Data Mining Analysis of Nitrate Occurrence in Ground Water in Central and Northeast Florida and an Overview of Stormwater Best Management Practices for Nitrate Control

Andy O'Reilly
Florida Integrated Science Center
U.S. Geological Survey

Ni-Bin Chang
Stormwater Management Academy
University of Central Florida

Marty Wanielista
Stormwater Management Academy
University of Central Florida

ABSTRACT

Nitrate concentrations have increased in many Upper Floridan aquifer springs since the 1950s, exceeding 1 mg/L in recent years at some springs. It is necessary to understand the occurrence and biogeochemical transformation of nitrate in ground water to develop effective measures to reduce ground-water contamination. The objective of this study was to develop a statistical model to predict the presence of nitrate possibly attributable to human influences. A database of nitrate concentrations and related water-quality, land-use, and geologic parameters for 570 wells sampled from 1990-2006 in central and northeast Florida was analyzed. For these purposes, nitrate is considered to be present when its concentration is equal to or greater than 0.2 mg/L and absent (at background concentrations) below this level. Five statistical models were developed consisting of three different types: classification tree, regression, and neural network. The classification tree was chosen as the preferred model because of its good predictive ability and ease of interpretation. Many of the associations identified by the classification tree model are consistent with the nitrification and denitrification processes that describe the biogeochemical transformation of other nitrogen species into nitrate and vice versa. Elevated nitrate concentrations were found to occur when dissolved oxygen concentrations were high and total organic carbon concentrations were low. These aerobic conditions with little carbon availability are not conducive for denitrification. Additionally, other factors were identified that may indicate possible sources of nitrate and aquifer vulnerability. Elevated nitrate levels were found to occur when potassium concentrations were high, possibly indicating fertilizer application as a source. Elevated nitrate levels were found to occur more commonly for unconfined aquifers than for confined aquifers, indicating that hydrogeologic conditions such as the presence of clay confining sediments may retard the movement of nitrogen-contaminated ground water. Stormwater runoff is one of the possible sources of nitrogen, among others such as septic tanks and land-based application of reclaimed water or fertilizer, which can contribute to elevated nitrate concentrations in ground water. As such, the results of this data mining analysis provided the impetus for development of novel technologies for nutrient removal using in-situ permeable reactive media units in stormwater retention ponds. Reactive

media of interest include tire crumb, sawdust, iron-amended resins, wood-fiber mulch, peat, leaf compost, naturally occurring clay or acid soils, zeolites, sulfur, polymers, and paper.

Key Words: Ground-water management, Nutrient removal, Permeable reactive media, Water Resources

Introduction

Nutrients in ground water may be discharged into surface waters by baseflow or direct discharge from springs. For streams draining an urban area of central Florida, isotope tracers were used to show that 76% of stream flow was attributed to ground water; this ground-water flow can have a substantial effect on the quality of the water flowing in the stream (Gremillion et al., 2000). There are currently (March 2007) about 1,250 water-body segments on the State of Florida impaired water bodies list (Florida Department of Environmental Protection, 2007). Of these waters, about 60% are classified as either lakes or streams. About 45% of the lakes and streams are impaired as measured by nutrients. The Florida Department of Environmental Protection (FDEP) also published a comprehensive integrated assessment of water quality (FDEP, 2006). Noted in this publication for many of the springs in the State was a nitrate-level increase (two to three times) over the past 20 years. Nitrate concentrations have increased in many Floridan aquifer springs since the 1950s, exceeding 1 mg/L in recent years at some springs (Spechler and Halford, 2001, p. 47; Phelps, 2004, p. 4). This trend indicates that human activities at the land surface likely are impacting ground-water quality, which is in turn impacting surface-water quality.

Two factors control the occurrence of nitrate in ground water: (1) transport of nitrogen from land surface into the aquifer; and (2) biogeochemical transformation of nitrogen. The occurrence and transport of nitrogen in the subsurface can depend on whether there is a source of nitrogen (e.g. fertilizer and animal or human waste) and the hydrogeologic properties of the aquifer (e.g. clay versus sand and depth of the aquifer). Two important processes that result in the transformation of nitrate are nitrification and denitrification. Nitrification is a process in which ammonium is oxidized and denitrification is a process in which nitrate is reduced. Nitrification is a microbiologically mediated process that occurs under aerobic (oxygen containing) conditions. Denitrification also is a microbiologically mediated process but occurs under anaerobic (oxygen depleted) conditions. Denitrification also requires the presence of an electron donor, which may commonly include organic carbon, iron, manganese, or sulfate.

Stormwater runoff is one possible source of nitrogen, among others such as septic tanks and land-based application of reclaimed water or fertilizer, which can contribute to elevated nitrate concentrations in ground water. On the other hand, throughout the eastern United States, from the Front Range of the Rocky Mountains to the Atlantic Ocean, bioavailable nitrogen has been falling in the rain since the industrial revolution (Smil, 1990; Vitousek et al., 1997). Nitrogen- and phosphorous-containing compounds are found in urban runoff primarily from highways (USEPA, 1999). Nitrates result both from vehicular exhaust on the road itself and from fertilization of landscaped areas beside the roads (German, 1989). Nitrate is very soluble and does not sorb well to soil components during infiltration (Spalding and Kitchen, 1988). In a trend that is expected to continue, these additions have been increasing (Brimblecombe and Stedman, 1982; Vitousek et al., 1997). Nitrogen, particularly nitrate nitrogen, easily moves from terrestrial ecosystems into surface and ground waters, including lakes, streams, rivers, and estuaries (Baker, 1992; Kahl et al., 1993; Peterjohn et al., 1996). Because nitrogen may be a limiting nutrient for plants, increased quantities of nitrogen in ecosystems alter competitive relationships among terrestrial and aquatic organisms (USEPA, 2005).

The purpose of this paper is to discuss the levels of nitrate in the ground water in central and northeast Florida and describe the development and interpretation of a statistical model to predict the presence of nitrate possibly attributable to human influences. Since the ground water in many locations is related to the surface waters, it is anticipated that the study of ground-water nitrate levels will provide valuable insight into nitrate levels in surface waters. In particular, an understanding of the occurrence and biogeochemical transformation of nitrate in ground water is desired so that effective measures to reduce ground-water contamination can be developed. In closing, this paper presents a discussion of how novel soil substitution technologies using permeable reactive media can be applied as part of comprehensive stormwater Best Management Practices (BMPs) to mitigate the nutrient impact on ground water and surface springs.

Methodology

Simple exploratory data analysis techniques were employed to gain a basic understanding of the data. These techniques consisted of (1) geographic analysis, (2) descriptive statistics, and (3) visual examination of bivariate distributions of the target variable and each input variable. Next, a data mining analysis was performed, involving statistical modeling using three different types of models. Finally, a comparison of the predictive performance of the statistical models was made and a preferred model was selected for further interpretation.

Study Sites

The hydrogeology of central and northeast Florida consists of three principal aquifers in order from shallowest to deepest: 1) surficial aquifer system (SAS), which is the aquifer exposed at land surface; 2) intermediate aquifer system/confining unit (IAC), which may constitute a sand/limestone aquifer or a clay confining unit depending on the local geology; and 3) Floridan aquifer system (FAS), which is a thick sequence of limestone formations that serve as the primary source of fresh water for drinking water and irrigation purposes. The SAS is an unconfined aquifer, meaning that it is not overlain by a low permeability, generally clay confining unit. The IAC is a confined aquifer where water-bearing units exist; otherwise the IAC functions as a confining unit. The FAS generally is a confined aquifer (confined by the overlying IAC), but in some areas is unconfined where the IAC is absent. Aquifer type is important in that an unconfined aquifer generally is more susceptible to contamination from land-use activities than a confined aquifer.

The geographic analysis consisted simply of mapping well locations coded by nitrate concentration ranges (Figure 1). Most wells were located in central Florida, with substantially fewer wells present in south-central and northeast Florida. The 570 wells tapped all three primary aquifers: SAS (277), FAS (263), and IAC (26). Four wells did not have aquifer data available. Wells with elevated nitrate concentrations (equal to or greater than 0.2 mg/L) generally are interspersed among wells with background concentrations (less than 0.2 mg/L). A notable exception is east Orange County where no wells have elevated nitrate concentrations.

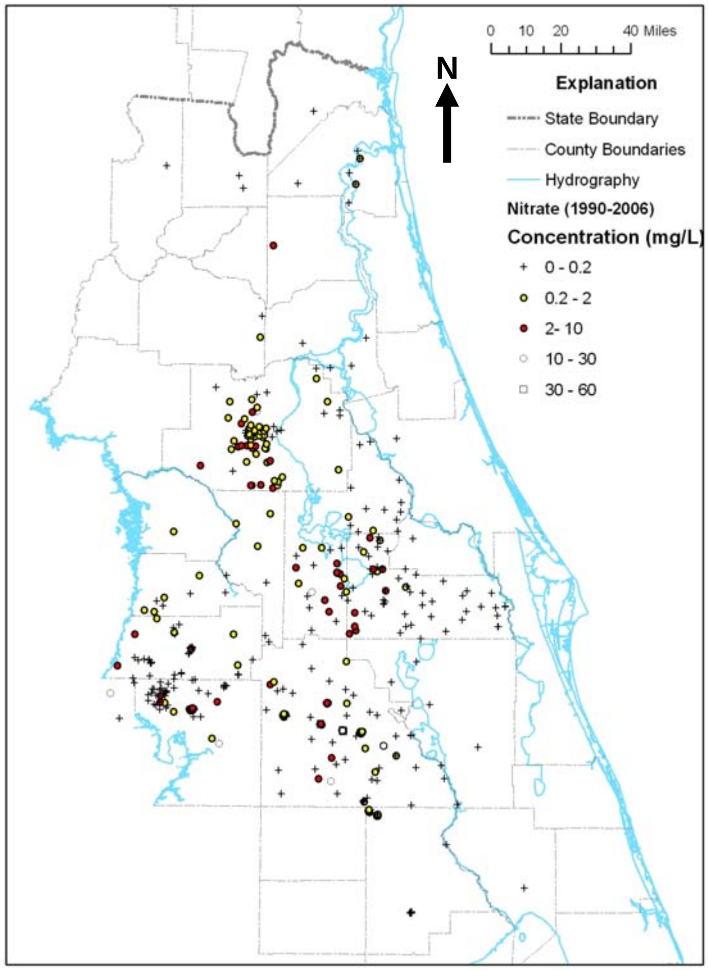

Figure 1. Nitrate concentrations in ground water, central and northeast Florida.

Database

A database of nitrate concentrations and related water-quality, land-use, and geologic parameters for 570 wells in central and northeast Florida was analyzed. For these purposes, nitrate is considered to be present when its concentration is equal to or greater than 0.2 mg/L and absent (at background concentrations) below this level. Madison and Brunett (1985, p. 95) identified 0.2 mg/L as the level below which nitrate concentrations probably represent natural background conditions, based on a statistical analysis of nitrate concentrations in nearly 124,000 wells throughout the United States. The database used in this study was compiled from several sources:

- Water-quality and well construction data were obtained from the U.S. Geological Survey National Water Information System (http://waterdata.usgs.gov/fl/nwis). Data were extracted for all wells in central and northeast Florida that had been sampled for nitrate at least once from January 1, 1990, through October 31, 2006.
- Land use at each well was obtained by a Geographic Information System (GIS) analysis of well location and a land-use map developed by Sepulveda (1999). This map represents estimated land-use conditions in 1995.
- Physiography and subsurface hydrogeology at each well was obtained by a GIS analysis of well location and relevant maps presented by Sepulveda (2002).

- Surface geology at each well was obtained by a GIS analysis of well location and a geologic map developed by Scott et al. (2001).

The final database consisted of 570 wells and 35 variables. Not all variables were used; the variables that were not pertinent to the analysis were rejected, e.g. latitude and longitude. The role, level, and description of each variable are presented in Table 1.

Table 1. Database variables.
[Level types: Binary, 2 discrete values; Interval, continuous numeric; Nominal, categorical; ICU, intermediate confining unit; UFA, Upper Floridan aquifer]

Variable Name	Role	Level	Description
ANC	Input	Interval	Acid Neutralizing Capacity
AQUIFER	Input	Nominal	Name of aquifer tapped by well
AQ_TYPE	Input	Nominal	Type of aquifer (confined or unconfined)
Ca	Input	Interval	Calcium concentration (mg/L)
Cl	Input	Interval	Chloride concentration (mg/L)
DEPTH_ICU	Input	Interval	Depth of ICU below land surface
DEPTH_UFA	Input	Interval	Depth of UFA below land surface
DO	Input	Interval	Dissolved oxygen concentration (mg/L)
DOC	Input	Interval	Dissolved organic carbon concentration (mg/L)
Fe	Input	Interval	Iron concentration (ug/L)
GEOH_UNIT	Rejected	Nominal	Code for geologic unit tapped by well
GEO_SERIES	Rejected	Nominal	Geologic unit series
ICU_THICK	Input	Interval	Thickness of ICU (ft)
K	Input	Interval	Potassium concentration (mg/L)
LAND_USE	Input	Nominal	Land use
LAT_DD	Rejected	Interval	Latitude of well
LONG_DD	Rejected	Interval	Longitude of well
Mg	Input	Interval	Magnesium concentration (mg/L)
Mn	Input	Interval	Manganese concentration (ug/L)
NO3_AVG	Rejected	Interval	Nitrate concentration (mg/L); arithmetic average, if multiple values exist during 1990-2006
NO3_RMSD	Rejected	Interval	Root mean square deviation of multiple nitrate concentrations used to compute NO3_AVG
NO3_YN	Target	Binary	Elevated nitrate (Y=1) if NO3_AVG $\geq$ 0.2; background nitrate (N=0) if NO3_AVG < 0.2
Na	Input	Interval	Sodium concentration (mg/L)
ORL_TPA	Rejected	Nominal	USGS office who collected data
pH	Input	Interval	pH
PHYSIOGRAPHY	Input	Nominal	Land physiography
SC	Input	Interval	Specific conductance (microsiemens per cm)
SO4	Input	Interval	Sulfate concentration (mg/L)
STAID	Rejected	Interval	USGS well identification number
SURF_GEOL	Input	Nominal	Code for geologic unit exposed at land surface
TDS	Input	Interval	Total dissolved solids (mg/L)
TOC	Input	Interval	Total organic carbon concentration (mg/L)
Temp	Input	Interval	Ground water temperature (degrees C)
WELLNUM	ID	Interval	Sequential well number
WELL_DEPTH	Input	Interval	Well depth (ft)

The target (response) variable NO3_YN was derived based on the average nitrate concentration (NO3_AVG, Table 1), thereby ignoring any temporal trend that might exist during the 16-year (1990-2006) data period. However, fewer than 20% (109) of wells were sampled three or more times and very few wells (39) were sampled four or more times, so little data exist for identifying temporal variability.

Statistical Modeling and Data Mining Analyses

Five statistical models were developed consisting of three different types – classification tree, regression, and neural network – to identify the model with the best predictive ability. A classification tree model was constructed based on CART principles (Breiman et al, 1984). Two logistic regression models, including both linear and nonlinear models, were constructed. Two neural network models were constructed with multilayer perceptron architectures with a different number of hidden layers. SAS Enterprise Miner (version 5.2) was used to manipulate the database; construct, train, and validate each model; and compare performance among all five models. In general, the modeling process can be summarized as follows:

- Exploratory data analysis and variable consolidation/imputation/transformation – This was performed on the entire database so that any rare occurrences would not be missed by subsampling. Nominal variables having five or more categories were consolidated using classification trees. Missing data were imputed using classification trees. Interval variables were transformed to maximize normality.
- Data partition – The database was divided into a training dataset (70%, 399 wells) and a validation dataset (30%, 171 wells). Due to the relatively small size of the database, a test dataset was not used.
- Classification tree model – The tree model was trained and validated on the partitioned database with no consolidation, imputation, or transformation of variables.
- Regression and Neural Network models – Variables in the partitioned database were consolidated, imputed, and transformed as necessary before training and validating regression and neural network models.
- Model comparison – All five models were compared to determine which yielded the best classification of wells with elevated nitrate concentrations (target variable NO3_YN = 1) and wells with background nitrate concentrations (NO3_YN = 0).

Results and Discussion

Data Mining and Statistical Analyses

Summary statistics were computed for all 25 input variables (Table 2). Most variables have some missing data; three variables have greater than 50% missing data (ANC, DOC, and TOC). Most of the interval variables are right skewed, some extremely so (e.g. Cl, Fe, and Na). One nominal variable has a large number of categories (SURF_GEOL).

Table 2. Summary statistics of the 25 input variables used in statistical models.

Variable	No. Missing	No. Categories	Minimum	Maximum	Mean	Stand. Dev.	Skewness	Kurtosis
ANC	355	--	1.0	391.0	128.7	85.6	0.4	-0.1
AQ_TYPE	4	2	--	--	--	--	--	--
AQUIFER	4	3	--	--	--	--	--	--
Ca	63	--	0.0	225.0	47.5	42.0	1.1	1.3
Cl	63	--	0.5	1200.0	29.4	107.9	8.3	74.6
DEPTH_ICU	0	--	2.5	241.6	52.0	41.2	1.4	2.1
DEPTH_UFA	0	--	2.5	628.0	129.8	127.0	1.8	3.5
DO	66	--	0.0	13.1	1.8	2.2	1.6	1.9
DOC	336	--	0.2	42.0	6.4	7.7	2.4	7.0
Fe	79	--	2.0	44200	906.3	2884.4	9.1	116.5
ICU_THICK	0	--	0.0	603.1	77.7	117.1	2.4	6.0
K	63	--	0.0	29.0	2.4	4.5	3.0	9.4
LAND_USE	0	7	--	--	--	--	--	--
Mg	63	--	0.1	100.0	6.9	9.1	4.5	30.7
Mn	214	--	0.1	820.0	31.3	87.1	5.5	36.4
Na	63	--	0.4	640.0	16.3	57.1	8.2	74.3
pH	9	--	3.8	9.8	6.4	1.2	-0.5	-0.9
PHYSIOGRAPHY	0	5	--	--	--	--	--	--
SC	8	--	3.0	5050.0	381.9	432.2	5.9	47.4
SO4	56	--	0.1	590.0	30.1	60.1	4.8	30.0
SURF_GEOL	0	17	--	--	--	--	--	--
TDS	223	--	24.0	3280.0	260.8	322.1	5.4	37.8
Temp	11	--	19.0	32.8	24.4	1.6	0.3	1.7
TOC	442	--	0.1	26.0	3.1	4.3	2.9	9.6
WELL_DEPTH	48	--	1.0	1450.0	156.6	252.7	2.7	8.4

An empirical cumulative probability distribution was developed for nitrate concentration (variable NO3_AVG) using the Weibull plotting-position formula (Figure 2). Nearly half the wells (47.9%) had nitrate concentrations of zero, indicating that the concentration was below the reporting limit established for the analytical laboratory method used. Approximately 38.4% (219) of the wells had nitrate concentrations above the background concentration. These wells were assigned a value of 1 for the variable NO3_YN, which implies the impact of nitrate. Bivariate distributions (on the target variable NO3_YN) were examined and some indicated promising explanatory value. For example, nearly all wells with dissolved oxygen concentrations greater than 4 mg/L also had elevated nitrate

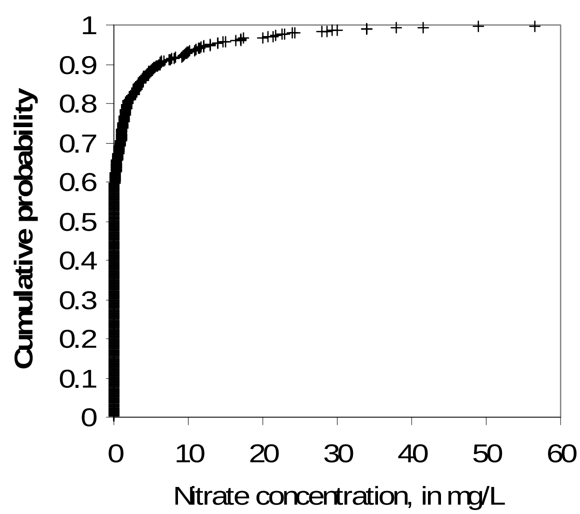

Figure 2. Empirical cumulative probability of nitrate concentration in ground water.

concentrations (Figure 3). Also, wells located in urban, agricultural, or transitional land uses were much more likely to have elevated nitrate concentrations than wells in range, wetland, forest, or mining land uses (Figure 4).

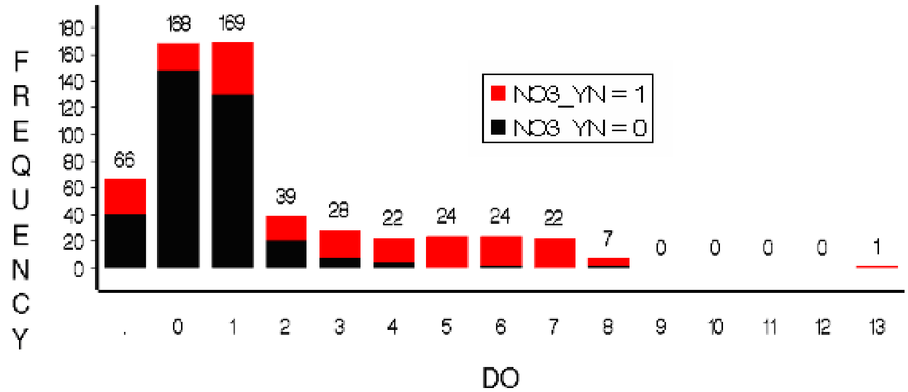

Figure 3. Bivariate histogram of dissolved oxygen concentration.

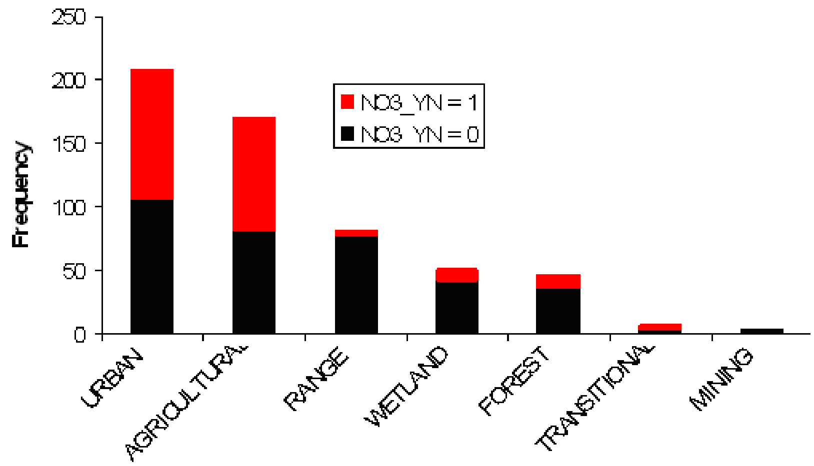

Figure 4. Bivariate histogram of land-use types.

Model Comparison

The primary objective of this study was to identify an interpretable model to gain some understanding of the processes that may be controlling the occurrence and transformation of nitrate in ground water. All models were trained using only the training dataset (399 wells). Each calibrated model was used to predict the target variable NO3_YN in the validation dataset (171 wells) and summary statistics of model fit for all five models were compared. Because the target variable is binary (decision prediction), the best model was selected based on validation misclassification rate. The validation misclassification rates for the tree and linear logistic regression models were identical at 12.28% (21 wells). Table 3 shows the better classification performance of the classification tree and linear logistic regression models compared to the other three models. Based on their lower misclassification rates and convenient interpretability, the tree and linear logistic regression models were preferred.

Table 3. Classification performance of the five predictive models.

Model	Role	False Negative	True Negative	False Positive	True Positive	Misclassification Rate (%)
Classification Tree	Train	20	222	24	133	
	Validate	11	95	10	55	12.28
Linear Regression	Train	38	231	15	115	
	Validate	15	99	6	51	12.28
Neural	Train	29	234	12	124	
	Validate	14	97	8	52	12.87
AutoNeural	Train	30	234	12	123	
	Validate	18	100	5	48	13.45
Nonlinear Regression	Train	34	234	12	119	
	Validate	18	96	9	48	15.79

The classification tree was selected as the best model for further interpretation because of its ability to accommodate missing data using surrogate splitting rules. Many of the important variables in the tree model are those used for surrogate rules rather than primary splitting rules (Table 4). A surrogate rule is one that is used in place of a primary rule when data are missing, thereby precluding imputation. Imputation of variables required for the regression models could possibly obscure associations or manifest false associations, because any imputation method is a type of model in itself with its own set of assumptions.

Table 4. Classification tree variable importance. [Nodes, number of nodes variable used in primary splitting rule; Surrogates, number of nodes variable used in surrogate splitting rule; Training, measure of relative importance based on training dataset; Validation, measure of relative importance based on validation dataset]

Variable	Nodes	Surrogates	Training	Validation	Importance
DO	4	1	1.000	1.000	
DEPTH_ICU	0	7	0.991	0.869	
Fe	1	4	0.885	0.709	
Temp	0	3	0.774	0.709	
WELL_DEPTH	0	3	0.669	0.617	
Na	0	4	0.667	0.595	
DEPTH_UFA	1	5	0.651	0.639	
SC	0	4	0.644	0.613	
Cl	1	4	0.629	0.503	
K	2	3	0.609	0.526	
SO4	1	2	0.528	0.424	
Mg	1	2	0.518	0.374	
PHYSIOGRAPHY	0	2	0.504	0.536	
AQ_TYPE	0	1	0.492	0.485	
pH	0	2	0.486	0.457	
DOC	1	0	0.482	0.389	
ICU_THICK	0	2	0.346	0.324	
TOC	1	0	0.290	0.119	
LAND_USE	0	1	0.189	0.278	
Ca	0	1	0.126	0.156	
TDS	0	2	0.114	0.052	
ANC	0	0	0.000	0.000	
AQUIFER	0	0	0.000	0.000	
Mn	0	0	0.000	0.000	
SURF_GEOL	0	0	0.000	0.000	

Classification Tree Model

A classification tree is a statistical modeling technique for modeling the predictive relation between a target variable and a set of explanatory variables whereby a dataset is recursively partitioned into increasingly homogenous (in terms of the target variable) subsets. The classification tree described herein was based on CART principles. This consisted of binary splits, Entropy-reduction splitting criterion, and surrogate rules for missing data. Also, a large tree was initially grown and the best subtree selected based on minimum misclassification rate.

The final tree consisted of 14 leaves (terminal nodes) with a validation misclassification rate of 12.28%. For simplified interpretability the tree model was pruned according to the 1-SE rule (Breiman et al., 1984, p. 78-80), yielding a smaller tree (9 leaves) with a validation misclassification rate of 14.62% (Figure 5).

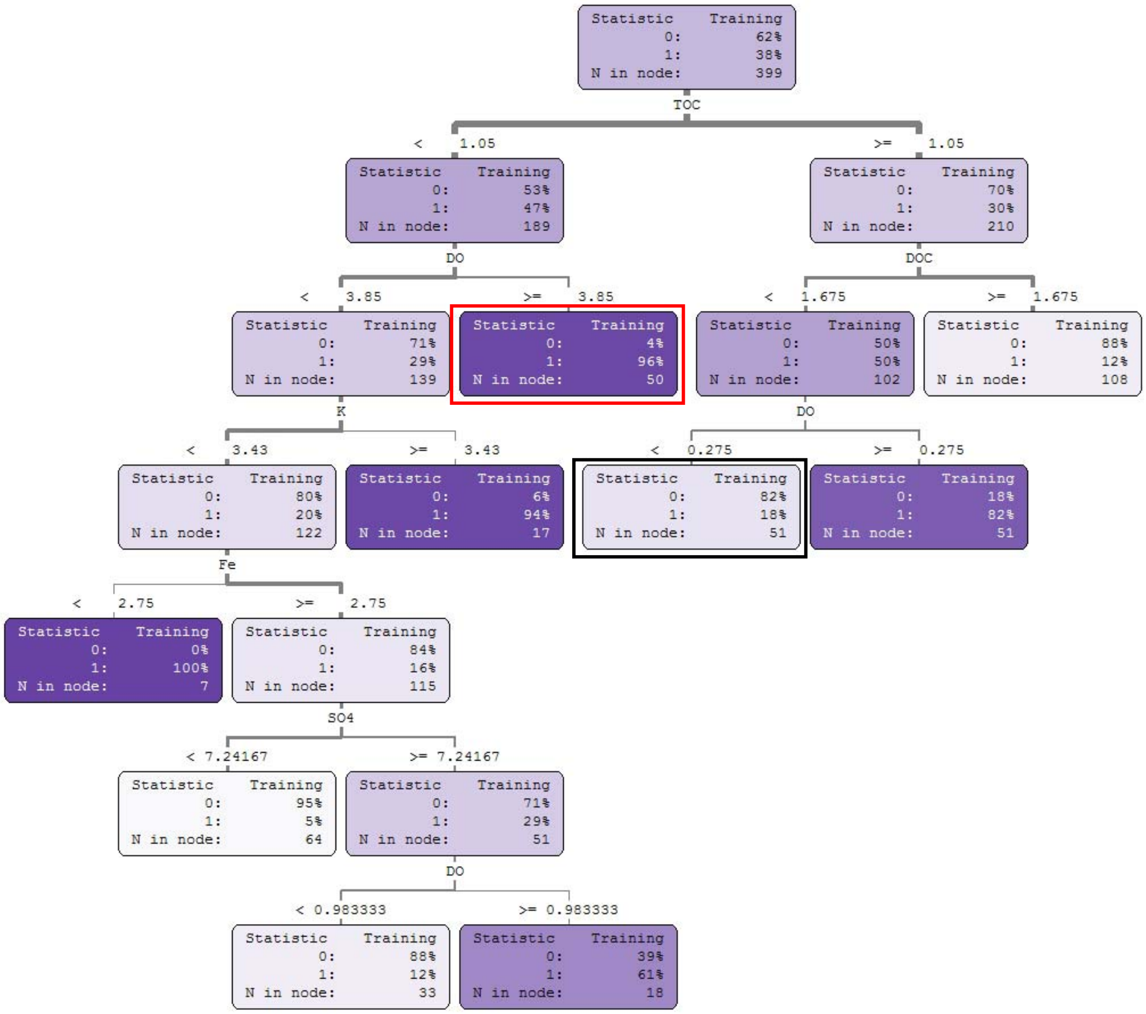

Figure 5. Final classification tree model (pruned using the 1-SE rule). Blue shading is proportional to the percentage of wells (in training dataset) with elevated nitrate concentration (NO3_YN=1), with light blue indicating a subset of wells with mostly background nitrate concentration and dark blue indicating a subset of wells with mostly elevated nitrate concentrations.

Many of the splitting rules are consistent with the nitrification and denitrification processes described earlier. For example, one rule indicates elevated nitrate concentrations (high primary outcome probability) when dissolved oxygen concentrations are high and total organic carbon concentrations are low (see tree leaf enclosed in the red box, Figure 5). These aerobic conditions with little carbon availability are not conducive for denitrification, thus nitrate present in the ground water is relatively stable. Under these conditions, other nitrogen species, if present, could be transformed into nitrate under aerobic conditions via the nitrification process. Also, high dissolved oxygen concentrations may indicate rapid movement of water through the ground-water flow system, precluding significant nitrate reduction based on reaction kinetics. Another example is the rule that indicates background concentrations of nitrate (low primary outcome probability) when dissolved oxygen concentrations are low, dissolved organic carbon concentrations are low, and total organic carbon concentrations are high (see tree leaf enclosed in the black box, Figure 5). These anaerobic conditions with high total organic carbon concentrations are conducive for denitrification, thus nitrate entering the ground water may be transformed to a different nitrogen species. A process-based explanation for the low dissolved organic carbon concentration is not apparent.

Other factors were identified that suggest potential nitrogen sources and aquifer vulnerability. For example, elevated nitrate concentrations occurred when potassium concentrations were high. Potassium concentrations do not influence the nitrogen transformation processes, but may likely be an indicator of fertilizer application. Potassium, in addition to nitrogen and phosphorus, is an important component of commercial fertilizers. A measure of aquifer vulnerability is expressed by variable AQ_TYPE, which is used in a surrogate splitting rule indicating elevated nitrate concentrations for unconfined aquifers and background concentrations for confined aquifers. Hydrogeologic conditions, such as the presence of clay confining sediments, may retard the movement of contaminated ground water into the aquifer. Interestingly, several of the variables important to the classification tree model (Table 4) are influential effects in the linear logistic regression model (DO, TOC, and K had the largest regression coefficient absolute magnitude). Further examination of the linear logistic regression model might prove instructive.

BMP Strategies

Permeable reactive barriers (PRBs) are a relatively new ground-water treatment technology where contaminants are converted into innocuous by-products in situ in the subsurface (Kietlinska and Renman, 2005). Biofilm-forming microbes can form biobarriers to inhibit nutrient migration in ground water. Also subsurface biofilms have the potential for biotransformation of nitrate to less harmful forms, such as nitrogen gas in an anaerobic environment as long as there are electron donors, thereby providing an *in situ* method for treatment of contaminated ground-water. One of the drawbacks of PRBs is their high initial capital cost. One method to reduce the total cost of PRBs is to reduce the cost of the reactive media. The last focus in this study is to provide a sound literature review leading to the identification of cost-effective reactive media that may be used in such barriers for ground-water remediation (see Table 5). Consequently, PRBs configured as part of the natural soil profile at some strategic locations in the watersheds may become one of the cost-effective BMPs for the attenuation of high nitrate concentrations in aquifers underlying developed regions.

Table 5. Summary of reactive media for the removal of nitrogen and phosphorous.

Reactive Media	Additional Environmental Benefits	References
Peat	Cu, Zn, Ni, and Mo	DeBusk and Langston, 1997; Braun-Howland, 2003; Kietlinska and Renman, 2005
Sawdust	Pesticide and phosphate	Gan et al., 2004
Paper		Kim et al., 2000
Lignocellulosic Materials		Tshabalala, 2002
Tire Crumb		Lisi et al., 2004
Sulfur/Limestone	TSS	DeBusk and Langston, 1997; Kim et al., 2000; Darbi et al., 2002; Zhang, 2002
Mulch/wood fiber	Polynuclear aromatic hydrocarbons	Kim et al., 2000; Jokela et al., 2002; Boving and Zhang, 2004; Ray et al., 2006
Compost	Heavy metal	Richman, 1997
Zeolites	Benzene, sulfate , chromate	Li, 2003
Cotton waste		Della Rocca et al., 2005
Perlite		www.perlite.net
Clay	phosphates, thiocyanates, cadmium, lead, nickel	Gálvez et al., 2003; Lazaridis, 2003
Shale and masonry sand		Forbes et al., 2005
Waste foundry sand	TCE, alachlor, and Metolachlor, Zinc	Benson, 2001
Acid soils (spodosols)		USDA, 2007
Opoka	Zinc	Braun-Howland, 2003
Wollastonite		DeBusk and Langston, 1997; Hedström, 2006
Iron sulfide (pyrite)		Tesoriero et al., 2000; Baeseman et al., 2006
Polyurethane porous media		Han et al., 2001
Clinoptilolite		Hedström, 2006
Blast furnace slag		Hedström, 2006
Emulsified edible oil substrate		Lieberman et al., 2005

Conclusion

This study involved the data mining analysis of 570 wells in central and northeast Florida to investigate variations in nitrate concentrations. Three types of data mining models, consisting of classification tree, logistic regression, and neural network models, were developed to explore the explanatory ability of 25 input variables to predict nitrate presence or absence. The classification tree was identified as the most suitable model for assessment, yielding an easily interpretable model with a misclassification rate of 14.62%. In general, the results of this analysis suggest that nitrate occurrence in ground water in central and northeastern Florida is controlled, to some degree, by biogeochemical transformation processes, nitrogen sources, and hydrogeologic conditions. Finally, a discussion of BMP strategies focusing on permeable reactive barriers (PRBs) as a cost-effective technology, may lead to the development of a suite of sustainable engineering technologies for in situ attenuation of nutrients in the subsurface.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Baeseman, J. L., Smith, R. L., and Silverstein, J., 2006. Denitrification potential in stream sediments impacted by acid mine drainage: effects of pH, various electron donors, and iron. Microbial Ecology, 51(2), 232-241.

Baker, L. A., 1992. Introduction to nonpoint source pollution in the United States and prospects for wetland use. Ecological Engineering, 1, 1-26.

Benson, C. H., 2001. Using Waste Foundry Sands as Reactive Media in Permeable Reactive Barriers, DNR Project #147.

Boving, T. B. and Zhang, W., 2004. Removal of aqueous-phase polynuclear aromatic hydrocarbons using aspen wood fibers. Chemosphere, 54, 831–839.

Braun-Howland,E., 2003. Validity Assessment of Methods to Distinguish Between Ruminant and Human Sources of Fecal Contamination in Watersheds, New York State Water Resources Institute Annual Technical Report-FY 2003, Cornell University, Ithaca, NY.

Brieman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., 1984. Classification and regression trees: Boca Raton, Fla., Chapman & Hall/CRC, 358 p.

Brimblecombe, P. and Stedman, D. H., 1982. Historical evidence for a dramatic increase in the nitrate component of acid rain. Nature 398, 460-462.

Darbi, A., Viraraghavan, T., Butler, R. and Corkal, D., 2002. Batch studies on nitrate removal from potable water. Water SA, 28 (3), 319-322.

DeBusk T. A. and Langston, M. A., 1997. An Evaluation of Filter Media for Treating Stormwater Runoff. Proceedings of the Fifth Biennial Stormwater Research Conference, pp 82-89.

Della Rocca, C., Belgiorno, V., and Meriç, S. 2005. Cotton-supported heterotrophic denitrification of nitrate-rich drinking water with a sand filtration post-treatment. Water SA 31(2), 229-236.

Gálvez, J. M., Gómez, M.A., Hontoria, E., and González-López, J., 2003. Influence of hydraulic loading and air flowrate on urban wastewater nitrogen removal with a submerged fixed-film reactor. Journal of Hazardous Materials, B101, 219–229.

Gan, Q., Allen, S. J.,and Matthews, R., 2004. Activation of waste MDF sawdust charcoal and its reactive dye adsorption characteristics. Waste Management, 24, 841–848.

German, E. R., 1989, Quantity and quality of stormwater runoff recharged to the Floridan Aquifer system through two drainage wells in the Orlando, Florida area. US Geological Survey – Water Supply paper 2344. Prepared in corporation with the Florida Department of Environmental Regulation. Denver, CO.

Gremillion, P., Gonyeau, A., and Wanielista, M., 2000. Application of alternative hydrograph separation models to detect changes in flow paths in a watershed undergoing urban development. Hydrological Processes, 14, 1485-1501.

Florida Department of Environmental Protection (FDEP), 2006. Integrated Water Quality Assessment for Florida", Bureau of Watershed Management, Tallahassee, Fla, May 2, 2006. http://www.dep.state.fl.us/water/tmdl/docs/2006_Integrated_Report.pdf

Florida Department of Environmental Protection (FDEP), 2007. Impaired Waters Working Document Lists, Bureau of Watershed Management, Tallahassee, Florida.

Forbes, M. G., Dickson, K. L, and Waller, W. T., 2005. Recovery and fractionation of phosphorus retained by lightweight expanded shale and masonry sand used as media in subsurface flow treatment wetlands. Environ. Sci. Technol. 39, 4621-4627.

Han, D. W., Yun, H. J., and Kim, D. J., 2001. Autotrophic nitrification and denitrification characteristics of an upflow biological aerated filter. Journal of Chemical Technology and Biotechnology, 76, 1112–1116.

Hedström, A., 2006. Reactive filter materials for ammonium and phosphorus sorption in small scale wastewater treatment, Doctoral Dissertation, Luleå University of Technology, ISSN: 1402-1544

Jokela, J. P. Y., Kettunen, R. H., Sormunen, K. M., and Rintala, J. A., 2002. Biological nitrogen removal from municipal landfill leachate: low-cost nitrification in biofilters and laboratory scale in-situ denitrification. Water Research, 36, 4079–4087.

Kahl, J. S., Norton, S. A., Fernandez, I. J., Nadelhoffer, K. J., Driscoll, C. T., and Aber, J. D., 1993. Experimental inducement of nitrogen saturation at the watershed scale. Environmental Science and Technology, 27, 565-568.

Kietlinska, A. and Renman, G. 2005. An evaluation of reactive filter media for treating landfill leachate, Chemosphere 61, 933–940.

Kim, H., Seagren, E. A., and Davis, A. P., 2000. Engineered Bioretention for Removal of Nitrate from Stormwater Runoff, WEFTEC (Water Environment Federation Technical Exhibition) 2000, Water Environment Federation (WEF), Grand Rapids, MI.

Lazaridis, N. K., 2003, Sorption removal of anions and cations in single batch systems by uncalcined and calcined Mg-Al-CO_3 hydrotalcite. Water, Air, & Soil Pollution, 146(1-4), 127-139.

Lieberman, M. T., Lindow, N. L., Borden, R. C., and Birk, G. M., 2005. Anaerobic Biodegradation and Biotransformation Using Emulsified Edible Oils in Contaminated Soils, Sediments and Water Science in the Real World Volume 9, edited by Edward J. Calabrese, Paul T. Kostecki and James Dragun, pp. 485-500.

Li Z., 2003. Use of surfactant-modified zeolite as fertilizer carriers to control nitrate release, Microporous and Mesoporous Materials, 61(1), 181-188.

Lisi, R. D., Park, J. K., and Stier, J. C., 2004, Mitigating nutrient leaching with a sub-surface drainage layer of granulated tires. Waste Management, 24, 831–839.

Madison, R.J., and Brunett, J.O., 1985, Overview of the occurrence of nitrate in ground water in the United States, in National Water Summary 1984—Hydrologic events, selected water-quality trends, and ground-water resources: U.S. Geological Survey Water-Supply Paper 2275, p. 93-105.

Peterjohn, W. T., Adams, M. B. and Gilliam, F. S., 1996. Symptoms of nitrogen saturation in two central Appalachian hardwood forest ecosystems. Biogeochemistry 35, 507-522.

Phelps, G. G., 2004. Chemistry of ground water in the Silver Springs basin, Florida, with emphasis on nitrate: U.S. Geological Survey Scientific Investigations Report 2004-5144, 54 p.

Ray, A. B., Selvakumar, A., and Tafuri,A. N., 2006. Removal of selected pollutants from aqueous media by hardwood mulch. Journal of Hazardous Materials, B136, 213–218.

Richman, M. 1997. Compost media capture pollutants from stormwater runoff. Water Environment and Technology, 9,21-22.

Scott, T. M., Campbell, K. M., and Rupert F. R., 2001, Geologic map of the state of Florida: Florida Geological Survey Map Series No. 146, 1 sheet.

Sepulveda, A.A., 1999. Land uses in the ecosystem management areas of Florida: Florida Geological Survey Map Series No. 144, 1 sheet.

Sepulveda, N., 2002, Simulation and ground-water flow in the intermediate and Floridan aquifer systems in peninsular Florida: U.S. Geological Survey Water-Resources Investigations Report 02-4009, 130 p.

Smil, V. 1990. Nitrogen and phosphorus. In The Earth as Transformed by Human Action; Global and Regional Changes in the Bioshpere over the Past 300 Years, B.L. Turner II, W.C. Clark, R.W. Kates, J.F. Richards, J.T. Mathews, and W.B. Meyer (Eds). Cambridge University Press, United Kingdom. 713pp.

Spalding, R. F. and Kitchen, L. A., 1988. Nitrate in the intermediate vadose zone beneath irrigation cropland. Ground Water Monitoring Review, 8(2), 89-95.

Spechler, R. S. and Halford, K. J., 2001. Hydrogeology, water quality, and simulated effects of ground-water withdrawals from the Floridan aquifer system, Seminole County and vicinity, Florida: U.S. Geological Survey Water-Resources Investigations Report 01-4182, 116 p.

Tesoriero, A. J., Liebscher, H., and Cox, S. E., 2000. Mechanism and rate of denitrification in an agricultural watershed: Electron and mass balance along groundwater flow paths. Water Resources Research, 36(6), 1545-1559.

Tshabalala, M. A., 2002. Use of Lignocellulosic Materials as Sorbents for Pesticide and Phosphate Residues, USDA Forest Service, Madison, WI 53705-2398.

United States Department of Agriculture (USDA), http://soils.usda.gov/technical/classification/orders/spodosols.html, Date accessed: March 2007.

United States Environmental Protection Agency (USEPA), 2005. Ecosystem Stress from Chronic Exposure to Low-levels of Nitrate, EPA/600/R-05/087, October 2005.

United States Environmental Protection Agency (USEPA), 1999. Stormwater Treatment At Critical Areas Evaluation of Filtration Media. EPA 600/R-00/010, Kuly, 1999.

Vitousek, P. M., Aber, J., Howarth, R.W., Likens, G. E., Matson, P.A., Schindler, D. W., Schlesinger, W. H., and Tilman, G. D., 1997. Human alteration of the global nitrogen cycle: causes and consequences. Ecology, 1, 1-16.

Zhang, T. C., 2002. Nitrate removal in sulfur: limestone pond reactors, J. Envir. Engrg., ASCE, 128(1), 73-84.