0,}\ and\ \beta_1, \beta_2, \ldots\ldots\ldots\ldots\ldots, \beta_n$.

3.4.4 Significance testing of β_i coefficients

Maximum Likelihood Estimation, MLE, is the preferred method used to test the significance of logit coefficients (Hosmer and Lemeshow, 2000). MLE considers the null hypothesis (H_0) indicating that the logistic regression coefficients, β_i ($\beta_{0,}\ and\ \beta_1, \beta_2, \ldots\ldots\ldots\ldots\ldots, \beta_n$), are not significantly different from zero. The idea behind MLE is to maximize the log likelihood (LL) by comparing the ratio of the maximum of the likelihood under H_0 to the maximum of the likelihood under H_a, the likelihood ratio test (Greene et al., 2004). This phenomenon reflects how likely (the odds) that the observed values of the dependent variable are estimated from the observed values of the explanatory variables. Likelihood is a conditional probability (e.g. P (Y|X), the probability of Y given X) and it varies from 0 to 1 like a probability.

3.4.5 Likelihood ratio test

The likelihood ratio test (G-statistic) tests the statistical significance of logit coefficients of logistic regression model (Hosmer and Lemeshow, 2000) and can be expressed as

$$G = -2\ln\left[\frac{Likelihood\ without\ the\ variable}{Likelihood\ with\ the\ variable}\right] \tag{3-13}$$

The G-statistic is chi square distributed under the hypothesis that β_i coefficients ($\beta_{0,}$ and $\beta_1, \beta_2, \ldots\ldots\ldots\ldots, \beta_n$) are equal to zero. The G-statistic compares estimated values with observed values of the response variable with and without different explanatory variables.

3.4.6 Wald statistic

Wald statistic was used as an alternative test to evaluate the significance of each of the explanatory variable for each logistic regression model. The Wald statistic was obtained for each of the β coefficient of the model. The statistic was calculated as the maximum likelihood estimate of the slope coefficients, β, to an estimate of its standard error (Hosmer and Lemeshow, 2000) and can be expressed as

$$W = \left[\frac{\hat{\beta}_i}{standard\ error\ \hat{\beta}_i}\right] \tag{3-14}$$

for i = 0, 1, 2…….k

3.4.7 Model goodness-of-fit

Wald and G-statistics were used to test the significance of each explanatory variable and finally then to build an optimal model. After the model was fitted, a global test of goodness-of-fit of the resulting model was performed using the Hosmer-Lemeshow (H-L) goodness-of-fit statistic ($\widehat{C}$) to see how well the model fit the data. The H-L test calculates probability values from the chi square distribution to test the fit of the model. The null hypothesis is that the model fits the data indicating higher p-values. The test divides the predicted probabilities into deciles of risks, generally 10 groups based on percentile rank and then compares a Pearson chi square from the 2x10 table of observed and expected estimated frequencies. The H-L test is expressed as (Hosmer and Lemeshow, 2000):

$$\widehat{C} = \sum_{k=1}^{10} \frac{\left(o_k - n_k \pi_k\right)}{n_k \pi_k \left(1 - \pi_k\right)} \tag{3-15}$$

Where n_k = the number of observations in the k^th decile,

o_k = the number of successes (Events exceeding thrersold) in the k^th decile,

π_k = the average of the estimated probabilities, and

$\widehat{C}$ = test statistic, approximated by the chi-square distribution with g-2 degree of freedom

The goodness-of-fit were also evaluated using coefficient of determination (R^2) value. The R^2 value is an indication of how well the plot of observed versus predicted values of deciles of risk fit the line 1:1. Linear regressions were plotted between predicted probabilities for deciles of risk used to calculate the H-L statistic versus observed probabilities of elevated nitrate concentrations for thresholds of 4 mg/l and 10 mg/l of NO_3-N.

3.4.8 Variable selection approach

A stepwise logistic regression technique was used to develop final multivariate logistic regression models (for 4 mg/l and 10 mg/l) from explanatory variables that were significant predictors at univariate models. SAS 9.3 statistical software (SAS, 2006) was used to perform logistic regression analysis. The stepwise logistic regression model starts with only the intercept and explanatory variables are added to the model one at a time. It was hypothesized that if the associated variable was significant at α= 0.3 level of significance; it was entered in the model. Hosmer and Lemeshow (2000) described the significance level of 0.05 as too stringent, often excluding important variables from the model. Hence, they propose to use the range from 0.15 to 0.25 and even 0.30. One or more entered variables were then tested for analysis for effects eligible for removal by using Wald chi-square . Variables those were not significant at α = 0.15 level of significance were removed from the model. The process stopped when variables did not meet aforementioned significance level of entry.

3.5 Stochastic modeling

Estimates of the major sources of uncertainty in predicting probability of nitrate concentration exceeding a threshold are useful for land managers to build aquifer protection strategies. Monte Carlo Simulation is widely accepted stochastic model

(Kaplan and McTernan, 1993) for propagating uncertainty of risk estimates associated with parameters (Doubilet et al., 1995; Critchfield and Willard, 1986; Thompson et al., 1992). Monte Carlo simulation understands complex stochastic systems and hence addresses the inherent variability of natural phenomena.

3.5.1 Monte Carlo simulation

The objective of using Monte Carlo simulation was to address the uncertainty associated with significant variables of final logit models (for exceeding 4 mg/l and 10 mg/l of NO_3-N) and the model output itself. Distributions to significant explanatory variables of logit model were fitted using @ Risk 4.5 software (Palisade, 2005). Monte Carlo simulation was applied to final logit models using cumulative density functions of fitted distribution parameters. This simulation algorithm is based on random draws with replacement from predefined statistical distributions. Essentially, the Monte Carlo simulation method is a method for evaluating an integral (Fishman, 1996)

$$\psi = E_{\pi}\left\{U\left(X\right)\right\} = \int U\left(x\right)\pi\left(x\right)dx \qquad (3\text{-}16)$$

where $E_{\pi}\{\ \}$ is the expectation, the probability of nitrate concentration exceeding threshold with respect to the probability density π. $U(\ )$ is a response function representing the logit model, and $\pi(x)$ represents the vector of all cumulative densities of the significant variables of the logit model. It involves random draws $X = x^{(j)}$ from the target distribution π.

The following are steps involved in the application of the Monte Carlo techniques (McTernan and Bonnett, 2002):

1. Select the appropriate cumulative probability distribution function for describing uncertainty in the significant explanatory variable(s) of logit model.
2. Select a random number from the distribution and use this as input to the model.
3. Run the model using the random number taken from the input distribution to calculate the output.
4. Repeat steps 2 and 3 for number (n) times.
5. Determine the cumulative probability distribution function of the output step 3.
6. Analyze the output distribution and utilize the statistics (i. e. mean and upper bound 95% confidence interval).

CHAPTER 4

RESULTS

This chapter is organized so as to provide results of all methods explained in Chapter 3. Results include analysis of water quality parameter selected, outputs from geostatistics and GIS, regression analysis, and stochastic modeling.

4.1 Results from analysis of water quality parameter selected

The mean and median NO_3-N concentrations for the entire study period (1997-2005) were 7.02 mg/l and 6.8 mg/l respectively. The 25th and 75th percentile values were 4.3 mg/l and 9.1 mg/l. Of the 821 samples of four wellfields in the study area, 21.3% had nitrate-N value less than 4 mg/l, 30.6% between 4 mg/l to 7 mg/l, 29.7% between 7 mg/l to 10 mg/l, and 18.4 % greater or equal than 10 mg/l which is the drinking water criteria set by U.S. EPA.

Nitrates in four wellfields in Cimarron terrace aquifer were analyzed separately. Average nitrate concentrations increased from 1997 to 1999 in all wellfields (Figs. 4-1 through 4-4), decreased in 2000, and again increased after 2000. This indicates that there were trends in nitrate concentrations over that study period. Time series plots of nitrate concentrations in all wellfields include connecting line of means, medians (horizontal lines), interquartile range boxes (25th and 75th percentiles), and outliers as asterisks.

Means largely varied from medians in Cleo Spring wellfield from the year 1997 to 2005. All wellfields data were combined together to examine the effect of nitrate in whole aquifer. For this, time series and box plots were drawn (Fig 4-5). The average nitrate in the aquifer showed a increasing trend from 1997 to 1999, decreasing trend from 1999 to 2000, again increasing trend from 2000 to 2002 and fairly constant after 2002.

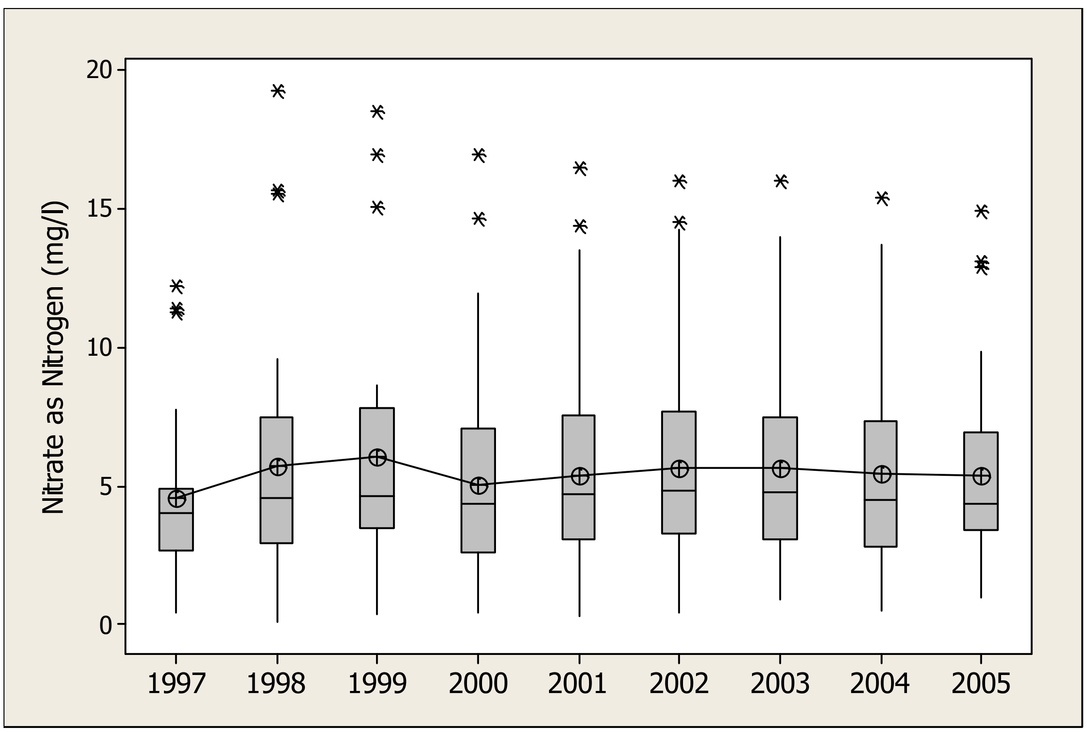


Figure 4-1. Time series, interquartile range, and median NO_3-N in Cleo Spring

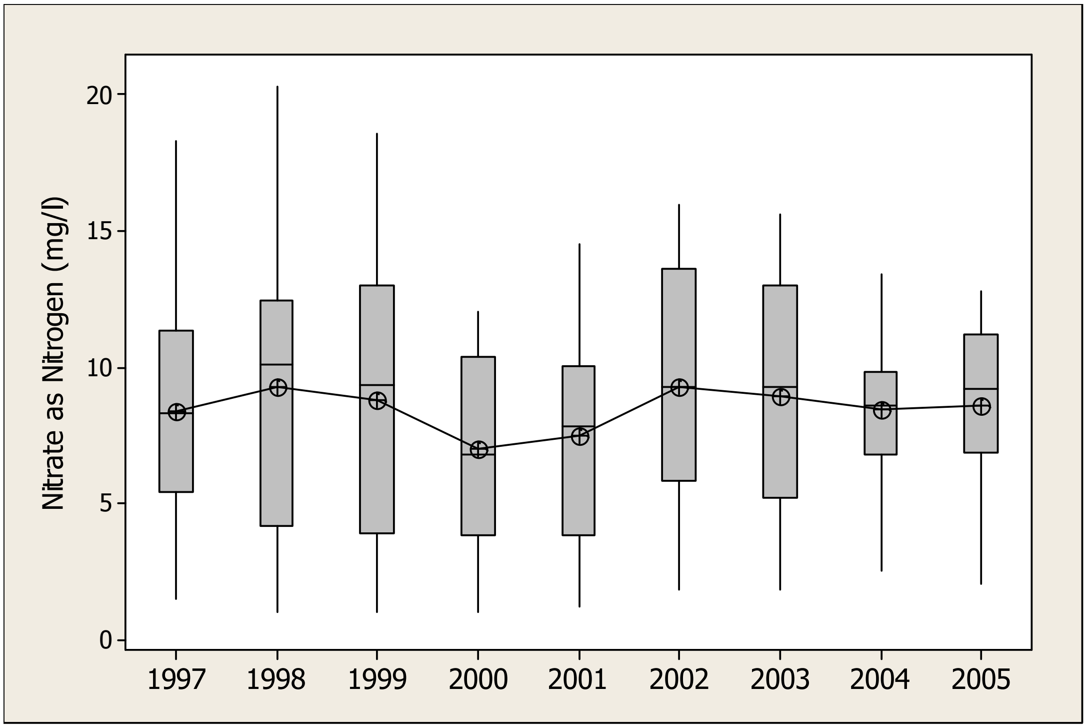


Figure 4-2. Time series, interquartile range, and median NO_3-N in Ames

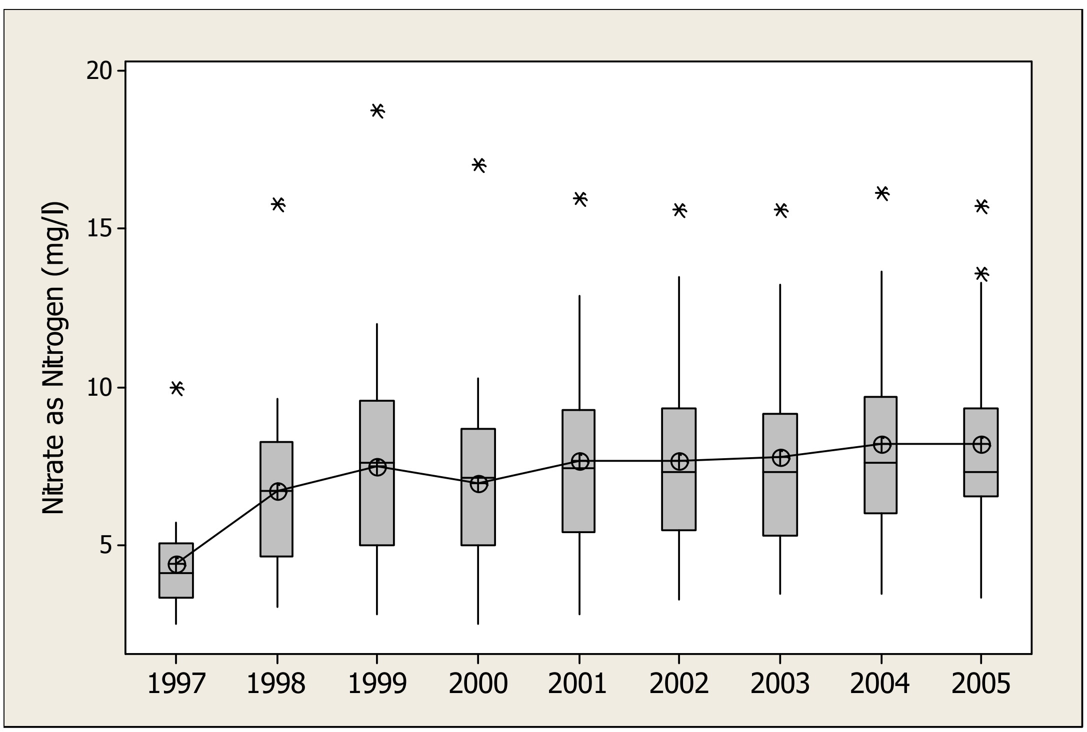


Figure 4-3. Time series, interquartile range, and median NO_3-N in Ringwood

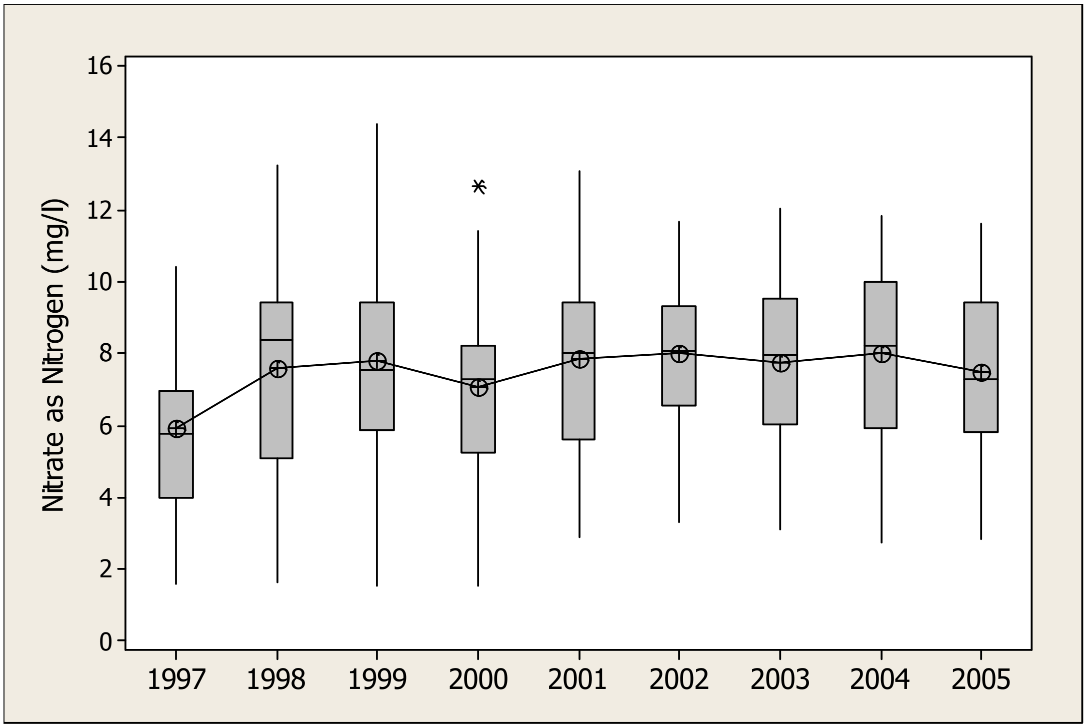


Figure 4-4. Time series, interquartile range, and median NO_3-N in Drummond

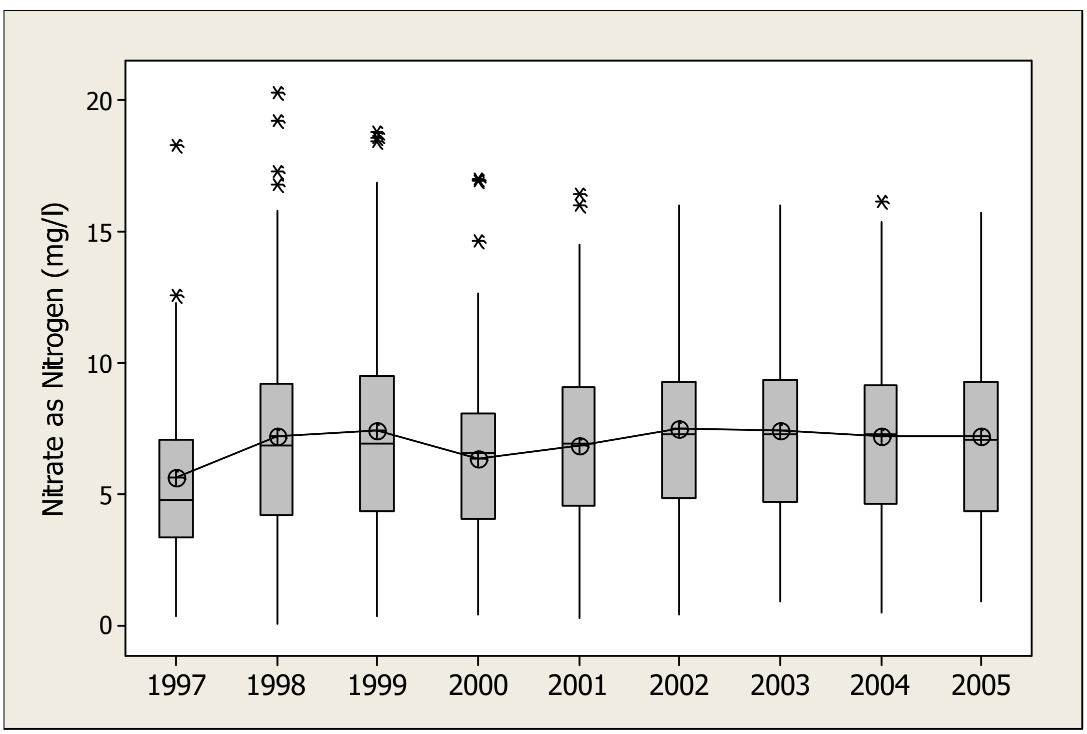


Figure 4-5. Time series, interquartile range, and median NO_3-N in the Cimarron terrace aquifer

Statistical summaries (box plots) of nitrate as nitrogen concentrations for Cleo Spring, Ames, Ringwood, and Drummond wellfields were compared in figure 4-6 below. Mean and median (50th percentile) concentration of nitrate in Ames wellfield were found to be higher than that in other three wellfields. Statistical difference of means from medians in four wellfields was evaluated by plotting 95% confidence interval boxes within the interquartile range boxes (Fig. 4-6). The mean and median nitrates in Ames, Ringwood, and Drummond wellfields were within the 95% confidence interval, indicating means were statistically same to medians for the entire study period. The mean and median nitrates in Cleo Spring were not within the 95% confidence interval, indicating statistical difference between mean and median.

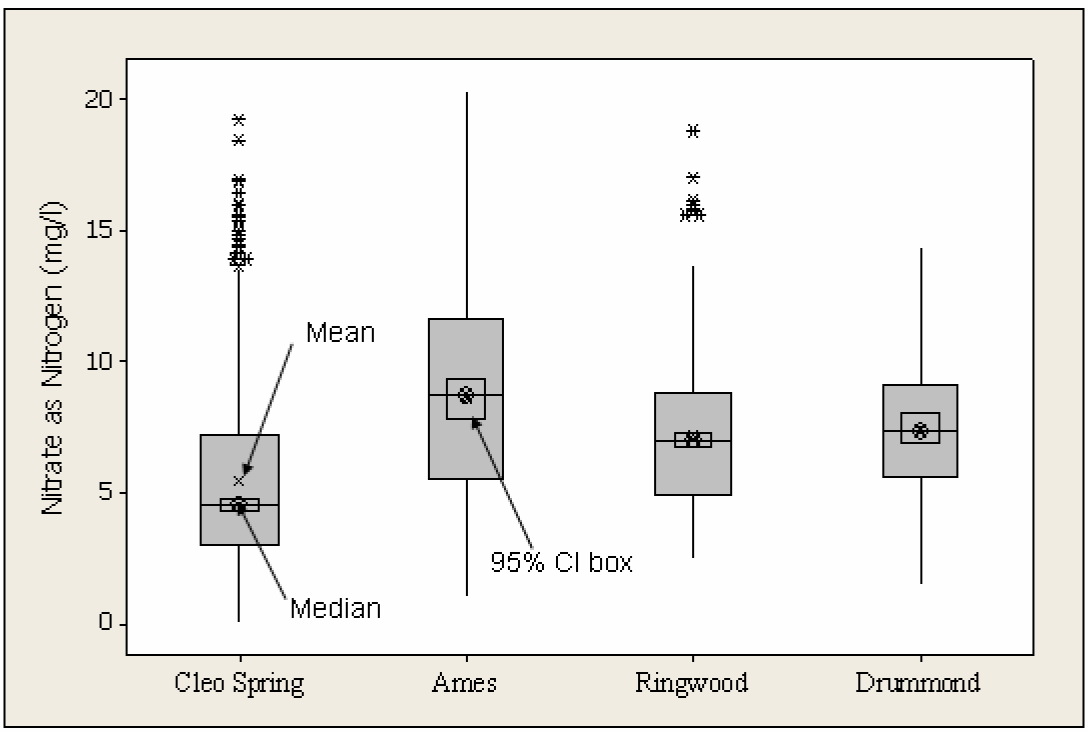


Figure 4-6. Statistical summary (box plots) of nitrate values in four wellfields

4.2 Geostatistical analysis

This section includes outputs of ordinary kriging, spatial-temporal trend analysis, and probabilistic assessment of nitrate using indicator kriging and GIS.

4.2.1 Ordinary kriging analysis

Semi-variograms were generated for nitrate concentration data from the sampling year 1997 to 2005 to quantify the spatial distribution. Trial and error processes were used to best fit the semi-variogram structure which included change of semi-variogram parameters such as ranges, sills, nuggets, and anisotropy. These parameters including fitted function are shown in table 4-1. This semi-variogram fitting process was observed whether this minimized the square differences between empirical semi-variogram values and the theoretical model. Various statistical criteria such as Mean Error (ME) and Root Mean Square-standardized Error (RMSE) were selected for cross-validation of variogram models. Cross- validation results are presented in table 4-2. These variogram models that provided the best cross-validation results were used in the estimation of groundwater nitrate concentrations at unsampled locations using ordinary kriging.

Table 4-1. Directional Semi-variogram model parameters

Year	Model	Range(m)	Direction	Partial Sill	Nuggets	Lag size (m)	No. of lags
1997	Spherical	4420.7	300	9.869	2.36	381.10	10
1998	Spherical	5666.5	337.3	15.651	7.7344	239.27	15
1999	Spherical	5753.0	344.7	14.56	7.93	302.41	12
2000	Spherical	4924.4	359.7	6.7068	6.6896	302.41	14
2001	Spherical	3807.3	350.8	9.1337	6.1638	235.05	16
2002	Spherical	4357.3	342.9	8.998	5.98	269.56	16
2003	Spherical	4409.1	342.9	7.9153	6.578	306.74	14
2004	Spherical	4841.8	345.2	8.992	6.3327	203.92	12
2005	Gaussian	1563.1	338.1	8.99	4.6	144.07	12

Table 4-2. Cross-validation results

	Cross Validation Criteria		
Year	ME	RMSE	Sample size
1997	-0.036	0.990	85
1998	-0.051	0.953	102
1999	-0.066	0.971	95
2000	-0.059	0.977	93
2001	-0.087	0.955	88
2002	-0.032	0.925	98
2003	-0.034	0.931	98
2004	-0.059	0.923	80
2005	-0.088	1.016	82

The spatial distribution of estimated nitrate concentrations obtained from kriging model for years 1997 and 2005 are shown in figures 4-7 and 4-8. Kriging estimates of nitrate as nitrogen for study period 1998-2004 can be found in Appendix B. The prediction maps show that there were high nitrate levels in the east region of the aquifer, especially eastward of Ames wellfield. This wellfield contained highest nitrate levels among four wellfields and the trend of nitrate was generally increasing from west to east. Standard error maps were also produced from ordinary kriging. Error maps for 1997 and 2005 are shown in figures 4-9 and 4-10. These maps indicated that errors largely varied from 1 mg/l to 5 mg/l of nitrate in the aquifer except where all wellfields were located because sampling wells were clustered in small areas and wellfields were located only in the central part of the Cimarron terrace aquifer.

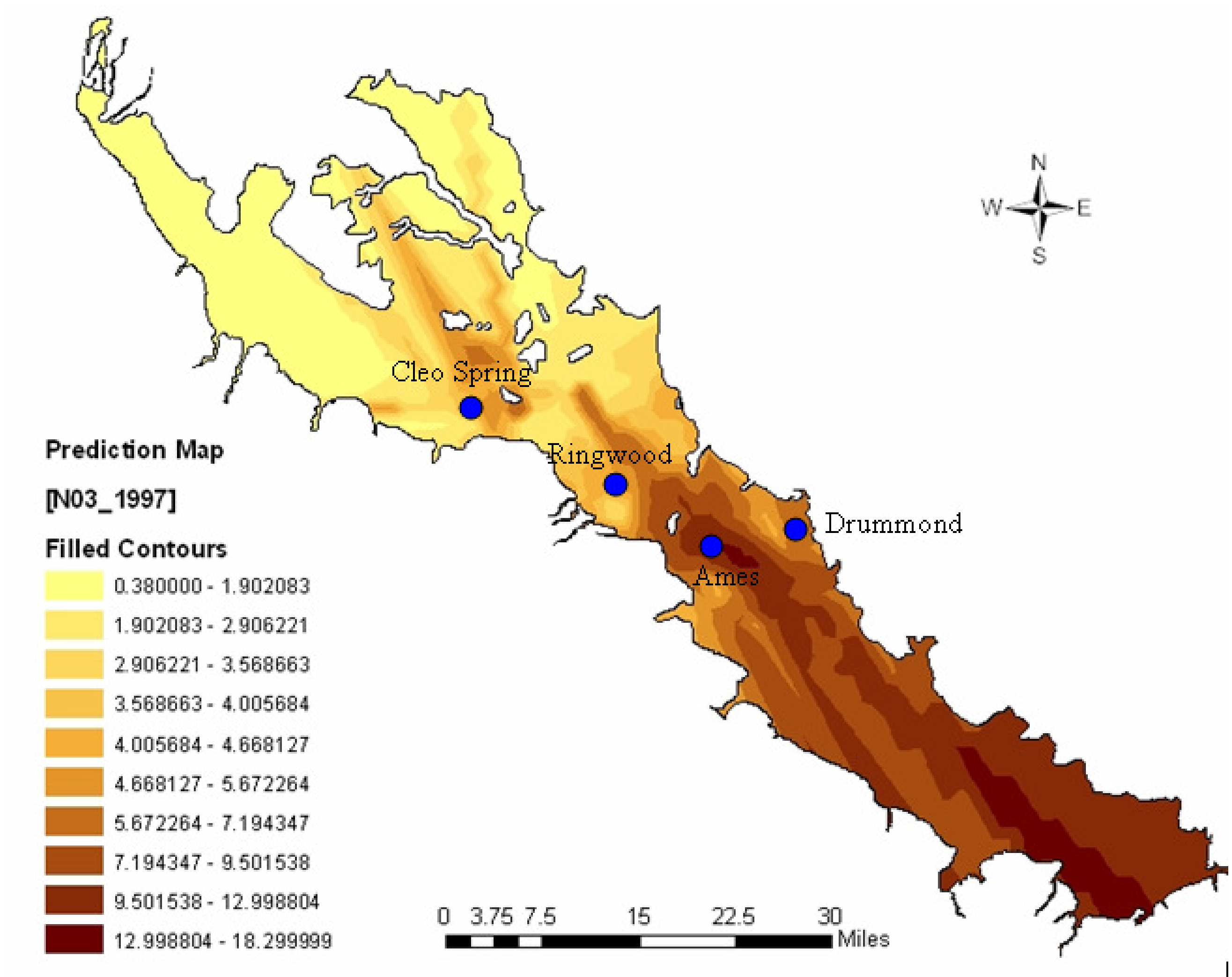


Figure 4-7. Kriging estimation of nitrate as nitrogen (mg/l) for the year 1997 in Cimarron terrace aquifer

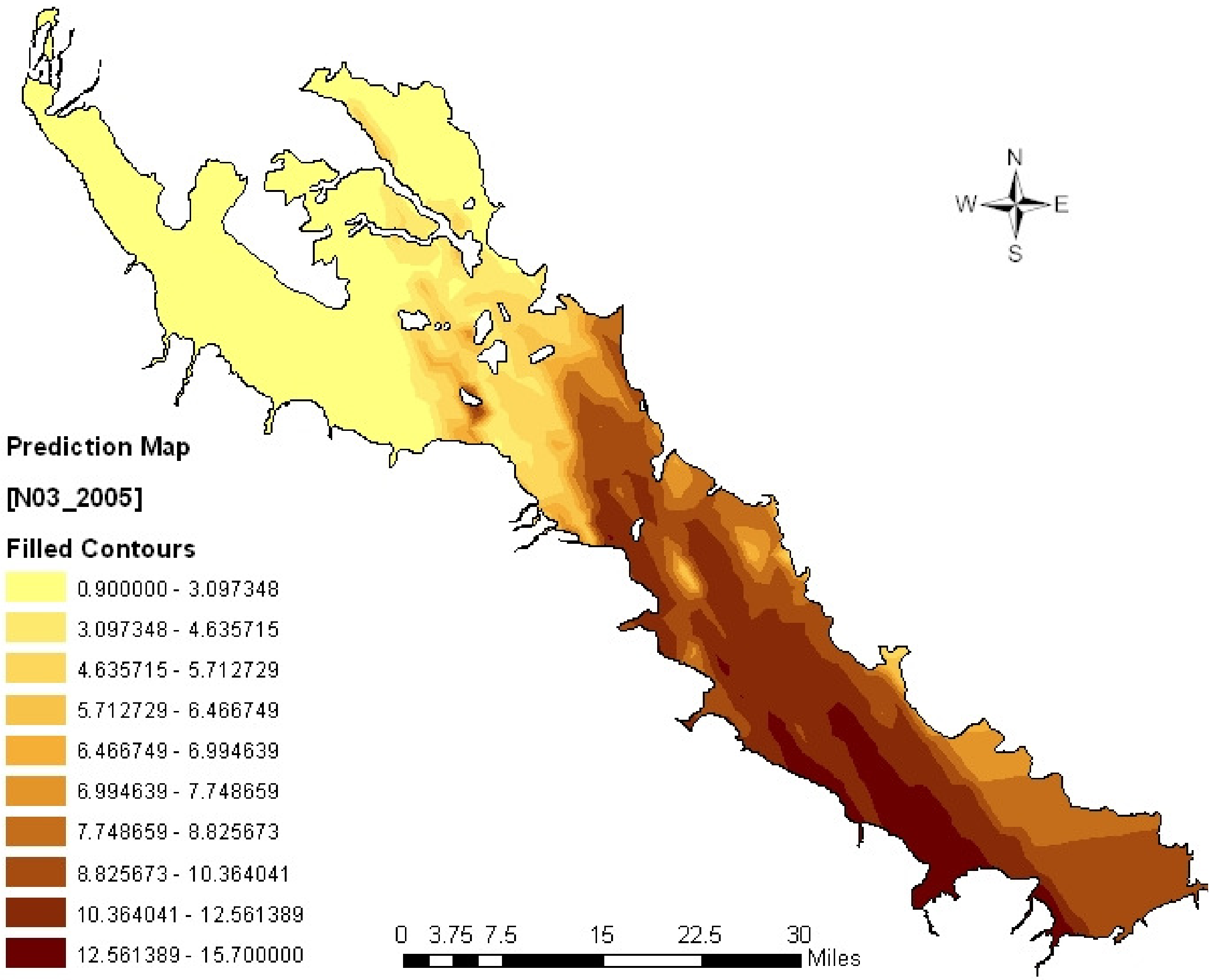


Figure 4-8. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2005 in Cimarron terrace aquifer

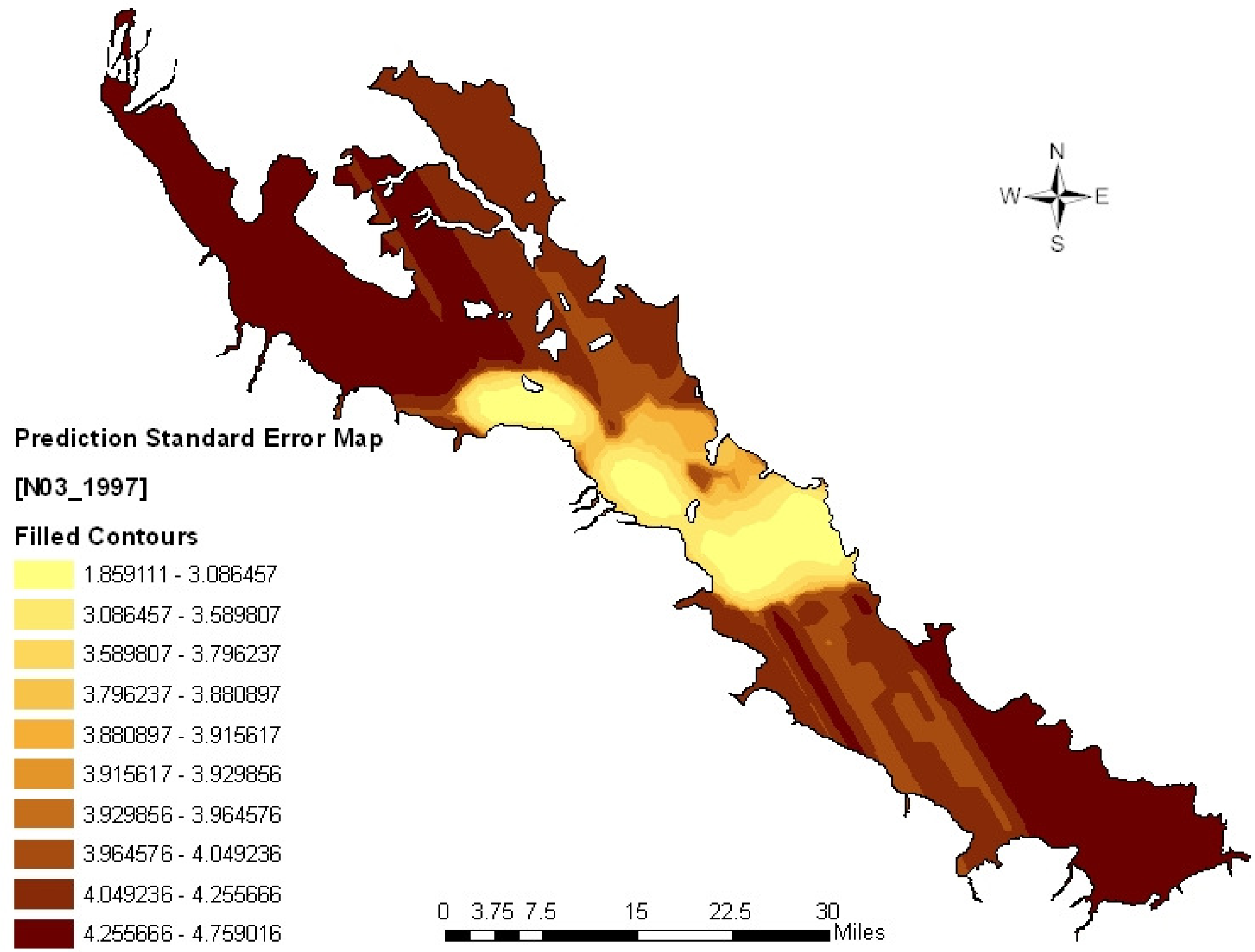


Figure 4-9. Kriging standard error map of nitrate as nitrogen (mg/l) for the year 1997 in Cimarron terrace aquifer

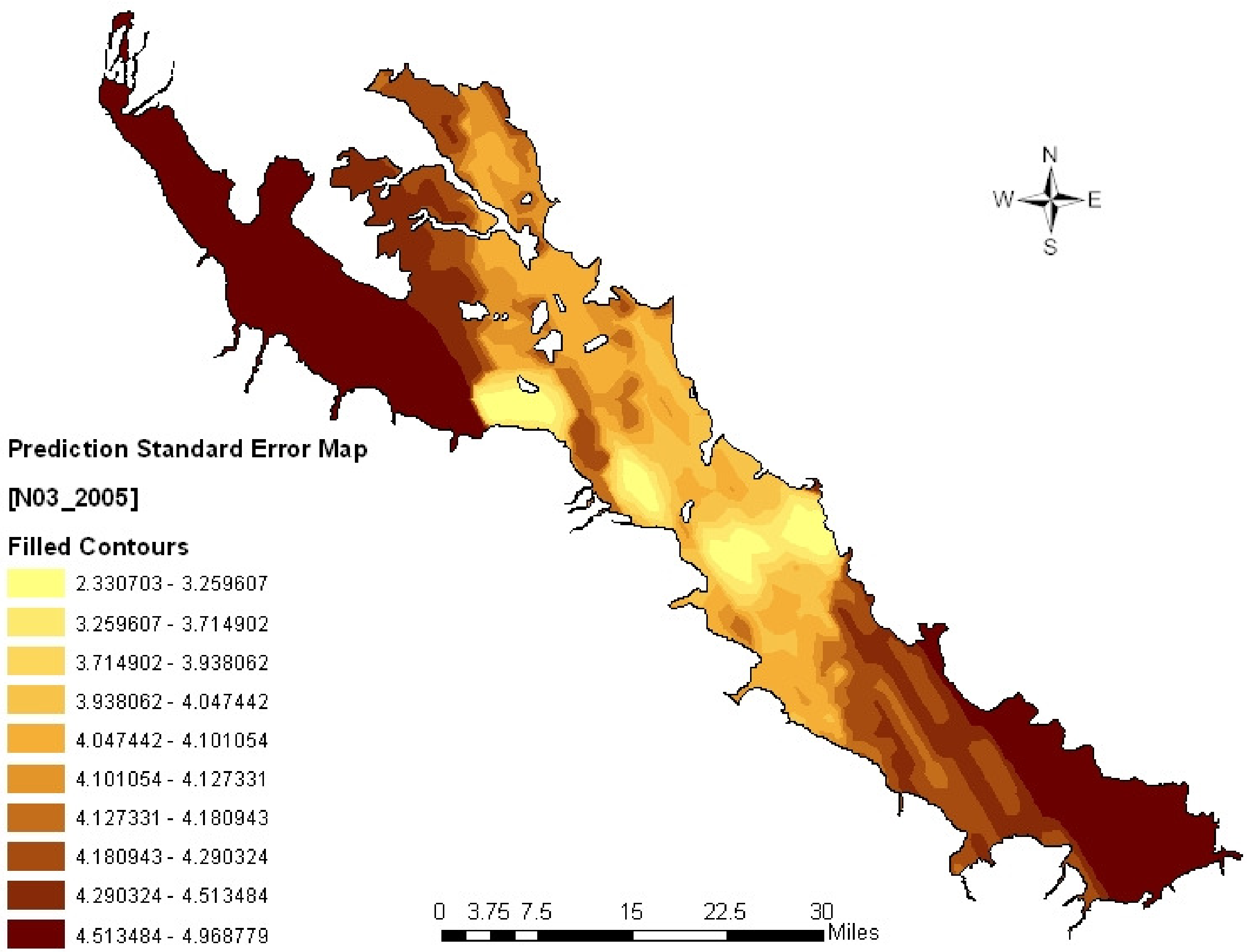


Figure 4-10. Kriging standard error map of nitrate as nitrogen (mg/l) for the year 2005 in Cimarron terrace aquifer

Areas of nitrate exceeding drinking water criteria in the aquifer were mapped from kriging estimation. The maximum area affected by nitrate was 444 square miles in 1998 while minimum area affected by nitrate was 54 square miles in 2000. Areas exceeding drinking water criteria (10 mg/l) of nitrate in the aquifer are shown in table 4-3. These area-wide maps of exceeding drinking water criteria are attached in Appendix C.

Table 4-3. Areas of nitrate concentrations exceeding 10 mg/l of nitrate as nitrogen

Years	Area of Nitrate Exceeding 10 mg/l (square miles)
1997	254
1998	444
1999	330
2000	54
2001	235
2002	394
2003	277
2004	396
2005	309

4.2.2 Spatial-temporal trend analysis

The spatial-temporal trend analysis was studied using kriged variance blocks in GIS. Using a global null hypothesis of a one tailed comparison test at 95% and 70% confidence intervals, areas defining statistically significant changes in nitrate between 1997 and 2005 were mapped in the Cimarron terrace aquifer (Fig 4-11 and 4-12). The negative sign indicates an increased nitrate levels while positive sign indicates a decreased. Areas of simple difference of nitrates were also mapped from kriging estimation (Fig. 4-13) and this map was compared with statistically significant nitrate change maps. A few differences between these two maps were observed in the central part of the aquifer. Areas of temporal trend were further studied as geographic areas of statistically increasing and decreasing trends in the aquifer from 1997 to 2005. It was observed that the area of increasing trend of nitrates from 1997 to 2005 at 95% and 70% confidence intervals is greater than half of the area of the aquifer. This area is depicted in figure 4-14 and 4-15. Similarly, statistical areas of nitrate that has been increased by at least 4 mg/l from 1997 to 2005 were also mapped. The area accounted for 12.11% of the aquifer at 95% confidence (Fig. 4-16) while at 70% confidence it was 12.43 % (Fig. 4-17). The result revealed that nitrates did not increase by 10 mg/l in the aquifer from 1997 to 2005.

Two levels of significance; 95% and 70%, were used to analyze the spatial-temporal trend. Abovementioned areas of statistically significant changes were quantified by hypothesizing a global null hypothesis of "two geographic means are equal" (two sided test). Using the equation for two sided test of two kriged mean blocks for 1997 and 2005, the map of Z_α, the amount of distribution falling in normal distribution, was obtained (Fig. 4-18). Based on α-values, pre-determined acceptance levels (0 to 1), a map of Z_α was reclassified as equal to or less than the lower confidence tail, -1.96 and -1.04 of Z_α at confidences 95% and 70% respectively, and equal to or greater than the upper confidence tail, 1.96 and 1.04 of Z_α at confidences 95% and 70% respectively in ArcGIS's Spatial Analyst. These classified areas were termed as statistically significant changes in nitrate at given thresholds. At 95% and 70% significances, areas of statistically significant changes in nitrate are shown in figures 4-19 and 4-20. At higher confidence, 95%, smaller area, only 0.4% of aquifer area was found while for 70% confidence, the area of nitrate change was 7.5% of the aquifer area.

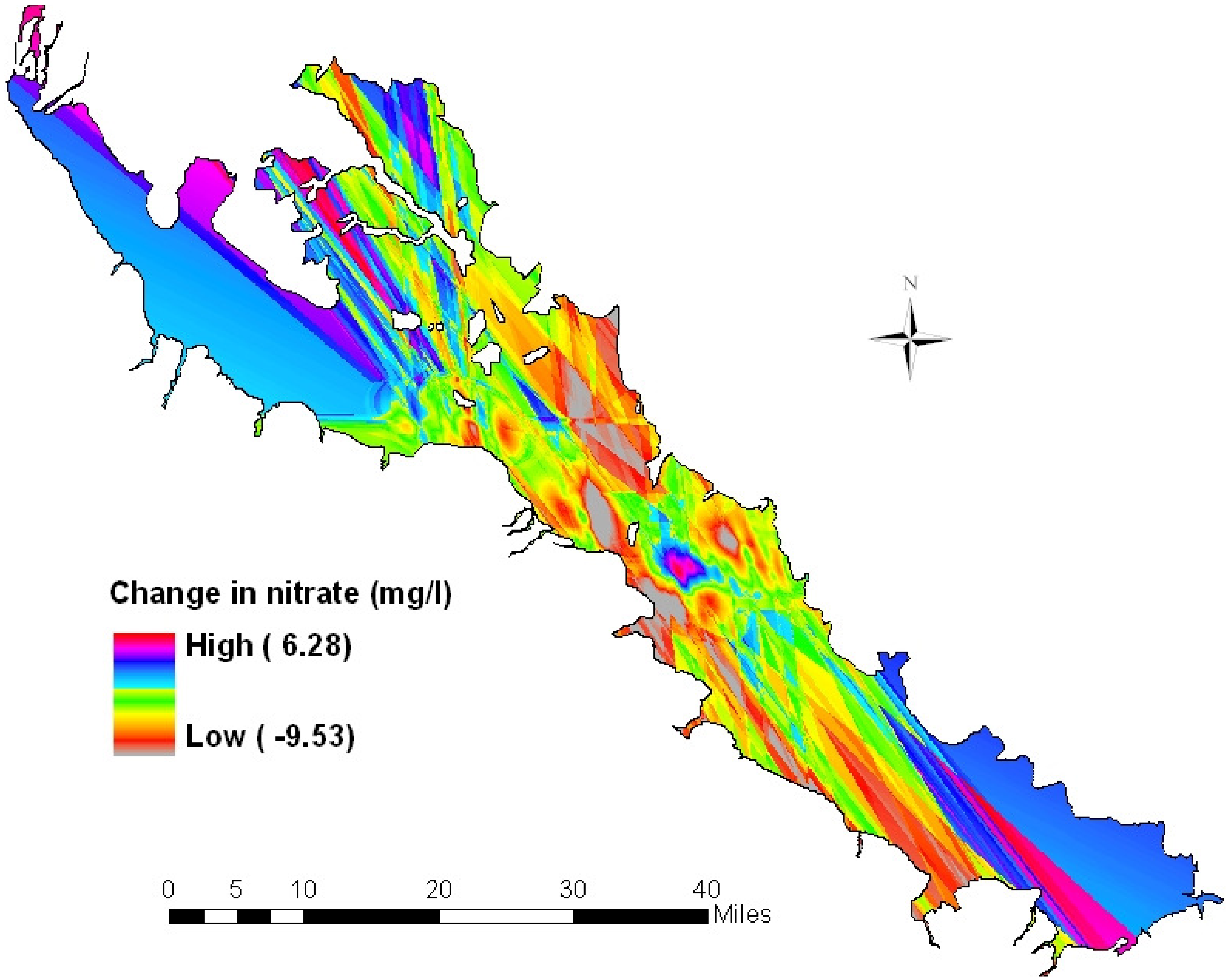


Figure 4-11. Areas defining statistically significant changes in nitrate (mg/l) 1997 – 2005 at 95 % confidence

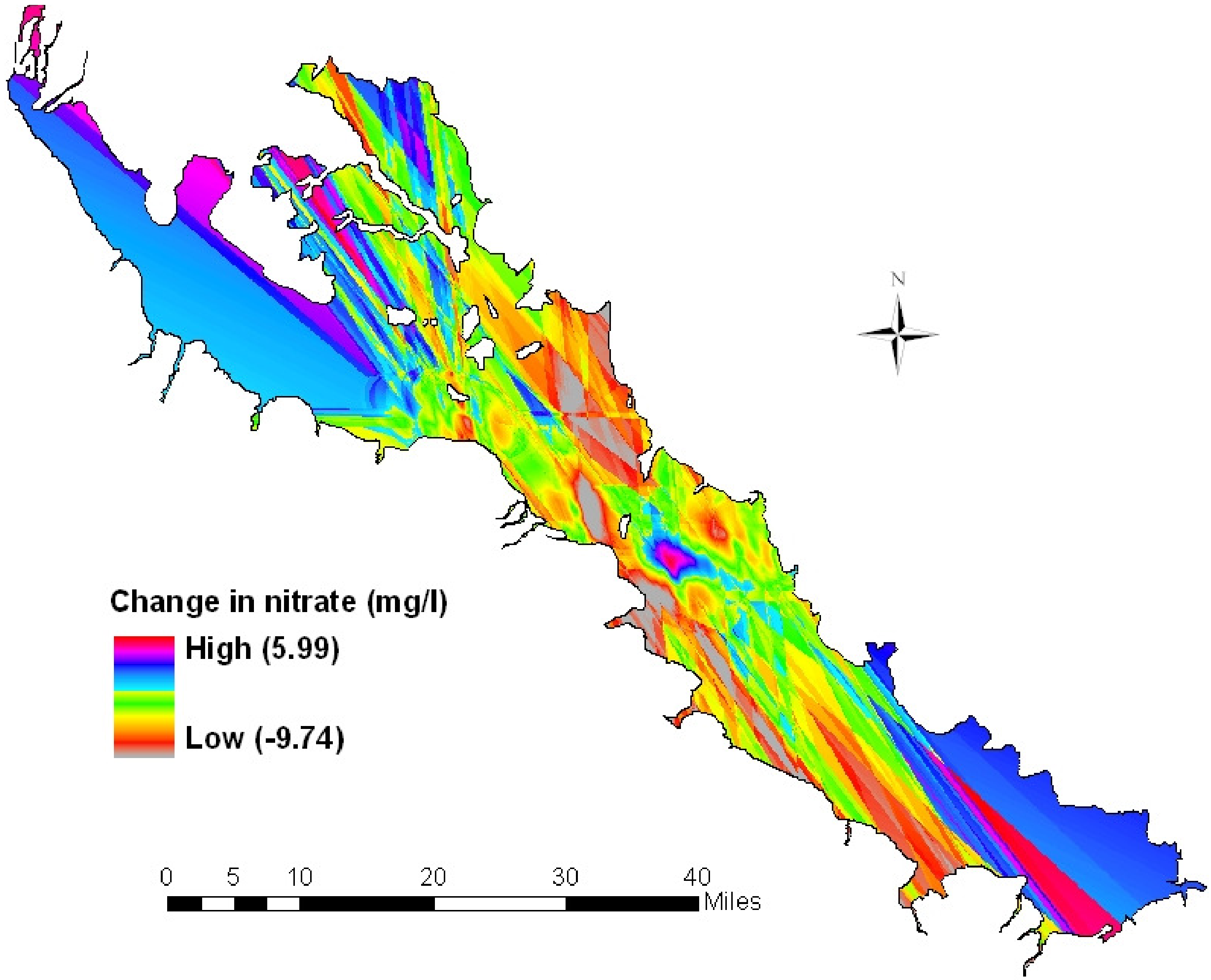


Figure 4-12. Areas defining statistically significant changes in nitrate (mg/l) 1997 – 2005 at 70 % confidence

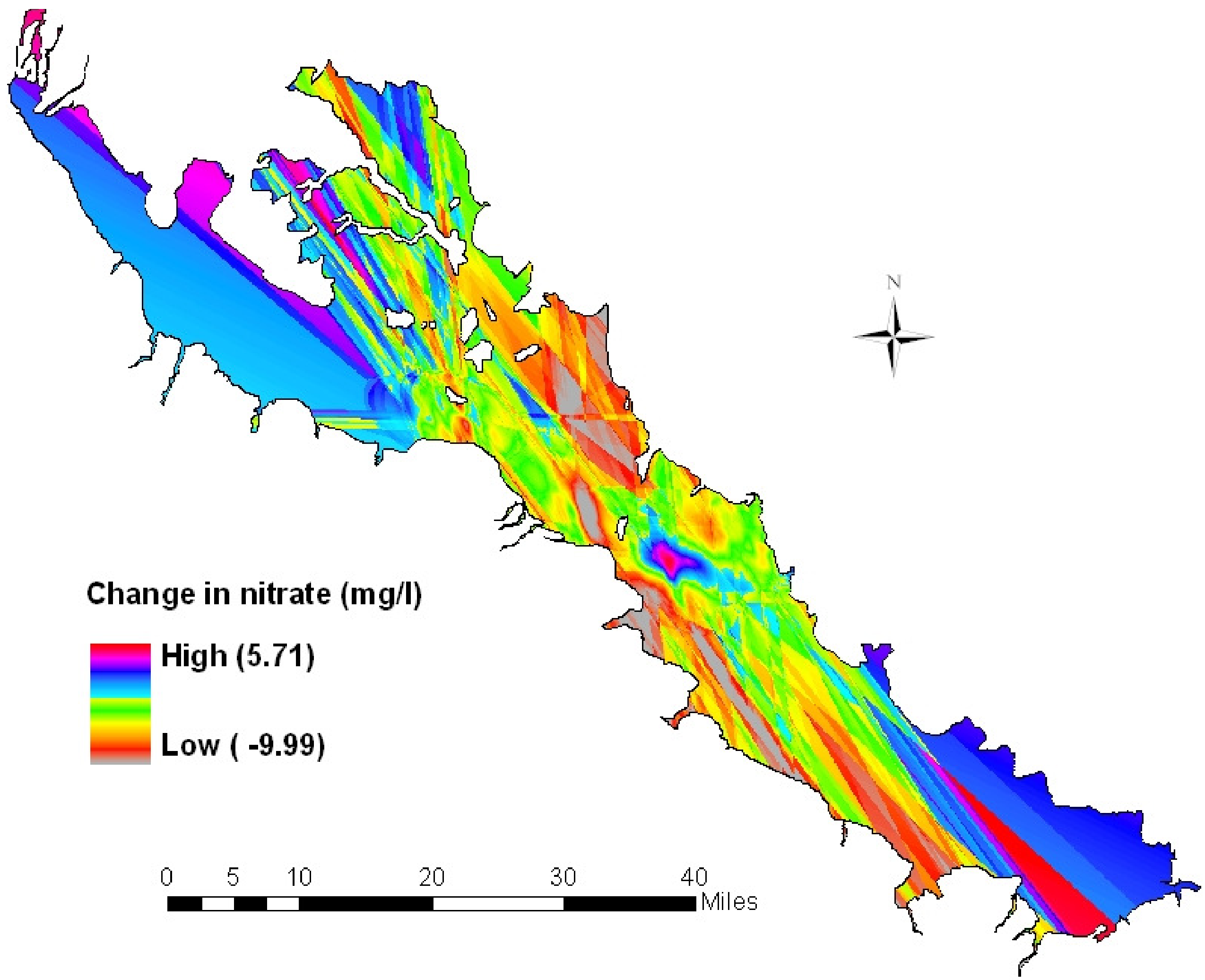


Figure 4-13. Areas of nitrate changes (simple difference, mg/l) 1997 - 2005

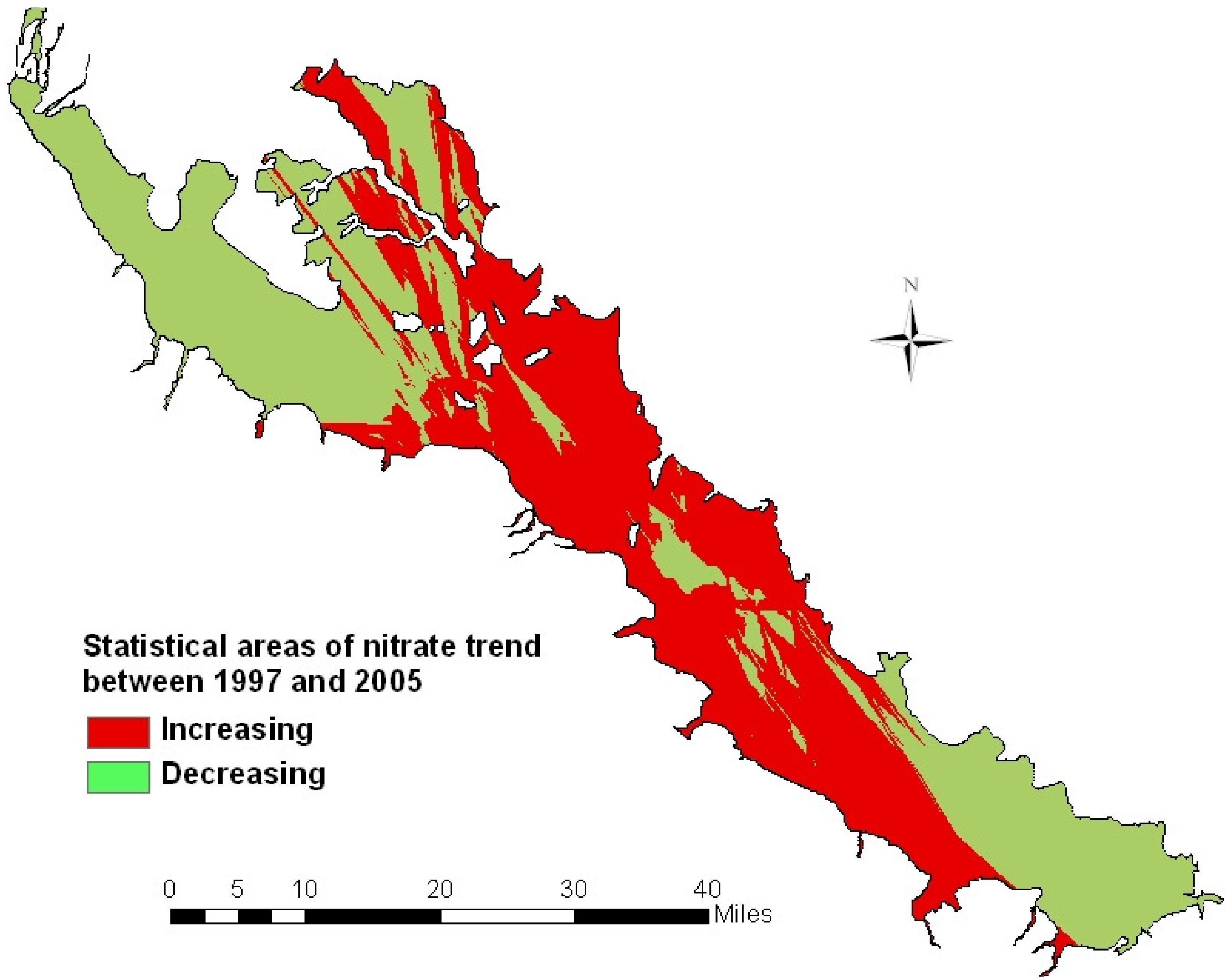


Figure 4-14. Statistical areas of nitrate trend 1997 - 2005 at 95 % confidence

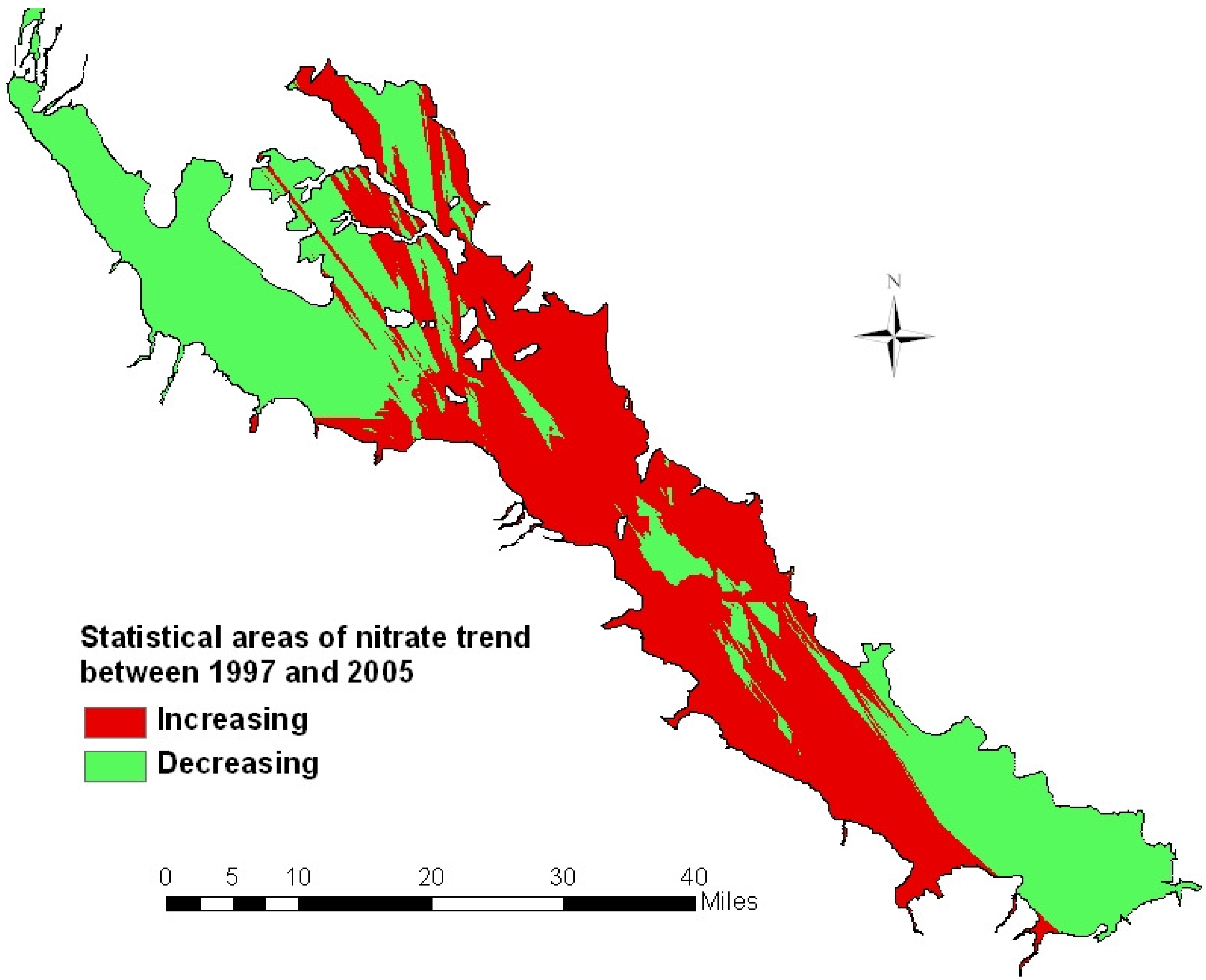


Figure 4-15. Statistical areas of nitrate trend 1997 - 2005 at 70 % confidence

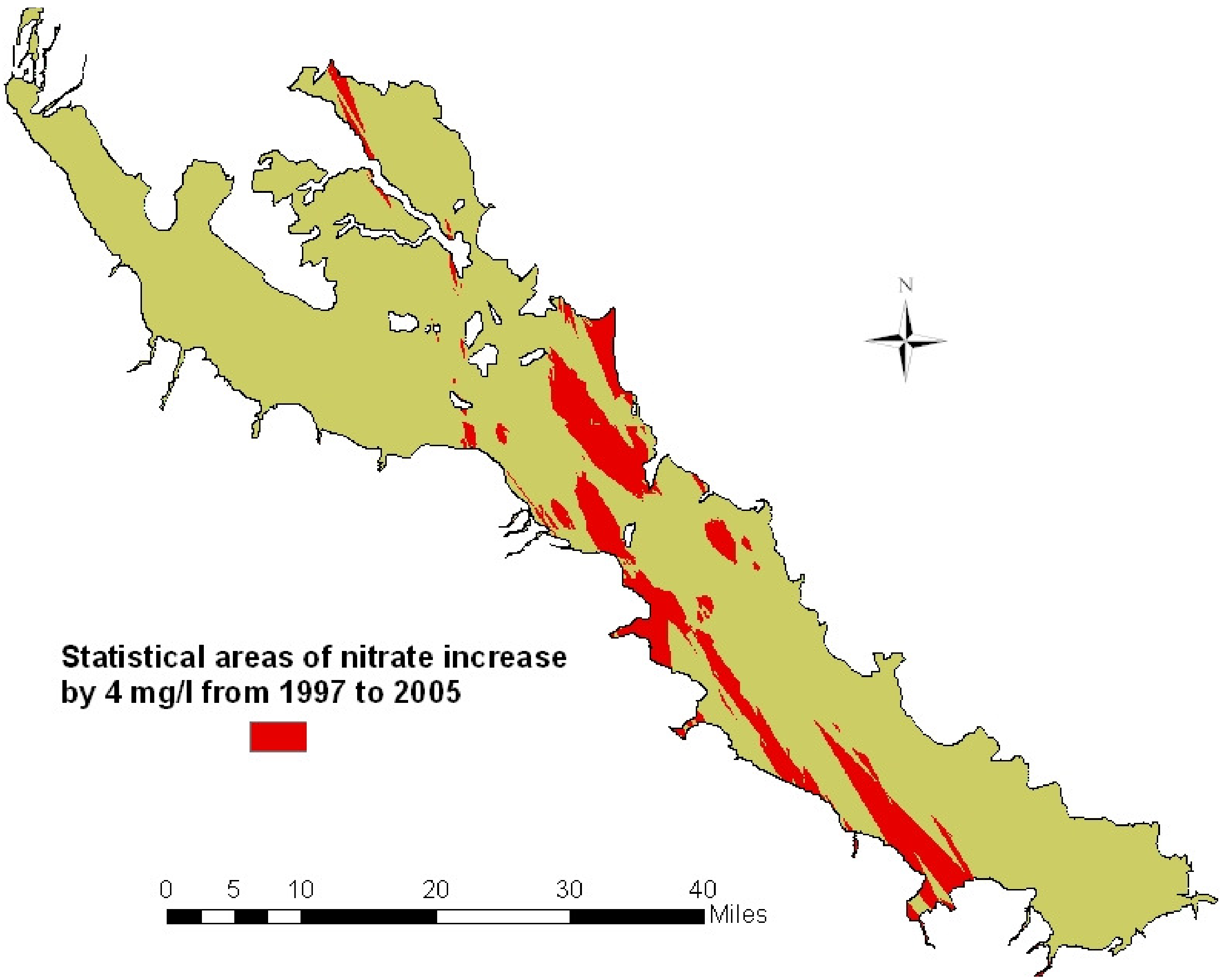


Figure 4-16. Statistical areas of nitrate that has increased by at least 4 mg/l from 1997 to 2005 at 95 % confidence

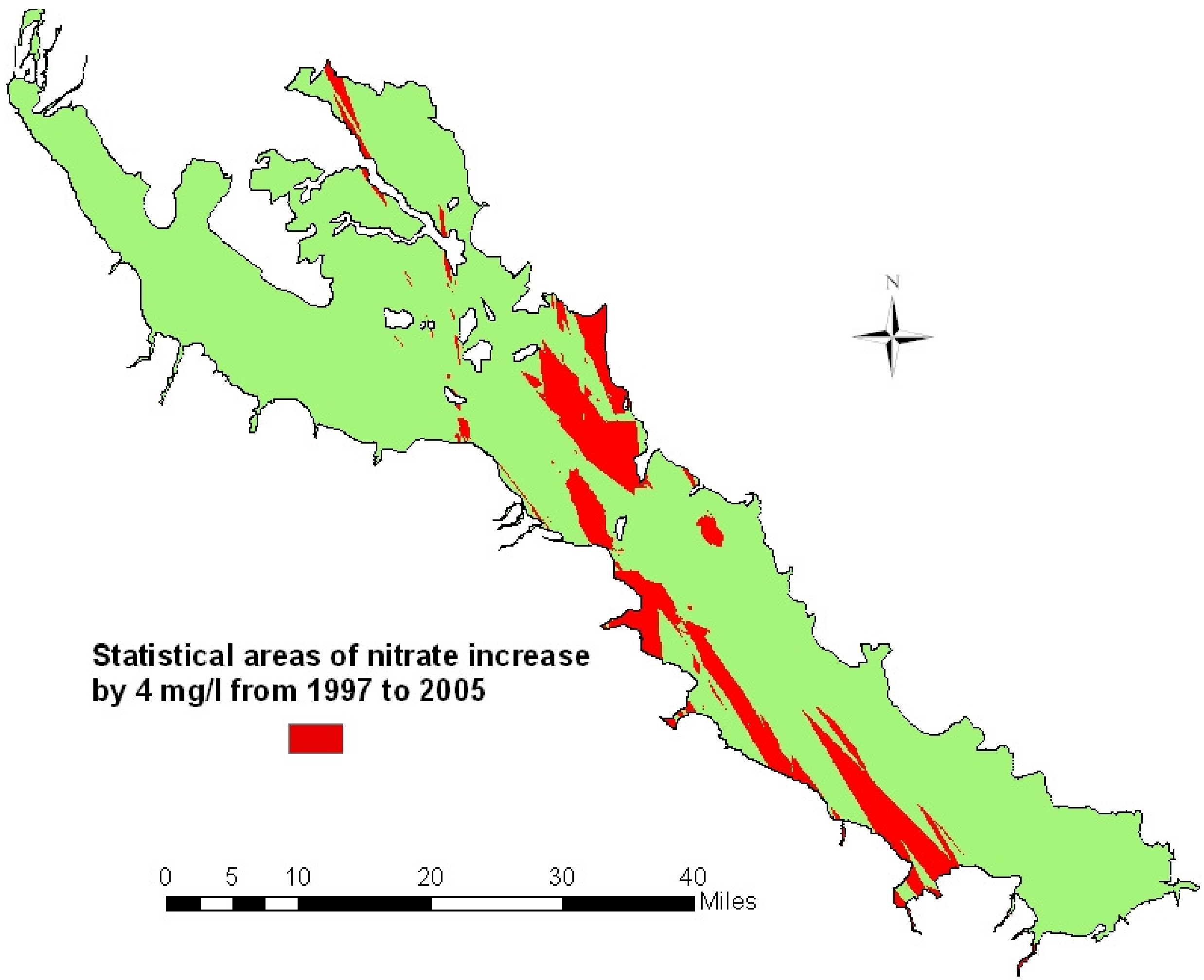


Figure 4-17. Statistical areas of nitrate that has increased by at least 4 mg/l from 1997 to 2005 at 70 % confidence

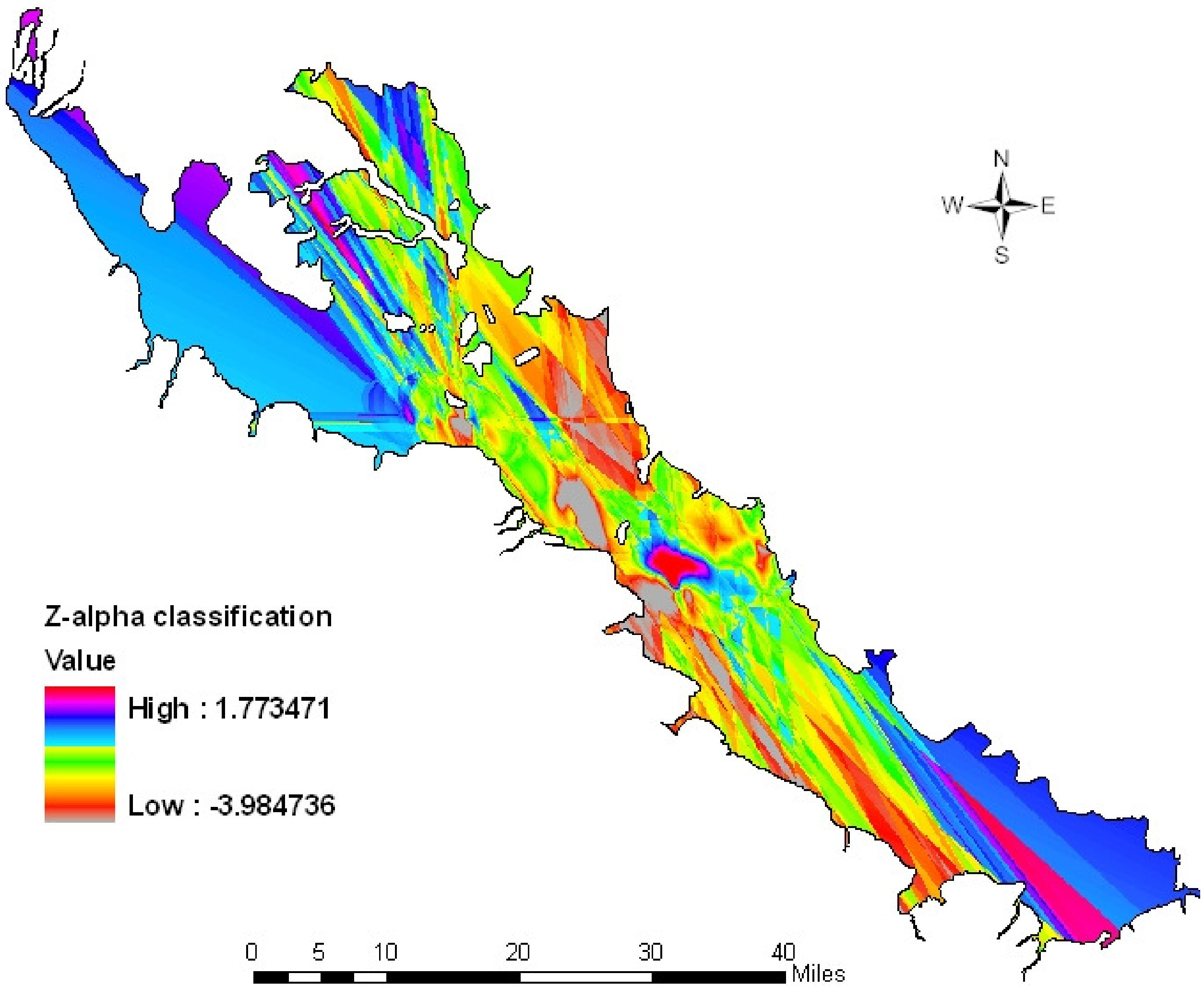


Figure 4-18. Areas of Z_{α} classification from two sided hypothetical test between 1997 and 2005

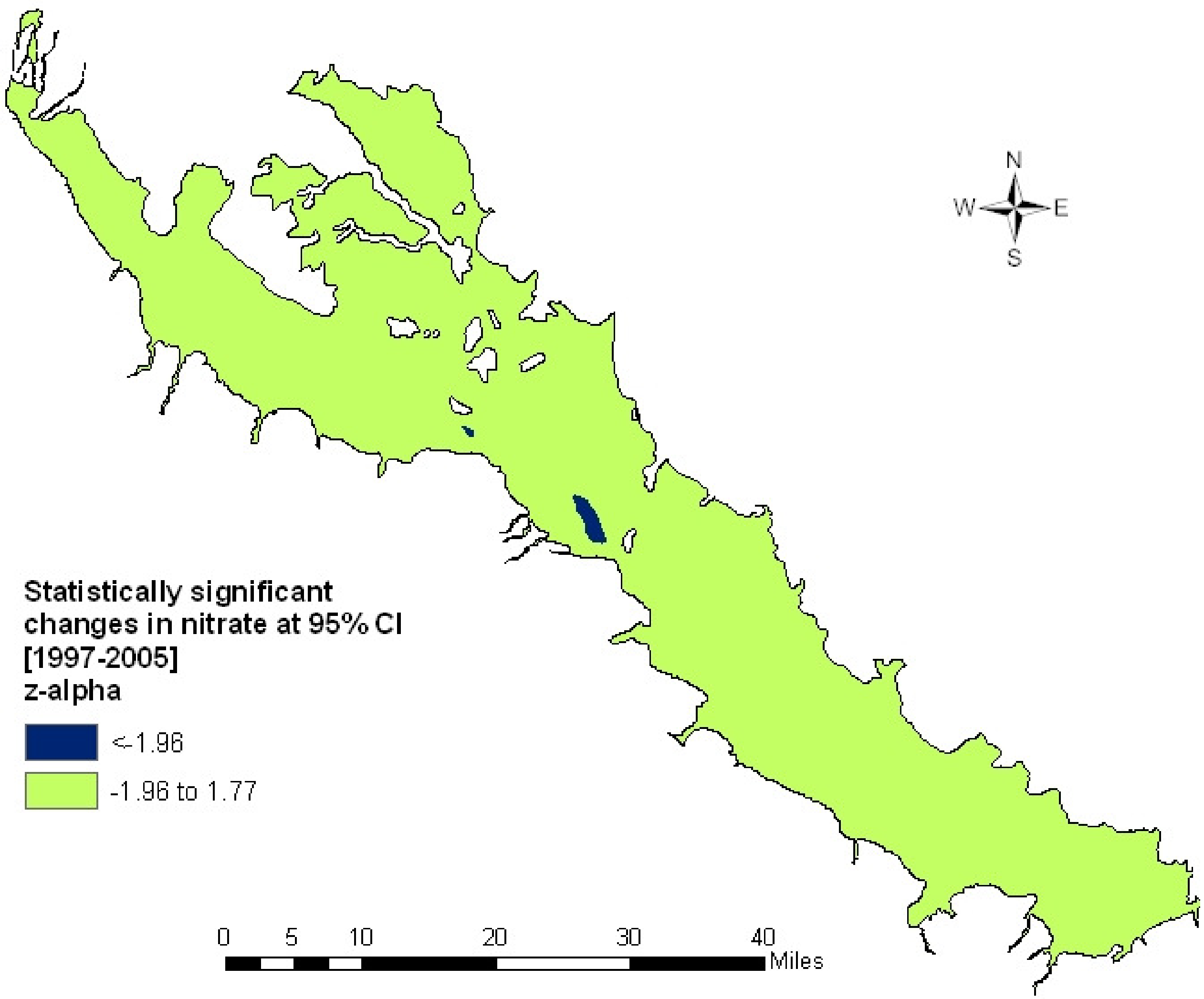


Figure 4-19. Areas of statistically significant changes from 1997 to 2005 at 95 % Confidence (<-1.96)

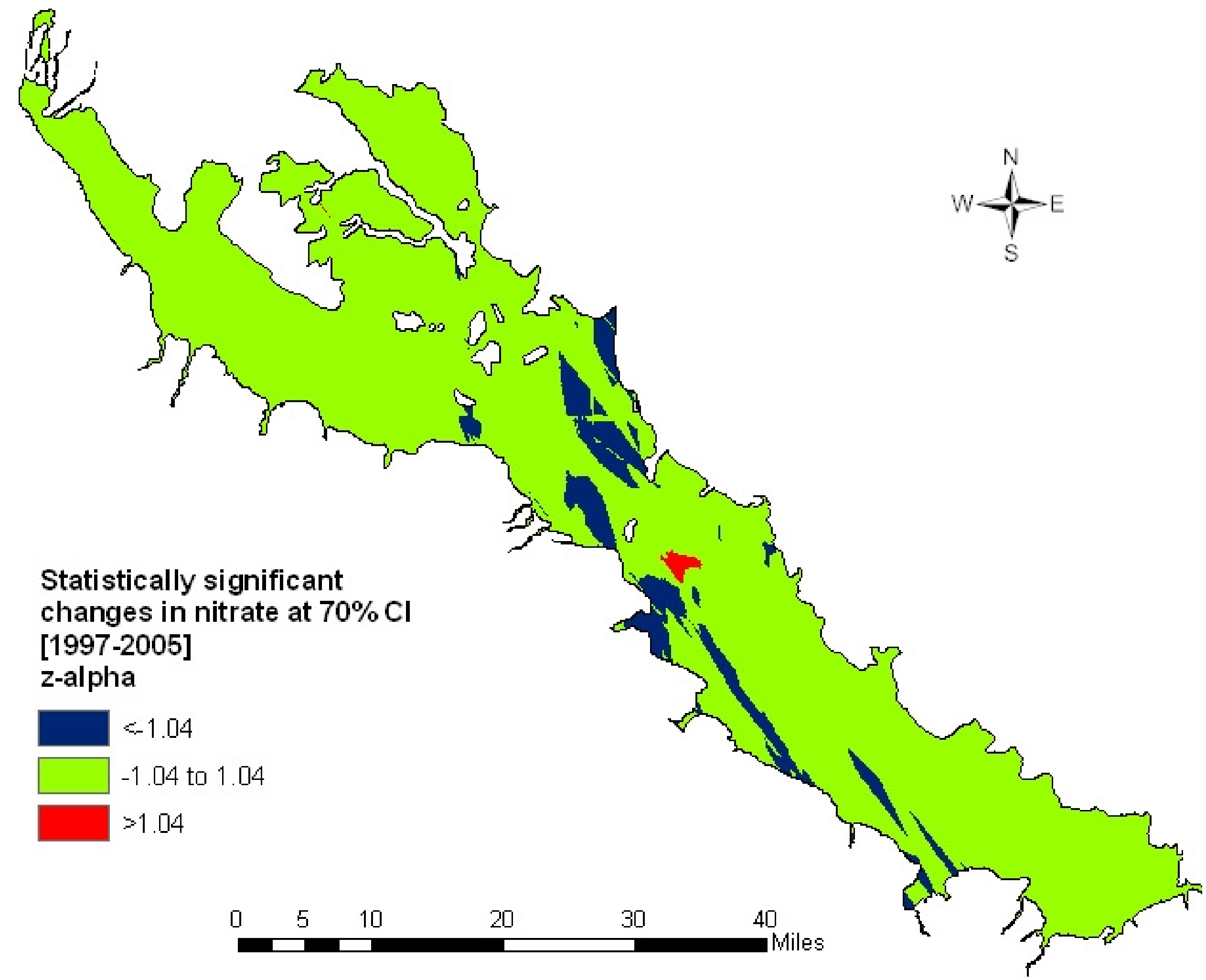


Figure 4-20. Areas of statistically significant changes from 1997 to 2005 at 70 % Confidence (<-1.04 and >1.04)

4.2.3 Indicator kriging analysis

Filled contours of nitrates were analyzed by using an approach of probability levels at thresholds of 4 mg/l and 10 mg/l of nitrate. Indicator kriging was used to estimate the spatial variability of a non-linear transform of the measured nitrate values. These indicator-transformed values produced estimates of the probability of nitrate occurrence such that given thresholds of 4 mg/l and 10 mg/l were exceeded at a given location during study period of 1997 to 2005. Cross-validation criteria from using spherical semi-variogram functions are shown in table 4-4 below. Probability of nitrate concentrations exceeding given thresholds for the year 1997 are shown in figures 4-21 and 4-22. It was found that the probability of exceeding a threshold of 4 mg/l was very high throughout the aquifer. Conversely, the probability of exceeding 10 mg/l of nitrate in the aquifer was low except for a small area in the central portion. Probability maps of all sampling periods can be found in appendix D and E.

The tendency of probability to remain geographically stationary or to change with time was also evaluated as chronic exceedance of nitrates from sampling year 1997 to 2005. These maps were generated by adding reclassified probability maps of all sampling periods in ArcGIS's Spatial Analyst. Two probability levels of 0.70 or more and 0.95 or more were chosen to define chronic exceedance of nitrates in the aquifer. Chronic exceedance maps of nitrates exceeding two probability levels of detecting equal to or greater than 4 mg/l and 10 mg/l are shown in figures 4-23 through 4-26. Areas of chronic exceedance of nitrate exceeding abovementioned probabilities of detecting equal or greater than 4 mg/l and 10 mg/l of nitrate are summarized in table 4-5. With given aquifer area of 1,242.5 square miles, these values can be transformed into percentages of area in the aquifer.

Table 4-4. Cross-validation results

Year	Cross Validation Criteria for nitrate exceeding 4 mg/l as N		Sample size
	ME	RMSE	
1997	0.007	0.995	85
1998	0.003	0.994	102
1999	0.004	1.023	95
2000	0.007	0.919	93
2001	-0.003	0.990	88
2002	0.003	1.007	98
2003	0.005	1.029	98
2004	0.007	1.011	80
2005	0.007	1.024	82
Year	Cross Validation Criteria for nitrate exceeding 10 mg/l as N		Sample size
	ME	RMSE	
1997	0.001	0.991	85
1998	-0.005	1.086	102
1999	-0.005	1.050	95
2000	-0.011	1.098	93
2001	-0.007	0.976	88
2002	-0.005	0.966	98
2003	-0.004	0.999	98
2004	0.0002	0.923	80
2005	-0.0032	0.940	82

Table 4-5. Areas of chronic exceedance frequencies of nitrate in the aquifer

	Areas detecting ≥4 mg/l of $N0_3$-N (sq. miles)		Areas detecting ≥10 mg/l of $N0_3$-N (sq. miles)	
Exceedance time	at ≥ 0.70 Probability	at ≥ 0.95 Probability	at ≥ 0.70 Probability	at ≥ 0.95 Probability
At least 11% of time	1,159.4	1,106.6	226.2	16.8
At least 22% of time	1,127.5	1,008.7	208.2	9.3
At least 33% of time	1,092.7	834.3	175.2	3.2
At least 44% of time	1,058.3	665.8	92.6	0.1
At least 56% of time	1,013.7	503.7	72.2	0.0
At least 67% of time	850.0	381.3	11.2	0.0
At least 78% of time	753.2	120.7	3.9	0.0
At least 89% of time	679.6	80.9	0.0	0.0
At least 100% of time	460.7	56.4	0.0	0.0

Note: area of aquifer = 1,242.5 square miles

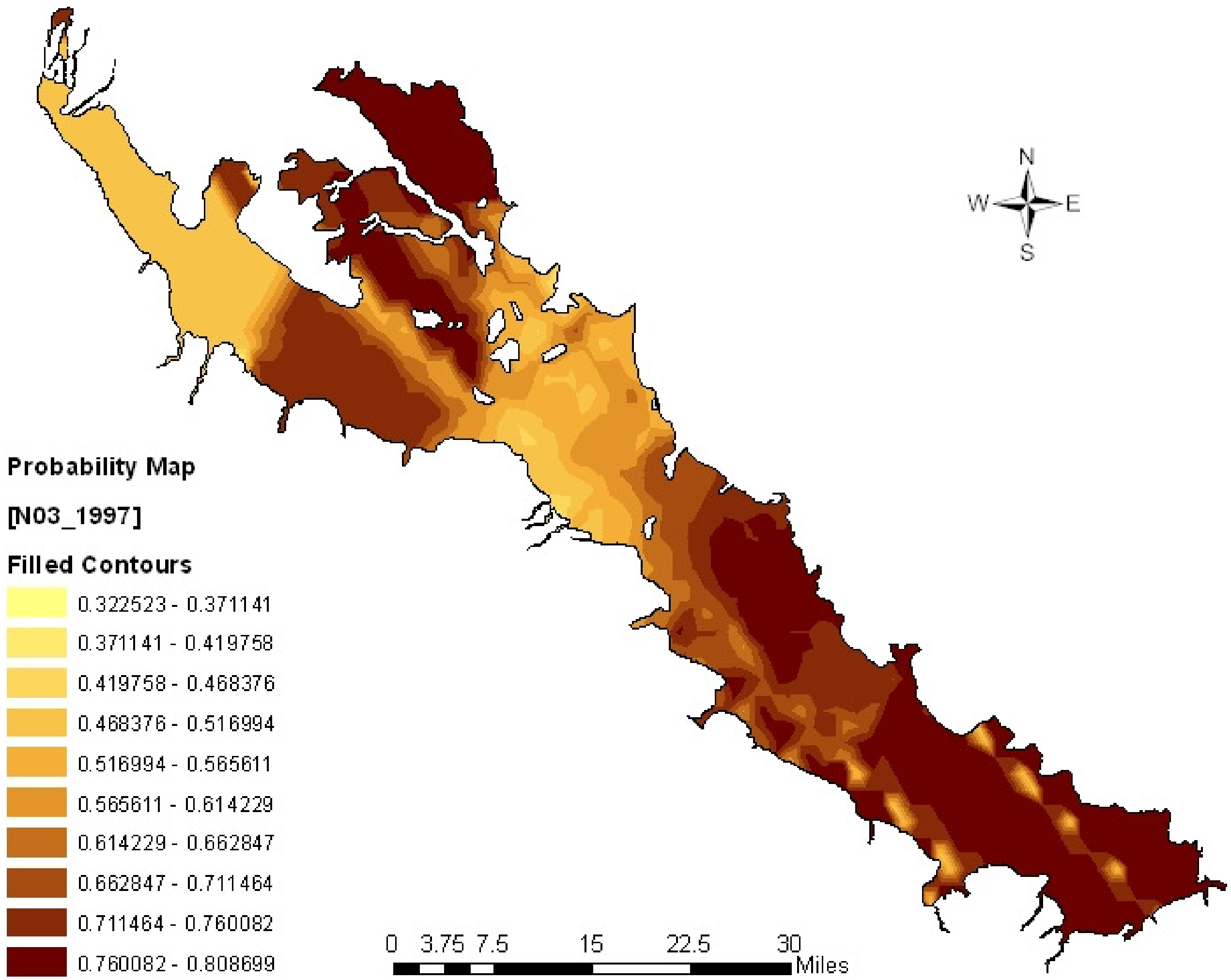


Figure 4-21. Probability map of nitrate concentration in 1997 at threshold of 4 mg/l

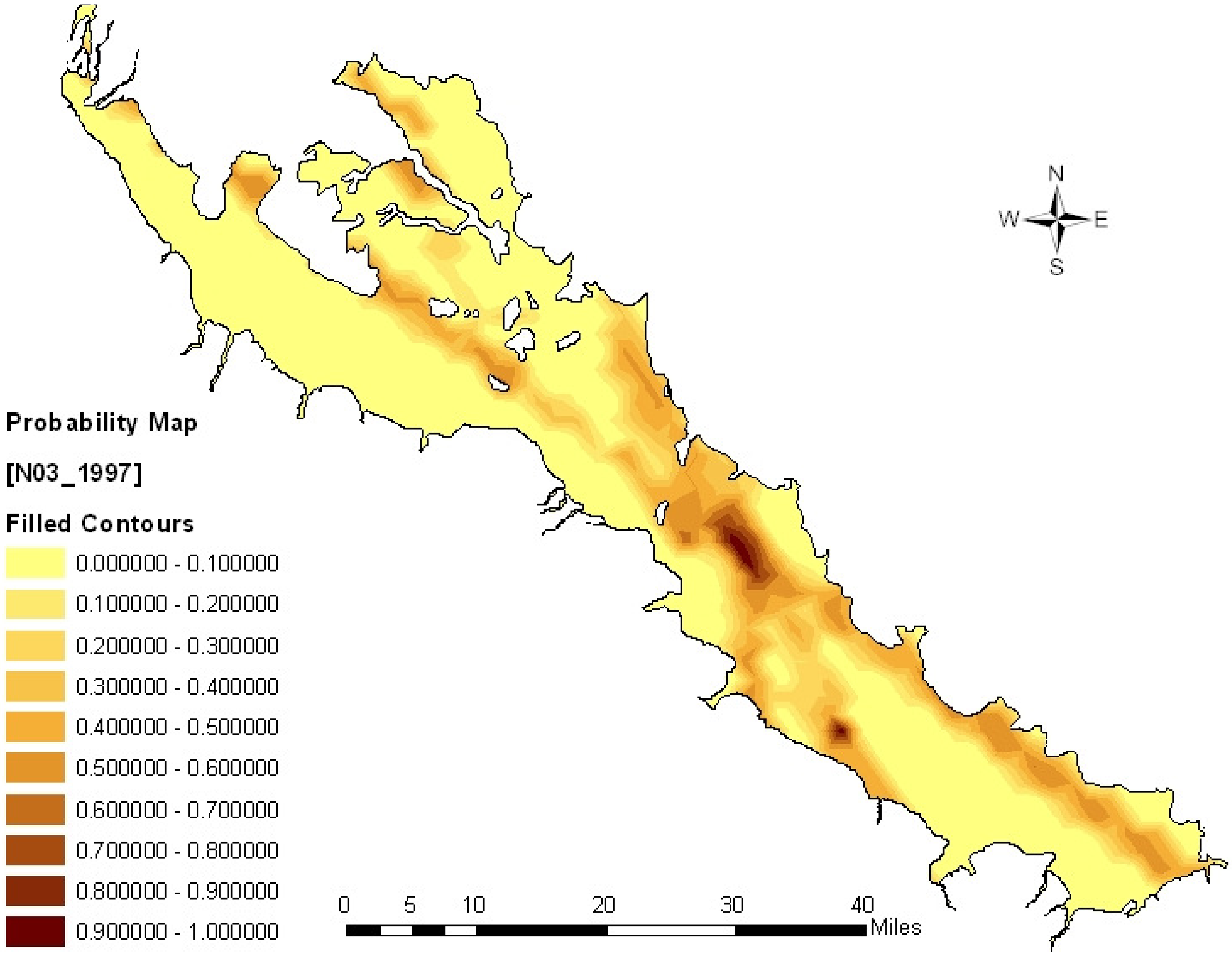


Figure 4-22. Probability map of nitrate concentration in 1997 at threshold of 10 mg/l

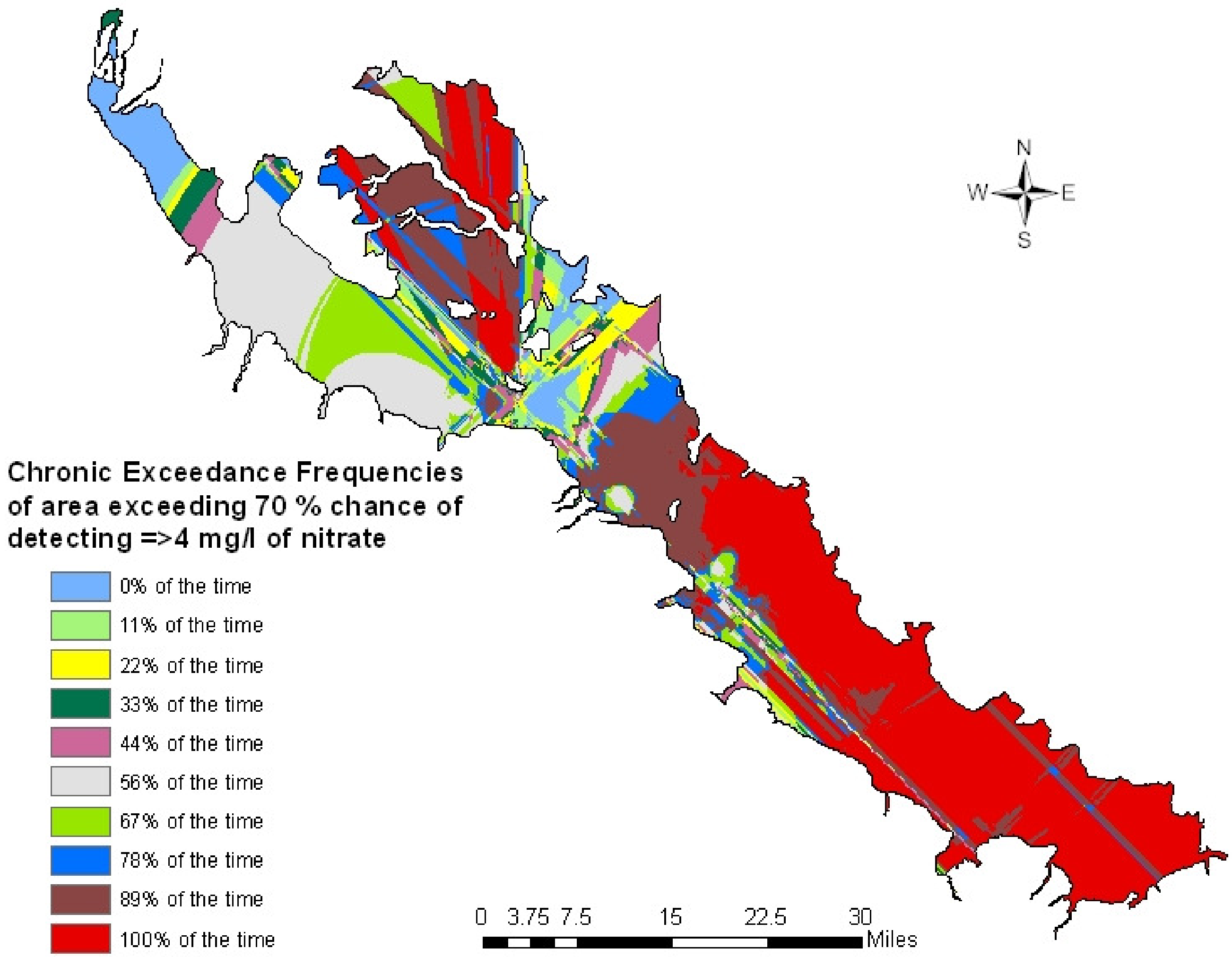


Figure 4-23. Chronic exceedance frequencies of area exceeding 70% chance of detecting ≥ 4 mg/l of nitrate

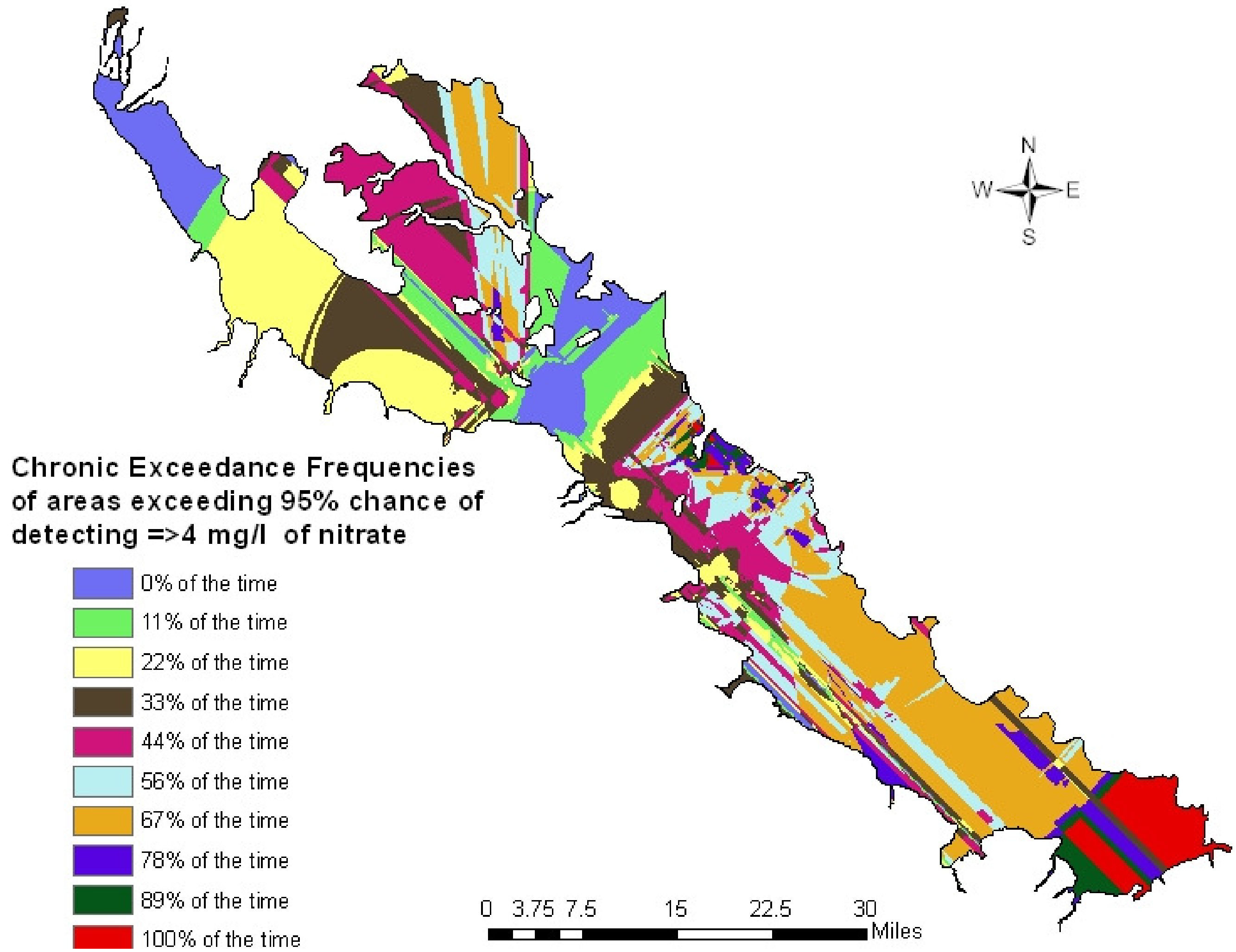


Figure 4-24. Chronic exceedance frequencies of area exceeding 95% chance of detecting ≥ 4 mg/l of nitrate

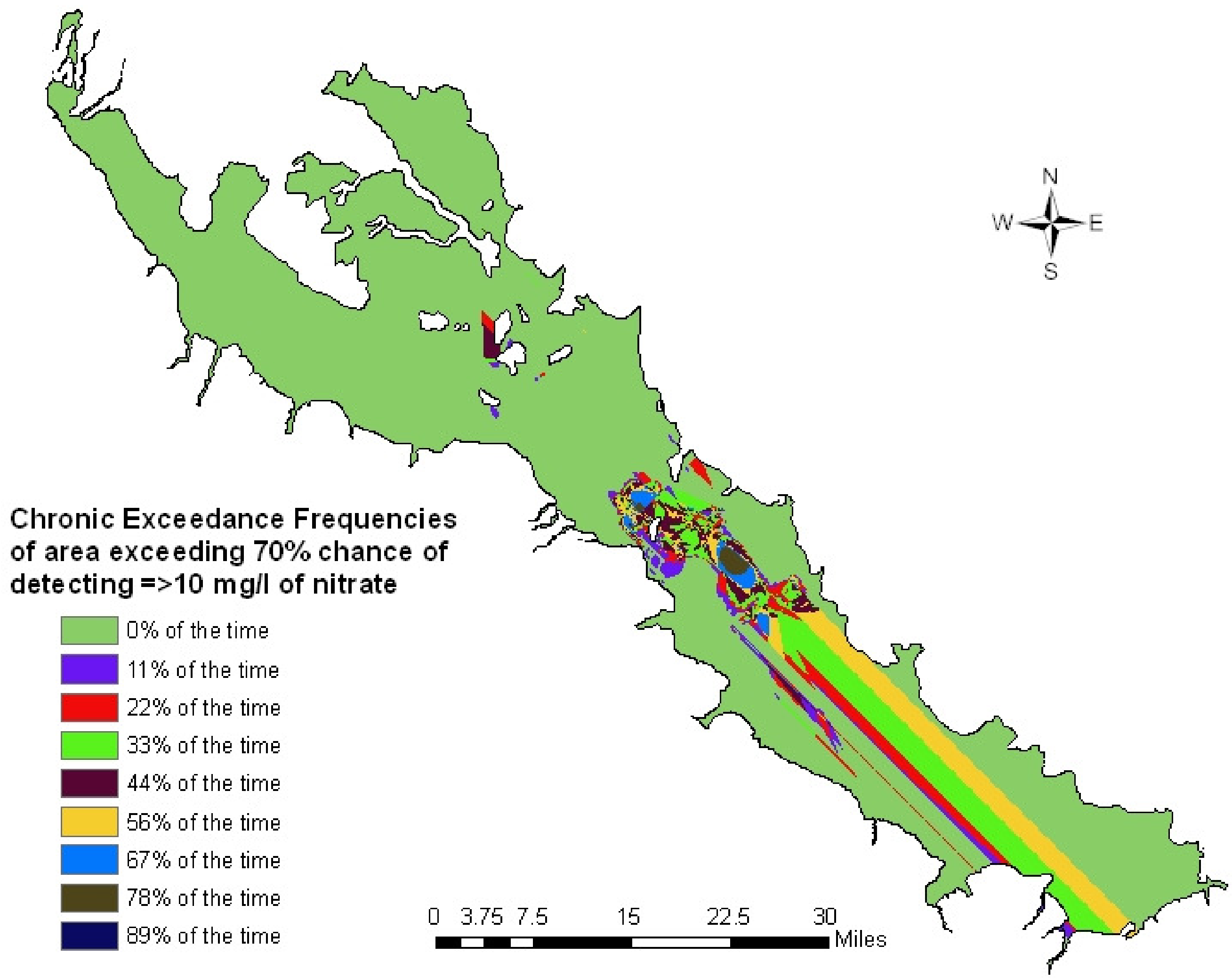


Figure 4-25. Chronic exceedance frequencies of area exceeding 70% chance of detecting ≥ 10 mg/l of nitrate

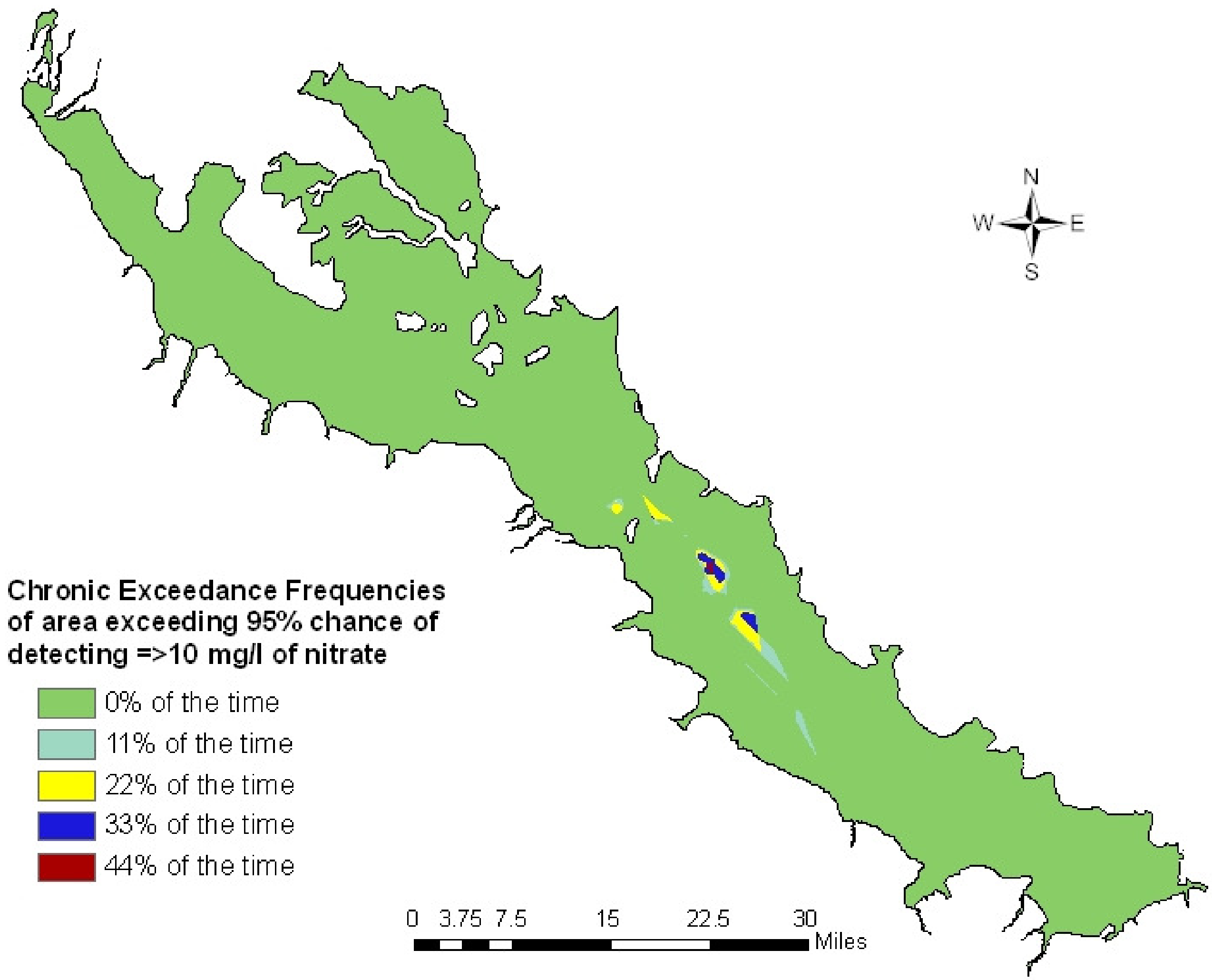


Figure 4-26. Chronic exceedance frequencies of area exceeding 95% chance of detecting ≥ 10 mg/l of nitrate

4.3 Logistic regression analysis

This section includes descriptions of variables, results of buffer analysis for establishing relationship between groundwater nitrate and land covers around sampled wells, and results of stepwise logistic regression models for specified thresholds of nitrate as nitrogen.

4.3.1 Descriptions of variables

Nitrate concentrations in the Cimarron terrace aquifer were highly variable. Nitrate data sampled in 2003 were used for logistic regression analysis. The figure 4-28 explains the distribution of nitrate concentrations in the aquifer. The mean and median concentrations were 7.42 mg/l and 7.28 mg/l respectively. Numbers of samples for each 0.5 mg/l of interval were plotted and are shown in figure 4-27. The range of nitrate concentration from 7 mg/l to 9.5 mg/l was frequently observed in the aquifer. Normality test of nitrate data were examined. The p-value of 0.062 (Fig 4-29) indicates that nitrate concentration in the Cimarron terrace aquifer generally followed a normal distribution. The empirical cumulative distribution function (ecdf) was also plotted against measured nitrate values (Fig 4-30). This plot closely followed the normal probability distribution.

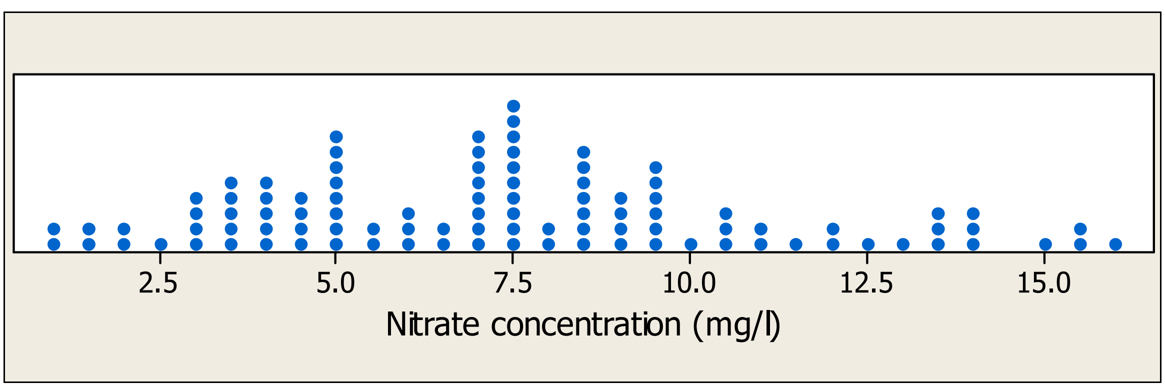


Figure 4-27. Frequencies of nitrate occurrence in the aquifer

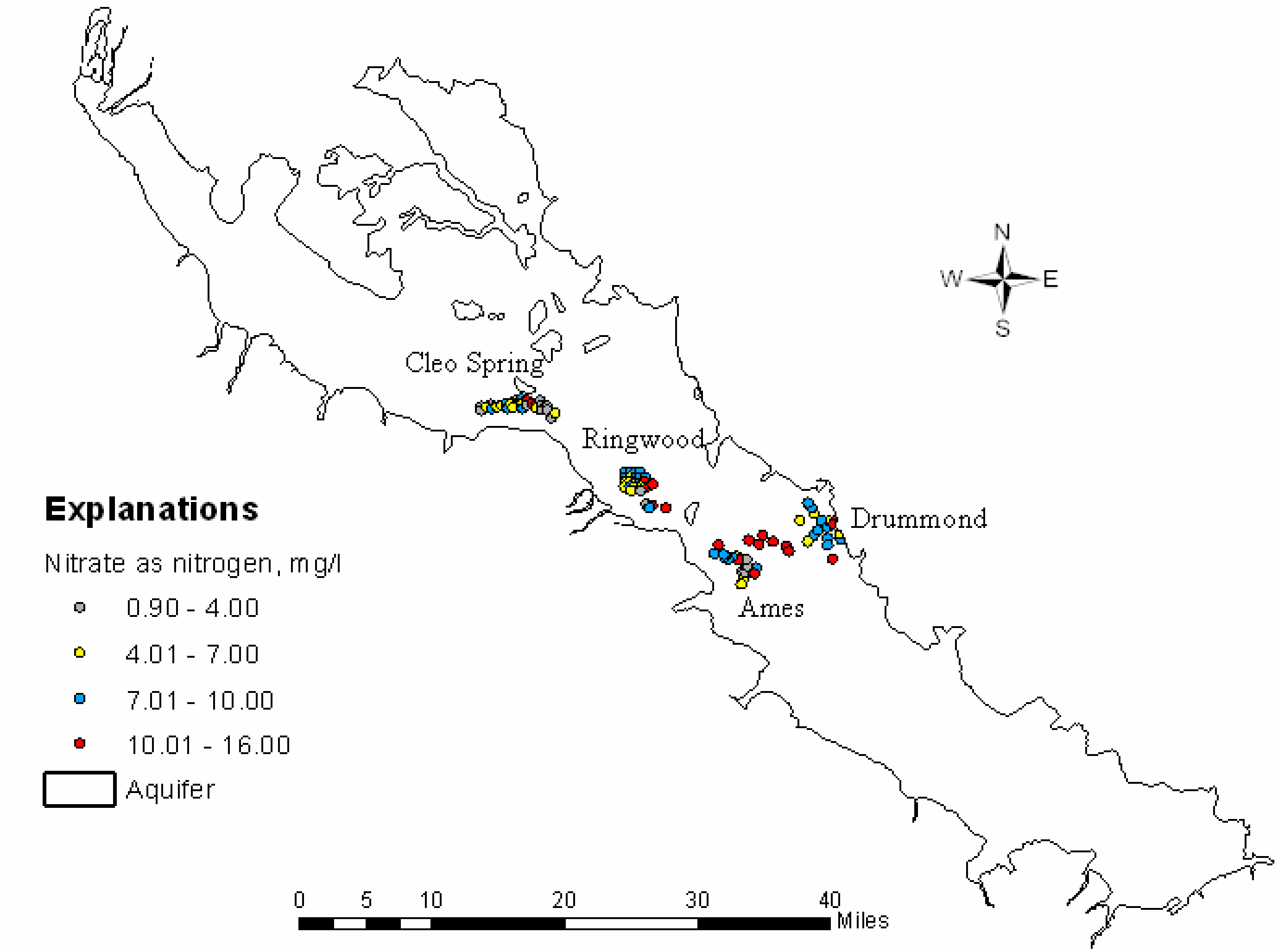


Figure 4-28. Location of sampling wells in the Cimarron terrace aquifer

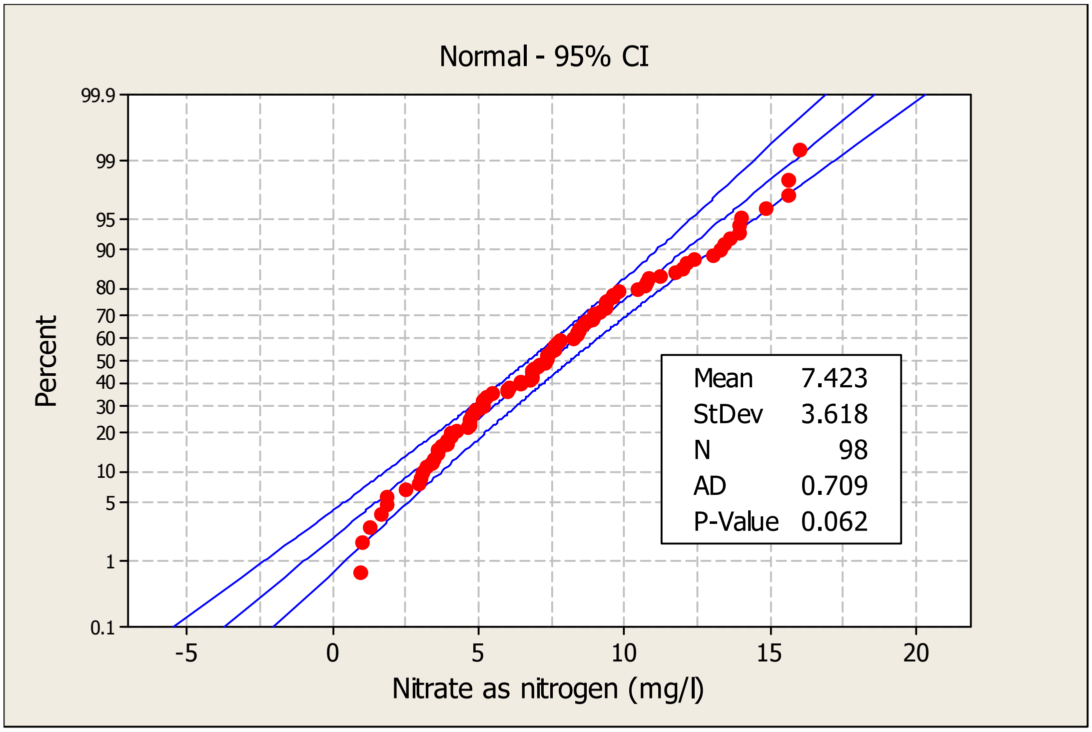


Figure 4-29. Normal probability plot of nitrate values

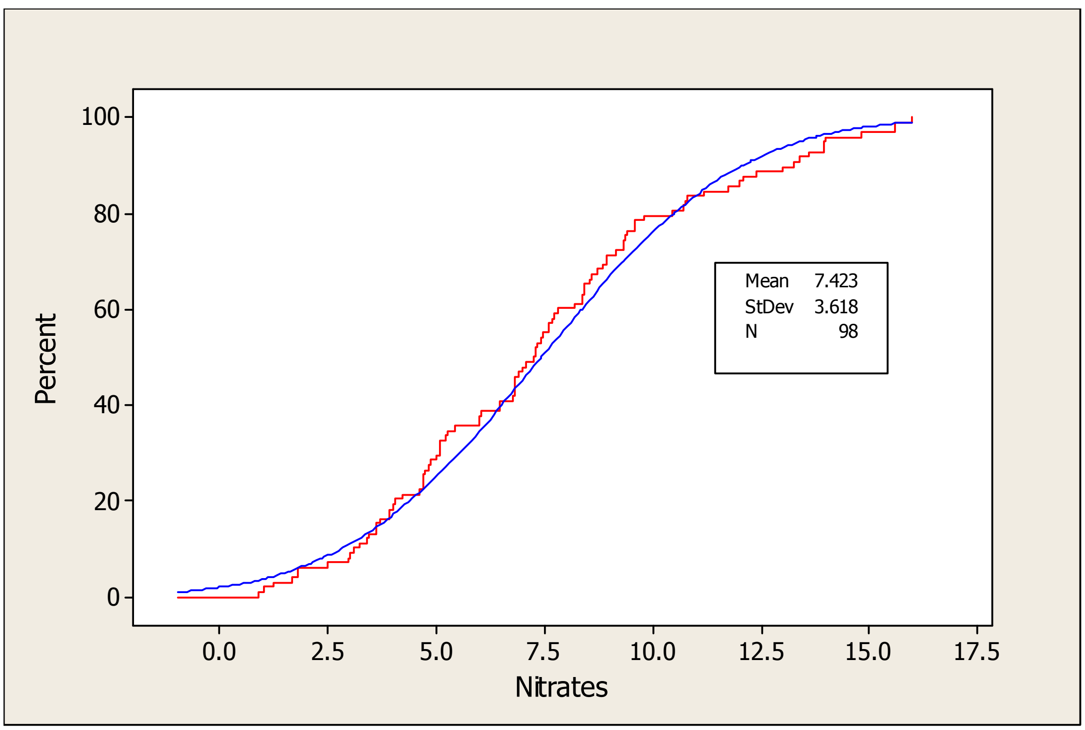


Figure 4-30. Emperical CDF of measured nitrate values

A total of 10 explanatory variables for nitrate contamination in the Cimarron terrace aquifer were initially evaluated. These variables include land covers such as developed land, forested-wetland, grassland, and cropland-pasture, percent of clay, sand, and silt by volumes, percent of organic matter by weight, fertilizer N, and depth to groundwater. Aggregated land covers are shown in figure 4-31. Cropland-pasture occupied 47% of aquifer area while developed land, forested-wetland, and grassland occupied 5.11%, 5.20%, and 41.09% of aquifer area respectively. The Cimarron terrace aquifer consists of high percentage of sand (Fig. 4-32). The percent of silt varies moderately (Fig. 4-33) and that of clay varies from 0 to 52% throughout the aquifer (Fig. 4-34). The percent of organic matter that represents denitrification potential in the aquifer ranges from 0 to 4.5% by weight (Fig 4-35). Fertilizer N of 110 kg/ha was apportioned equally to agricultural and developed land to account for residential fertilizer use. Depth to seasonally high water table that is the unsaturated zone thickness ranged from 6 to 71 feet with 30 feet median depth.

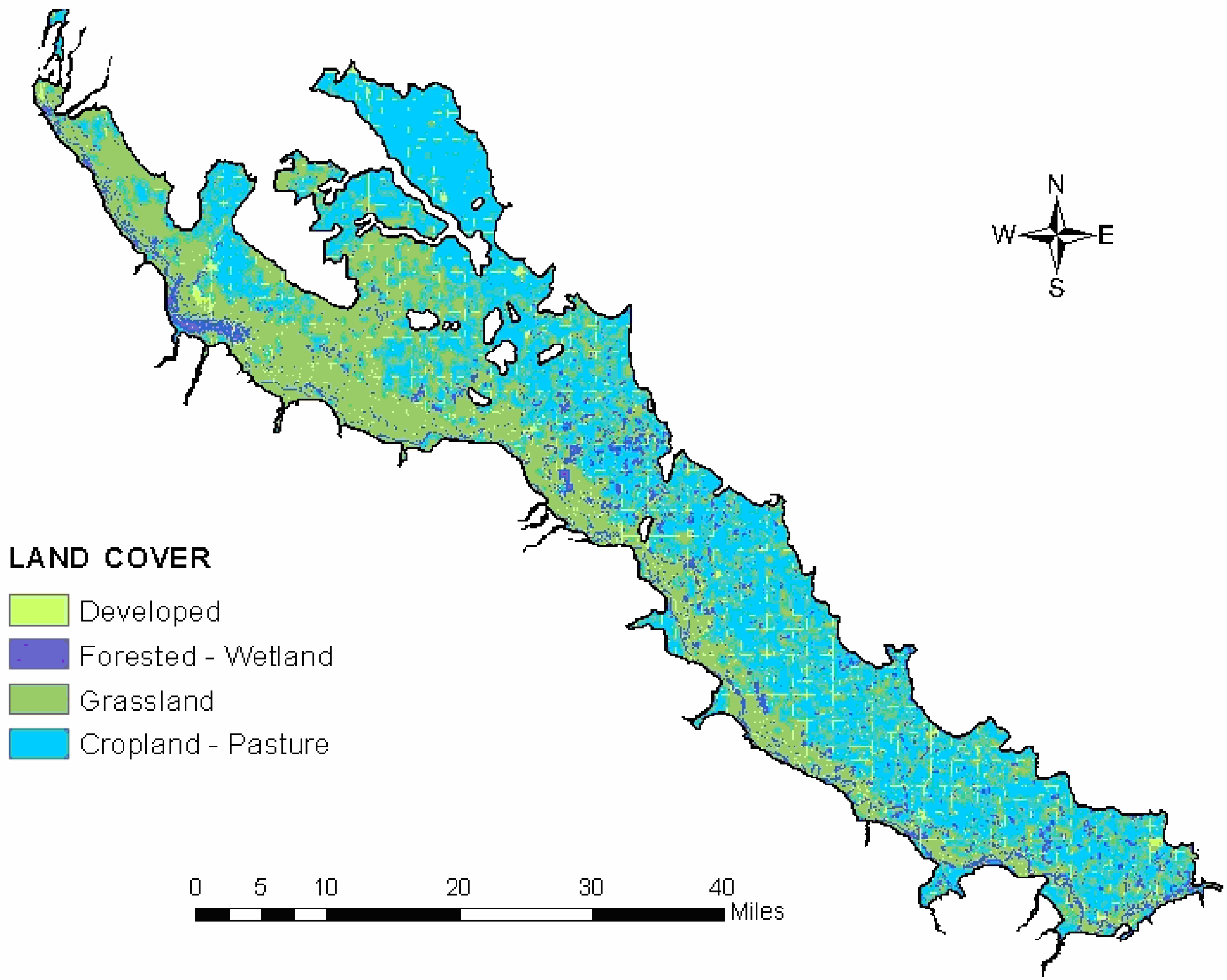


Figure 4-31. Distribution of land cover (aggregated) based NLCD 2001 database

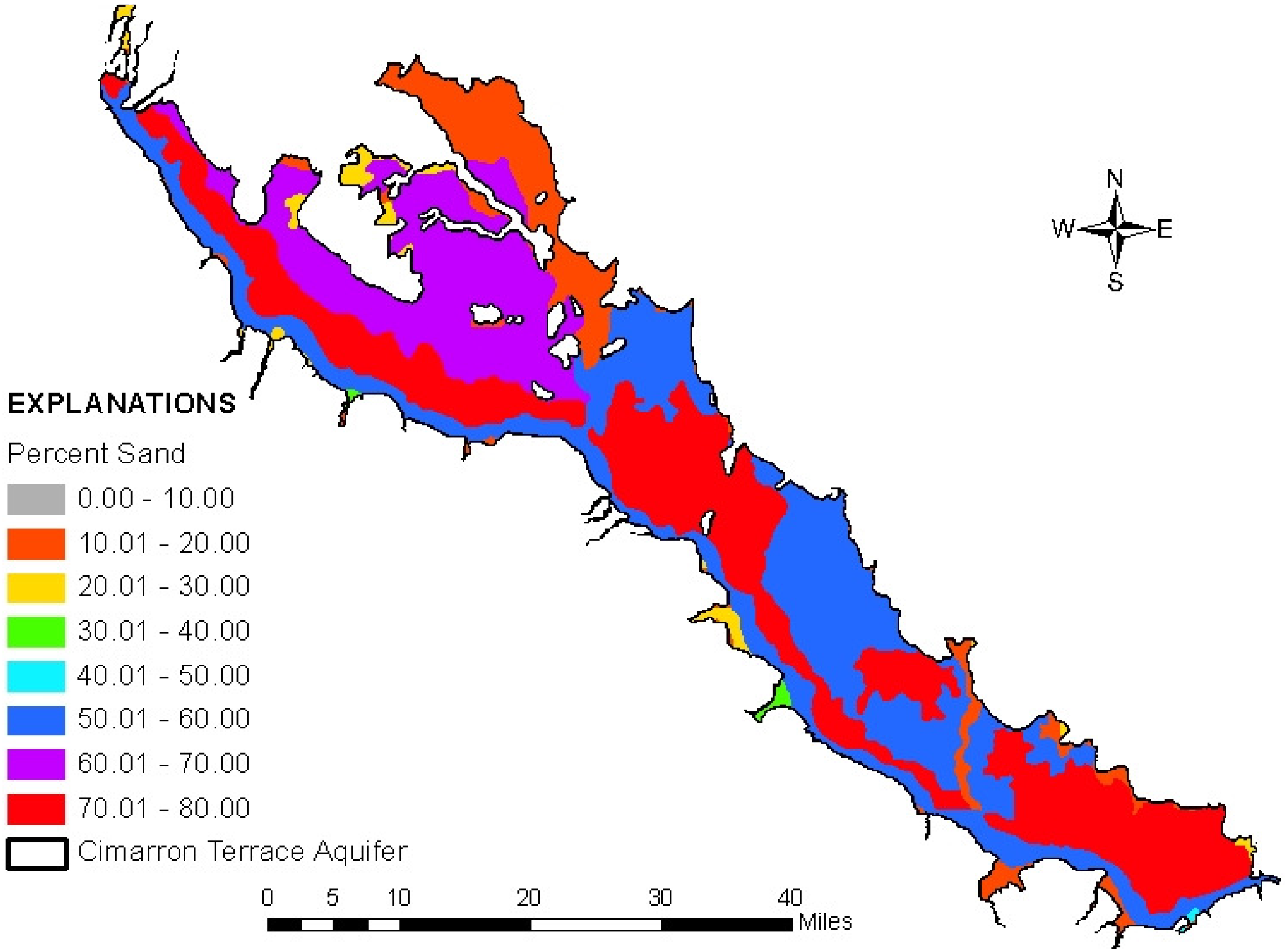


Figure 4-32. Average soil profile sand content derived from STATSGO database

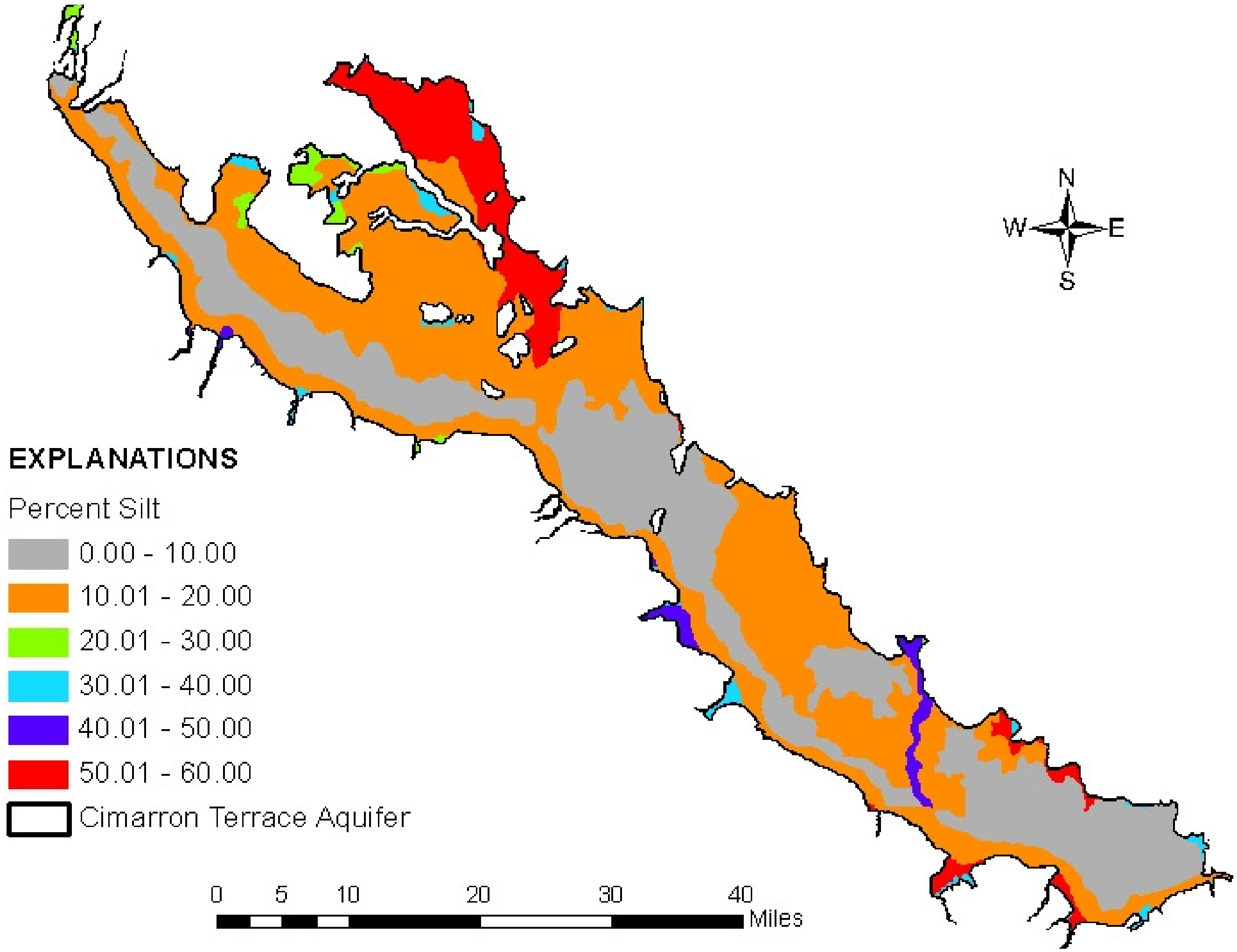


Figure 4-33. Average soil profile silt content derived from STATSGO database

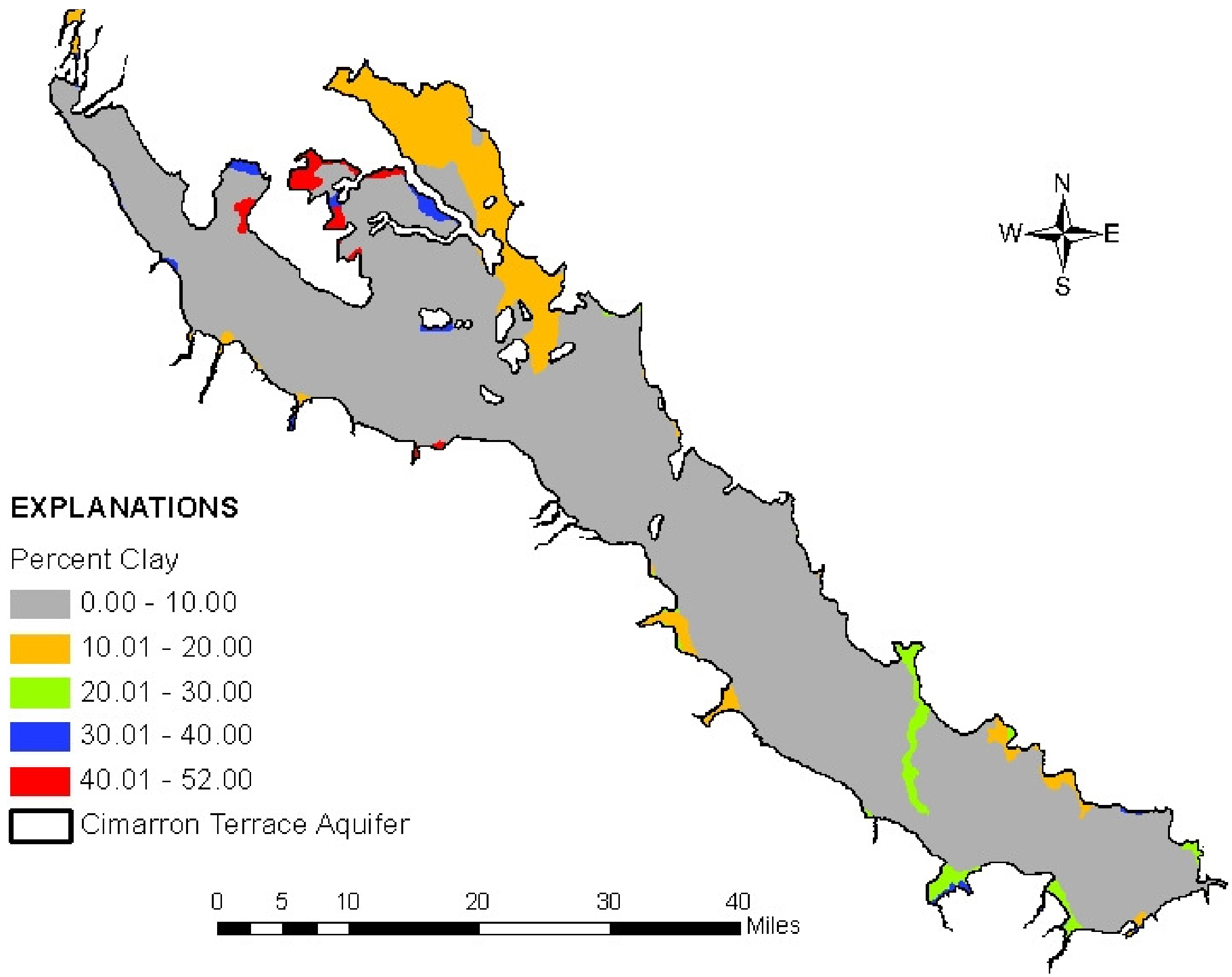


Figure 4-34. Average soil profile clay content derived from STATSGO database

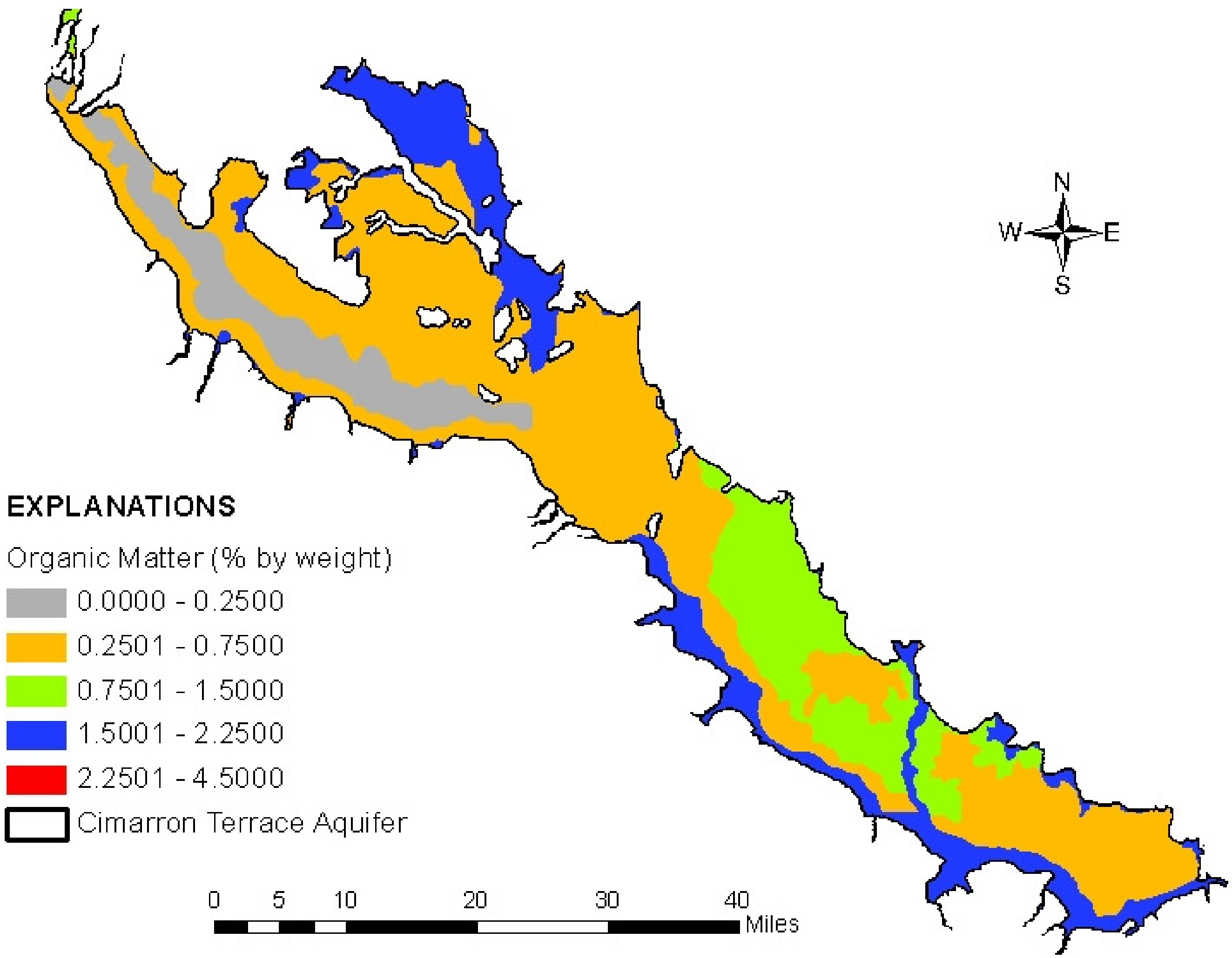


Figure 4-35. Average soil profile organic matter derived from STATSGO database

4.3.2 Areas of well influence

A logistic regression analysis was performed to establish a relationship between groundwater nitrate and the type of land cover around sampled wells. Land cover types as shown in figure 4-33 were statistically analyzed for various well buffers (100m, 250m, 500m, 750m, 1,000m, 1,500m, 1,750m). Land covers at 1,000 meters radial distance defined the best fit logistic model for both 4 mg/l and 10 mg/l of thresholds by examining G-statistic and Wald statistic. A typical GIS buffer of 1,000 m radial distance is shown in figure 4-36. The G-statistic values of 13.49 and 16.82, and Wald statistic values of 10.96 and 13.39 at thresholds 4 mg/l and 10 mg/l respectively were the highest statistics at 1,000 m radius of well influence. G-statistics and Wald statistics along with p-values for various well buffers are presented in tables 4-6 and 4-7. These statistics are plotted in figures 4-37 and 4-38 for better visualization.

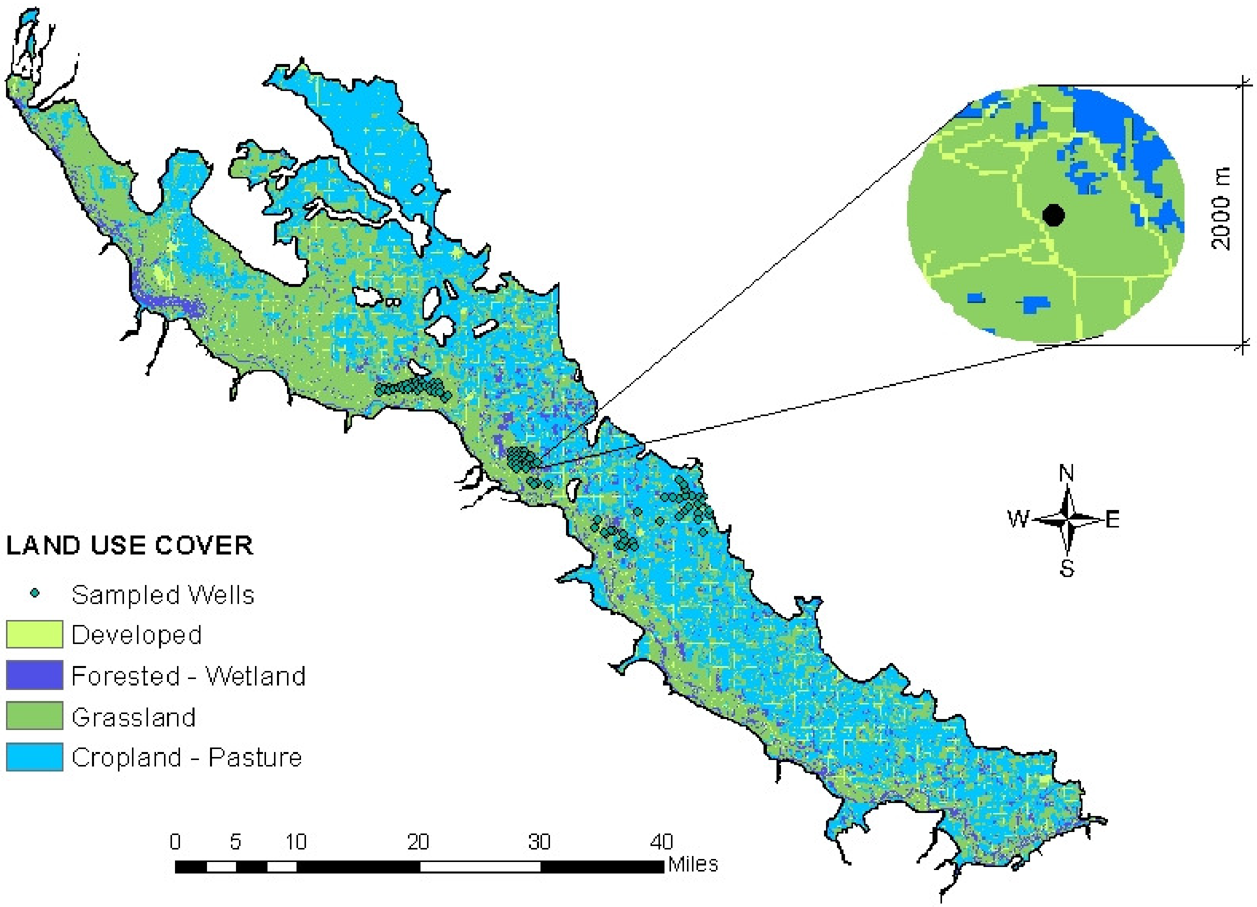


Figure 4-36. Extractions of land cover variables within a statistical area of well influence around each groundwater well sampled in the Cimarron terrace aquifer

Table 4-6. G and Wald statistics for land cover types at various buffers at threshold 4 mg/l

Radius, meter	Likelihood ratio test		Wald test	
	G-statistic	p-value	Wald statistic	p-value
100	1.664	0.797	1.56	0.815
250	3.50	0.477	3.09	0.5424
500	11.83	<0.018	9.58	<0.0481
750	11.511	<0.021	9.073	0.0593
1000	13.49	<0.009	10.96	< 0.0269
1500	11.81	<0.018	9.89	< 0.0423
1750	11.30	< 0.023	9.27	0.0546

Table 4-7. G and Wald statistics for land cover types at various buffers at threshold 10 mg/l

Radius, meter	Likelihood ratio test		Wald test	
	G-statistic	p-value	Wald statistic	p-value
100	4.1679	0.3838	4.0821	0.3950
250	2.691	.6108	2.71	0.60
500	6.385	.1721	5.839	0.211
750	6.971	0.137	6.586	0.159
1000	16.82	<0.002	13.39	<0.0095
1500	9.067	0.0594	8.543	0.07
1750	9.08	0.059	8.31	0.08

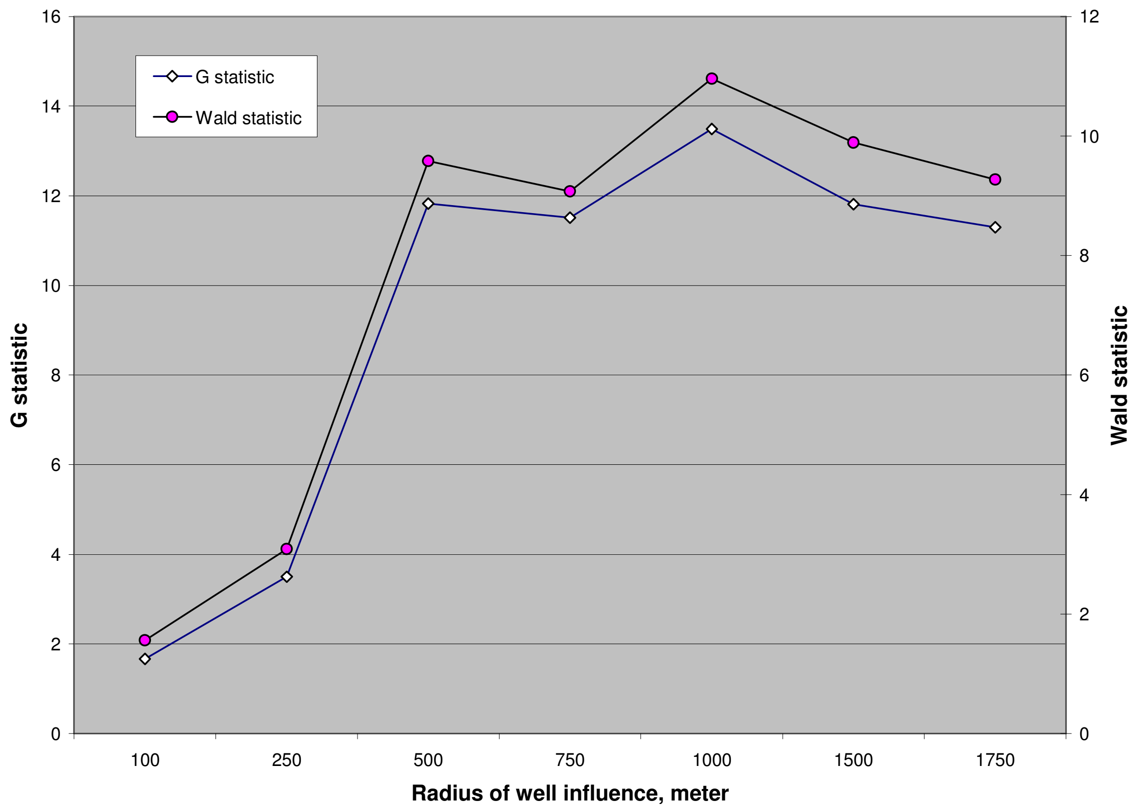


Figure 4-37. Best radius of well influence for the best-fit-model using land cover variables for nitrate concentrations in groundwater exceeding a threshold of 4 mg/l

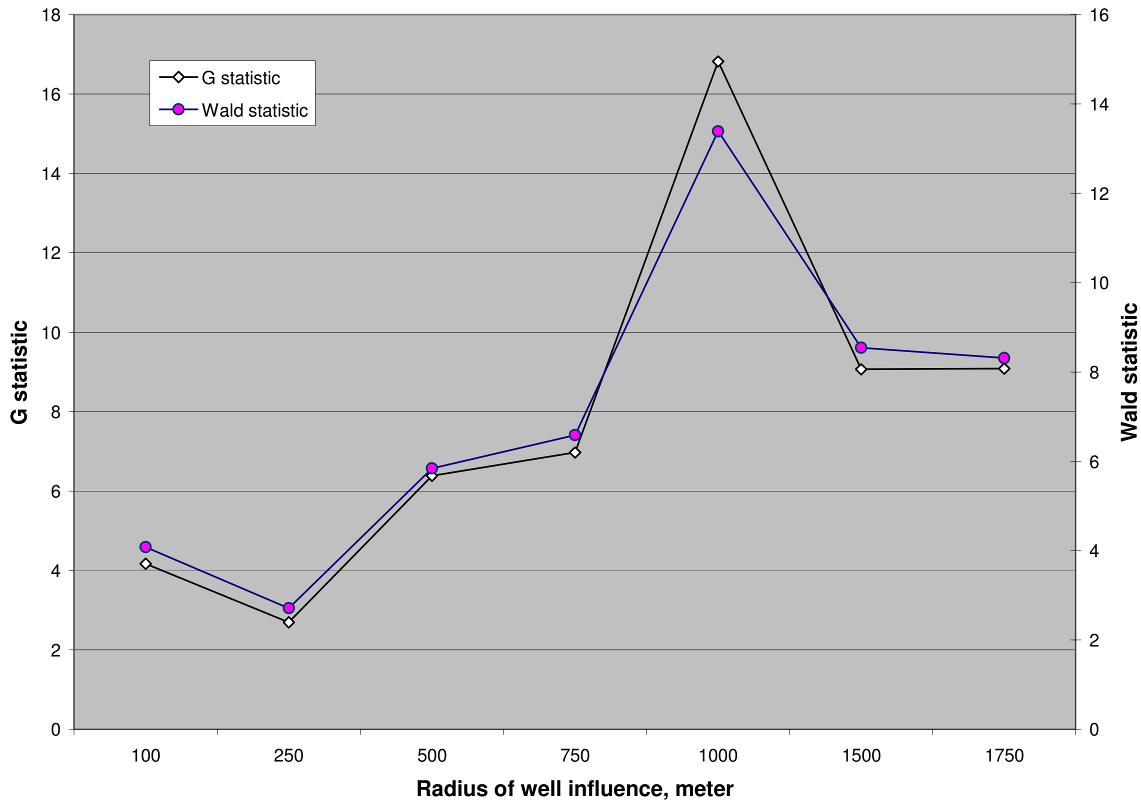


Figure 4-38. Best radius of well influence for the best-fit-model using land cover variables for nitrate concentrations in groundwater exceeding a threshold of 10 mg/l

The best radius of well influence (1,000 m) was applied to all explanatory variables that were the function of area. The descriptive statistics of all explanatory variables within a statistical area of well influence along with nitrate data are summarized in table 4-8 below.

Table 4-8. Dependent and explanatory variables and their descriptive statistics within a statistical area of well influence

Variable	Unit or Categories	Minimum	Median	Maximum	Quartiles	
					25%	75%
Dependent						
Nitrate as nitrogen*	Milligram per liter	0.90	7.28	16.00	4.71	9.34
Explanatory						
Developed land	Percentage	1.33	5.52	11.24	4.25	6.52
Forested-wetland	Percentage	0.00	4.82	23.84	1.22	9.51
Grassland	Percentage	4.95	71.27	96.79	43.73	83.38
Cropland-pasture	Percentage	0.00	14.40	88.01	0.14	48.20
Sand	Percentage by volume	21.84	73.91	78.18	59.00	77.42
Silt	Percentage by volume	4.27	7.73	45.88	6.13	14.09
Clay	Percentage by volume	2.57	4.91	14.50	3.62	9.00
Soil organic matter	Percentage by weight	0.05	0.75	1.90	0.31	1.50
Inorganic fertilizer application	kg/sq. mile	2.87	27.06	125.69	9.50	70.30
Depth to water table*	Feet	5.82	30.19	71.00	18.75	38.61

* Well point data

4.3.3 Univariate models

Various explanatory variables were tested individually if they were significant predictors of nitrate contamination in the groundwater. Most of the potential explanatory variables were significantly related with thresholds 4 mg/l and 10 mg/l of nitrate (Tables 4-9 and 4-10). The p-value ≤ 0.05 of Likelihood ratio test and Wald statistic were used to screen variables for inclusion in stepwise multivariate logistic models. In addition to this, coefficient of determination (R^2 value) that describes Hosmer-Lemeshow (H-L) goodness-of-fit of observed versus predicted value of deciles risk was also used. Variables that did not meet the criteria of p-value were tested with R^2 value of 0.8. Results of univariate logistic regression analysis at thresholds 4 mg/l and 10mg/l are shown in figures 4.8 and 4.9 respectively. From the univariate model at threshold 4 mg/l of nitrate, forested-wetland variable was dropped because it did not meet any of abovementioned criteria. Similarly, these explanatory variables were tested for threshold of 10 mg/l of nitrate as nitrogen. Developed land, forested-wetland, and log of percent sand failed to screening test and hence dropped.

Table 4-9. Results of univariate logistic regression analysis to evaluate the significance of each explanatory variable exceeding 4 mg/l of nitrate as nitrogen

Explanatory variables	Likelihood ratio test (p-value)	Wald statistic (p-value)	Coefficient of determination (R^2-value)	Hosmer-Lemeshow p-value
Developed land	0.0017	0.0035	0.72	0.2915
Forested-wetland	0.7225	0.7258	0.49	0.5137
Grassland	0.0041	0.0143	0.32	0.0626
Cropland- pasture	0.0149	0.0373	0.84	0.0898
Log of percent sand	0.8988	0.8979	0.85	0.0051
Log of percent silt	0.2312	0.2538	0.88	0.0045
Log of percent clay	0.1124	0.1285	0.84	0.0005
Organic matter	0.0284	0.0376	0.98	0.0025
Log of fertilizer N	0.004	0.0065	0.69	0.6492
Log of depth to water table	0.2795	0.2762	0.85	0.9826

Table 4-10. Results of univariate logistic regression analysis to evaluate the significance of each explanatory variable exceeding 10 mg/l of nitrate as nitrogen

Explanatory variables	Likelihood ratio test (p-value)	Wald statistic (p-value)	Coefficient of determination (R^2-value)	Hosmer-Lemeshow p-value
Developed land	0.7306	0.7307	0.1	0.0559
Forested-wetland	0.9718	0.9717	0.07	0.6802
Grassland	0.0076	0.0086	0.39	0.0434
Cropland- pasture	0.0122	0.0127	0.15	0.034
Log of percent sand	0.3781	0.3638	0.68	0.5953
Log of percent silt	0.2525	0.2519	0.81	0.8269
Log of percent clay	0.1707	0.1713	0.9	0.6366
Organic matter	0.0629	0.0683	0.95	0.8739
Log of fertilizer N	0.001	0.003	0.65	0.4195
Log of depth to water table	0.0581	0.0597	0.35	0.3125

4.3.4 Multivariate models

Explanatory variables that passed these screening tests were selected for multivariate logistic models at threshold 4 mg/l and 10 mg/l levels. Stepwise logistic regression techniques were used to select most significant variables for final multivariate models. Each explanatory variable was entered in the model if it met the significance level of 0.3. The variable that met the criterion of entry (SLENTRY) was again tested to stay in the model at significance level of 0.15 (SLSTAY). Final multivariate models were built at significance level of 0.05 of both Likelihood ratio test and Wald test.

For model at threshold 4 mg/l of nitrate, the result revealed that developed land, fertilizer N, and depth to water table were most significant variables (Table 4-11). The p-value of likelihood ratio test and Wald test were found to be 0.0018 and 0.0052. Similarly, significant variables of multivariate logistic regression model at threshold 10 mg/l of nitrate were percent of clay, fertilizer N, and depth to water table (Table 4-12). All significant variables of the final multivariate logistic model at threshold 4 mg/l were found to be positively correlated while percent of clay and depth to water table were negatively correlated to the occurrence of nitrate exceeding 10 mg/l. The final multivariate logistic regression model at threshold 10 mg/l was significant at p-value ≥ 0.05 (Table 4-12)

Table 4-11. Results of the multivariate logistic regression model at threshold 4 mg/l

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0412	3	0.0018
Wald	12.7432	3	0.0052

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5924	2.2181	4.2868	0.0384
Developed land (dvlp)	1	33.2097	17.8786	3.4503	0.0632
Fertilizer N (nitro)	1	1.2746	0.6883	3.429	0.0641
Depth to WT (dwt)	1	2.0057	1.3181	2.3154	0.1281

H-L Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
10.8835	8	0.2084

Table 4-12. Results of the multivariate logistic regression model at threshold 10 mg/l

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.9512	3	0.0005
Wald	10.8908	3	0.0123

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9949	2.2953	0.1879	0.6647
Clay	1	-4.8939	2.4022	4.1505	0.0416
Fertilizer N (nitro)	1	3.9986	1.3806	8.3889	0.0038
Depth to WT (dwt)	1	-2.0419	1.2656	2.6032	0.106

H-L Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
4.6003	8	0.7993

The Hosmer-Lemeshow (H-L) goodness-of-fit was used to evaluate final multivariate logistic models by comparing average predicted versus observed probabilities of deciles of risk. The H-L p-values of 0.20 and 0.799 at threshold 4 mg/l and 10 mg/l respectively indicated that fitted models were acceptable. The coefficients of determination (R^2 values) between observed and predicted probabilities were 0.72 and 0.79 at thresholds 4 mg/l and 10 mg/l respectively (Figs. 4-39 and 4-40). This indicates that final multivariate logistic regression models predicted probabilities of nitrate exceeding given thresholds very well.

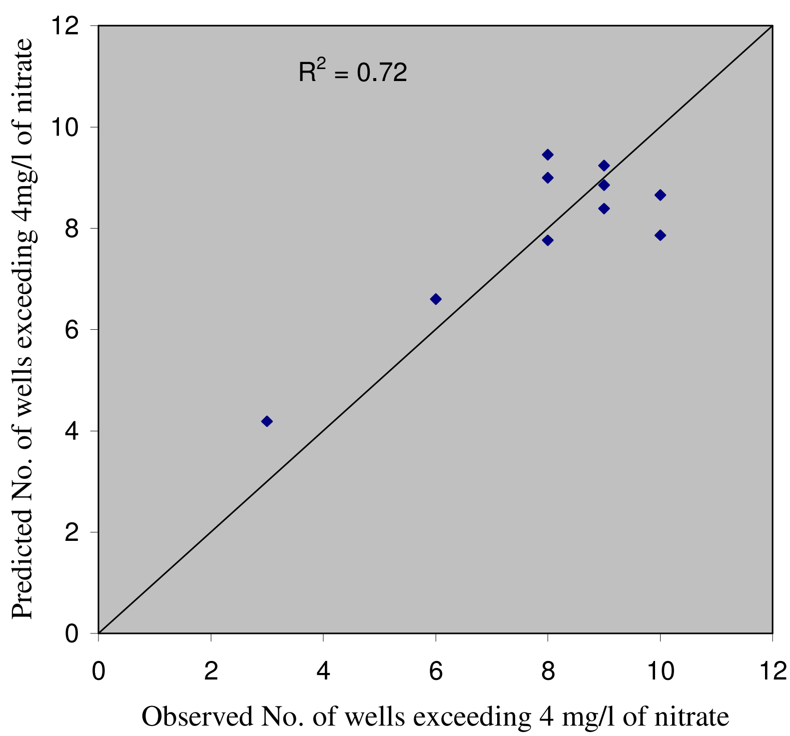


Figure 4-39. Predicted versus observed number of wells exceeding 4 mg/l of nitrate for deciles of risk

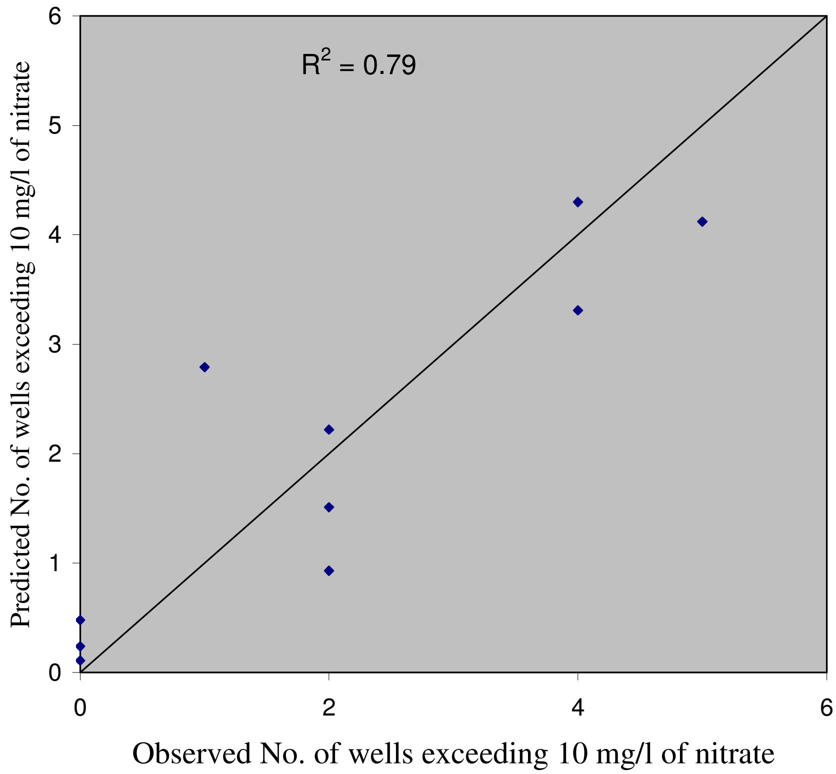


Figure 4-40. Predicted versus observed number of wells exceeding 10 mg/l of nitrate for deciles of risk

4.4 Results of stochastic modeling

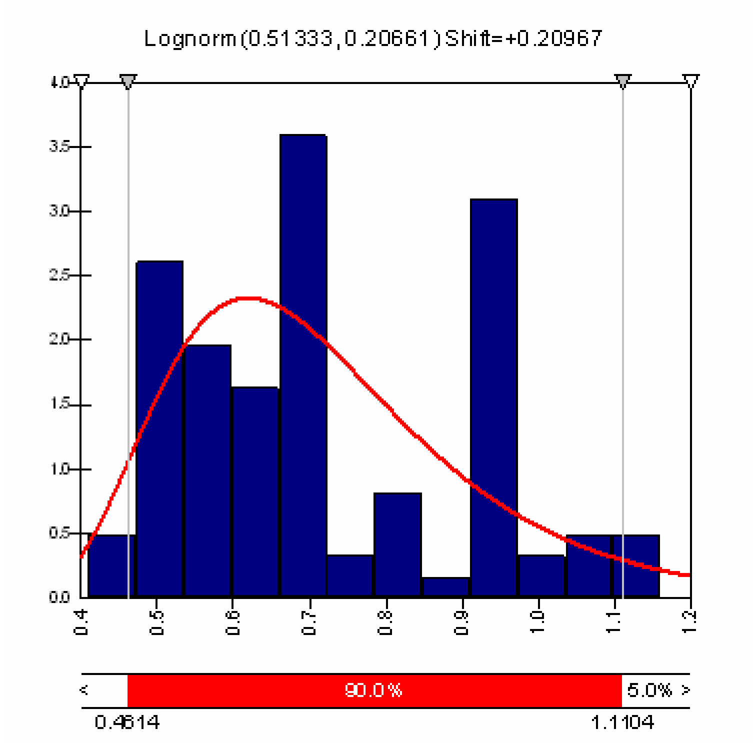
To address the uncertainty of logit model's probability estimates associated with parameters, a Monte Carlo technique was used to propagate parameter uncertainty. Each of the significant variable of logistic regression models that were considered uncertain were selected to fit them to a probability distribution. The @RISK's fitting distribution option was used to find the best fit curves (Palisade, 2005). The results of fitting distribution to all significant variables of logit models are shown in figures 4-41 through 4-44. From the probability distribution curves mean and 95th percentile values were calculated and these values are shown in table 4-13 below.

Estimated coefficients of most significant variables of regression models were replaced by probability distribution functions. Monte Carlo simulation was run for 5,000 iterations to find cumulative probability curves of chance of nitrate occurrence at given thresholds. The cumulative probability of chance of nitrate exceeding 4 mg/l and 10 mg/l in the groundwater are shown in figures 4-45 and 4-46 respectively. Mean probabilities of chance of exceeding 4 mg/l and 10 mg/l were determined as 0.8214 and 0.2581 respectively while 95th percentile probabilities were 0.9679 and 0.8030 respectively.

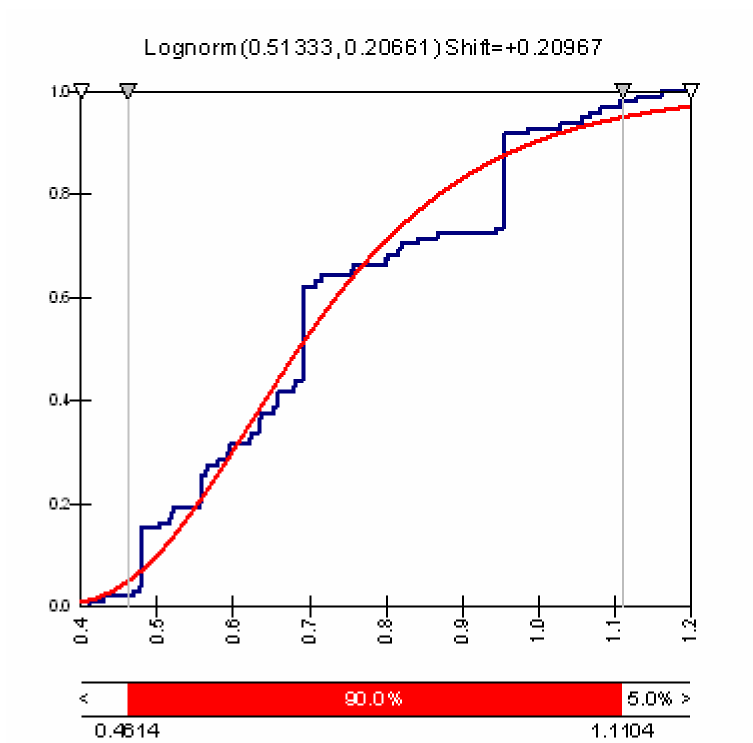
Along with the Monte Carlo simulation, sensitivity analysis method was used to identify important variables, whose uncertainty was a driving factor in the overall uncertainty of risk estimates for nitrate occurrence at given thresholds. A rank correlation coefficient between each input variable and the associated risk output was computed to measure the importance of each parameter to the overall uncertainty. Fertilizer N was found to be the most sensitive variable with 61.1% and 81.7% positive correlation for nitrate contribution in the groundwater at thresholds 4 mg/l and 10 mg/l of nitrate as nitrogen respectively. Other variables were also observed as significant driving factors because Developed land and Depth to water table have positive correlation of 59.9% and 43.3 % for model exceeding 4 mg/l of nitrate in groundwater. On the other hand, Clay was found to be another most influential driving factor with 37.8% negative impact for nitrate occurrence at threshold 10 mg/l while Depth to water table was 18.1% negatively sensitive to changes in output. Results of sensitivity analysis are shown in figure 4-47.

Table 4-13. Fitted distribution of significant variables with 95th percentile values

Variable	Mean	95th percentile	Distribution
Clay	0.5133	1.1104	Lognormal
Developed land	0.0535	0.0847	Logistic
Depth to WT	19.974	1.733	Beta General
Fertilizer N	0.9412	2.067	Beta General

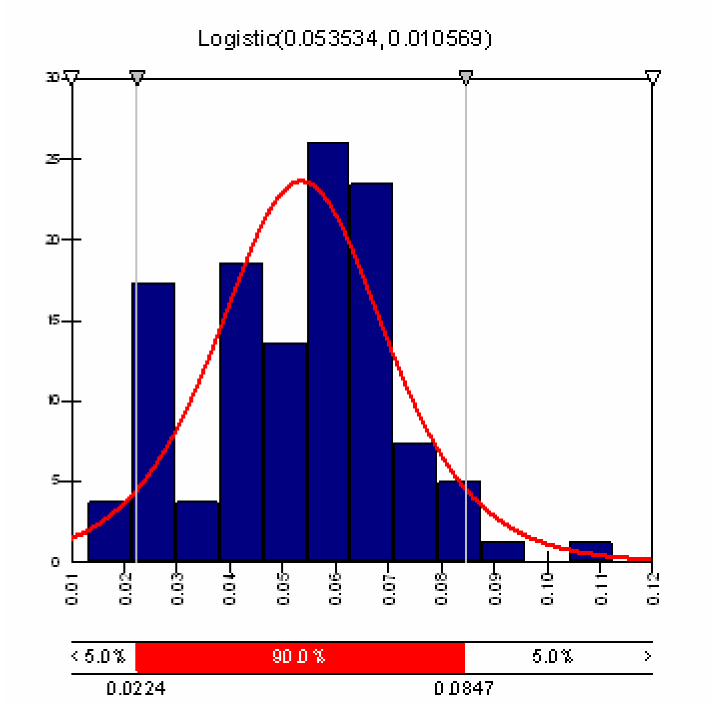


(a)

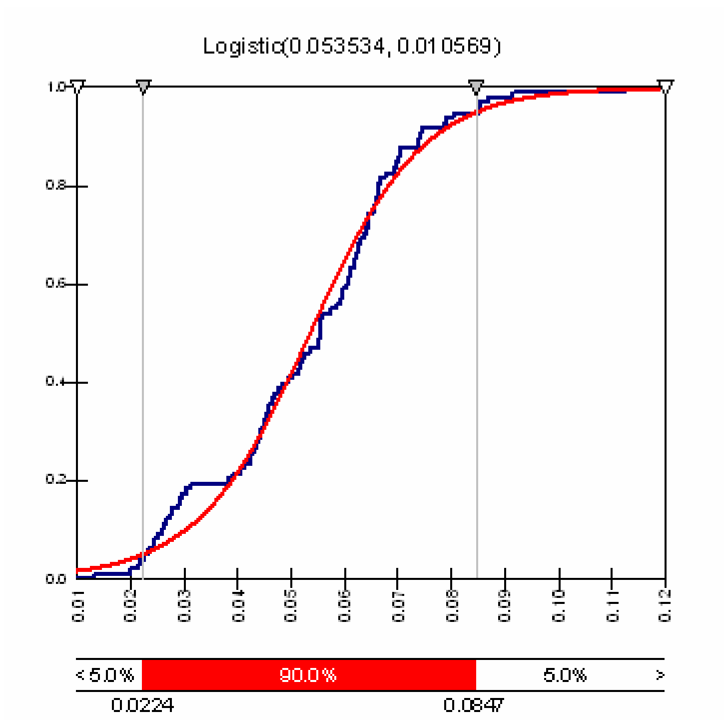


(b)

Figure 4-41. Best fit curve for clay (a) Probability distribution (b) Cumulative distribution

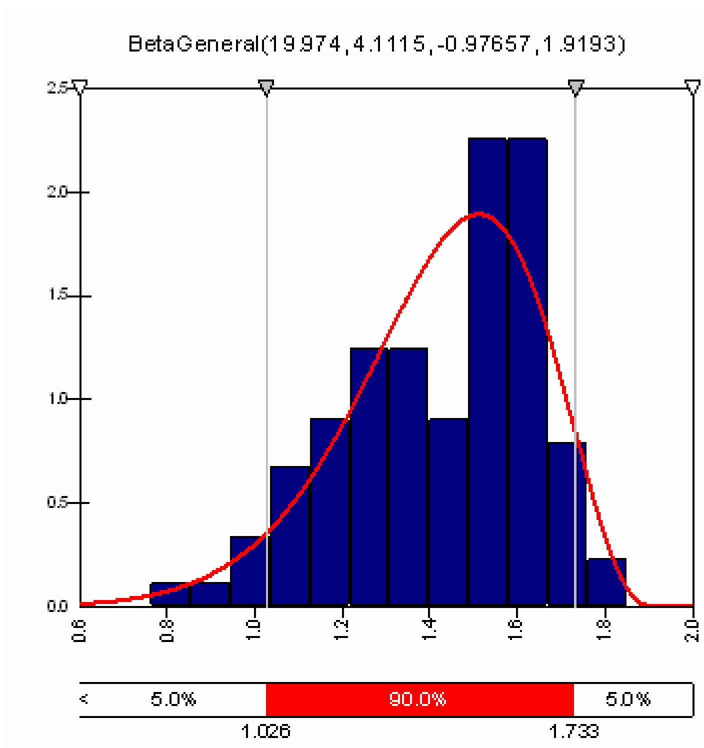


(a)

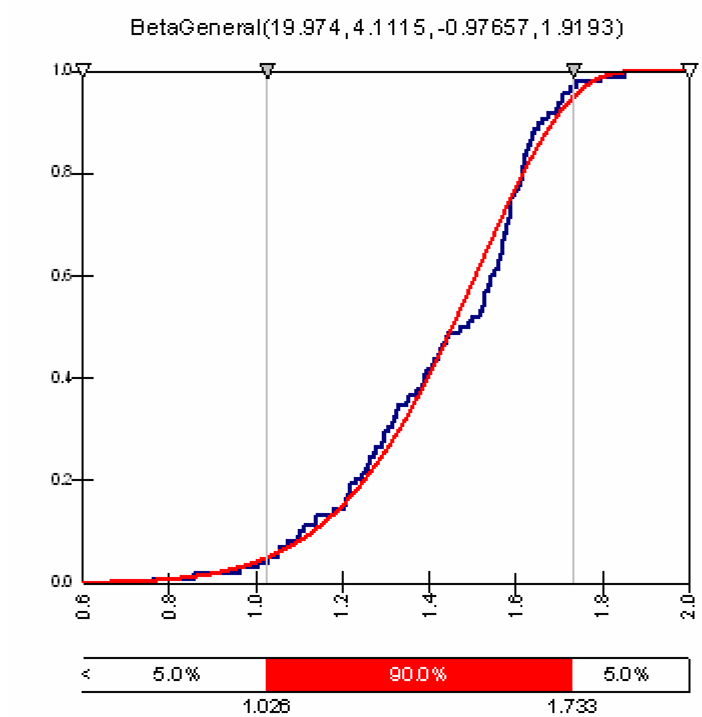


(b)

Figure 4-42. Best fit curve for Developed land (a) Probability distribution (b) Cumulative distribution

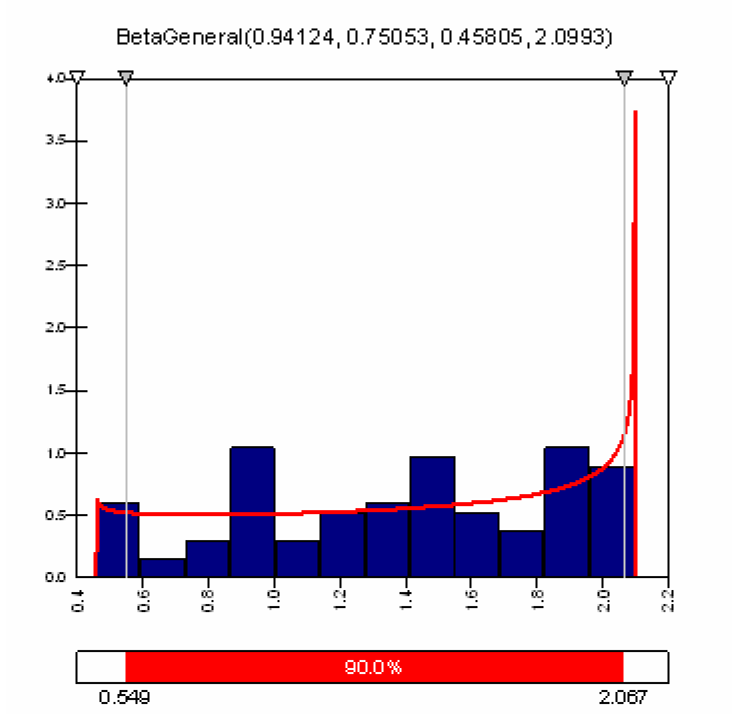


(a)

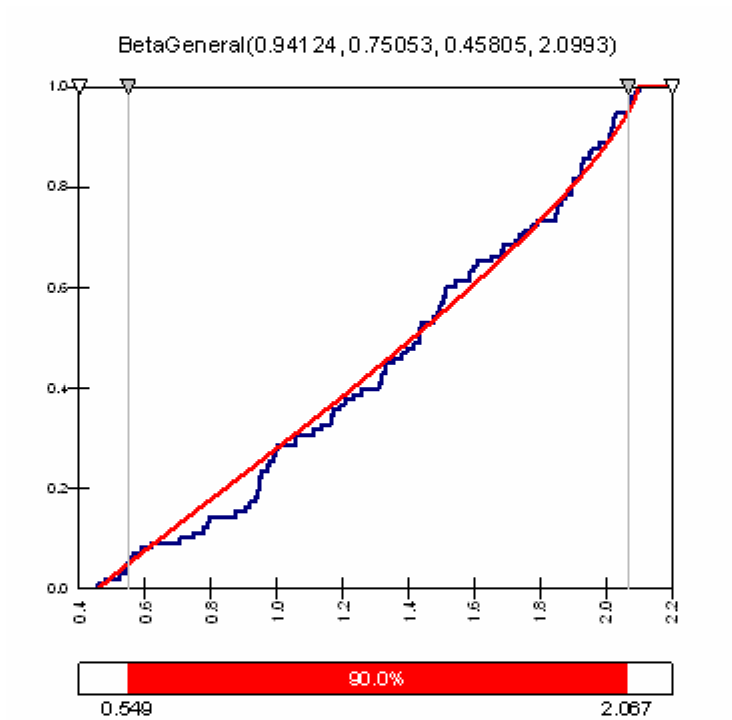


(b)

Figure 4-43. Best fit curve for Depth to WT (a) Probability distribution (b) Cumulative distribution



(a)



(b)

Figure 4-44. Best fit curve for Fertilizer N (a) Probability distribution (b) Cumulative distribution

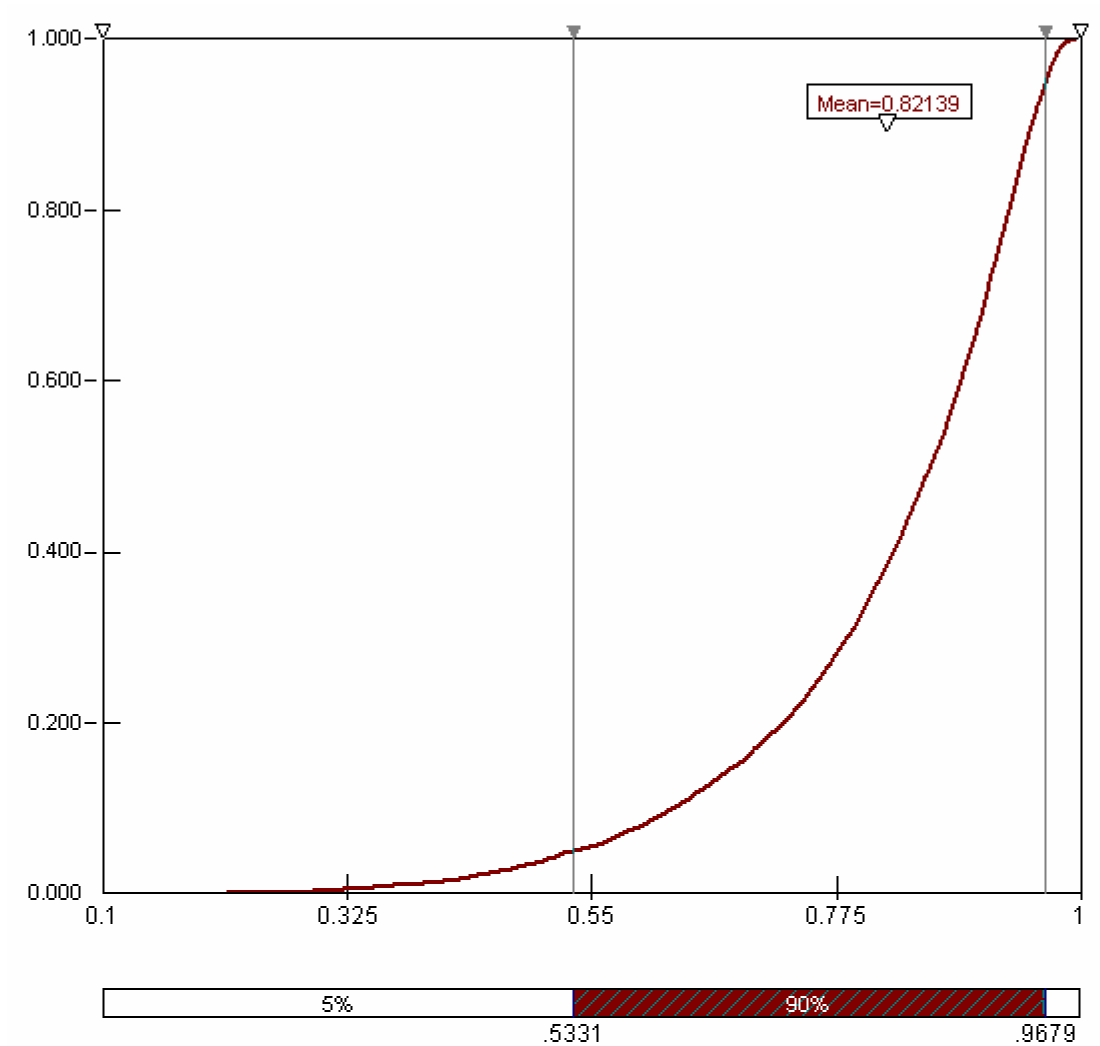


Figure 4-45. Probability curve of chance of occurring nitrate at threshold 4 mg/l

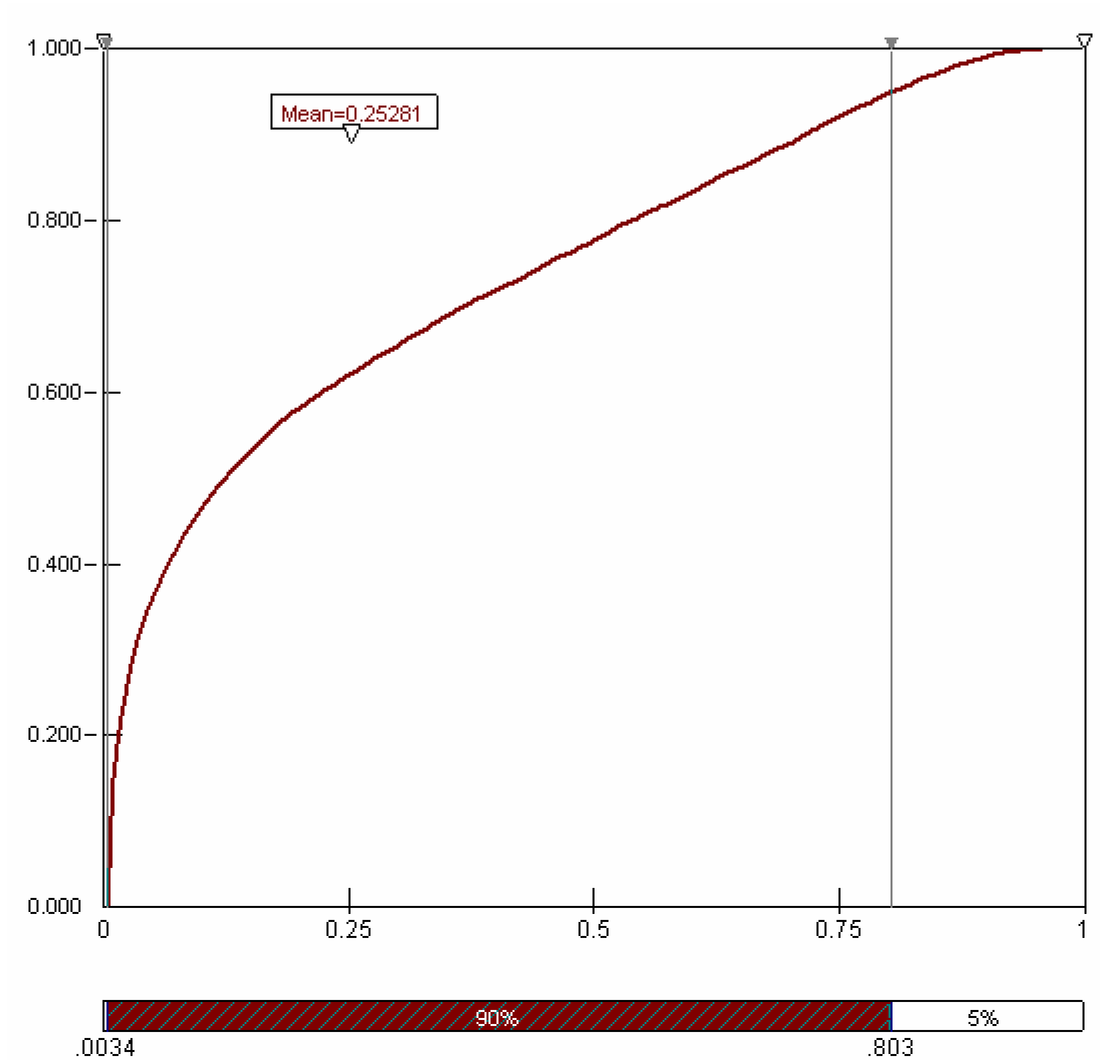


Figure 4-46. Probability curve of chance occurring of nitrate at threshold 10 mg/l

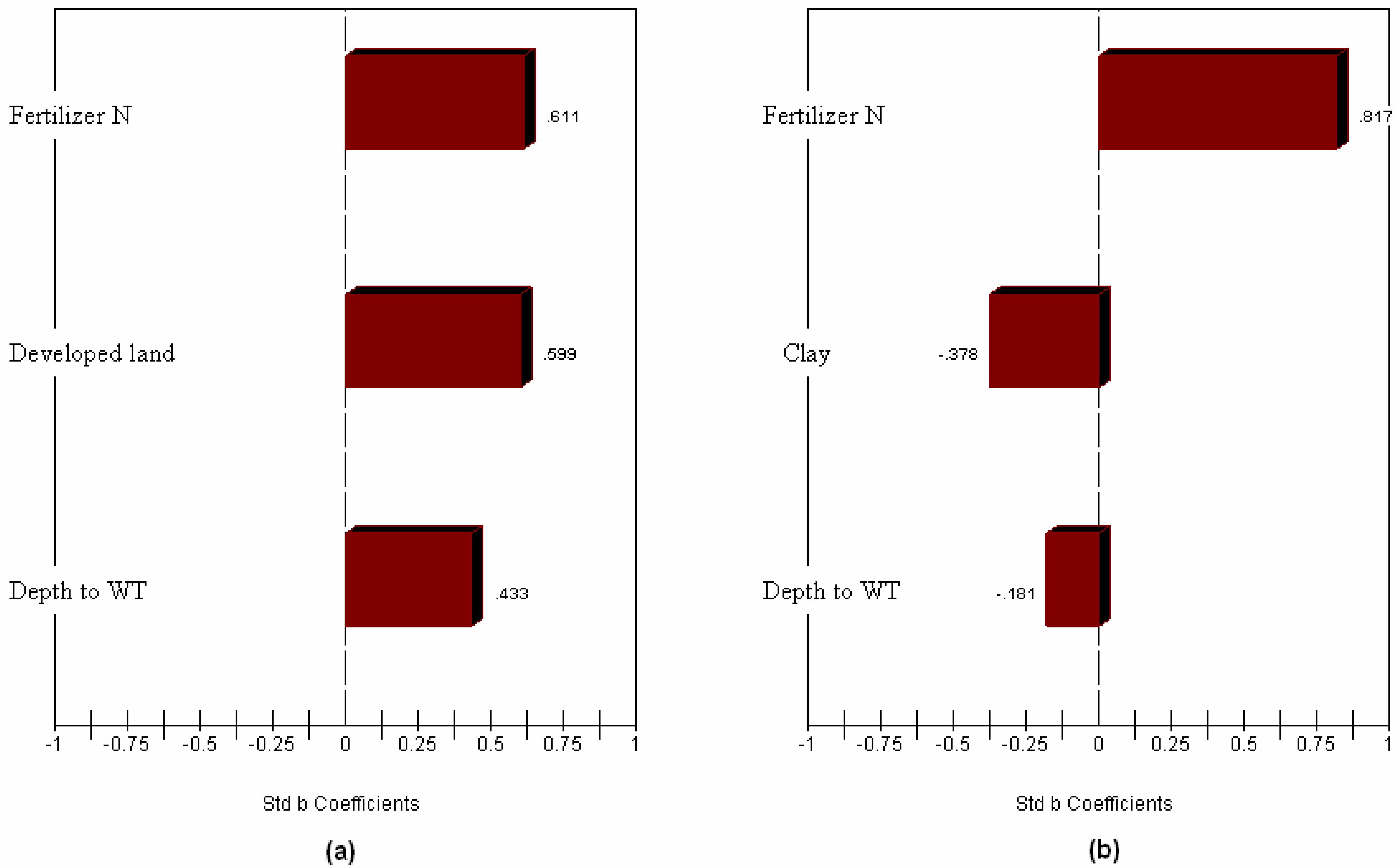


Figure 4-47. Sensitivity analysis of input variables to probability of nitrate contamination exceeding (a) 4 mg/l (b) 10 mg/l

CHAPTER 5

DISCUSSION

This chapter explains the results and discusses the findings with respect to statement of problems. The discussion includes results of nitrate data analysis, geostatistics, logistic regression, and stochastic modeling.

5.1 Discussion on results of nitrate data analysis

Time series plots of nitrate (Figs. 4-1 through 4-4) indicated that the average nitrate concentration in wells, located at wellfields, were fairly constant except that in Ringwood wellfield. Average nitrate concentration in Ringwood wellfield increased by 4 mg/l from the year 1997 to 2005. On the other hand, mean nitrate was found to be statistically different from median nitrate in wells, located at Cleo Spring wellfield, while Ringwood and Cleo Spring wellfields are located in areas dominated by grassland. Out of 18.4% samples that exceeded 10 mg/l, 8% were within Cleo Spring and Ringwood wellfields. This result indicates that the high-density residential growth in these areas has contributed to the nitrate levels to increase. Nitrate concentrations largely varied among wellfields. The sampling data included wells located in the central part of the aquifer which are operated by the City of Enid, Oklahoma. These sampling data used in this study presented an issue of lack of sampling in an appropriate scale to detect variation.

5.2 Discussion on results of geostatistics

Filled contour maps of the average nitrate concentration (Figs. 4-7 and 4-8) in the aquifer shows greater concentration in shallow wells at areas close to Ames wellfield. These sites are located at areas of intensive agriculture and high percentage of sand (50-80%) as sub-surficial geology. Forty eight percent of wells in Ames wellfield exceeded the drinking water criteria set by EPA. This indicates the surface water detention and sub-surface leaching during high rainfall periods. The trend of nitrate increase was consistent with the direction of groundwater movement which was from northwest to southeast over the study period. This is to be expected as nitrate is a highly soluble constituent. The directional anisotropy varied from 300^0 to 360^0 which also resembles the increasing trends of nitrate in the direction of groundwater movement.

Standard error maps (Figs. 4-9 and 4-10) indicated that errors were largely varied throughout the aquifer except within the central part of the aquifer where four wellfields used by the City of Enid are located. Nitrate data used in this study presented an issue of lack of sampling in an appropriate scale to detect variation. Hence, kriging was used as the optimal estimator. Kriging estimation is highly dependent on the number of control points and their proximity to one another.

Maps of areas of statistical difference between 1997 and 2005 were also observed at 95% and 70% confidences (Figs. 4-11 and 4-12.). These statistical maps of nitrate difference were compared with simple difference of kriging estimate. Even the slightest statistically significant change does not go unnoticed while creating maps at 95% confidence interval. At higher confidence level, there is a higher risk of mistakenly concluding that change has occurred when it has not really occurred, and not identifying change where it has really occurred. Hence, to be in a conservative side during trend identification, a lower confidence level may be beneficial. At given thresholds, where the magnitude of kriging variance relative to the magnitude of change that was observed between comparison period (i.e. from 1997 to 2005), a confidence of 70% appeared to be more appropriate that provided a conservative trade-off between the risk to miss potential real change while filtering out a large proportion of the apparent changes portrayed in a simple difference map of kriged concentrations in figure 5-1 below. Areas of significant nitrate change were quantified statistically at 70% confidence (Fig. 4-20) and the area accounted for 7.5% areas of the aquifer.

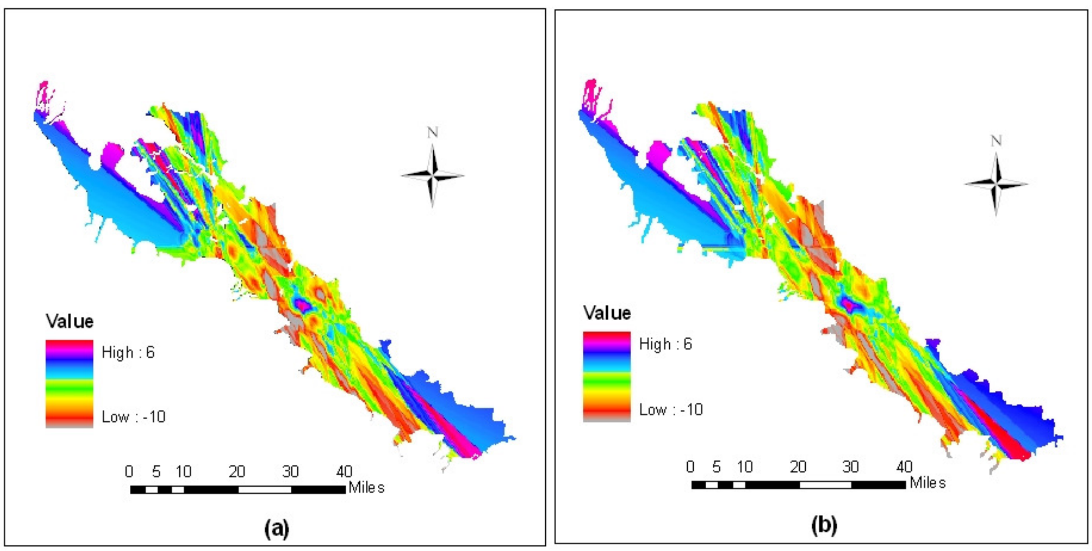
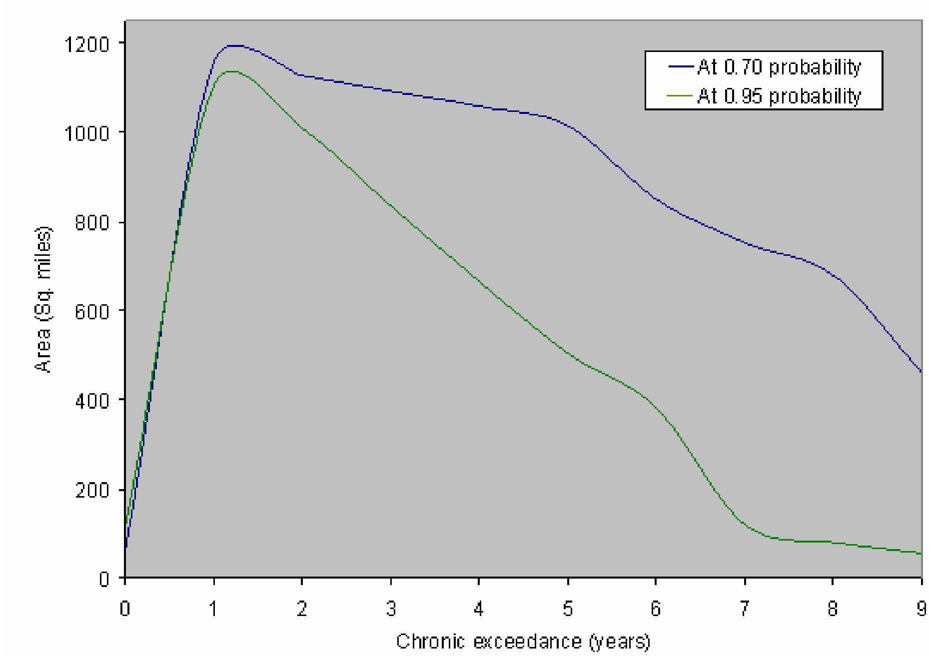


Figure 5-1. Difference in nitrate concentration (mg/l) between 1997 and 2005 (a) at 70% Confidence (b) apparent change (simple difference)

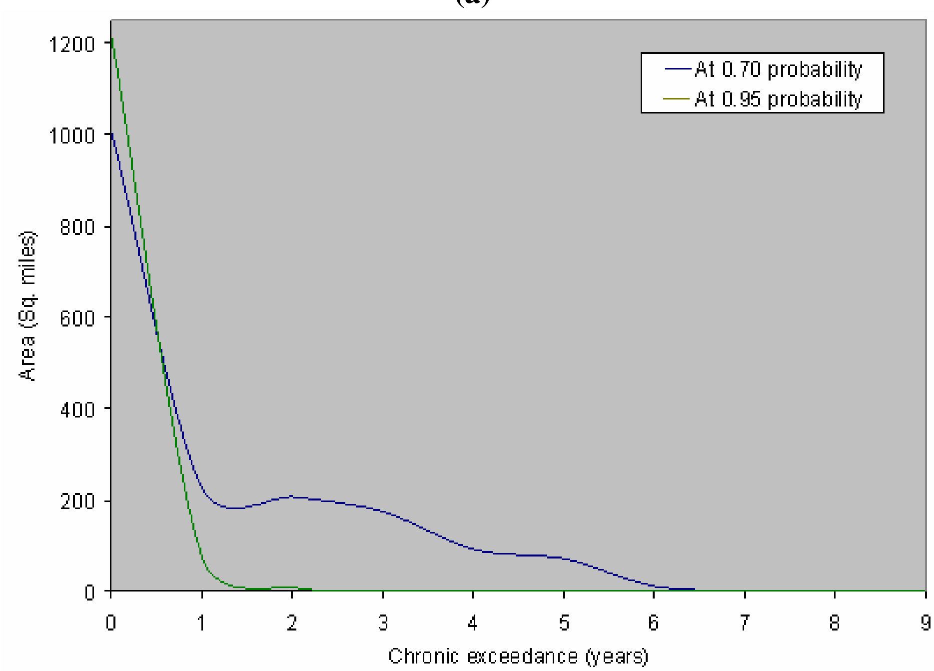
Areas of increasing trend of nitrate at 95% and 70% confidence intervals covered more than 50% of aquifer area. Statistically nitrate increase by at least 4 mg/l from 1997 to 2005 (Figs. 4-16 and 4-17) shows that the central part of the aquifer where wellfields are located, were subject to various activities including increase or change in agricultural and developed lands, and fertilizer N applied to residential and agricultural use etc.

The indicator kriging output, illustrated in figures 4-21 through 4-22, defined the probability of nitrate level being above or below given threshold levels at a given location. These maps can be useful for water managers to build aquifer protection strategies. These probability maps not only portray the uncertainty in contaminant levels but are also useful means to identify locations for additional samples. It was apparent from the figure 4-21 that the probability of nitrate contamination at threshold 4 mg/l was higher in some areas of the aquifer while the probability was low in others. Additional sampling should not be located either in higher or in lower probability areas. However, the additional sampling is worthwhile in areas with intermediate values to better refine the probability regions with very likely (or very unlikely) exceedance of specified thresholds. Similarly, the probability at threshold 10 mg/l of nitrate (Fig. 4-22) was higher in central part of the aquifer, especially around Ames wellfield. This also mimics the intensive agriculture practice used in the past in that region.

A particular probability level for extent and severity of contamination at the field is required for site characterization. This indicates the willingness of the analyst to accept the risk of an incorrect decision based on site characterization and is termed as the probability cutoff or digline (Rautman and Istok, 1996; Smith and Williams, 1996). The probability cutoff for a particular project is determined by various factors such as consequences of this decision on human life, property values, and the environment (Istok and Rautman, 1996). Probability cutoffs or diglines of 0.70 and 0.95 were chosen to define chronic exceedance of nitrates exceeding 4 mg/l and 10 mg/l. The areas of chronic exceedances (Table 4-5) were plotted to observe the effect of cutoff or digline on exceedance of nitrate areas in the aquifer. As the probability increased, the detection of areas of chronic exceedences decreased for particular threshold. The 0.70 and 0.95 probability cutoffs for detecting equal or greater than 4 mg/l and10 mg/l versus chronic exceedance were plotted and are shown in figure 5-2 below. The areas of chronic exceedance at 0.70 probability cutoff at threshold of 10 mg/l were less than that of 0.70 probability cutoff. As the threshold value increased, the probability level decreased, transforming the nature of probability cutoff curve from convex downward to convex upward (Fig. 5-2).



(a)



(b)

Figure 5-2. Probability cutoff curves of chronic exceedance areas at nitrate thresholds (a) 4 mg/l (b) 10 mg/l

5.3 Discussion on results of logistic regressions

Results of logistic regression analysis for establishing a relationship between nitrate and land cover types around sampled wells indicate that the best radius of well influence was 1,000 meters. The likelihood ratio (G-statistic) and Wald statistic were maximized at 1,000 meters of well radius. The analysis of radius of well influence for each mg/l of nitrate was not evaluated; however, results at thresholds of 4 mg/l and 10 mg/l indicates that the best radius of well influence was unchanged regardless of the threshold level that was chosen. Similar type of study was done by Greene et al. (2004) to find the impact of various threshold levels to the best radius of well influence. They also concluded the same result.

The best multivariate logistic models resulted from the variables that were screened from univariate logistic models. The statistical significances of likelihood ratio p-value and Wald p-value were within 5% significance level (Tables 4-11 and 4-12) for logistic models at thresholds 4 mg/l and 10 mg/l of nitrate respectively. Developed land that can be termed as high-density residential in the Cimarron terrace aquifer, and fertilizer N were positively correlated with occurrence of nitrate at threshold 4 mg/l, indicating that increasing values of these variables lead to higher probability of nitrate contamination in the aquifer. The positive correlation of the percentage of high residential land suggests that the sources of nitrates from septic system and fertilizers applied to lawns caused elevated levels of nitrates in the aquifer. Similar results were observed by other investigators (e.g. Tesoriero and Voss, 1997; Gardner and Vogel, 2005). The cropland-pasture was not the significant variable but the fertilizer N load apportioned equally to agriculture land and developed land that accounts for residential fertilizer use was significant for contributing nitrate threshold of 4 mg/l. This result also mimics the fertilizer N applied to residential use and effluent from septic system. Masoner and Mashburn (2004) studied the nitrogen isotopes of 45 wells sampled in the Cimarron terrace aquifer. According to them, of the 28 wells in the agricultural areas, 18 wells were in the mixed sources category (combination of synthetic fertilizer, septic or manure waste sources), and 1 was in the septic source category. Similarly, of the 17 wells in grassland areas, 4 wells were in the mixed category, and 1 well was in septic source category. Percent of clay in the aquifer was negatively correlated with nitrate concentration exceeding 10 mg/l. The percent organic matter that represents dinitrification is generally correlated with clay content. Hence, clay was negatively correlated with nitrates at threshold of 10 mg/l in the aquifer.

The coefficient for depth to water table was positive (2.005) for multivariate logistic model at threshold 4 mg/l of nitrate in the groundwater while negative (-2.041) at threshold of 10 mg/l. The first result indicates that as the depth increases, the nitrate level also increases. Similar observations were noticed by Nolan (2001). The 66% of sampled wells in Cleo Spring and 92% in Ringwood wellfield exceeded the threshold of 4 mg/l at grassland and forest. Land covers in 1992, derived from STATSGO soil database, showed about 7% of pasture/hay area in the aquifer while in 2001 the area of pasture/hay was almost zero (Table 2-1). This evidence also provides the possible denitrification process in the aquifer. The negative slope coefficient for depth to water table from logistic regression at threshold 10 mg/l of nitrate indicates that most of well samples in Ames wellfield (48%) that have exceeded this threshold are located in agricultural lands. These locations are most likely found on high percentage of sand (Fig. 4-32). A plot of nitrate concentrations versus depth to water tables is shown in figure 5-3 below. This scatter plot indicates that nitrate concentration except few outliers increased with depth to water table until the concentration reached about 5 mg/l and after which the relationship between nitrates and the depth to water table became vice-versa. This also indicates that the depth to water table in the Cimarron terrace aquifer is playing as double characters with the occurrence of various nitrate levels.

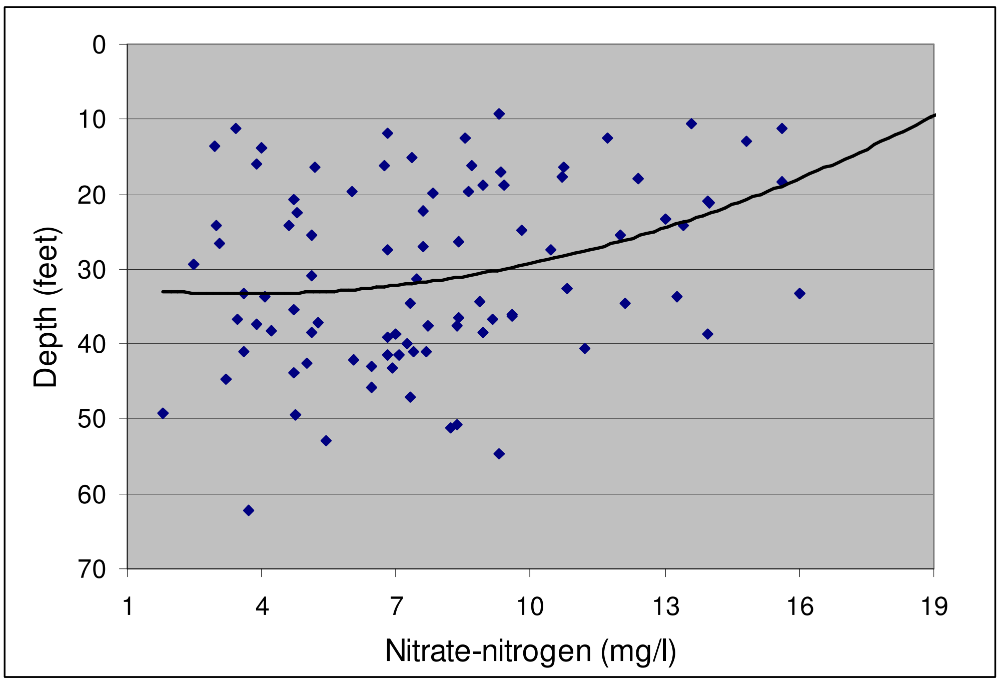


Figure 5-3. Relationship between groundwater nitrate concentrations and depth to water table

Two multivariate logistic models at thresholds 4 mg/l and 10 mg/l of nitrate were evaluated using Hosmer-Lemeshow (H-L) goodness-of-fit by comparing predicted versus observed probabilities of deciles of risk. The coefficients of determination (R^2) were 0.72 and 0.79 for final models at thresholds 4 mg/l and 10 mg/l respectively, indicating that the models fit the data well (Figs. 4-39 and 4-40). H-L p-values of 0.20 and 0.79 for models exceeding 4 mg/l and 10 mg/l of nitrate respectively indicate that the models are acceptable. The global null hypothesis was that models fit the data and hence higher p-values indicate a better fit.

5.4 Discussion on results of stochastic modeling

The variability for a parameter can be represented as a probability distribution function (pdf) or a cumulative distribution function (cdf). The pdf shows the likelihood that the value for a random sample will occur within a very small interval. The shape of variables as pdf can greatly affect the outcome of a Monte Carlo analysis and hence, an appropriate shape should carefully be selected. The variability of variables such as clay, developed land, fertilizer N, and depth to water table determined by using @RISK's fit distribution option were represented as lognormal, logistic, beta general , and beta general respectively. Mean and 95% likelihood value of these variables are shown in figures 4-41 through 4-44.

The benefits of using Monte Carlo simulation over deterministic model is that a single value for each of the model's input parameter is used to calculate a single output parameter in deterministic model while in Monte Carlo method, each of the parameter is assigned a distribution (pdf or cdf). The output of the model is calculated many times based on a new random value selected from the probability distribution for each of the input parameter each time. The probability curve of output is generated and the probability of occurrence of any particular value can be calculated. The outputs of Monte Carlo simulation are illustrated in figures 4-45 through 4-46 for models exceeding 4 mg/l and 10 mg/l of nitrate respectively. From these cumulative distributions, the probability of nitrate occurrence at specified thresholds can be determined as less than or equal to specified value. Mean value and 95 % chance of occurring nitrate at thresholds 4 mg/l were 0.8214 and 0.9679 respectively. Similarly, values at threshold 10 mg/l of nitrates were 0.2581 and 0.8030 respectively. This indicated that as the nitrate threshold level increased the probability of occurring nitrate in the groundwater decreased. Slopes of Monte Carlo output, the cumulative distribution curves, were compared. For model exceeding 4 mg/l, as the likelihood of occurrence of nitrate increased, the probability curve changed to steeper, indicating an increase of uncertainty of model with the increase in probability levels. For model at threshold 10 mg/l, the effect was just vice-versa to that of model exceeding 4 mg/l.

Sensitivity analysis of models to changes in specific parameters about which there is a high degree of uncertainty was performed. Fertilizer N was the most sensitive (61.1%) variable along with developed land (59.9%) that can be termed as high-density residential land in the Cimarron terrace aquifer and the depth to water table (43.3%) for model exceeding 4 mg/l of nitrate in the groundwater. Similarly, Fertilizer N was the most sensitive variable (81.7%) along with clay (-37.8%) for model exceeding 10 mg/l of nitrate. Clay was found to be negatively correlated with occurrence of nitrate exceeding Drinking Water Criteria (i.e. 10 mg/l) set by EPA as already explained in logistic regression analysis. The depth to water table was found to be negatively correlated (-18.1%) with the output exceeding 10 mg/l of nitrate.

CHAPTER 6

CONCLUSIONS

This chapter summarizes the findings of geostatistical, statistical, and stochastic methods used in this research, with respect to the problem statement and general focus of the investigation. Some recommendations have also been suggested in the latter section of this chapter.

6.1 Conclusions

These results led to the following conclusions:

1. Results of trend analysis of nitrates revealed that the average concentration of nitrate in Ames, Cleo spring, and Drummond wellfields were fairly constant from 2001 to 2005 while for Ringwood wellfield there were increasing trend of average nitrate over the same time period. Ames wellfield was the most severely affected wellfield with average concentration of 8.6 mg/l.
2. This research has attempted to predict the spatial distribution and uncertainty of groundwater nitrate in the Cimarron terrace aquifer. Spatial nitrate maps of ordinary kriging showed high levels of nitrate in and around the Ames wellfield and eastward from this wellfield. Average nitrate concentrations exceeding Drinking Water Criteria, 10 mg/l nitrate as nitrogen, were mapped in the aquifer.

3. Quantitative, statistical, and means of kriged blocks of nitrates defining the statistical significance of geographic apparent changes between 1997 and 2005 were assessed using GIS and some statistical hypothesis tests. Most of statistical changes of nitrates were found in central part of the aquifer. Increasing and decreasing trends at 70% and 95% confidences were mapped and areas were calculated. At 95 % and 70% significances, areas of statistically significant changes in nitrate are shown in figures 4-19 and 4-20. At higher confidence, 95%, smaller area, only 0.4% of aquifer area was found while for confidence 70%, the area of nitrate changes was 7.5% of the aquifer area. Areas of increasing trend were found to be more than one-half of the aquifer area.
4. This research has presented an approach for probabilistic assessment of groundwater nitrate that can explicitly be used in site characterization. This approach is based on use of indicator kriging to predict whether the concentration would exceed 4 mg/l and 10 mg/l of nitrate in the groundwater.
5. Maps of chronic exceedance frequencies of nitrate were generated for 0.70 and 0.95 probability cutoffs or diglines by combining indicator probability maps of 9 years. This approach can be used as a tool for delineating nitrate management areas that are statistically meaningful.
6. Logistic regression was used to establish the relationship between nitrate and the land cover types around sample wells. The best radius of well influence so that the land use could show significant effects on nitrate concentration in wells was determined to be 1,000 meters. This area of well influence did not change regardless of the threshold level that was chosen.

7. Stepwise logistic regression approach revealed that the percentages of Developed land or that can be termed as high-density residential land, Fertilizer N, Depth to water table were significant predictors of groundwater nitrate concentration in excess of 4 mg/l while percent of Clay, Fertilizer N, Depth to water table were significant predictors in excess of 10 mg/l.
8. This investigation has attempted to address the variability and uncertainty of significant variables of logit model and the model itself. This approach is based on Monte Carlo simulation. The probability curve of Monte Carlo output was generated based on assigned distribution (cdf) of each significant variable. These probability curves can be used to calculate the likelihood of occurrence of any particular value (chance of occurring nitrate) as less than or equal to specified value.
9. From Monte Carlo simulation, sensitivity analysis of models to changes in specific parameters about which there was a high degree of uncertainty was performed. The Fertilizer N was the most sensitive (61.1%) along with depth to water table (59.9%) and fertilizer N (43.3%) for model exceeding 4 mg/l of nitrate in the groundwater. Similarly, Fertilizer N was the most sensitive variable (81.7%) along with clay (-37.8%) and depth to water table (-18.1%) for model exceeding 10 mg/l of nitrate.

6.2 Recommendations

Following recommendations are made from this research:

1. This research used nitrate samples of highly clustered wells in four wellfields of the Cimarron terrace aquifer. As mentioned earlier, the radius of well influence in the aquifer was 1,000 meters. Hence, it is recommended that monitoring wells should be selected or established at least 1,000 meters of spacing throughout the aquifer to perform fully convincible statistical and probabilistic assessments. Result of semi-variogram fitting, the ‘Range’ revealed that nitrates were spatially correlated within 1,500 meters to 6,000 meters from the year 1997 through 2005.
2. Spatial distribution or probability maps of nitrates of this research can be utilized in prioritizing implementation of nitrate management areas and also to establish further sampling locations.
3. The sustainable approach to nutrients management must be adopted, and it must be ensured that agricultural activities do not degrade the groundwater quality. All efforts to reduce fertilizer N in residential and agricultural use should be encouraged.
4. The sub-surficial geology of the Cimarron terrace aquifer taken from STATSGO database indicated that the porosity varied from 0.4 to 1. Porosities greater than 0.4 is regarded as a fracture or fissure. Because of this reason it was not possible to take this variable as categorical value in logistic regression. This might be the possible cause of elevated nitrates in Cleo Spring and Ringwood wellfields where these wellfields are located in areas dominated by grasslands. Most of wells have exceeded the concentration of 4 mg/l of nitrate in this area. On the other hand, the

clays can crack when they shrink after wetting periods and help nitrates to leach in deep groundwater. Area of mixed-grass prairie that covered the Cimarron terrace aquifer in the past could play the important role. The deep, interconnected root holes of prairie grasses prevent surface run-off and hence helped to increased infiltration. It is recommended that combined maps of high clay with some field surveys late in the dry portions of the year can be used to determine if there are areas more prior to fracturing.

REFERENCES

Adams, G. P., and Bergman, D. L., 1996. Geohydrology of alluvium and terrace deposits of the Cimarron river from Freedom to Guthrie, Oklahoma: U.S. Geological Survey Water-Resources Investigations Report 95-4066, 57 p.

Allard, D., 1998. Geostatistical classification and class kriging: Journal of Geographic Information and Decision Analysis, v.2, no.2, p.77-90.

Aller, L., T. Bennett, J. H. Lehr, and J. D. Petty, 1985. DRASTIC- A standardized system for evaluating groundwater pollution potential using hydrogeologic settings: U.S. EPA Report EPA/600/2-85/018, p.163.

Atkins, P. M. and C. D. Lloyd, 1998. Mapping precipitation in Switzerland with ordinary kriging and indicator kriging: Journal of Geographic Information and Decision Analysis, v.2, no.2, p.65-76.

Bosch, H., A. Rosefield, R. Houston, H. Shipman, F. Woodward, 1950. Mathemoglobinemia and Minnesota well supplies: Journal of the American Water Works Association, v.42, p.161-170.

Boyce, J. S., J. Muir, A. P. Edwards, E. C. Seim, and R. A. Oslor, 1976. Geologic nitrogen in Pleistocene loess of Nebraska: Journal of Environmental Quality, v. 5, no.1, p.93-96.

Burkart, M. R. and J. D. Stoner, 2002. Nitrate in Aquifer beneath Agricultural Systems: Water Science and Technology, v. 45, no.9, p.19-29.

Buttafuoco, G., A. Castrignano, and M. Stelluti, 2000. Accounting for local uncertainty in agricultural management decision making: 7th ICCTA International Congress for computer Technology in agriculture, Florence, Nov. 15-18, 1998, p. 510-517.

Cain, D., D. R. Helsel, and S. E. Ragone, 1989. Preliminary evaluations of regional ground-water quality in relation to land use: Ground Water, v.27, no.2, p.230-244.

Castrignano, A., M. Mazzoncini, L. Giugliarini, and R. Risality, 1999. A geostatistical approach to management decision making in alkali soil improvement: Italian Journal of Agronomy, v.2, no.2, p.133-140.

Cinnirella, S., G. Buttafuoco, and N. Pirrone, 2005. Stochastic analysis to assess the spatial distribution of groundwater nitrate concentrations in the Po catchment (Italy): Environmental Pollution, v.33, p.69-580.

Comly, H. H., 1945. Cyanosis in infants caused by nitrates in well water: Journal of American Medical Assiciation, v. 129, p.112-116.

Cooper, R. M., and J. D. Istok, 1988a. Geostatistics applied to groundwater contamination. I. Methodology: Journal of Environmental Engineering, ASCE, v.114, no.2, p.270-286.

Cooper, R. M., and J. D. Istok, 1988b. Geostatistics applied to groundwater contamination. II. Application. Journal of Environmental Engineering, ASCE, v.114, no.2, p.287-299.

Critchfield, G. C., and K. E. Willard, 1986. Probabilistic analysis of decision trees using Monte Carlo simulation: Medical Decision Making, v.6, p.85-92.

David, M., 1977. Geostatistical ore reserve estimation: Elsevier Scientific Publishing Company, Ney York, NY. p.364.

Doubilet, P., C. B. Begg, M. C. Weinstein, and P. Braun, 1995. Probabilistic sensitivity analysis using Monte Carlo simulation- A practical approach: Medical Decision Making, v.5, p.157-177.

ESRI, 2006. ArcGIS 9.1, Environmental Systems Research Institute, Redlands, CA.

Fishman, G. S., 1996. Monte Carlo-Concepts, algorithms, and applications: Springer-Verlag New York, Inc.

Gardner, K. K. and R. M. Vogel, 2005. Predicting ground water nitrate concentration from land use: Ground Water, v.43, no.3, p.343-352.

Goovaerts, P., 1994. Comparative performance of indicator algorithms for modeling conditional probability distribution functions: Mathematical Geology, v.26, no.3, p.389-411.

Goovaerts, P., 1997. Geostatistics for natural resources elevation: Oxford University press, UK.

Greene, E. A., A. E. LaMotte, and K. Culliran, 2004. Groundwater vulnerability of nitrate contamination at multiple thresholds in mid-Atlantic region using spatial probability models: USGS, scientific investigation report 2004-5118.

Halvorson, J. J., L. Jeffery, R. Smith, and I. Papendic, 1996. Integration of multiple soil parameters to evaluate soil quality- A field example: Biology and Fertility of Soil, v. 21, no.3, p.207-214.

Hamerlinck, J. D., and C. S. Ameson, 1998. Wyoming ground water vulnerability assessment handbook: University of Wyoming, Report SDVC 98-01, < http://www.wygisc.uwyo.edu/groundwater/vol2_front.pdf>.

Hay, L. E., and W. A. Battaglin, 1990. Effects of land use buffer size on Spearman's rank partial correlations of land use and shallow groundwater quality: U.S. Geological Survey Water-resources Investigations Report 89-4163.

Helsel, D. R. and R. M. Hirsch, 2002. Statistical methods in water resources: U. S. Geological Survey, Techniques of Water-Resources Investigations Book 4, Chapter A3, p.395.

Hendry, M. J., R. G. L. Macready, and W. D. Gould, 1984. Distribution, source and evolution of nitrate in glacial till in southern Alberta, Canadian: Journal of Hydrology, v.70, p.177-198.

Hong, N, J. G. White, R. Weisz, C. R. Crozvex, M. L. Gumpertz, and D. K. Cassel, 2006. Remote Sensing-informed variable- rate nitrogen management of wheat and corn: Agronomy and groundwater outcomes: Journal of Agronomy, v.98, p.327-338.

Hosmer, D. W. and S. Lemeshow, 2000. Applied logistic regression, second edition: A Wiley-Interscience Publication, New York.

Issaks, E. H and R. M. Srivastava, 1989. An Introduction to applied geostatistics: New York, oxford University Press.

Istok, J. D, J. D. Smith, and A. L. Flint, 1993. Multivariate geostatistical analysis of groundwater contamination. A case history: Ground Water, v.31, no.1, p.63-74.

Istok, J. D. and C. A. Rautman, 1996. Probability assessment of ground-water contamination: 2 results of case study: Ground Water, v.36, no.6, p.1050-1064.

Journel, A. G. and C. J. Huijbregts, 1978. Mining geostatistics. New york, N.Y.: Academic press.

Kaplan, E. and W. F. McTernan, 1993. Overview of the Risk Assessment Process in Relation for Groundwater Contamination: The Environmental Professional, v.15, p.334-340.

Kapoor, A., and T. Viravaghavan, 1997. Nitrate removal from drinking water- review: Journal of Environmental Engineering, v.123, no.4, p. 371-380.

Kitanidis, P. K., 1997, Introduction to Geostatistics, Applications to Hydrogeology: Cambridge University Press, p. 249.

Koterba, M. T, W. L Banks, and R. J. Shedlock, 1993. Pesticide in shallow ground water in the Delmarva Peninsula: Journal Environmental Quality, v.22, no.3, p.500-518.

Krige, D. G, 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand: Journal of the Chemical, Metallurgical and Mining Society of South Africa, v.56, no.6, p.119-139.

Krivoruchko, K., 2001. Using linear and non-linear kriging interpolators to produce probability maps: Report presented at annual conference of the International association for mathematical geology, Cancun, Mexico, p.16.

Lamotte, A. E. and E. A. Greene, 2007. Spatial analysis of landuse and shallow groundwater vulnerability in the watershed adjacent to Assateague Island National Seashore, Maryland and Virginia, USA: Environmental Geology, v.52, p.1413-1421.

Lanza, L. G., J. A. Ramirez and E. Todini, 2001. Stochastic rainfall interpolation and downscaling: Hydrology and Earth System Sciences, v.5, no.2, p.139-143.

Lemeshow, S., D. Teres, J. S. Avrunin, and H. Pastidates, 1988. Predicting the outcome of intensive care unit patients: Journal of the American Statistical Association, v.83. p.348-356.

Livingston, M. L. and D. C. Corey, 1998. Agricultural Nitrate Contamination of Groundwater: An evaluation of environmental policy: Journal of the American Water Resources Association, v.34, no.6, p.1311-1317.

Lynch, S. D., A. G. Reynders, R. E. Schule, 1994. Preparing input data for a national-scale groundwater vulnerability map of Southern Africa: Water SA, v.20, no.3, p. 239-246.

Masoner, J. R., and Mashburn, S. L., 2004. Water quality and possible sources of nitrate in the Cimarron terrace aquifer, Oklahoma, 2003: U.S. Geological Survey Scientific Investigation Report 04-5221.

McKay, P. M and J. A. Cherry, 1989. Groundwater Contamination: Pump and Treat Remediation: environmental Science and Technology, v.23, no.6, p.630- 636.

McTernan, W. F. and B. V. Bonnett, 2002. Using Artificial Neural Network models to determine contaminant sources: Oklahoma State University, Report EA TIET 02-003.

Minitab Inc., 2003, Minitab version 14, USA.

Moore, S. F. and D. B. McLaughlin, 1980. Mapping contaminated soil plumes by kriging: Proceedings of U.S EPA, National Conference on Management of Uncontrolled hazardous wastes, Washington, DC.

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

Nolan, B. T., 2001. Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States: Ground Water, v. 39, no.2, p.290-299.

Nolan, B. T., and J. D. Stoner, 2000. Nutrients in ground waters of the conterminous

Nolan, B. T., B. C. Ruddy, K. J. Hitt, and D. R. Helsel, 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., K. J. Hitt, and B. C. Ruddy, 1998. A national look at nitrate contamination in ground water: Water Conditioning and Purification, v.39, p.76-79.

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

Palisade Corporation, 2005. @Risk, http;//www.palisade.com. Newfield, New York.

Rautman, C. A. and J. D. Stock, 1996. Probability assessment of ground-water contamination: 1 results of case study: Ground Water, v.34, no.5, p.899-909.

Reed, E., J. Mogg, J. Barclay, and G. Penden, 1952. Groundwater resources of the terrace deposits along the northeast side of the Cimarron river, Alfalfa, Garfield, kingfisher, and major counties, Oklahoma, Oklahoma and planning and resources Board, Division of water resources bulletin no. 9: In Reely, B. T., 1992, a linked optimization-simulation aquifer management model, Ph.D. Dissertation, Oklahoma State University, Stillwater, Oklahoma, p. 96.

Reely, B. T., 1992. A linked optimization-simulation aquifer management model, Ph.D. Desertation, Oklahoma State University, Stillwater, Oklahoma, p.94.

Rupert, M. G, 2001. Calibration of the DRASTIC groundwater vulnerability mapping method: Ground Water, v.39. no.4, p.625-630.

Rupert, M. G., 2003. Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate in groundwater in Colorado: U.S. Geological Survey Open File Report 03-4269.

SAS Institute, 2006. SAS online 9.1.3, SAS Inst., Cary, NC.

Scanlon, B. R., R. C. Reedy, and K. S. Kier, 2003. Evaluation of nitrate contamination in major porous media aquifers in Texas: University of Texas at Austin, final report submitted to Texas Commission on Environmental Qualify.

Shukla, S., S. Mostaghimi, V. O. Shanholt, M. C. Collins, and B.B Ross, 2000. A county-level assessment of groundwater contamination by pesticides: Ground Water Monitoring and Review, v. 20, no. 1, p. 104-119.

Smith, L. S., and R. E. Williams, 1996. Examination of methods for evaluating remining a mine waste site. Part II. Indicator kriging for selective remediation: Engineering Geology, v.43, p.23-30.

Solley, W. B, R. R. Pierce, H. A. Perlman, 1995. Estimated use of water in the United States in 1990: U.S. Geological Survey Circular 1200.

Spalding, R. F., and M. F. Exner, 1993. Occurrence of nitrate in groundwater- A review: Journal of Environmental Quality, v.22, no.3, p.392-402.

Squillance, P. J., J. C. Scott, M. J. Morgan, B. T. Nolan, and D. W. Kolpin, 2002. VOCs, pesticides, nitrate, and their mixtures in groundwater used for drinking water in the United States: Environmental Science Technology, v.36, no.9, p.1923-1930.

Storm, D. E., M. White, and S. Stoodley, 2003. Stillwater creek modeling and land cover classification, Oklahoma State University: A Report submitted to Oklahoma Conservation Commission.

Storm, D. E., G. David, R. Tejral, P. Kenkel, and M. S. Gregory, 2000. Estimating watershed level nonpoint source loading for the state of Oklahoma: A report submitted to U.S. EPA, OCC TASK #78 - FY 1996 319(h) TASK #210 - Output #3.

Stumm,W. and J. J. Morgan , 1996. Aquatic Chemistry, 3rd Edition: Wiley Interscience, NY.

Teso, R. R, M. P. Poe, T. Younglove, and P. M. McCool, 1996. Use of logistic regression and GIS modeling to predict ground vulnerability to pesticides: Journal of Environmental Quality, v.25, p.425-432.

Tesoriero, A. J. and F. D. Voss, 1997. Predicting the probability of elevated nitrate concentrations in the Puget sound basis: Implication for aquifer susceptibility and vulnerability: Ground Water, v.35, no.6, p.1029-1039.

Thompson, K. M., D. E. Burmaster and E.A. Crouch, 1992. Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments: Risk Analysis, v. 12, p.53-63.

U.S. Department of Agriculture, 1996. Extension toxicology network, pesticide information profiles: In Masoner, J. R. and S. L.Mashburn, 2004, Water quality and possible sources of nitrate in the Cimarron Terrace aquifer, Oklahoma: USGS, Scientific Investigative Reports 2004-5221.

U.S. EPA, 1996. Drinking water regulations and health advisories, Washington, D.C., U.S. EPA, Office of water: US Government printing office.
United States: Environmental Science and technology, v. 34, no.7, p.1156-1165.

Ward, M. H, S. D. Mark, K. P. Cantor, D. D. Weisenburger, A. Correa- Villasenor, and , S. H. Zahm, 1996. Drinking Water Nitrate and Risk of Non-Hodgkin's Lymphoma: Epidemiology, v.7, p. 465-471.

APPENDIX A

NITRATE DATA

Table A-1. Annual average nitrate as nitrogen (mg/l)

Well ID	1997	1997	1997	1997	1997	1997	1997	1997	1997
D1		8.2				7.0	3.6		4.0
D2			7.3	8.0		9.2	8.2		
D3	8.0	9.5	9.5	8.1	9.0	9.1	8.9	9.0	
D5	6.1	9.4	9.7	8.7	9.0	9.3	9.8	10.1	10.0
D6	5.4	6.9	7.0	6.4	6.9	7.1	6.8	7.0	7.3
D7	8.9	13.2	14.4	12.6	13.1	11.7	12.0	11.8	11.6
D8	10.4	12.9	12.4	11.4	10.7	10.8	10.5	10.5	
D9	1.6	1.6	1.5	1.5	2.9	3.3	3.1	2.7	2.9
D10	3.7	5.3	5.3	4.7	5.5	5.8	4.9	4.9	4.7
D12	4.8	5.5	5.5	5.3	5.7	5.5	5.4	5.6	5.8
D18	6.5	9.1	9.4	8.5	9.6	8.5	8.9	9.1	8.8
D19	3.7	4.5	4.8	4.9	5.4	6.4	6.0		
D20	6.9	8.5	8.4	7.7	8.6	8.4	8.5	8.8	8.7
D21	5.7	7.1	7.1	6.8	7.3	7.7	7.6	7.4	7.4
D23	8.5	12.5	11.7	10.7	10.8	10.6	10.7	11.3	11.1
D25	6.8	8.7	8.2	6.7	8.0	7.5	7.7	7.6	7.2
D26		3.6	7.0	3.1					
D27	5.5	8.8	8.7	8.1		7.1	6.9	6.8	7.1
D28		8.6	7.8	7.8		10.0	9.6		
D29	6.8	9.4	9.4	8.0	9.2	9.1	9.3	9.5	9.4
D31	5.4	3.9	6.2	6.1					
D32	3.4	4.4	4.4	4.2	4.7				
D33	4.0	5.6	6.0	5.9	7.3	6.0	6.0	5.6	5.9
A1	11.4	12.9	13.0	10.4	10.0	11.5	11.7	11.8	11.7
A2	7.0	10.3	10.7	10.8	9.6	9.9	9.3	9.4	9.2
A3	3.7	4.1	4.1	3.8		3.8	3.4		
A4	5.4	7.0	7.0		7.6	6.6	6.8		8.3
A5	3.4	3.6	3.7		8.6	5.8	5.2		
A6	1.5	1.7	1.7	1.6	1.8	2.4	2.5	2.5	2.0
A7	3.2	2.9	2.9	5.1	2.8	5.6	4.0		
A8		3.5	2.7	2.7	3.8	3.0	3.6		3.5
A9		1.0	1.0	1.0	1.2	1.8	1.8		
A11	8.7	11.8	12.2			14.8	13.6		11.2

Table A-1. Cont.

A12	7.1	10.1		6.8					
A13	12.6	14.9				14.0			
A14	10.9	14.3	14.2			12.9	13.0		11.0
A15		4.3	4.4		4.4				
A16	12.2	17.3	16.3			13.6	12.4		11.2
A17	12.3	16.8		11.7	14.5	16.0	15.6		
A18	18.3	20.3	18.6			15.3	14.8		12.8
A19		8.5	13.0			12.8	14.0		11.7
A20		6.7	6.7	6.2	6.6	6.6	6.8	6.8	6.1
A21	8.9	11.1	10.5			8.7	10.7		
A22	5.8	7.9	8.0	6.3	7.8	6.4	7.6		
A23		10.3		9.1		9.3	9.4	9.6	
A24	7.2	8.9	9.4	8.4	10.0	8.7	8.6		9.2
A25	11.3	12.0	15.0	12.1	14.0	14.1	13.4	13.4	
A27							7.8	8.6	7.8
A29	8.3	10.3	9.8	9.1	9.9	9.4	9.4	9.8	9.2
A30								7.6	7.6
A32								7.4	7.9
A33								5.9	5.9
R1	5.6	8.2	8.5	7.8	7.9	7.5	7.4	7.5	6.8
R2	4.3	7.3	7.7	7.4	7.6	7.3	7.3	7.6	7.1
R3		7.4	7.6	7.7	9.1	9.0	9.2	9.0	9.2
R4	4.5		8.3	7.8	8.1	8.1	8.4	8.6	8.4
R5	5.7	7.9	7.9	7.2	7.3	7.3	7.3	7.3	6.8
R6	4.5	6.9	6.9	6.8	7.0	6.8	6.8	7.0	
R7		6.1	8.8	7.1	7.5	7.5	7.3	7.7	7.3
R8	5.6	9.2	9.9	8.7	9.9	9.3	9.0	9.1	8.9
R9	5.2	8.3	9.6	8.7	9.3	9.4	9.6	10.2	9.3
R10	3.4	6.6	6.7	6.6	6.9	7.3	7.0	7.4	6.7
R11	3.7	6.3	6.6	6.8	6.8	7.0	6.5	6.6	6.5
R12	3.5	5.7	6.0	5.7	5.9	6.2	6.5	6.7	6.5
R13	4.0	7.0	7.7	7.3	8.1	7.8	7.7	7.8	7.9
R14	4.7	7.3	10.5	9.4	12.2	11.6	12.1	13.6	13.6
R15	4.1	4.8	5.6	5.0	5.4	5.3	5.3	5.0	
R16	3.3	5.2	6.1	6.6	7.4	7.5	8.4		
R17		3.6	3.8	3.3	4.6	4.5	4.7	4.7	4.9
R18	3.3	4.5	4.2	4.2	4.6	4.9	5.1	5.5	5.5
R19	3.3	4.3	4.2	4.1	4.8	5.5	5.1	5.4	5.1
R20	5.0	8.5	9.6	8.9	10.8	10.5	10.8	12.0	11.7
R21		9.2	12.0	10.3	12.9	13.5	13.3	13.7	13.3
R22	3.4	3.9	3.6	3.6	3.9	4.2	4.2		
R24	3.3	3.3	2.8	2.5	3.5	3.3	3.5	3.6	3.8

Table A-1. Cont.

R25	2.5	3.0	2.8	2.7	2.8	3.5	3.7	3.4	3.3
R26	4.7	9.6	10.6	9.0	10.3	10.8	11.2	11.9	12.2
R27	2.8	4.7	5.0	5.3	5.4	6.3	7.1	8.0	7.8
R28	10.0	15.8	18.8	17.0	16.0	15.6	15.6	16.2	15.7
CS1	4.52	4.47	3.8	2.74	3.92	3.24	2.95	2.0	1.8
CS2	0.9	1		0.63	1.36	1.4	1.25	1.0	1.1
CS3	4.85	6.38	4.3	3.3	5.24	4.4	3.9	3.9	4.1
CS4	3.08	3.63	4.4	5.16	5.4	4.64	4.7	4.2	3.6
CS5		8.3		7.87	7.9	7.9	7.35	7.1	6.9
CS6	7.7	9.53	8.12	7.34	7.88	7.65	6.75	6.4	6.0
CS8	4.88	5.27	4.9	4.6	4.76	4.8	4.6	5.3	4.4
CS9		8.12	8.6	6.84	7.3	7.68	7.45	7.4	
CS10	6.1	8.02	8.4	7.66	7.72	8.08	8.7	9.1	9.8
CS11	4.06	4.82	4.45	4.28	4.68	4.92	5	4.9	4.8
CS12		4.74	5.6	4.34	4.56	4.88	4.75	4.5	4.4
CS13	4.02	4.57	4.64	4.41	4.68	4.68	4.7	4.3	4.3
CS14	3.8	4.47	4.83	4.7	4.88	5	4.8	4.7	4.8
CS15	4.78	6.86	7.42	7.26	7.86	7.88	8.35	8.3	7.3
CS16	4.03	5.4	5.9	5.43	5.7	5.75	6.05	5.6	5.3
CS17	4.7	6.8	6.2	5.87	6.84	6.64	8.4	8.7	9.6
CS18	11.26	19.23	18.5	16.94	16.44	15.96	16	15.4	14.9
CS19	2.63	2.8	3.2	2.74	3	3.36	3.2	3.3	3.4
CS20	11.37	15.63	16.9	11.93	13.5	14.48	13.95	13.4	13.0
CS21	12.2	15.52	15.03	14.66	14.36	14.2	13.95	13.7	12.9
CS22	2.7	3.06	3.02	3.96	3.27	4.05	4.05	4.1	4.0
CS23	2.65	3.9	3.66	3.13	3.1	3.7	3.9	4.0	3.4
CS24		2.7		0.87	0.88	0.4			4.0
CS25	2.24	3.42	3.08	2.46	3.04	2.55	3	2.6	2.6
CS26	1	0.65		1.04	0.7	1.8	1.65	1.4	1.1
CS27	1.45	1.38	1.42	1.1	1.52	1.96	1.8	1.5	1.5
CS28		0.07	0.37	0.44	0.4		1	0.7	
CS29	0.38	0.12	0.34	0.48	0.3	0.8	0.9	0.5	0.9
CS30	3.28	3.6	4.08	3.99	4.24	4.76	5.1	4.4	4.3

APPENDIX B

KRIGING ESTIMATE OF NITRATE FROM THE YEAR 1998 TO 2004

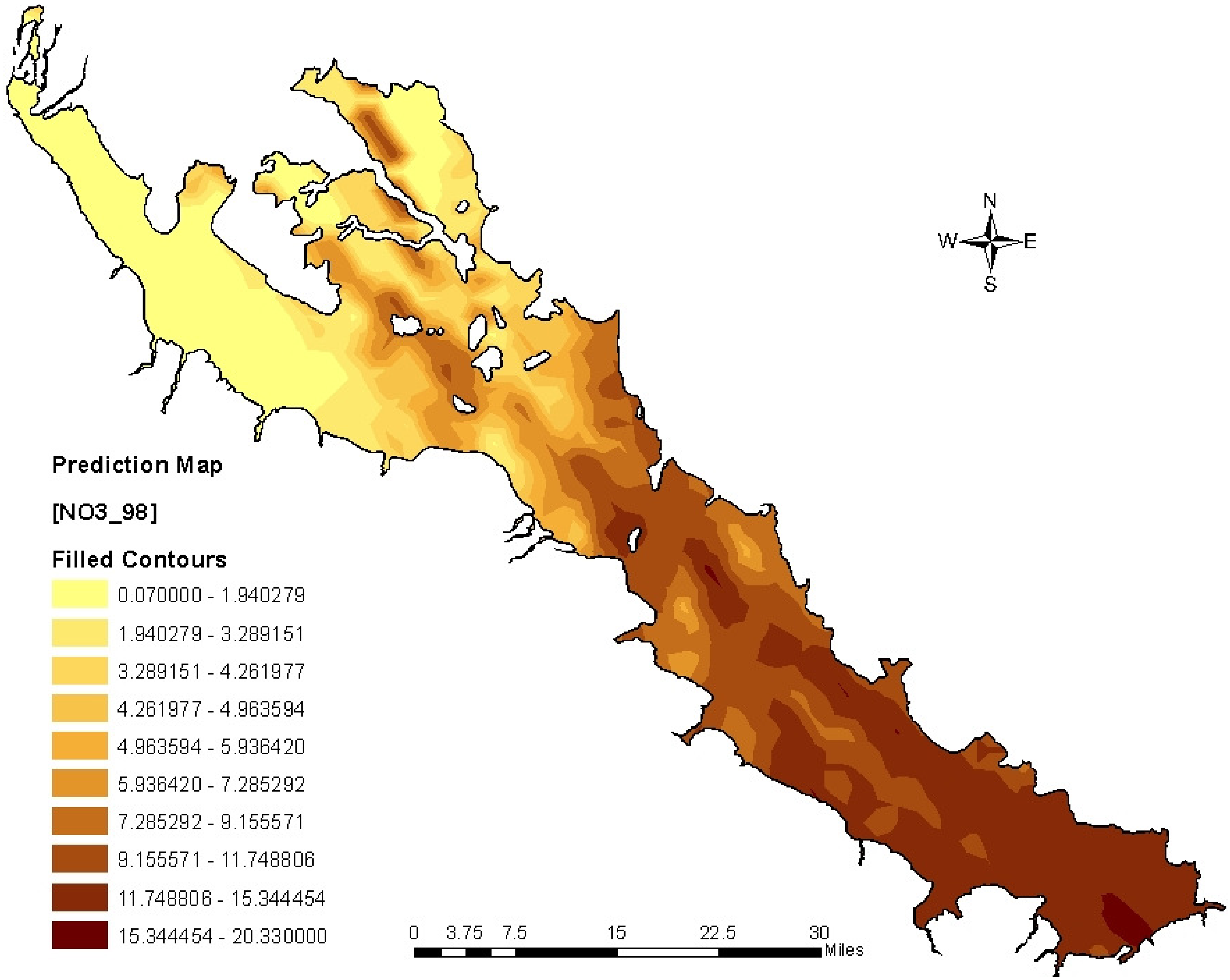


Figure B-1. Kriging estimation of nitrate as nitrogen (mg/l) for the year 1998 in Cimarron terrace aquifer

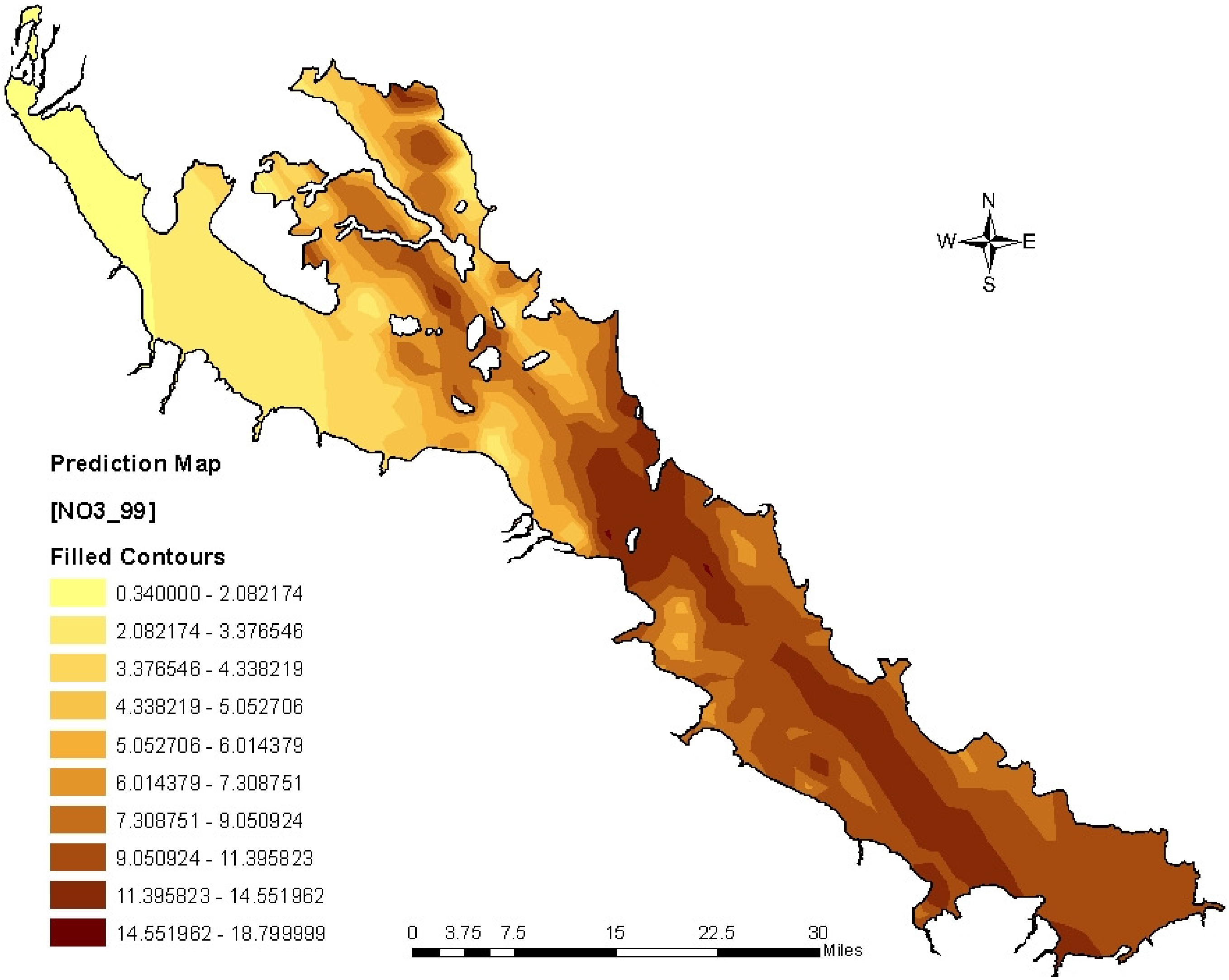


Figure B-2. Kriging estimation of nitrate as nitrogen (mg/l) for the year 1999 in Cimarron terrace aquifer

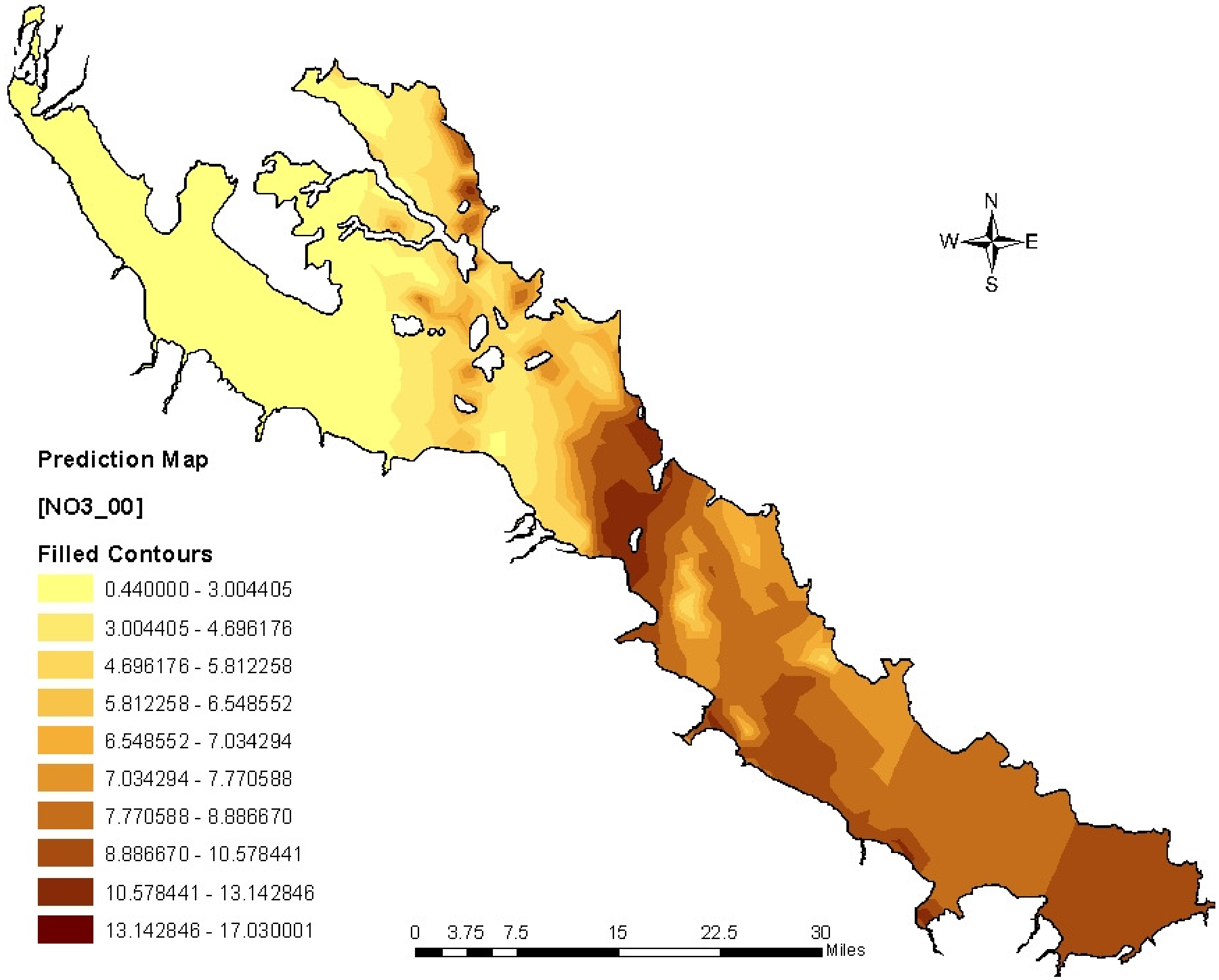


Figure B-3. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2000 in Cimarron terrace aquifer

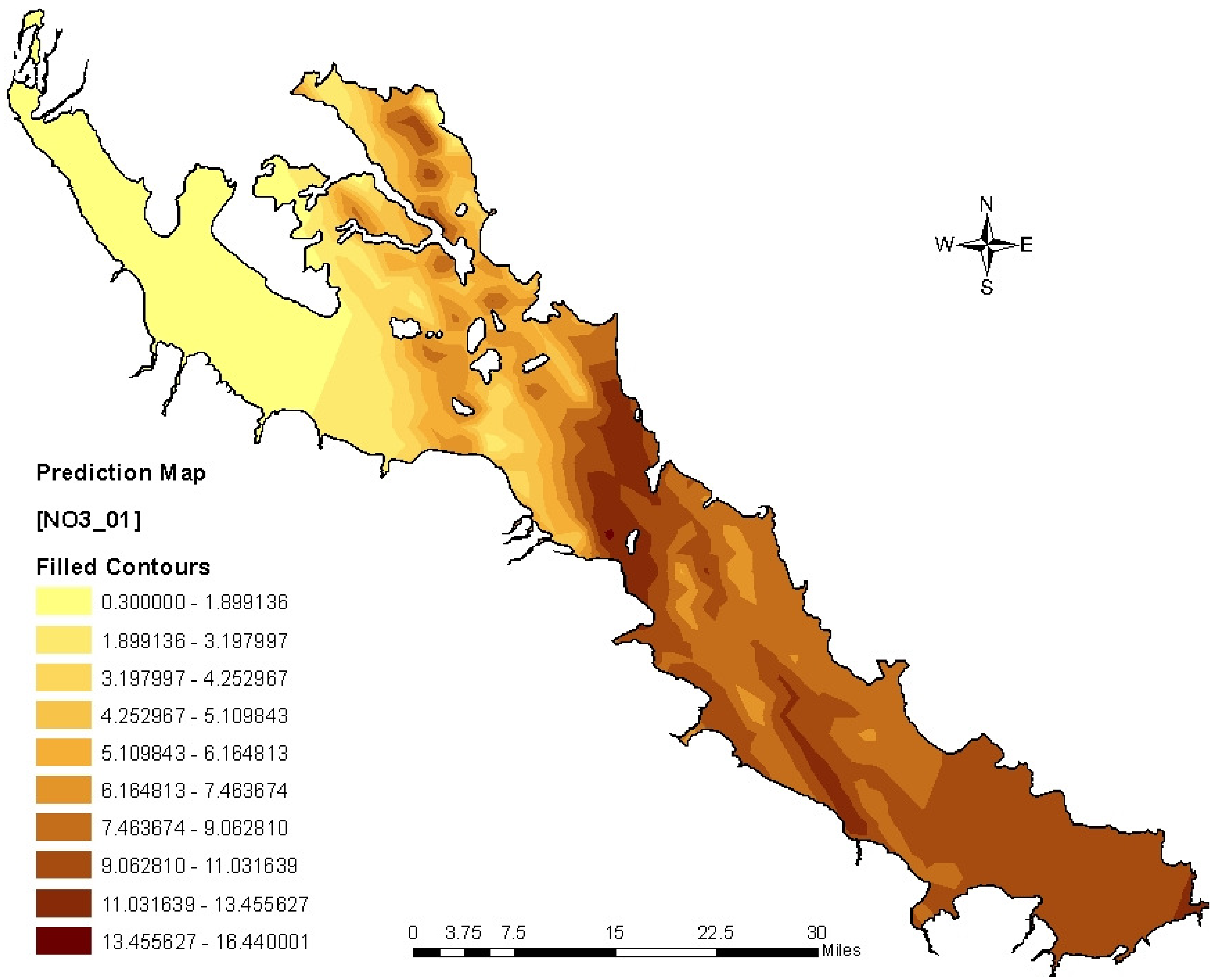


Figure B-4. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2001 in Cimarron terrace aquifer

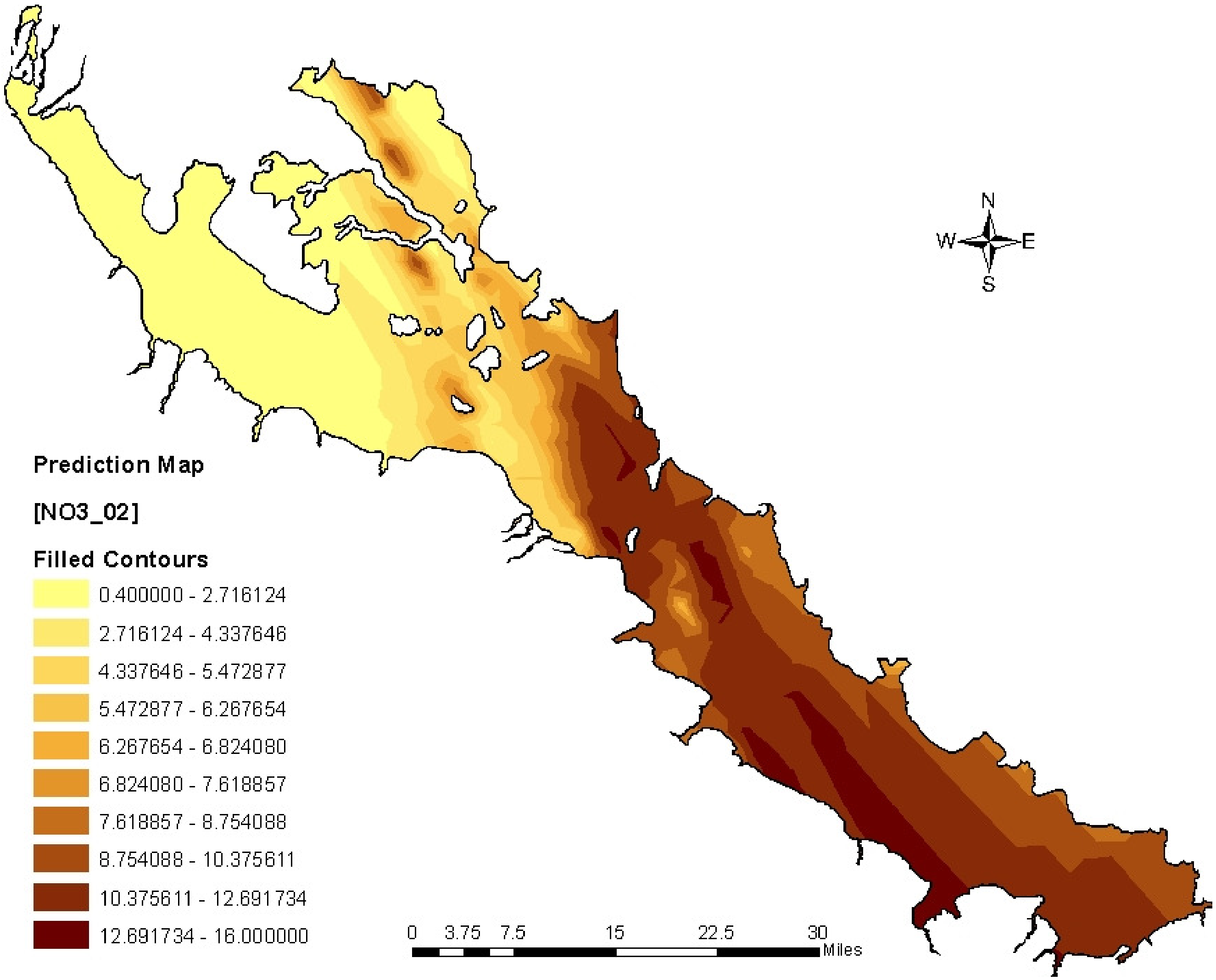


Figure B-5. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2002 in Cimarron terrace aquifer

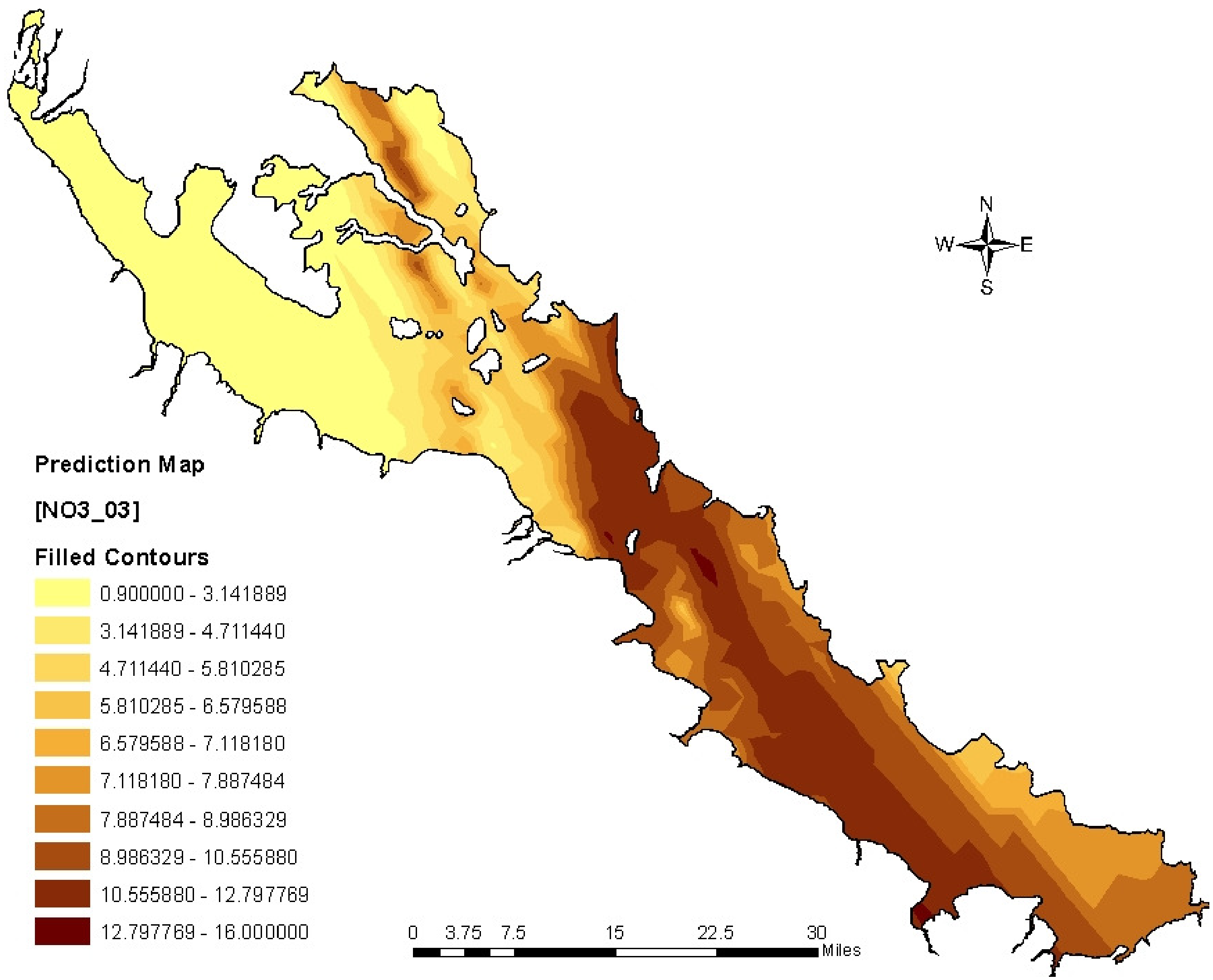


Figure B-6. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2003 in Cimarron terrace aquifer

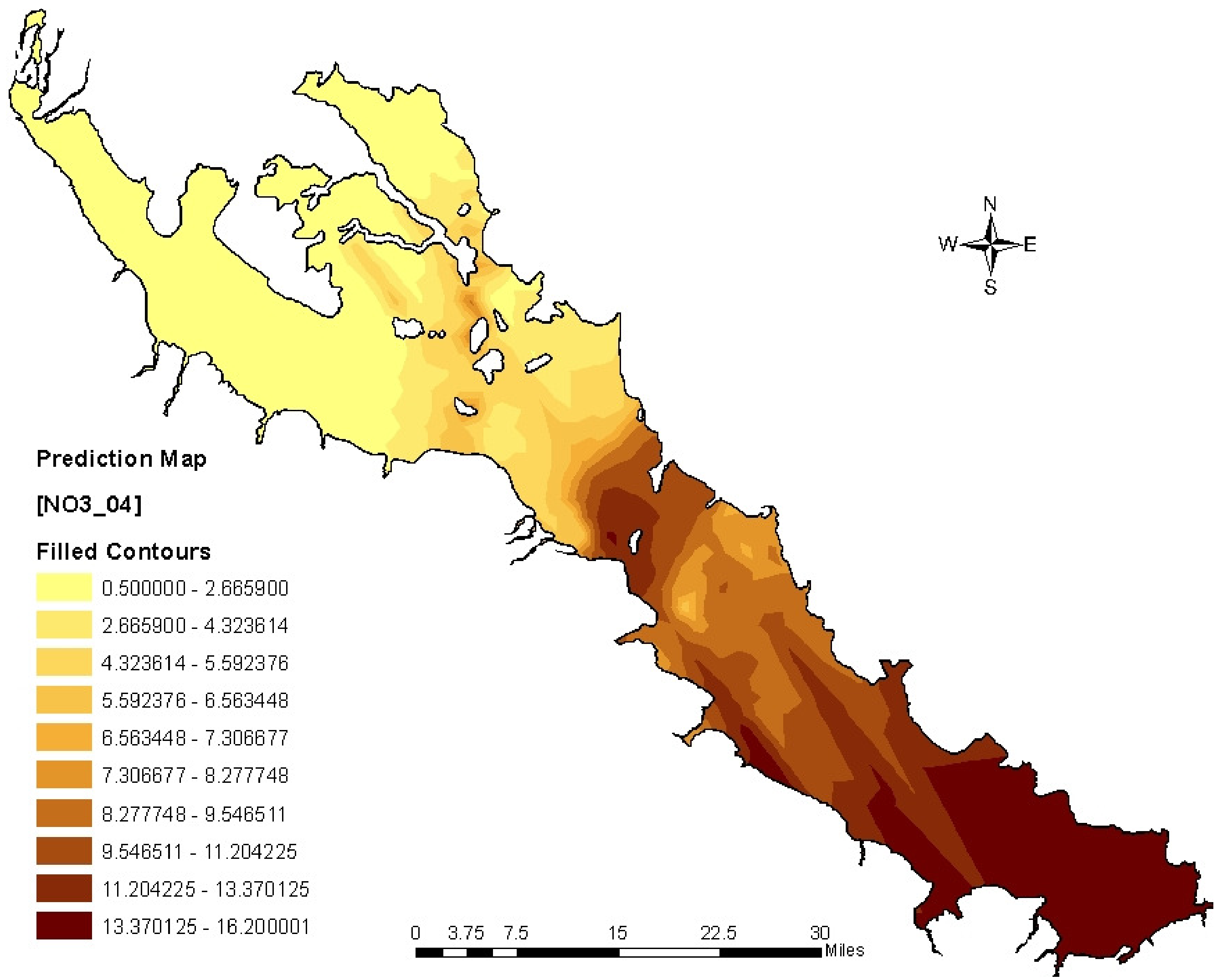


Figure B-7. Kriging estimation of nitrate as nitrogen (mg/l) for the year 2004 in Cimarron terrace aquifer

APPENDIX C

AREAS EXCEEDING 10 MG/L OF NITRATE AS NITROGEN

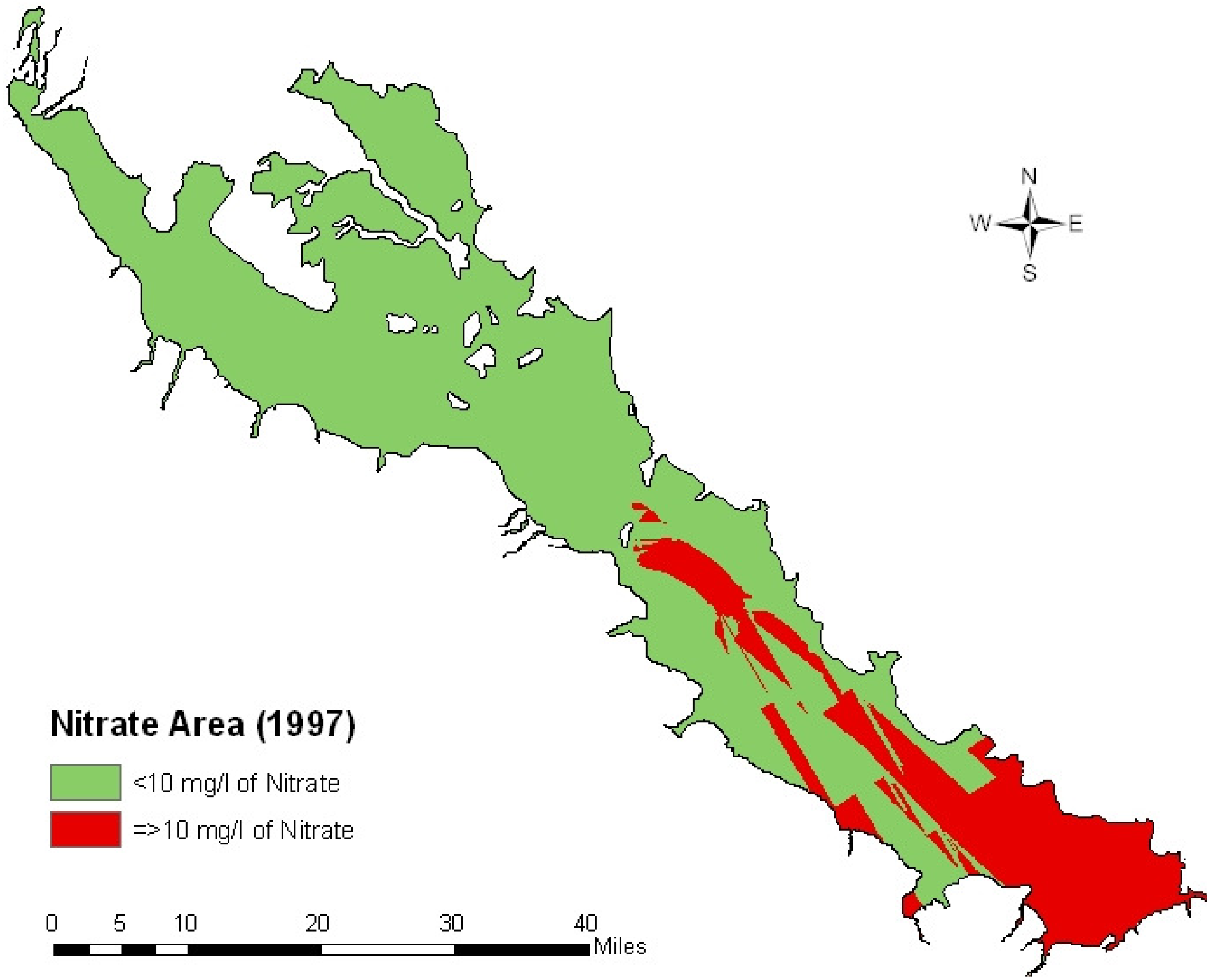


Figure C-1. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 1997

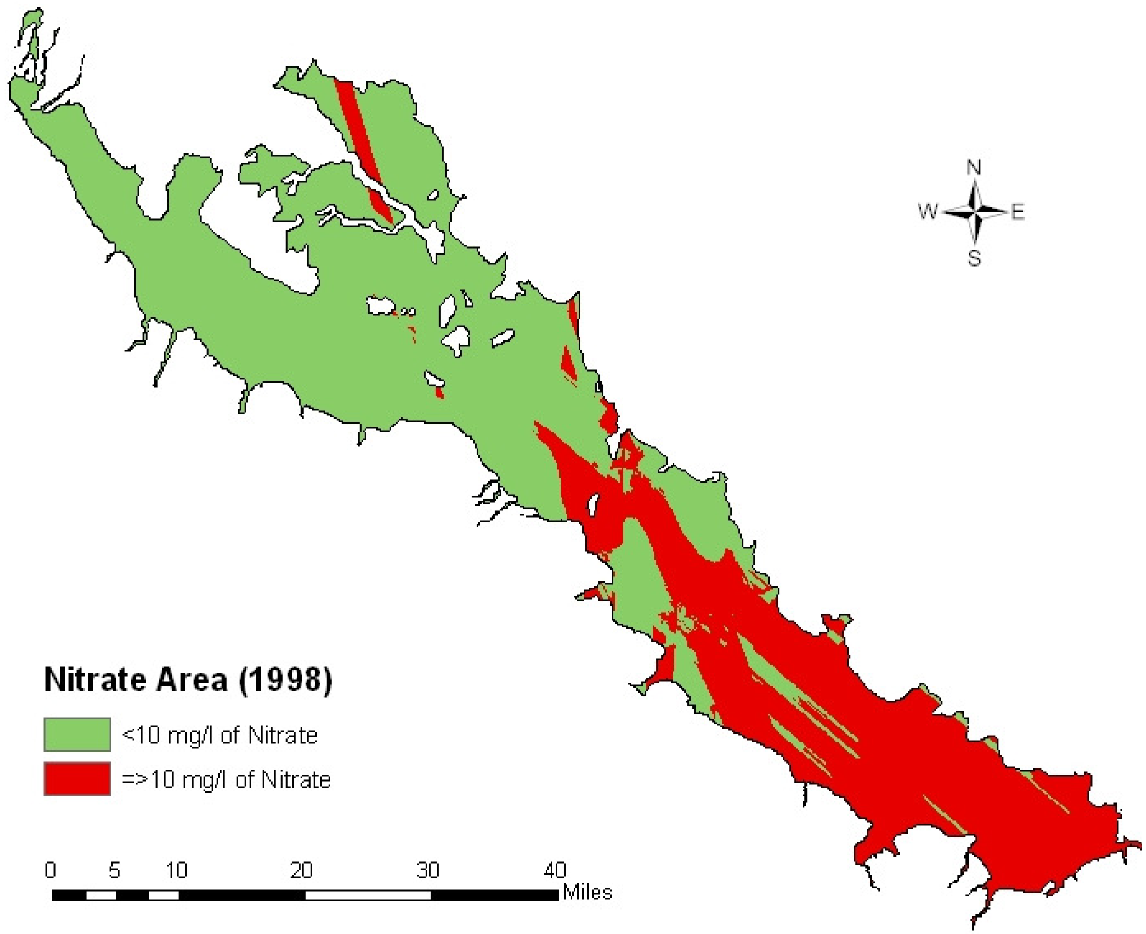


Figure C-2. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 1998

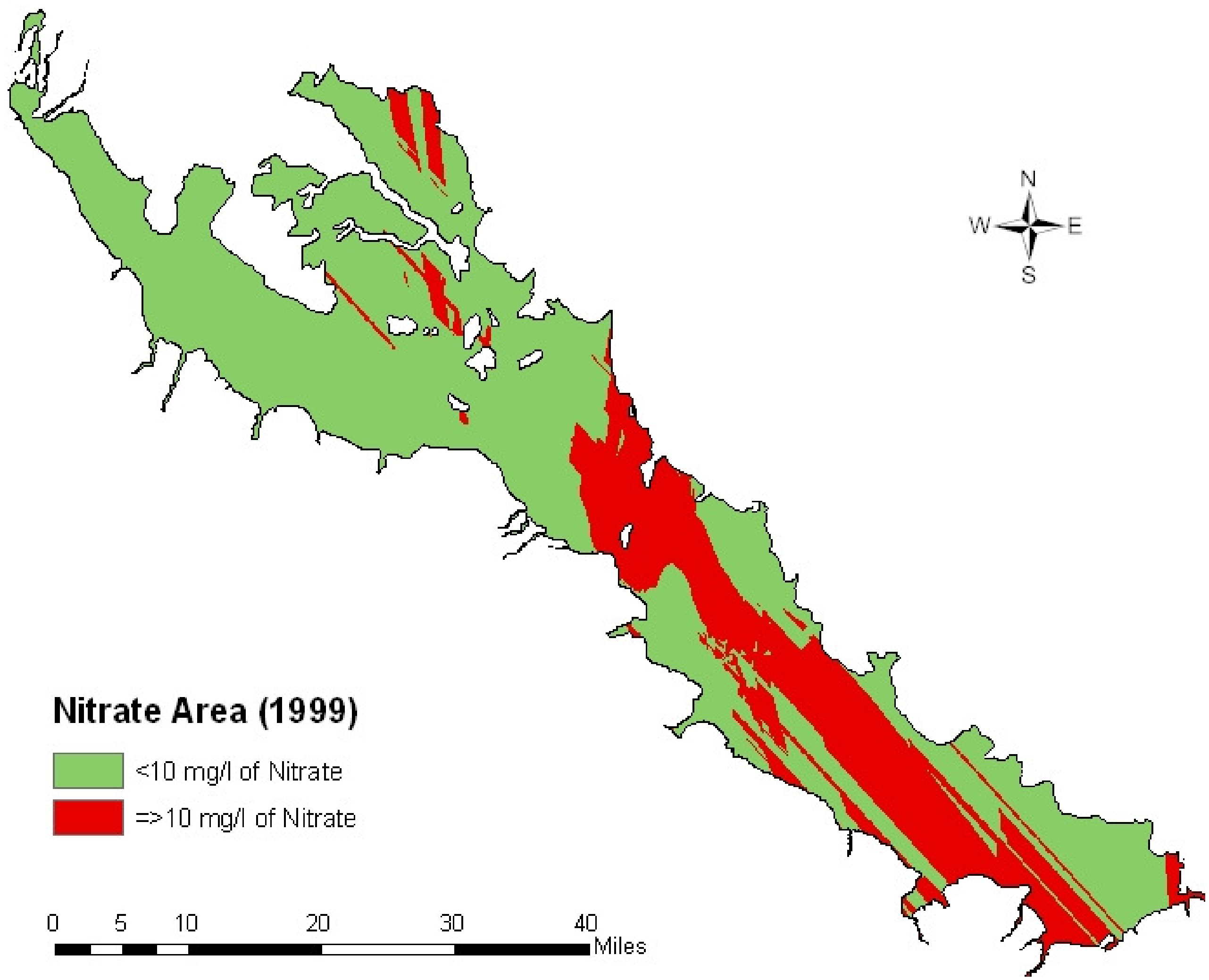


Figure C-3. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 1999

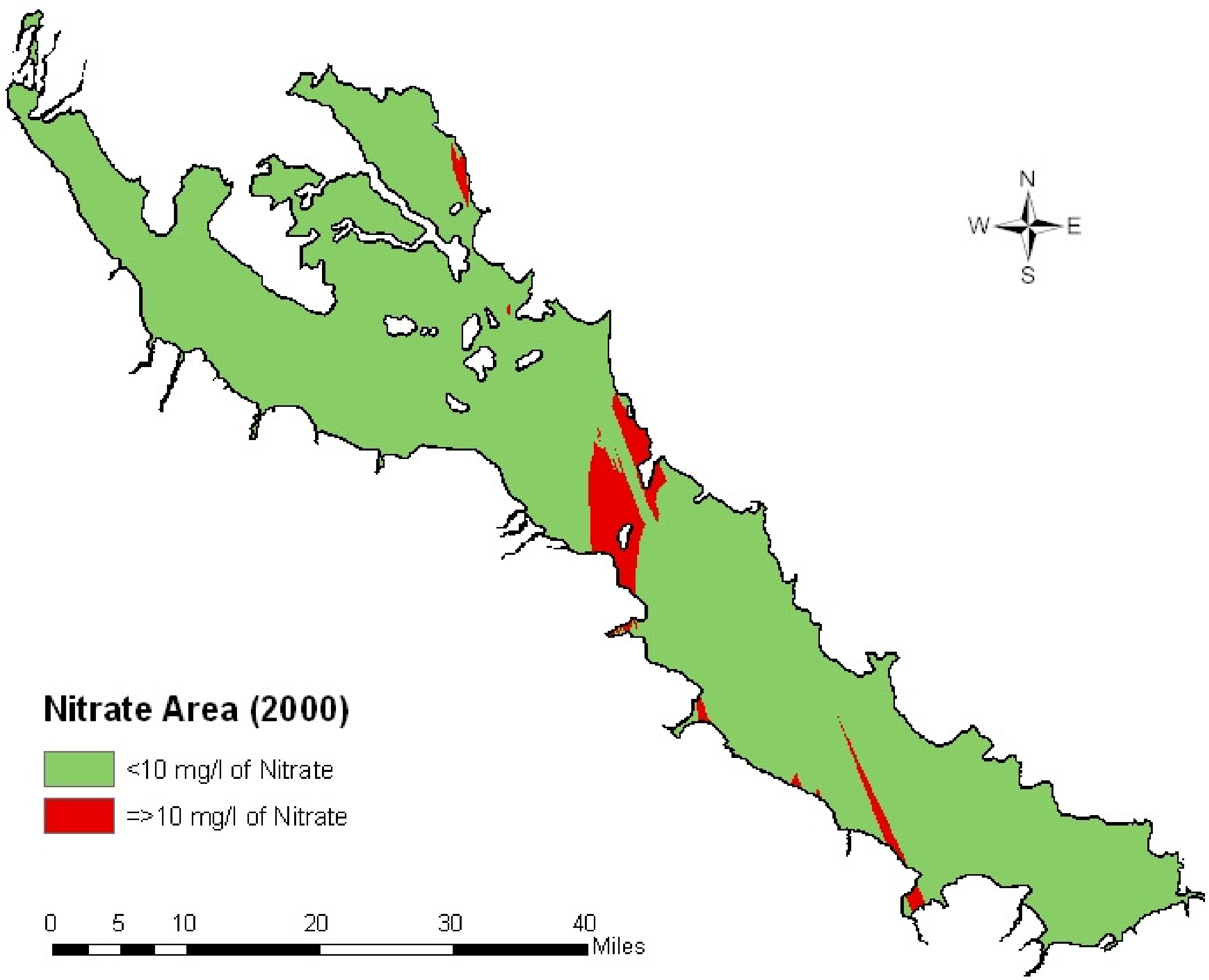


Figure C-4. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2000

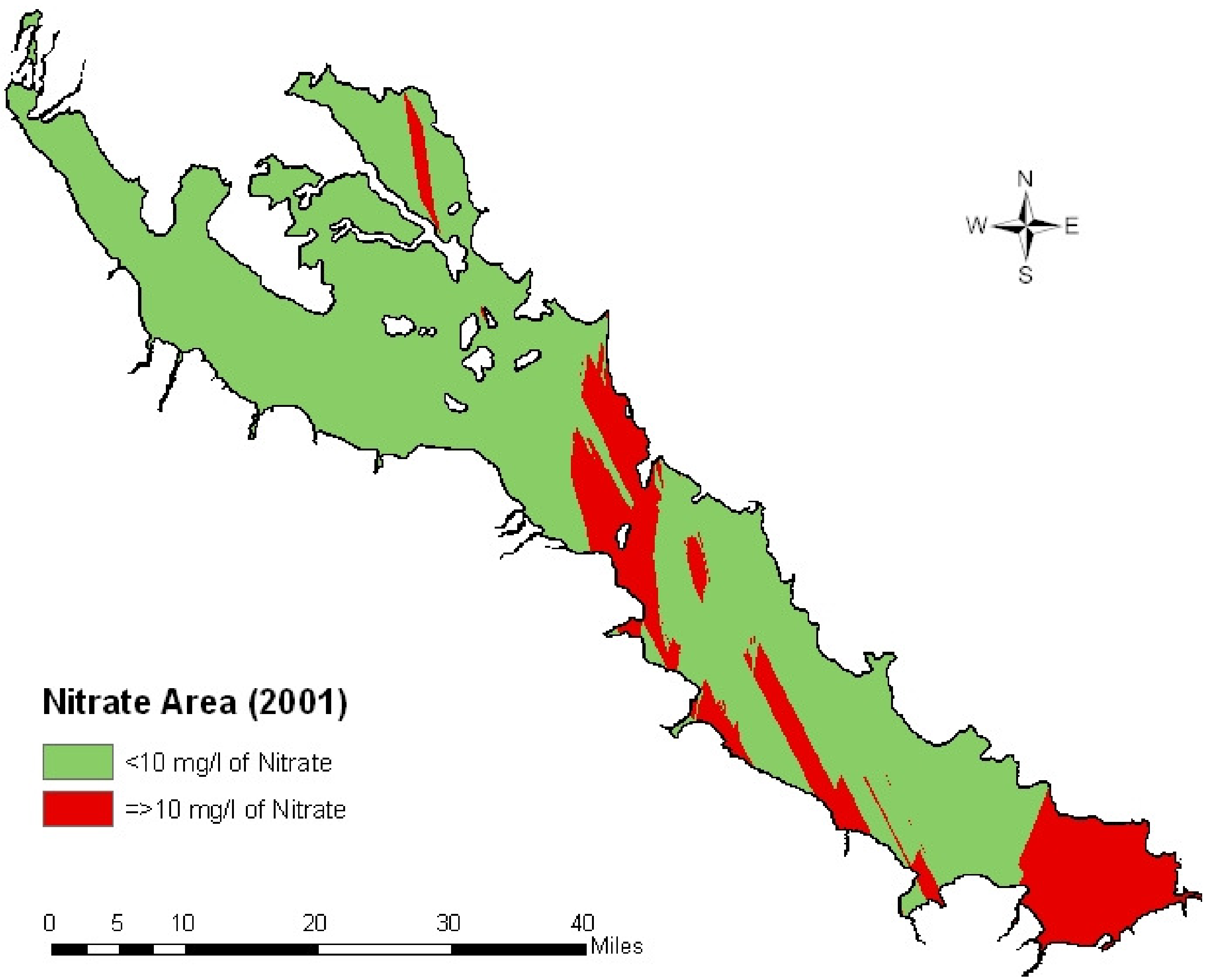


Figure C-5. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2001

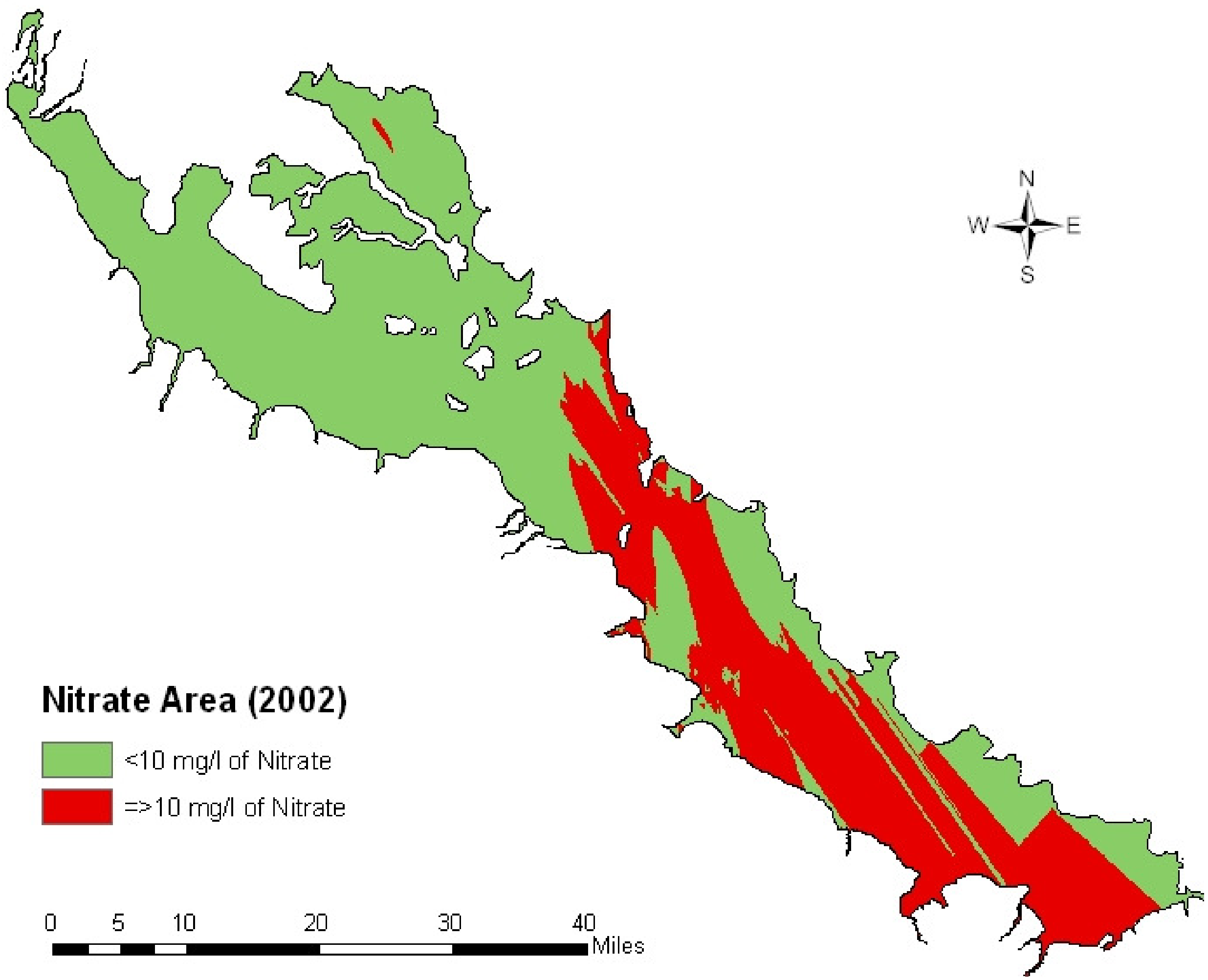


Figure C-6. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2002

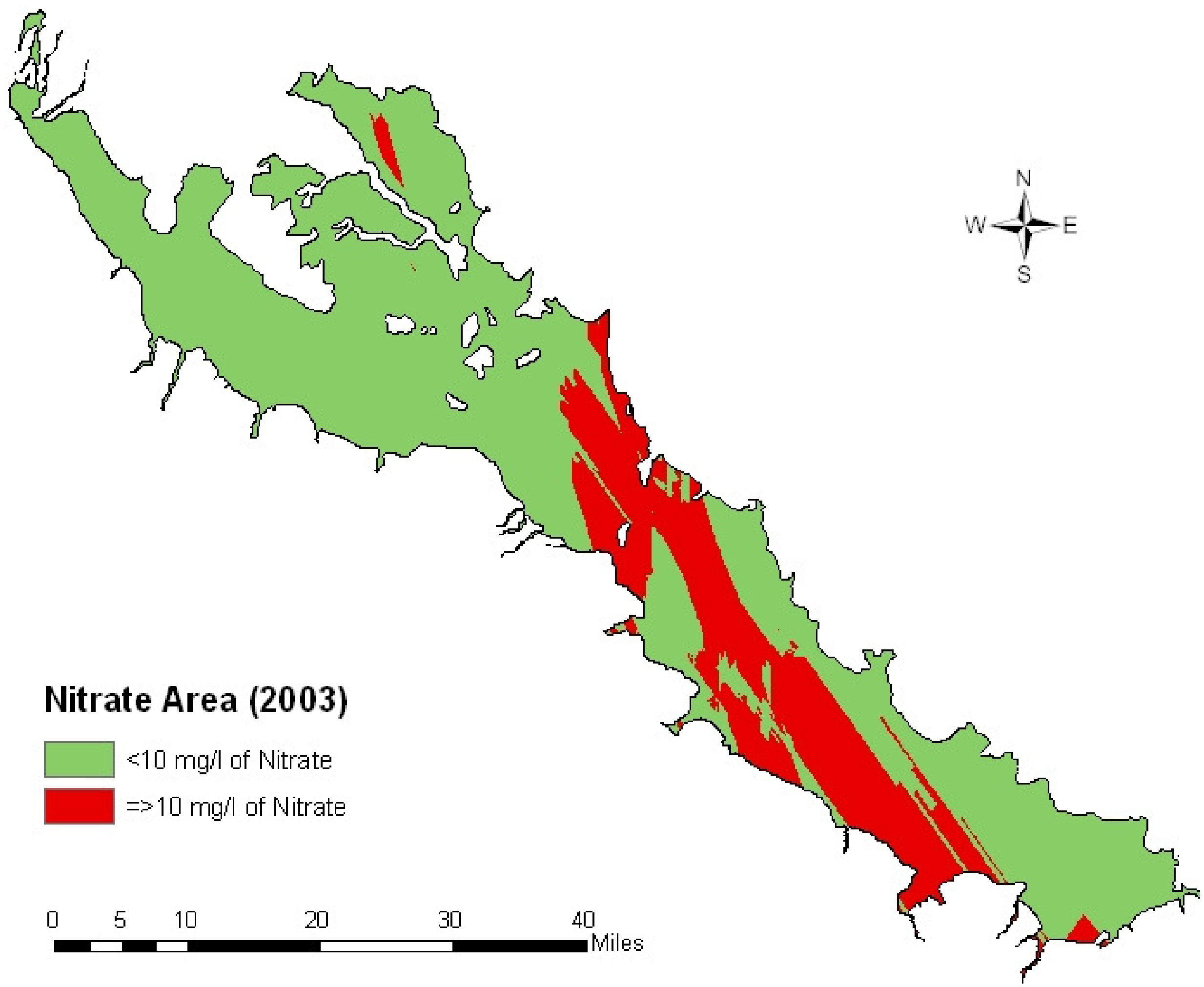


Figure C-7. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2003

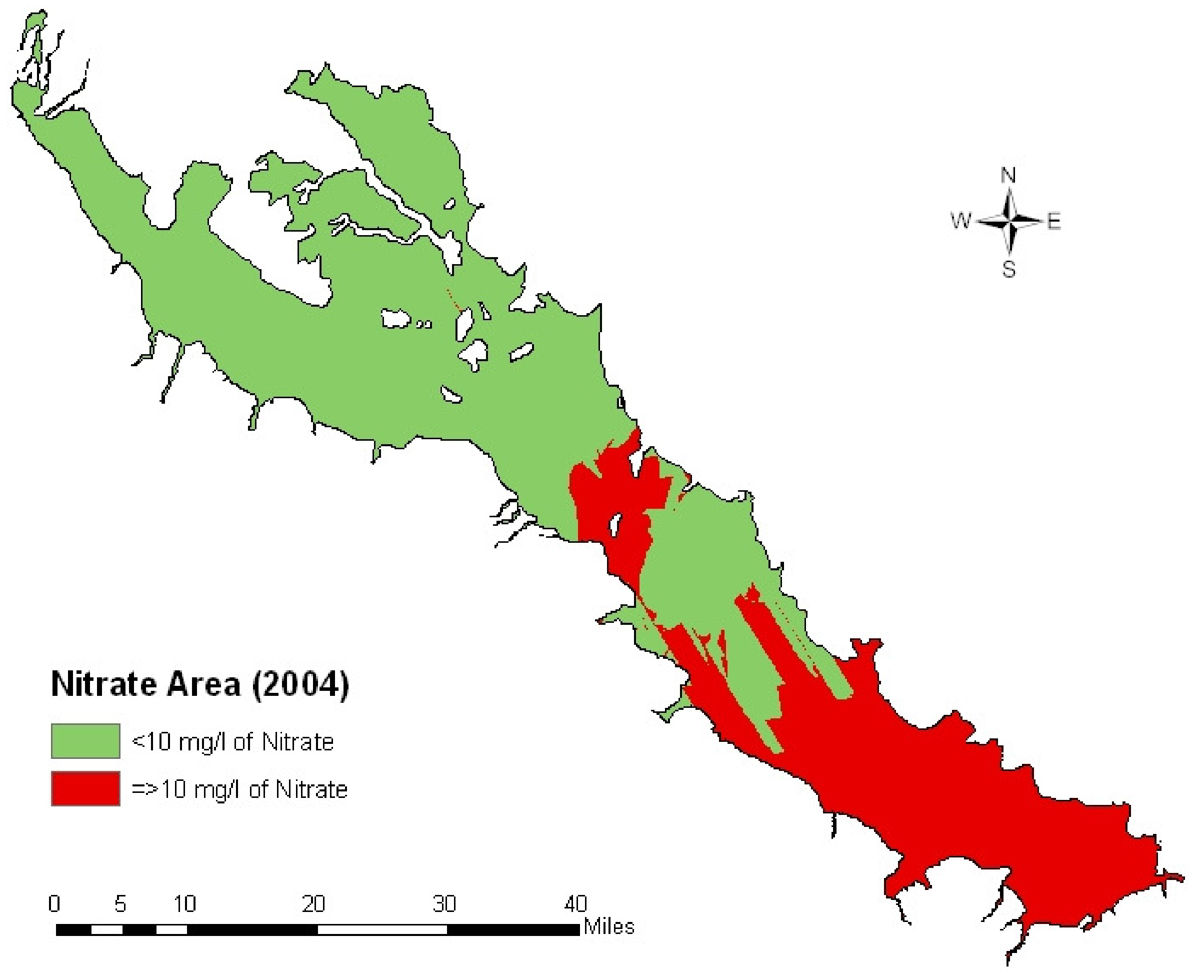


Figure C-8. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2004

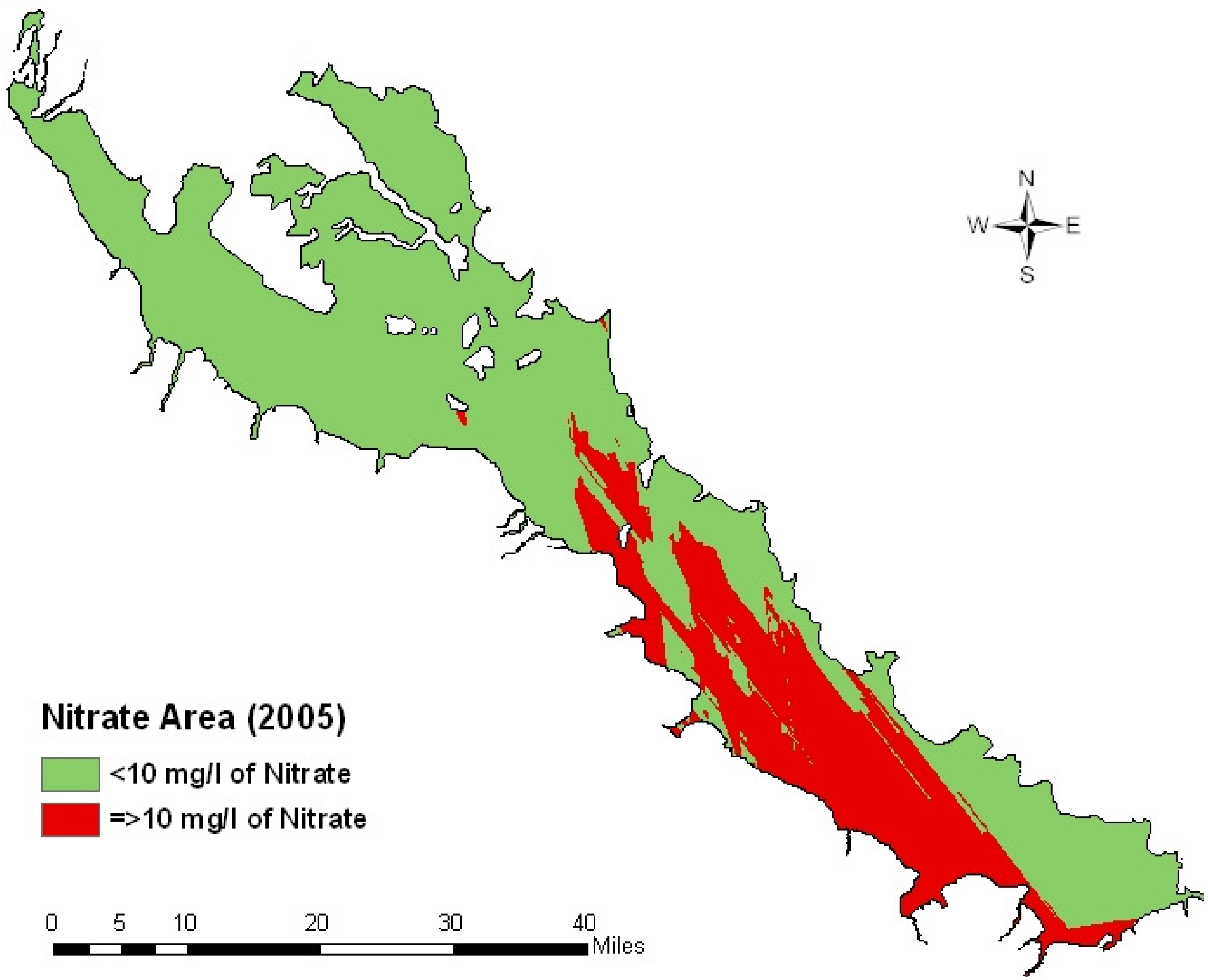


Figure C-9. Kriging estimate of nitrate areas exceeding drinking water criteria, 10 mg/l of nitrate as N, in 2005

APPENDIX D

PROBABILITY MAPS OF NITRATE EXCEEDING 4 MG/L OF NITRATE AS NITROGEN

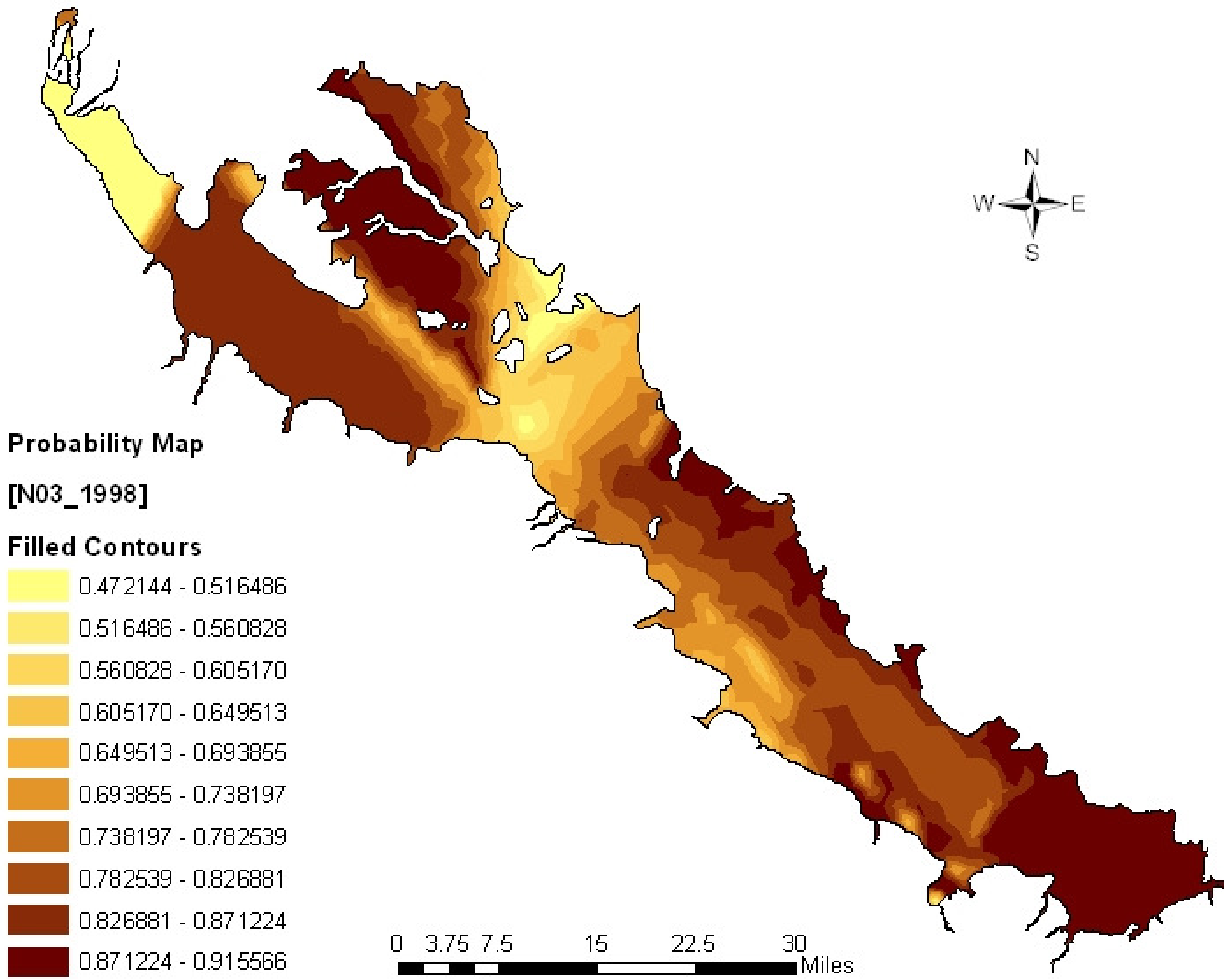


Figure D-1. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 1998

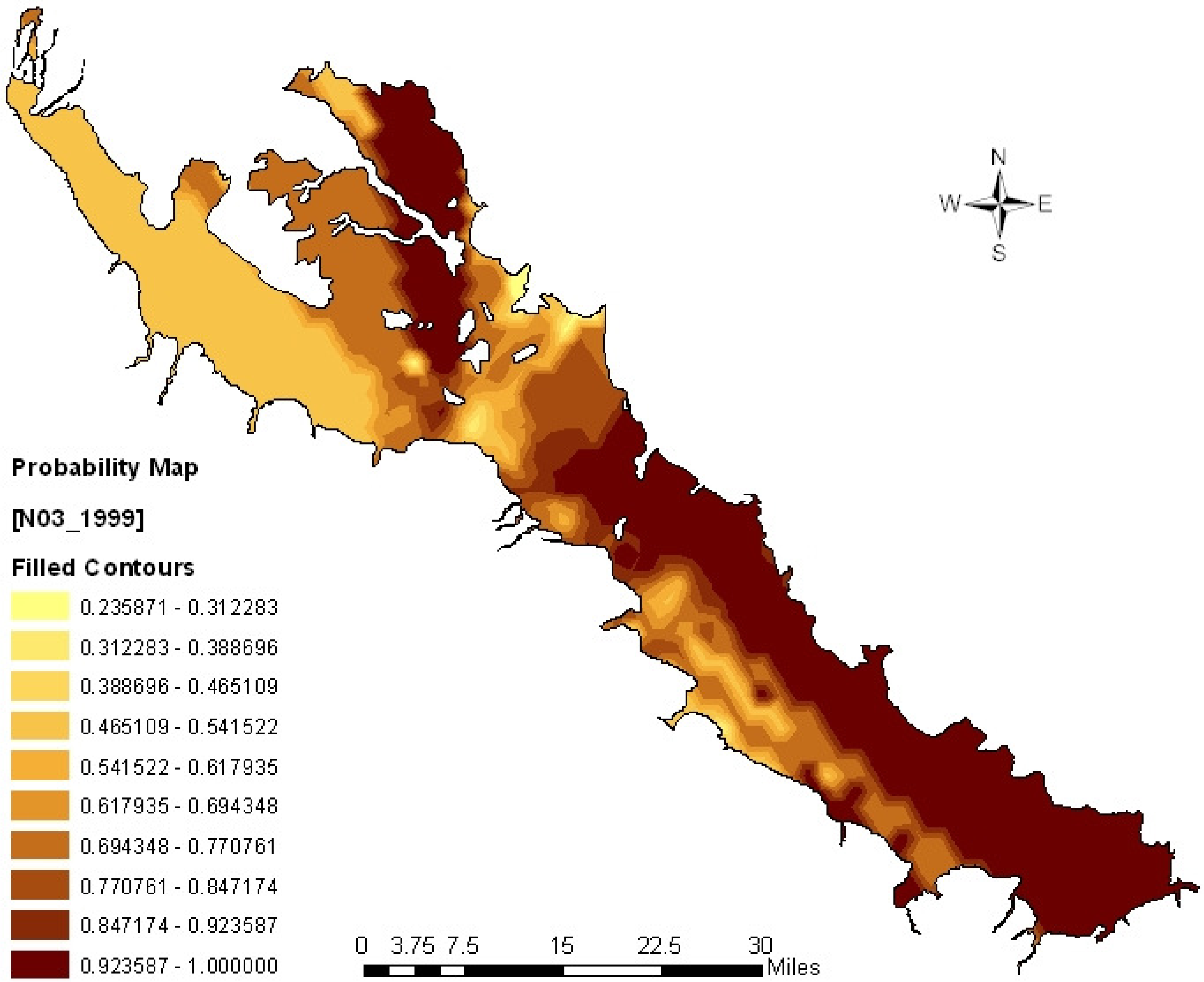


Figure D-2. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 1999

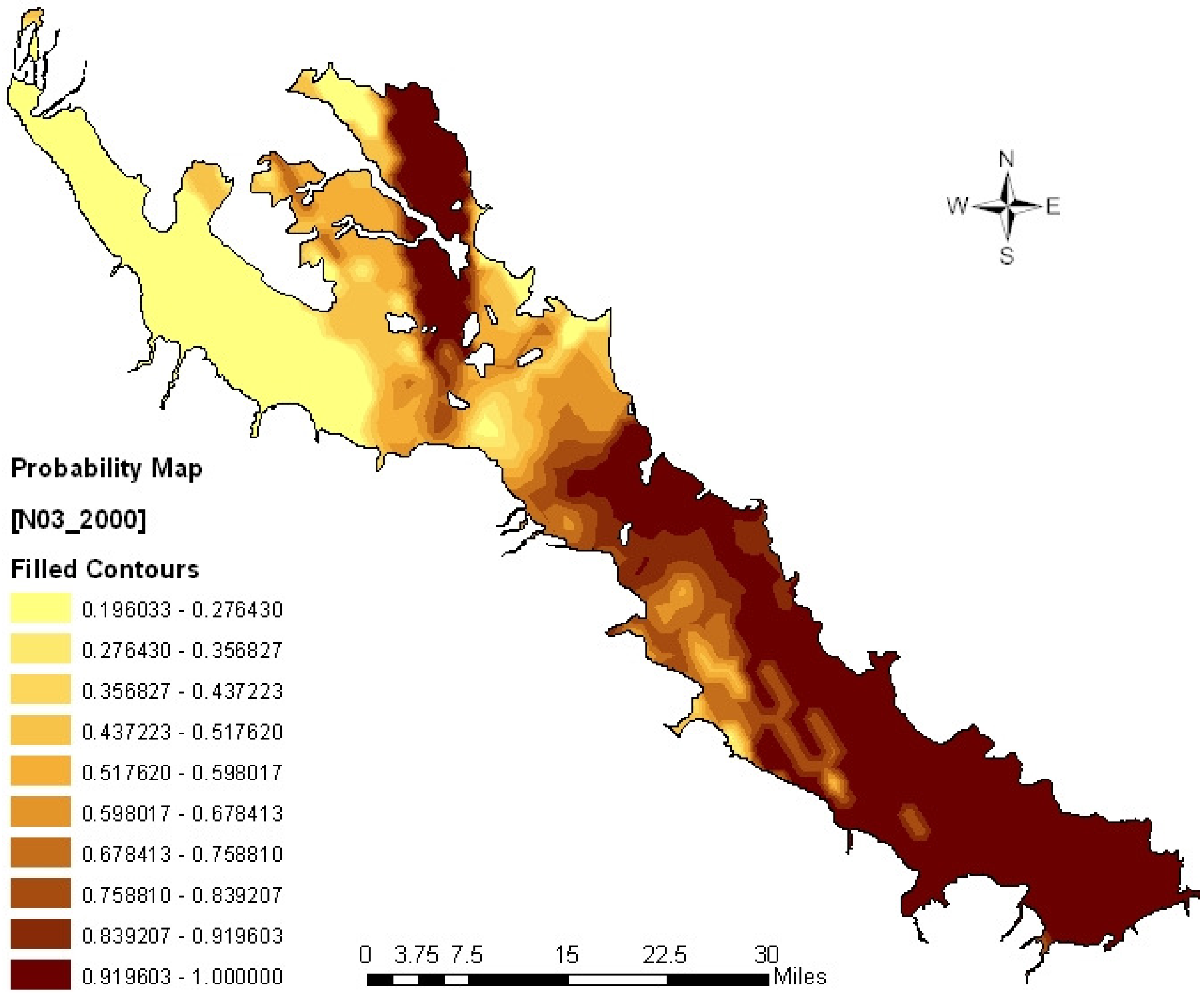


Figure D-3. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2000

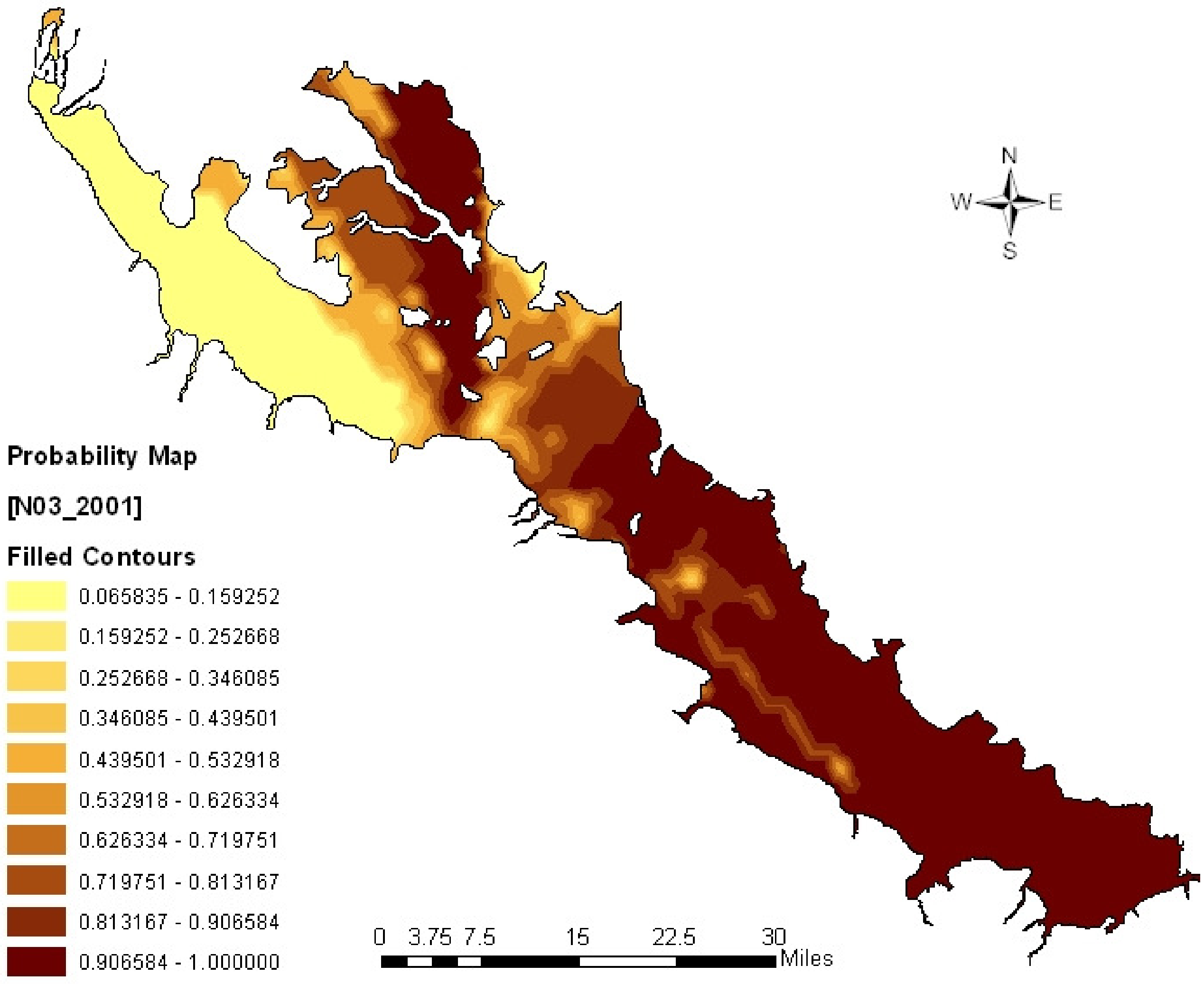


Figure D-4. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2001

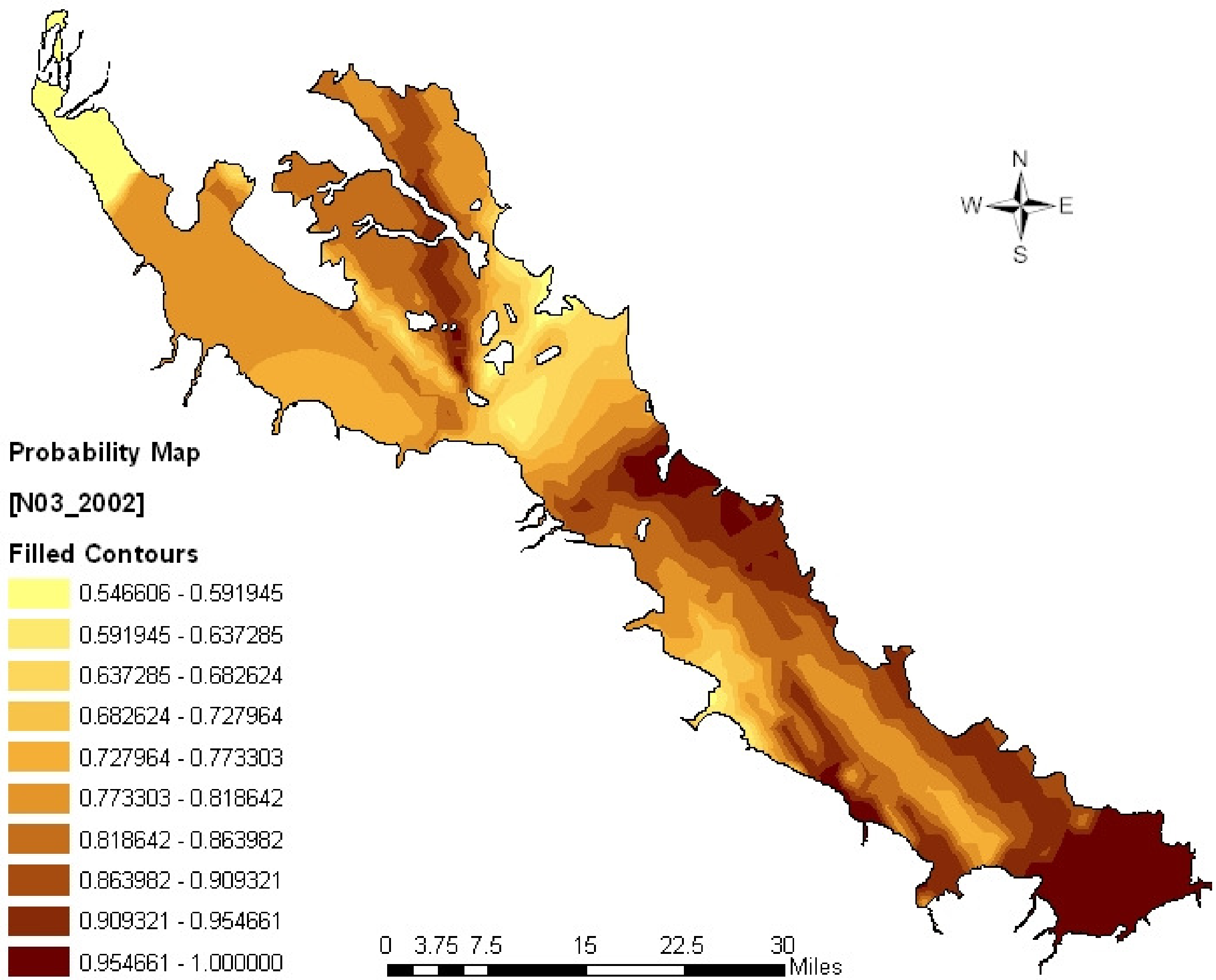


Figure D-5. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2002

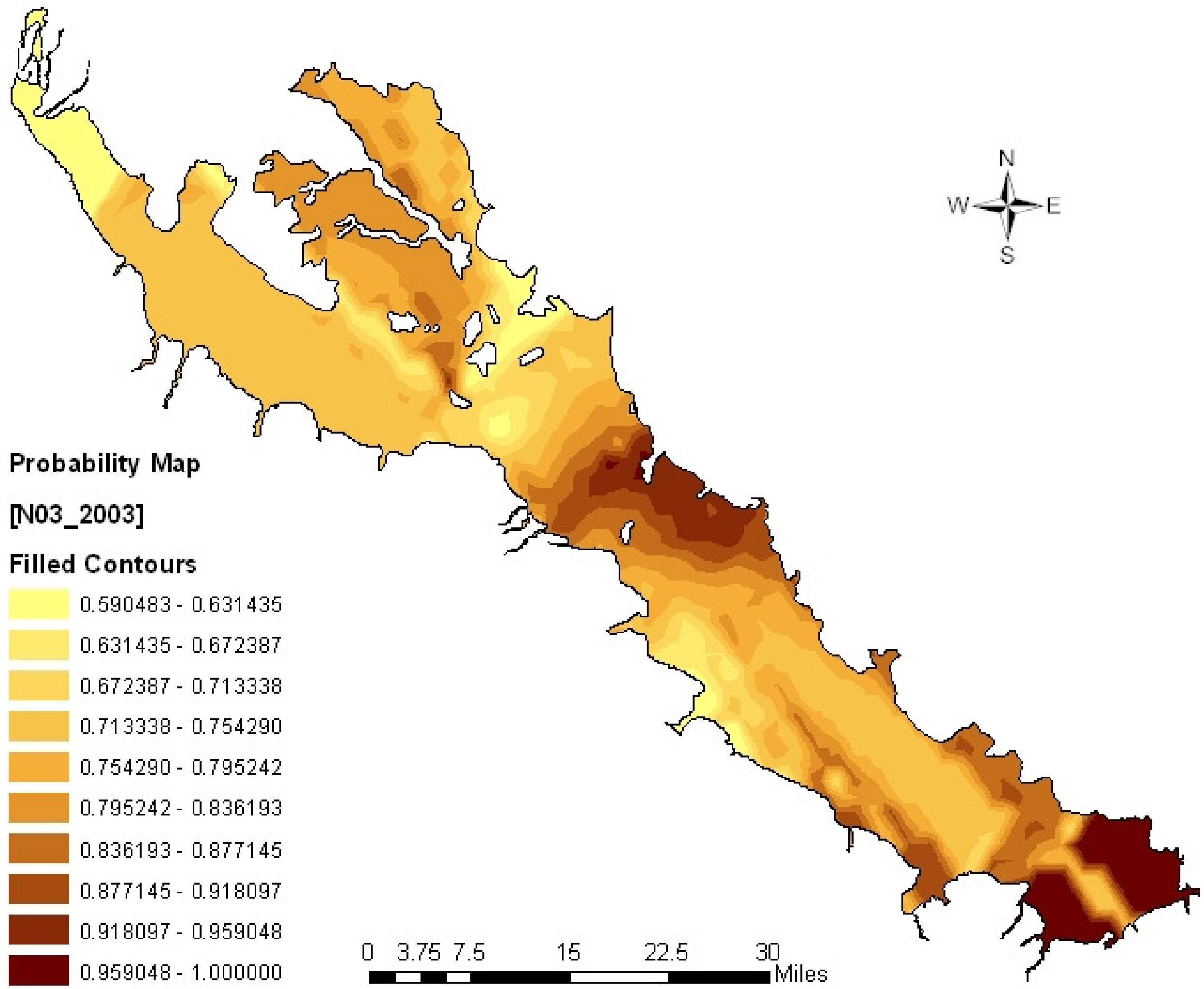


Figure D-6. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2003

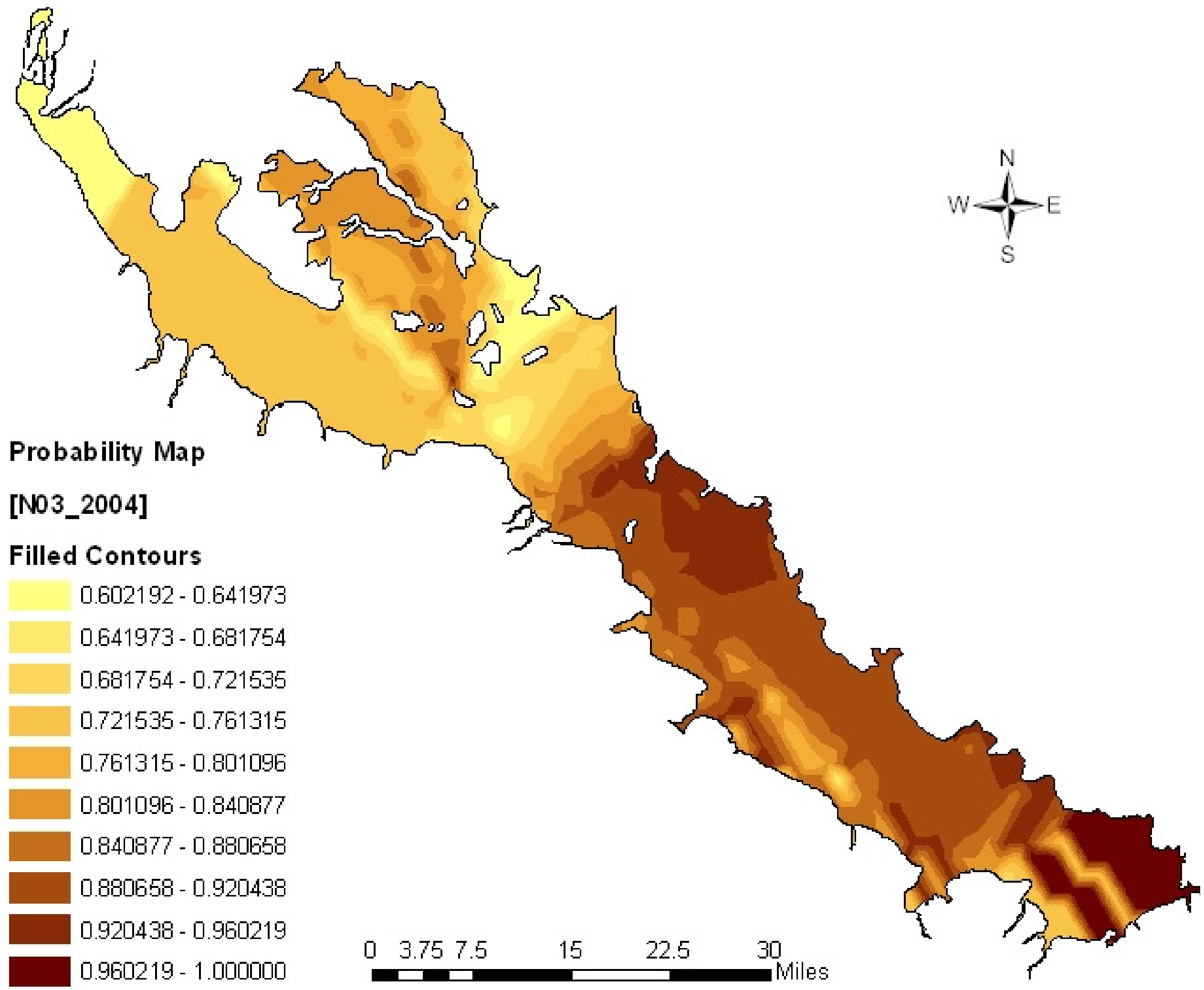


Figure D-7. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2004

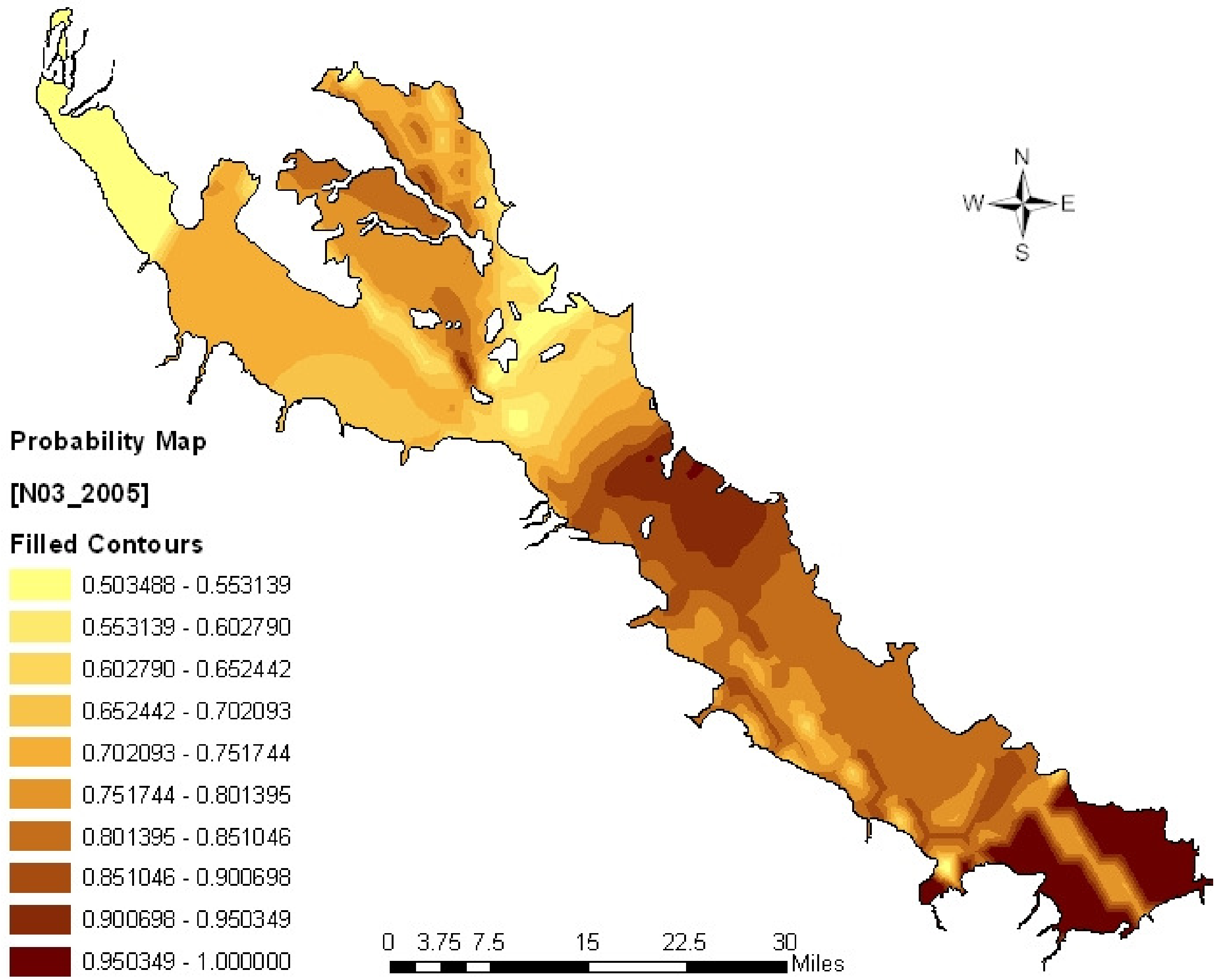


Figure D-8. Indicator probability map at threshold 4 mg/l of nitrate data sampled in 2005

APPENDIX E

PROBABILITY MAPS OF NITRATE EXCEEDING 10 MG/L OF NITRATE AS NITROGEN

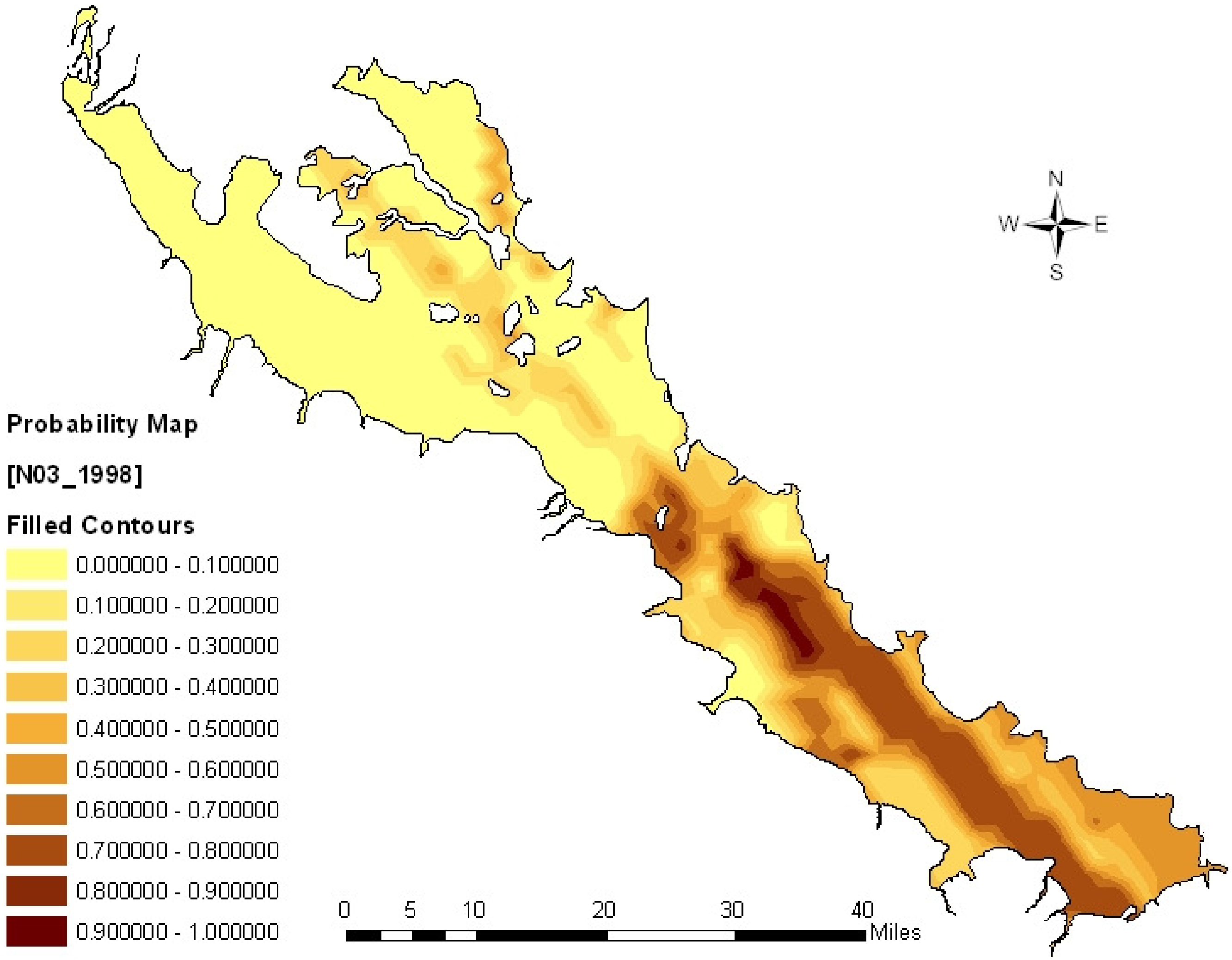


Figure E-1. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 1998

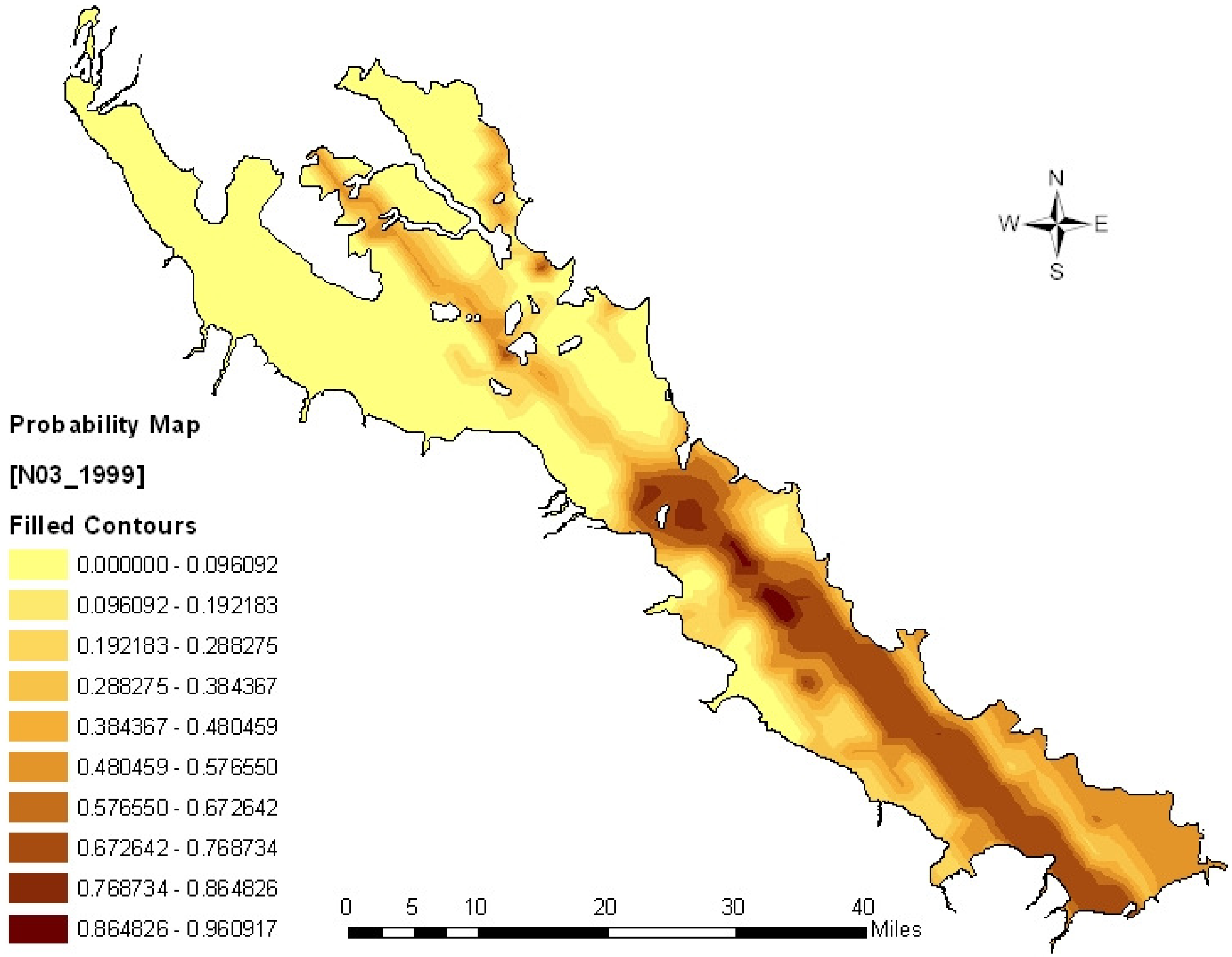


Figure E-2. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 1999

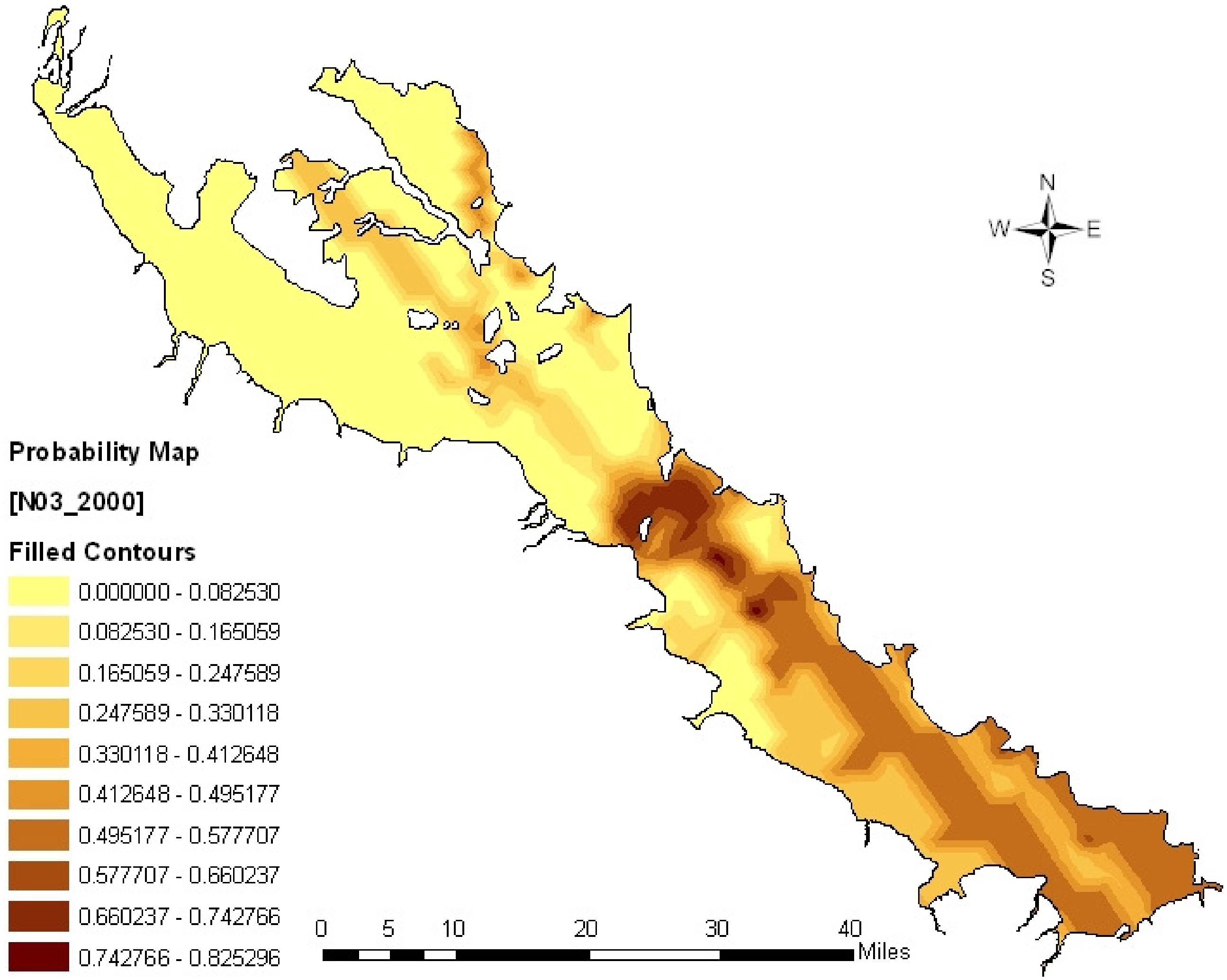


Figure E-3. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2000

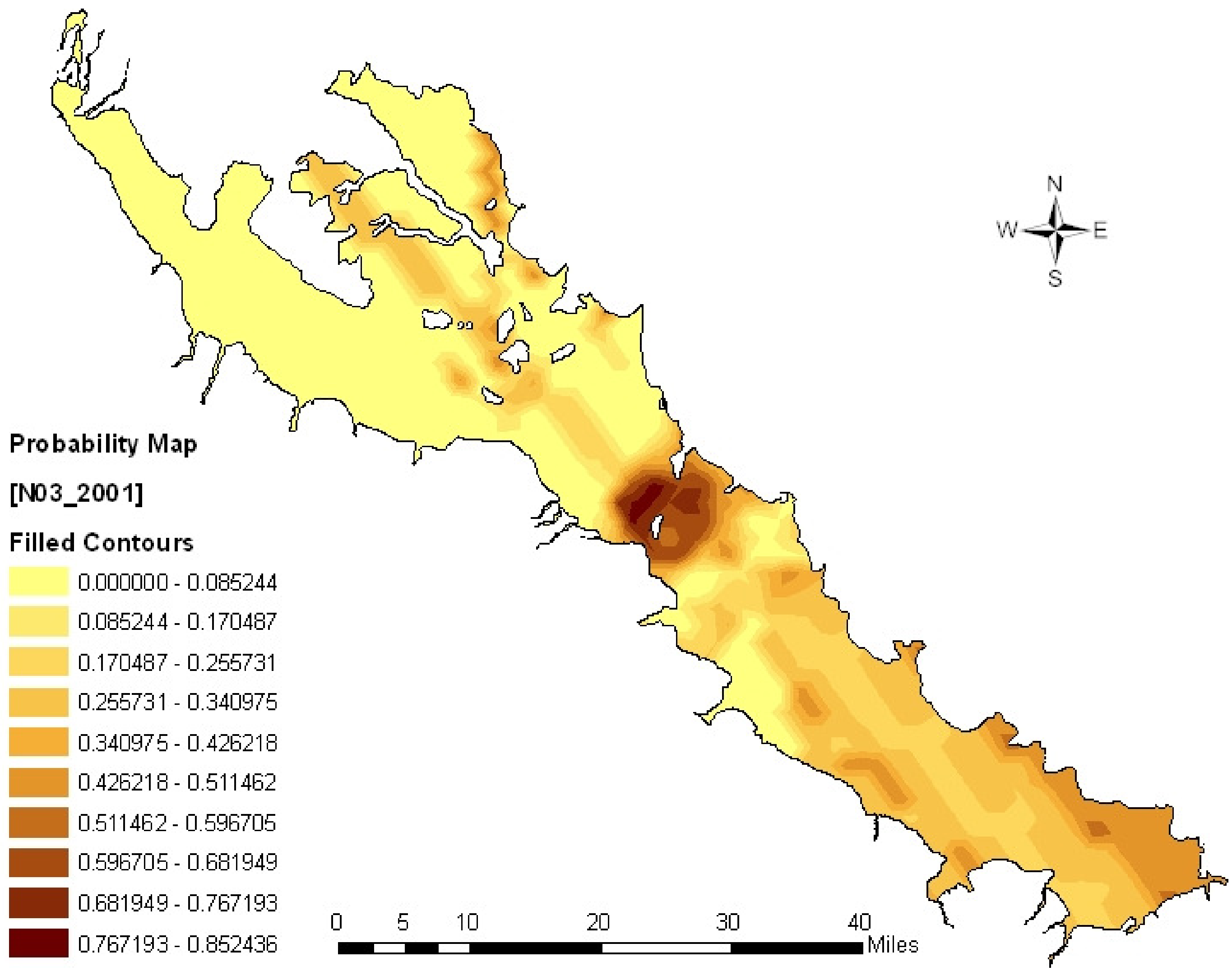


Figure E-4. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2001

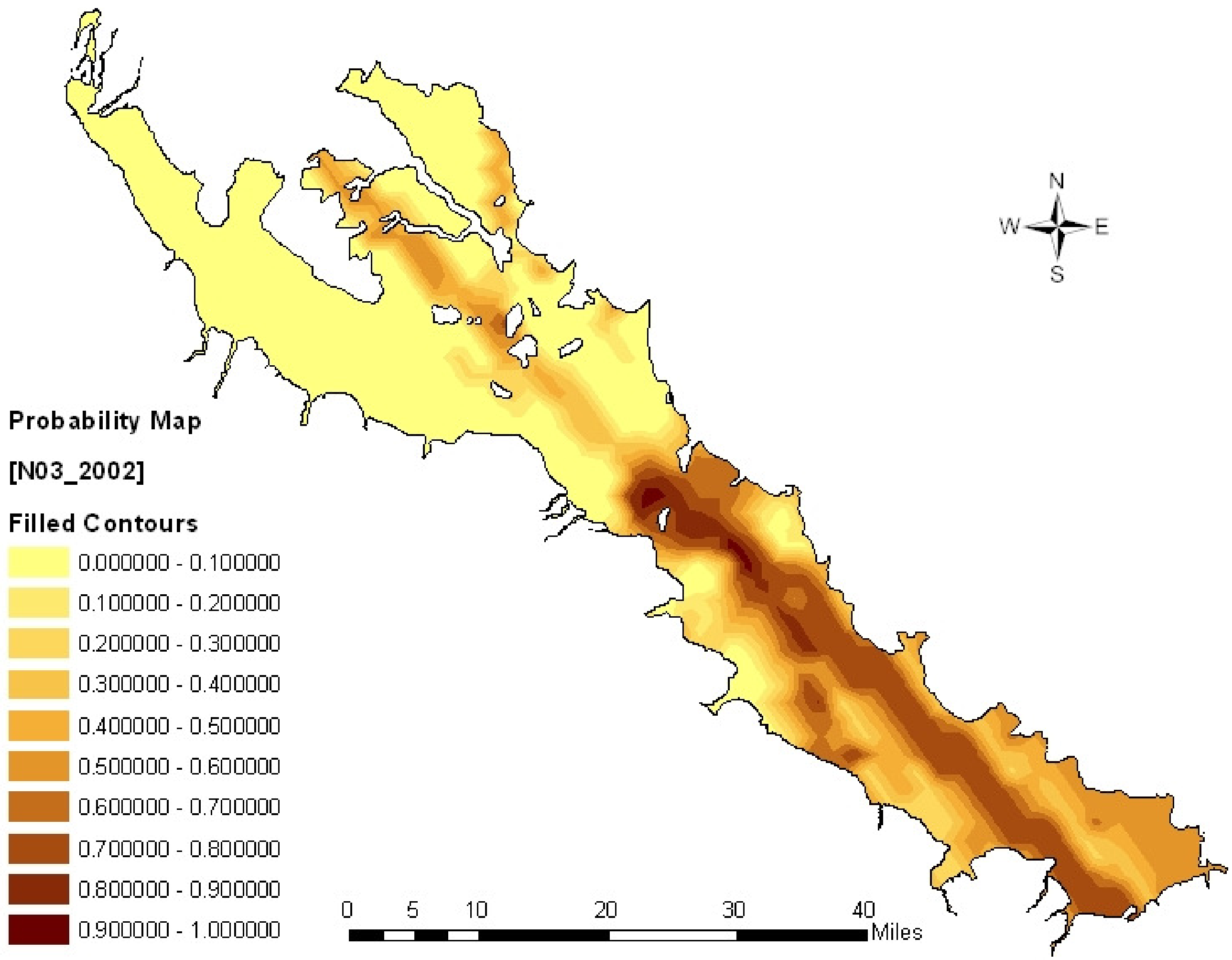


Figure E-5. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2002

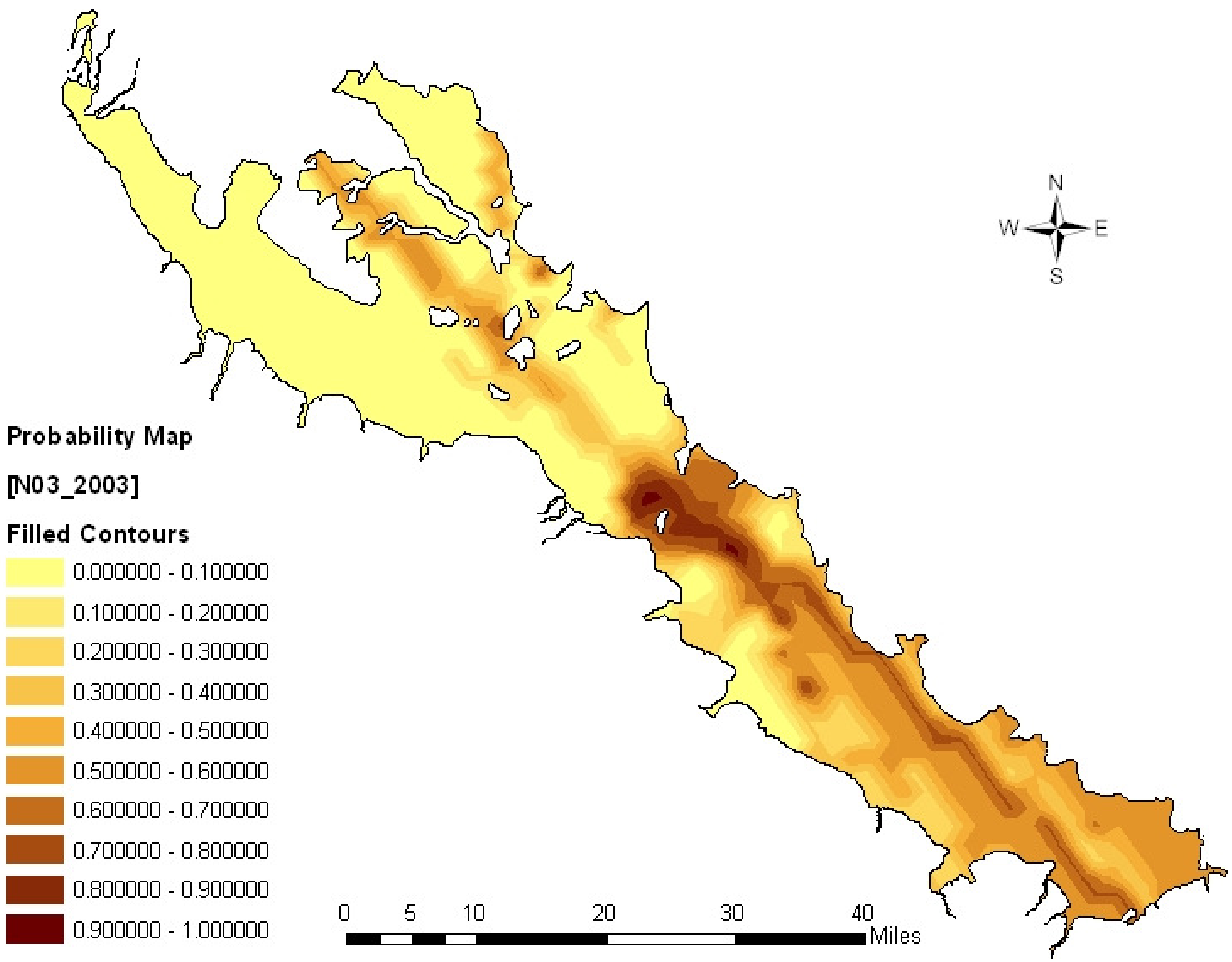


Figure E-6. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2003

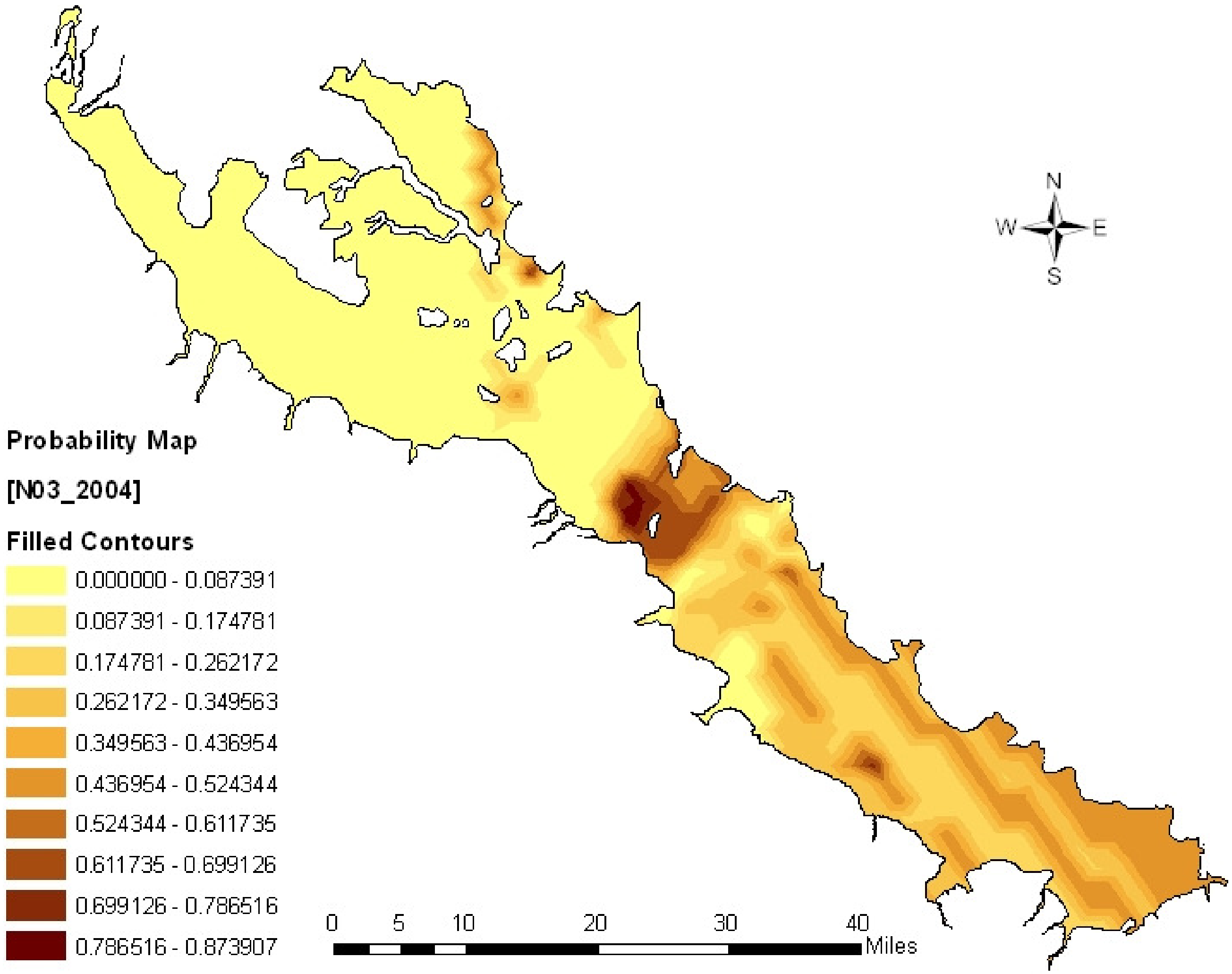


Figure E-7. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2004

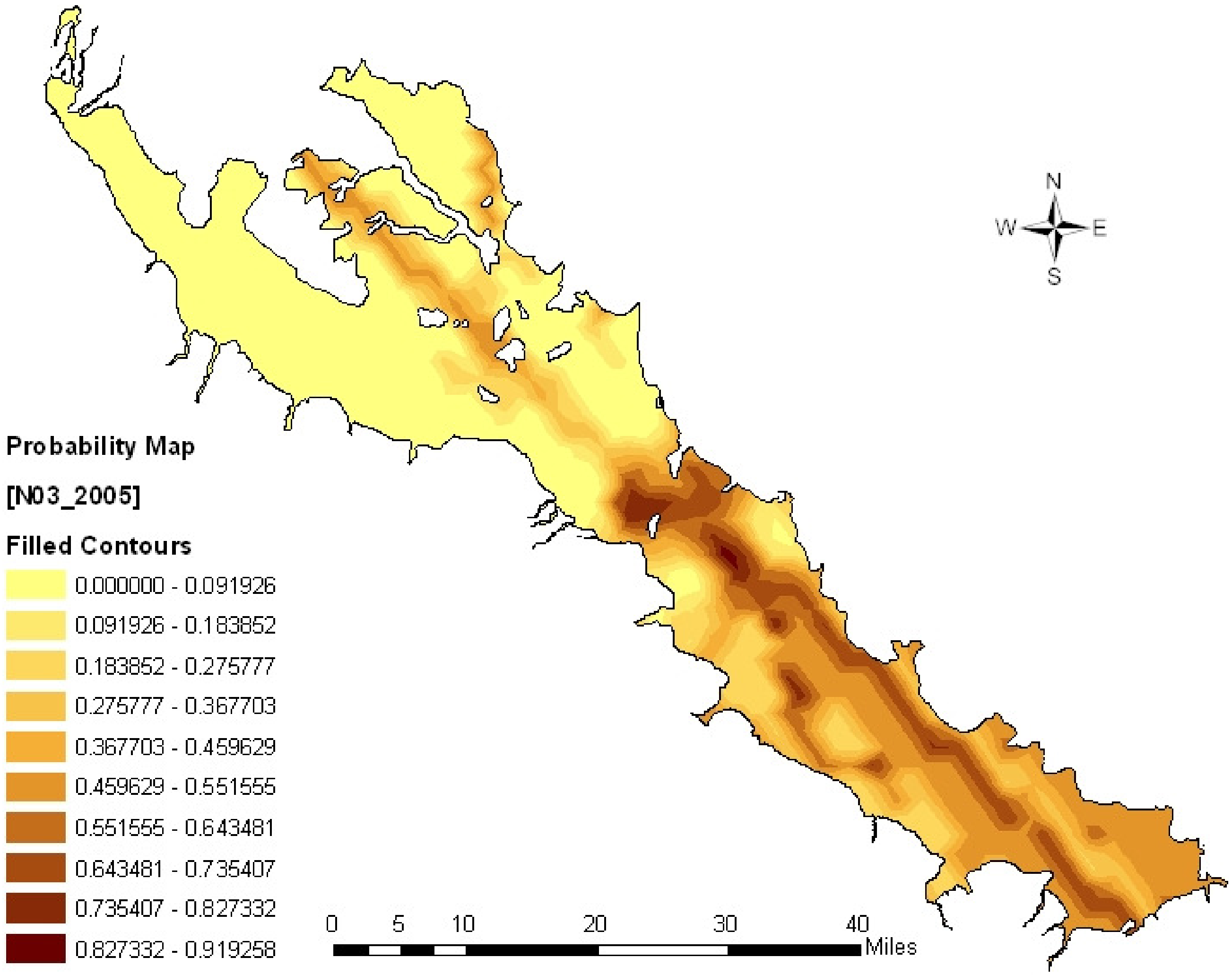


Figure E-8. Indicator probability map at threshold 10 mg/l of nitrate data sampled in 2005

APPENDIX F

LAND USE PROPORTIONS FROM BUFFER ANALYSIS

Table F-1. Percent of land uses at 100 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.175	0.000	0.825	0.000	11.7
A2	0.079	0.000	0.503	0.418	9.3
A3	0.000	0.000	0.659	0.341	3.4
A4	0.079	0.000	0.680	0.241	6.8
A5	0.275	0.000	0.725	0.000	5.2
A6	0.339	0.000	0.373	0.288	2.5
A7	0.143	0.000	0.857	0.000	4.0
A8	0.045	0.153	0.802	0.000	3.6
A9	0.249	0.000	0.751	0.000	1.8
A11	0.000	0.000	0.000	1.000	13.6
A14	0.085	0.167	0.421	0.328	13.0
A16	0.000	0.000	0.000	1.000	12.4
A17	0.106	0.000	0.415	0.479	15.6
A18	0.000	0.045	0.955	0.000	14.8
A19	0.048	0.302	0.265	0.386	14.0
A20	0.339	0.000	0.661	0.000	6.8
A21	0.127	0.003	0.497	0.373	10.7
A22	0.042	0.000	0.720	0.238	7.6
A23	0.156	0.000	0.317	0.526	9.4
A24	0.000	0.606	0.394	0.000	8.6
A25	0.000	0.056	0.926	0.019	13.4
A27	0.074	0.000	0.926	0.000	7.8
A29	0.003	0.000	0.997	0.000	9.4
CS1	0.000	0.243	0.757	0.000	3.0
CS2	0.000	0.190	0.810	0.000	1.3
CS3	0.177	0.000	0.122	0.701	3.9
CS4	0.093	0.000	0.907	0.000	4.7
CS5	0.000	0.000	1.000	0.000	7.4
CS6	0.000	0.000	0.905	0.095	6.8
CS8	0.000	0.000	1.000	0.000	4.6
CS9	0.000	0.000	1.000	0.000	7.5
CS10	0.000	0.000	0.000	1.000	8.7

Table F-1. Cont.

CS11	0.272	0.000	0.728	0.000	5.0
CS12	0.000	0.000	1.000	0.000	4.8
CS13	0.000	0.103	0.820	0.077	4.7
CS14	0.167	0.000	0.833	0.000	4.8
CS15	0.000	0.000	1.000	0.000	8.4
CS16	0.000	0.000	1.000	0.000	6.1
CS17	0.000	0.000	0.495	0.505	8.4
CS18	0.000	0.122	0.878	0.000	16.0
CS19	0.000	0.000	1.000	0.000	3.2
CS20	0.000	0.000	1.000	0.000	14.0
CS21	0.000	0.000	1.000	0.000	14.0
CS22	0.000	0.000	1.000	0.000	4.1
CS23	0.000	0.000	1.000	0.000	3.9
CS25	0.000	0.013	0.987	0.000	3.0
CS26	0.106	0.000	0.894	0.000	1.7
CS27	0.000	0.000	1.000	0.000	1.8
CS28	0.000	0.000	1.000	0.000	1.0
CS29	0.029	0.000	0.971	0.000	0.9
CS30	0.000	0.000	1.000	0.000	5.1
D1	0.140	0.000	0.082	0.778	3.6
D2	0.164	0.000	0.836	0.000	8.2
D3	0.000	0.000	0.294	0.706	8.9
D5	0.156	0.000	0.415	0.429	9.8
D6	0.262	0.000	0.000	0.738	6.8
D7	0.286	0.000	0.669	0.045	12.0
D8	0.000	0.000	0.000	1.000	10.5
D9	0.050	0.000	0.056	0.894	3.1
D10	0.127	0.000	0.201	0.672	4.9
D12	0.317	0.000	0.000	0.683	5.4
D18	0.167	0.000	0.122	0.712	8.9
D19	0.217	0.000	0.783	0.000	6.0
D20	0.138	0.000	0.646	0.217	8.5
D21	0.151	0.000	0.000	0.849	7.6
D23	0.114	0.000	0.000	0.886	10.7
D25	0.275	0.000	0.045	0.680	7.7
D27	0.138	0.000	0.000	0.862	6.9
D28	0.135	0.000	0.000	0.865	9.6
D29	0.000	0.000	0.000	1.000	9.3
D33	0.370	0.000	0.556	0.074	6.0
R1	0.000	0.000	1.000	0.000	7.4
R2	0.000	0.000	1.000	0.000	7.3
R3	0.053	0.302	0.646	0.000	9.2

Table F-1. Cont.

R4	0.000	0.045	0.955	0.000	8.4
R5	0.000	0.000	1.000	0.000	7.3
R6	0.000	0.294	0.706	0.000	6.8
R7	0.000	0.180	0.820	0.000	7.3
R8	0.000	0.000	1.000	0.000	9.0
R9	0.000	0.000	1.000	0.000	9.6
R10	0.000	0.000	1.000	0.000	7.0
R11	0.172	0.000	0.828	0.000	6.5
R12	0.000	0.728	0.272	0.000	6.5
R13	0.000	0.487	0.513	0.000	7.7
R14	0.188	0.000	0.812	0.000	12.1
R15	0.280	0.000	0.720	0.000	5.3
R16	0.283	0.183	0.534	0.000	8.4
R17	0.196	0.243	0.561	0.000	4.7
R18	0.000	0.000	1.000	0.000	5.1
R19	0.000	0.063	0.937	0.000	5.1
R20	0.021	0.071	0.907	0.000	10.8
R21	0.130	0.188	0.683	0.000	13.3
R22	0.058	0.000	0.942	0.000	4.2
R24	0.000	0.460	0.540	0.000	3.5
R25	0.000	0.135	0.865	0.000	3.7
R26	0.000	0.048	0.952	0.000	11.2
R27	0.000	0.000	1.000	0.000	7.1
R28	0.177	0.000	0.823	0.000	15.6

Table F-2. Percent of land uses at 250 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.121	0.024	0.734	0.120	11.7
A2	0.051	0.046	0.232	0.671	9.3
A3	0.082	0.048	0.602	0.268	3.4
A4	0.039	0.000	0.578	0.383	6.8
A5	0.180	0.000	0.760	0.060	5.2
A6	0.199	0.000	0.337	0.464	2.5
A7	0.159	0.000	0.768	0.073	4.0
A8	0.048	0.115	0.837	0.000	3.6
A9	0.158	0.061	0.775	0.006	1.8
A11	0.027	0.008	0.009	0.955	13.6
A14	0.086	0.169	0.344	0.401	13.0
A16	0.093	0.000	0.061	0.846	12.4
A17	0.104	0.000	0.397	0.499	15.6
A18	0.044	0.178	0.778	0.000	14.8
A19	0.096	0.139	0.086	0.679	14.0
A20	0.186	0.000	0.811	0.003	6.8
A21	0.146	0.064	0.552	0.238	10.7
A22	0.125	0.000	0.592	0.283	7.6
A23	0.092	0.000	0.632	0.276	9.4
A24	0.048	0.423	0.529	0.000	8.6
A25	0.046	0.123	0.673	0.158	13.4
A27	0.029	0.216	0.755	0.000	7.8
A29	0.025	0.000	0.975	0.000	9.4
CS1	0.000	0.121	0.879	0.000	3.0
CS2	0.000	0.158	0.842	0.000	1.3
CS3	0.085	0.000	0.415	0.500	3.9
CS4	0.040	0.000	0.901	0.059	4.7
CS5	0.000	0.000	1.000	0.000	7.4
CS6	0.096	0.000	0.641	0.262	6.8
CS8	0.080	0.000	0.920	0.000	4.6
CS9	0.040	0.000	0.863	0.096	7.5
CS10	0.000	0.000	0.128	0.872	8.7
CS11	0.090	0.000	0.910	0.000	5.0
CS12	0.000	0.000	0.998	0.002	4.8
CS13	0.000	0.159	0.620	0.221	4.7
CS14	0.092	0.008	0.901	0.000	4.8

Table F-2. Cont.

CS15	0.006	0.000	0.994	0.000	8.4
CS16	0.000	0.000	1.000	0.000	6.1
CS17	0.000	0.000	0.483	0.517	8.4
CS18	0.000	0.284	0.716	0.000	16.0
CS19	0.000	0.068	0.932	0.000	3.2
CS20	0.000	0.000	1.000	0.000	14.0
CS21	0.000	0.000	1.000	0.000	14.0
CS22	0.000	0.000	1.000	0.000	4.1
CS23	0.070	0.012	0.918	0.000	3.9
CS25	0.000	0.044	0.956	0.000	3.0
CS26	0.069	0.000	0.931	0.000	1.7
CS27	0.000	0.000	1.000	0.000	1.8
CS28	0.028	0.000	0.972	0.000	1.0
CS29	0.077	0.000	0.923	0.000	0.9
CS30	0.067	0.020	0.912	0.000	5.1
D1	0.074	0.000	0.296	0.630	3.6
D2	0.135	0.006	0.109	0.751	8.2
D3	0.000	0.000	0.405	0.595	8.9
D5	0.098	0.000	0.507	0.395	9.8
D6	0.196	0.000	0.000	0.804	6.8
D7	0.147	0.000	0.594	0.260	12.0
D8	0.000	0.000	0.071	0.929	10.5
D9	0.068	0.000	0.601	0.331	3.1
D10	0.073	0.000	0.341	0.586	4.9
D12	0.088	0.000	0.036	0.876	5.4
D18	0.076	0.000	0.216	0.708	8.9
D19	0.160	0.000	0.780	0.060	6.0
D20	0.063	0.002	0.642	0.293	8.5
D21	0.126	0.000	0.035	0.839	7.6
D23	0.072	0.000	0.013	0.915	10.7
D25	0.155	0.034	0.325	0.485	7.7
D27	0.115	0.000	0.036	0.850	6.9
D28	0.077	0.000	0.015	0.907	9.6
D29	0.080	0.000	0.002	0.918	9.3
D33	0.252	0.051	0.467	0.230	6.0
R1	0.000	0.000	1.000	0.000	7.4
R2	0.027	0.000	0.973	0.000	7.3
R3	0.091	0.141	0.769	0.000	9.2
R4	0.041	0.076	0.883	0.000	8.4
R5	0.082	0.000	0.918	0.000	7.3

Table F-2. Cont.

R6	0.000	0.182	0.818	0.000	6.8
R7	0.000	0.226	0.774	0.000	7.3
R8	0.056	0.000	0.944	0.000	9.0
R9	0.025	0.000	0.975	0.000	9.6
R10	0.000	0.045	0.955	0.000	7.0
R11	0.106	0.029	0.864	0.000	6.5
R12	0.058	0.670	0.272	0.000	6.5
R13	0.096	0.268	0.635	0.000	7.7
R14	0.115	0.009	0.876	0.000	12.1
R15	0.127	0.025	0.848	0.000	5.3
R16	0.135	0.114	0.751	0.000	8.4
R17	0.084	0.405	0.511	0.000	4.7
R18	0.028	0.116	0.857	0.000	5.1
R19	0.027	0.166	0.807	0.000	5.1
R20	0.022	0.171	0.807	0.000	10.8
R21	0.095	0.241	0.664	0.000	13.3
R22	0.082	0.000	0.918	0.000	4.2
R24	0.000	0.304	0.696	0.000	3.5
R25	0.000	0.112	0.888	0.000	3.7
R26	0.000	0.031	0.969	0.000	11.2
R27	0.010	0.043	0.947	0.000	7.1
R28	0.134	0.000	0.829	0.037	15.6

Table F-3. Percent of land uses at 500 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.085	0.066	0.617	0.232	11.7
A2	0.047	0.070	0.299	0.584	9.3
A3	0.072	0.122	0.619	0.187	3.4
A4	0.026	0.063	0.623	0.288	6.8
A5	0.101	0.030	0.705	0.163	5.2
A6	0.102	0.019	0.360	0.519	2.5
A7	0.107	0.073	0.659	0.161	4.0
A8	0.064	0.216	0.721	0.000	3.6
A9	0.101	0.247	0.576	0.076	1.8
A11	0.034	0.027	0.184	0.755	13.6
A14	0.061	0.108	0.360	0.471	13.0
A16	0.059	0.023	0.097	0.820	12.4
A17	0.079	0.005	0.394	0.521	15.6
A18	0.069	0.096	0.621	0.214	14.8
A19	0.054	0.067	0.217	0.661	14.0
A20	0.142	0.002	0.686	0.170	6.8
A21	0.124	0.033	0.556	0.287	10.7
A22	0.132	0.000	0.656	0.212	7.6
A23	0.080	0.035	0.701	0.185	9.4
A24	0.083	0.140	0.720	0.056	8.6
A25	0.071	0.103	0.629	0.197	13.4
A27	0.052	0.197	0.751	0.000	7.8
A29	0.028	0.020	0.951	0.000	9.4
CS1	0.005	0.089	0.906	0.000	3.0
CS2	0.011	0.142	0.816	0.031	1.3
CS3	0.034	0.000	0.636	0.330	3.9
CS4	0.023	0.014	0.837	0.125	4.7
CS5	0.041	0.000	0.852	0.107	7.4
CS6	0.080	0.000	0.564	0.356	6.8
CS8	0.083	0.000	0.752	0.165	4.6
CS9	0.022	0.000	0.747	0.230	7.5
CS10	0.011	0.008	0.498	0.482	8.7
CS11	0.064	0.015	0.921	0.000	5.0
CS12	0.034	0.016	0.882	0.068	4.8
CS13	0.061	0.084	0.685	0.170	4.7
CS14	0.044	0.116	0.753	0.087	4.8
CS15	0.039	0.008	0.953	0.000	8.4
CS16	0.025	0.011	0.959	0.004	6.1
CS17	0.017	0.000	0.581	0.401	8.4

Table F-3. Cont.

CS18	0.000	0.104	0.742	0.154	16.0
CS19	0.000	0.045	0.955	0.000	3.2
CS20	0.009	0.008	0.951	0.031	14.0
CS21	0.005	0.000	0.995	0.000	14.0
CS22	0.012	0.000	0.988	0.000	4.1
CS23	0.049	0.013	0.938	0.000	3.9
CS25	0.000	0.024	0.976	0.000	3.0
CS26	0.050	0.021	0.929	0.000	1.7
CS27	0.008	0.018	0.974	0.000	1.8
CS28	0.021	0.009	0.969	0.000	1.0
CS29	0.067	0.000	0.933	0.000	0.9
CS30	0.039	0.015	0.946	0.000	5.1
D1	0.039	0.000	0.311	0.650	3.6
D2	0.090	0.019	0.255	0.636	8.2
D3	0.051	0.000	0.490	0.458	8.9
D5	0.045	0.000	0.547	0.407	9.8
D6	0.110	0.020	0.000	0.870	6.8
D7	0.075	0.000	0.346	0.579	12.0
D8	0.034	0.000	0.185	0.781	10.5
D9	0.039	0.000	0.654	0.308	3.1
D10	0.074	0.000	0.451	0.475	4.9
D12	0.034	0.000	0.030	0.936	5.4
D18	0.038	0.000	0.214	0.748	8.9
D19	0.078	0.000	0.730	0.192	6.0
D20	0.047	0.013	0.528	0.412	8.5
D21	0.081	0.000	0.040	0.879	7.6
D23	0.037	0.000	0.123	0.840	10.7
D25	0.084	0.028	0.335	0.553	7.7
D27	0.062	0.000	0.060	0.879	6.9
D28	0.063	0.040	0.052	0.845	9.6
D29	0.038	0.015	0.044	0.903	9.3
D33	0.124	0.103	0.227	0.546	6.0
R1	0.010	0.017	0.973	0.000	7.4
R2	0.017	0.024	0.959	0.000	7.3
R3	0.058	0.051	0.891	0.000	9.2
R4	0.067	0.063	0.870	0.000	8.4
R5	0.035	0.000	0.964	0.000	7.3
R6	0.026	0.076	0.898	0.000	6.8
R7	0.072	0.083	0.846	0.000	7.3
R8	0.078	0.001	0.920	0.000	9.0
R9	0.020	0.000	0.914	0.066	9.6
R10	0.046	0.083	0.872	0.000	7.0

Table F-3. Cont.

R11	0.074	0.159	0.767	0.000	6.5
R12	0.090	0.431	0.478	0.000	6.5
R13	0.067	0.195	0.737	0.000	7.7
R14	0.090	0.037	0.868	0.005	12.1
R15	0.077	0.040	0.882	0.000	5.3
R16	0.128	0.165	0.706	0.000	8.4
R17	0.056	0.441	0.503	0.000	4.7
R18	0.041	0.256	0.703	0.000	5.1
R19	0.078	0.126	0.797	0.000	5.1
R20	0.087	0.241	0.672	0.000	10.8
R21	0.072	0.192	0.560	0.176	13.3
R22	0.072	0.051	0.876	0.000	4.2
R24	0.041	0.172	0.787	0.000	3.5
R25	0.035	0.056	0.909	0.000	3.7
R26	0.047	0.026	0.908	0.018	11.2
R27	0.013	0.019	0.969	0.000	7.1
R28	0.146	0.024	0.649	0.180	15.6

Table F-4. Percent of land covers at 750 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.059	0.055	0.563	0.324	11.7
A2	0.067	0.043	0.285	0.605	9.3
A3	0.072	0.110	0.582	0.235	3.4
A4	0.053	0.081	0.633	0.233	6.8
A5	0.096	0.116	0.602	0.186	5.2
A6	0.059	0.076	0.420	0.445	2.5
A7	0.081	0.110	0.547	0.262	4.0
A8	0.063	0.225	0.704	0.009	3.6
A9	0.082	0.268	0.576	0.074	1.8
A11	0.037	0.042	0.316	0.605	13.6
A14	0.056	0.055	0.332	0.557	13.0
A16	0.068	0.011	0.129	0.793	12.4
A17	0.075	0.038	0.291	0.595	15.6
A18	0.056	0.093	0.433	0.418	14.8
A19	0.045	0.088	0.269	0.597	14.0
A20	0.109	0.061	0.648	0.183	6.8
A21	0.098	0.096	0.521	0.284	10.7
A22	0.103	0.044	0.679	0.174	7.6
A23	0.067	0.069	0.664	0.200	9.4
A24	0.081	0.097	0.748	0.073	8.6
A25	0.083	0.075	0.614	0.227	13.4
A27	0.077	0.125	0.785	0.013	7.8
A29	0.034	0.023	0.943	0.000	9.4
CS1	0.019	0.068	0.894	0.019	3.0
CS2	0.036	0.090	0.874	0.000	1.3
CS3	0.035	0.021	0.738	0.206	3.9
CS4	0.016	0.015	0.886	0.082	4.7
CS5	0.057	0.000	0.774	0.168	7.4
CS6	0.078	0.681	0.241	0.000	6.8
CS8	0.071	0.000	0.677	0.253	4.6
CS9	0.027	0.006	0.759	0.208	7.5
CS10	0.037	0.008	0.733	0.222	8.7
CS11	0.051	0.019	0.906	0.024	5.0
CS12	0.032	0.018	0.843	0.108	4.8
CS13	0.060	0.064	0.703	0.173	4.7
CS14	0.045	0.074	0.674	0.207	4.8
CS15	0.041	0.027	0.876	0.056	8.4
CS16	0.021	0.027	0.877	0.075	6.1

Table F-4. Cont.

CS17	0.031	0.029	0.652	0.288	8.4
CS18	0.006	0.051	0.728	0.215	16.0
CS19	0.007	0.048	0.901	0.044	3.2
CS20	0.021	0.035	0.866	0.078	14.0
CS21	0.024	0.001	0.953	0.021	14.0
CS22	0.016	0.001	0.982	0.000	4.1
CS23	0.036	0.010	0.911	0.043	3.9
CS25	0.020	0.016	0.964	0.000	3.0
CS26	0.029	0.015	0.956	0.000	1.7
CS27	0.011	0.011	0.979	0.000	1.8
CS28	0.017	0.018	0.965	0.000	1.0
CS29	0.051	0.007	0.942	0.000	0.9
CS30	0.037	0.013	0.950	0.000	5.1
D1	0.043	0.000	0.404	0.553	3.6
D2	0.065	0.012	0.248	0.675	8.2
D3	0.059	0.008	0.440	0.492	8.9
D5	0.029	0.000	0.517	0.453	9.8
D6	0.083	0.009	0.071	0.837	6.8
D7	0.065	0.000	0.207	0.727	12.0
D8	0.041	0.000	0.306	0.652	10.5
D9	0.040	0.000	0.611	0.348	3.1
D10	0.065	0.000	0.435	0.499	4.9
D12	0.039	0.003	0.043	0.916	5.4
D18	0.035	0.000	0.310	0.655	8.9
D19	0.052	0.000	0.622	0.325	6.0
D20	0.053	0.006	0.421	0.520	8.5
D21	0.054	0.004	0.062	0.881	7.6
D23	0.045	0.000	0.170	0.785	10.7
D25	0.057	0.079	0.231	0.634	7.7
D27	0.048	0.003	0.126	0.823	6.9
D28	0.054	0.018	0.066	0.862	9.6
D29	0.037	0.018	0.039	0.906	9.3
D33	0.096	0.100	0.128	0.676	6.0
R1	0.025	0.032	0.943	0.000	7.4
R2	0.021	0.044	0.936	0.000	7.3
R3	0.041	0.046	0.913	0.000	9.2
R4	0.064	0.043	0.865	0.028	8.4
R5	0.022	0.056	0.922	0.000	7.3
R6	0.036	0.051	0.913	0.000	6.8
R7	0.058	0.135	0.807	0.000	7.3
R8	0.063	0.035	0.902	0.000	9.0
R9	0.027	0.025	0.815	0.134	9.6

Table F-4. Cont.

R10	0.056	0.073	0.871	0.000	7.0
R11	0.068	0.186	0.746	0.000	6.5
R12	0.078	0.299	0.624	0.000	6.5
R13	0.072	0.208	0.720	0.000	7.7
R14	0.061	0.054	0.818	0.068	12.1
R15	0.063	0.048	0.889	0.000	5.3
R16	0.089	0.175	0.736	0.000	8.4
R17	0.075	0.304	0.621	0.000	4.7
R18	0.067	0.258	0.676	0.000	5.1
R19	0.063	0.187	0.750	0.000	5.1
R20	0.067	0.267	0.623	0.043	10.8
R21	0.059	0.190	0.489	0.263	13.3
R22	0.091	0.082	0.827	0.000	4.2
R24	0.068	0.153	0.780	0.000	3.5
R25	0.026	0.059	0.916	0.000	3.7
R26	0.054	0.044	0.851	0.052	11.2
R27	0.024	0.024	0.952	0.000	7.1
R28	0.125	0.031	0.570	0.274	15.6

Table F-5. Percent of land covers at 1,000 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.047	0.039	0.432	0.482	11.7
A2	0.061	0.057	0.311	0.571	9.3
A3	0.062	0.095	0.545	0.298	3.4
A4	0.070	0.101	0.561	0.268	6.8
A5	0.092	0.176	0.536	0.196	5.2
A6	0.068	0.095	0.438	0.399	2.5
A7	0.059	0.104	0.501	0.337	4.0
A8	0.064	0.189	0.608	0.139	3.6
A9	0.064	0.208	0.588	0.140	1.8
A11	0.055	0.053	0.400	0.492	13.6
A14	0.050	0.048	0.321	0.580	13.0
A16	0.062	0.006	0.213	0.719	12.4
A17	0.060	0.074	0.246	0.619	15.6
A18	0.046	0.078	0.395	0.482	14.8
A19	0.060	0.107	0.305	0.528	14.0
A20	0.087	0.172	0.594	0.147	6.8
A21	0.080	0.117	0.579	0.224	10.7
A22	0.079	0.114	0.641	0.166	7.6
A23	0.070	0.092	0.678	0.161	9.4
A24	0.074	0.097	0.750	0.079	8.6
A25	0.066	0.064	0.692	0.177	13.4
A27	0.074	0.107	0.791	0.028	7.8
A29	0.064	0.032	0.903	0.000	9.4
CS1	0.023	0.066	0.888	0.023	3.0
CS2	0.030	0.051	0.907	0.012	1.3
CS3	0.041	0.042	0.778	0.138	3.9
CS4	0.024	0.053	0.837	0.086	4.7
CS5	0.055	0.004	0.790	0.151	7.4
CS6	0.065	0.005	0.754	0.176	6.8
CS8	0.061	0	0.710	0.229	4.6
CS9	0.044	0.003	0.801	0.152	7.5
CS10	0.045	0.020	0.778	0.157	8.7
CS11	0.043	0.022	0.882	0.053	5.0
CS12	0.029	0.028	0.813	0.130	4.8
CS13	0.052	0.043	0.722	0.182	4.7
CS14	0.051	0.043	0.661	0.245	4.8
CS15	0.045	0.033	0.807	0.115	8.4
CS16	0.030	0.045	0.785	0.141	6.1
CS17	0.049	0.062	0.715	0.174	8.4

Table F-5. Cont.

CS18	0.013	0.038	0.701	0.248	16.0
CS19	0.020	0.048	0.824	0.108	3.2
CS20	0.022	0.040	0.824	0.114	14.0
CS21	0.024	0.011	0.879	0.085	14.0
CS22	0.022	0.011	0.968	0.000	4.1
CS23	0.027	0.015	0.876	0.082	3.9
CS25	0.029	0.011	0.960	0.000	3.0
CS26	0.026	0.008	0.966	0.000	1.7
CS27	0.026	0.016	0.958	0.000	1.8
CS28	0.023	0.014	0.963	0.000	1.0
CS29	0.045	0.007	0.948	0.000	0.9
CS30	0.038	0.022	0.940	0.000	5.1
D1	0.044	0.000	0.474	0.482	3.6
D2	0.055	0.016	0.272	0.657	8.2
D3	0.048	0.004	0.370	0.578	8.9
D5	0.063	0.007	0.406	0.524	9.8
D6	0.066	0.024	0.207	0.703	6.8
D7	0.057	0.000	0.237	0.706	12.0
D8	0.061	0.001	0.331	0.606	10.5
D9	0.044	0.000	0.593	0.363	3.1
D10	0.055	0.000	0.437	0.508	4.9
D12	0.069	0.009	0.104	0.818	5.4
D18	0.052	0.000	0.405	0.543	8.9
D19	0.041	0.000	0.554	0.406	6.0
D20	0.045	0.003	0.370	0.582	8.5
D21	0.042	0.003	0.117	0.838	7.6
D23	0.042	0.012	0.184	0.762	10.7
D25	0.067	0.052	0.169	0.712	7.7
D27	0.052	0.014	0.107	0.827	6.9
D28	0.056	0.011	0.064	0.869	9.6
D29	0.062	0.008	0.050	0.880	9.3
D33	0.079	0.094	0.115	0.712	6.0
R1	0.027	0.100	0.874	0.000	7.4
R2	0.028	0.038	0.934	0.000	7.3
R3	0.043	0.053	0.873	0.031	9.2
R4	0.055	0.033	0.846	0.066	8.4
R5	0.025	0.085	0.890	0.000	7.3
R6	0.047	0.080	0.873	0.000	6.8
R7	0.065	0.145	0.790	0.000	7.3
R8	0.059	0.083	0.857	0.001	9.0
R9	0.047	0.054	0.760	0.138	9.6
R10	0.062	0.086	0.852	0.000	7.0

Table F-5. Cont.

R11	0.066	0.154	0.780	0.000	6.5
R12	0.066	0.211	0.723	0.000	6.5
R13	0.070	0.200	0.730	0.000	7.7
R14	0.056	0.093	0.751	0.100	12.1
R15	0.066	0.041	0.893	0.000	5.3
R16	0.070	0.159	0.770	0.000	8.4
R17	0.085	0.214	0.701	0.000	4.7
R18	0.074	0.215	0.711	0.000	5.1
R19	0.064	0.230	0.703	0.003	5.1
R20	0.053	0.238	0.606	0.102	10.8
R21	0.059	0.186	0.449	0.305	13.3
R22	0.085	0.135	0.780	0.000	4.2
R24	0.074	0.184	0.742	0.000	3.5
R25	0.031	0.052	0.917	0.000	3.7
R26	0.063	0.066	0.817	0.054	11.2
R27	0.039	0.027	0.917	0.017	7.1
R28	0.112	0.057	0.520	0.311	15.6

Table F-6. Percent of land covers at 1,500 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.040	0.032	0.301	0.627	11.7
A2	0.075	0.072	0.353	0.500	9.3
A3	0.050	0.103	0.452	0.395	3.4
A4	0.076	0.155	0.450	0.319	6.8
A5	0.072	0.185	0.478	0.264	5.2
A6	0.054	0.118	0.547	0.280	2.5
A7	0.053	0.110	0.502	0.335	4.0
A8	0.058	0.118	0.425	0.400	3.6
A9	0.061	0.140	0.490	0.310	1.8
A11	0.059	0.066	0.405	0.469	13.6
A14	0.051	0.039	0.297	0.613	13.0
A16	0.056	0.010	0.223	0.711	12.4
A17	0.051	0.080	0.264	0.605	15.6
A18	0.058	0.066	0.344	0.533	14.8
A19	0.049	0.164	0.311	0.475	14.0
A20	0.063	0.240	0.546	0.151	6.8
A21	0.065	0.180	0.548	0.207	10.7
A22	0.069	0.195	0.597	0.139	7.6
A23	0.065	0.090	0.662	0.182	9.4
A24	0.068	0.091	0.722	0.119	8.6
A25	0.055	0.098	0.674	0.172	13.4
A27	0.069	0.115	0.710	0.107	7.8
A29	0.057	0.087	0.840	0.017	9.4
CS1	0.030	0.107	0.800	0.063	3.0
CS2	0.033	0.072	0.872	0.022	1.3
CS3	0.042	0.053	0.765	0.140	3.9
CS4	0.038	0.081	0.800	0.081	4.7
CS5	0.042	0.050	0.820	0.087	7.4
CS6	0.049	0.025	0.781	0.145	6.8
CS8	0.045	0.013	0.800	0.142	4.6
CS9	0.045	0.019	0.812	0.123	7.5
CS10	0.048	0.040	0.776	0.137	8.7
CS11	0.042	0.040	0.779	0.140	5.0
CS12	0.050	0.020	0.807	0.123	4.8
CS13	0.040	0.032	0.795	0.133	4.7
CS14	0.048	0.026	0.783	0.143	4.8
CS15	0.043	0.048	0.671	0.238	8.4
CS16	0.045	0.048	0.725	0.182	6.1

Table F-6. Cont.

CS17	0.047	0.078	0.780	0.096	8.4
CS18	0.043	0.030	0.641	0.286	16.0
CS19	0.036	0.054	0.731	0.178	3.2
CS20	0.029	0.027	0.772	0.171	14.0
CS21	0.031	0.026	0.836	0.108	14.0
CS22	0.037	0.037	0.885	0.041	4.1
CS23	0.043	0.016	0.792	0.148	3.9
CS25	0.032	0.013	0.914	0.041	3.0
CS26	0.030	0.018	0.951	0.000	1.7
CS27	0.030	0.012	0.955	0.002	1.8
CS28	0.032	0.008	0.959	0.000	1.0
CS29	0.042	0.028	0.926	0.004	0.9
CS30	0.038	0.011	0.950	0.000	5.1
D1	0.037	0.001	0.399	0.563	3.6
D2	0.049	0.009	0.305	0.638	8.2
D3	0.061	0.014	0.334	0.591	8.9
D5	0.051	0.009	0.404	0.536	9.8
D6	0.045	0.023	0.304	0.629	6.8
D7	0.048	0.008	0.320	0.624	12.0
D8	0.057	0.003	0.405	0.535	10.5
D9	0.044	0.495	0.461	0.000	3.1
D10	0.044	0.467	0.489	0.000	4.9
D12	0.064	0.016	0.219	0.700	5.4
D18	0.037	0.001	0.375	0.587	8.9
D19	0.030	0.496	0.474	0.000	6.0
D20	0.029	0.027	0.772	0.171	8.5
D21	0.038	0.007	0.211	0.744	7.6
D23	0.035	0.033	0.149	0.782	10.7
D25	0.055	0.034	0.223	0.688	7.7
D27	0.052	0.056	0.125	0.766	6.9
D28	0.046	0.019	0.073	0.862	9.6
D29	0.056	0.011	0.061	0.872	9.3
D33	0.053	0.063	0.131	0.753	6.0
R1	0.034	0.159	0.805	0.001	7.4
R2	0.045	0.087	0.831	0.037	7.3
R3	0.052	0.058	0.810	0.080	9.2
R4	0.049	0.069	0.757	0.125	8.4
R5	0.050	0.151	0.799	0.000	7.3
R6	0.048	0.112	0.840	0.000	6.8
R7	0.053	0.107	0.840	0.000	7.3
R8	0.056	0.120	0.762	0.061	9.0
R9	0.067	0.118	0.692	0.123	9.6

Table F-6. Cont.

R10	0.049	0.125	0.827	0.000	7.0
R11	0.056	0.112	0.832	0.000	6.5
R12	0.070	0.122	0.808	0.000	6.5
R13	0.060	0.147	0.753	0.040	7.7
R14	0.057	0.177	0.672	0.095	12.1
R15	0.061	0.101	0.838	0.000	5.3
R16	0.063	0.117	0.820	0.000	8.4
R17	0.074	0.124	0.802	0.000	4.7
R18	0.068	0.169	0.740	0.023	5.1
R19	0.060	0.218	0.667	0.054	5.1
R20	0.056	0.225	0.561	0.159	10.8
R21	0.062	0.202	0.466	0.270	13.3
R22	0.078	0.122	0.800	0.001	4.2
R24	0.075	0.190	0.735	0.001	3.5
R25	0.050	0.127	0.802	0.021	3.7
R26	0.051	0.107	0.762	0.080	11.2
R27	0.055	0.124	0.788	0.034	7.1
R28	0.088	0.082	0.526	0.304	15.6

Table F-7. Percent of land covers at 1,750 m well buffer

well id	developed land	forested-wetland	grassland	cultivated land	nitrate as nitrogen, mg/l
A1	0.049	0.027	0.255	0.669	11.7
A2	0.084	0.076	0.363	0.478	9.3
A3	0.060	0.099	0.434	0.407	3.4
A4	0.069	0.138	0.462	0.332	6.8
A5	0.079	0.167	0.481	0.274	5.2
A6	0.060	0.124	0.550	0.266	2.5
A7	0.061	0.111	0.495	0.332	4.0
A8	0.062	0.096	0.431	0.411	3.6
A9	0.066	0.133	0.465	0.336	1.8
A11	0.061	0.071	0.409	0.460	13.6
A14	0.049	0.032	0.276	0.643	13.0
A16	0.055	0.012	0.255	0.678	12.4
A17	0.050	0.070	0.233	0.647	15.6
A18	0.066	0.063	0.319	0.553	14.8
A19	0.049	0.168	0.302	0.481	14.0
A20	0.062	0.226	0.553	0.159	6.8
A21	0.065	0.188	0.542	0.205	10.7
A22	0.065	0.200	0.590	0.146	7.6
A23	0.066	0.097	0.681	0.157	9.4
A24	0.066	0.104	0.705	0.126	8.6
A25	0.059	0.121	0.645	0.174	13.4
A27	0.065	0.111	0.713	0.111	7.8
A29	0.057	0.100	0.798	0.045	9.4
CS1	0.032	0.101	0.773	0.094	3.0
CS2	0.032	0.099	0.816	0.054	1.3
CS3	0.043	0.065	0.742	0.150	3.9
CS4	0.037	0.079	0.779	0.106	4.7
CS5	0.039	0.056	0.816	0.089	7.4
CS6	0.041	0.035	0.798	0.126	6.8
CS8	0.045	0.021	0.815	0.119	4.6
CS9	0.053	0.025	0.792	0.130	7.5
CS10	0.049	0.037	0.784	0.131	8.7
CS11	0.039	0.039	0.762	0.160	5.0
CS12	0.052	0.018	0.811	0.120	4.8
CS13	0.043	0.036	0.801	0.120	4.7
CS14	0.043	0.033	0.805	0.120	4.8
CS15	0.044	0.045	0.649	0.263	8.4
CS16	0.049	0.044	0.700	0.207	6.1

Table F-7. Cont.

CS17	0.048	0.075	0.767	0.110	8.4
CS18	0.044	0.036	0.645	0.275	16.0
CS19	0.048	0.056	0.718	0.178	3.2
CS20	0.041	0.036	0.738	0.185	14.0
CS21	0.038	0.026	0.808	0.128	14.0
CS22	0.036	0.047	0.860	0.057	4.1
CS23	0.042	0.021	0.776	0.161	3.9
CS25	0.037	0.016	0.890	0.057	3.0
CS26	0.038	0.028	0.925	0.009	1.7
CS27	0.034	0.016	0.943	0.006	1.8
CS28	0.033	0.009	0.958	0.000	1.0
CS29	0.038	0.029	0.914	0.020	0.9
CS30	0.038	0.025	0.922	0.016	5.1
D1	0.041	0.001	0.385	0.572	3.6
D2	0.055	0.011	0.286	0.648	8.2
D3	0.056	0.013	0.343	0.589	8.9
D5	0.054	0.018	0.394	0.534	9.8
D6	0.052	0.029	0.303	0.616	6.8
D7	0.051	0.016	0.402	0.531	12.0
D8	0.052	0.011	0.431	0.506	10.5
D9	0.047	0.456	0.497	0.000	3.1
D10	0.046	0.001	0.476	0.477	4.9
D12	0.059	0.023	0.236	0.682	5.4
D18	0.042	0.003	0.366	0.589	8.9
D19	0.045	0.440	0.515	0.000	6.0
D20	0.047	0.003	0.361	0.590	8.5
D21	0.046	0.010	0.255	0.689	7.6
D23	0.039	0.036	0.172	0.753	10.7
D25	0.058	0.034	0.257	0.651	7.7
D27	0.062	0.053	0.119	0.766	6.9
D28	0.054	0.040	0.088	0.818	9.6
D29	0.053	0.019	0.081	0.847	9.3
D33	0.062	0.054	0.139	0.744	6.0
R1	0.044	0.170	0.769	0.017	7.4
R2	0.047	0.103	0.784	0.065	7.3
R3	0.052	0.086	0.762	0.101	9.2
R4	0.049	0.081	0.708	0.162	8.4
R5	0.048	0.161	0.788	0.003	7.3
R6	0.051	0.118	0.822	0.008	6.8
R7	0.052	0.097	0.830	0.021	7.3
R8	0.057	0.125	0.736	0.082	9.0
R9	0.063	0.135	0.664	0.137	9.6

Table F-7. Cont.

R10	0.051	0.149	0.799	0.000	7.0
R11	0.055	0.107	0.838	0.000	6.5
R12	0.061	0.104	0.826	0.009	6.5
R13	0.062	0.148	0.740	0.050	7.7
R14	0.063	0.174	0.666	0.096	12.1
R15	0.060	0.121	0.819	0.000	5.3
R16	0.061	0.103	0.836	0.000	8.4
R17	0.070	0.110	0.820	0.000	4.7
R18	0.065	0.170	0.726	0.039	5.1
R19	0.062	0.200	0.665	0.072	5.1
R20	0.062	0.213	0.570	0.156	10.8
R21	0.057	0.217	0.486	0.239	13.3
R22	0.072	0.105	0.820	0.003	4.2
R24	0.068	0.190	0.730	0.012	3.5
R25	0.057	0.158	0.750	0.035	3.7
R26	0.056	0.123	0.733	0.088	11.2
R27	0.057	0.153	0.747	0.043	7.1
R28	0.080	0.077	0.567	0.276	15.6

APPENDIX G

LOGISTIC REGRESSION SAS CODES FOR ESTABLISHING BEST RADIUS OF WELL INFLUENCE

```
DM 'LOG; CLEAR; OUTPUT; CLEAR;';
DATA ONE;
      infile 'F:\stat\n1.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF1;
      SET ONE;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF1; RUN;

proc logistic data=buff1;
      model nitr_cat = devd frst gras plnt;
        run;
proc logistic data=buff1;
      model nitr_cat2=devd frst gras plnt;
      run;
DATA TWO;
      infile 'F:\stat\n2.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF2;
      SET TWO;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF2; RUN;

proc logistic data=buff2;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff2;
      model nitr_cat2=devd frst gras plnt;
   run;
DATA FIVE;
      infile 'F:\stat\n5.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF5;
```

```
SET FIVE;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF5; RUN;

proc logistic data=buff5;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff5;
      model nitr_cat2=devd frst gras plnt;
   run;
DATA SEVEN;
      infile 'F:\stat\n7.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF7;
SET SEVEN;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF7; RUN;

proc logistic data=buff7;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff7;
      model nitr_cat2=devd frst gras plnt;
   run;
DATA TEN;
      infile 'F:\stat\n10.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF10;
SET TEN;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF10; RUN;

proc logistic data=buff10;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff10;
      model nitr_cat2=devd frst gras plnt;
   run;
DATA FIFT;
      infile 'F:\stat\n15.prn';
      INPUT devd frst gras plnt nitrt;
```

```
proc print;
run;
DATA BUFF15;
SET FIFT;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF15; RUN;

proc logistic data=buff15;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff15;
      model nitr_cat2=devd frst gras plnt;
   run;
DATA SEVTN;
      infile 'F:\stat\n17.prn';
      INPUT devd frst gras plnt nitrt;
proc print;
run;
DATA BUFF17;
SET SEVTN;
IF nitrt >= 4.00 THEN NITR_CAT=0;
IF nitrt < 4.00 THEN NITR_CAT=1;
if nitrt >= 10.00 then nitr_cat2=0;
if nitrt < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=BUFF17; RUN;

proc logistic data=buff17;
      model nitr_cat = devd frst gras plnt;
   run;
proc logistic data=buff17;
      model nitr_cat2=devd frst gras plnt;
   run;
```

APPENDIX H

LOGISTIC REGRESSION OUTPUTS FOR ESTABLISHING BEST RADIUS OF WELL INFLUENCE

The SAS System

The LOGISTIC Procedure

Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	98
Number of Observations Used	98

Response Profile (at threshold 4 mg/l)

Ordered Value	NITR_CAT	Total Frequency
1	0	80
2	1	18

Probability modeled is NITR_CAT=0.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Response Profile (at threshold 10 mg/l)

Ordered Value	nitr_cat2	Total Frequency
1	0	20
2	1	78

Probability modeled is nitr_cat2=0.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

<u>Model Information</u>

Data Set	WORK.BUFF100m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	101.812
SC	98.061	114.736
-2 Log L	93.476	91.812

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.6643	4	0.7972
Score	1.5977	4	0.8092
Wald	1.5613	4	0.8157

The SAS System

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-810.1	993.1	0.6654	0.4146
devd	1	813.4	993.1	0.6709	0.4128
frst	1	810.9	992.9	0.6670	0.4141
gras	1	811.4	993.1	0.6675	0.4139
plnt	1	811.8	993.1	0.6681	0.4137

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	56.8	Somers' D	0.195
Percent Discordant	37.3	Gamma	0.207
Percent Tied	5.9	Tau-a	0.059
Pairs	1440	c	0.598

<u>Model Information</u>

Data Set	WORK.BUFF100m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	105.010
SC	103.763	117.935
-2 Log L	99.178	95.010

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.1679	4	0.3838
Score	4.3523	4	0.3604
Wald	4.0821	4	0.3950

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1358.0	815.9	2.7700	0.0960
devd	1	1354.7	815.6	2.7587	0.0967
frst	1	1355.5	815.2	2.7649	0.0964
gras	1	1356.5	815.9	2.7643	0.0964
plnt	1	1357.3	815.9	2.7675	0.0962

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	59.6	Somers' D	0.229
Percent Discordant	36.7	Gamma	0.238
Percent Tied	3.7	Tau-a	0.075
Pairs	1560	c	0.615

Model Information

Data Set	WORK.BUFF250m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	99.973
SC	98.061	112.898
-2 Log L	93.476	89.973

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.5026	4	0.4775
Score	3.2603	4	0.5152
Wald	3.0927	4	0.5424

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	262.8	736.9	0.1272	0.7213
devd	1	-257.8	736.8	0.1224	0.7264
frst	1	-260.3	736.8	0.1248	0.7239
gras	1	-262.0	736.9	0.1264	0.7222
plnt	1	-260.5	736.9	0.1250	0.7237

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.0	Somers' D	0.253
Percent Discordant	36.7	Gamma	0.256
Percent Tied	1.3	Tau-a	0.077
Pairs	1440	c	0.626

Model Information

Data Set	WORK.BUFF250m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	106.487
SC	103.763	119.412
-2 Log L	99.178	96.487

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.6910	4	0.6108
Score	2.8108	4	0.5900
Wald	2.7149	4	0.6066

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	133.6	638.2	0.0438	0.8342
devd	1	-135.8	638.1	0.0453	0.8314
frst	1	-133.2	638.2	0.0436	0.8346
gras	1	-135.4	638.2	0.0450	0.8320
plnt	1	-134.1	638.2	0.0442	0.8336

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	59.4	Somers' D	0.212
Percent Discordant	38.1	Gamma	0.218
Percent Tied	2.5	Tau-a	0.070
Pairs	1560	c	0.606

Model Information

Data Set	WORK.BUFF500m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	91.644
SC	98.061	104.569
-2 Log L	93.476	81.644

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.8317	4	0.0186
Score	10.9922	4	0.0267
Wald	9.5832	4	0.0481

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	990.9	534.1	3.4417	0.0636
devd	1	-972.1	533.3	3.3223	0.0683
frst	1	-991.4	534.0	3.4474	0.0634
gras	1	-990.7	534.2	3.4394	0.0637
plnt	1	-989.0	534.2	3.4272	0.0641

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	76.1	Somers' D	0.527
Percent Discordant	23.4	Gamma	0.530
Percent Tied	0.5	Tau-a	0.160
Pairs	1440	c	0.764

Model Information

Data Set	WORK.BUFF500m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	102.792
SC	103.763	115.717
-2 Log L	99.178	92.792

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.3858	4	0.1721
Score	6.3036	4	0.1776
Wald	5.8394	4	0.2115

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	700.4	494.7	2.0041	0.1569
devd	1	-698.6	495.1	1.9912	0.1582
frst	1	-700.8	494.7	2.0068	0.1566

gras	1	-702.8	494.9	2.0162	0.1556
plnt	1	-700.9	494.8	2.0068	0.1566

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.4	Somers' D	0.333
Percent Discordant	33.1	Gamma	0.335
Percent Tied	0.5	Tau-a	0.109
Pairs	1560	c	0.667

Model Information

Data Set	WORK.BUFF750m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	91.965
SC	98.061	104.889
-2 Log L	93.476	81.965

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.5114	4	0.0214
Score	10.4792	4	0.0331
Wald	9.0730	4	0.0593

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	232.8	495.1	0.2210	0.6383
devd	1	-202.8	495.1	0.1677	0.6821
frst	1	-234.3	494.7	0.2243	0.6358
gras	1	-233.0	495.1	0.2215	0.6379
plnt	1	-230.8	495.2	0.2173	0.6411

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.2	Somers' D	0.490
Percent Discordant	25.2	Gamma	0.493
Percent Tied	0.6	Tau-a	0.149
Pairs	1440	c	0.745

Model Information

Data Set	WORK.BUFF750m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	102.206
SC	103.763	115.131
-2 Log L	99.178	92.206

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.9715	4	0.1374
Score	7.1521	4	0.1281
Wald	6.5865	4	0.1594

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	368.2	427.3	0.7425	0.3888
devd	1	-364.2	428.2	0.7232	0.3951
frst	1	-369.5	426.8	0.7497	0.3866
gras	1	-370.6	427.3	0.7521	0.3858
plnt	1	-368.5	427.2	0.7438	0.3884

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999

frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	67.6	Somers' D	0.362
Percent Discordant	31.5	Gamma	0.365
Percent Tied	0.9	Tau-a	0.119
Pairs	1560	c	0.681

<u>Model Information</u>

Data Set	WORK.BUFF1000m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	89.979
SC	98.061	102.904
-2 Log L	93.476	79.979

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	13.4968	4	0.0091
Score	12.6469	4	0.0131
Wald	10.9660	4	0.0269

The SAS System

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-313.9	616.6	0.2591	0.6108
devd	1	361.3	619.6	0.3400	0.5598
frst	1	309.4	616.7	0.2517	0.6159
gras	1	312.9	616.5	0.2575	0.6118
plnt	1	314.8	616.6	0.2607	0.6096

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.8	Somers' D	0.501
Percent Discordant	24.7	Gamma	0.504
Percent Tied	0.6	Tau-a	0.152
Pairs	1440	c	0.751

Model Information

Data Set	WORK.BUFF1000m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	92.357
SC	103.763	105.282
-2 Log L	99.178	82.357

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.8203	4	0.0021
Score	16.7483	4	0.0022
Wald	13.3918	4	0.0095

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1769.8	653.7	7.3304	0.0068
devd	1	-1791.8	654.8	7.4893	0.0062
frst	1	-1763.9	653.8	7.2784	0.0070
gras	1	-1771.8	653.8	7.3442	0.0067
plnt	1	-1768.4	653.6	7.3193	0.0068

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	<0.001
frst	<0.001	<0.001	<0.001
gras	<0.001	<0.001	<0.001
plnt	<0.001	<0.001	<0.001

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.0	Somers' D	0.483
Percent Discordant	25.7	Gamma	0.484
Percent Tied	0.3	Tau-a	0.158
Pairs	1560	c	0.741

Model Information

Data Set	WORK.BUFF1500m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	91.662
SC	98.061	104.587
-2 Log L	93.476	81.662

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.8137	4	0.0188
Score	11.3051	4	0.0233
Wald	9.8936	4	0.0423

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-144.4	459.7	0.0986	0.7535
devd	1	205.3	452.9	0.2054	0.6504
frst	1	142.2	459.9	0.0956	0.7572
gras	1	142.5	460.2	0.0959	0.7568
plnt	1	144.4	460.1	0.0985	0.7536

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	73.1	Somers' D	0.467
Percent Discordant	26.3	Gamma	0.470
Percent Tied	0.6	Tau-a	0.142
Pairs	1440	c	0.734

Model Information

Data Set	WORK.BUFF1500m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	100.110
SC	103.763	113.035
-2 Log L	99.178	90.110

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.0673	4	0.0594
Score	9.5954	4	0.0478
Wald	8.5431	4	0.0736

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-42.4677	433.3	0.0096	0.9219
devd	1	43.3894	430.5	0.0102	0.9197
frst	1	42.2657	433.5	0.0095	0.9223
gras	1	39.8037	433.5	0.0084	0.9268
plnt	1	42.9088	433.5	0.0098	0.9212

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.0	Somers' D	0.486
Percent Discordant	25.4	Gamma	0.488
Percent Tied	0.5	Tau-a	0.159
Pairs	1560	c	0.743

Model Information

Data Set	WORK.BUFF1750m
Response Variable	NITR_CAT
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	92.174

SC	98.061	105.099
-2 Log L	93.476	82.174

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.3015	4	0.0234
Score	10.3499	4	0.0349
Wald	9.2724	4	0.0546

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-360.3	510.3	0.4987	0.4801
devd	1	434.0	514.5	0.7117	0.3989
frst	1	355.7	510.4	0.4857	0.4859
gras	1	357.8	510.1	0.4922	0.4830
plnt	1	359.4	510.1	0.4965	0.4811

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	>999.999	<0.001	>999.999
frst	>999.999	<0.001	>999.999
gras	>999.999	<0.001	>999.999
plnt	>999.999	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.9	Somers' D	0.443
Percent Discordant	27.6	Gamma	0.445
Percent Tied	0.4	Tau-a	0.134
Pairs	1440	c	0.722

<u>Model Information</u>

Data Set	WORK.BUFF1750m
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	100.095
SC	103.763	113.020
-2 Log L	99.178	90.095

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.0825	4	0.0591
Score	9.3364	4	0.0532
Wald	8.3128	4	0.0808

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	317.4	481.9	0.4338	0.5101
devd	1	-317.5	478.7	0.4398	0.5072
frst	1	-316.1	482.6	0.4291	0.5124
gras	1	-320.2	482.1	0.4413	0.5065
plnt	1	-317.0	482.1	0.4323	0.5109

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
devd	<0.001	<0.001	>999.999
frst	<0.001	<0.001	>999.999
gras	<0.001	<0.001	>999.999
plnt	<0.001	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	73.4	Somers' D	0.473
Percent Discordant	26.1	Gamma	0.476
Percent Tied	0.5	Tau-a	0.155
Pairs	1560	c	0.737

APPENDIX I

EXPLANATORY VARIABLES WITHIN A STATISTICALLY SIGNIFICANT RADIUS OF WELL INFLUENCE

Table I-1. Explanatory variables

dvlp	frst	grss	cult	sand	silt	clay	orga	nitro	Dwt*	well id
0.047	0.039	0.432	0.482	1.77	1.15	0.95	1.50	1.85	1.10	A1
0.061	0.057	0.311	0.571	1.77	1.15	0.95	1.50	1.93	0.96	A2
0.062	0.095	0.545	0.298	1.82	1.06	0.87	1.20	1.68	1.05	A3
0.070	0.101	0.561	0.268	1.88	0.84	0.64	0.83	1.65	1.07	A4
0.092	0.176	0.536	0.196	1.86	0.82	0.56	1.00	1.58	1.22	A5
0.068	0.095	0.438	0.399	1.88	0.91	0.72	0.83	1.79	1.47	A6
0.059	0.104	0.501	0.337	1.87	0.95	0.76	0.90	1.72	1.14	A7
0.064	0.189	0.608	0.139	1.84	1.04	0.84	1.13	1.43	1.61	A8
0.064	0.208	0.588	0.140	1.85	1.01	0.82	1.05	1.44	1.69	A9
0.055	0.053	0.400	0.492	1.78	1.14	0.94	1.46	1.86	1.03	A11
0.050	0.048	0.321	0.580	1.77	1.15	0.95	1.50	1.92	1.37	A14
0.062	0.006	0.213	0.719	1.77	1.15	0.95	1.50	2.02	1.25	A16
0.060	0.074	0.246	0.619	1.77	1.15	0.95	1.50	1.96	1.26	A17
0.046	0.078	0.395	0.482	1.77	1.15	0.95	1.50	1.85	1.11	A18
0.060	0.107	0.305	0.528	1.85	1.00	0.80	1.01	1.89	1.32	A19
0.087	0.172	0.594	0.147	1.89	0.89	0.69	0.75	1.49	1.44	A20
0.080	0.117	0.579	0.224	1.89	0.89	0.69	0.75	1.61	1.25	A21
0.079	0.114	0.641	0.166	1.89	0.89	0.69	0.75	1.51	1.43	A22
0.070	0.092	0.678	0.161	1.89	0.86	0.66	0.75	1.49	1.23	A23
0.074	0.097	0.750	0.079	1.89	0.82	0.62	0.75	1.31	1.29	A24
0.066	0.064	0.692	0.177	1.89	0.80	0.59	0.75	1.51	1.38	A25
0.074	0.107	0.791	0.028	1.89	0.83	0.63	0.75	1.13	1.30	A27
0.064	0.032	0.903	0.000	1.89	0.88	0.68	0.75	0.93	1.28	A29
0.023	0.066	0.888	0.023	1.87	0.67	0.48	0.05	0.79	1.14	CS1
0.030	0.051	0.907	0.012	1.86	0.79	0.56	0.12	0.75	1.00	CS2
0.041	0.042	0.778	0.138	1.87	0.67	0.48	0.05	1.38	1.20	CS3

Table I-1. Cont.

0.024	0.053	0.837	0.086	1.87	0.67	0.48	0.05	1.17	1.31	CS4
0.055	0.004	0.790	0.151	1.83	0.95	0.68	0.26	1.44	1.18	CS5
0.065	0.005	0.754	0.176	1.83	0.98	0.71	0.30	1.51	1.21	CS6
0.061	0	0.710	0.229	1.84	0.92	0.65	0.23	1.59	1.38	CS8
0.044	0.003	0.801	0.152	1.84	0.88	0.62	0.19	1.42	1.50	CS9
0.045	0.020	0.778	0.157	1.80	1.09	0.80	0.44	1.43	1.21	CS10
0.043	0.022	0.882	0.053	1.87	0.67	0.48	0.05	1.11	1.63	CS11
0.029	0.028	0.813	0.130	1.86	0.79	0.56	0.12	1.33	1.69	CS12
0.052	0.043	0.722	0.182	1.86	0.79	0.56	0.12	1.50	1.55	CS13
0.051	0.043	0.661	0.245	1.81	1.04	0.76	0.37	1.60	1.35	CS14
0.045	0.033	0.807	0.115	1.86	0.72	0.52	0.12	1.33	1.71	CS15
0.030	0.045	0.785	0.141	1.87	0.67	0.48	0.05	1.36	1.63	CS16
0.049	0.062	0.715	0.174	1.86	0.79	0.56	0.12	1.47	1.42	CS17
0.013	0.038	0.701	0.248	1.86	0.72	0.52	0.12	1.54	1.52	CS18
0.020	0.048	0.824	0.108	1.87	0.67	0.48	0.05	1.23	1.65	CS19
0.022	0.040	0.824	0.114	1.87	0.67	0.48	0.05	1.26	1.32	CS20
0.024	0.011	0.879	0.085	1.87	0.67	0.48	0.05	1.17	1.59	CS21
0.022	0.011	0.968	0.000	1.87	0.67	0.48	0.05	0.46	1.53	CS22
0.027	0.015	0.876	0.082	1.87	0.67	0.48	0.05	1.16	1.57	CS23
0.029	0.011	0.960	0.000	1.87	0.67	0.48	0.05	0.59	1.38	CS25
0.026	0.008	0.966	0.000	1.86	0.79	0.56	0.12	0.54	1.32	CS26
0.026	0.016	0.958	0.000	1.87	0.67	0.48	0.05	0.54	1.31	CS27
0.023	0.014	0.963	0.000	1.86	0.79	0.56	0.12	0.48	1.26	CS28
0.045	0.007	0.948	0.000	1.79	1.12	0.82	0.47	0.78	0.76	CS29
0.038	0.022	0.940	0.000	1.85	0.83	0.59	0.16	0.71	1.41	CS30
0.044	0.000	0.474	0.482	1.68	1.37	1.07	1.63	1.85	1.52	D1
0.055	0.016	0.272	0.657	1.77	1.15	0.95	1.50	1.98	1.71	D2
0.048	0.004	0.370	0.578	1.77	1.15	0.95	1.50	1.92	1.54	D3
0.063	0.007	0.406	0.524	1.77	1.15	0.95	1.50	1.89	1.39	D5
0.066	0.024	0.207	0.703	1.74	1.26	0.99	1.55	2.01	1.62	D6
0.057	0.000	0.237	0.706	1.52	1.56	1.11	1.78	2.01	1.41	D7
0.061	0.001	0.331	0.606	1.63	1.45	1.06	1.68	1.95	1.44	D8
0.044	0.000	0.593	0.363	1.46	1.60	1.13	1.83	1.73	1.42	D9
0.055	0.000	0.437	0.508	1.34	1.66	1.16	1.90	1.88	0.86	D10

Table I-1. Cont.

0.069	0.009	0.104	0.818	1.77	1.15	0.95	1.50	2.07	1.72	D12
0.052	0.000	0.405	0.543	1.72	1.29	1.03	1.58	1.90	1.27	D18
0.041	0.000	0.554	0.406	1.61	1.47	1.08	1.70	1.77	1.29	D19
0.045	0.003	0.370	0.582	1.77	1.15	0.95	1.50	1.92	1.10	D20
0.042	0.003	0.117	0.838	1.77	1.15	0.95	1.50	2.07	1.35	D21
0.042	0.012	0.184	0.762	1.77	1.15	0.95	1.50	2.03	1.22	D23
0.067	0.052	0.169	0.712	1.77	1.15	0.95	1.50	2.02	1.61	D25
0.052	0.014	0.107	0.827	1.77	1.15	0.95	1.50	2.07	1.64	D27
0.056	0.011	0.064	0.869	1.77	1.15	0.95	1.50	2.09	1.56	D28
0.062	0.008	0.050	0.880	1.77	1.15	0.95	1.50	2.10	1.74	D29
0.079	0.094	0.115	0.712	1.77	1.15	0.95	1.50	2.02	1.85	D33
0.027	0.100	0.874	0.000	1.89	0.89	0.69	0.75	0.55	1.61	R1
0.028	0.038	0.934	0.000	1.89	0.89	0.69	0.75	0.57	1.60	R2
0.043	0.053	0.873	0.031	1.89	0.89	0.69	0.75	0.99	1.56	R3
0.055	0.033	0.846	0.066	1.89	0.89	0.69	0.75	1.21	1.57	R4
0.025	0.085	0.890	0.000	1.89	0.86	0.66	0.75	0.52	1.54	R5
0.047	0.080	0.873	0.000	1.89	0.89	0.69	0.75	0.79	1.59	R6
0.065	0.145	0.790	0.000	1.89	0.89	0.69	0.75	0.94	1.67	R7
0.059	0.083	0.857	0.001	1.89	0.89	0.69	0.75	0.90	1.59	R8
0.047	0.054	0.760	0.138	1.89	0.89	0.69	0.75	1.39	1.56	R9
0.062	0.086	0.852	0.000	1.89	0.80	0.59	0.75	0.92	1.59	R10
0.066	0.154	0.780	0.000	1.89	0.86	0.66	0.75	0.94	1.63	R11
0.066	0.211	0.723	0.000	1.89	0.89	0.69	0.75	0.95	1.66	R12
0.070	0.200	0.730	0.000	1.89	0.89	0.69	0.75	0.97	1.58	R13
0.056	0.093	0.751	0.100	1.89	0.89	0.69	0.75	1.32	1.54	R14
0.066	0.041	0.893	0.000	1.88	0.71	0.50	0.75	0.95	1.57	R15
0.070	0.159	0.770	0.000	1.89	0.78	0.58	0.75	0.97	1.56	R16
0.085	0.214	0.701	0.000	1.89	0.83	0.63	0.75	1.06	1.64	R17
0.074	0.215	0.711	0.000	1.89	0.89	0.69	0.75	1.00	1.58	R18
0.064	0.230	0.703	0.003	1.89	0.89	0.69	0.75	0.95	1.49	R19
0.053	0.238	0.606	0.102	1.89	0.89	0.69	0.75	1.32	1.51	R20
0.059	0.186	0.449	0.305	1.89	0.89	0.69	0.75	1.69	1.53	R21
0.085	0.135	0.780	0.000	1.88	0.73	0.52	0.75	1.06	1.58	R22
0.074	0.184	0.742	0.000	1.89	0.83	0.63	0.75	0.99	1.56	R24

Table I-1. Cont.

0.031	0.052	0.917	0.000	1.88	0.65	0.43	0.75	0.62	1.79	R25
0.063	0.066	0.817	0.054	1.88	0.68	0.47	0.75	1.19	1.61	R26
0.039	0.027	0.917	0.017	1.88	0.63	0.41	0.75	0.87	1.62	R27
0.112	0.057	0.520	0.311	1.88	0.77	0.57	0.75	1.75	1.05	R28

* Well reading data

Notations
dvlp= developed land
frst= forested-wetland
grss=grassland
cult= cultivated land
sand=log(percent sand)
silt=log(percent silt)
clay=log(percent clay)
orga=percent organic matter by weight
nitro=log(nitrogen-fertilizer kg/sq. mile)
dwt=log(depth to water table ,ft)

APPENDIX J

SAS CODES FOR UNIVARIATE LOGISTIC AND STEPWISE LOGISTIC REGRESSIONS

For univariate logistic regression

```
DM 'LOG; CLEAR; OUTPUT; CLEAR;';
DATA ONE;
        infile 'E:\stepwise\univariate\dvlp.prn';
        INPUT dvlp nitra;
proc print;
run;
DATA developed;
        SET ONE;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=developed; RUN;

proc logistic data=developed;
    model nitr_cat1= dvlp/ lackfit rsq;
    run;
proc logistic data=developed;
    model nitr_cat2=dvlp/ lackfit rsq;
    run;
DATA TWO;
        infile 'E:\stepwise\univariate\frst.prn';
        INPUT frst nitra;
proc print;
run;
DATA forested;
        SET two;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=forested; RUN;
proc logistic data=forested;
    model nitr_cat1= frst/ lackfit rsq;
    run;
```

```
proc logistic data=forested;
     model nitr_cat2=frst/ lackfit rsq;
     run;
DATA three;
          infile 'E:\stepwise\univariate\grss.prn';
          INPUT grss nitra;
proc print;
run;
DATA grass;
          SET three;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=grass; RUN;

proc logistic data=grass;
     model nitr_cat1= grss/ lackfit rsq;
     run;
proc logistic data=grass;
     model nitr_cat2=grss/ lackfit rsq;
     run;
DATA four;
          infile 'E:\stepwise\univariate\cult.prn';
          INPUT cult nitra;
proc print;
run;
DATA cultivated;
          SET four;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=cultivated; RUN;

proc logistic data=cultivated;
     model nitr_cat1= cult/ lackfit rsq;
     run;
proc logistic data=cultivated;
     model nitr_cat2=cult/ lackfit rsq;
     run;
DATA five;
          infile 'E:\stepwise\univariate\sand.prn';
          INPUT sand nitra;
```

```
proc print;
run;
DATA sandp;
        SET five;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=sandp; RUN;

proc logistic data=sandp;
    model nitr_cat1= sand/ lackfit rsq;
    run;
proc logistic data=sandp;
    model nitr_cat2=sand/ lackfit rsq;
    run;
DATA six;
        infile 'E:\stepwise\univariate\silt.prn';
        INPUT silt nitra;
proc print;
run;
DATA siltp;
        SET six;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=siltp; RUN;

proc logistic data=siltp;
    model nitr_cat1= silt/ lackfit rsq;
    run;
proc logistic data=siltp;
    model nitr_cat2=silt/ lackfit rsq;
    run;
DATA seven;
        infile 'E:\stepwise\univariate\clay.prn';
        INPUT clay nitra;
proc print;
run;
DATA clayp;
        SET seven;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
```

```
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=clayp; RUN;

proc logistic data=clayp;
     model nitr_cat1= clay/ lackfit rsq;
     run;
proc logistic data=clayp;
     model nitr_cat2=clay/ lackfit rsq;
     run;
DATA eight;
          infile 'E:\stepwise\univariate\orga.prn';
          INPUT orga nitra;
proc print;
run;
DATA organic;
          SET eight;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=organic; RUN;

proc logistic data=organic;
     model nitr_cat1= orga/ lackfit rsq;
     run;
proc logistic data=organic;
     model nitr_cat2=orga/ lackfit rsq;
     run;
DATA nine;
          infile 'E:\stepwise\univariate\nitro.prn';
          INPUT nitro nitra;
proc print;
run;
DATA nitrogen;
          SET nine;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=nitrogen; RUN;
proc logistic data=nitrogen;
     model nitr_cat1= nitro/ lackfit rsq;
```

```
      run;
proc logistic data=nitrogen;
      model nitr_cat2=nitro/ lackfit rsq;
      run;
DATA ten;
            infile 'E:\stepwise\univariate\dwt.prn';
            INPUT dwt nitra;
proc print;
run;
DATA depthwt;
            SET ten;
IF nitra >= 4.00 THEN nitr_cat1=0;
IF nitra < 4.00 THEN nitr_cat1=1;
if nitra >= 10.00 then nitr_cat2=0;
if nitra < 10.00 then nitr_cat2=1;
RUN;
PROC PRINT DATA=depthwt; RUN;
proc logistic data=depthwt;
      model nitr_cat1= dwt/ lackfit rsq;
      run;
proc logistic data=depthwt;
      model nitr_cat2=dwt/ lackfit rsq;
      run;
```

<u>For Stepwise logistic regression</u>

For 4 mg/l

```
DM 'LOG; CLEAR; OUTPUT; CLEAR;';
DATA ONE;
            infile 'C:\THESIS_GIS\logistic\four.prn';
            INPUT dvlp grss cult sand silt clay orga nitro dwt nitra;
proc print;
run;
DATA BUFF10;
            SET ONE;
IF nitra >= 4.00 THEN nitra_cat1=1;
IF nitra < 4.00 THEN nitra_cat1=0;
RUN;
title 'Stepwise Regression on groundwater variables';
   proc logistic data=BUFF10 descending outest=betas covout;
      model nitra_cat1=dvlp grss cult sand silt clay orga nitro dwt
                  / selection=stepwise
                    slentry=0.3
                    slstay=0.15
```

```
                details
                lackfit;
    output out=pred p=phat lower=lcl upper=ucl
          predprobs=(individual crossvalidate);
  run;
  proc print data=betas;
    title2 'Parameter Estimates and Covariance Matrix';
  run;
  proc print data=pred;
    title2 'Predicted Probabilities and 95% Confidence Limits';
  run;
```

For 10 mg/l

```
DM 'LOG; CLEAR; OUTPUT; CLEAR;';
DATA ONE;
        infile 'C:\THESIS_GIS\logistic\tenmg.prn';
        INPUT grss cult silt clay orga nitro dwt nitra;
proc print;
run;
DATA BUFF10;
        SET ONE;
IF nitra >= 10.00 THEN nitra_cat1=1;
IF nitra < 10.00 THEN nitra_cat1=0;
RUN;
title 'Stepwise Regression on groundwater variables';
  proc logistic data=BUFF10 descending outest=betas covout;
    model nitra_cat1=grss cult silt clay orga nitro dwt
              / selection=stepwise
                slentry=0.3
                slstay=0.15
                details
                lackfit;
    output out=pred p=phat lower=lcl upper=ucl
          predprobs=(individual crossvalidate);
  run;
  proc print data=betas;
    title2 'Parameter Estimates and Covariance Matrix';
  run;
  proc print data=pred;
    title2 'Predicted Probabilities and 95% Confidence Limits';
  run;
```

APPENDIX K

UNIVARIATE LOGISTIC REGRESSION OUTPUTS

The SAS System
The LOGISTIC Procedure

Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	98
Number of Observations Used	98

Response Profile (at threshold 4 mg/l)

Ordered Value	nitr_cat1	Total Frequency
1	0	80
2	1	18

Probability modeled is nitr_cat1=0.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Response Profile (at threshold 10 mg/l)

Ordered Value	nitr_cat2	Total Frequency
1	0	20
2	1	78

Probability modeled is nitr_cat2=0.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

<u>Model Information</u>

Data Set	WORK.DEVELOPED
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	87.638
SC	98.061	92.808
-2 Log L	93.476	83.638

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.8380	1	0.0017
Score	9.2789	1	0.0023
Wald	8.5345	1	0.0035

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8367	0.7797	1.1516	0.2832
dvlp	1	48.1282	16.4744	8.5345	0.0035

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	>999.999	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	70.5	Somers' D	0.423
Percent Discordant	28.2	Gamma	0.429
Percent Tied	1.3	Tau-a	0.128
Pairs	1440	c	0.711

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat1 = 0 Observed	nitr_cat1 = 0 Expected	nitr_cat1 = 1 Observed	nitr_cat1 = 1 Expected
1	11	6	6.18	5	4.82
2	10	6	6.55	4	3.45
3	9	6	6.94	3	2.06
4	10	9	8.01	1	1.99
5	11	11	9.33	0	1.67
6	11	11	9.69	0	1.31
7	10	7	9.00	3	1.00
8	10	9	9.14	1	0.86
9	10	9	9.36	1	0.64
10	6	6	5.81	0	0.19

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
9.6367	8	0.2915

Model Information

Data Set	WORK.DEVELOPED
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	103.059
SC	103.763	108.229
-2 Log L	99.178	99.059

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.1185	1	0.7306
Score	0.1186	1	0.7305
Wald	0.1185	1	0.7307

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6116	0.7762	4.3106	0.0379
dvlp	1	4.6760	13.5858	0.1185	0.7307

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	107.343	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	47.2	Somers' D	-.013
Percent Discordant	48.5	Gamma	-.014
Percent Tied	4.3	Tau-a	-.004
Pairs	1560	c	0.493

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	11	3	2.00	8	9.00
2	10	0	1.87	10	8.13
3	9	1	1.76	8	7.24
4	10	2	1.99	8	8.01
5	11	3	2.24	8	8.76
6	11	6	2.29	5	8.71
7	10	2	2.11	8	7.89
8	10	1	2.14	9	7.86
9	10	0	2.20	10	7.80
10	6	2	1.40	4	4.60

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
15.1709	8	0.0559

Model Information

Data Set	WORK.FORESTED
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	97.350
SC	98.061	102.520
-2 Log L	93.476	93.350

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.1261	1	0.7225
Score	0.1233	1	0.7254
Wald	0.1230	1	0.7258

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.3967	0.3704	14.2177	0.0002
frst	1	1.4829	4.2280	0.1230	0.7258

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
frst	4.406	0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	51.5	Somers' D	0.097
Percent Discordant	41.9	Gamma	0.103
Percent Tied	6.6	Tau-a	0.029
Pairs	1440	c	0.548

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat1 = 0 Observed	nitr_cat1 = 0 Expected	nitr_cat1 = 1 Observed	nitr_cat1 = 1 Expected
1	11	9	8.82	2	2.18

2	9	7	7.23	2	1.77
3	10	6	8.05	4	1.95
4	11	11	8.89	0	2.11
5	10	7	8.12	3	1.88
6	10	9	8.14	1	1.86
7	10	9	8.20	1	1.80
8	10	8	8.24	2	1.76
9	10	8	8.37	2	1.63
10	7	6	5.93	1	1.07

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
7.2138	8	0.5137

Model Information

Data Set	WORK.FORESTED
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	103.177
SC	103.763	108.346
-2 Log L	99.178	99.177

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0012	1	0.9718
Score	0.0013	1	0.9718
Wald	0.0013	1	0.9717

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3701	0.3605	14.4468	0.0001
frst	1	0.1386	3.9085	0.0013	0.9717

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
frst	1.149	<0.001	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	30.4	Somers' D	0.046

Percent Discordant	25.8	Gamma	0.082
Percent Tied	43.7	Tau-a	0.015
Pairs	1560	c	0.523

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat2 = 0 Observed	nitr_cat2 = 0 Expected	nitr_cat2 = 1 Observed	nitr_cat2 = 1 Expected
1	11	2	2.23	9	8.77
2	9	1	1.82	8	7.18
3	10	2	2.03	8	7.97
4	11	1	2.24	10	8.76
5	10	3	2.04	7	7.96
6	10	3	2.04	7	7.96
7	10	4	2.04	6	7.96
8	10	1	2.05	9	7.95
9	10	2	2.06	8	7.94
10	7	1	1.45	6	5.55

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
5.7056	8	0.6802

Model Information

Data Set	WORK.GRASS
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	89.244
SC	98.061	94.414
-2 Log L	93.476	85.244

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.2318	1	0.0041
Score	6.9063	1	0.0086
Wald	5.9999	1	0.0143

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	4.0621	1.1650	12.1585	0.0005
grss	1	-3.6674	1.4972	5.9999	0.0143

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
grss	0.026	0.001	0.481

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.5	Somers' D	0.433
Percent Discordant	28.1	Gamma	0.435
Percent Tied	0.4	Tau-a	0.131
Pairs	1440	c	0.717

Partition for the Hosmer and Lemeshow Test

		nitr_cat1 = 0		nitr_cat1 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	10	4	6.43	6	3.57
2	11	8	7.63	3	3.37
3	10	9	7.36	1	2.64
4	10	9	7.69	1	2.31
5	10	9	7.99	1	2.01
6	10	9	8.37	1	1.63
7	10	6	8.87	4	1.13
8	10	9	9.28	1	0.72
9	10	10	9.56	0	0.44
10	7	7	6.83	0	0.17

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
14.8247	8	0.0626

<u>Model Information</u>

Data Set	WORK.GRASS
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	96.002
SC	103.763	101.172
-2 Log L	99.178	92.002

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.1756	1	0.0074
Score	7.5133	1	0.0061
Wald	6.9021	1	0.0086

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.1443	0.5907	0.0597	0.8070
grss	1	-2.5691	0.9779	6.9021	0.0086

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
grss	0.077	0.011	0.521

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.2	Somers' D	0.429
Percent Discordant	28.3	Gamma	0.431
Percent Tied	0.4	Tau-a	0.141
Pairs	1560	c	0.714

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	10	0	0.92	10	9.08
2	11	1	1.17	10	9.83
3	10	2	1.21	8	8.79
4	10	0	1.35	10	8.65
5	10	1	1.50	9	8.50
6	10	3	1.76	7	8.24
7	10	2	2.22	8	7.78
8	10	4	2.87	6	7.13
9	10	7	3.71	3	6.29
10	7	0	3.29	7	3.71

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
15.9293	8	0.0434

<u>Model Information</u>

Data Set	WORK.CULTIVATED
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	91.547
SC	98.061	96.717
-2 Log L	93.476	87.547

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.9288	1	0.0149
Score	4.9537	1	0.0260
Wald	4.3378	1	0.0373

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.9325	0.3318	7.9009	0.0049
cult	1	2.9886	1.4349	4.3378	0.0373

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
cult	19.858	1.193	330.642

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.0	Somers' D	0.346
Percent Discordant	28.4	Gamma	0.378
Percent Tied	8.6	Tau-a	0.105
Pairs	1440	c	0.673

Partition for the Hosmer and Lemeshow Test

		nitr_cat1 = 0		nitr_cat1 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	24	17	17.22	7	6.78
2	10	8	7.35	2	2.65
3	10	8	7.74	2	2.26
4	10	7	7.97	3	2.03
5	10	10	8.22	0	1.78
6	11	7	9.73	4	1.27
7	10	10	9.29	0	0.71
8	13	13	12.47	0	0.53

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
10.9551	6	0.0898

Model Information

Data Set WORK.CULTIVATED
Response Variable nitr_cat2
Number of Response Levels 2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	96.896
SC	103.763	102.066
-2 Log L	99.178	92.896

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.2813	1	0.0122
Score	6.7230	1	0.0095
Wald	6.2119	1	0.0127

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.0211	0.3994	25.6057	<.0001
cult	1	2.2536	0.9042	6.2119	0.0127

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
cult	9.522	1.618	56.030

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.3	Somers' D	0.451
Percent Discordant	27.2	Gamma	0.453
Percent Tied	0.4	Tau-a	0.148
Pairs	1560	c	0.725

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat2 = 0 Observed	nitr_cat2 = 0 Expected	nitr_cat2 = 1 Observed	nitr_cat2 = 1 Expected
1	24	0	2.81	24	21.19
2	10	1	1.24	9	8.76
3	10	4	1.42	6	8.58
4	10	0	1.56	10	8.44
5	10	3	1.73	7	8.27
6	11	4	2.61	7	8.39
7	10	4	3.15	6	6.85
8	13	4	5.48	9	7.52

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
13.6312	6	0.0340

Model Information

Data Set	WORK.SANDP
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	97.460
SC	98.061	102.630
-2 Log L	93.476	93.460

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0162	1	0.8988
Score	0.0165	1	0.8978
Wald	0.0165	1	0.8979

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.8446	5.0457	0.0280	0.8671
sand	1	0.3540	2.7580	0.0165	0.8979

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
sand	1.425	0.006	317.190

Association of Predicted Probabilities and Observed Responses

Percent Concordant	32.8	Somers' D	-.031
Percent Discordant	35.8	Gamma	-.045
Percent Tied	31.4	Tau-a	-.009
Pairs	1440	c	0.485

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat1 = 0 Observed	nitr_cat1 = 0 Expected	nitr_cat1 = 1 Observed	nitr_cat1 = 1 Expected
1	8	6	6.43	2	1.57
2	18	18	14.64	0	3.36
3	10	7	8.16	3	1.84
4	10	6	8.18	4	1.82
5	14	8	11.46	6	2.54
6	10	8	8.19	2	1.81
7	10	9	8.20	1	1.80
8	18	18	14.76	0	3.24

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
18.5198	6	0.0051

Model Information

Data Set	WORK.SANDP
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	102.401
SC	103.763	107.571
-2 Log L	99.178	98.401

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.7770	1	0.3781
Score	0.8594	1	0.3539
Wald	0.8247	1	0.3638

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	2.6453	4.4072	0.3603	0.5484
sand	1	-2.1969	2.4192	0.8247	0.3638

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
sand	0.111	<0.001	12.739

Association of Predicted Probabilities and Observed Responses

Percent Concordant	50.2	Somers' D	0.129
Percent Discordant	37.3	Gamma	0.147
Percent Tied	12.5	Tau-a	0.042
Pairs	1560	c	0.564

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	18	4	3.25	14	14.75
2	10	1	1.82	9	8.18
3	10	2	1.85	8	8.15
4	12	2	2.26	10	9.74
5	10	2	1.92	8	8.08
6	10	0	2.02	10	7.98
7	2	1	0.44	1	1.56
8	18	6	4.02	12	13.98
9	8	2	2.43	6	5.57

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
5.5320	7	0.5953

Model Information

Data Set	WORK.SILTP
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	96.042
SC	98.061	101.212
-2 Log L	93.476	92.042

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.4336	1	0.2312
Score	1.3279	1	0.2492
Wald	1.3024	1	0.2538

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.0919	1.2259	0.0056	0.9403
silt	1	1.5200	1.3319	1.3024	0.2538

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
silt	4.572	0.336	62.200

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.7	Somers' D	0.290
Percent Discordant	33.7	Gamma	0.301
Percent Tied	3.6	Tau-a	0.088
Pairs	1440	c	0.645

Partition for the Hosmer and Lemeshow Test

		nitr_cat1 = 0		nitr_cat1 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	14	7	10.53	7	3.47
2	13	10	10.09	3	2.91
3	12	11	9.54	1	2.46
4	2	2	1.61	0	0.39
5	18	18	14.56	0	3.44
6	10	6	8.31	4	1.69
7	3	2	2.57	1	0.43
8	18	18	15.53	0	2.47
9	8	6	7.26	2	0.74

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
20.5243	7	0.0045

Model Information

Data Set	WORK.SILTP
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	101.896
SC	103.763	107.066
-2 Log L	99.178	97.896

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.2820	1	0.2575
Score	1.3407	1	0.2469
Wald	1.3125	1	0.2519

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.5361	1.0741	5.5746	0.0182
silt	1	1.2218	1.0665	1.3125	0.2519

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
silt	3.393	0.420	27.442

Association of Predicted Probabilities and Observed Responses

Percent Concordant	53.3	Somers' D	0.162
Percent Discordant	37.1	Gamma	0.180
Percent Tied	9.7	Tau-a	0.053
Pairs	1560	c	0.581

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat2 = 0 Observed	nitr_cat2 = 0 Expected	nitr_cat2 = 1 Observed	nitr_cat2 = 1 Expected
1	14	2	2.13	12	11.87
2	13	3	2.17	10	10.83
3	12	1	2.15	11	9.85
4	2	0	0.38	2	1.62
5	18	4	3.42	14	14.58
6	10	1	2.09	9	7.91
7	3	1	0.71	2	2.29
8	18	6	4.39	12	13.61
9	8	2	2.57	6	5.43

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
3.5783	7	0.8269

<u>Model Information</u>

Data Set	WORK.CLAYP
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	94.956
SC	98.061	100.126
-2 Log L	93.476	90.956

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.5197	1	0.1124
Score	2.3894	1	0.1222
Wald	2.3103	1	0.1285

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1011	1.0442	0.0094	0.9229
clay	1	2.2895	1.5063	2.3103	0.1285

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	9.870	0.515	188.996

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.1	Somers' D	0.274
Percent Discordant	34.7	Gamma	0.284
Percent Tied	3.3	Tau-a	0.083
Pairs	1440	c	0.637

Partition for the Hosmer and Lemeshow Test

		nitr_cat1 = 0		nitr_cat1 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	15	8	10.88	7	4.12
2	10	7	7.56	3	2.44
3	11	10	8.60	1	2.40
4	7	7	5.63	0	1.37
5	18	18	14.66	0	3.34
6	10	5	8.45	5	1.55
7	19	19	16.89	0	2.11
8	8	6	7.31	2	0.69

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
24.0203	6	0.0005

Model Information

Data Set	WORK.CLAYP
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	101.301
SC	103.763	106.471
-2 Log L	99.178	97.301

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.8770	1	0.1707
Score	1.9106	1	0.1669
Wald	1.8713	1	0.1713

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.6662	1.0106	6.9601	0.0083
clay	1	1.7615	1.2877	1.8713	0.1713

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	5.821	0.467	72.626

Association of Predicted Probabilities and Observed Responses

Percent Concordant	53.7	Somers' D	0.171
Percent Discordant	36.6	Gamma	0.189
Percent Tied	9.7	Tau-a	0.056
Pairs	1560	c	0.585

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	15	3	2.06	12	12.94
2	10	1	1.52	9	8.48
3	11	2	1.84	9	9.16
4	7	0	1.27	7	5.73
5	18	4	3.42	14	14.58
6	10	1	2.18	9	7.82
7	19	7	5.16	12	13.84
8	8	2	2.54	6	5.46

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
4.2969	6	0.6366

Model Information

Data Set	WORK.ORGANIC
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	92.672
SC	98.061	97.842
-2 Log L	93.476	88.672

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.8035	1	0.0284
Score	4.5807	1	0.0323
Wald	4.3210	1	0.0376

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.6872	0.4291	2.5646	0.1093
orga	1	1.1213	0.5394	4.3210	0.0376

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
orga	3.069	1.066	8.833

Association of Predicted Probabilities and Observed Responses

Percent Concordant	61.0	Somers' D	0.303
Percent Discordant	30.8	Gamma	0.330
Percent Tied	8.2	Tau-a	0.092
Pairs	1440	c	0.651

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat1 = 0 Observed	nitr_cat1 = 0 Expected	nitr_cat1 = 1 Observed	nitr_cat1 = 1 Expected
1	12	6	8.13	6	3.87
2	10	7	6.97	3	3.03
3	6	5	4.47	1	1.53
4	35	33	28.76	2	6.24
5	9	5	7.77	4	1.23
6	18	18	16.46	0	1.54
7	8	6	7.44	2	0.56

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
18.3928	5	0.0025

Model Information

Data Set	WORK.ORGANIC
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	99.719
SC	103.763	104.889
-2 Log L	99.178	95.719

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.4586	1	0.0629
Score	3.4449	1	0.0634
Wald	3.3223	1	0.0683

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.1379	0.5268	16.4712	<.0001
orga	1	0.8695	0.4771	3.3223	0.0683

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
orga	2.386	0.937	6.077

Association of Predicted Probabilities and Observed Responses

Percent Concordant	52.1	Somers' D	0.231
Percent Discordant	29.0	Gamma	0.285
Percent Tied	18.9	Tau-a	0.076
Pairs	1560	c	0.616

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	12	2	1.32	10	10.68
2	10	1	1.17	9	8.83
3	6	0	0.82	6	5.18
4	35	7	6.46	28	28.54
5	9	2	2.05	7	6.95
6	18	6	5.45	12	12.55
7	8	2	2.73	6	5.27

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
1.8164	5	0.8739

<u>Model Information</u>

Data Set	WORK.NITROGEN
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	89.186
SC	98.061	94.356
-2 Log L	93.476	85.186

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.2902	1	0.0040
Score	8.1877	1	0.0042
Wald	7.3990	1	0.0065

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.6167	0.7627	0.6538	0.4188
nitro	1	1.6550	0.6084	7.3990	0.0065

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
nitro	5.233	1.588	17.244

Association of Predicted Probabilities and Observed Responses

Percent Concordant	70.7	Somers' D	0.419
Percent Discordant	28.8	Gamma	0.421
Percent Tied	0.5	Tau-a	0.127
Pairs	1440	c	0.709

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat1 = 0 Observed	nitr_cat1 = 0 Expected	nitr_cat1 = 1 Observed	nitr_cat1 = 1 Expected
1	10	5	5.76	5	4.24
2	10	7	6.92	3	3.08
3	10	9	7.34	1	2.66
4	10	8	7.95	2	2.05
5	10	8	8.39	2	1.61
6	10	9	8.64	1	1.36
7	10	8	8.94	2	1.06
8	10	8	9.20	2	0.80
9	11	11	10.26	0	0.74
10	7	7	6.60	0	0.40

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
5.9827	8	0.6492

Model Information

Data Set	WORK.NITROGEN
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	92.049
SC	103.763	97.219
-2 Log L	99.178	88.049

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.1285	1	0.0009
Score	10.0520	1	0.0015
Wald	8.8237	1	0.0030

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.4430	1.1446	15.0680	0.0001
nitro	1	2.0390	0.6864	8.8237	0.0030

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
nitro	7.683	2.001	29.498

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.1	Somers' D	0.448
Percent Discordant	27.3	Gamma	0.451
Percent Tied	0.6	Tau-a	0.147
Pairs	1560	c	0.724

Partition for the Hosmer and Lemeshow Test

Group	Total	nitr_cat2 = 0 Observed	nitr_cat2 = 0 Expected	nitr_cat2 = 1 Observed	nitr_cat2 = 1 Expected
1	10	0	0.36	10	9.64
2	10	0	0.64	10	9.36
3	10	0	0.81	10	9.19
4	10	3	1.19	7	8.81
5	10	2	1.62	8	8.38
6	10	2	1.97	8	8.03
7	10	3	2.58	7	7.42
8	10	4	3.40	6	6.60
9	11	5	4.33	6	6.67
10	7	1	3.09	6	3.91

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
8.1444	8	0.4195

<u>Model Information</u>

Data Set	WORK.DEPTHWT
Response Variable	nitr_cat1
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	96.307
SC	98.061	101.477
-2 Log L	93.476	92.307

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.1694	1	0.2795
Score	1.2048	1	0.2724
Wald	1.1855	1	0.2762

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.2742	1.6217	0.0286	0.8657
dwt	1	1.2559	1.1535	1.1855	0.2762

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dwt	3.511	0.366	33.674

Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.7	Somers' D	0.128
Percent Discordant	42.9	Gamma	0.130
Percent Tied	1.4	Tau-a	0.039
Pairs	1440	c	0.564

Partition for the Hosmer and Lemeshow Test

		nitr_cat1 = 0		nitr_cat1 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	10	7	7.26	3	2.74
2	10	8	7.71	2	2.29
3	10	8	7.90	2	2.10
4	10	8	8.05	2	1.95
5	10	8	8.21	2	1.79
6	10	9	8.38	1	1.62
7	10	8	8.45	2	1.55
8	10	9	8.50	1	1.50
9	10	9	8.56	1	1.44
10	8	6	6.97	2	1.03

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
1.9473	8	0.9826

<u>Model Information</u>

Data Set	WORK.DEPTHWT
Response Variable	nitr_cat2
Number of Response Levels	2

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	99.587
SC	103.763	104.757
-2 Log L	99.178	95.587

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.5906	1	0.0581
Score	3.7227	1	0.0537
Wald	3.5465	1	0.0597

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.5953	1.5601	1.0456	0.3065
dwt	1	-2.1178	1.1246	3.5465	0.0597

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dwt	0.120	0.013	1.090

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.9	Somers' D	0.348
Percent Discordant	32.1	Gamma	0.351
Percent Tied	1.0	Tau-a	0.114
Pairs	1560	c	0.674

Partition for the Hosmer and Lemeshow Test

		nitr_cat2 = 0		nitr_cat2 = 1	
Group	Total	Observed	Expected	Observed	Expected
1	10	0	1.14	10	8.86
2	11	0	1.50	11	9.50
3	10	2	1.46	8	8.54
4	10	1	1.54	9	8.46
5	11	4	1.88	7	9.12
6	10	3	2.02	7	7.98
7	10	2	2.33	8	7.67
8	10	4	2.61	6	7.39
9	10	3	3.11	7	6.89
10	6	1	2.41	5	3.59

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
9.3642	8	0.3125

APPENDIX L

STEPWISE LOGISTIC REGRESSION OUTPUTS

At threshold 4 mg/l

Stepwise Regression on groundwater variables

Predicted Probabilities and 95% Confidence Limits

The LOGISTIC Procedure

Model Information

Data Set	WORK.BUFF10
Response Variable	nitra_cat1
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	98
Number of Observations Used	98

Response Profile

Ordered Value	nitra_ cat1	Total Frequency
1	1	80
2	0	18

Probability modeled is nitra_cat1=1.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 93.476

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.4917	0.2609	32.6944	<.0001

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
20.2275	9	0.0166

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
dvlp	1	9.2789	0.0023
grss	1	6.9063	0.0086
cult	1	4.9537	0.0260
sand	1	0.0041	0.9492
silt	1	1.3805	0.2400
clay	1	2.2492	0.1337
orga	1	4.5593	0.0327
nitro	1	8.1920	0.0042
dwt	1	1.2671	0.2603

Step 1. Effect dvlp entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	87.638
SC	98.061	92.808
-2 Log L	93.476	83.638

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.8380	1	0.0017
Score	9.2789	1	0.0023
Wald	8.5345	1	0.0035

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8367	0.7797	1.1516	0.2832
dvlp	1	48.1282	16.4744	8.5345	0.0035

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	>999.999	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant	70.5	Somers' D	0.423
Percent Discordant	28.2	Gamma	0.429
Percent Tied	1.3	Tau-a	0.128
Pairs	1440	c	0.711

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
11.0916	8	0.1966

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	8.5345	0.0035

The LOGISTIC Procedure

NOTE: No effects for the model in Step 1 are removed.

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	2.2182	0.1364
cult	1	2.7135	0.0995
sand	1	0.0967	0.7559
silt	1	0.0163	0.8984
clay	1	0.1227	0.7261
orga	1	0.4199	0.5170
nitro	1	2.7968	0.0944
dwt	1	1.5901	0.2073

Step 2. Effect nitro entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	86.807
SC	98.061	94.562
-2 Log L	93.476	80.807

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12.6690	2	0.0018
Score	12.9227	2	0.0016
Wald	11.2295	2	0.0036

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5726	0.8938	3.0959	0.0785
dvlp	1	34.4587	17.3519	3.9437	0.0470
nitro	1	1.0799	0.6578	2.6951	0.1007

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	1.568	>999.999
nitro	2.944	0.811	10.688

Association of Predicted Probabilities and Observed Responses

Percent Concordant	73.4	Somers' D	0.469
Percent Discordant	26.5	Gamma	0.469
Percent Tied	0.1	Tau-a	0.142
Pairs	1440	c	0.734

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
9.9518	7	0.1913

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	3.9437	0.0470
nitro	1	2.6951	0.1007

NOTE: No effects for the model in Step 2 are removed.

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	0.0395	0.8425
cult	1	0.2879	0.5916
sand	1	2.9318	0.0869
silt	1	1.1312	0.2875
clay	1	0.7684	0.3807
orga	1	0.0670	0.7958
dwt	1	2.4443	0.1180

Step 3. Effect sand entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	86.616
SC	98.061	96.956
-2 Log L	93.476	78.616

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	14.8598	3	0.0019
Score	14.9591	3	0.0019
Wald	12.6252	3	0.0055

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-11.3673	6.3295	3.2254	0.0725
dvlp	1	30.1298	18.2660	2.7209	0.0990
sand	1	5.0284	3.1947	2.4773	0.1155
nitro	1	1.7176	0.8019	4.5879	0.0322

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	0.003	>999.999
sand	152.685	0.291	>999.999
nitro	5.571	1.157	26.824

Association of Predicted Probabilities and Observed Responses

Percent Concordant	75.3	Somers' D	0.510
Percent Discordant	24.3	Gamma	0.512
Percent Tied	0.4	Tau-a	0.154
Pairs	1440	c	0.755

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
5.7336	6	0.4537

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	2.7209	0.0990
sand	1	2.4773	0.1155
nitro	1	4.5879	0.0322

NOTE: No effects for the model in Step 3 are removed.

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	0.1929	0.6605
cult	1	1.1211	0.2897
silt	1	0.1347	0.7136
clay	1	0.0347	0.8522
orga	1	0.5013	0.4789
dwt	1	1.2942	0.2553

Step 4. Effect dwt entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	87.337
SC	98.061	100.262
-2 Log L	93.476	77.337

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.1389	4	0.0028
Score	16.5692	4	0.0023
Wald	13.6040	4	0.0087

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-11.1443	6.4185	3.0147	0.0825
dvlp	1	29.9108	18.5879	2.5894	0.1076
sand	1	3.7227	3.4301	1.1779	0.2778
nitro	1	1.6976	0.8071	4.4235	0.0354
dwt	1	1.5535	1.3844	1.2591	0.2618

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	0.001	>999.999
sand	41.375	0.050	>999.999
nitro	5.461	1.123	26.562
dwt	4.728	0.314	71.301

Association of Predicted Probabilities and Observed Responses

Percent Concordant	75.2	Somers' D	0.506
Percent Discordant	24.6	Gamma	0.507
Percent Tied	0.2	Tau-a	0.153
Pairs	1440	c	0.753

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
4.8675	5	0.4323

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	2.5894	0.1076
sand	1	1.1779	0.2778
nitro	1	4.4235	0.0354
dwt	1	1.2591	0.2618

The LOGISTIC Procedure

Step 5. Effect sand is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	86.435
SC	98.061	96.775
-2 Log L	93.476	78.435

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0412	3	0.0018
Score	15.2811	3	0.0016
Wald	12.7432	3	0.0052

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5924	2.2181	4.2868	0.0384
dvlp	1	33.2097	17.8786	3.4503	0.0632
nitro	1	1.2746	0.6883	3.4290	0.0641
dwt	1	2.0057	1.3181	2.3154	0.1281

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	0.160	>999.999
nitro	3.577	0.928	13.788
dwt	7.431	0.561	98.407

Stepwise Regression on groundwater variables
The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.9	Somers' D	0.442
Percent Discordant	27.7	Gamma	0.444
Percent Tied	0.4	Tau-a	0.134
Pairs	1440	c	0.721

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
6.5998	6	0.3594

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	3.4503	0.0632
nitro	1	3.4290	0.0641
dwt	1	2.3154	0.1281

NOTE: No effects for the model in Step 5 are removed.

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	0.0306	0.8612
cult	1	0.5368	0.4638
sand	1	1.2848	0.2570
silt	1	0.3624	0.5472
clay	1	0.2212	0.6381
orga	1	0.0640	0.8003

Step 6. Effect sand entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	87.337
SC	98.061	100.262
-2 Log L	93.476	77.337

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.1389	4	0.0028
Score	16.5692	4	0.0023
Wald	13.6040	4	0.0087

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-11.1443	6.4185	3.0147	0.0825
dvlp	1	29.9108	18.5879	2.5894	0.1076
sand	1	3.7227	3.4301	1.1779	0.2778
nitro	1	1.6976	0.8071	4.4235	0.0354
dwt	1	1.5535	1.3844	1.2591	0.2618

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	0.001	>999.999
sand	41.375	0.050	>999.999
nitro	5.461	1.123	26.562
dwt	4.728	0.314	71.301

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

Percent Concordant	75.2	Somers' D	0.506
Percent Discordant	24.6	Gamma	0.507
Percent Tied	0.2	Tau-a	0.153
Pairs	1440	c	0.753

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
4.8675	5	0.4323

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	2.5894	0.1076
sand	1	1.1779	0.2778
nitro	1	4.4235	0.0354
dwt	1	1.2591	0.2618

Step 7. Effect sand is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	95.476	86.435
SC	98.061	96.775
-2 Log L	93.476	78.435

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0412	3	0.0018

Score	15.2811	3	0.0016
Wald	12.7432	3	0.0052

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5924	2.2181	4.2868	0.0384
dvlp	1	33.2097	17.8786	3.4503	0.0632
nitro	1	1.2746	0.6883	3.4290	0.0641
dwt	1	2.0057	1.3181	2.3154	0.1281

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
dvlp	>999.999	0.160	>999.999
nitro	3.577	0.928	13.788
dwt	7.431	0.561	98.407

Association of Predicted Probabilities and Observed Responses

Percent Concordant	71.9	Somers' D	0.442
Percent Discordant	27.7	Gamma	0.444
Percent Tied	0.4	Tau-a	0.134
Pairs	1440	c	0.721

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
6.5998	6	0.3594

The LOGISTIC Procedure

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
dvlp	1	3.4503	0.0632
nitro	1	3.4290	0.0641
dwt	1	2.3154	0.1281

NOTE: No effects for the model in Step 7 are removed.

NOTE: Model building terminates because the last effect entered is removed by the Wald statistic criterion.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	dvlp		1	1	9.2789		0.0023
2	nitro		1	2	2.7968		0.0944
3	sand		1	3	2.9318		0.0869
4	dwt		1	4	1.2942		0.2553
5		sand	1	3		1.1779	0.2778
6	sand		1	4	1.2848		0.2570
7		sand	1	3		1.1779	0.2778

Partition for the Hosmer and Lemeshow Test

Group	Total	nitra_cat1 = 1 Observed	nitra_cat1 = 1 Expected	nitra_cat1 = 0 Observed	nitra_cat1 = 0 Expected
1	10	3	4.19	7	5.81
2	10	6	6.60	4	3.40
3	10	10	7.86	0	2.14
4	10	9	8.39	1	1.61
5	10	10	8.66	0	1.34
6	10	9	8.85	1	1.15
7	10	8	9.00	2	1.00
8	10	9	9.24	1	0.76
9	10	8	9.46	2	0.54
10	8	8	7.76	0	0.24

The LOGISTIC Procedure

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
10.8835	8	0.2084

At threshold 10 mg/l

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Model Information

Data Set	WORK.BUFF10
Response Variable	nitra_cat1
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	98
Number of Observations Used	98

Response Profile

Ordered Value	nitra_ cat1	Total Frequency
1	1	20
2	0	78

Probability modeled is nitra_cat1=1.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 99.178

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3610	0.2506	29.4849	<.0001

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
16.7627	7	0.0190

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	7.5133	0.0061
cult	1	6.7230	0.0095
silt	1	1.3581	0.2439
clay	1	1.8843	0.1698
orga	1	3.4292	0.0641
nitro	1	10.0581	0.0015
dwt	1	3.7178	0.0538

Step 1. Effect nitro entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	92.042
SC	103.763	97.212
-2 Log L	99.178	88.042

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.1355	1	0.0008
Score	10.0581	1	0.0015
Wald	8.8283	1	0.0030

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.4443	1.1448	15.0719	0.0001
nitro	1	2.0397	0.6865	8.8283	0.0030

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
nitro	7.688	2.002	29.524

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.2	Somers' D	0.449
Percent Discordant	27.3	Gamma	0.451
Percent Tied	0.4	Tau-a	0.147
Pairs	1560	c	0.725

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
8.8470	6	0.1824

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
nitro	1	8.8283	0.0030

NOTE: No effects for the model in Step 1 are removed.

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	1.1532	0.2829
cult	1	2.6137	0.1059
silt	1	3.0269	0.0819
clay	1	3.9631	0.0465
orga	1	1.0222	0.3120
dwt	1	2.1181	0.1456

Step 2. Effect clay entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	89.901
SC	103.763	97.656
-2 Log L	99.178	83.901

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.2764	2	0.0005
Score	11.9476	2	0.0025
Wald	9.4065	2	0.0091

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.0431	1.3105	9.5181	0.0020
clay	1	-4.5460	2.3597	3.7114	0.0540
nitro	1	3.9650	1.3411	8.7409	0.0031

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	0.011	<0.001	1.082
nitro	52.719	3.806	730.324

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.8	Somers' D	0.498
Percent Discordant	25.0	Gamma	0.499
Percent Tied	0.2	Tau-a	0.163
Pairs	1560	c	0.749

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
4.5508	5	0.4731

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
clay	1	3.7114	0.0540
nitro	1	8.7409	0.0031

NOTE: No effects for the model in Step 2 are removed.

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	0.2843	0.5939
cult	1	1.0958	0.2952
silt	1	0.0993	0.7527
orga	1	1.3738	0.2412
dwt	1	2.6942	0.1007

Step 3. Effect dwt entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	89.227
SC	103.763	99.566
-2 Log L	99.178	81.227

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.9512	3	0.0005
Score	14.4566	3	0.0023
Wald	10.8908	3	0.0123

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9949	2.2953	0.1879	0.6647
clay	1	-4.8939	2.4022	4.1505	0.0416
nitro	1	3.9986	1.3806	8.3889	0.0038
dwt	1	-2.0419	1.2656	2.6032	0.1067

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	0.007	<0.001	0.831
nitro	54.522	3.643	816.018
dwt	0.130	0.011	1.550

Association of Predicted Probabilities and Observed Responses

Percent Concordant	78.1	Somers' D	0.564
Percent Discordant	21.7	Gamma	0.565
Percent Tied	0.1	Tau-a	0.185
Pairs	1560	c	0.782

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
1.9128	4	0.7518

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
clay	1	4.1505	0.0416
nitro	1	8.3889	0.0038
dwt	1	2.6032	0.1067

NOTE: No effects for the model in Step 3 are removed.

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Effects Eligible for Entry

Effect	DF	Score Chi-Square	Pr > ChiSq
grss	1	0.0117	0.9137
cult	1	0.1516	0.6970
silt	1	0.0123	0.9118
orga	1	1.4604	0.2269

Step 4. Effect orga entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	89.774
SC	103.763	102.698
-2 Log L	99.178	79.774

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	19.4042	4	0.0007
Score	16.6115	4	0.0023
Wald	12.2964	4	0.0153

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.8448	2.7612	0.0936	0.7596
clay	1	-8.7098	4.0885	4.5384	0.0331
orga	1	1.5338	1.2831	1.4290	0.2319
nitro	1	3.8243	1.3837	7.6384	0.0057
dwt	1	-2.1047	1.2838	2.6878	0.1011

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	<0.001	<0.001	0.498
orga	4.636	0.375	57.322
nitro	45.801	3.041	689.791
dwt	0.122	0.010	1.509

Association of Predicted Probabilities and Observed Responses

Percent Concordant	80.0	Somers' D	0.602
Percent Discordant	19.8	Gamma	0.603
Percent Tied	0.2	Tau-a	0.198
Pairs	1560	c	0.801

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
0.4610	3	0.9274

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
clay	1	4.5384	0.0331
orga	1	1.4290	0.2319
nitro	1	7.6384	0.0057
dwt	1	2.6878	0.1011

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Step 5. Effect orga is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	101.178	89.227
SC	103.763	99.566
-2 Log L	99.178	81.227

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.9512	3	0.0005
Score	14.4566	3	0.0023
Wald	10.8908	3	0.0123

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9949	2.2953	0.1879	0.6647
clay	1	-4.8939	2.4022	4.1505	0.0416
nitro	1	3.9986	1.3806	8.3889	0.0038
dwt	1	-2.0419	1.2656	2.6032	0.1067

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
clay	0.007	<0.001	0.831
nitro	54.522	3.643	816.018
dwt	0.130	0.011	1.550

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

Percent Concordant	78.1	Somers' D	0.564
Percent Discordant	21.7	Gamma	0.565
Percent Tied	0.1	Tau-a	0.185
Pairs	1560	c	0.782

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
1.9128	4	0.7518

Analysis of Effects Eligible for Removal

Effect	DF	Wald Chi-Square	Pr > ChiSq
clay	1	4.1505	0.0416
nitro	1	8.3889	0.0038
dwt	1	2.6032	0.1067

NOTE: No effects for the model in Step 5 are removed.

NOTE: Model building terminates because the last effect entered is removed by the Wald statistic criterion.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	nitro		1	1	10.0581		0.0015
2	clay		1	2	3.9631		0.0465
3	dwt		1	3	2.6942		0.1007
4	orga		1	4	1.4604		0.2269
5		orga	1	3		1.4290	0.2319

Stepwise Regression on groundwater variables

The LOGISTIC Procedure

Partition for the Hosmer and Lemeshow Test

		nitra_cat1 = 1		nitra_cat1 = 0	
Group	Total	Observed	Expected	Observed	Expected
1	10	0	0.11	10	9.89
2	10	0	0.24	10	9.76
3	10	0	0.48	10	9.52
4	10	2	0.93	8	9.07
5	10	2	1.51	8	8.49
6	10	2	2.22	8	7.78
7	10	1	2.79	9	7.21
8	10	4	3.31	6	6.69
9	10	5	4.12	5	5.88
10	8	4	4.30	4	3.70

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
4.6003	8	0.7993

VITA

MOTI LAL K.C.

Candidate for the Degree of

Master of Science

Thesis: GIS-BASED STATISTICAL, GEOSTATISTICAL, AND STOCHASTIC ANALYSES OF NITRATE CONTAMINATION IN THE CIMARRON TERRACE AQUIFER IN OKLAHOMA

Major Field: Civil Engineering

Biographical:

Personal Data: Born in Nepal, on Oct. 30, 1976, the son of Jay and Radhika Khatri

Education: Graduated from Amrit Science Campus in July1995, Kathmandu Nepal; received Certificate in Geographic Information System at Oklahoma State University in May 2007; received Bachelor's Degree in Civil Engineering from Pulchowk Campus, Institute of Engineering, Tribhuvan University, Lalitpur, Nepal in October 2002; completed the requirements for the Master of Science Degree in Civil Engineering (Emphasis in Environmental and Water Resources) at Oklahoma State University, Stillwater, Oklahoma in December 2007.

Experience: Employed with Thapathali Campus, Institute of Engineering, Tribhuvan University, Kathmandu, Nepal from June 2003 to May 2005 as an assistant lecturer and from May 2005 to July 2005 as a lecturer; employed by Oklahoma State University, Oklahoma Infrastructure Consortium, Department of Civil and Environmental Engineering as a graduate research assistant from August 2005 to September 2007.

Memberships: Member- Nepal Engineers' Association; Member-Nepal Engineering Council; Charter Member- Rotary Club of Kathmandu Sunrise, Kathmandu, Nepal.

Name: Moti Lal K.C. Date of Degree: December 2007

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: GIS-BASED STATISTICAL, GEOSTATISTICAL, AND STOCHASTIC ANALYSES OF NITRATE CONTAMINATION IN THE CIMARRON TERRACE AQUIFER IN OKLAHOMA

Pages in Study: 263 Candidate for the Degree of Master of Science

Major Field: Civil Engineering (Emphasis in Environmental and Water Resources)

Summary:

Geostatistics and GIS were used to identify areas affected by nitrate contamination. Ordinary kriging estimates of groundwater quality monitoring data were integrated into GIS to provide a quantitative, statistical, and weighted means of nitrates defining the statistical significance of geographic apparent change between 1997 and 2005. The central portion of the aquifer was found to be mostly affected by nitrates. The tendency of probability to remain geographically stationary or to change with time was also evaluated as chronic exceedance frequencies of nitrates by using indicator kriging and GIS. Chronic exceedance frequencies of nitrates exceeding 0.70 and 0.95 probability levels of detecting equal to or greater than 4 mg/l and 10 mg/l of NO_3-N were mapped. At threshold 10 mg/l of NO_3-N, nitrates were frequently observed in the central portion of the aquifer. From logistic regressions, the best radius of well influence so that the land use could show significant effects on nitrate concentration in wells was determined to be 1,000 meters.

Stepwise logistic regression approach revealed that the percentages of Developed land or that can be termed as high-density residential land, Fertilizer N, Depth to water table were significant predictors of groundwater nitrate concentration in excess of 4 mg/l while percent of Clay, Fertilizer N, Depth to water table were significant predictors in excess of 10 mg/l. Goodness-of-fit tests indicated that models fitted the data well, and predicted and observed probabilities of nitrate exceeding 4 mg/l and 10 mg/l were strongly correlated (R^2 = 0.72 and 0.79 respectively). The probability curves of Monte Carlo simulation indicated that mean probabilities of occurrence of nitrate in excess of 4 mg/l and 10 mg/l were 0.8214 and 0.2581 respectively. In addition to Monte Carlo simulation, sensitivity analysis method was used to identify important variable, whose uncertainty was a driving factor in the overall uncertainty of risk estimates for nitrate occurrence at given thresholds. Fertilizer N was found to be the most sensitive variable with 61.1% and 81.7% positive impacts for nitrate contribution in the groundwater at thresholds 4 mg/l and 10 mg/l of nitrate as nitrogen respectively.

ADVISER'S APPROVAL: Dr. William F. McTernan and Dr. Avdhesh K. Tyagi