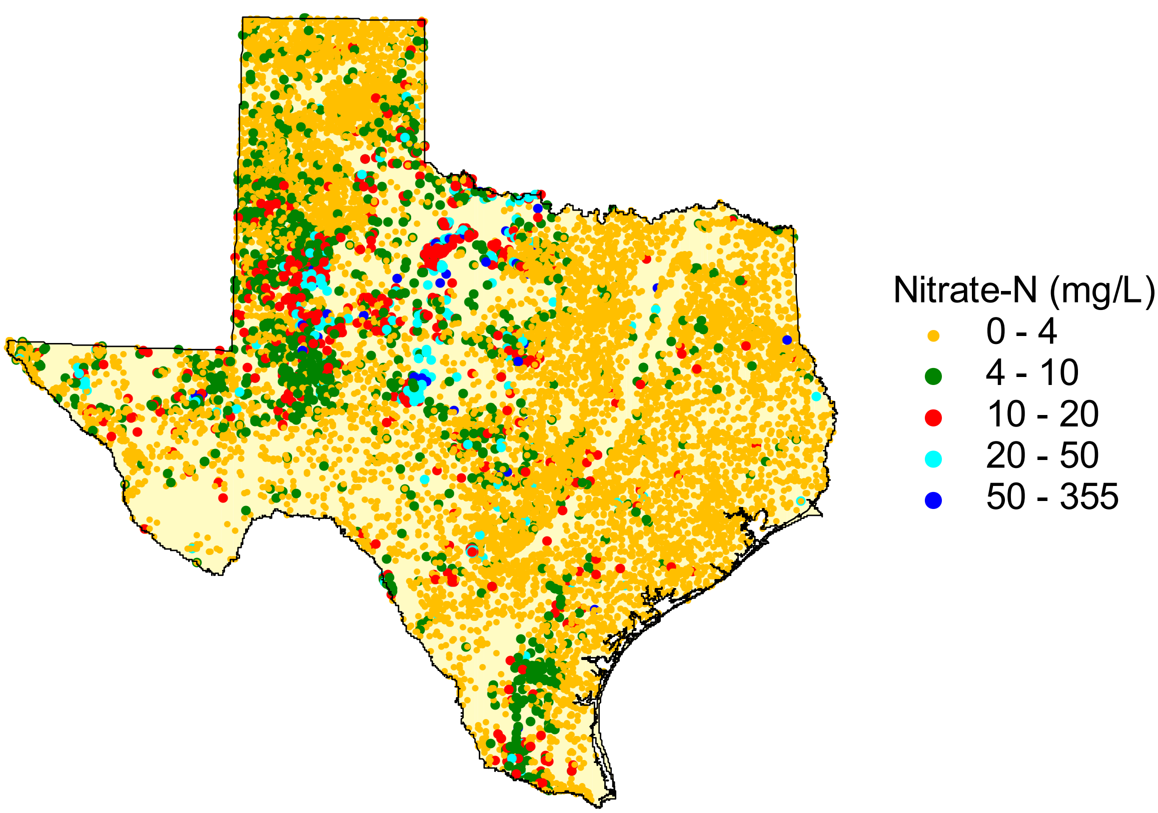
EVALUATION OF NITRATE CONTAMINATION IN MAJOR POROUS MEDIA AQUIFERS IN TEXAS

Bridget R. Scanlon, Robert C. Reedy, and Katherine S. Kier



Final Report prepared for Texas Commission on Environmental Quality

Bureau of Economic Geology
Scott W. Tinker, Director
John A. and Katherine G. Jackson School of Geosciences
The University of Texas at Austin
Austin, Texas 78713-8924

2003

ABSTRACT

Nitrate is one of the most pervasive contaminants in groundwater in Texas, exceeding maximum contaminant levels in many aquifers in the State. The purpose of this study was to assess controls on nitrate contamination in major porous media aquifers in the state by comparing groundwater nitrate concentration data with nitrogen loading and aquifer susceptibility parameters. Attributes characterizing nitrogen loading included atmospheric deposition, inorganic and organic fertilizers, land use, proxies for sewage and septic input, population density, precipitation, and irrigation. Attributes characterizing aquifer susceptibility to contamination included percent land surface slope, percent well drained soils, clay content, and organic matter content. Multivariate logistic regression was used to relate the probability of nitrate concentrations in shallow wells (≤ 30 m) exceeding a pre-specified threshold value of 4 mg/L nitrate-N with potential explanatory variables representing nitrogen loading and aquifer susceptibility. The final regression model included precipitation, percent agricultural land, low density residential land, and soil organic matter. Observed and predicted probabilities of elevated nitrate concentrations were highly correlated in calibration and validation data sets (R^2, 0.96; 0.98). The inverse relationship between precipitation and nitrate concentration may be related to dilution in high precipitation areas and possibly evapoconcentration in low precipitation areas. Although nitrate loading is not explicitly represented in the final model, percent agricultural land may be considered a proxy for nitrogen loading from agricultural sources and low density residential land use may be considered a proxy for septic tank effluent. Percent organic matter may reflect the influence of denitrification in some regions. This GIS and logistic regression analysis described in this study provides valuable insights into controls on the distribution of nitrate concentrations in groundwater and should be supplemented in future studies with field sampling to ground reference the GIS and logistic regression analysis of this study and to assess the impact of different processes such as dilution and denitrification on nitrate concentrations.

INTRODUCTION

Nitrate is the most pervasive contaminant in groundwater in Texas and in the U.S. (TCEQ, 2002; Nolan et al., 2002). In this report all nitrate concentrations are expressed as elemental nitrogen and the term nitrate is used to refer to nitrate-nitrogen. Nitrate exceeds the maximum contaminant level (MCL) of 10 mg/L set by the EPA (1995) under the Safe Drinking Water Act in many domestic and public water supply wells, particularly in the Southern High Plains, Seymour, Trinity, and southern Gulf Coast aquifers. Although public water supply wells are routinely monitored, domestic wells are not regulated or monitored routinely.

High nitrate concentrations in groundwater can have adverse health impacts. Methemoglobinemia in infants is a potentially fatal disease and results from low oxygen levels in the blood caused by ingestion of high nitrate groundwater (Spalding and Exner, 1993). A total of eight spontaneous abortions in four women in Indiana (1991 – 1994) may be related to high nitrate concentrations (19 – 29 mg/L nitrate) in domestic well water in rural regions of Indiana (Centers for Disease Control and Prevention, 1996). Increased risk of non-Hodgkin's lymphoma has been related to nitrate concentrations ≥4 mg/L nitrate in community water supply wells in Nebraska (Ward et al., 1996). Toxicological studies indicate that multi-contaminant exposure may have a much greater impact on health than exposure to single pure contaminants because of additive or synergistic interactions among compounds (Squillace et al., 2002). Adverse health impacts are much greater for mixtures of nitrate and pesticides (Porter et al., 1999) and suggest that the MCL for nitrate may be reduced in the future, which would greatly affect water availability in Texas. Nitrate concentrations ≥2 mg/L in groundwater are considered to be impacted by human activities (Mueller and Helsel, 1996).

Nitrate is highly soluble in water and is not prone to ion exchange (Stumm and Morgan, 1996). The anionic form of nitrate does not sorb onto clay particles which are also negatively charged under normal pH conditions. Nitrate also cannot be lost through volatilization because it is nonvolatile. The high solubility and mobility of nitrate results in nitrate being readily leached through the soil zone to underlying aquifers.

Nitrate is not affected by chlorination, the most common method of treating most public water. It can be removed from water by reverse osmosis, although this is an expensive process. Additional treatment technologies include ion exchange and denitrification (Kapoor and Viraraghavan, 1997). Commonly water supply companies try to reduce nitrate concentrations by blending water with groundwater/or surface water that contain~~s~~ low nitrate concentrations. Another water treatment option involves extending wells to greater depths where nitrate concentrations are often lower (McMahon et al., 2003).

Potential sources of nitrate contamination in groundwater include atmospheric deposition, natural sources, inorganic fertilizer, organic fertilizer or manure, septic tanks, and leaking sewer systems. Natural sources result from nitrogen fixation by legumes. Many previous studies have attempted to relate the distribution of nitrate in groundwater to various potential sources. Evaluation of nitrate contamination and relationship to explanatory variables has been conducted on a national scale in many previous studies (Nolan et al., 2002; Squillace et al., 2002). Original studies used geographic information systems (GIS) overlay analysis and statistical analysis to determine risk of nitrate contamination in shallow aquifers (Nolan et al., 1997). Univariate analysis of nitrate contamination and potential explanatory variables (Nolan and Stoner, 2000) was generally unsatisfactory because there was considerable unexplained variation when each variable was considered. Equal weighting applied to potential explanatory variables in GIS overlays and univariate statistical methods does not describe complex interrelationships between various explanatory variables and nitrate contamination. Use of logistic regression represents considerable advancement in assessing the risk of nitrate contamination in various aquifers because it incorporates a large number of potential explanatory variables and assigns weights to these variables based on slope coefficients determined from measured data (Nolan et al., 2002).

The purpose of this study is to evaluate potential sources and processes controlling nitrate contamination in major porous media aquifers in Texas. Potential explanatory variables used for nitrate contamination include nitrogen loading (e.g. atmospheric deposition, inorganic and organic fertilizer, leaking septic tanks and sewers, sludge applications, concentrated animal feeding operations (CAFOs)), aquifer susceptibility to contamination (soil drainage characteristics, soil clay content and organic matter content), and other factors.

Terminology

The terms susceptibility, vulnerability, and risk are used to describe the potential for aquifer contamination (Evans and Maidment, 1995). Susceptibility of an aquifer to contamination represents the ability of contaminants to reach an aquifer but does not include any information about contaminant source or loading. Vulnerability includes susceptibility combined with contaminant loading. The contamination risk includes the probability that the contaminant is present in the aquifer and can be quantified. 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 location above the uppermost aquifer" Rupert (2003). EPA (1993) includes factors such as water table depth, geology, and soils into hydrogeologic sensitivity which is grouped with contaminant loading into aquifer vulnerability. Groundwater vulnerability mapping has been conducted in many areas using the DRASTIC model (Aller et al., 1985). The DRASTIC model includes: Depth to water, net Recharge, Aquifer media, Soil media, Topography, Impact of vadose zone media, and hydraulic Conductivity of the aquifer. The questionable success of DRASTIC in predicting groundwater vulnerability (Koterba et al., 1993; Rupert, 2001) has been attributed to the subjective point rating system that is based on best professional judgement and the lack of calibration to actual groundwater quality data. To overcome some of these problems with traditional DRASTIC mapping, Rupert (2001) calibrated the vulnerability point ratings to measured nitrate concentrations in ground water using nonparametric statistical tests. The logistic regression approach overcomes some of the deficiencies of traditional vulnerability mapping also by calibrating to actual contaminant concentration data.

Logistic Regression

Logistic regression is widely used in social sciences research and for epidemiological studies to assess risk. The use of logistic regression to produce probability maps of groundwater contamination with potential explanatory variables has increased in the past decade. Logistic regression has been used in several national assessments of nitrate and pesticide contamination (Nolan et al., 1998; 2002; Nolan and Stoner, 2000; Nolan, 2001). The most recent study reported by Nolan et al. (2002) indicated that nitrogen fertilizer loading, percent cropland-pasture, log of human population density, percent well drained soils, depth to seasonally high water table and presence/absence of unconsolidated sand and gravel aquifers were important in explaining groundwater elevated nitrate concentrations in shallow wells (≤ 4 mg/L) in the U.S. More localized evaluation of contamination has also been conducted in Colorado by Rupert (2003).

The approach used to evaluate nitrate contamination has been to represent groundwater nitrate concentration as a bivariate dependent variable by selecting a threshold nitrate concentration to represent nitrate concentrations that exceed natural background levels. Threshold nitrate concentrations have ranged from 2 mg/L (Rupert, 1998); 3 mg/L (Squillace et al., 2002), 4 mg/L (Nolan, 2002), and 5 mg/L (Rupert, 2003) in different studies. Groundwater nitrate concentrations are then related to various explanatory variables that include nitrogen loading (atmospheric deposition, organic and inorganic fertilizer application, sewage systems, septic tanks) and parameters related to loading (precipitation, irrigation) and ability of soils to transmit contaminants from the land surface (well drained soils, land surface slope, clay content, and soil organic matter).

Logistic regression is used to predict binary dependent variables using independent variables and to assess the percent of variance in the dependent variable that can be explained by the independents and to determine the relative importance of different independent variables

(Kleinbaum, 1994; Hosmer and Lemeshow, 2001). One of the primary differences between logistic regression and ordinary linear regression is that the dependent variable is the probability of being in a category (i.e. > 4 mg/L NO_3) rather than the measured value of the dependent variable. Ordinary least squares regression cannot be used with binary dependent variables because variables have to be normally distributed and binary variables do not fit this requirement. Logistic regression is much less stringent than ordinary least squares regression and does not assume a linear relationship between the independent and dependent variables, does not require variables to be normally distributed, and does not require homoscedasticity (uniform variance with X). Ordinary regression is used to predict a continuous dependent variable from one or many independent variables (x, predictors) by finding values of b_0, b_1, b_2 etc.

$$y = a + b_1 * x + b_2 * x_2 + ... \tag{1}$$

Logistic regression is used when the dependent variable is limited to 2 values (e.g. presence or absence of nitrate concentrations with respect to a threshold concentration, 3, 4, or 5 mg/L). The resultant equation from logistic regression is used to determine the probability of the occurrence of the dependent variable as a function of the independent variables. The odds ratio is the probability of occurrence of an event, e.g. probability of exceeding a threshold value, divided by the probability of the event not occurring.

$$Odds\ ratio = \frac{P}{1-P} \tag{2}$$

The odds ratio provides information on the number of times the outcome occurs or does not occur when the predictor is increased by 1 unit. The odds ratio is constrained between 0 and 1. To make the odds of an event occurring relative to the odds of an event not occurring symmetrical, the natural log is used. If P is greater than 0.5, ln(P/1-P) is positive whereas if P < 0.5, ln(P/1-P) is negative. In logistic regression, the dependent variable is a logit (i.e. natural log of the odds ratio) (Helsel and Hirsch, 1992):

$$\ln(odds\ ratio) = \log it(P) = \ln\left(\frac{P}{1-P}\right) \tag{3}$$

In logistic regression logit(P) is a linear function of the independent variables. Odds ratios can be converted back to probabilities as follows:

$$\ln\left(\frac{P}{1-P}\right) = b_0 + b\ X; \qquad \frac{P}{1-P} = e^{b_0 + bX}; \qquad P = \frac{e^{b_0 + bX}}{1 + e^{b_0 + bX}} \quad or \quad P = \frac{1}{1 + e^{-(b_0 + bX)}} \tag{4}$$

where P is probability of a 1 or the occurrence of a contaminant concentration greater than a threshold value in our case. If ln(odds) is linearly related to X, then P and X are nonlinearly related and form an S shaped curve. The variance is P(1-P) and is not constant with X (i.e. not homoscedastic). The variance is a maximum at P = 0.5 and approaches zero as P approaches 1 or 0.

Model parameters are generally chosen to maximize the goodness of fit between the measured and simulated values. In ordinary least squares regression, the sum of squared distances of the data points to the regression line are minimized to estimate the coefficients in the regression equation. In logistic regression, there is no mathematical solution to produce least squares estimates of parameters. Maximum likelihood estimation optimizes the fit by maximizing the log likelihood (LL) which represents how likely it is (the odds that the measured values of the dependent variable may be predicted from the measured values of the independent variables). A likelihood is a conditional probability (e.g. P(YIX) the probability of Y given X) or probability of the measured values of the dependent variable may be predicted from the observed values of the independent variables. The likelihood varies from 0 to 1 like any probability. Because the probability is a small number the natural log of this number is used which varies from 0 to minus infinity. The natural log of this number is generally multiplied by -2 to make the number positive. The null hypothesis is that the LR coefficients are zero. Therefore, the statistic -2LL (-2 log likelihood) is a badness of fit indicator. An iterative process is used to determine the parameters of the logistic regression (i.e. b_0 and b) to maximize the likelihood (conditional probability of the data given parameter estimates) of the sample data until convergence is achieved. When large samples are used, -2LL is chi-square distributed.

The log-likelihood test of a model, also called the model chi-square test or likelihood ratio test or G statistic tests the statistical significance of coefficients in the logistic regression model (Hosmer and Lemeshow, 1989).

$$G = -2(L_{\text{int}} - L_{\text{mod}\,el}) \tag{5}$$

where L_{int} is log-likelihood of intercept only and L_{model} is log-likelihood of model with explanatory variables. The G statistic is chi square distributed and the null hypothesis is that the slope coefficients for the explanatory variables are 0. The G statistic is used to compare predicted values with observed values of the dependent variable with and without different explanatory variables. A well fitting model will have a low p value.

The Wald statistic can be used to test the significance of individual logistic regression coefficients for independent variables. The Wald statistic is the ratio of the unstandardized logit coefficient to its standard error. If logit coefficients are large, the standard error is inflated which lowers the Wald statistic and may result in false negatives (i.e. effect not significant when it is) (Menard, 2002).

The Hosmer and Lemeshow's (HL) goodness of fit test provides a test of the overall model (Hosmer and Lemeshow, 2001). The HL test differs from the G statistic in that only one model is evaluated in the HL test whereas the G statistic compares models with and without specific explanatory variables. The data are divided into deciles based on predicted probabilities and chi squares are computed from observed and expected frequencies. A probability value is calculated from the chi square distribution with 9 degrees of freedom to test the fit of the logistic regression model. The null hypothesis is that the model fits the data; therefore, higher p values indicate a better fit.

The goodness of fit was also evaluated using linear regression of predicted probabilities for deciles of risk used to calculate the HL statistic versus observed probabilities of elevated nitrate concentrations. Higher coefficients of determination (R^2) values indicate better fits. In addition, predicted and observed probabilities were plotted and compared with a 1:1 line with a zero intercept. A perfect fit between predicted and observed probabilities would plot along the 1:1 line.

Nitrogen Cycle

The nitrogen cycle includes the transport of nitrogen from the atmosphere through various chemical and biological transformations in the subsurface and returning back to the atmosphere (Fig. 1). The original source of much of the nitrogen in the subsurface is inorganic nitrogen from atmospheric deposition. Gaseous nitrogen is converted to solid forms of nitrogen either by (1) atmospheric fixation, (2) biological fixation, or (3) industrial fixation. Atmospheric fixation results from lightning. Biological fixation occurs when bacteria, cyanobacteria and/or actinomycetes fix nitrogen. Certain plants, e.g. legumes (soybeans, alfalfa, peas) form symbiotic relationships with nitrogen fixing bacteria in nodules within the plants. In return for nitrogen, the bacteria receive carbohydrates from the plant. Industrial fixation involves formation of ammonia which can be applied directly as a fertilizer but is generally processed further to urea ($CO(NH_2)_2$) and ammonium nitrate (NH_4NO_3). The primary source of nitrogen in soils is from organic matter in plant and animal residues. Although organic forms of nitrogen are readily available in soils, plants cannot use these forms of nitrogen and rely on soil bacteria (decomposers) to convert organic forms of nitrogen to inorganic forms which plants can take up through their roots. Mineralization is the process of converting nitrogen found in organic matter from ammonia (NH_3) to ammonium salts (NH_4^+). The positively charged ammonium ion can be sorbed onto negatively charged clay particles and transported as a colloid. Alternatively, ammonia can be converted to nitrate in a two step process termed nitrification: bacteria of the genus Nitrosomonas converts ammonia to nitrite (NO_2^-) and the genus Nitrobacter converts nitrite to nitrate. Both steps involve odixation. Nitrate can be lost from the system by (1) plant uptake, (2) runoff, (3) leaching, and (4) denitrification. Runoff transports nitrogen in the soil to the streams whereas leaching transports nitrates below the root zone to underlying aquifers. Nitrate in the hydrologic cycle can return to the oceans where it is denitrified and returned to the atmosphere. The term denitrification refers to reduction of nitrate to nitrogen gas (N_2) or nitrous oxide gas (N_2O). Denitrification can take place in anaerobic soils or aquifers and the resultant gases can diffuse through the subsurface into the atmosphere. Denitrifying bacteria use nitrates as an alternative to oxygen for the final electron acceptor in their respiration.

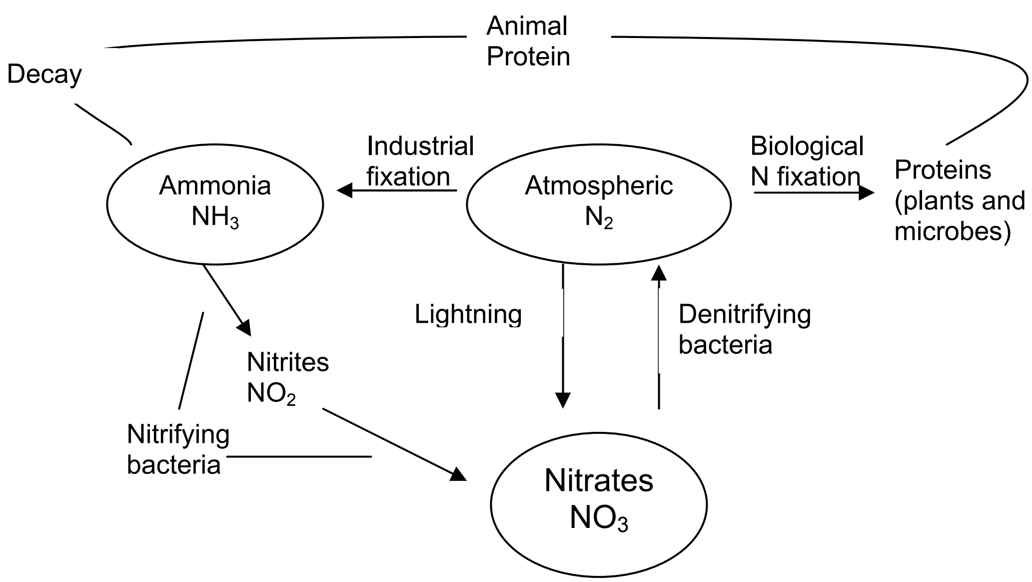


Figure 1. Schematic of nitrogen cycle.

Anthropogenic influences on the nitrogen cycle include increased atmospheric deposition as result of fossil fuel combustion and forest fires, application of nitrogen fertilizers (inorganic and organic (animal manure, sewage)) forms to crops, livestock ranching and concentrated animal feeding operations, and sewage waste and septic tank effluent. Urea in fertilizers and manure on the land surface can volatilize and escape into the atmosphere.

Previous Studies of Nitrate Contamination in Texas

Studies of nitrate in soils and groundwater have been conducted in different regions of Texas. Kreitler (1975) developed a nitrogen isotope technique for distinguishing different sources of nitrate contamination. The study was conducted in Runnels County which had the highest average nitrate concentrations in groundwater in the state (53 mg/L NO_3). Nitrate derived from mineralization of organic nitrogen in cultivated soils in Runnels County had $\delta^{15}N$ of 2 to 8‰ whereas nitrate associated with animal waste near farmhouse barnyard complexes had $\delta^{15}N$ of +10 to +22‰ (Kreitler, 1975; Kreitler and Jones, 1975). Kreitler and Browning (1983) used nitrogen isotope analysis of nitrate in groundwater to show that nitrate in the Edwards aquifer could be attributed to naturally occurring nitrogen compounds in the recharge streams. The $\delta^{15}N$ of 73 groundwater samples ranged from +1.9 to +10‰ which was similar to the range found in the recharge streams (+1 to +8.3‰). The lack of enrichment in $\delta^{15}N$ indicated that animal waste sources of nitrogen were not present. Studies of nitrate contamination in alluvial fan aquifers in central Texas (siliceous gravel aquifers in Seymour and carbonate gravel aquifers in Lockhart and Taylor regions) indicated that fertilizer was the primary source of nitrate in cultivated fields whereas animal wastes were the dominant sources in domestic well water (Kreitler, 1979). The $\delta^{15}N$ of 20 soil nitrate samples ranged from +2 to +14‰ which is more enriched than that of the fertilizer (-7.4 to +1.9‰) because of volatilization of the ammonium based fertilizers. Studies of groundwater nitrate in the Seymour aquifer by Bartolino (1994) suggested that high nitrate concentrations in groundwater (late 1950s) predate the use of fertilizers which began in the mid 1960s and may be attributed to symbiotic nitrogen fixation by mesquite trees which were replaced by crops. The cultivation process oxygenated the soils converting organic nitrogen to nitrate and increased recharge which leaches the nitrate to the underlying aquifer.

Playas in the southern High Plains have been used for industrial wastewater, sewage and feedlot runoff. Previous studies have evaluated nitrate loading related to wastewater discharge from CAFOs to playas. Fryar et al. (2000) summarized much of the previous research. Nitrate nitrogen concentrations in core extracts decreased from 189 mg/kg at 0.3 m depth to 1.5 mg/kg at 1.2 m depth beneath a playa receiving waste water discharge from a feedlot (Lehman et al., 1970). Resampling of this site 5 yr later showed similar reductions in nitrate concentrations; however, chloride concentrations had increased by factors of 2 to 5 (Clark, 1975). Similar decreases in nitrate concentrations with depth were recorded beneath other playas adjacent to feedlots (Stewart et al., 1994; Daniel, 1997). Nitrate reduction from CAFO runoff may be attributed to sealing of surface soils caused by deposition of suspended solids and denitrification (Lehman and Clark, 1975; Roswell et al., 1985; Barrington and Broughton, 1988). A study of wells adjacent to 26 feedlots indicated that the highest nitrate concentration was 9.5 mg/L NO_3. Fryar et al. (2000) conducted a detailed study of the unsaturated and saturated zone in the Southern High Plains that showed denitrification in the unsaturated zone. High $\delta^{15}N$ values (> 12.5‰) in groundwater and correlations between $\delta^{15}N$ and the natural log of nitrate concentrations suggest denitrification; however, high O_2 concentrations in groundwater indicate that denitrification in groundwater is unlikely. The presence of denitrifying bacteria in cores, soil gas $\delta^{15}N$ values < 0‰, and decreases in NO_3/Cl and SO_4/Cl^- ratios with depth in cores indicate that denitrification occurs in the upper unsaturated zone.

METHODS

In this study the relationship between elevated nitrate concentrations in groundwater and potential explanatory variables was examined using GIS and logistic regression. Parameters related to nitrogen loading included precipitation, irrigation, atmospheric deposition, inorganic and organic fertilizer application, CAFO and sludge application locations, low density and high density residential and agricultural land use, and population density. Aquifer susceptibility to nitrate contamination was examined using land surface slope, percent well drained soils, clay content, organic matter content, available soil water content, and depth to seasonally high water table.

Groundwater nitrate concentrations were obtained from the TWDB database on ambient groundwater quality. Other forms of nitrogen, such as ammonia (NH_3) or nitrite (NO_2) were not included in this study because they have relatively low concentrations as a result of reduced mobility, increased chemical instability, and reduced loadings relative to nitrate (Nolan et al., 2002). All nitrate concentrations in this study are reported as elemental nitrogen. The detection limit for nitrate in the database was 0.1 mg/L. The TWDB database includes information on the well location and depth, drill date, primary water use (domestic, irrigation industrial, commercial), water quality sampling time, and major ion chemistry. To avoid overrepresentation of wells that were sampled multiple times, the TWDB database was screened for the most recent water quality sample between 1980 and 2002. This time period was used to provide the greatest number of records and because no time trends were obvious from the data. The resultant set of sampled wells contained 14,985 records.

Average annual precipitation was based in precipitation data from 1961 – 1990 and was gridded at 60 m resolution. The distribution of irrigated land was based on land use/land cover data developed by the USGS and published in 1994. Nitrate loading from atmospheric deposition was based on wet deposition of nitrate in precipitation from the National Atmospheric Deposition Program (NADP, http://nadp.sws.uiuc.edu/). There are 5 NADP stations in Texas that are fairly uniformly distributed throughout the state. Data from 1980 through 2000 were averaged and contoured to determine nitrate loading for each well. Inorganic and organic fertilizers are considered major sources of nitrate in the subsurface. Annual nitrate loading from inorganic commercial fertilizer was obtained from county fertilizer sales for 1998 using a procedure similar to that described in Battaglin and Goolsby (1994) and Goolsby et al. (1999). Nitrogen fertilizer loading was apportioned equally to agricultural (MRLC 61, 71, 81, 82, 83, and 84) and urban (21, 22, and 85) land uses. Urban fertilizer use represents addition of fertilizer to residential lawns, parks, and golf courses. Nitrate loading from organic sources (manure) was estimated from the animal population in counties based on the 1998 Census of Agriculture statistics and the amount of manure produced by each animal (Lander et al., 1998). Organic fertilizers were only applied to agricultural cultivated regions.

A statewide coverage of permitted concentrated animal feeding operations (CAFO) was obtained from the Texas Commission on Environmental Quality. Texas Institute for Applied Environmental Research (TIAER) provided CAFO coverage for dairies located in the Trinity River watershed. The two datasets were combined and duplicate CAFOs were omitted. Additional information on CAFOs was obtained from a land use/land cover developed by the USGS and published in 1994. The majority of these CAFOs are located in east Texas where poultry is the dominant type of CAFO. The TCEQ and TIAER datasets contained information on the location of the CAFOs, the permit number, the owner/operator(s), the permitted number of animal units it contains, and the type of animals. The TCEQ rules and regulations in Chapter 321 Subchapter B § 321.21- 321.49 define a CAFO as any animal feeding operation which the executive director designates as a significant contributor of pollution or any animal feeding operation defined, in the most basic instances as: "any new and existing operations which stable and confine and feed or maintain for a total of 45 days or more in any 12-month period more than the numbers of animals specified in any of the following categories: 1,000 cattle; 700 mature dairy cattle; 2,500 swine > 55 pounds or 10,000 weaned ≤ 55 pounds; 500 horses; 10,000 sheep; 55,000 turkeys; 100,000 laying hens or broilers where the facility has unlimited continuous flow watering systems; 30,000 hens or broilers when the facility has a liquid waste handling system; 5,000 ducks; or 1,000 animal units from a combination of slaughter steers and heifers, mature dairy cattle, swine > 55 pounds, and sheep. The TCEQ rules and regulations outlined for controlling CAFOs are even more specific for operations that discharge pollutants into waters. The combined TCEQ and TIAER datasets represent 516 CAFO locations for which there is latitude/longitude information. Of these CAFOs, six are poultry, 310 dairy, 164 feedlot, two sheep, and 34 swine. In addition to CAFOs, the location of TCEQ permitted Class B waste water treatment plant sludge and domestic septage were also evaluated relative to groundwater nitrate concentrations. Population density was also included as a potential explanatory variable for nitrate contamination. County statistics on population density were obtained for 1990 census.

Information on land use was obtained from National Land Cover Data (NLCD) (Vogelmann et al., 2001; USGS, 2000). This product was developed by the Multi-Resolution Land Characteristics (MRLC) consortium that included EPA, USGS, NOAA and USFS, NASA and BLM and development of the NLCD was directed by USGS and EPA. The NLCD was derived from images acquired by LandSat Thematic Mapper sensor from the early to mid-1990s and other data sources. NLCD is a 21-class land cover classification scheme (Table 1). The spatial resolution of the data is 30 m. Urban land use settings included low and high-density residential settings. Low density residential setting is used as an indicator or surrogate of septic systems and high density residential setting as a surrogate for leaking sewers. Agricultural land use includes orchards/vineyards (61), pasture/hay (81), row crops (82), small grains (83), and fallow (84).

Table 1. National Land Cover Data (NLCD) classification scheme.

Number	*Land Use*
21	Low density residential
22	High density residential
23	Commercial/industrial/transportation
41	Deciduous forest
42	Evergreen forest
43	Mixed forest
51	Shrubland
61	Orchards/vineyards
71	Grasslands/herbaceous
81	Pasture/hay
82	Row crops
83	Small grains
84	Fallow
85	Urban/recreational grasses
91	Woody wetlands
92	Emergent herbaceous wetlands

Much of the information required to asses aquifer susceptibility to contamination was based on the State Soil Geographic (STATSGO) database (USDA, 1994). The STATSGO database is mapped at a 1:250,000 scale. STATSGO mapped units consist of from 3 to 21 soil series or components in Texas. The average slope of the land surface and the average water content for each soil series were area weighted for each STATSGO map unit. A depth weighted average clay content and percent organic matter were calculated for each soil profile and these values were area weighted for each map unit. Soils are subdivided into 4 drainage classes (A, B, C, and D) that are grouped into well drained (A, B) and poorly drained (C, D). A single value of drainage was provided for each soil series and these values were area weighted for each map unit. Values of these soil parameters for each well were based on the map unit in which it was located.

Logistic Regression

Logistic regression was used to determine which potential explanatory variables are important in predicting the probability of groundwater nitrate concentrations greater than a pre-specified threshold nitrate concentration of 4 mg/L. A value of 4 mg/L was chosen because it has been previously related to the occurrence of non-Hodgkin's lymphoma (Ward et al., 1996) and has been used in previous national assessments of nitrate contamination (Nolan et al., 2002). Using a value of 4 mg/L provides an opportunity for identifying potential problem areas before they exceed the MCL. The nitrate database used for logistic regression was filtered from the general TWDB database of ~ 15,000 for the most recent well sampled between 1980 and 2002. The regression analysis focused on shallow wells (≤ 30 m) which resulted in 969 sampled wells.

A total of 18 potential explanatory variables was considered in the logistic regression analysis. Land cover was modeled as percentage of a specified land cover classification within a pre-specified circular buffer region of each well. Different radial distances around well locations were examined to determine the optimal value for land use percentage calculations and nitrogen loading. Percentages by land use category were calculated for buffer distance values of 100, 250, 500, 1000, 2000, and 4000 m. Preliminary univariate logistical regression models were constructed for each resulting value and McFadden's ρ^2 statistic for each model was plotted against the distance and examined for the optimal (maximum) correlation. The McFadden's statistic is:

$$\rho^2 = 1 - \frac{LL(b_0, b)}{LL(b_0)} \qquad (6)$$

where the second term on the right represents the ratio of the log-likelihood of the model with intercept and slope variables (b_0,b) to that of the model without the slope (i.e., an intercept-only model). The rho squared statistic is analogous to the r squared statistic in linear regression and ranges form 0 – 1; however, the rho squared statistic is generally much lower than R^2. The results indicated that 2000 m resulted in the best correlation (Figure 2).

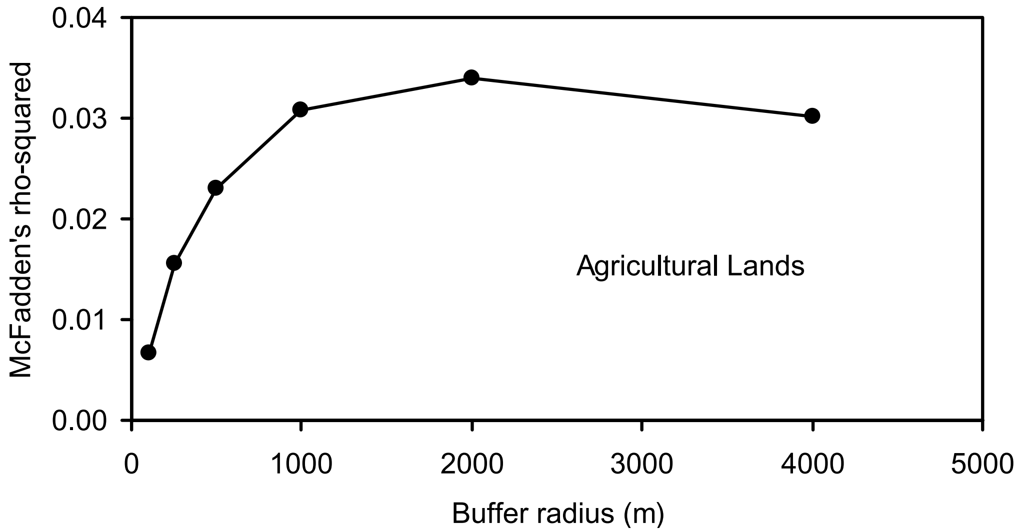


Figure 2: Well buffer radius versus McFadden's ρ^2 statistic for percentage of agricultural land use.

Development of a logistic regression model is a multi-step process that consists of examining individual variables and combinations of variables to determine the best fitting model that describes the dependent variable, nitrate concentrations > 4 mg/L in shallow wells (≤ 30 m). The first step consists of a univariate analysis of each potential explanatory variable and the Wald p-value is examined to determine if the variable has a significant effect on the dependent variable. The p value indicates whether the slope of the relationship is significantly different from 0. The null hypothesis is that there is no relationship between the variables. The p value reflects the significance of the relationship: smaller p values indicate increasing significance. For example a p value of 0.05 indicates a 95 percent significance level whereas a p value of 0.001 indicates a 99 percent significance level. The significant variables are then grouped in various combinations using a forward and/or backward iterative process to identify the combination of variables that best predicts the dependent variable. The statistical significance of each interim model is evaluated using the Hosmer-Lemeshow (HL) statistic, which measures goodness-of-fit and compares the observed and predicted deciles of risk, along with the Wald p-values of the individual variables included in the model. At different points in the process, variables that were initially significant may become insignificant when combined with other variables, and vice-versa. Thus the final model is tested by sequentially including variables that had been previously eliminated to test whether they should be included in the final model. In general, individual variables should have a Wald p-value ≤0.05 to remain in the model. Finally, the selected variables are examined for linearity in the logit (i.e., plots of variable values versus the logit should be approximately linear). The logistic regression model is developed using a calibration data set and then tested using a validation data set. The nitrate data set used for logistic regression was screened from the general TWDB data set to exclude samples from wells that did not include information on well depth. The screening also included only wells ≤ 30 m deep to evaluate nitrate contamination in the shallow zone. This resulted in a total of 969 wells that met the screening criteria, of which 235 wells were used as a validation data set and the remaining 734 wells were used to construct the logistic regression model. The validation data set was selected using a spreadsheet-generated, evenly-distributed random number between 0 and 1 and all wells with an associated random number value ≤ 0.25 were included in the validation data set.

RESULTS AND DISCUSSION

Characteristics of Nitrate Concentrations in Major Aquifers

The TWDB data base for all aquifers (major and minor) in Texas with the most recent reported nitrate concentration from 1980 to 2002 consisted of ~ 15,000 records. Approximately 7 percent of these records exceeded the EPA MCL of 10 mg/L nitrate as nitrogen. This set of records was filtered to represent only major aquifers that included well depth, which resulted in approximately 8,500 records. Nitrate concentrations are highly variable in the major aquifers in the state (Fig. 3; Table 2). Median values of nitrate were used to represent the central tendency because they are less sensitive to outliers than means. Sample range was represented by the interquartile range rather than standard deviation for similar reasons. Most wells are used for domestic water, public water supply, irrigation, and stock purposes. Less than 7 percent of the wells are used for industrial and commercial purposes in each of the major aquifers (App. A). Median nitrate concentrations range from less than the detection limit to 13 mg/L in the Seymour aquifer. The percent of wells that exceeded the MCL in each aquifer ranged from about 1 to 66 percent. Maximum nitrate concentrations were much greater than median concentrations in all aquifers and reflect outliers in the data. Median nitrate concentrations in domestic wells are greater than those in public water supply wells (App. A). This difference in nitrate concentrations may reflect the shallower depth of domestic wells. Differences in median nitrate concentrations between unconfined and confined aquifers are not very high; however, maximum nitrate concentrations are much greater in unconfined aquifers relative to confined aquifers (Table 2, App. A).

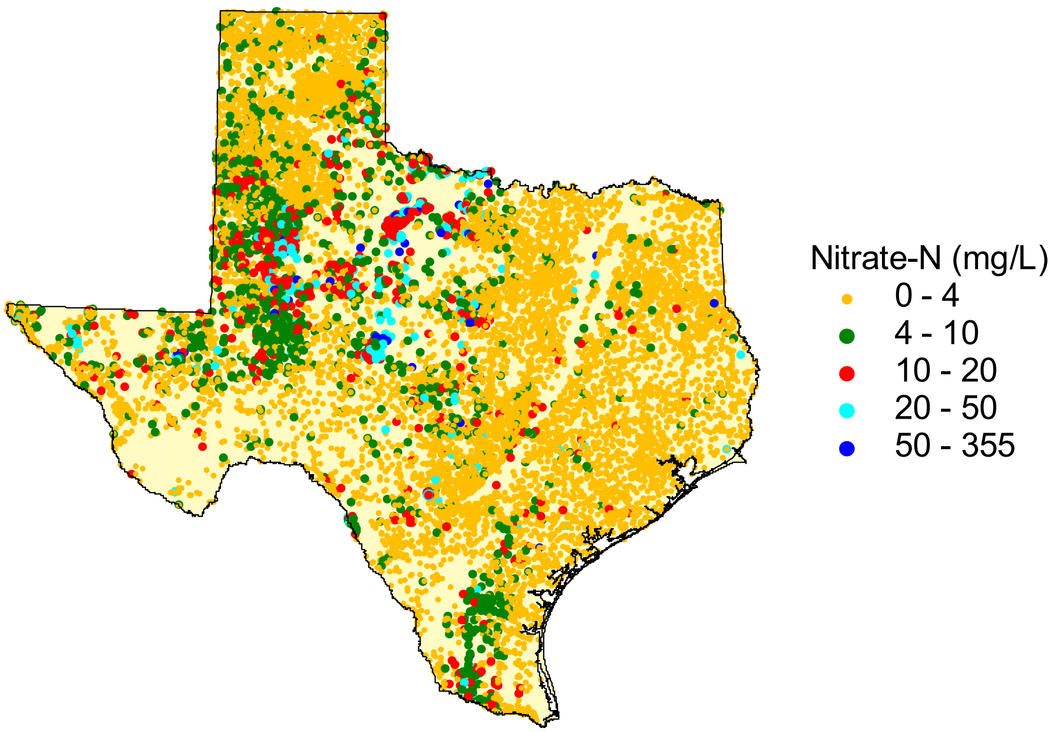


Figure 3. Nitrate concentrations in the most recent samples collected between 1980 and 2002 from wells in all aquifers in Texas based on the TWDB ambient groundwater monitoring database. A total of 14,985 sampled wells are represented.

Table 2. Number of nitrate analyses from wells in each of the major aquifers representing the most recent sample collected between 1980 and 20002; number of analyses ≤ 4 mg/L, ≥10 mg/L (EPA MCL), percent of samples ≥ 10 mg/L, median concentration, minimum and maximum concentrations, and 10, 25, 75, and 90 percentile values based on 10,322 analyses for the major aquifers. (Cen. Pec. All., Cenozoic Pecos Alluvium; Ed.-Trin. Plat., Edwards Trinity Plateau; HMB, Hueco Mesilla Bolson).

	No.	*≤ 4mg/L*	*>10mg/L*	*%*	*Median*	*Min*	*Max*	*10th %*	*25th %*	*75th %*	*90th %*
Carrizo-Wilcox	1339	1302	14	1	<0.02	0	70.65	<0.01	<0.01	0.06	0.35
Unconfined	582	555	7	1	<0.03	0	70.65	<0.01	<0.01	0.13	1.12
Confined	757	747	8	1	<0.02	0	27.37	<0.01	<0.01	<0.05	0.14
Cen. Pec. All.	185	132	19	10	1.45	0	174.55	0.03	0.57	4.61	10.38
Edwards BFZ	629	570	8	1	1.50	<0.01	22.96	<0.02	0.50	2.07	3.73
Ed.-Trin. Plat.	1039	728	61	6	2.35	<0.01	213.54	0.1	1.07	4.53	7.45
Gulf Coast	1752	1578	53	3	<0.05	0	50.82	<0.01	<0.02	0.7	3.98
HMB	274	250	3	1	1.21	0	16.21	0.05	0.39	2.14	3.77
Ogallala	3206	2422	284	9	2.01	0	94.54	0.24	0.41	2.09	3.37
Seymour	236	32	155	66	12.95	0.08	334.92	3.23	8.24	19.39	26.31
Trinity	1662	1530	42	3	0.05	0	60.96	<0.01	<0.01	0.63	2.92
Unconfined	888	772	37	4	0.05	0	60.96	<0.01	<0.02	0.62	2.88
Confined	774	758	5	1	<0.02	0	20.29	<0.01	<0.01	0.1	0.57

The median well depth for each aquifer ranges from 13 to 198 m (Table 3). Well depths were shallowest in the Seymour aquifer. A plot of nitrate concentrations versus well depth indicates that there is a lot of variability in the data (Fig. 4). The locally weighted scatterplot smooth (LOWESS) line indicates that nitrate concentrations decrease with depth in the aquifer. The break in slope of the LOWESS line at 74 m indicates that reduction in nitrate concentrations with depth is much greater in the shallow zone and is much less at greater depths.

Table 3. Median well depth for each of the major aquifers and for unconfined and confined portions of the Carrizo Wilcox and Trinity aquifers.

Aquifer Name	*Aquifer*	*Unconfined*	*Confined*
	Well depth (m)		
Carrizo-Wilcox	167	96	240
Cenozoic Pecos Alluvium	76		
Edwards (BFZ)	152		
Edwards-Trinity Plateau	82		
Gulf Coast	121		
Hueco Mesilla Bolson	198		
Ogallala	81		
Seymour	13		
Trinity	79	137	287

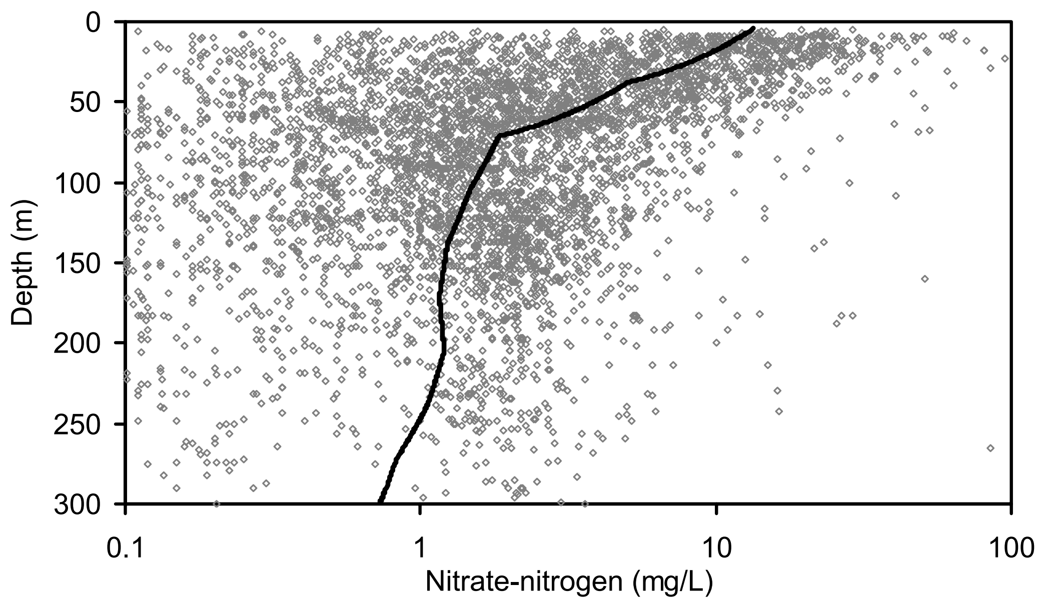


Figure 4: Relationship between groundwater nitrate concentrations and well depth. Line generated using LOWESS smoothing with f=0.2.

There is no obvious trend in nitrate concentrations over time in many of the major aquifers. A preliminary assessment of temporal trends was conducted by evaluating median nitrate concentrations for each decade since 1940 to present in counties that had high nitrate concentrations in four of the major aquifers (Table 4). Although the data do not indicate any obvious trends, the number of samples for each county was quite variable and may affect the analysis.

Table 4. Median nitrate concentrations and associated number of samples in parenthesis for counties with high nitrate concentrations in four of the major aquifers.

Major Aquifer	*Time*	*Median nitrate (mg/L) (no. of samples)*			
Gulf Coast		Duval	Hidalgo	Starr	
	1940-1949	3.95 (4)	1.14 (168)	8.35 (18)	
	1950-1959		0.51 (84)	4.06 (60)	
	1960-1969	4.06 (69)	4.06 (11)	2.37 (2)	
	1970-1979	4.97 (109)	0.28 (18)	6.09 (6)	
	1980-1989	5.61 (39)	0.62 (40)	4.69 (13)	
	1990-1999	4.96 (44)	0.96 (51)	7.8 (42)	
	2000+	5.25 (24)	0.8 (17)	9.64 (9)	
Ogallala		Dawson	Gaines	Lynn	
	1930-1939	4.51 (39)			
	1940-1949		1.25 (12)	2.93 (70)	
	1950-1959		1.06(2)	4.97 (4)	
	1960-1969	2.11 (4)	1.13 (766)	4.06 (27)	
	1970-1979	2.93 (16)	2.26 (37)	3.43 (27)	
	1980-1989	7.27 (30)	3.19 (30)	9.33 (53)	
	1990-1999	8.1 (35)	3.55 (77)	10.69 (135)	
	2000+	13.09 (17)	4.32 (32)	14.19 (15)	
Seymour		Baylor	Haskell	Knox	Wilbarger
	1940-1949	10.5 (4)	17.16 (39)	15.12 (6)	21.11 (16)
	1950-1959	7 (31)	12.87 (22)	8.8 (68)	3.16 (14)
	1960-1969	8.01 (121)	12.64 (72)	9.59 (89)	7.9 (57)
	1970-1979	10.77 (48)	13.77 (512)	12.19 (516)	9.03 (377)
	1980-1989	12.91(13)	17.2 (26)	15.06 (27)	15.52 (24)
	1990-1999	18.72 (8)	15.89 (23)	15.73 (83)	11.99 (22)
	2000+		14.24 (10)	19.39 (9)	3.82 (9)
Trinity		Comanche	Eastland	Erath	
	1940-1949	1.25 (4)	4.85 (8)	0.45 (16)	
	1950-1959	2.93 (7)	4.85 (12)	0.55 (52)	
	1960-1969	1.2 (85)	2.71 (37)	0.68 (35)	
	1970-1979	1.69 (241)	2.93 (177)	0.45 (87)	
	1980-1989	4.63 (35)	4.35 (29)	0.47 (17)	
	1990-1999	2.88 (48)	4.14 (25)	0.57 (132)	

Description of Potential Explanatory Variables

A total of 18 potential explanatory variables for nitrate contamination were examined in this study (Table 5). The variables can be broadly divided into three categories; sources of nitrogen, aquifer susceptibility, and other. Potential natural nitrogen sources include atmospheric deposition while potential anthropogenic nitrogen sources include inorganic and organic fertilizers, CAFOs, sewage sludge application locations, and residential septic tank or sewer system leakage. As no state-wide database of septic and sewer systems is currently available, the percentages of low and high density residential land use and population density were used as proxies for these in the analysis. Aquifer susceptibility variables include land surface slope, percent well drained soils, depth to the seasonally high water table, and percent clay content and organic matter content in soils.

Table 5. Potential explanatory variables included in logistic regression analysis.

Variable	*Median*	*Minimum*	*Maximum*	*Interquartile Range*	*Number*
Nitrogen Sources					
Precipitation, *mm/yr*	656	272	1449	334	734
Distance to CAFO location, *km*	21.4	0.1	100.0	58.6	734
Distance to sludge application location, *km*	68.3	1.6	336.5	74.9	734
NADP nitrate-nitrogen deposition, *kg/ha*	1.1	0.6	2.2	0.5	734
Inorganic fertilizer nitrate, *kg/ha*	7.0	0.0	85.2	18.6	734
Organic fertilizer (manure) nitrate, *kg/ha*	3.0	0.0	92.3	7.3	734
Total fertilizer nitrate, *kg/ha*	12.1	0.6	143.6	25.4	734
Low density residential land use within 2000 m, *%*	0	0	36.6	0.02	734
High density residential land use within 2000 m, *%*	0	0	22.4	0.00	734
Agricultural land use within 2000 m, *%*	48.1	0.0	100.0	62.7	734
Population density, *people/km^2*	1.4	0.0	73.6	2.2	734
Aquifer Susceptibility					
Average land surface slope, *%*	1.97	0.50	16.15	2.16	734
Well drained soils (A, B), *%*	46	0	100	87	734
Depth to seasonally high water table, *m*	1.83	0	1.83	0.28	734
Average soil clay content, *%*	28	5	64	10	734
Average soil organic matter content, *%*	0.52	0.04	2.96	0.44	734
Average soil available water content, *%*	13.7	4.7	17.6	14.5	734
Other					
Total dissolved solids, *mg/L*	663	16	10493	755	725

The climate in Texas ranges from semiarid to arid in the west to humid in the east. Long-term average annual precipitation ranges from 15 cm/yr in west Texas to 150 cm/yr in east Texas (Fig. 5). The precipitation bands are generally north south. Irrigation occurs predominantly in the High Plains aquifer. Some parts of the Seymour, southern Carrizo-Wilcox, and parts of the Gulf Coast aquifers are also heavily irrigated (Fig. 6).

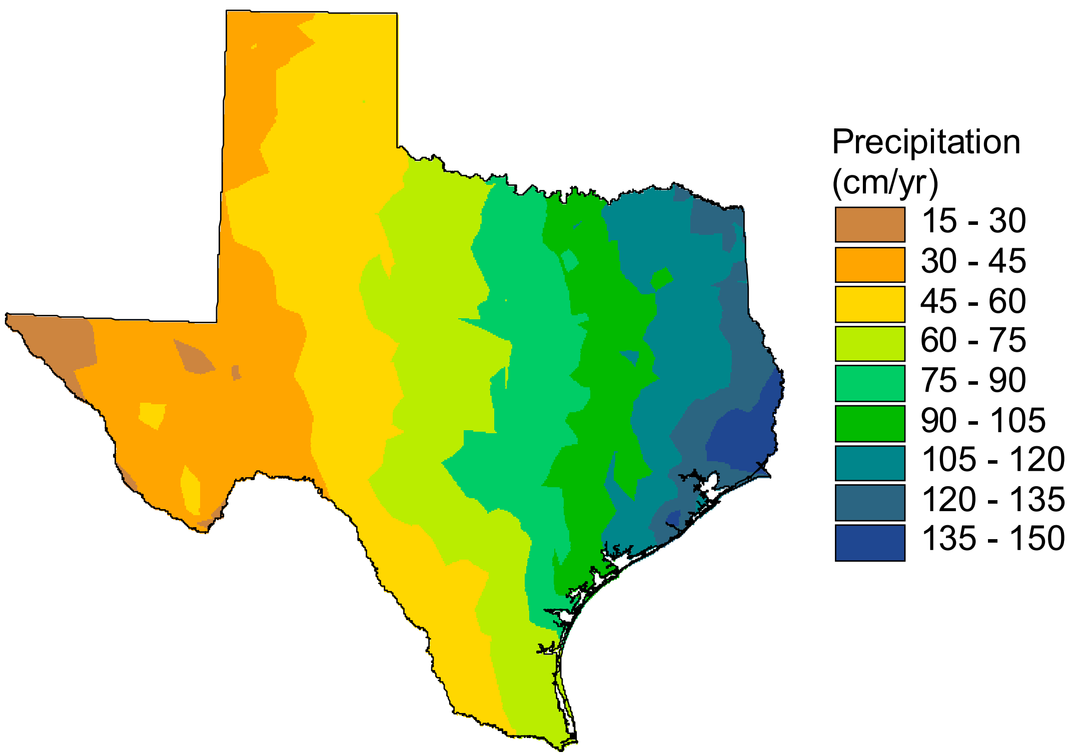


Figure 5. Long-term (1961-1990) average annual precipitation.

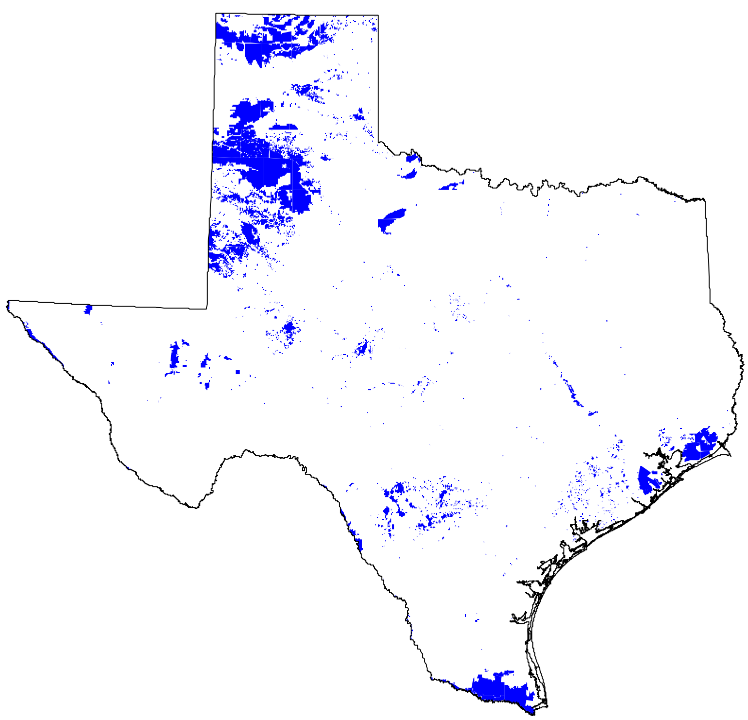


Figure 6. Distribution of irrigated land in Texas (coverage provided by TWDB).

Nitrate loading includes atmospheric deposition (Fig. 7), inorganic and organic fertilizers (Figs. 8, 9), CAFOs, sewage sludge application sites (Fig. 10), and leaking sewer and septic systems. Atmospheric wet deposition of nitrate ranged from 0.4 to 2.2 kg/ha (Fig. 7). The increasing deposition of nitrate from west to east generally follows the trends in increasing precipitation. Deposition includes nitrate concentrations in precipitation, which are generally fairly uniform (0.14 - 0.27 mg/L nitrate for 2000) times annual precipitation rate; therefore, the increasing nitrate deposition from west to east Texas generally reflects increasing precipitation.

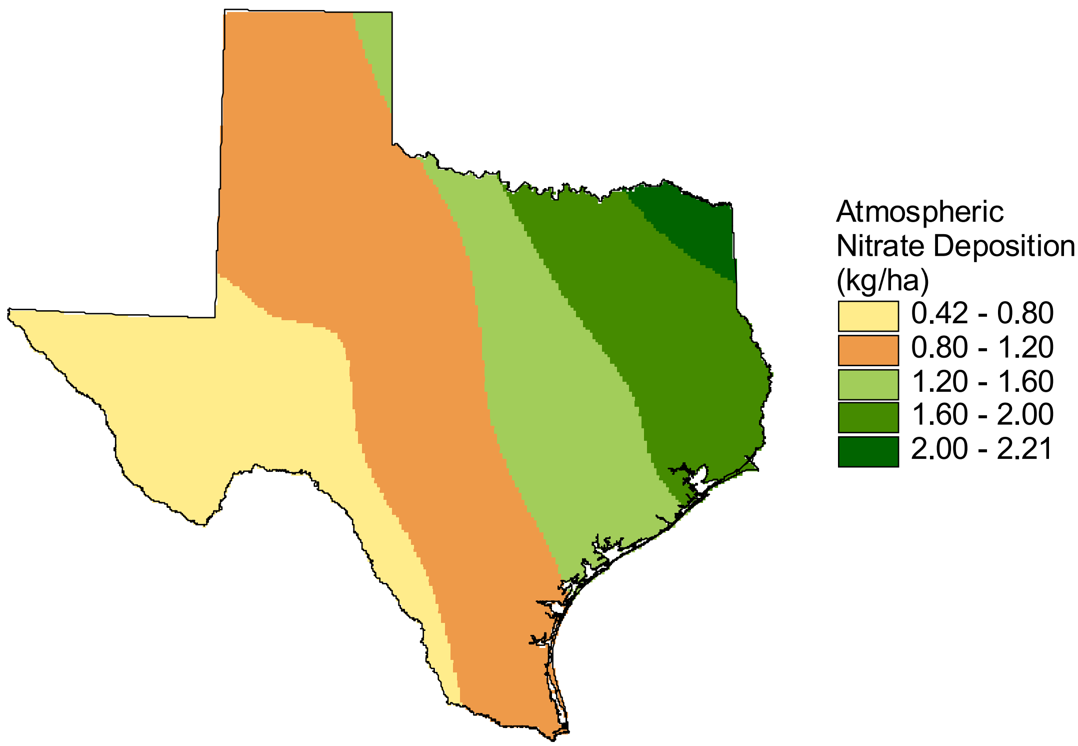


Figure 7. Average annual atmospheric nitrate deposition based on NADP data.

Inorganic fertilizer loading ranges from 0 throughout much of the southwestern extent of the state to values of 20 to 40 kg/ha in many regions in the High Plains and some parts of central Texas (Fig. 8). The fertilizer loading values were estimated from county fertilizer sales; therefore, the patterns often reflect those of the counties. Nitrogen loading from organic fertilizer or manure was generally lower than that of inorganic fertilizer but the regions where it is applied are similar in extent. High loading of organic fertilizer is found in the High Plains, north east Texas, east Texas, and central Texas. Organic fertilizer loading generally correlates with the distribution of CAFOs. The highest concentration of CAFOs is in the central and southern High Plains, the outcrop area of the Trinity aquifer, and east Texas. CAFO types range from cattle feedlots and hogs in the High Plains, dairy in the outcrop of the Trinity aquifer and predominantly poultry in east Texas. Most of the sewage sludge application sites from waste water and septic tanks are located in the eastern half of the state.

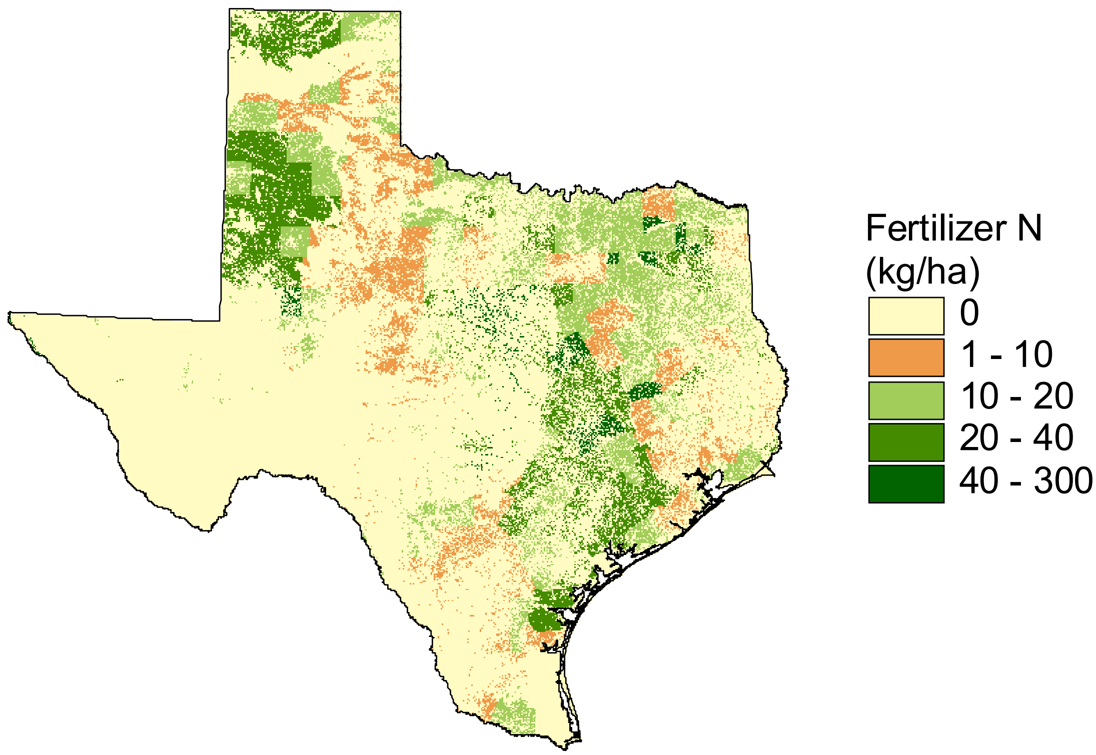


Figure 8. Spatial distribution of inorganic nitrogen fertilizer application.

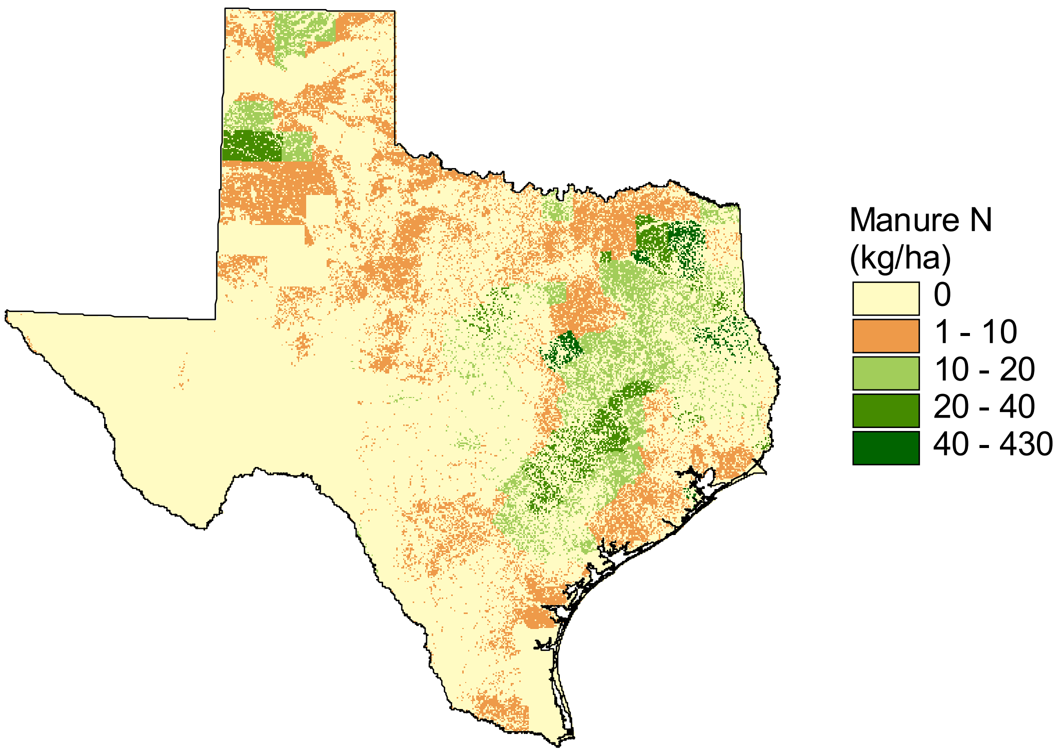


Figure 9. Spatial distribution of organic fertilizer (manure) application.

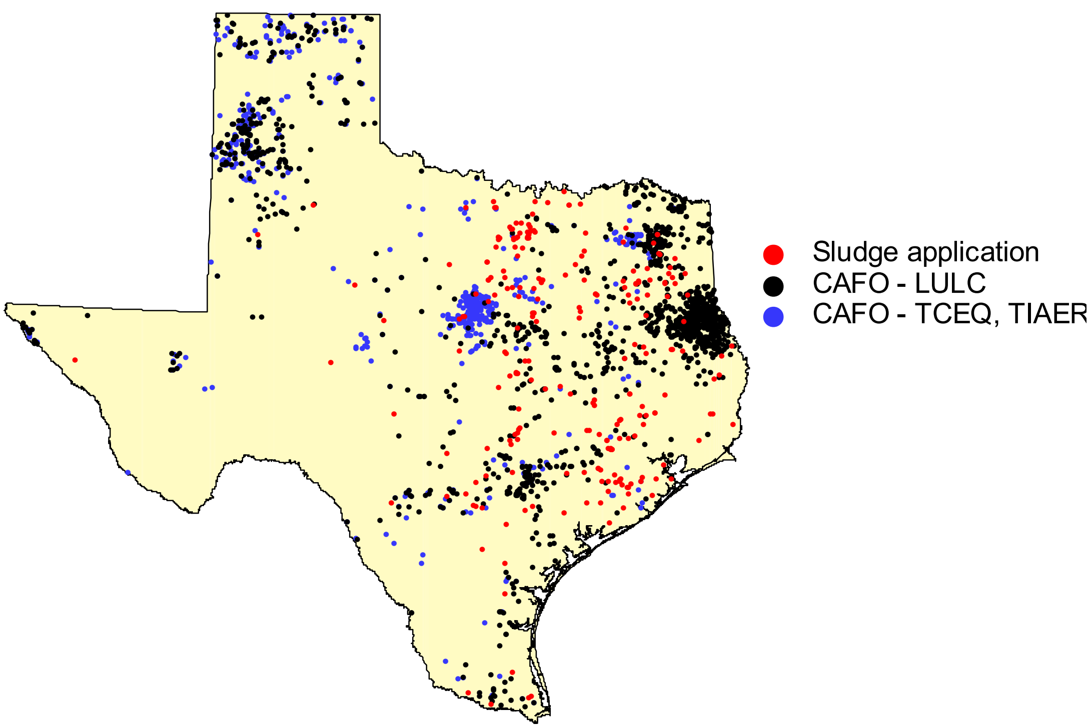


Figure 10. Distribution of concentrated animal feeding operations (CAFOs) based on data from TCEQ, TIAER, and USGS and permitted sludge application based on data from TCEQ.

The land cover map of Texas indicates that agriculture is focused in the High Plains and Rolling Plains and parts of central Texas and the Gulf Coast (Fig. 11). West Texas is dominated by shrubland with scattered grasslands which extend into the Edwards Trinity aquifer region. East Texas is dominated by forested lands. Major urban regions are located in Dallas, Austin, San Antonio, and Houston.

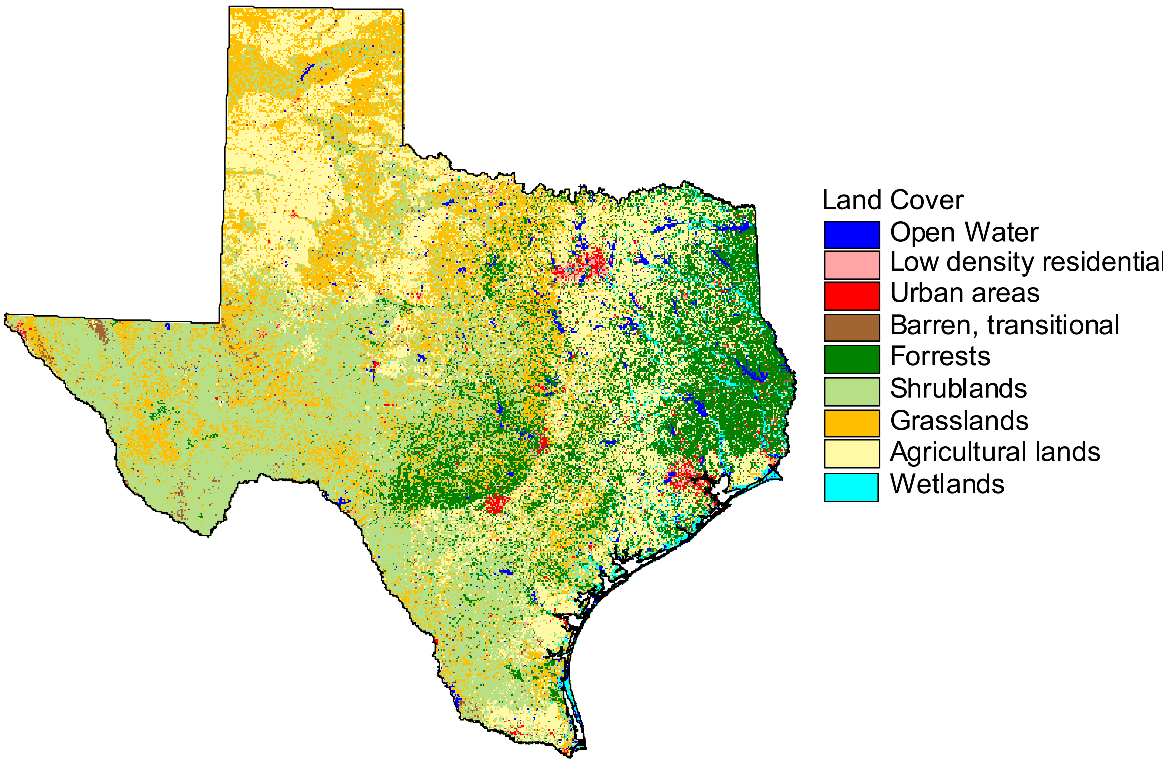


Figure 11. Distribution of land use based on National Land Cover Data.

Potential explanatory variables obtained form the STATGO database include land surface slope, percent well drained soils, depth to seasonally high water table in the upper 2 m zone, percent clay content, organic matter, and available water content. The percent well drained soils include hydrologic groups A and B from the STATSGO database. Well drained soils occur primarily in the High Plains (80 – 100%) and also in the southwestern Gulf Coast (Fig. 12). A map of average clay content in the upper 1.5 to 2.0 (Fig. 13) shows some general trends: low clay content in west Texas (Trans Pecos and Cenozoic Pecos Alluvium regions), high clay content in the central High Plains decreasing in the southern High Plains, generally high clay content in central Texas, low clay content in east Texas, high clay content in the central and northern portions of the Gulf Coast and low clay content in the southwestern Gulf Coast. The trends in clay content generally follow the underlying geology. Soil organic matter ranges from 0.03 to 3.00 percent and generally parallels the map of clay content (Fig. 14).

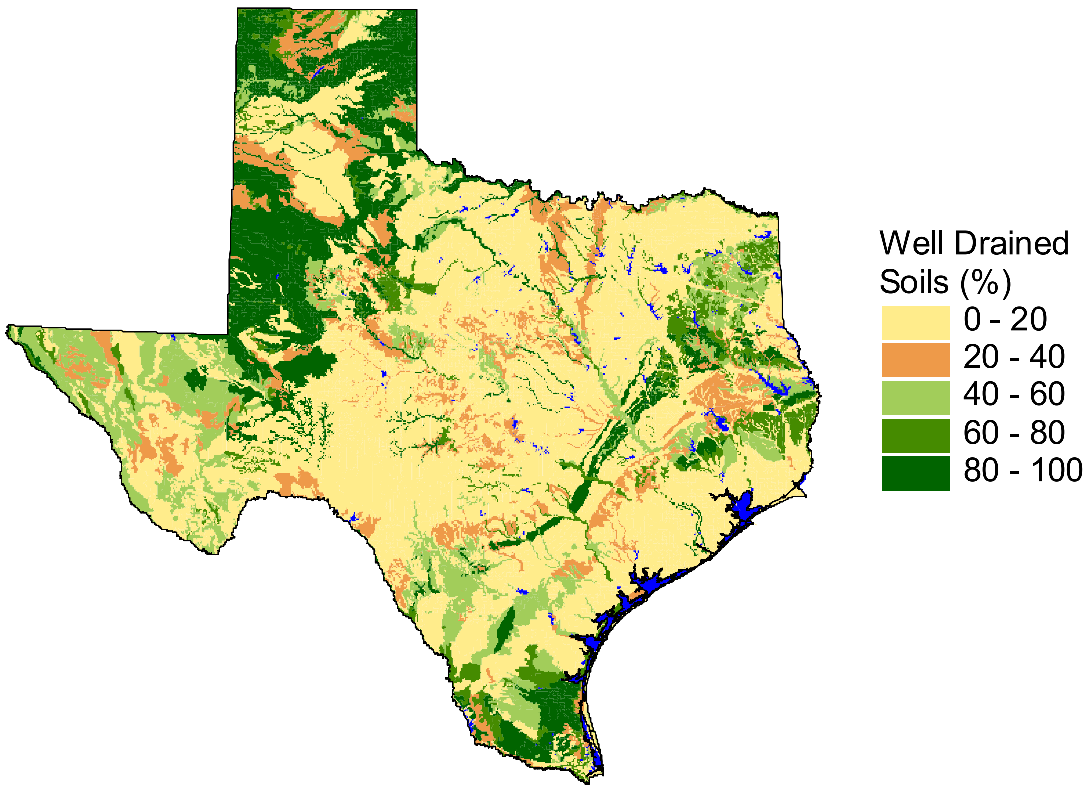


Figure 12. Percentage of well drained soils (A, B) derived from STATSGO database.

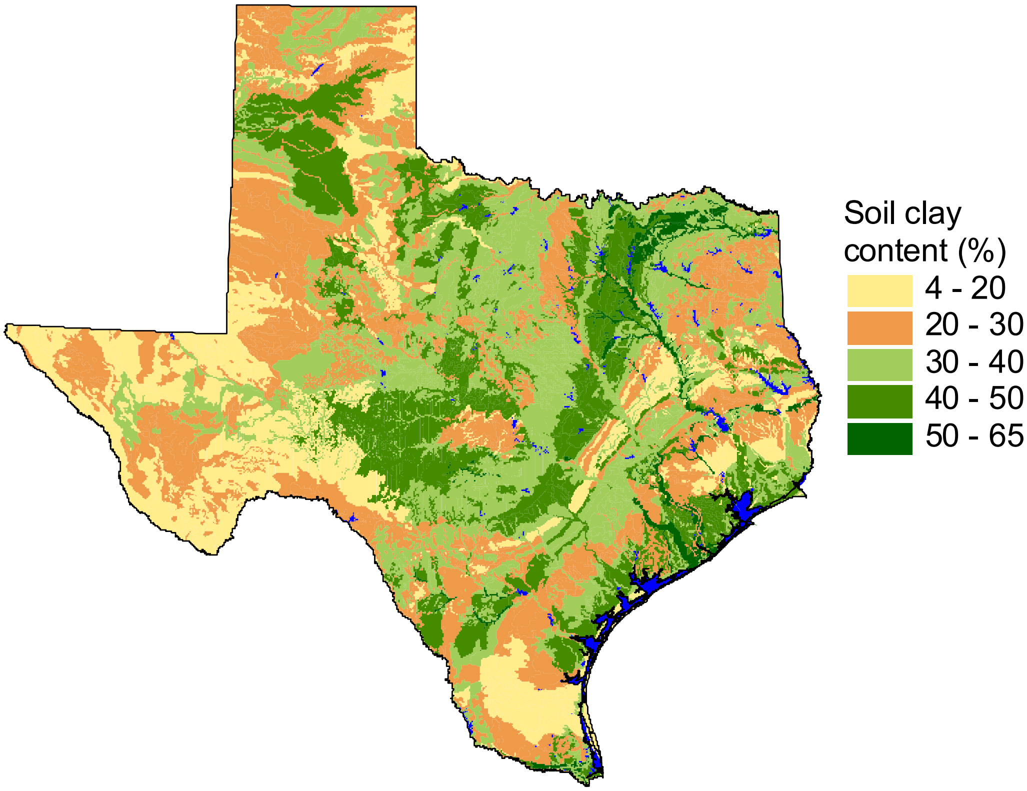


Figure 13. Average soil profile clay content derived from STATSGO database.

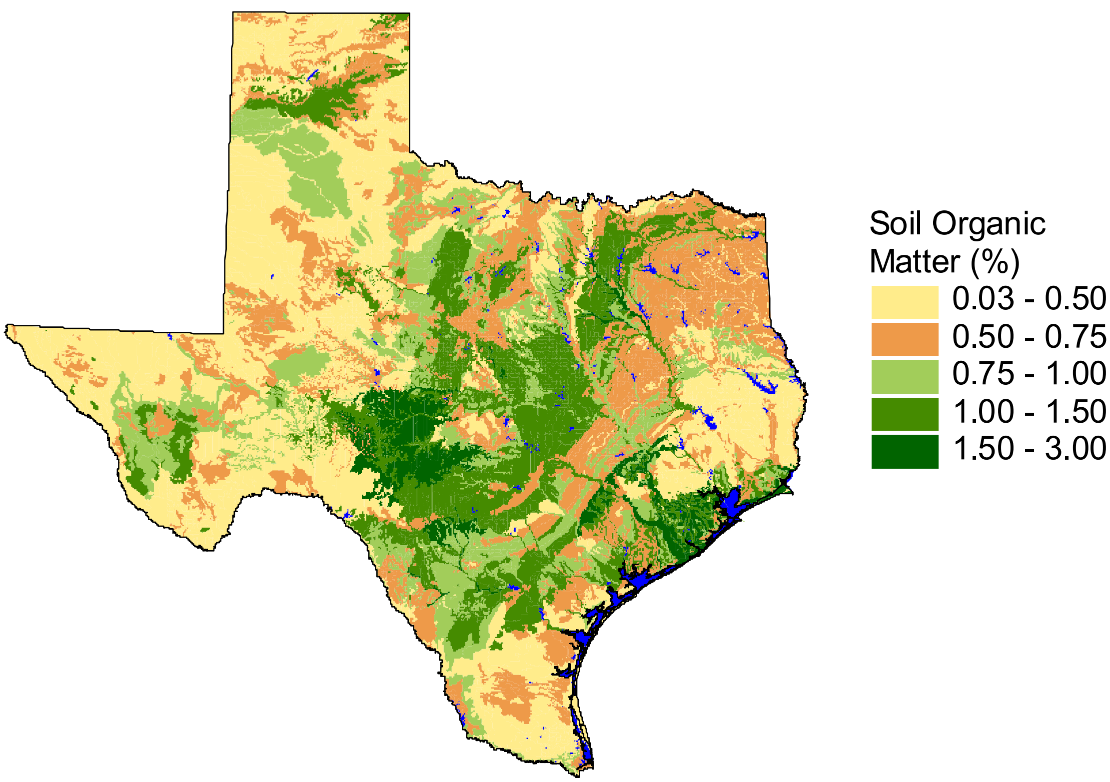


Figure 14. Average soil profile organic matter derived from STATSGO database.

Qualitative comparison of groundwater nitrate concentration data (Fig. 3) and the various potential explanatory variables reveals several trends. High nitrate concentrations in the Southern High Plains correspond to fairly high inorganic nitrogen fertilizer applications, low organic fertilizer applications, low CAFO concentration, well drained soils (80 – 100 %), low clay content, and low organic matter content. These relationships suggest that the nitrogen loading may not be particularly high in this region but that the aquifer susceptibility to contamination is high because of the well drained soils and low clay content. Another area of high nitrate concentrations is the Seymour aquifer. Nitrogen loading in this region is not obviously high: low to moderate inorganic fertilizer application, low organic fertilizer application, and very few CAFOs. Aquifer susceptibility to contamination is also variable: moderate to well drained soils, low to moderate clay content, and low organic matter content. Previous studies indicated that the source of high nitrate in this region is natural resulting from nitrogen fixation by mesquite and other plants being released during cultivation and aeration (Bartolino, 1994). Therefore, the most obvious relationship with high nitrate concentrations may be with agricultural land. Increasing efficiency of irrigation may result in evapoconcentration and increasing concentration of drainage water below the root zone. Another area of high nitrate concentrations is in the outcrop area of the Trinity aquifer, e.g. Erath, Comanche, and Eastland counties. This is an area of dense CAFOs for the dairy industry. The CAFOs are mostly concentrated in Erath county; however, nitrate concentrations in groundwater in this county are lower than those in Comanche county. Comparison of data between these two counties indicates that wells in Comanche county are shallower and are mostly domestic wells whereas many of those in Erath county are deeper and are used for public water supply. These factors may account for some of the differences in nitrate concentrations. The region of high nitrate concentrations in the southern Gulf Coast is not obviously associated with high nitrogen loading: inorganic and organic fertilizer loading is low to moderate and the CAFO density is not very high. However, aquifer susceptibility to contamination may be high because the percent of well drained soils is high and percent clay content and organic matter is low. This qualitative evaluation of nitrate concentrations relative to nitrogen loading and aquifer susceptibility to contamination is a useful prerequisite to formal statistical analysis to provide insights into controls on nitrate contamination. The analysis suggests that there is no single factor that can explain high nitrate concentrations in the various aquifers and controls on nitrate contamination can vary from nitrogen loading/aquifer susceptibility to a combination of both.

Nitrate Logistic Regression Model

Most of the potential explanatory variables were significantly related to the outcome variable during the univariate analysis (Table 6). Variables not significant to $p \leq 0.05$ included manure nitrate loading, low- and high density residential land use within 2000 m, and average soil available water content.

Table 6. Results of univariate statistical analysis to evaluate the significance of each explanatory variable in explaining nitrate concentrations in groundwater. (n is the number of observations for the calibration data set).

Variable	*Coefficient*	*Wald p*	*n*
Nitrogen Sources			
Precipitation, *mm/yr*	-0.00359	<0.0001	734
Distance to CAFO location, *km*	0.0263	<0.0001	734
Distance to sludge spreading location, *km*	0.0044	0.0019	734
NADP nitrate-nitrogen deposition, *kg/ha*	-2.82	<0.0001	734
Fertilizer nitrate, *kg/ha*	0.00058	<0.0001	734
Manure nitrate, *kg/ha*	-0.00021	0.1300	734
Total nitrate, *kg/ha*	0.000249	0.0003	734
Low density residential land use within 2000 m, *%*	0.0191	0.3413	734
High density residential land use within 2000 m, *%*	-0.0021	0.9575	734
Agricultural land use within 2000 m, *%*	0.0302	<0.0001	734
Population density, *people/km²*	-0.08	0.0001	734
Aquifer Susceptibility			
Average land surface slope, *%*	-0.35	<0.0001	734
Well drained soils, *%*	0.017	<0.0001	734
Depth to seasonally high water table, *m*	2.73	<0.0001	734
Average soil clay content, *%*	-0.03	0.0005	734
Average soil organic matter content, *%*	-2.27	<0.0001	734
Average soil available water content, *%*	0.07	0.0725	734
Other			
Total dissolved solids, *mg/L*	0.00060	<0.0001	725

Multivariate models were then developed using both forward (stepwise) and backward elimination techniques. Forward modeling is performed by sequentially adding variables in a stepwise fashion, starting with the most significant variable. At each step, the significances of the remaining variables are calculated and the most significant remaining variable is then included in the next model. This process continues until all of the variables have been sequentially examined in relation to the (growing) combined model. Also, during the process, a pre-specified threshold significance level is used to determine if a variable can be included in the model. A threshold (model entry) value of 0.2 was used in this analysis. Backward elimination modeling is essentially the reverse of the forward process, where all of the variables are initially included and the least significant variable is eliminated sequentially. Again, a threshold (model exit) elimination value of 0.2 was used in this analysis.

The best multivariate model resulted from the variables for (1) the percentage of agricultural lands with a 2000 m radius of the well, (2) annual average precipitation, (3) the average percentage of soil organic matter, and (4) the percentage of low density residential land use within a 2000 m radius of the well (Table 7). The statistical significance of the Wald p value is high for agricultural land and precipitation and lower for percent organic matter and low density residential land use. Both of the land use variables have positive slope coefficients, indicating that increasing values for these variables lead to higher probability of nitrate contamination in wells ≤ 30 m deep. Conversely, increasing precipitation and soil organic matter content values result in lower probability of elevated nitrate concentrations.

The relationship between agricultural land use and elevated nitrate concentrations may generally reflect the impact of cultivation on nitrate contamination (e.g. Seymour aquifer) in addition to associated inorganic and organic fertilizer loading associated with agricultural land. The inverse relationship between average annual precipitation and groundwater nitrate concentrations is similar to that found by Evans and Maidment (1995) and may reflect the impact of high recharge and dilution in humid regions and possibly evapoconcentration in the shallow subsurface in semiarid and arid regions resulting in increased nitrogen loading. The inverse relationship between soil organic matter content and elevated nitrate concentrations may reflect denitrification associated with high organic matter content and/or an embedded effect of percent well drained soils on elevated nitrate concentrations because percent organic matter is generally correlated with clay content and negatively correlated with percent well drained soils. The model accurately characterizes elevated nitrate concentrations in shallow wells (≤ 30 m deep) at the state-wide scale.

Table 7. Results of the multivariate logistic regression model.

Variable	*Coefficient*	*Wald p value*
Intercept	1.4391	<0.0001
Agricultural land within 2000 m, %	0.0305	<0.0001
Precipitation, mm/yr	-0.0326	<0.0001
Average soil organic matter content, %	-0.7201	0.0173
Low density residential land within 2000 m, %	0.0475	0.0515

The Hosmer-Lemeshow (HL) goodness-of-fit test evaluates the overall model fit by comparing average predicted versus observed probabilities for deciles of risk. The HL p-value of 0.217 indicates that the fitted model is generally acceptable. The coefficient of determination between observed and predicted probabilities is high (R^2 = 0.977) (Figure 15).

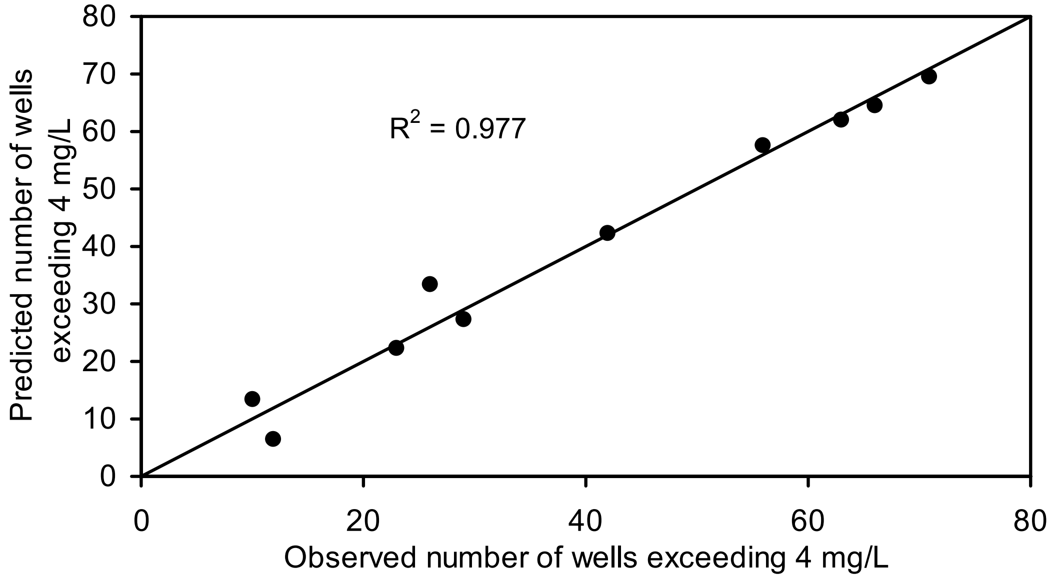


Figure 15: Predicted versus observed number of wells ≤ 30 m deep with nitrate concentrations exceeding 4 mg/L for deciles of risk using the model data set (n=734).

The logistic regression model parameters were used to calculate the probability of nitrate exceeding 4 mg/L for the validation data set. The fit of the model was evaluated by comparing average predicted and observed probabilities for deciles of risk (Figure 16). The coefficient of determination (R^2 = 0.959) indicates that the model predicts the observed probabilities of nitrate exceeding 4 mg/L very well.

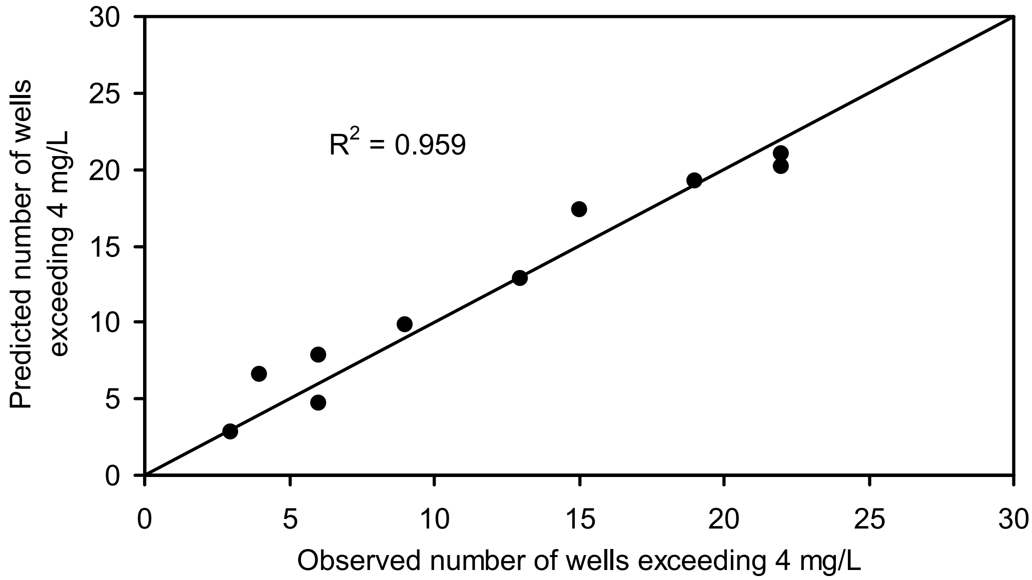


Figure 16: Predicted versus observed number of wells ≤ 30 m deep with nitrate concentrations exceeding 4 mg/L for deciles of risk using the validation data set (n=235).

The ability of the model to predict nitrate concentrations varied for different aquifers (Fig. 17). The model underpredicted observed exceedances in some aquifers (Carrizo-Wilcox, Gulf Coast, and Seymour aquifers) whereas the model overpredicted exceedances in the High Plains, Cenozoic Pecos Alluvium, and much of the Trinity aquifers. The number of sampled wells is limited in some aquifers (Carrizo Wilcox, Cenozoic Pecos Alluvium and Gulf Coast aquifers) and may affect the analysis.

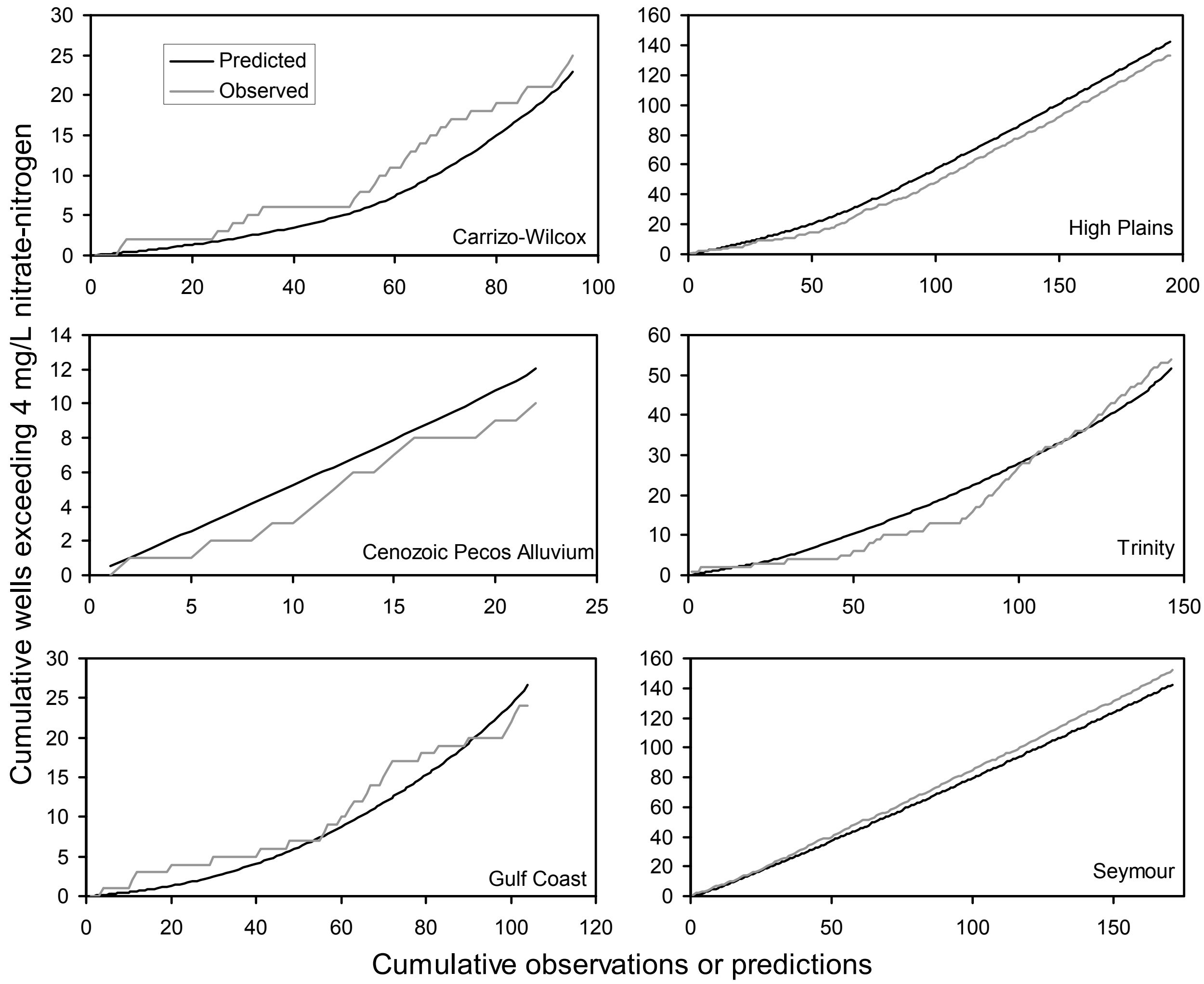


Figure 17. Cumulative number of sampled wells exceeding 4 mg/L nitrate relative to cumulative number of observations or predictions.

STUDY LIMITATIONS

It is important to recognize the limitations of the various data sources and analysis to better understand the findings from this analysis. Much of the analysis focused on evaluating impacts of nitrogen loading and aquifer susceptibility on the distribution of nitrate in groundwater. The dataset on groundwater nitrate concentrations covered the 1980 – 2002 time period. The use of data for such an extended time period could potentially introduce effects of temporal variability in nitrate on the spatial analysis of nitrate in this study. However, preliminary evaluation indicated that there were no obvious temporal trends in the nitrate data (Table 4). Nitrogen loading data from fertilizer was restricted to county fertilizer sales records which may not be highly accurate and does not provide the detailed spatial coverage of nitrogen loading for this study. Nitrogen fertilizer application may be quite different for irrigated and nonirrigated agriculture and it would be useful to include detailed information on this in the analysis. Nitrogen loading from manure is calculated using a number of assumptions including counts for different types of animals and per animal production of manure and losses due to volatilization. The reliability of the manure estimates depends on the validity of the various assumptions that were used in developing these statistics. The inorganic and organic fertilizer loading values were based on data from 1997 and 1998; however, the groundwater nitrate data cover the period from 1980 – 2002. It would be interesting to evaluate temporal variability in fertilizer loading during that time and incorporate this information into the analysis. No information is available on the distribution of septic tanks, another potential source of nitrate. Using low density residential setting from the NLCD data may or may not serve as an appropriate proxy for the distribution of septic tanks. In addition, information on the location of sewer networks is also lacking. Accurate information on the distribution of septic tanks and sewers would allow a more thorough evaluation on their potential contribution to nitrate contamination.

Information on CAFOs was restricted to permitted CAFOs and available data in a 1994 land use/land cover dataset. Accurate location information on all CAFOs, regardless of size, would be very valuable in evaluating potential relationships with nitrate contamination. In addition, information on sludge amounts and application rates adjacent to CAFOs and water water treatment sludge application sites would be very useful in evaluating potential nitrogen loading form these sites to underlying aquifers. Monitoring temporal variability in nitrate transport beneath these sites could help to develop optimal sludge application rates and amounts to minimize aquifer contamination. Best management practices could be developed based on field monitoring of nitrate transport.

In addition to evaluation of nitrogen loading, much of the analysis focused on evaluating aquifer susceptibility to contamination. Data sources for assessing aquifer susceptibility focused on the attributes of the soil profile provided by the STATSGO database. The applicability of these data in areas of thick unsaturated zones is questionable. It would be very useful if information on these types of parameters, such as drainage characteristics, percent organic matter, percent clay could be extended from the soil zone to underlying aquifers to better understand aquifer susceptibility issues.

Because many of the aquifers in Texas are overlain by fairly thick unsaturated zones, particularly in the High Plains, it is very important to characterize the distribution of nitrates in the unsaturated zone for different climate conditions, soils, vegetation coverage, and land use.

Vertical profiles of nitrate in the unsaturated zone would allow us to better predict future concentrations in underlying aquifers and could be used in addition to surface loading data.

While this analysis focused on nitrogen loading and aquifer susceptibility issues, other factors, such as recharge, dilution, evapoconcentration, and denitrification may also play an important role in controlling the distribution of nitrate in groundwater. Generally low nitrate concentrations in east Texas may reflect higher recharge and associated dilution in this humid setting, or denitrification. The density of CAFOs in this region is fairly high; however, most of the CAFOs are poultry and may have lower nitrogen outputs than other CAFOs. Irrigation systems in the 1960s and 1970s, such as furrow irrigation, were fairly inefficient with up to 50 percent of the water draining below the root zone. In the last decade, much more efficient irrigation systems have been developed and are being used, for example the low energy precision application (LEPA) system are considered to be 95 – 98% efficient with only 2 to 5% of the water returning to the aquifer. This increased efficiency results in much more evapoconcentration of nutrients near the land surface and may ultimately result in higher nitrate concentrations in aquifers if the nitrate is not taken up by crops. Monitoring nitrate concentrations in the unsaturated zone is critical for evaluating the potential impacts of these land management practices on potential contamination of underlying aquifers. Denitrification is a very important process for reducing nitrate loading to aquifers and has been documented in unsaturated zones beneath playas near Amarillo, Texas (Fryar et al., 2000). Large reductions in nitrate concentrations beneath and adjacent to CAFOs has also been attributed to denitrification (Clark, 1975; Stewart et al., 1994; Daniel, 1997). However, evaluation of this process requires detailed field studies and sampling for nitrogen gas, nitrogen isotopes, and other parameters. Regionalizing the results from point based measurements would require evaluation of the applicability of the point measurements beyond the local scale.

FUTURE STUDIES

The reconnaissance study described in this report focused on groundwater nitrate data available from the ambient groundwater monitoring program conducted by the Texas Water Development Board. Although this database includes public water supply wells, an additional database focused solely on public water supply systems is available through the Texas Commission on Environmental Quality and should also be evaluated using similar approaches. The work described in this study focused on wells in the upper 30 m; however, this analysis should be extended to wells of greater depth. Future studies that could improve the quality of inputs to the statistical analyses should also be done. Improving the accuracy of input parameters should increase the reliability of the model predictions.

The GIS and statistical analysis discussed in this report should be linked to focused field studies that assess different aspects of nitrate transport and other processes. An understanding of nitrate concentrations in the unsaturated zone would greatly improve our understanding of nitrate inputs to aquifers. Limited studies of nitrate concentrations in the unsaturated zone have been conducted in the High Plains (Bruce et al., 2000; Scanlon et al., 2003; Fryar et al, 2000). Weighing lysimeter drainage from USDA Agricultural Research Services in Bushland and Uvalde provide another potential source of nitrate concentrations in water draining below the root zone. Much more extensive evaluation of nitrate in unsaturated systems should be conducted to understand relationships between nitrate concentrations and land use, soils, climate, and other factors. Monitoring temporal variability in nitrate concentrations in unsaturated systems would allow us to understand plant uptake better, and assess processes such as evapoconcentration that could impact long-term nitrate loading to aquifers. These types of measurement and monitoring programs are an essential component of precision agriculture to asses nutrient needs by crops and impacts of agriculture on nitrate loading. The GIS and field studies should also be supplemented by physical flow and transport modeling to assess various processes that could potentially impact nitrate concentrations, such as temporal variability in climate, nitrate loading, plant uptake, and recharge. Various levels of modeling could be conducted ranging from simple 1 dimensional models to complex 3 dimensional models.

To assess the potential impacts of different processes such as recharge, dilution, and denitrification, focused field studies should be conducted to evaluate these processes. In addition to conducting these studies in irrigated and nonirrigated agricultural settings, these studies could also be conducted in areas where CAFO and waste water treatment plant sludge is being applied to understand the fate of nitrate in these regions. Areas with differing amounts of organic matter in soils should also be evaluated. These studies should include nitrogen gas analyses, nitrogen isotope studies, and modeling analyses.

Although this study focused on groundwater nitrate, future studies should evaluate linkages between groundwater nitrate distribution and nutrient loading in surface water bodies that could impact dissolved oxygen and also result in eutrophication. The Total Maximum Daily Load (TMDL) program focuses on nutrient and dissolved oxygen issues in surface water bodies but generally ignores potential inputs from groundwater. Because many of the problems arise during periods of low flow, groundwater input may be significant and should be assessed. Nitrogen loading in bays and estuaries is also a critical issue and inputs from groundwater to these systems should also be addressed. Many studies have indicated that riparian zones can greatly reduce nitrate loading from surface runoff and groundwater inflow to streams (Lowrance et al., 1984; Hill, 1996; NRC, 2002; Simpkins, 2002). The distribution of these riparian zones in Texas should be delineated. Riparian zones can also be constructed and managed for this purpose.

Although this study included a preliminary assessment of temporal trends, much more detailed evaluation of temporal trends in nitrate should be conducted. Understanding the impacts of current land use practices on nitrate input and characterizing nitrate concentrations in unsaturated systems will allow us to better predict future concentrations in aquifers and develop sustainable land use practices that minimize further increases and potentially decrease nitrate concentrations in groundwater.

CONCLUSIONS

Nitrate is the most pervasive contaminant in groundwater in Texas. The percent of wells exceeding the maximum contaminant level (MCL) of 10 mg/L nitrate as nitrogen ranged from 1% in the Edwards (BFZ), Hueco Mesilla Bolson, and Carrizo Wilcox aquifers to 66% in the Seymour aquifer. Nitrate contamination was greatest in the Seymour, Southern High Plains, and Southern Gulf Coast aquifers. Nitrate levels were greater in unconfined aquifers relative to confined aquifers. Nitrate concentrations decreased with depth with a distinct break in the LOWESS curve at 74 m depth. The reduction in nitrate concentrations with depth may reflect stratification in water chemistry in aquifers.

Multivariate logistic regression was used to determine controls on the spatial distribution of nitrate concentrations in major porous media aquifers by relating the probability of elevated nitrate concentrations (≥ 4mg/L nitrate) to nitrogen loading and aquifer susceptibility parameters. Nitrogen loading was represented by atmospheric deposition, inorganic and organic fertilizers, CAFO and sludge application locations, proxies for sewage and septic input, and precipitation and irrigation in GIS. Aquifer susceptibility was represented by percent well drained soils, percent clay content, organic matter content, and available water content. The final logistic regression model included precipitation, percent agricultural land, low density residential land, and soil organic matter. Observed and predicted probabilities of elevated nitrate concentrations were highly correlated in calibration and validation data sets (R^2, 0.96; 0.98). The inverse relationship between precipitation and nitrate concentration may be related to dilution in high precipitation areas and possibly evapoconcentration in low precipitation areas. Although nitrate loading is not explicitly represented in the final model, percent agricultural land may be considered a proxy for nitrogen loading from agricultural sources and low density residential land use may be considered a proxy for septic tank effluent. Percent organic matter may reflect the influence of denitrification in some regions. Future studies should include field sampling and analysis to evaluate the influence of different processes such as dilution and denitrification on nitrate concentrations. Such field sampling could serve to ground reference GIS and logistic regression analysis. This reconnaissance study provides valuable insights into controls on the distribution of nitrate contamination in major porous media aquifers in the state.

ACKNOWLEDGMENTS

We would like to thank John Meyer and Greg Rogers from the TCEQ SWAP program for providing GIS coverages of aquifers, precipitation, nitrogen loading, and population statistics. We very much appreciate discussions with Randy Ulery and GIS guy USGS for providing valuable insights into data inputs and analyses.

REFERENCES

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

Barrington, S. F., and R. S. Broughton. 1988. Designing earthen storage facilities for manure. Can. Agric. Eng. **30**:289-292.

Bartolino, J. R. 1994. Nitrate nitrogen in the Seymour Aquifer, Knox County, Texas. GSA Abstracts with Programs, 1994, Spring meeting: p.156.

Battaglin, W. A., and D. A. Goolsby. 1994. Spatial data in GIS format on agricultural chemical use, land use, and cropping practices in the United States. U.S. Geological Survey Water Resources Investigation Report 94-4176, 87 p.

Bruce, B. W., K. F. Dennehy, K. M. Ellett, J. Gurdak, P. B. McMahon, R. C. Reedy, B. R. Scanlon, and M. A. Sophocleous. 2002. Evaluation of recharge fluxes to the High Plains aquifer, Colorado, Kansas, Texas, and Nebraska. GSA Abs. with Programs, Abs. No. 43-1.

Centers for Disease Control and Prevention. 1996. Spontaneous abortions possibly related to ingestion of nitrate-contaminated well water -- LaGrange County, Indiana, 1991-1994. Morbidity and Mortality Weekly Report **45**.

Clark, R. N. 1975. Seepage beneath feedyard runoff catchments. Pages 289-295 *in* Managing Livestock Wastes, Proc. Third Intl. Symposium on Livestock Wastes. Am. Soc. Civil. Engin., St. Joseph, MI.

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

Daniel, J. A. 1997. Effectiveness of animal waste containment in a Texas playa. Environ. Eng. Geosci. **3**.

Evans, T. A., and D. R. Maidment. 1995. A spatial and statistical assessment of the vulnerability of Texas groundwater to nitrate contamination. Center for Research in Water Resources, Technical Report 260, 259 p.

Fryar, A. E., S. A. Macko, W. F. Mullican, III, K. D. Romanak, and P. C. Bennett. 2000. Nitrate reduction during ground-water recharge, Southern High Plains, Texas. J. Contam. Hydrol. **40**:335-363.

Goolsby, D. A., W. A. Battaglin, G. B. Lawrence, R. S. Atrz, B. T. Aulenback, and R. P. Hooper. 1999. Flux and sources of nutrients in the Mississippi-Atchafalaya River Basin: Topic 3. Rept. for Integrated Assessment of Hypoxia in the Gulf of Mexico: NOAA Coastal Decision Analysis Series No. 17, NOAA Coastal Ocean Program, SIlver Spring, MD., 130 p.

Helsel, D. R., and R. M. Hirsch. 1992. Statistical Methods in Water Resources, Amsterdam.

Hill, A. R. 1996. Nitrate removal in stream riparian zones. J. Env. Qual. **25**:743-755.

Hosmer, D. W. J., and S. Lemeshow. 2001. Applied Logistic Regression. John Wiley & Sons, New York.

Kapoor, A., and T. Viraraghavan. 1997. Nitrate removal from drinking water - review. J. Env. Engin. **123**:371-380.

Kleinbaum, D. G. 1994. Logtistic Regression: A Self-Learning Text. Springer-Verlag, New York.

Koterba, M. T., W. S. L. Banks, and R. J. Shedlock. 1993. Pesticides in shallow groundwater in the Delmarva Peninsula. J. Env. Qual. **22**:500-518.

Kreitler, C. W. 1975. Determining the source of nitrate in groundwater by nitrogen isotope studies. Bureau of Economic Geology, Univ. Texas at Austin, Rept. Inv. No. 83, 57 p.

Kreitler, C. W. 1979. Nitrogen-isotope ratio studies of soils and groundwater nitrate from alluvial fan aqufiers in Texas. J. Hydrol. **42**:147-170.

Kreitler, C. W., and L. Browning. 1983. Nitrogen isotope analysis of groundwater nitrate in carbonate aquifers: natural sources versus human pollution. J. of Hydrol. **61**:285-301.

Kreitler, C. W., and D. Jones. 1975. Natural soil nitrate: the cause of the nitrate contamination of groundwater in Runnels County, Texas. Ground Water **13**:53-61.

Lander, C. H., D. Moffitt, and K. Alt. 1998. Nutrients available from livestock manure relative to crop growth requirements. U. S. Department of Agriculture, Natural Resources Conservation Service, Resource Assessment and Strategic Planning Working Paper 98-1. http://www.nhq.nrce.usda.gov/land/pubs/nlweb.html.

Lehman, O. R., and R. N. Clark. 1975. Effect of cattle feedyard runoff on soil infiltration rates. J. Env. Qual. **4**:437-439.

Lehman, O. R., B. A. Stewart, and A. C. Mathers. 1970. Seepage of feedyard runoff water impounded in playas. Texas Agric. Expt station, Texas A&M Univ. **MP-944, p. 1 - 7.**:3-7.

Lowrance, R., R. Todd, J. Fail, O. Hendrickson, R. Leonard, and L. Asmussen. 1984. Riparian forests as nutrient filters in agricultural watersheds. Bioscience **34**:374-377.

McMahon, P. B., K. F. Dennehy, K. M. Ellett, M. A. Sophocleous, R. L. Michel, and D. B. Hurlbut. 2003. Water movement through thick unsaturated zones overlying the central High Plains aquifer, southwestern Kansas, 2000-2001. USGS Water REsour. Inv. Rept. 03-4171.

Nolan, B. T. 2001. Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States. Ground Water **39**:290-299.

Nolan, B. T., K. J. Hitt, and B. C. Ruddy. 2002. Probability of nitrate contamination of recently recharged groundwaters in the conterminous United States. Env. Sci & Technol. **36**:2138-2145.

Nolan, B. T., B. C. Ruddy, K. J. Hitt, and D. R. Helsel. 1997. Risk of nitrate in groundwaters of the United States - A national perspective. Env. Sci. & Technol. **31**:2229-2236.

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

Nolan, B. T., and J. D. Stoner. 2000. Nutrients in groundwaters of the conterminous United States. Environ. Sci. Technol. **34**:1156-1165.

Porter, W. P., J. W. Jaeger, and I. H. Carlson. 1999. Endocrine, immune, and behavioral effects of aldicarb (carbamate), atrazine (triazine) and nitrate (fertilizer) mixtures at groudnwater concentrations. Toxicology and Industrial Health **15**:133-150.

Roswell, J. G., M. H. Miller, and P. H. Groenevelt. 1985. Self-sealing of earthen liquid manure storage ponds: II Rate and mechanism. J. Env. Qual. **14**:539-543.

Rupert, M. G. 1998. Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate (NO2+NO3-N) in groun water in the Idaho part of the upper Snake River Basin. U.S. Geol. Surv. Water Resour. Inv. Rept. 98-4203, 32 p.

Rupert, M. G. 2001. Calibration of the DRASTIC ground water vulnerability mapping method. Ground Water **39**:625-630.

Rupert, M. G. 2003. Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate in ground water in Colorado. U.S. Geol. Surv. Open File Rept. 02-4269.

Scanlon, B. R., R. C. Reedy, and K. E. Keese. 2003. Estimation of groundwater recharge in Texas related to aquifer vulnerability to contamination. Bureau of Economic Geology, Univ. of Texas at Austin, Final Contract Report, 84 p.

Simpkins, W. W., T. R. Wineland, R. J. Andress, D. A. Johnston, and G. C. Caron. 2002. Hydrogeological constraints on riparian buffers for reduction of diffuse pollution: examples from the Bear Creek watershed in Iowa. Water Science and Technology: v. 45, 61-68.

Spalding, R. F., and M. E. Exner. 1993. Occurrence of nitrate in groundwater - a review. J. Env. Qual. **22**:391-402.

Squillace, P. J., J. C. Scott, M. J. Moran, B. T. Nolan, and D. W. Kolpin. 2002. VOCs, pesticides, nitrate, and their mixtures in groundwater used for drinking water in the United States. Environ. Sci. Technol. **36**:1923-1930.

Stewart, B. A., S. J. Smith, A. N. Sharpley, J. W. Naney, T. McDonald, M. G. Hickey, and J. M. Seweeten. 1994. Nitrate and other nutrients associated with playa storage of feedlot wastes. *in* L. V. Urban and A. W. Wyatt, editors. Proc. Playa Basin Symposium, Texas Tech Univ. Lubbock.

Stumm, W., and J. J. Morgan. 1996. Aquatic Chemistry, 3rd edition. Wiley Interscience, New York.

Sweeten, J. M., T. H. Marek, and D. McReynolds. 1995. Groundwater quality near two cattle feedlots in the Texas High Plains: a case study. Appl. Engineer. Agricul. **11**:845-850.

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

U.S. EPA. 1995. Drinking water regulations and health advisories. Office of Water, Washington, D.C.

U.S. Geological Survey. 2000. National land cover dataset. U.S. Geol. Surv. Fact Sheet 108-00, 3 p.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. van Driel. 2001. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engin. and Remote Sensing **67**:650-662.

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 **7**:465-471.

Appendix A. Number of sampled wells for nitrate, number of samples with nitrate concentrations ≤ 4 mg/L and ≥10 mg/L, percent of samples ≥ 10 mg/L, median, minimum, and maximum nitrate concentrations, 10, 25, 75, and 90th percentiles. Major aquifers, subdivided by well categories and summed for all wells in each aquifer.

Carrizo-Wilcox	No. samples	Samples ≤ 4mg/l	Samples ≥10mg/l	% ≥10mg/l	Median	Min	Max	10th %	25th %	75th %	90th %
Commercial	1	1	0	0.00	0.08	0.08	0.08				
Industrial	50	49	0	0.00	<0.01	<0.01	5.92	<0.01	<0.01	0.05	0.28
Domestic	388	370	4	1.03	<0.04	0.00	70.65	<0.01	<0.01	0.20	1.18
Public	687	685	1	0.15	<0.02	0.00	10.59	<0.01	<0.01	0.04	0.10
Irrigation	125	117	6	4.80	<0.03	0.00	27.37	<0.01	<0.01	0.07	0.49
Stock	88	80	3	3.41	<0.04	<0.01	26.60	<0.01	<0.02	0.12	1.98
All Wells	1339	1302	14	1.05	<0.02	0.00	70.65	<0.01	<0.01	0.06	0.35
Unconf.											
Commercial	1	1	0	0.00	0.08	0.08	0.08				
Industrial	29	28	0	0.00	<0.01	<0.01	5.92	<0.01	<0.01	0.05	0.420
Domestic	237	221	2	0.84	<0.05	0.00	70.65	<0.01	<0.01	0.40	2.46
Public	251	250	1	0.40	<0.02	0.00	8.18	<0.01	<0.01	<0.05	0.12
Irrigation	26	22	3	11.54	<0.05	<0.01	15.49	<0.01	<0.01	0.43	8.47
Stock	38	33	1	2.63	<0.05	<0.01	16.09	<0.01	<0.02	0.53	5.34
All Wells	582	555	7	1.20	<0.03	0.00	70.65	<0.01	<0.01	0.13	1.12
Conf.											
Commercial	0	0	0								
Industrial	21	21	0	0.00	<0.02	<0.01	0.50	<0.01	<0.01	<0.05	0.42
Domestic	151	149	2	1.32	<0.03	<0.01	14.69	<0.01	<0.01	0.06	0.23
Public	436	435	1	0.23	<0.02	0.00	10.59	<0.01	<0.01	<0.04	0.09
Irrigation	99	95	3	3.03	<0.02	0.00	27.37	<0.01	<0.01	0.06	0.38
Stock	50	47	2	4.00	<0.04	<0.01	26.60	<0.01	<0.02	<0.05	0.29
All Wells	757	747	8	1.06	<0.02	0.00	27.37	<0.01	<0.01	<0.05	0.14

Appendix A. Number of sampled wells for nitrate, number of samples with nitrate concentrations ≤ 4 mg/L and ≥10 mg/L, percent of samples ≥ 10 mg/L, median, minimum, and maximum nitrate concentrations, 10, 25, 75, and 90th percentiles. Major aquifers, subdivided by well categories and summed for all wells in each aquifer.

Cen. Pec. All.	No. samples	Samples ≤ 4mg/l	Samples ≥10mg/l	% ≥10mg/l	Median	Min	Max	10th %	25th %	75th %	90th %
Commercial	0	0	0								
Industrial	15	10	1	6.67	1.98	<0.01	14.89	<0.01	0.58	4.82	7.15
Domestic	25	19	3	12.00	1.75	<0.02	60.80	0.03	0.19	3.32	12.10
Public	50	49	0	0.00	1.07	<0.02	8.17	0.44	0.90	1.43	2.02
Irrigation	41	20	10	24.39	4.02	<0.02	85.44	0.75	1.86	9.99	23.18
Stock	54	34	5	9.26	0.96	0.00	174.55	<0.02	0.04	5.22	9.17
All Wells	185	132	19	10.27	1.45	0.00	174.55	0.03	0.57	4.61	10.38
All Wells	185	132	19	10.27027	1.45	0	174.55	0.03	0.57	4.61	10.38
Edwards (BFZ)											
Commercial	12	12	0	0.00	1.61	<0.02	3.98	<0.03	0.43	2.58	3.60
Industrial	29	23	3	10.34	1.60	<0.01	22.96	<0.02	0.06	2.75	10.94
Domestic	161	137	5	3.11	0.90	<0.01	21.90	<0.01	<0.04	1.98	4.84
Public	351	329	0	0.00	1.63	<0.01	9.48	<0.05	1.12	2.06	2.80
Irrigation	42	36	0	0.00	1.67	<0.01	8.22	<0.01	0.62	2.55	4.14
Stock	34	33	0	0.00	0.85	<0.01	7.35	0.09	0.12	1.52	2.32
All Wells	629	570	8	1.27	1.50	<0.01	22.96	<0.02	0.50	2.07	3.73
Ed. Trin. Plat											
Commercial	0										
Industrial	31	27	0	0.00	1.73	<0.05	5.89	<0.05	1.04	3.07	5.17
Domestic	323	221	18	5.57	2.42	<0.01	213.54	0.27	1.31	4.83	7.52
Public	173	123	5	2.89	2.11	<0.01	25.88	0.09	1.07	4.23	6.58
Irrigation	115	61	20	17.39	3.61	<0.01	46.92	0.58	1.74	7.09	17.93
Stock	397	296	18	4.53	2.09	<0.01	17.98	<0.05	0.68	4.02	6.75
All Wells	1039	728	61	5.87	2.35	<0.01	213.54	0.10	1.07	4.53	7.45

Appendix A. Number of sampled wells for nitrate, number of samples with nitrate concentrations ≤ 4 mg/L and ≥10 mg/L, percent of samples ≥ 10 mg/L, median, minimum, and maximum nitrate concentrations, 10, 25, 75, and 90th percentiles. Major aquifers, subdivided by well categories and summed for all wells in each aquifer.

Gulf Coast	No. samples	Samples ≤ 4mg/l	Samples ≥10mg/l	% ≥10mg/l	Median	Min	Max	10th %	25th %	75th %	90th %
Commercial	35	34	0	0.00	<0.02	<0.01	5.98	<0.01	<0.01	0.20	1.28
Industrial	71	65	1	1.41	<0.02	<0.01	23.58	<0.01	<0.02	0.43	2.83
Domestic	552	466	34	6.16	0.20	0.00	50.82	<0.02	<0.02	1.88	5.98
Public	725	709	2	0.28	<0.02	0.00	19.83	<0.01	<0.02	0.09	0.50
Irrigation	118	110	2	1.69	0.11	0.00	19.35	<0.02	<0.02	0.59	1.77
Stock	251	194	14	5.58	0.40	<0.01	40.98	<0.01	<0.02	3.68	8.04
All Wells	1752	1578	53	3.03	<0.05	0.00	50.82	<0.01	<0.02	0.70	3.98
High Plains											
Commercial	14	7	5	35.71	4.38	0.86	63.75	1.45	2.50	15.86	23.43
Industrial	45	38	4	8.89	1.90	<0.01	80.63	0.16	0.70	2.49	5.95
Domestic	727	473	130	17.88	2.17	0.00	80.97	0.34	0.87	6.99	14.99
Public	464	371	20	4.31	1.88	<0.01	94.54	0.55	1.21	3.43	6.41
Irrigation	1497	1168	95	6.35	2.05	0.00	34.22	0.58	1.17	3.62	7.83
Stock	459	365	30	6.54	1.84	<0.01	65.43	0.31	0.88	3.70	6.49
All Wells	3206	2422	284	8.86	2.01	0.00	94.54	0.24	0.41	2.09	3.37
Hueco Bolson											
Commercial	1	1	0	0.00	0.09	0.09	0.09				
Industrial	23	23	0	0.00	0.82	0.02	1.99	0.09	0.26	1.38	1.65
Domestic	10	9	0	0.00	2.29	0.90	4.80	0.90	0.90	2.87	3.98
Public	213	194	3	1.41	1.24	0.00	16.21	<0.03	0.41	2.26	3.79
Irrigation	20	18	0	0.00	0.99	<0.01	6.91	<0.02	0.28	1.73	2.31
Stock	7	5	0	0.00	1.27	0.09	4.86	0.14	0.65	3.20	4.40
All Wells	274	250	3	1.09	1.21	0.00	16.21	0.05	0.39	2.14	3.77

Appendix A. Number of sampled wells for nitrate, number of samples with nitrate concentrations ≤ 4 mg/L and ≥10 mg/L, percent of samples ≥ 10 mg/L, median, minimum, and maximum nitrate concentrations, 10, 25, 75, and 90th percentiles. Major aquifers, subdivided by well categories and summed for all wells in each aquifer.

Seymour	No. samples	Samples ≤ 4mg/l	Samples ≥10mg/l	% ≥10mg/l	Median	Min	Max	10th %	25th %	75th %	90th %
Commercial	1	0	1	100.00	17.79	17.79	17.79				
Industrial	5	2	2	40.00	5.78	0.08	23.21	0.53	1.20	11.49	18.53
Domestic	120	18	83	69.17	13.82	0.20	104.79	2.27	8.68	21.10	30.63
Public	48	7	30	62.50	11.31	0.84	31.18	3.82	8.37	14.50	19.78
Irrigation	42	4	24	57.14	12.49	1.71	28.00	5.26	7.88	19.14	23.04
Stock	20	1	15	75.00	13.21	0.70	334.92	6.65	10.09	21.53	46.97
All Wells	236	32	155	65.68	12.95	0.08	334.92	3.23	8.24	19.39	26.31
Trinity											
Commercial	11	10	0	0.00	0.11	<0.01	5.70	<0.01	<0.01	0.85	1.14
Industrial	43	41	0	0.00	0.02	<0.01	5.06	<0.01	<0.01	0.17	1.60
Domestic	588	515	26	4.42	0.15	0.00	60.96	<0.01	<0.02	1.31	4.89
Public	841	829	2	0.24	<0.02	0.00	26.41	<0.01	<0.02	1.12	4.75
Irrigation	108	77	9	8.33	0.59	0.00	52.56	<0.01	<0.02	4.59	8.16
Stock	71	58	5	7.04	0.61	<0.01	30.50	<0.01	0.10	2.50	7.44
All Wells	1662	1530	42	2.53	0.05	0.00	60.96	<0.01	<0.01	0.63	2.92
Unconf.											
Commercial	7	6	0	0.00	0.85	<0.01	5.70	<0.01	<0.01	0.99	2.96
Industrial	10	8	0	0.00	0.91	<0.01	5.06	<0.01	0.10	2.28	4.18
Domestic	413	351	21	5.08	0.32	0.00	60.96	<0.01	<0.02	1.79	5.92
Public	322	311	2	0.62	0.10	0.00	26.41	<0.01	<0.01	0.91	1.90
Irrigation	84	54	9	10.71	2.15	0.00	52.56	<0.02	0.11	5.41	10.12
Stock	52	42	5	9.62	0.95	<0.01	30.50	0.03	0.29	3.16	9.51
All Wells	888	772	37	4.17	0.05	0.00	60.96	<0.01	<0.02	0.62	2.88
Conf.											
Commercial	4	4	0	0.00	0.06		0.65	<0.01	<0.01	0.25	0.49
Industrial	33	33	0	0.00	<0.02	<0.01	1.64	<0.01	<0.01	0.08	0.19
Domestic	175	164	5	1.21	0.05	0.00	20.29	<0.01	<0.01	0.54	2.46
Public	519	518	0	0.00	<0.02	0.00	5.84	<0.01	<0.01	0.04	0.29
Irrigation	24	23	0	0.00	<0.01	<0.01	8.50	<0.01	<0.01	<0.02	0.04
Stock	19	16	0	0.00	0.13	<0.01	7.18	<0.01	<0.02	0.84	6.45
All Wells	774	758	5	0.65	<0.02	0.00	20.29	<0.01	<0.01	0.10	0.57