7-1-2001

Application of Neuro-Fuzzy Technique+2:9s to Predict Ground Water Vulnerability in Northwest Arkansas

B. Dixon
University of Arkansas, Fayetteville

H. D. Scott
University of Arkansas, Fayetteville

J. V. Brahana
University of Arkansas, Fayetteville

A. Mauromoustakos
University of Arkansas, Fayetteville

Arkansas Water Resources Center

APPLICATION OF NEURO-FUZZY TECHNIQUES TO PREDICT GROUND WATER VULNERABILITY IN NORTHWEST ARKANSAS

Final Report for Award #1434HQ96GR02658

Submitted to
ARKANSAS WATER RESOURCES CENTER

June, 2001

Authors

B. Dixon, Research Specialist, H.D. Scott, Professor, Crop, Soil, and Environmental Sciences, J.V. Brahana, Professor, J.C. Dixon, Professor, Department of Geosciences, A. Mauromoustakos, Associate Professor, Agriucltural Statistics, University of Arkansas, Fayetteville, Arkansas

Publication No. PUB-183

Arkansas Water Resources Center
112 Ozark Hall
University of Arkansas
Fayetteville, Arkansas 72701

APPLICATION OF NEURO-FUZZY TECHNIQUES TO PREDICT GROUND WATER VULNERABILITY IN NORTHWEST ARKANSAS

B. Dixon[1], H. D. Scott[2], J. V. Brahana[3], A. Mauromoustakos[4], and J. C. Dixon[3]

[1] Research Specialist, Crop, Soil and Environmental Science

[2] University Professor, Crop, Soil and Environmental Science

[3] Professors, Geosciences

[4] Associate Professor, Agricultural Statistics

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Contamination of ground water has been a major concern in recent years of local, state and federal agencies involved with the management, quality, and quantity of water and their relationships with human health. The Springfield Plateau aquifer, which lies beneath the study area in northwest Arkansas, has been shown to have higher nitrate-N (NO_3-N) concentrations than the national median. The dominant landuse (LULC) of this area is agriculture (primarily pasture/cattle and woodlands) and an encroaching urbanization. The major sources of nitrogen in the study area are poultry/cattle wastes, inorganic fertilizers (Peterson et. al., 1998) and septic filter fields. Many of the soils in the Ozark Region are highly permeable and well drained and the geology is karst. The probability of pollution occurring at a given location is a function not only of its hydrogeologic setting but also of anthropogenic pollution in the area (Evans, 1990).

Delineation of vulnerable areas and selective applications of animal wastes/fertilizer in those areas can minimize contamination of ground water. Assessment of ground water vulnerability from agricultural inputs and the understanding of spatial and temporal variabilities of the most important parameters affecting vulnerability is imperative before undertaking monitoring, rehabilitation or regulatory efforts at the watershed or regional scales. However, assessment of ground water vulnerability or delineation of the monitoring zones is not easy since contamination depends upon numerous and complex interacting parameters. Uncertainty is inherent in all methods of assessing ground water vulnerability and arises from errors in obtaining data, the natural spatial and temporal variability of the hydrogeologic parameters in the field, and in the numerical approximation and computerization (National Research Council, 1993).

Existing ground water vulnerability assessment methods may be grouped into three categories: overlay and index, statistical, and process-based simulation models. Overlay and index methods have been developed because of limitations in process-based models and lack of monitoring data required for statistical methods (National Research Council, 1993). Despite common use at the regional scale, overlay and index methods do not have built in mechanisms to deal with uncertainties, nor do these models consider spatial and temporal variability of parameters affecting ground water vulnerability. Most models based on overlay and index methods use physiographic parameters and do not consider LULC and management.

Corwin, et.al., (1996) suggested that an integrated system of advanced information technologies such as Global Positioning System (GPS), Geographic Information systems (GIS), geostatistics, remote sensing, solute transport modeling, neural networks (NN), fuzzy logic, and uncertainty analysis could provide a framework from which real-time or simulated assessment of non point source (NPS) pollution can be made.

This research not only expanded modeling capability of an overlay and index models by addressing meaningful and relevant interactions of soil properties and LULC on ground water quality of watersheds in a karst region, but also used GPS, GIS, remote sensing, NN and fuzzy logic. The Neuro-fuzzy models developed in this study have the inherent capability to deal with uncertainties in the data, tolerate imprecision and can extract information from incomplete datasets. Expert knowledge, which is a valuable source of information on the physical, chemical and biological parameters that are hard to measure, as well as experimental information were used in this research project.

GOALS AND OBJECTIVES

The overall purpose of this research was to innovatively extend the capability of overlay and index methods of modeling ground water vulnerability to agricultural chemicals in northwest Arkansas by using Neuro-fuzzy techniques in a Geographic Information Systems (GIS) platform. The specific objective was to develop models using Neuro-fuzzy techniques with GIS to predict ground water vulnerability in a relatively large watershed in northwest Arkansas having mixed LULC, and variably permeable soils over the Boone Formation.

JUSTIFICATION FOR THIS RESEARCH

Development of a simple and flexible model that has inherent capabilities to deal with uncertainty and incomplete data sets in a GIS will be useful for delineation of ground water vulnerability at the regional scale. A Neuro-fuzzy system is a fuzzy system that is trained by a learning algorithm from NN theory.

Neuro-fuzzy modeling is an approach where the fusion of NN and Fuzzy Logic find their strengths. These two techniques complement each other. This approach employs heuristic learning strategies derived from the domain of NN theory to support the development of a fuzzy system. It is possible to completely map NN knowledge to Fuzzy Logic (Khan, 1999). A marriage between the NN and Fuzzy Logic techniques should help overcome the short comings of both techniques The fuzzy system provides a computational approach to human thinking and behavior (Zadeh, 1965). Nauck et. al. (1997) mentioned that experts can do their job without developing and using a mathematical model. Experts obtain experience by solving problems. The use of fuzzy systems allows experts to specify the actions or the relationships among the input parameters in the form of linguistic rules. These rules are translated into a framework of fuzzy set theory providing a calculus which can simulate the opinion and behavior of the expert (Nauck et al., 1997). The translation into fuzzy set theory is not formalized and arbitrary choices can be made while defining the shape of the membership functions. This posses uncertainties in the process of building a fuzzy system, however, these uncertainties mostly causes a minor system design problem that can be overcome by the heuristic tuning processes (Nauck et al., 1997). The major contribution of fuzzy set theory is that it tolerates imprecision of the real world. Unlike two valued logic, this approach uses multivalued logic to represent real world situations where ‘classes’ do not have sharp boundaries. It is designed to deal with properties that are represented as a matter of degree (Yager et al., 1987).

NN are systems that try to make use of some of the known or expected organizing principles of the human brain. The neural nets are composed of neurons and are connected by weights. NN can solve difficult problems by using a learning procedure that depends on the NN models and the given problem, but are black box (Nauck et al., 1997). NN are multi-input and multi-output nonlinear models and can represent the complex interactions among the input/output parameters (Fausett, 1994). NN outperform conventional statistical techniques in extracting information when datasets are diverse and relationships are poorly defined. Although in recent years NN has been successfully used in solving difficult hydrological and environmental problems, it is not possible to determine how the solution was found due to the inherent black box nature of the NN. Nauck et al., (1997) mentioned that it is also not possible to insert prior knowledge to a NN, however, in a Neuro-fuzzy system the combination of the learning ability of the NN and the linguistic rule handling of the fuzzy systems results in a self tuning system which is interpretable and in which prior knowledge can be inserted.

Simple and easily available parameters were used in the models in this research to ensure global application of the models. The model developed used soil properties such as hydrologic group, soil structure, and thickness of the soil horizons and LULC as input data. In selection of parameters it was assumed that since the underlying geology is the Boone Formation, which is highly fractured, the variability of water and contaminant transmitting properties of soils as well as attenuation processes of the overlying soils will govern the vulnerability of the ground water. Once the contaminant moves beyond the soil zone, it will eventually reach the ground water due to extensive presence of fractures, low consumption capabilities and the considerable amounts of water flowing in this humid region. Therefore, only soils and LULC related parameters were used in this research. However, in other regions without the Boone Formation, adoption of this modeling approach may need to include more geological inputs along with the soil properties and LULC, because geology may pose some resistance to contaminant transport.

The output from the Neuro-fuzzy model was displayed in the form of a map that shows regions of ground water in the Savoy Experimental Watershed (SEW) and its surrounding area having more or less potential vulnerability to NO_3-N. Also, a table was developed to present the areal extent of the vulnerability categories. Coincidence reports were generated indicating interactions among major model parameters.

The Neuro-fuzzy model predictions were compared with the water quality data sets (field data) to obtain information on relative suitability of the modeling techniques for predicting ground water vulnerability in the watershed. It should be noted that due to the point nature of the water quality data and inherent variability associated with the water quality data, a comparison of well data (point) and vulnerability maps (spatial) is not suitable for determining the best modeling approach in an absolute sense. However, the research showed how a Neuro-fuzzy technique can be used in a GIS to predict ground water vulnerability.

Methodologies employed in this project are applicable and readily transferable to other watersheds to delineate ground water vulnerability. This research served as a model for regional-scale assessments of ground water vulnerability in other states, which may reduce their monitoring costs and increase the accuracy of the prediction. For a more detailed discussion of the results the reader is referred to the work reported by Dixon (2001).

STUDY AREA

Location

The Neuro-fuzzy models were developed based on the characteristics of the four sub-basins of the Illinois River Watershed and SEW. The study area, which is located east of the Arkansas-Oklahoma border (Figure 1), is an intensely monitored watershed in northwest Arkansas. SEW is a University of Arkansas (U of A) property of approximately 1250 hectares located in the Illinois River watershed, 24 km west of the U of A campus in Fayetteville. The study area (watershed) has an area of about 109,000 ha (270,000 acres). An integrated research effort between the U of A, Arkansas Water Resources Center (AWRC), the U.S. Geological Survey (USGS), Arkansas Department of Environmental Quality (ADEQ), Agricultural Research Service (USDA-ARS), and the Natural Resources and Conservation Service (USDA-NRCS) has been established for the SEW.

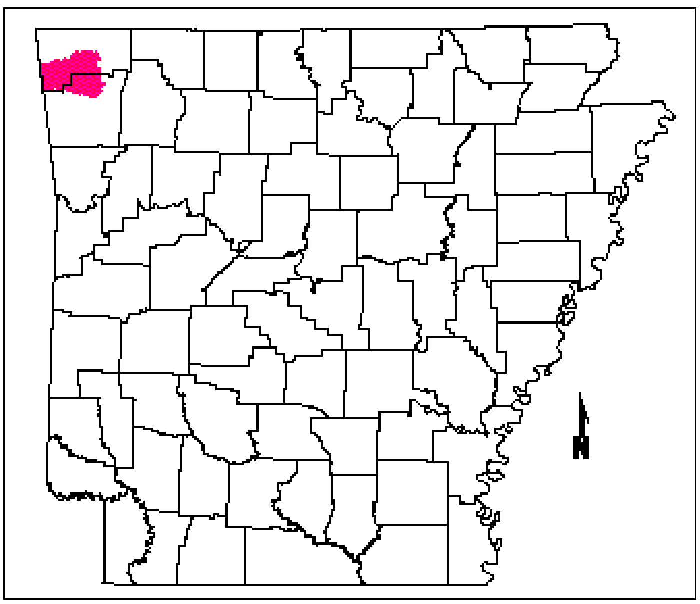

Figure 1. Location of the study area in northwest Arkansas.

Physical Characteristics of the Study Area

Sub-basins

The study area (watershed) has diverse soils, LULC and karst terrane and is affected by urban encroachment. There are four sub-basins in the study area. The names of the sub-basins and their 10 digit hydrologic units are: Osage Creek (1111010303), Clear Creek (1111010302), Flint Creek (1111010306) and Middle Illinois River (1111010305) (Figure 2). The Osage Creek sub-basin covers about 52% or 57,052 ha of the study area, followed by the Clear Creek sub-basin, which covers about 18% or 19,839 ha of the study area. The Flint Creek and Middle Illinois River sun-basins occupy about 16% (17,927 ha) and 13% (14,460 ha) of the study area, respectively.

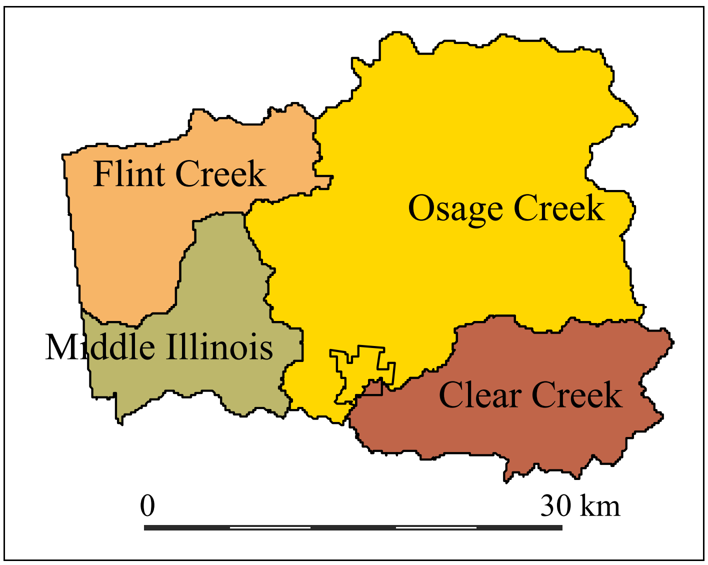

Figure 2. Location of the major watersheds and SEW in the study area.

<u>Soils</u>

The study area has 44 soil series. The two most dominant soil series of the study area are Nixa and Captina. These soils occupy about 21% and 18% of the study area, respectively (Table 1). Nixa soils are found mainly in the north and northeast part of the watershed whereas Captina soils are found more in the eastern part (Figure 3). Small patches of Captina soils are also found in the western part of the study area. In addition, Clarksville soils comprise about 16% of the study area and are found in the central part of the watershed. Peridge soils comprise about 2% of the study area and are found mainly along the streams valleys. A few patches of Peridge soil also occur in the eastern part of the watershed.

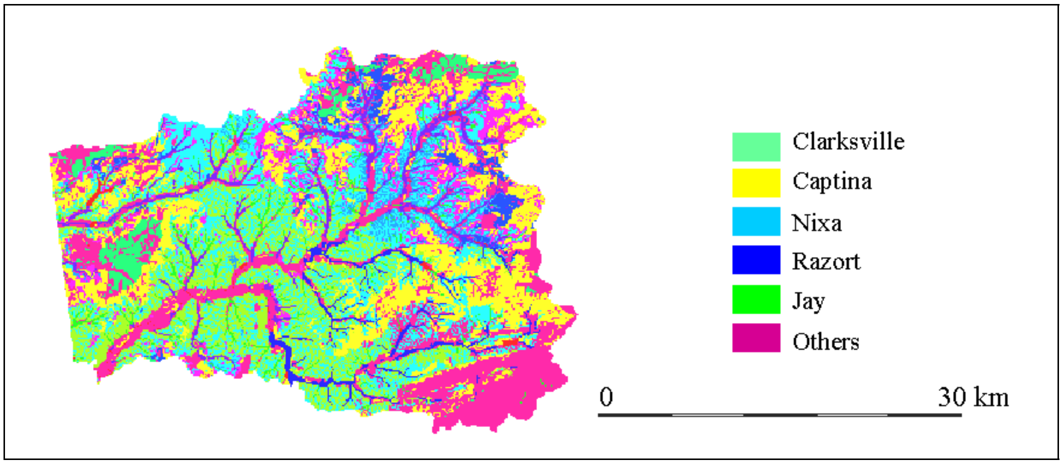

Figure 3. Spatial distribution of major soil series in the watershed.

Table 1. Areal distribution of the major soils in the watershed.

Major Series	acres	ha	%	Hydrologic Group	Pedality classes
Captina	49,684	20,107	18.4	C	Moderately high
Clarksville	42,903	17,363	15.9	B	Very high
Elsah	6,182	2,502	2.3	B	Moderately high
Jay	8,206	3,321	3.1	C	High
Nixa	56,476	22,856	20.9	C	High
Noark	7,064	2,859	2.6	B	Moderately high
Peridge	6,422	2,599	2.4	B	Low
Razort	6,029	2,440	2.2	B	Moderately high
Secesh	11,317	4,580	4.2	B	High
Tonti	18,774	7,596	6.9	C	Moderately High
Other soil series	55,759	22,565	20.7	B,C,D	
Water	1,186	480	0.4	N.A.	
Total	270,002	109,268	100		

<u>LULC</u>

The watershed is characterized by mixed LULC. Agriculture, particularly pasture, covers about 64% of the study area (Table 2). About 23% of the study area is covered by forests that are found in the central part of the watershed. Urban LULC, which covers about 10% of the study area, is found mainly in eastern part of the study area (Figure 4).

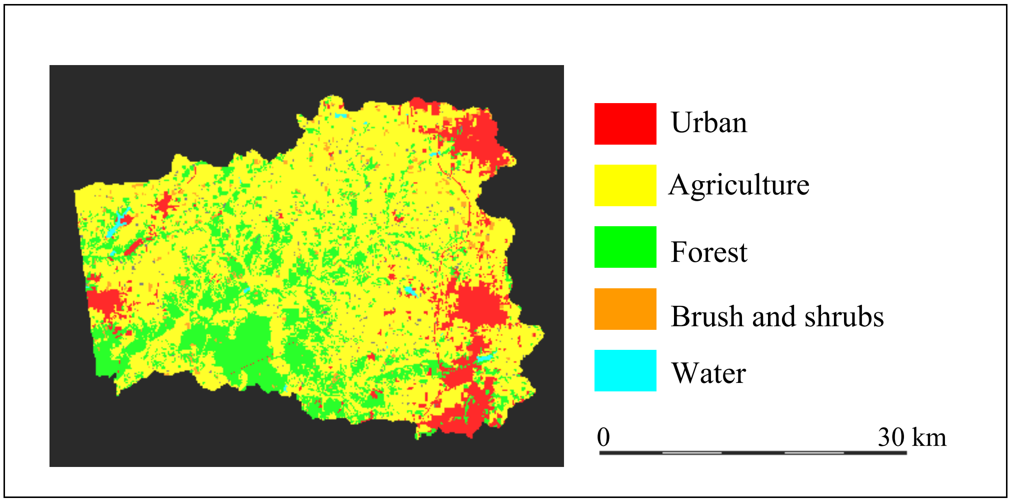

Figure 4. Spatial distribution of LULC in the watershed.

Table 2. Areal distribution of LULC in the watershed.

LULC	acres	ha	%
Urban	28,241	11,429	10.5
Agriculture	171,922	69,576	63.7
Shrubs and Brush	5,288	2,140	1.9
Forest	61,022	24,694	22.6
Water	1,120	454	0.4
Confined Animal Operation	2,409	975	0.9
Total	270,002	109,268	100

Geology

The major rock unit, which occupies 88% of the study area, is the Mississippian age Boone Formation (Figure 5). The Boone Formation is a limestone with varying amounts of densely interbedded chert ranging between 30 and 60% by volume. In NW Arkansas, the Boone Formation typically occurs between 300 – 350 feet or 91 – 106 m

(Croneis, 1930). The Upper Mississippian Formation, which is a combination of Pitkin limestone, Fayetteville shale and the Batesville sandstone, occupies about 8% of the study area (Table 3). This geological formation is found in patches all over the study area. The primary porosity of the Boone Formation is low but the secondary porosity of this formation is high due to the presence of numerous fractures (Curtis, 2000; Chitsazan, 1980; Razaie, 1979).

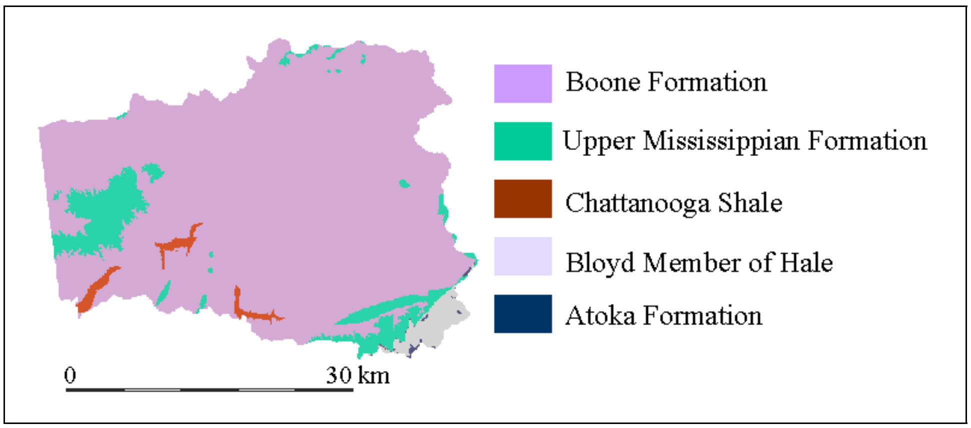

Figure 5. Spatial distribution of geology in the watershed.

Table 3. Areal distribution of geology in the watershed.

Geology	acres	ha	%
Atoka Formation	397	161	0.1
Bloyd Member of the Hale	5,806	2,350	2.2
Cane Hill Member of Hale	114	46	0
Upper Mississippian Formation	22,793	9,224	8.4
Boone Formation	237,159	95,976	87.8
Chattanooga Shale	3,733	1,511	1.5
Total	270,002	109,268	100

Slopes

The slopes in the watershed vary from 0 to > 31 degrees. The slopes were classified according to the scheme used by Hays (1995). About 39% of the study area has slopes that are classed as nearly level (0 – 2 degrees) and are found further away from the stream beds (Figure 6). Gently sloping and strongly sloping categories of slopes occupy about 27- and 23 % of the study area, respectively (Table 4).

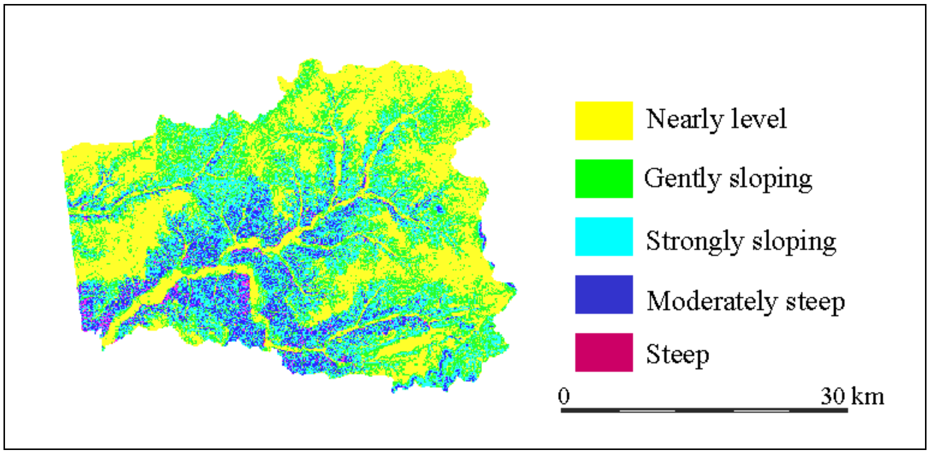

Figure 6. Spatial distribution of slope in the watershed.

Table 4. Areal distribution of slopes in the watershed.

Slopes (degrees)	acres	ha	%
Nearly level (0 - 2)	104,142	42,146	38.6
Gently sloping (3 - 4_	74,144	30,006	27.5
Strongly sloping (5 - 9)	62,334	25,226	23.1
Moderately steep (10 - 16)	25,224	10,208	9.3
Steep (17 - 30)	4,099	1,659	1.5
Very steep (> 31)	59	23	0
Total	270,002	109,268	100

<u>Elevation</u>

Elevation in the watershed ranges from 280 to 526 m. The higher elevations occur away from the stream valleys whereas the lower elevations occur along the stream valleys (Figure 7). About 19% of the study area has an elevation > 400 m (Table 5). Elevations ranging from 280 to 340 m occupy about 13% of the study area and are found along the stream beds.

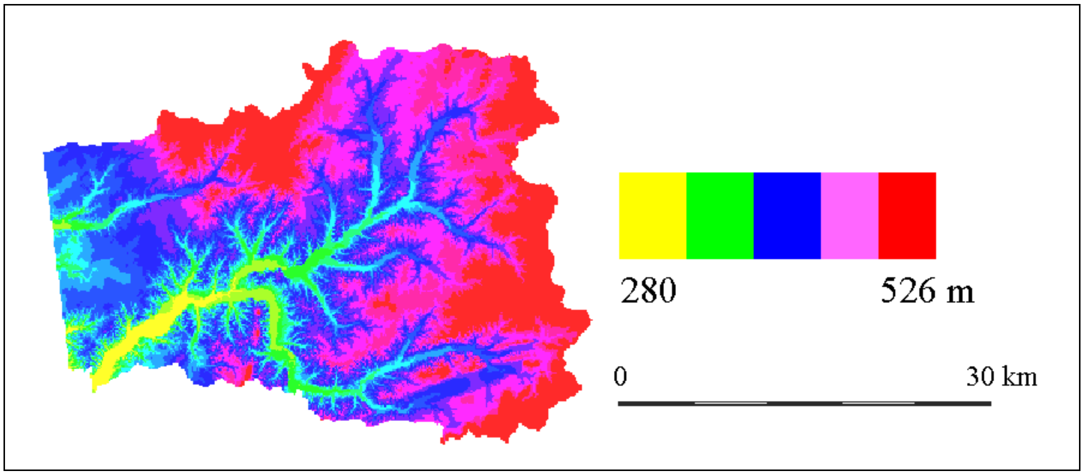

Figure 7. Spatial distribution of elevation in the watershed.

Table 5. Areal distribution of the elevation in the watershed.

Elevation (m)	acres	ha	%
280 - 299	3,973	1,608	1.5
300 - 310	4,440	1,797	1.6
311 - 320	6,212	2,512	2.3
321 - 330	8,157	3,301	3.1
331 - 340	11,786	4,770	4.4
341 - 350	19,731	7,985	7.3
351 - 360	29,336	11,872	10.9
361 - 370	29,729	12,031	11
371 - 380	34,102	13,801	12.6
381 - 390	40,067	16,216	14.8
391 - 400	31,962	12,935	11.8
> 400	50,507	20,440	18.7
Total	270,002	109,268	100

METHODOLOGY

Data Development

This study used both primary and secondary data layers as model inputs. The primary data layers were used to generate coincidence reports for the models and secondary data were used as inputs in the models (Tables 6 - 7). The primary data layers included (i) soil series, (ii) LULC, (iii) geology, and (iv) digital elevation models (DEMs). The secondary data layers included (i) soil hydrologic group, (ii) depth of the soil profile, and (iii) soil structure of the A horizon.

Table 6. Description of primary data layers.

Primary Data layers	Source	Scale/resolution	Comments
Watershed boundaries	NRCS	1:100,000	Digital
Location of Springs/wells	Field determined	N. A.	GPS
Water Quality data	Collected at SEW and surroundings	N. A.	ADEQ and AWRC Publication
Geology	Arkansas Geological Commission	1:24,000	Digital
Soils	NRCS* and Iowa State	1:24,000	Mylar for primary and Tabular for secondary attributes
LULC	Oklahoma State University (1985) and CAST* (1992)	1:24,000 30 m	Mylar Digital
DEMs	USGS	30 m	Digital

**CAST : Center for Advanced Spatial Technologies, University of Arkansas*

Table 7. Primary and secondary data layers and their use in the study.

Primary data	Secondary data	Model Inputs		Validation	Fine-Tune
		Preliminary	Final		
Soils	Hydrologic groups	✔	✔		
	Structure	✔	✔		
	Depth of the profile	✔	✔		
LULC	LULC	✔	✔		✔
DEMs	Slope				✔
	Elevation				
Location of wells/springs				✔	
Water quality				✔	

The Soil Survey Geographic Database (SSURGO) level soil maps were obtained from the Natural Resources Conservation Service (NRCS) and digitized in the Soil Physics Laboratory at the University of Arkansas. Tabular data for soil hydrologic groups were obtained from the SSURGO database for each soil map unit. Soil map units were reclassified into appropriate categories to generate maps for hydrologic groups (Table 1). The data layers for the soil map units were then reclassified into a series level map. This step was necessary because SSURGO level data did not contain information on depth of the profile and soil structure that are required by the Neuro-fuzzy models. Soils data for depth of the profile and soil structure were obtained from the Official Soil Series Description database of Iowa State (http://www.statlab.iastate.edu/soils/nsdaf/; viewed 6/16/00). The soil series map was reclassified to generate maps for soil structure and depth of the profile. Soil structure, specifically pedality, was classified according to the classification scheme of pedality points developed by Lin et, al. (1999) to indicate water transmitting properties of the soils. The depth of the profile was estimated by excluding Cr and R horizons. The GRASS (4.2) command 'r.reclass' were used for all reclassification routines. The secondary data layers used in the models are in the form of maps of (i) soil structure surface horizon (A), (ii) total depth of the horizon, (iii) hydrologic group and (iv) LULC. For more detailed discussion of the input data layers see Dixon (2001).

The water quality data were obtained from two different sources: ADEQ and AWRC. The water quality data provided by ADEQ were collected with respect to storm events during 1998 and 1999 for 24 different wells/springs. These data were analyzed by ADEQ lab personnel for about 40 different ions and compounds. In this study, NO_3-N is the water quality parameter used to compare the validity of the models. The discharge records and contamination level of NO_3-N data for springs and only contamination data for wells were stored in a relational table. Quality control and quality assurance (QC/QA) procedures were followed according to the internal guidelines of the agency. In addition, AWRC provided a set of historical data consisting of 20 wells (Smith and Steele, 1990) for the study area. These wells were sampled during the wet season of 1990 and analyzed for NO_3-N and other ions and compounds in the AWRC water quality lab. The addition of historical data added temporal variability as well as uncertainty in the data. However, the data from AWRC complemented the data set provided by ADEQ. All of the ADEQ wells/springs were clustered in one zone of the watershed whereas the historical data were not. In essence, although addition of historical data (Smith and Steele, 1990) added temporal variability in the data set it reduced the clustering of the locations of the wells if only the ADEQ data set were used.

Model Development

The software NEFCLASS-J developed by Nauck, et. al, (1999) was used in the study for Neuro-fuzzy techniques (Figure 8). This software was written in JAVA and the output function was customized to tie the model in a GIS. From the view point of the NEFCLASS architecture and the flow of data, the fuzzy sets were trained by a backpropagation-like algorithm (Nauck et. al., 1999). Backpropagation means errors from computation are propagated backwards through the architecture from the output units towards the input units. However, the NEFCLASS does not use the gradient descent approach common with NN.

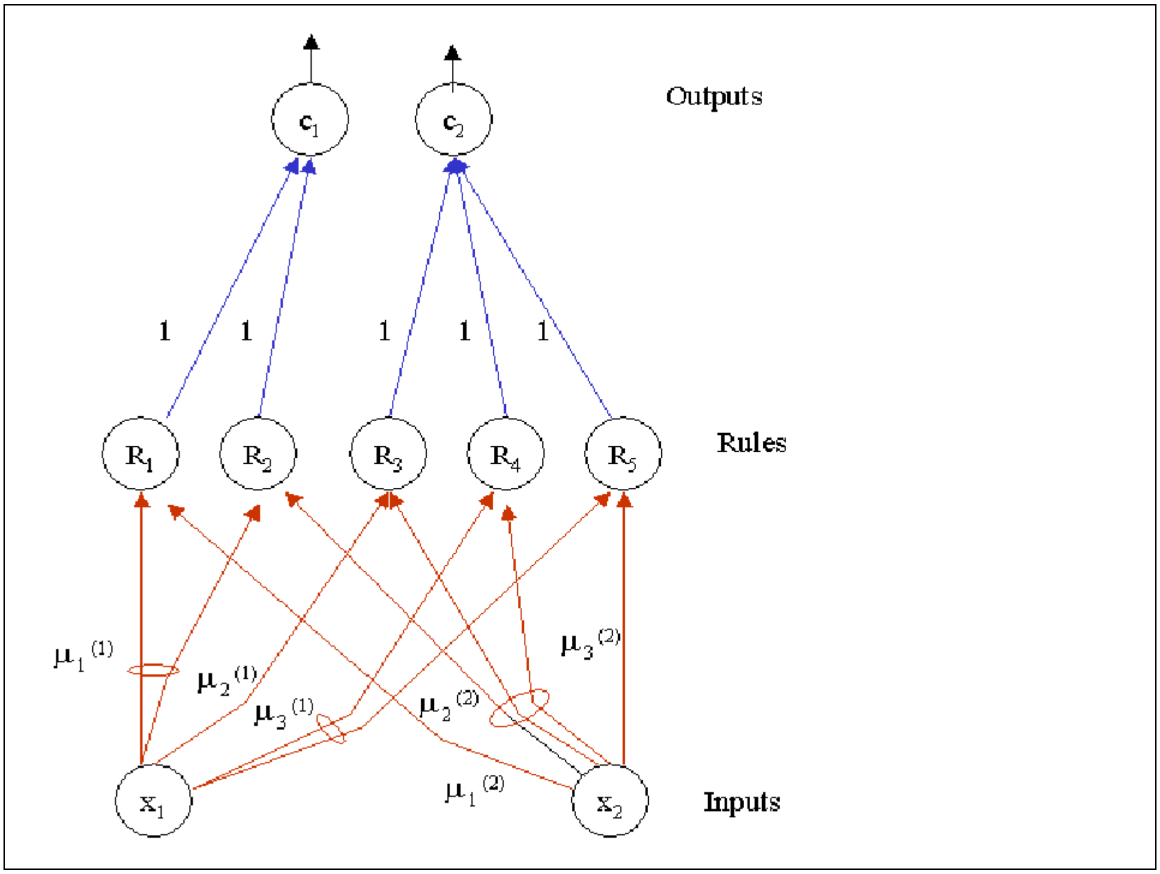

Figure 8. Schematics of the Neuro-fuzzy model used in this study, x=inputs, μ = weight, R = rules and C=output class.

For this research, NEFCLASS-J was used with trapezoidal membership functions. This parameterized membership function, which was chosen because of its simplicity, not only reduces system design time, but facilitates the automated tuning of the system by desired changes (Yen and Langari, 1998). The trapezoid membership function uses four parameters and the peak of the membership function is 1. The desired change of the membership function for the trapezoidal function can be obtained from the widening or narrowing of the membership function and the corresponding changes of the related parameters (Yen and Langari, 1998). This simple membership function facilitated an important principle underlying the theories of fuzzy logic: exploring cost effective approximate solutions. All training datasets were reclassified according to the following ground water vulnerability categories: high, moderately high, moderate and low potential. Application of Neuro-fuzzy techniques was a two step process. First, the training data set was used with the Neuro-fuzzy software to generate classifiers, fuzzy sets and rule bases. Second, once the software was trained, the application data set was used with the Neuro-fuzzy models. The outputs from the models then were used to generate maps.

The training data sets were obtained for the entire watershed from the GIS software GRASS command r.stats. Four input parameters to the command r.stats included the data sets used in the Neuro-fuzzy model, viz. soil hydrologic groups, LULC, depth of the profile and structure (pedality) of the soils as it indicates water transmitting properties. This GRASS command generated a table presenting all of the possible combinations of input parameters in the data sets. There were 202 patterns found for the entire watershed from the four input parameters. The output table from GRASS was imported in Dbase (IV) and the data were classified based on expert's opinion. The format of the input Dbase table is presented in Table 8. The first 4 columns represented input parameters such as soil hydrologic group, LULC, depth of the soil horizon and soil structure. The remaining four columns indicated the classifier. All of the training data sets were reclassified according to the high, moderately high, moderate and low potential for ground water vulnerability.

Table 8. Example of the classifier for potential vulnerability categories used in the training data sets.

Hydrologic Group	LULC	Depth*	Soil Structure**	High	Moderately High	Moderate	Low
10	10	15	38	1	0	0	0
10	10	18	38	1	0	0	0
10	10	36	53	1	0	0	0
10	10	60	38	0	1	0	0
10	10	66	34	0	1	0	0
10	10	72	20	0	0	1	0
10	10	72	38	0	1	0	0
10	10	78	34	0	1	0	0
10	10	97	34	0	0	1	0
10	50	72	20	0	0	0	1
10	50	78	34	0	0	0	1

*depth in inches
** pedality points

RESULTS AND DISCUSSION

Spatial Characteristics of the Model Inputs

This section includes discussion on characteristics of the four model input parameters derived from soil characteristics. All of the four model input soil parameters and LULC are plausible parameters that influence the transport of water and contaminants through the profile to the ground water.

LULC

The watershed is characterized by mixed LULC. Agriculture, particularly pasture, covers about 64% of the study area (Table 2). About 23% of the study area is covered by forests that are found in the central part of the watershed. Urban LULC, which covers about 10% of the study area, is found mainly in eastern part of the study area (Figure 4).

Soil Hydrologic Groups

The soils in the watershed were classified into three soil hydrologic groups B, C and D (Figure 9). About 54% of the land area in the watershed was in soil hydrologic group C (Table 9). Hydrologic group C indicates that Ksat is moderately low and internal free water occurrence is deeper than shallow (Soil Division Staff, 1993). Hydrologic group B covers about 40% of the watershed and occurs along the stream valleys. Soil hydrologic group B indicates Ksat is moderately high and free water occurrence is deep or very deep (Soil Division Staff, 1993). Hydrologic group D, which occurred in small patches across the watershed and indicates low Ksat value, occupies slightly more than 5% of the land area.

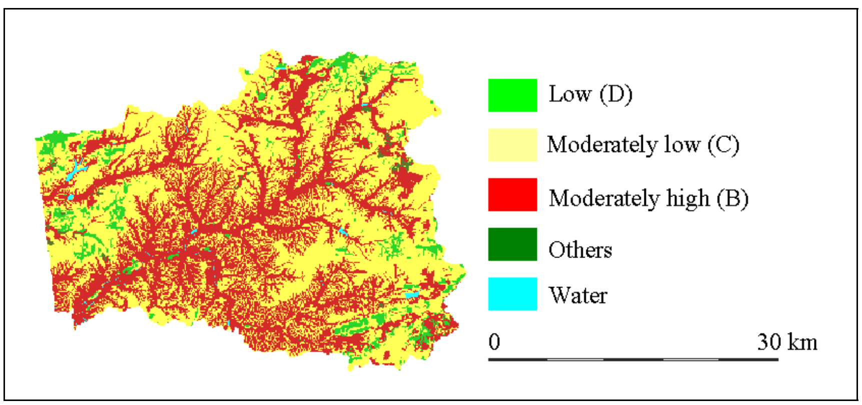

Figure 9. Spatial distribution of soil hydrologic groups in the watershed.

Table 9. Areal distribution of soil hydrologic groups in the watershed.

Soil Hydrologic Groups	acres	ha	%
Low (D)	14,118	5,713	5.2
Moderately low (C)	145,950	59,065	54
Moderately high (B)	106,869	43,249	39.6
Others	1,880	761	0.7
Water	1,185	480	0.5
Total	270,002	109,268	100

Depth of the Soil Profile

The depth of the soil profiles was estimated from the soil series description for the solum thickness excluding the Cr and R horizons. About 83% of the study area has deep or very deep soil profiles (Table 10). Deep soil profiles are found all over the watershed whereas very deep soils occur along the stream valleys (Figure 10). Moderately deep soils comprise 15% of the study area and occur along the stream valleys. Moderately shallow soils (31 – 50 inches) are found in small patches across the watershed and occupies about 1% of the study area.

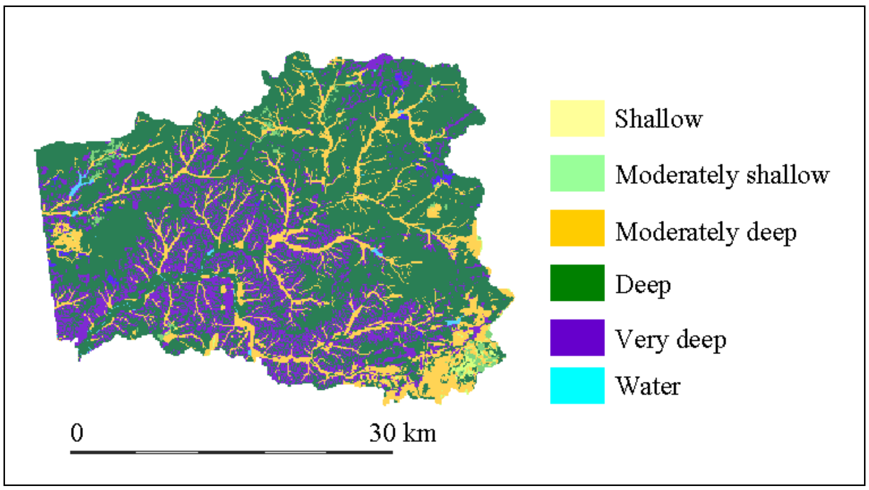

Figure 10. Spatial distribution of depth of the soil profile in the watershed.

Table 10. Areal distribution of depth of the profile in the watershed.

Depth (inches)	acres	ha	%
Shallow (9 - 30)	1,221	494	0.5
Moderately shallow (31 - 50)	2,920	1,180	1
Moderately deep (55 - 69)	39,355	15,927	15
Deep (70 – 85)	168,088	68,025	62
Very deep (> 85)	55,352	22401	21
Others	1,880	761	0.1
Water	1,186	480	0.4
Total	270,002	109,268	100

Soil Structure (pedality)

The structure (pedality) of the soil varies within the soil profile. For the Neuro-fuzzy model of the watershed, only the structural properties or pedality of the surface horizon (A) was used. The soil structure of the A horizon was reclassified according to the pedality points suggested by Lin et. al., (1999). The final pedality points were obtained for each soil by adding all of the points for ped size, ped shape and ped grade. Pedality points were regrouped to indicate water transmission potential in the profile. Five pedality groups were generated based on the total pedality points. For example, pedality point ranging from 10 – 17 was considered low (Table 11). Low pedality points are found in small patches across the watershed (Figure 11). High (40 – 50) and very high (>51) pedality points together occupy about 49% of the watershed and are found mainly in the central part of the watershed. Most of the high pedality points are associated with coarse textured soils found closer to the main stream bed of the watershed. Table 11 shows pedality points for the major soils in the study area.

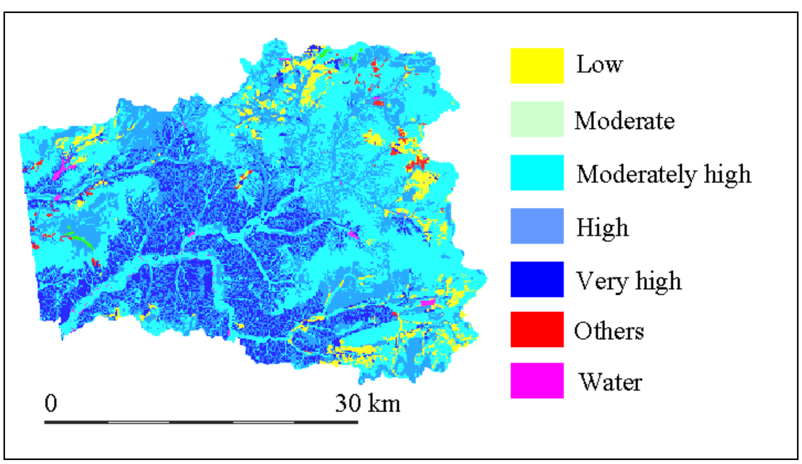

Figure 11. Spatial distribution of soil structure (pedality) in the watershed.

Table 11. Areal distribution of pedality point for soils in the watershed.

Soil Structure (Pedality Points)	acres	ha	%
Low (14 - 17)	14,201	5,747	5.2
Moderate (20 - 30)	507	205	0.2
Moderately high (31 - 40)	117,758	47,656	43.6
High (40 - 50)	89,991	36,419	33.3
Very high (> 51)	44,480	18,001	16.5
Water	1,186	480	0.5
Others	1,879	760	0.7
Total	270,002	109,268	100

Coincidence Analyses for Model Parameters

Coincidence reports were used to tabulate the mutual occurrence of the categories for any two raster map layers with respect to each other. Nine sets of coincidence reports were generated. The coincidence reports facilitated training of the classifier, formulation and interpretation of the rule bases and the explanation of the fuzzy sets and classifier used in the study. Not only were the input parameters used in the coincidence analysis, but the 'fine tuning' parameter such as slope was also used in the coincidence analyses.

Depth of the Soil Profile and Slope

Depth of the soil profile and slopes are important parameters, since depth of the profile influences attenuation processes and slope influences runoff and infiltration of water. Coincidence analyses between depth of the soil profiles and slope help fine tune the classifier. For example, for the Neuro-fuzzy models operators classify the training data set according to the expert's opinion. This classification process involved assigning values to the outcome of a set of parameters based on the underlying assumptions. One underlying assumption of using depth of the soil profile as a plausible parameter is that the deeper the profile, the lower the ground water vulnerability or conversely the shallower the profile, the higher the ground water vulnerability. However, if the watershed is characterized by shallow profiles with steep slopes, then runoff could be a more dominating factor than infiltration, hence chances of the underlying ground water contamination should be lower. On the other hand, if a deep soil profile coincides with nearly level slopes, then chances of water infiltrating and transmitting through the profile may be higher, hence the potential contamination of the ground water will increase. If the operator does not include this situation in the training data set while assigning values, then the classification results will be less than desirable.

In this watershed it was found that a higher proportion of the deep soils coincided with nearly level slopes followed by gentle slopes. About 31,456- and 20,431 ha of deep soil profiles coincided with nearly level and gentle slopes, respectively (Table 12). Only about 387 and 639 ha of moderately shallow soils coincided with nearly level and gentle slopes, respectively. About 241 ha of shallow soil profiles coincided with strong slopes. These coincidence patterns were included in the training data set for the classifier.

Table 12. Mutual occurrence of slopes and depth of the soil profiles in the watershed. Results are given in ha.

	Slopes						
Depth	Nearly level	Gentle	Strong	Moderate	Steep	Very steep	Total
Shallow	55	161	241	29	8	0	494
Moderately shallow	387	639	151	4	0	0	1,181
Moderately deep	5,737	5,147	3,572	1,288	182	0	15,926
Deep	31,456	20,431	12,789	3,064	284	2	68,026
Very deep	3,768	3,247	8,382	5,803	1,179	22	22,401
Others	370	324	55	8	3	0	760
Water	372	56	36	12	4	0	480
Total	42,145	30,005	25,226	10,208	1,660	24	109,268

Depth of the Profile and LULC

As mentioned earlier, training data for the classifier requires experts' opinion. The initial assumption for the ground water contamination model is that the ground water in agricultural areas are more vulnerable to NO_3-N than in forested areas. When agricultural (pasture) areas are associated with shallow to moderately shallow soil profiles, chances of ground water contamination are much higher than where the agricultural land is associated with very deep soils. Training of the classifier should have examples of all possible combinations of the input parameters to yield a reliable vulnerability map.

In the watershed a higher proportion of agricultural land (pasture) was associated with deep soils (47,818 ha) followed by moderately deep soils (10,641 ha). About 951 ha of pasture was associated with moderately shallow soils (Table 13). A considerable proportion of the forest LULC coincided with very deep soils (11,521 ha) and deep soils (9,926 ha). Only about 2,906 ha of the forests coincided with moderately deep soils.

Table 13. Mutual occurrence of depth of the profile and LULC in the watershed. The results are given in ha.

	LULC						
Depth	**Urban**	**Agriculture**	**Shrubs**	**Forest**	**Water**	**Confined Animal**	**Total**
Shallow	104	198	15	176	0	1	494
Moderately shallow	125	951	30	70	0	6	1,182
Moderately deep	2,049	10,642	205	2,906	42	84	15,928
Deep	7,854	47,818	1,592	9,926	42	791	68,023
Very deep	1,129	9,361	275	11,521	38	76	22,400
Others	151	539	23	30	0	18	761
Water	18	66	0	67	329	0	480
Total	11,430	69,575	2,140	24,696	451	976	109,268

Depth of the Soil Profile and Hydrologic Groups

The majority of the land area in soil hydrologic group B coincided with very deep soils followed by moderately deep soils (Table 14). About 54,358 ha of deep soils coincided with hydrologic group C. Soil hydrologic group D coincided with both deep and very deep soils. About 494 ha of hydrologic group D also coincided with a shallow soil profile.

The coincidence report between depth of the soil profiles and soil hydrologic groups helps fine tune the classifier. For example, for the Neuro-fuzzy models, operators need to classify the training data set. The initial underlying assumptions used in the training data were that a combination of shallow soil profile and hydrologic group B indicates a greater potential for ground water contamination than a combination of very deep soil profile and soil hydrologic group D. However, if the majority of the watershed had deep and very deep soil profiles that coincided with hydrologic group B, then the potential for ground water contamination will be somewhere between the previous assumption. The classifier for the Neuro-fuzzy models should accommodate this situation, otherwise, classification results will be less than appropriate.

Table 14. Mutual occurrence of soil profile depth and soil hydrologic groups in the watershed. The results are given in ha.

	Hydrologic Group					
Depth	**D**	**C**	**B**	**Others**	**Water**	**Total**
Shallow	494	0	0	0	0	494
Moderately shallow	18	0	1,164	0	0	1,182
Moderately deep	334	2,993	12,601	0	0	15,928
Deep	3,236	54,358	10,429	0	0	68,023
Very deep	1,631	1,713	19,056	0	0	22,400
Others	0	0	0	761	0	761
Water	0	0	0	0	480	480
Total	5,713	59,064	43,250	761	480	109,268

Soil Structure and LULC

The plausible parameters, used in this study, are influential in the transmission of water and contaminants through the profile. During the development of the training data set for the classifier it was assumed that coincidence of high pedality points and agricultural activity (pasture) would increase the chance of ground water contamination several times more than the coincidence between high pedality points and forests. About 34,525 ha and 23,335 ha of land in agriculture coincided with moderately high and high pedality points, respectively (Table 15). Significant amounts of forested land coincided with both very high and high pedality points (Table 15). About 7,357 ha of urban LULC is also associated with moderately high pedality points.

Table 15. Mutual occurrence of LULC and soil structure in the watershed. The results are given in ha.

Pedality Points	LULC Urban	Agriculture	Shrubs	Forest	Water	Confined Operation	Total
Low (14 - 17)	685	4,601	197	203	8	53	5,747
Moderate (20 - 30)	20	166	14	4	0	0	204
Moderately high (31 - 40)	7,357	34,525	934	4,324	37	476	47,653
High (40 - 50)	2,799	23,335	768	9,101	44	374	36,421
Very high (> 51)	400	6,344	201	10,967	35	55	18,002
Water	18	66	0	67	329	0	480
Others	151	539	23	30	0	18	761
Total	11,430	69,576	2,137	24,696	453	976	109,268

Depth of the Soil Profile and Soil Structure (Pedality Points)

The initial assumptions used for the classifier were as following: deep soil profiles and low pedality points indicate low contamination potential and high pedality points and shallow profiles indicate higher ground water contamination potential. In the watershed soils with deep profiles coincided with moderately high (37,627 ha) and high pedality points (27,117 ha). About 3,076 ha of low pedality points coincided with deep soil profiles (Table 16). Soils with very high pedality points coincided with very deep soils (17,690 ha). These coincidence patterns were included in the classifier.

Table 16. Mutual occurrence of soil structure (pedality points) and depth of the profile in the water shed. The results are given in ha.

	Pedality Points							
Depth (inches)	**Low**	**Moderate**	**Moderately high**	**High**	**Very high**	**Water**	**Others**	**Total**
Shallow (9 - 30)	0	0	494	0	0	0	0	494
Moderately shallow (31 - 50)	1,164	0	0	0	18	0	0	1,182
Moderately deep (55 - 69)	1,499	0	6,189	7,947	291	0	0	15,926
Deep (70 - 85)	3,076	205	37,627	27,117	0	0	0	68,025
Very deep (> 85)	9	0	3,345	1,356	17,690	0	0	22,400
Others	0	0	0	0	0	0	761	761
Water	0	0	0	0	0	480	0	480
Total	5,748	205	47,655	36,420	17,999	480	761	109,268

Soil Structure (Pedality Points) and Soil Hydrologic Groups

The underlying assumptions for the training classifier was that a coincidence between higher pedality points and soil hydrologic group B will increase potential for water and contaminant transport. Whereas a combination of soil hydrologic group D and low pedality points will transmit less water when all other factors are assumed to be same. Any variation of these combinations should be accommodated in the training data set for the classifier.

The mutual occurrence between soil structure (pedality points) and soil hydrologic groups in the watershed is presented in Table 17. Soil hydrologic group C coincided with moderately high and high pedality points. About 17,983- and 11,924 ha of soils with very high and moderately high pedality categories, respectively coincided with soil hydrologic group B. Only about 476- and 5,270 ha of the low pedality category coincided with hydrologic groups C and B, respectively.

Table 17. Mutual occurrence of soil hydrologic groups and pedality points in the watershed. The results are given in ha.

	Soil hydrologic groups					
Pedality Points	**D**	**C**	**B**	**Others**	**Water**	**Total**
Low (14 - 17)	0	476	5,270	0	0	5,746
Moderate (20 – 30)	205	0	0	0	0	205
Moderately high (31 - 40)	5,491	30,241	11,924	0	0	47,656
High (40 - 50)	0	28,347	8,072	0	0	36,419
Very high (> 51)	18	0	17,983	0	0	18,001
Water	0	0	0	0	480	480
Others	0	0	0	761	0	761
Total	5,714	59,064	43,249	761	480	109,268

Slope and LULC

For the training classifier, initially it was assumed that the agricultural (pasture) LULC will increase the chance of contamination more than any other LULC since fertilizers and animal wastes are directly applied to these fields. However, application of contaminant is not the only factor responsible for transport of the contaminant to the ground water. Water transports contaminants to the ground water. Coincidence between LULC and slope provided additional important information on the relation between agricultural land and the slope category. Agricultural land on steeper slopes should transport less contaminants down the profile than the agricultural land on level slopes. When all other factors are the same, the steeper the slope, the higher the runoff, hence, less infiltration and contaminant transport. Coincidence analysis provided important insights into the slope -LULC relation in the watershed.

The majority of the agricultural land in the watershed is associated with nearly level (0-2 degrees) to gentle (3 – 4 degrees) slopes. Therefore these areas are more vulnerable than the areas with strong slopes. About 13,261 ha of agricultural land is associated with strong slopes of 5 – 9 degrees. The majority of the forests coincided with strong slopes followed by moderately steep slopes (Table 18). About 6,306 ha of urban LULC coincided with nearly level slopes. Confined animal operations in the watershed coincided with nearly level slopes (418 ha) followed by gentle slopes (335 ha).

Table 18. Mutual occurrence of slopes and LULC in the watershed. The results are given in ha.

	Slopes						
LULC	Nearly level	Gentle	Strong	Moderate	Steep	Very steep	Total
Urban	6,306	3,212	1,584	313	15	0	11,430
Agriculture	31,528	21,596	13,261	2,901	290	1	69,577
Shrubs	1,087	544	374	122	12	0	2,139
Forest	2,484	4,266	9,765	6,817	1,340	23	24,695
Water	325	52	51	24	1	0	453
Confined operation	418	335	190	30	1	0	974
Total	42,148	30,005	25,225	10,207	1,659	24	109,268

Slope and Hydrologic Groups

Coincidence between soil hydrologic groups and slopes provided valuable information to train the classifier. Theoretically, soil hydrologic group B should transmit more water through the profile than soil hydrologic group D. However, if more land with soil hydrologic group B coincides with relatively steep slopes then, these areas might not be as vulnerable as they were assumed to be. Thus, these scenarios needed to be incorporated in the interpretation of the results from the classifier.

The majority of the land area in soil hydrologic group C coincided with nearly level slopes (0 – 2 degrees). About 18,413 ha and 11,457 ha of land with hydrologic group C coincided with gentle and strong slopes, respectively. Almost equal land area with soil hydrologic group B coincided with nearly level and gentle slope, respectively. The majority of the soil hydrologic group D coincided with nearly level slopes (Table 19).

Table 19. Mutual occurrence of slopes and soil hydrologic units in the watershed. The results are given in ha.

	Soil hydrologic groups					
Slopes (degrees)	**D**	**C**	**B**	**Others**	**Water**	**Total**
Nearly level (0 - 2)	4,781	25,928	10,694	370	372	42,145
Gently sloping (3 - 4_	628	18,413	10,584	324	56	30,005
Strongly sloping (5 - 9)	266	11,457	13,413	55	36	25,227
Moderately steep (10 - 16)	31	3,031	7,126	8	12	10,208
Steep (17 - 30)	8	235	1,409	3	4	1,659
Very steep (> 31)	0	1	23	0	0	24
Total	5,714	59,065	43,249	760	480	109,268

<u>Soil Hydrologic Group and LULC</u>

For the training data, initially it was assumed that wherever land area with soil hydrologic group B coincided with agriculture, chances of ground water contamination would be higher than the coincidence between land area with soil hydrologic group D and agricultural land use. The coincidence report provided insights into the relation between soil hydrologic groups and LULC in the watershed and was used in the interpretation of the classifier. The majority of the land area in agriculture (40,178 ha) coincided with soil hydrologic group C (Table 20). About 24,807 ha of agriculture coincided with hydrologic group B. a greater proportion of forested land coincided with hydrologic group B (15, 480 ha) followed by hydrologic group C (8,782 ha). About 3,986 ha of agricultural land coincided with category D.

Table 20. Mutual occurrence of soil hydrologic group and LULC in the watershed. The results are given in ha.

	Soil Hydrologic Group					
LULC	D	C	B	Others	Water	Total
Urban	1,233	8,013	2,015	151	18	11,430
Agriculture	3,986	40,178	24,807	539	66	69,576
Shrubs	139	1,356	621	23	0	2,139
Forest	334	8,782	15,480	30	67	24,693
Water	2	20	103	0	329	454
Confined Animal Operation	20	716	222	18	0	976
Total	5,714	59,065	43,248	761	480	109,268

Neuro-fuzzy Model

The Neuro-fuzzy model was developed using the trapezoidal membership functions. A dataset consisting of 202 combinations of patterns from input data was used to train the net. These patterns are also referred to as ‘cases’. The application data set for the watershed consisted of 2,662,528 rows and four columns of input data consisting of hydrologic groups, LULC, depth of the profile and soil structure for the study area. Four fuzzy sets were developed for each input parameter.

Characteristics of the Neuro-fuzzy Model

The parameter settings used for the model are presented in the Table 21. The software NEFCLASS-J developed by Nauck (1999) was used for this study.

Table 21. The parameter settings for the Neuro-fuzzy model.

Parameters	Settings
Training data file	BasinsCLASSIF.dat
Number of fuzzy sets	4
Type of fuzzy sets	Trapezoidal
Aggregation function	Maximum
Interpretation of classification results	Winner takes all (WTA)
Size of rule base	Automatically determined
Learning rule procedure	Best per class
Fuzzy sets constraints	(i) Keep relative order (ii) Always overlap
Rule weights	Not used
Learning rate	0.1
Validation	Single test (50%). 50% of the data withheld from the training
Stop Control	Maximum number of epochs = 100 Minimum number of epochs = 0 Number of epochs after optimum = 10 Admissible classification errors = 0

The training data set was composed of 202 rows. Out of 202 rows 46 cases were classified as class 1 (high), 72 cases as class 2 (moderately high), 65 cases as class 3 (moderate) and 19 cases as class 4 (low). The validation technique ‘single test’ which randomly divides the data into two sets according to a given percentage value, was used to develop the model. The smaller part is used for training and larger part is used for validation. The training process used 49% and the validation process used 51% of all cases presented in the data sets. A total of 41 possible rules were found. The optimal consequents were determined. ‘Best per class’ rule learning strategy was used. The maximum numbers of rules were determined automatically. ‘Best per class’ was used for trimming the rule base. ‘Best per class’ option selects under the constraints of the size of the rule base the best per class. A final rule base with 29 rules was created. This rule base covered all patterns. Performance on training and validation data are presented in Tables 22 - 24. Table 22 indicates that about 15% of the training data with high category coincided with high category, about 24% of moderately high category coincided with moderately high category, 21% of moderate category coincided with moderate and 5% of the low category coincided with low. For the training data sets, correct classification was 65 (65%) and number of misclassified entries were 35 (35%). For the validation data sets (49% of the training sets), the correct classification was 41 (40%) and number of misclassified entries were 61 (59%). Examples of the rule bases and fuzzy sets used in this study are presented in Appendix A. Statistical characteristics and correlation analysis of the training data sets are presented in Tables 25 and 26. Tables 27 and 28 presented statistical characteristics and correlation analysis of application data, which consisted of 2,660,864 rows.

Table 22 . Performance of the training data (%) for ground water vulnerability classes.

Vulnerability classes	High	Moderately high	Moderate	Low	Not classified	Total
High	*15*	8	0	0	0	23
Moderately high	5	*24*	7	0	0	36
Moderate	0	5	*21*	2	4	32
Low	0	2	2	*5*	0	0
Total	20	39	30	7	4	100

Table 23. Performance of the validation data (%) for ground water vulnerability classes.

Vulnerability classes	High	Moderately high	Moderate	Low	Non Classified	Total
High	10 (9.80%)	8 (7.84%)	1 (0.98%)	0 (0.00%)	4 (3.92%)	23 (22.55%)
Moderately high	4 (3.92%)	19 (18.63%)	4 (3.92%)	3 (2.94%)	6 (5.88%)	36 (35.29%)
Moderate	2 (1.96%)	8 (7.84%)	11 (10.78%)	1 (0.98%)	11 (10.78%)	33 (32.35%)
Low	0 (0.00%)	4 (3.92%)	5 (4.90%)	1 (0.98%)	0 (0.00%)	10 (9.80%)
Total	16 (15.69%)	39 (38.24%)	21 (20.59%)	5 (4.90%)	21 (20.59%)	102 (100.00%)

Table 24. Characteristics of the training sets.

Learning Procedure	Patterns	Misclassification	Errors
Training	100	23	55
Validation	102	61	79

Table 25. Statistics for training data.

Input Variables	mean	std. deviation	minimum	maximum	missing
Var 1 Hydrologic Group	24.55	9.23	10	50	0
Var 2 LULC	32.77	16.98	10	60	0
Var 3 Depth	74.09	26.43	9	151	0
Var 4 Sructure	38.17	11.43	14	53	0

Table 26. Correlation for training data.

Input variables	1	2	3	4	Class
Hydrologic groups (1)	1	0.05	*0.48*	0.23	0.26
LULC (2)		1	0.05	-0.01	-0.21
Depth (3)			1	0.23	0.47
Structure (4)				1	-0.28

Table 27. Statistics for application data.

Input Variables	mean	std. deviation	minimum	maximum	missing
Var 1 Hydrologic Group	10.81	12.52	0	50	0
Var 2 LULC	11.01	13.85	0	60	0
Var 3 Depth	35.5	40.05	0	151	0
Var 4 Structure	19.13	21.91	0	53	0

Table 28. Correlation for application data.

Input variables	var 1	var 2	var 3	var 4	class
var 1 Hydrologic group	1	0.86	0.94	0.92	0
var 2 LULC		1	0.86	0.88	0
var 3 Depth			1	0.95	0
var 4 Structure				1	0

Spatial Distribution of Neuro-fuzzy Model

The spatial distribution of ground water vulnerability predicted from the Neuro-fuzzy model using the four input parameters, soil hydrologic groups, LULC, depth of the soil profile and soil structure (pedality points) is shown in Figure 12 and summarized in Table 29. The high ground water vulnerability coincided with regions of hydrologic group B, agricultural land use, deep soils and high soil structure or pedality points (Appendix B). Moderately high ground water vulnerability categories comprised 25% of the watershed and was distributed across watershed. The moderate vulnerability category coincided with urban LULC and soil hydrologic group C.

Non-classified categories occurred where the rule base and input data contradicted each other. An example of contradictory information was the occurrence of soil hydrologic group C with high pedality points. Details of contradictory information are presented in Appendix B. Fine tuning of the training data sets was required to improve the model's prediction from contradictory data. About 13% of the watershed area was not classified by the preliminary model. This could be attributed to the validation technique used in the model, i.e. the single test. The results of the learning processes provide information on the smallest error caused during all propagations through the classifier and the error of the training data of the same cycle. This does not cross-validate error during the training processes.

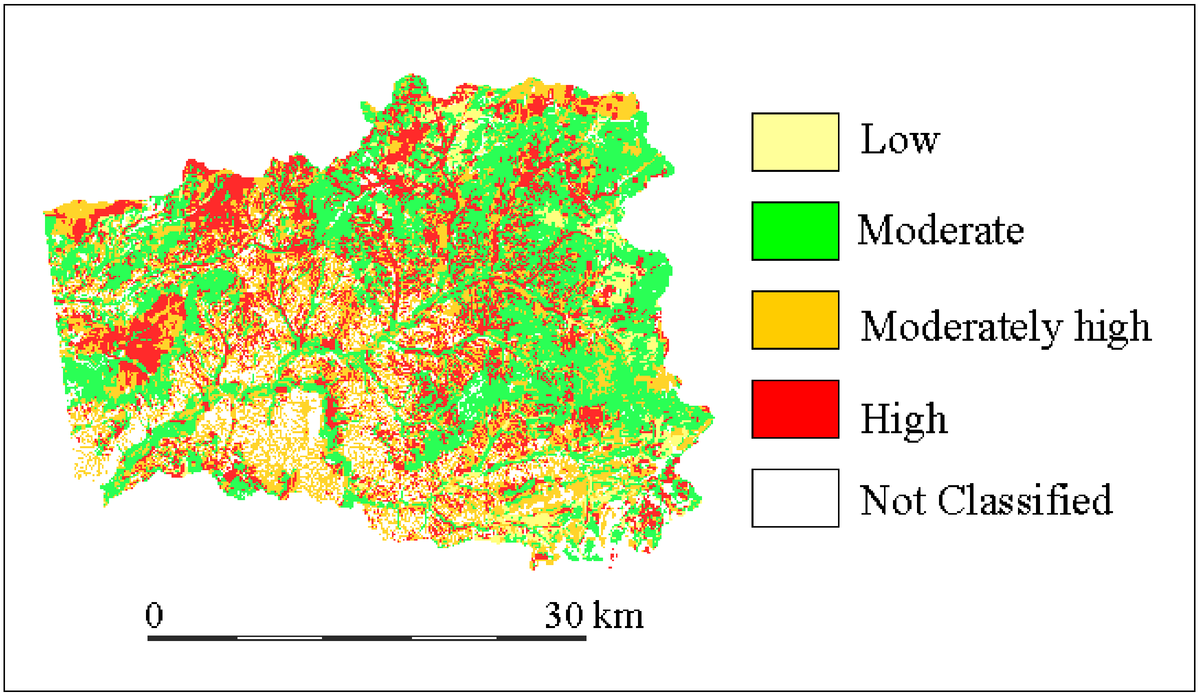

Figure 12. Spatial distribution of ground water vulnerability from the Neuro-fuzzy models in the watershed.

Table 29. Areal distribution of ground water vulnerability in the watershed.

Vulnerability	acres	ha	%
Non Classified (0)	37,375	15,125	13.8
High (1)	59,617	24,127	22
Moderately high (2)	68,975	27,914	25.8
Moderate (3)	93,277	37,748	34.5
Low (4)	10,758	4,354	3.9
Total	270,002	109,268	100

Coincidence Reports for the Neuro-fuzzy Model

Coincidence reports provided information on the mutual occurrence of the ground water vulnerability categories and the physical characteristics of the watershed. Seven sets of coincidence reports were prepared and are shown in Figures 13 – 19. A higher proportion of the highly vulnerable categories coincided with the soil hydrologic group C. Theoretically, highly vulnerable areas were expected to coincide with soil hydrologic group B. However, in this watershed almost equal areas of soil hydrologic groups B and C coincided with moderately high vulnerability category (Figure 13).

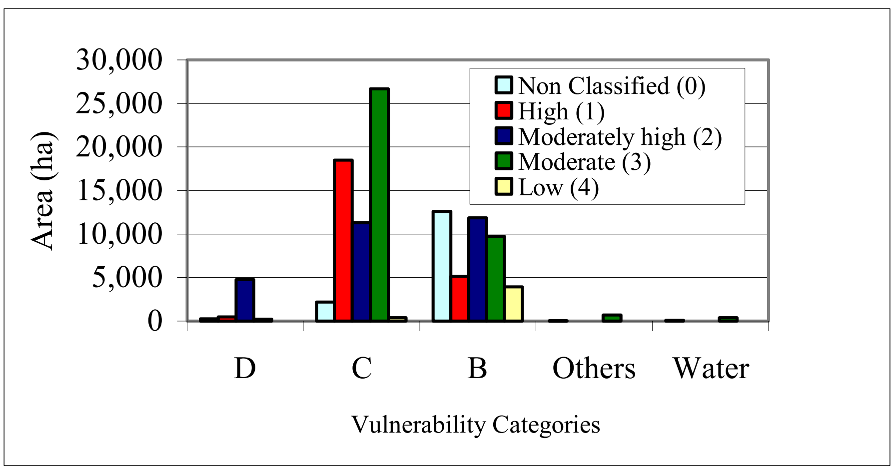

Figure 13. Mutual occurrence of soil hydrologic groups and Neuro-fuzzy-based ground water vulnerability categories.

As expected, a higher proportion of the highly vulnerable areas coincided with agricultural landuse. About 32,000 ha of agricultural land also coincided with moderately vulnerable categories (Figure 14).

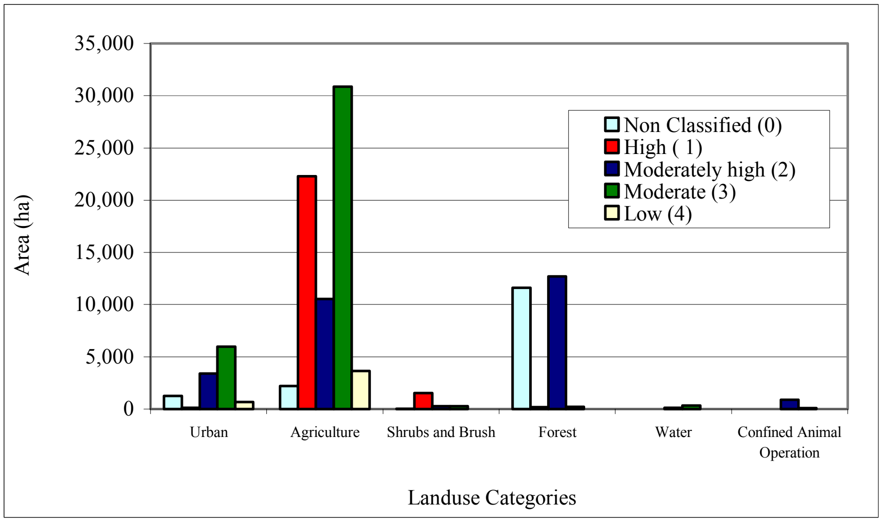

Figure 14. Mutual occurrence of LULC and Neuro-fuzzy-based ground water vulnerability categories.

A higher proportion of soils with deep profiles coincided with moderately high ground water vulnerable areas. About 10,000 ha of the very deep soils coincided with moderately high vulnerability categories. About 32,000 ha of the moderate vulnerability categories also coincided with deep soils (Figure 15).

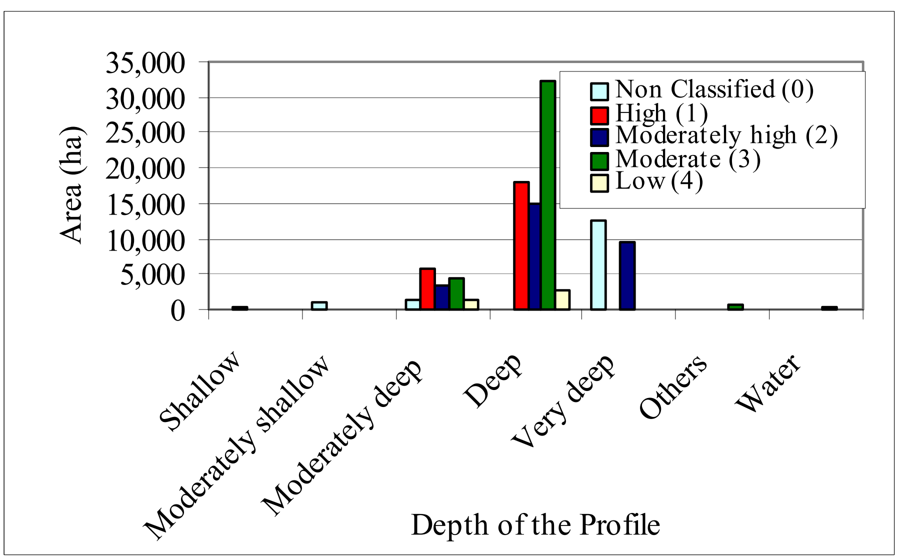

Figure 15. Mutual occurrence of ground water vulnerability and depth of the soil profile.

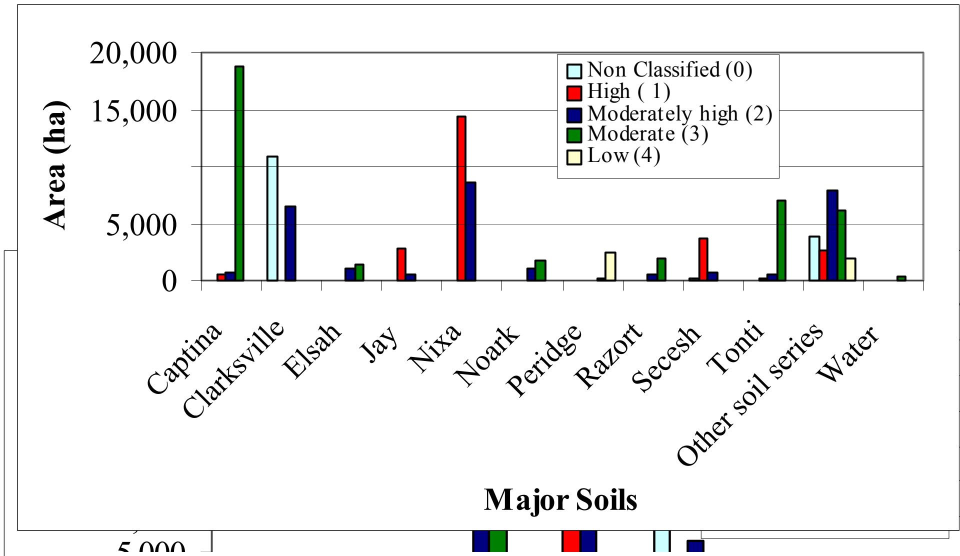

Figure 17. Mutual occurrence of major soils and vulnerability categories.

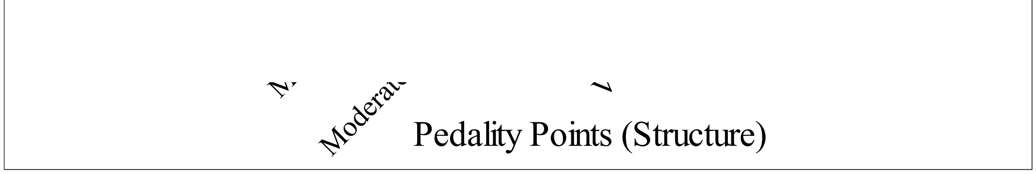

Figure 16. Mutual occurrence of pedality points (soil structure) and ground water vulnerability categories.

Almost all of the high ground water vulnerability category coincided with soil structure or high pedality points with high water transmitting capabilities through the profiles (Figure 16). The majority of the moderately vulnerable categories coincided with moderately high pedality points. Almost equal moderately high vulnerability category coincided with high and moderately high pedality points.

Figure 17 shows mutual occurrence between dominant soils and ground water vulnerability categories in the watershed. The locations of the Nixa soils coincided with high or moderately high vulnerability categories. Almost all of Captina soil area was classified as moderately vulnerable. About 6,000 ha of the Clarksville soil area was classified as moderately high (Figure 17).

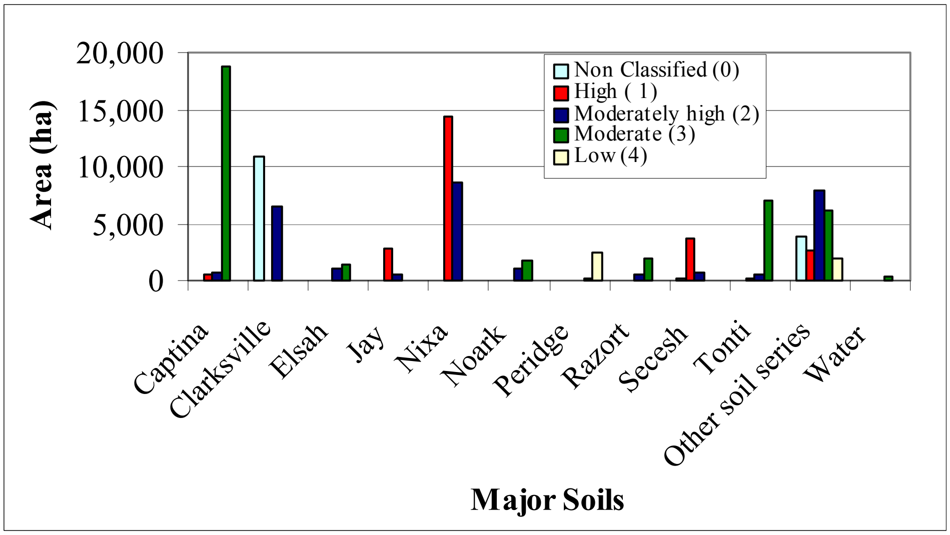

Figure 17. Mutual occurrence of major soils and vulnerability categories.

Almost equal area with the highly vulnerable category coincided with nearly level, gentle and strong slopes (Figure 18). The nearly level slope category also coincided with moderately vulnerable category. Areas with a combination of high vulnerability and nearly level slope have greater contamination potential than areas with high vulnerability but strong slopes.

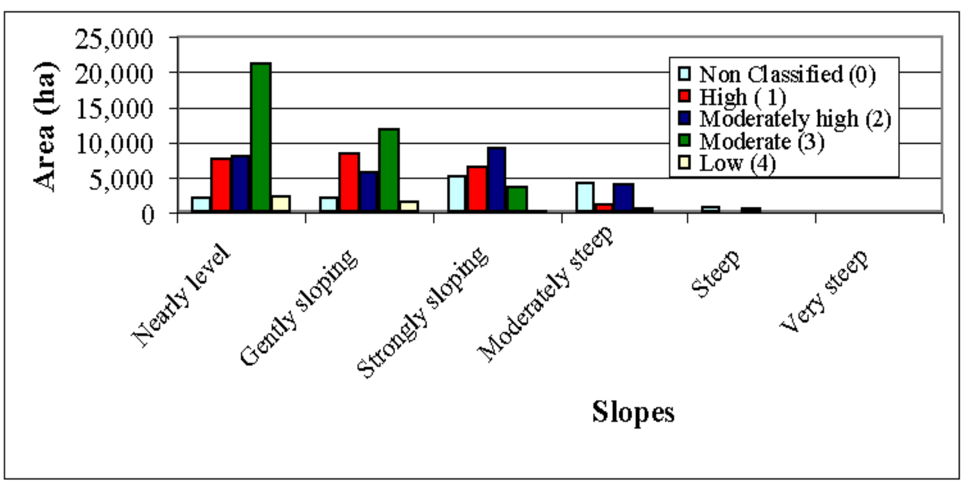

Figure 18. Mutual occurrence of slopes and ground water vulnerability categories.

Almost all of the highly vulnerable area coincided with the Boone geological Formation followed by the Upper Mississippian Formation. About 32,000 ha of the watershed mapped as the Boone Formation coincided with the moderate vulnerability categories. The Boone Formation covers about 87% of the study area (Figure 19). Primary and secondary porosity of the Boone formation are <1% and about 3.5 %, respectively (J. V. Brahana, Personal Communication).

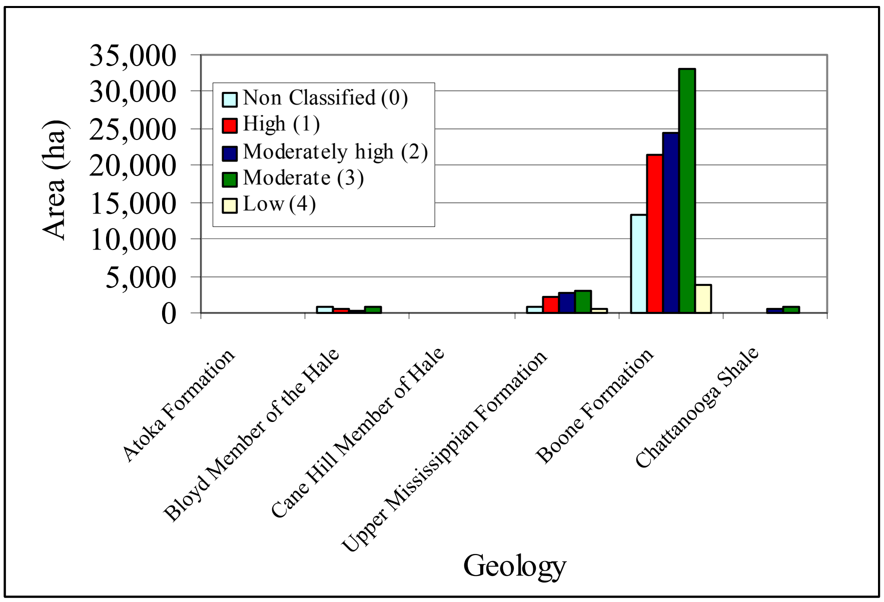

Figure 19. Mutual occurrence between geology and ground water vulnerability categories.

Coincidence between the Neuro-fuzzy Model and Field Data.

Water quality data for NO_3-N collected from 44 springs/wells in the watershed were also used in the study. Of the 44 wells, 24 analyses were provided by ADEQ and 20 analyses by AWRC. Three sets of coincidence reports were generated. First, coincidence was conducted between the data for 24 of the wells provided by ADEQ and vulnerability categories. Second, coincidence reports were generated between all of the 44 wells and the ground water vulnerability categories. Third, coincidence reports were generated between the NO_3-N contamination level classes and ground water vulnerability categories for all of the 44 wells.

For the first coincidence report, out of 24 wells, the location of two springs/wells coincided with the high vulnerability categories having a nitrate-N contamination level ranging from 14.5 to 16 mg/l. Seven wells/springs with nitrate-N concentration levels ranging from 0.02 – 6 mg/l coincided with the moderately high vulnerability category (Figure 20). The vulnerability category moderate also coincided with 7 wells/springs and nitrate-N concentrations ranged between 0.56 – 7.7 mg/l. This could be related to the spatial distribution of the wells and the vulnerability categories. The available springs/well data obtain from ADEQ were spatially biased and represent only a small portion of the watershed. Moreover, spatially, the moderately high category occupied about 26% whereas the moderate category occupied 35% of the watershed.

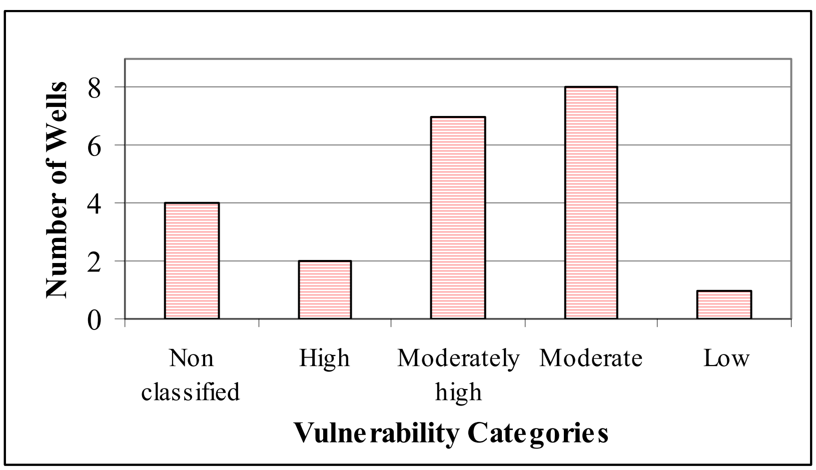

Figure 20. Coincidence of vulnerability classes predicted from the model and well locations.

The low ground water vulnerability category occupied only 4% of the watershed. Ten out of 24 wells sampled in the study area coincided with a soil profile depth of 60 inches and 7 wells coincided with profile depth of 96 inches. Eight wells coincided with pedality points of 38 and 53, respectively. Only one well coincided with low pedality points 17.

The results obtained from the second set of coincidence report analyses are shown in the Figure 21. It is interesting to note that in spite of adding 20 more wells to the water quality data set the general trend of the coincidence remained the same. Like the ADEQ data set, the moderate vulnerability category had the highest number of wells (16). Six and ten wells coincided with non-classified and high category, respectively. Compared to the data shown in Figure 20, only one more well was added to the low vulnerability category when the coincidence was generated with the entire data set of 44 wells. Two out of 44 wells coincided with low pedality points 14 and one well coincided with pedality point 17. Eight wells coincided with pedality points of 34. About 10 and 9 wells coincided with pedality points of 38 and 49, respectively. High pedality points of 53 coincided with 9 wells. About 10 wells coincided with 60 inches deep soil profile. About 17 wells coincided with a soil profile between 72 – 80 inches and 11 wells coincided with profile depth of greater than 96 inches. About 15 and 26 wells coincided with soil hydrologic groups C and B, respectively.

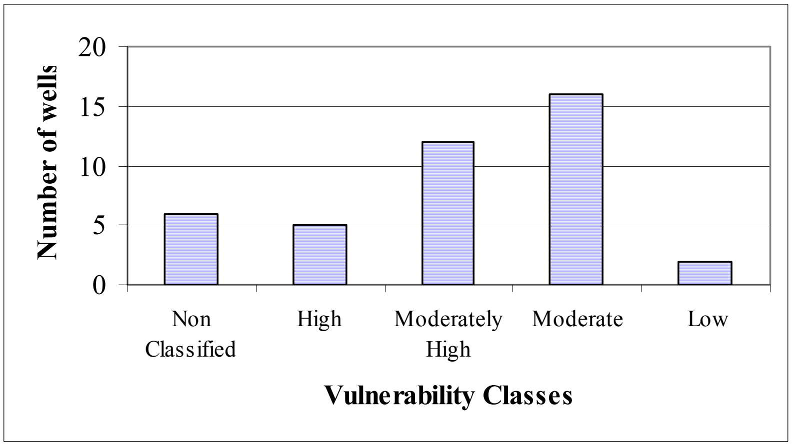

Figure 21. Coincidence of vulnerability classes predicted from the model and the entire well data for nitrate-N contamination.

A third set of coincidence reports was generated between vulnerability categories and the classes nitrate-N contamination data for all 44 wells (Figure 22). The nitrate-N contamination data were arbitrarily classified into four categories: low (<0.5 mg/l), moderate (0.5 – 3 mg/l), moderately high (3 – 10 mg/l) and high (> 10 mg/l). Two wells were classified as high concentration, one well coincided with high vulnerability and the

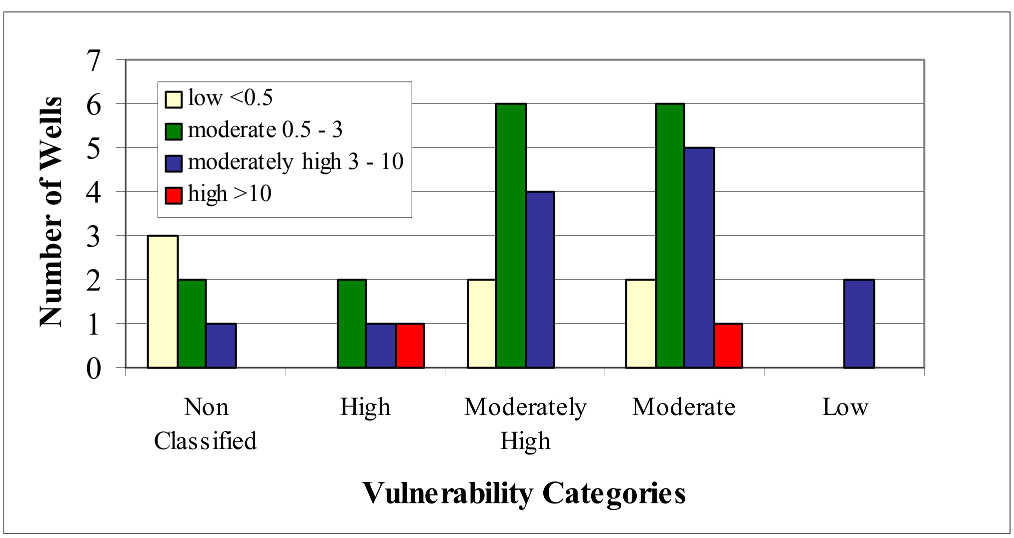

Figure 22. Coincidence between well contamination level and vulnerability categories

other wells coincided with the moderate vulnerable category. Relatively higher number of wells with moderately high contamination level coincided with moderate vulnerability category followed by moderately high vulnerability category (Figure 22). Almost equal numbers of wells with moderate contamination level coincided with moderate and moderately high vulnerability categories. Two wells with moderately high contamination level coincided with low vulnerability area. Location and nitrate-N contamination levels (mg/l) of wells are shown in Figure 23.

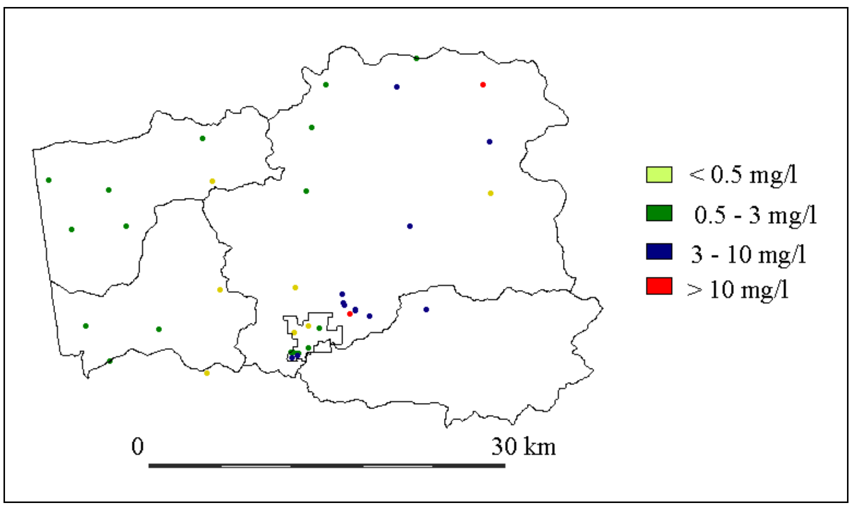

Figure 23. Location and nitrate-N contamination levels of wells in the watershed.

Coincidence analyses between model inputs and well contamination data are presented in Figure 24. Only two wells sampled in the study area were categorized in the highly contaminated category. One of each associated with urban and agricultural LULC, moderately deep and deep soil profile, moderately high and high soil structure and one each with hydrologic groups B and C (Figure 24). The majority of the moderately high contamination levels are associated with agricultural landuse, moderately deep soil profile, soil hydrologic group B and moderately high pedality points. As mentioned earlier, the well contamination data were not collected during the same time nor by the same agencies and were complied from different sources, therefore, this data set contained some additional uncertainty and variability. As a result, the comparison between well data and vulnerability categories should not be considered ‘absolute’ parameter in determining applicability of the model.

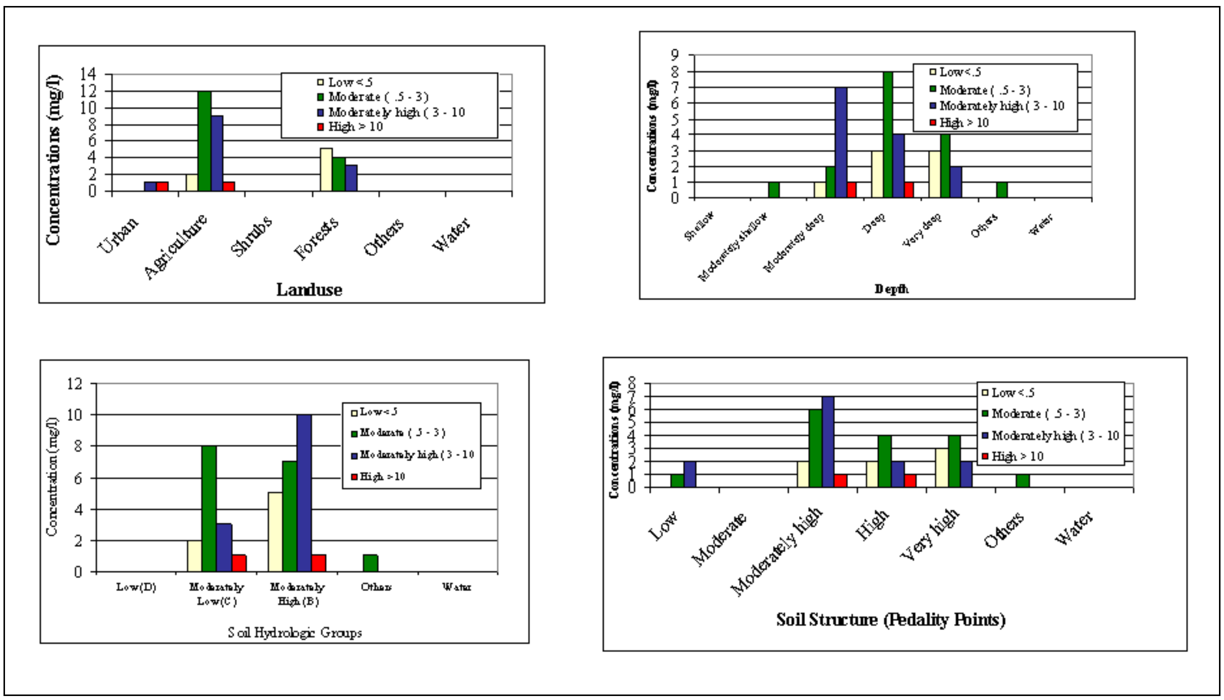

Figure 24. Coincidence between well contamination level and model inputs.

The data set for water quality was not considered to be adequate to determine the ability of the models to predict ground water vulnerability since it had inherent uncertainty. This brings out an interesting aspect of solute transport modeling on a regional scale as pointed out by Burrough (1996). He mentioned that most of the environmental data are collected on a project basis rather than in a systematic way which poses a problem in development of solute transport models in a regional scale. One of the goals of this research was to examine the usefulness of existing data in regional scale modeling of ground water vulnerability since this will reduce cost of modeling. Digging new wells to validate the model will be cost prohibitive.

Moreover, well/spring data are point data, a validation technique that, compared point data with spatial data predicted by the model, has inherent uncertainty. Further studies are required with a larger set of water quality data collected all over the watershed. Use of geostatistical tools to generate contamination surface and comparison of that surface with the map generated by the vulnerability model could be useful in the assessment of the model performance. Further studies are needed to determine the applicability of Neuro-fuzzy techniques in modeling ground water vulnerability on a regional scale. This modeling approach should be used again, preferably in areas where water quality data are already available with less temporal variability and adequately characterize the watershed or the study area.

CONCLUSIONS

Application of Neuro-fuzzy techniques to the prediction of ground water vulnerability does not provide exact solutions. Fuzzy systems, which are used to exploit the tolerance for imprecise solutions, are useful because they are easy to use, handle and understand. Use of the NEFCLASS-J tool provided all necessary statistics. No further statistical tools or computation with statistical tools were required. The model output showed higher coincidence with soil structure parameters (pedality points) and agricultural landuse than any other parameter. The preliminary model needed to be fine tuned though fine tuning of rulebase and classifier. From this research it is evident that the tool NEFCLASS-J could not automatically create the classifier. It supports the user but it cannot do all the work because a precise and interpretable fuzzy classifier can hardly be found by an automatic learning process. The NEFCLASS-J needs experts' opinion and tuning.

The water quality data were not sufficient to characterize the watershed. Locations of the wells/spring used in the study to validate the model had a spatial bias, and therefore, were not practically useful in validating the Neuro-fuzzy model. Further study is needed.

SUMMARY

This research used Neuro-fuzzy techniques to predict ground water vulnerability in northwest Arkansas. These techniques allowed incorporation of expert's opinion in the models, which is a valuable source of information particularly for the parameters that are hard to measure and vary over space and time. The models developed in this research used simple soil parameters including depth of the soil profile, soil hydrologic groups and pedality points of the A horizon and LULC to ensure global scope of the model. Since the underlying geology of the watershed is primarily the Boone Formation, which is highly fractured, it was assumed that any contaminants that reached the Boone Formation would also move to the ground water because this formation poses little hydraulic resistance to flow.

About 22% of the watershed was classified as highly vulnerable area and almost all of the highly vulnerable areas coincided with agricultural landuse, moderately deep and deep soils, soil hydrologic group C and high pedality points. As expected, the vulnerability map showed high coincidence patterns between the highly vulnerable areas and LULC and soil structure (pedality points) as expected. However, the coincidence patterns between the vulnerability categories and depth of the soil profile and soil hydrologic groups did not show the expected patterns of coincidence. The expected patterns were: the shallower the profile, the higher the vulnerability; or the higher the Ksat of a soil hydrologic group, the higher the vulnerability (i.e. soil hydrologic group B is more vulnerable than C). These discrepancies could be attributed to the fact that the majority of the agricultural landuse coincided with soil hydrologic group C (40,178 ha) followed by soil hydrologic group B (24,807 ha). About 47,818 ha of the watershed with deep soils coincided with agricultural landuse. Only two wells sampled in the study area had NO_3-N concentration of greater than 10 mg/l. One of each coincided with highly and moderately vulnerable areas, agricultural and urban landuse, moderately deep and deep soil profile, soil hydrologic groups B and C and pedality points of moderately high and high. As mentioned earlier due to the inherent uncertainty associated with the well contamination data, the data set should not be considered as 'absolute' parameter in determining performance of the model.

The proposed methodology has potential in facilitating modeling ground water vulnerability at a regional scale. This methodology can be used for other regions, however, this approach would require incorporation of appropriate input parameters suitable for the region. For example, if the geology of an area is different from the study area, geological factors should be incorporated to account for potential resistance to water and contaminants transport processes. This study is the first step toward incorporation of Neuro-fuzzy techniques in a GIS and would require improvements for wider range of application.

REFERENCES

Burrough, P. A. 1996. Opportunities and Limitations of GIS-based Modeling of Solute Transport at the Regional Scale. In (D. Corning and K. Loague eds.), Application of GIS to the non-point source pollutants in the vadose zone, SSSA, Special Publication # 48, Madison, WI.

CAST (1992). Arkansas Gap Project. Center for Advanced Spatial Technologies. Available at

http://www.cast.uark.edu/gap (viewed May 2000)

Chitsazan, Manouchehr. 1980. Hydrogeologic evaluation of the Boone-St. Joe carbonate aquifer. M.S. Thesis, Department of Geology, University of Arkansas, Fayetteville.pp.140.

Corwin, D. L, K. Loague and T. R. Ellsworth. 1996. Introduction to non-point source pollution in the vadose zone with advanced information technologies. In (D. L. Corwin, K. Loague and T. R. Eliisworth, eds.) Assessment of non-point source pollution in the vadose zone. Geophysical Monograph 108. AGS, Washington D.C.

Croneis, C. 1930. Geology of the Paleozoic region with special reference to oil and gas possibilities. Arkansas Geologic Survey Bulletin # 3. 457 pp.

Curtis, Darrin L. 2000. An integrated rapid hydrogeologic approach to delineate areas affected by advective transport in mantled karst, with an application to Clear Creek Basin, Washington County, Arkansas. Ph.D. Dissertation, University of Arkansas, Fayetteville. 121 p.

Dixon B., 2001. Application of Neuor-fuzzy tecniques to predict ground water vulnerability in NW Arkansas. Doctoral Dissertation. University of Arkansas. Environmental Dynamics.

Evans, B. M. 1990. A GIS-based approach to evaluating regional ground water pollution potential with DRASTIC. J. Soil and Water Conservation. 45(2): 242-245.

Fausett, L. 1994. Fundamentals of Neural Networks. Prentice Hall, NJ, p. 460.

Hays, T. S. 1995. Digital Characterization of the Illinois River Watershed and simulated phosphorus loading in representative sub-basins. M.S. Thesis. Department of Agronomy, University of Arkansas, Fayetteville.

Khan, E. 1999. Neural fuzzy based intelligent systems and applications In (Jain, L.C and N. M. Martin eds.) Fusion of neural networks, fuzzy sets and genetic algorithms: industrial applications. CRC Press, Washington, D.C. p. 331.

Lin, H. S., K. J. McInnes, L.P. Wilding, and C. T. Hallmark. 1999. Effects of Soil Morphology on hydraulic Properties: I. Quantification of Soil Morphology. Soil Science Society of America Journal.63:948- 953.

National Research Council. (NRC). 1993. Ground water vulnerability assessment: contamination potential under conditions of uncertainty. Committee on Techniques for Assessing Ground Water Vulnerability, Water Science and technology Board, Commission on Geosciences, Environment, and Resources. National Academy Press, Washington D.C. p.179.

Nauck, D., F. Klawonn, R. Kruse. 1997. Foundations on Neuro-Fuzzy Systems. John Wiley, Chichester.

Nauck, U. and R. Kruse. 1999. Design and Implementation of a Neuro-fuzzy data analysis tool in JAVA. Manual. Technical University of Braunschweig. Germany.

Peterson, J.C., Adamski, J.C., Bell, Davis, J.V., Femmer, S.R., Freiwald, D.A., and Joseph, R.L., 1998, U.S. Geological Survey Circular 1158, on line at < URL: http://water.usgs.gov/pubs/circ1158>

Razaie, Nassir. 1979. The hydrogeology of the Boone-St. Joe aquifer of Benton County, Arkansas. M.S. Thesis, Department of Geology, University of Arkansas, Fayetteville. pp. 147.

Smith, C. R. and K. F. Steele. 1990. Nitrate concentrations of ground water in Benton County, Arkansas. AWRC Pub. # 73. p. 48.

Soil Division Staff. 1993. Soil Survey Manual. USDA Handbook. No.18. U.S. Gov. Print. Office, Washington, DC.

Yager, R. R., Ovchinnikov, S., R. M. Tong and H. T. Nguyen. 1987. Coping with the imprecision of real world: an interview with L. A. Zadeh. In (Yager, R. R., Ovchinnikov, S., R. M. Tong and H. T. Nguyen, Eds.) Fuzzy sets and applications: selected papers by L. A. Zadeh. John Wiely and Sons, NY.

Yen, J. and Langari R. 1998. Fuzzy Logic: intelligence, control and information. Prentice Hall. NJ. pp. 548.

Zadeh, A. L. 1965. Fuzzy sets. Information and Control. 8:338-353

APPENDIX A

% NEFCLASS rule file file created by NEFCLASS-J 1.0 (c) Ulrike Nauck, Braunschweig, 1999

% Filename: Z:\Iberia\Barnali_home\dissertation\model\neuro_fuzz\basins\objective 1\obj1_rul.txt
% This file was created at April 15, 2001 7:34:13 AM CDT

if Var 1 is very small and Var 2 is very large and Var 3 is very small and Var 4 is large then Class 1
if Var 1 is small and Var 2 is very large and Var 3 is small and Var 4 is large then Class 2
if Var 1 is very small and Var 2 is very small and Var 3 is very small and Var 4 is very large then Class 1
if Var 1 is small and Var 2 is small and Var 3 is small and Var 4 is very large then Class 1
if Var 1 is small and Var 2 is very large and Var 3 is small and Var 4 is very large then Class 2
if Var 1 is small and Var 2 is small and Var 3 is small and Var 4 is large then Class 1
if Var 1 is very small and Var 2 is very small and Var 3 is very small and Var 4 is large then Class 1
if Var 1 is very small and Var 2 is large and Var 3 is very small and Var 4 is large then Class 1
if Var 1 is very small and Var 2 is very small and Var 3 is large and Var 4 is large then Class 2
if Var 1 is small and Var 2 is large and Var 3 is small and Var 4 is large then Class 2
if Var 1 is small and Var 2 is very small and Var 3 is large and Var 4 is very large then Class 2
if Var 1 is very small and Var 2 is small and Var 3 is small and Var 4 is large then Class 2
if Var 1 is small and Var 2 is very large and Var 3 is large and Var 4 is very large then Class 2
if Var 1 is small and Var 2 is small and Var 3 is large and Var 4 is very large then Class 2
if Var 1 is very small and Var 2 is large and Var 3 is small and Var 4 is very small then Class 2
if Var 1 is small and Var 2 is large and Var 3 is small and Var 4 is very large then Class 2
if Var 1 is small and Var 2 is very small and Var 3 is small and Var 4 is large then Class 3
if Var 1 is small and Var 2 is very large and Var 3 is small and Var 4 is very small then Class 3
if Var 1 is small and Var 2 is large and Var 3 is small and Var 4 is very small then Class 3
if Var 1 is very small and Var 2 is very large and Var 3 is large and Var 4 is large then Class 3
if Var 1 is very small and Var 2 is small and Var 3 is large and Var 4 is large then Class 3
if Var 1 is very small and Var 2 is small and Var 3 is small and Var 4 is very small then Class 3
if Var 1 is very large and Var 2 is very large and Var 3 is very large and Var 4 is very large then Class 3
if Var 1 is small and Var 2 is large and Var 3 is large and Var 4 is very small then Class 3
if Var 1 is very large and Var 2 is very small and Var 3 is very large and Var 4 is very large then Class 3
if Var 1 is small and Var 2 is very small and Var 3 is small and Var 4 is very small then Class 4
if Var 1 is small and Var 2 is small and Var 3 is small and Var 4 is very small then Class 3
if Var 1 is small and Var 2 is very small and Var 3 is large and Var 4 is very small then Class 4
if Var 1 is very small and Var 2 is very large and Var 3 is small and Var 4 is large then Class 4

Var 1 = Soil hydrologic groups

Var 2 = Landuse (LULC)

Var 3 = Depth of the soil profile

Var 4 = Soil Structure or pedality points

APPENDIX B

365254.33885602(E) 4007029.56321023(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

371195.19890633(E) 4015371.53764205(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

376385.95036443(E) 4001358.22088068(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (4)Forested Land
depth.total.rcl in mitra.ill4 (5)Very deep
structure.layr1.rcl.final in mitra.ill4 (5)Very high
objective1_trapez.rcl in mitra.ill4 (0)no data

399519.29934823(E) 4010090.28764205(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (3)

398499.15166282(E) 4002498.49076705(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (1)Urban
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (3)

393428.41757947(E) 3996437.05610795(N)
hydrogrp.fuzzy.input in mitra.ill4 (10)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (2)

364264.1955143(E) 4000848.10014205(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (4)Forested Land
depth.total.rcl in mitra.ill4 (5)Very deep
structure.layr1.rcl.final in mitra.ill4 (5)Very high
objective1_trapez.rcl in mitra.ill4 (0)no data

373235.49427715(E) 4014111.23934659(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (4)Forested Land
depth.total.rcl in mitra.ill4 (5)Very deep

structure.layr1.rcl.final in mitra.ill4 (5)Very high
objective1_trapez.rcl in mitra.ill4 (0)no data

365644.39532397(E) 4016001.68678977(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (3)

360093.59174161(E) 4007299.62713068(N)
hydrogrp.fuzzy.input in mitra.ill4 (10)
lulc.85.30m.study in mitra.ill4 (1)Urban
depth.total.rcl in mitra.ill4 (3)Moderately deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (0)no data

361473.79155128(E) 4003638.76065341(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (3)

399129.24288028(E) 4002738.54758523(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (1)Urban
depth.total.rcl in mitra.ill4 (5)Very deep
structure.layr1.rcl.final in mitra.ill4 (3)Moderately high
objective1_trapez.rcl in mitra.ill4 (2)

381066.62797983(E) 4017772.10582386(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (3)Moderately deep
structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

391178.09180284(E) 4018552.29048295(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (3)Moderately deep
structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

369304.92525396(E) 4014501.33167614(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

371705.27274904(E) 4016901.89985795(N)
hydrogrp.fuzzy.input in mitra.ill4 (20)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (4)Deep

structure.layr1.rcl.final in mitra.ill4 (4)High
objective1_trapez.rcl in mitra.ill4 (1)

392048.2177698(E) 3995026.72230114(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (2)Agricultural Land
depth.total.rcl in mitra.ill4 (3)Moderately deep
structure.layr1.rcl.final in mitra.ill4 (1)Low
objective1_trapez.rcl in mitra.ill4 (4)

397869.06044536(E) 4011830.69957386(N)
hydrogrp.fuzzy.input in mitra.ill4 (30)
lulc.85.30m.study in mitra.ill4 (1)Urban
depth.total.rcl in mitra.ill4 (4)Deep
structure.layr1.rcl.final in mitra.ill4 (1)Low
objective1_trapez.rcl in mitra.ill4 (4)

APPENDIX C

Source: ADEQ

Northing	Easting	Well ID	NO_3-N concentration (mg/l)
379426.8131	3997642.065	#1	0
379427.7591	3997674.411	#2	2
379442.5775	3997679.138	#3	2
383676.0906	4001220.973	#4	15
383025.2901	4003017.085	#5	4
383090.7579	4002307.403	#6	6
383237.7978	4002089.712	#7	8
385273.4467	4001045.764	#8	7
384130.8208	4001584.739	#9	7.77
380323.5933	4000168.522	#10	0.063
379123.9042	3999552.647	#11	0.063
378906.7469	3997772.214	#12	1.89
378928.6386	3997312.415	#13	6.12
379060.8084	3997756.214	#14	0.562
381184.4485	3999960.158	#15	2.017
380286.028	3998246.619	#16	0.605
379408.9446	3997543.078	#17	4.53
384055.444	4001554.922	#18	6.32
384055.444	4001554.922	#19	7.7
383676.0906	4001220.973	#20	16
383115.3328	4002276.256	#21	5.9
389854.5038	4001603.141	#22	4.2

Source: Smith and Steele (1990)

Northing	Easting	Well ID	No3-N concentration (mg/l)
381706.1053	4021772.388	#28	0.74
387463.6712	4021572.535	#29	9.99
394498.1283	4021792.331	#30	26.74
359149.5084	4013294.147	#32	0.53
364054.1198	4012445.619	#33	1.74
380178.314	4012239.541	#34	0.73
373135.5475	4003463.38	#42	0.09
379285.1649	4003622.724	#43	0.01
368183.2069	3999869.205	#49	1.27
362161.7228	4000116.197	#50	0.59
365501.742	4009186.966	#51	1.21
361099.8299	4008793.707	#52	0.89
394985.5117	4016732.444	#53	7.68
388551.6091	4009169.845	#54	4.86
380507.1051	4017998.1	#55	0.53
371735.4163	4016982.835	#56	
389117.7165	4024170.758	#58	1.68
372527.5885	4013149.532	#66	0.1
395078.7028	4012047.185	#67	0.08
374718.4312	4033458.801	#72	0.3