Land coverand landscape as predictors of groundwater contamination: a neural-network modelling approach applied to dobrogea, romania

Article *in* Journal of environmental protection and ecology · January 2010

6 authors, including:

Richard Ross Shaker
Ryerson University
48 PUBLICATIONS **253** CITATIONS

Mirela Bucur
Universitatea Ovidius Constanţa
12 PUBLICATIONS **24** CITATIONS

Daciana Sava
30 PUBLICATIONS **93** CITATIONS

Journal of Environmental Protection and Ecology 11, No 1, 337–348 (2010)

Environmental informatics

LAND COVER AND LANDSCAPE AS PREDICTORS OF GROUNDWATER CONTAMINATION: A NEURAL-NETWORK MODELLING APPROACH APPLIED TO DOBROGEA, ROMANIA

R. SHAKER[a], L. TOFAN[b], M. BUCUR[b], S. COSTACHE[b], D. SAVA[b], T. EHLINGER[c]*

[a]*Department of Geography, University of Wisconsin–Milwaukee, Milwaukee, WI, USA*
[b]*Faculty of Natural and Agricultural Sciences, Ovidius University, Constanta, Romania*
[c]*Department of Biological Sciences, University of Wisconsin – Milwaukee, P. O. Box 413, Milwaukee, WI. 53 201, USA*
E-mail: ehlinger@uwm.edu

Abstract. Drinking water contamination is a major problem facing urban and rural communities worldwide. A study exploring the linkages between ground water quality and landscape metrics was performed using 55 ground water wells found within 20 watershed sub-basins in the Dobrogea region of Romania. The goals of this research were: (1) to establish a statistical model that could be used to predict the risk of ground water contamination; and (2) increase the understanding of the relationship between landscape composition and configuration on ground water quality. The relationships between land cover, landscape and ground water nitrate level were established using artificial neural-network statistical analysis. Analysis revealed that the percent land cover in watersheds as forest, agriculture and artificial (urban) surface, combined with area weighted mean contiguity index, like-patch adjacencies, and area weighted proximity index were able to predict over 62% of the variation in nitrate levels in wells. This approach shows promise for being an effective predictor of risk of contamination for drinking wells.

Keywords: groundwater, nitrate contamination, landscape analysis, neural network modelling, Romania.

AIMS AND BACKGROUND

The structure, function, and dynamics of contemporary ecosystems are profoundly influenced by human activities[1], and understanding the mechanisms responsible for environmental changes requires the integration of both natural and human processes. Pervasive ecological changes have occurred as a result of human activities[1]. Changes in land cover through the appropriation of natural landscapes to

* For correspondence.

provide for human needs is one of the most pervasive alterations resulting from human activity. While the ecological and sociological effects of land conversion for agricultural uses are well studied, the long-term effects on human health are constantly under investigation.

Nitrate contamination is possibly the most wide-spread contaminant of water globally[2]. In the United States, ground water provides drinking water for more than one-half of the nation population; mostly rural communities and a majority of cities[3]. In terms of the health and well-being of humans, increased nitrate levels in ground water can lead to numerous concerns, including nutrient enrichment of surface waters and carcinogenetic effects in animals and humans[4]. The United States Protection Agency (USEPA) has stated that nitrate concentrations greater than 10 mg/l can have long-term effects on sensitive populations[4]. The Maximum Acceptable Concentration (MAC) for nitrate (NO_3–N) in drinking water in Canada and the United States is currently set at: 10 mg/l (Refs 5 and 6). The European Environmental Agency (EEA) has set a maximum concentration of nitrate in groundwater at 25 and 50 mg/l for country and regional levels, respectively[7].

Sources of increased nitrate levels in ground water have been linked to nitrogen-based fertilisers, animal and human wastes, and to a lesser extent, industrial waste, waste waters, and landfills[8]. It has been suggested that agricultural practices are the main cause of elevated nitrate in groundwater and intensive agriculture is considered to be the main source of water pollution by nitrate in Europe and North America.

A variety of techniques have been used in the past to estimate ground water contamination, from simple overlay/index approaches to complex simulation models[9]. Overlay/index methods have proven to be successful at assessing vulnerability of shallow groundwater contamination and are better suited to accommodate uncertainty in data than other methods[10]. The overlay/index methods involve the statistical analysis of physiological attributes of the landscape and the combining and assigning of weights to each property to obtain a final score[9]. This method is driven largely by data availability and expert scientific judgment[10].

Sources of ground water contamination have been linked to nitrogen-based fertilisers, animal and human waste, and to a lesser extent, industrial, waste water, and landfills[8]. Specifically, nitrate contamination has been suggested to be the most widespread contamination of water globally[2].

Our ability to accurately model non-point source nutrients in ground water is fair at best, and has been suggested to need further investigation[9,11]. The goals of this research were twofold: (1) to establish a statistical model that could be used to locate ground water contamination for remediation; and (2) increase the understanding of the relationship between landscape composition/configuration and ground water quality.

EXPERIMENTAL

The Dobrogea region of Eastern Romania is located between the Danube river to the west and the Black sea to the east. The region is divided into two major basins; the Litoral basin contains all the waterways that flow east into the Black sea, and the Dunabe basin contains all the waterways that flow west into the Danube river. This area of Romania, like many places throughout the world, experienced significant landscape alterations since the agrarian revolution albeit intensifying during the communist period when fertiliser applications to agricultural lands increased from 20 kg/ha to over 160 kg/ha throughout Romania (Fig. 1). Application rates dropped drastically following the 1989 revolution, however, the legacy impacts of overfertilisation on groundwater quality remain.

Fig. 1. Fertiliser application rate (mean kg/ha per year) for Romania from 1961–2001. Datasource: http://earthtrends.wri.org

A study by Constanta County Public Health Direction sampled drinking wells from 55 locations within 20 sub-watersheds from Dobrogea in 2007. Methods for water sampling, conservation, transport, storage and identification for samples followed procedures specified in S.R. 2852/1994 – 'Drinkable Water' by the Romanian government. Samples were collected from wells with a bucket from a depth of 0–30 cm below the water surface. Samples for nitrate analysis were fixed with 2 ml sulphuric acid, placed in refrigerated boxes, at –8°C and transported to the laboratory as soon as possible (maximum 6 h) after sampling. If the samples could not be taken to the laboratory in the specified time or could not be analysed immediately, they were stored in a refrigerator under the same condition for a maximum of 24 h. Nitrate levels were measured by photometric methods and reported as concentration in mg/l of NO_3–N. Spatial distribution of sampling locations and relative nitrate concentrations are shown in Fig. 2.

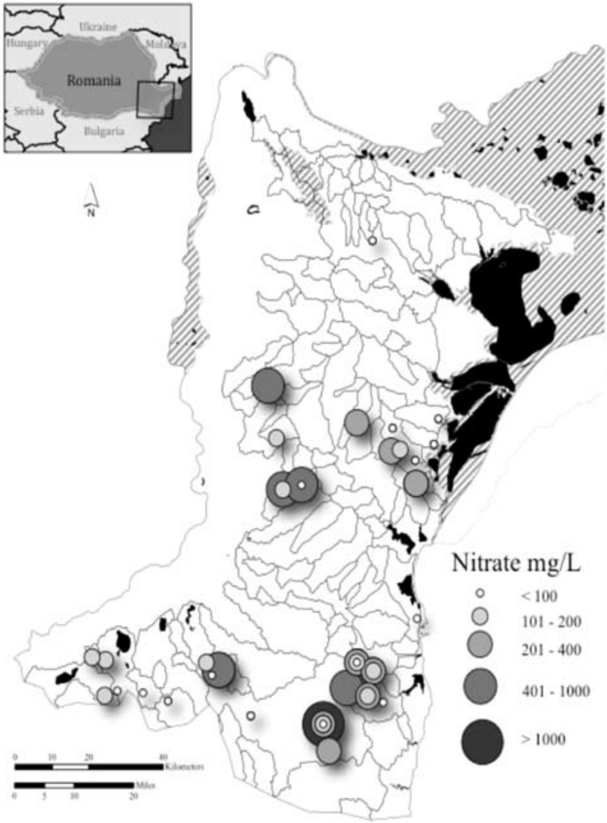

Fig. 2. Location of the 55 ground water well sites within 20 corresponding sub-basin watershed landscapes. Relative nitrate concentrations are indicated by the size and colour of the symbol. Sampling locations in close proximity are indicated by over-laid symbols

ESRI ArcGIS 9.3 and FRAGSTATS Version 3.0 (Ref. 12) were used to calculate the selected land cover and landscape ecology metrics for the 20 watershed sub-basin landscapes. Generalised land cover data for 2000 (percent agriculture, forest, wetland, water and artificial/urban) for each sub-basin were calculated at 30 × 30 m resolution. Using previously stated data for agricultural lands, 25 landscape ecology metrics were calculated by FRAGSTATS to quantify landscape composition and configuration (Table 1). These landscape ecology metrics were chosen based on literature relevance[13–16] and statistical strength. All data were transformed to correct for non-normality prior to statistical analysis. Proportional (percent) and length data were transformed using the negative arcsine and log10 transformations, respectively. Predictor variables were standardised using a z-transformation to set all parameters to a mean of 0 and variance of 1.

Table 1. List of land cover (A) and landscape metrics (B) used in analyses. Full descriptions can be found in Ref. 12

Metric	Abbreviation	Name
(A)	LAND COVER METRICS	
	TO_AREA	total sub-basin area
	Agriculture[1]	percentage of landscape agriculture
	Artificial[1]	percentage of landscape artificial surface
	Wetland	percentage of landscape wetland
	Forest[1]	percentage of landscape forest
	Water	percentage of landscape water
(B)	LANDSCAPE METRICS	
AREA/EDGE/DENSITY		
	PD	patch density
	LPI	largest patch index
	ED	edge density
	AREA_AM	area weighted mean patch area
	GYRATE_AM	area weighted mean of radius of gyration distribution
	nLSI	normalised landscape shape index
	CA	total class area
SHAPE		
	SHAPE_MN	mean shape index
	SHAPE_AM	area weighted mean shape index
	FRAC_AM	area weighted mean fractal dimension index
	FRAC_SD	standard deviation of mean fractal dimension index
	PARA_MN	mean perimeter area ratio
	CONTIG_MN	mean contiguity index
	CONTIG_AM[1]	area weighted mean contiguity index
PROXIMITY/ISOLATION		
	PROX_AM **[1]	area weighted mean proximity index
CONTATION/INTERSPERSION		
	PLADJ[1]	percentage of like adjacencies
	SPLIT	splitting index
	AI	aggregation index
CONNECTIVITY		
	COHESION*	patch cohesion index

[1] Metric was chosen for nitrate model based on backward stepping multiple regression; * calculated using a fixed edge depth of 50 m; ** calculated using a search radius of 100 m.

We first reduced the number of predictor variables by removing those that exhibited auto-correlation. This was done by using a backward-stepping multiple regression with nitrate as the dependent variable, and retaining the predictor parameters from the best-fit model to be used in the remaining analyses. Regressions were conducted separately for land cover variables and landscape ecology metrics.

Next, an artificial neural network (JMP version 6.02)(Ref. 17) was used to predict nitrate levels for well locations based upon the retained land cover and landscape variables. The model was run with 3 hidden nodes and 5-fold cross validation.

RESULTS AND DISCUSSION

The distribution of nitrate concentrations from the 55 wells sampled is presented in Fig. 3, with levels ranging from less than 20 mg/l to almost 1000 mg/l. Data exhibit an approximately log-normal distribution, with 70% of the observations exceeding the 50 mg/l standard for drinking water (Fig. 3C). Mean nitrate concentration for the 20 sub-basins is presented in Fig. 4; with the highest mean just less than 500 mg/l.

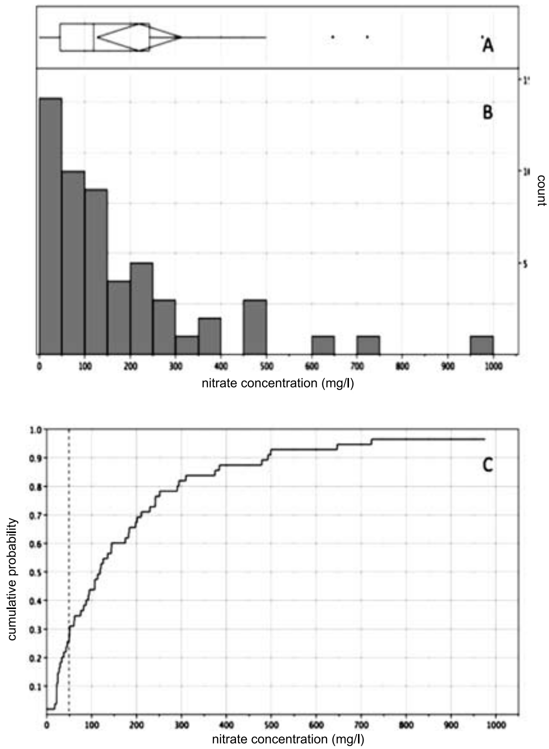

Fig. 3. Distribution of groundwater nitrate concentrations (mg/l) detected in samples collected from Dobrogea Romania in 2007: box plot showing quartiles and confidence diamond for mean value (A), frequency distribution (B), and cumulative probability distribution (C)

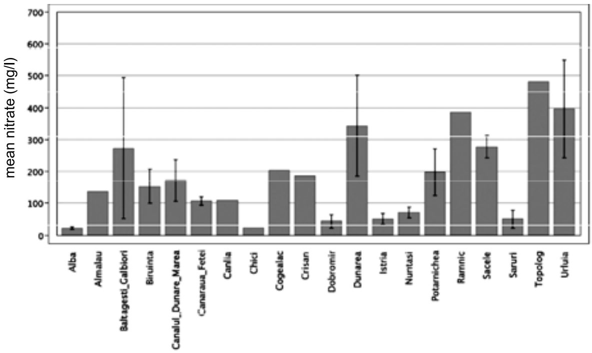

Fig. 4. Mean groundwater nitrate concentrations in groundwater samples collected from wells in Dobrogea, Romania in 2007. Data are means ± 1 SE for each sub-basin used in the analysis

As expected, agriculture was the dominant land cover for all sub-basins (Fig.5), ranging from 50 to over 95% of the surface area. Forested and artificial/urban land covers were the next most abundant, exhibiting maximum cover of 43 and 22%, respectively (Fig. 5). Percentages of wetland and water were lower than other categories, but exhibited great variability among sub-basins (Fig. 5).

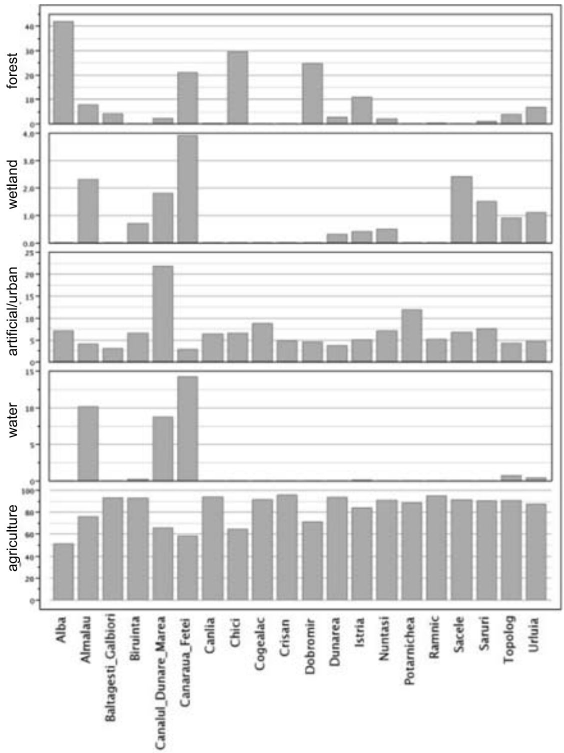

Fig. 5. Percentages of major land cover classes for the 20 sub-basins included in the study

Backward-stepping multiple regressions of nitrate concentration against land cover metrics revealed that the combination of percent agriculture, forested, and artificial/urban land cover provided the best predictive model (R^2=0.24, p<0.01). Similarly, multiple regression with landscape metrics resulted in 3 metrics selected in the best-fit model ($R^2 = 0.28, p < 0.01$). The 3 landscape metrics selected included Area Weighted Mean Contiguity Index (CONTIG_AM), Percent Like Adjacent Index Agriculture (PLADJ), and Area Weighted Proximity Index (PROX_AM).

The 6 metrics selected were then entered into a supervised neural network with 3 hidden nodes to predict nitrate levels (Fig. 6A). The resulting model (Table 2) was able to predict 64% of the variation in nitrate levels among samples (Fig. 6B).

Table 2. Parameter estimates for the three hidden layers for the neural network model predicting nitrate levels ($R^2 = 0.64$, $p = 0.001$, see Fig. 6A for model diagram)

Parameter	Parameter estimates		
	H:1	H:2	H:3
Intercept	–3.19	2.04	–0.57
ASN_FOREST	–2.40	2.35	–1.71
ASN_AGLAND	1.18	–2.59	1.01
ASN_ARTIF	–2.05	–0.91	–3.94
L10_CONTIG_AM	–3.49	0.90	0.88
ASN_PLADJ	0.41	3.56	–2.04
L10_PROX_AM	0.27	0.27	2.50
LOG_NITRATE	–5.77	–2.64	3.40
LOG_NITRATE:Intercept	1.10		

The 3 nodes differently captured variation in nitrate levels (Table 2) and the significant non-linearity of the relationships between landscape, land cover and water quality (Fig. 6C). Increased nitrate levels were associated with low forest and high agricultural land cover. Nitrate peaked at moderately low artificial/urban land cover and then exhibited a moderate increase again at high artificial/urban land cover. With regard to landscape ecology metrics, increased CONTIG_AM explained increased nitrate in the lower range, whereas increased PLADJ explained decreasing nitrate levels from very high to low. Changes in PROX_AM capture nitrate variation in the middle range of concentrations.

A combination of 6 metrics predicted nitrate loadings in groundwater using neural networks (Table 2 and Fig. 6). Three metrics (percent forest, percent agriculture, and percent artificial surfaces) were related to watershed land cover and three metrics (CONTIG_AM, PLADJ, and PROX_AM) were related to watershed agriculture land composition and configuration. Each of these metrics exhibited strong non-linearity with nitrate concentrations and each had predictive ability in different ranges of nitrate contamination (note different curves in Fig. 6C).

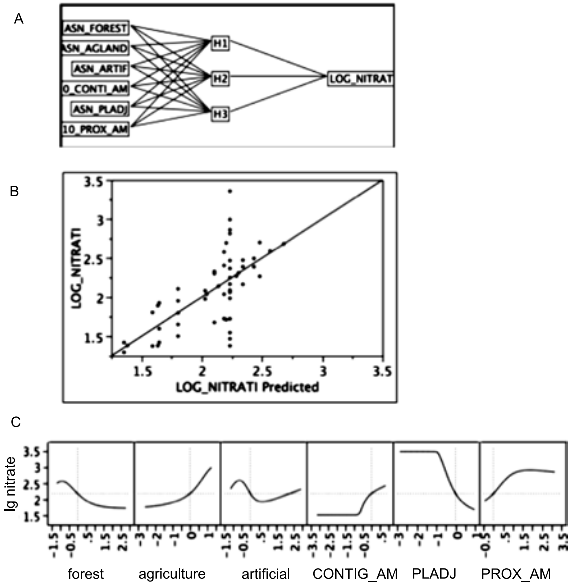

Fig. 6. Neural network diagram (A), observed versus predicted plot for nitrate (B), and parameter response profiles for predictor variables from neural network (C)

Percent forest, *agricultural and artificial/urban land*. A negative relationship was found between percent forest and nitrate loadings in groundwater. In Dobrogea, there is a strong negative correlation between the amount of forested and agricultural land in basins. The question becomes whether a loss of forest is responsible (e.g. through impacts on hydrogeology and geochemistry), or whether the correlated increase in agriculture is the dominant driver. Other studies have shown agricultural lands are associated with higher rates of nitrate loading into groundwater. As supported by previous groundwater nitrate research, this phenomenon can be attributed to the over-application of natural and synthetic fertilisers. Furthermore, historic over-application of fertilisers can have a legacy effect in groundwater. However, the nonlinearity of the response profiles for Forest and Agriculture in the neural network (Fig. 6C) that even small amounts of forest can have a positive impact on water quality, and that the negative impact of agriculture occurs mostly at the highest levels of agriculture.

These finding suggest that efforts to restore forestland within basins can have a positive impact at lowering nitrate risk to groundwater, even without large changes in agricultural land. In addition, this pattern suggests that much of the variation in nitrate levels may be the result of past historically high levels of fertiliser application.

The neural network indicated a strong nonlinear relationship between artificial surfaces and nitrate levels, with a peak effect at a low level of artificial surface

(Fig. 6C). To some extent these levels of artificial surfaces can reflect agricultural development in the basin. For example, at the earliest stages of agricultural development, a minimum level of artificial surfaces would be required for operations and worker residences – so historically one may see a concomitant increase in artificial and agricultural land cover. At the other end, increases in human population and development would expect that the percentage of agricultural land would decrease as urbanisation increases. The modest increase in nitrate levels at higher artificial surface levels could indicate a relationship with urban sewage systems and leakages/runoff into groundwater.

Agricultural landscape configuration. Low area weighted mean contiguity index (CONTIG_AM) scores for basins were associated with low nitrate levels. However, nitrate levels rose above a threshold CONTIG_AM score (Fig. 6C). CONTIG_AM is a shape metric that is related to the spatial connectedness, or contiguity, of cells within a grid-cell patch. This provides an index of patch boundary configuration and patch shape[18,19]. Area-weighted mean (AM), in this case, equals the total area of all agricultural patches divided by the total area of each watershed landscape. As CONTIG_AM metric increases, the size and shape of the agricultural patches increases for each sub-basin landscape. This can be interpreted as the larger the size and shape of the agricultural patches, the more potential for nitrate contamination in groundwater. Above a threshold, the larger the agricultural patches within basins, the greater the potential for moderately increased levels of nitrate in groundwater.

Percent like adjacent index agriculture (PLADJ) exhibited a relationship opposite to that of CONTIG_AM. Nitrate levels were highest at low PLADJ, and then dropped abruptly above a threshold level (Fig. 6C). PLADJ is a configuration metric that is related to land cover type connectivity[19] and equals the sum of the number of cell-like adjacencies for agriculture land, divided by the total number of cell adjacencies in the basin landscape (multiplied by 100 to convert to a percentage). As PLADJ metric increases, the connectivity of the agricultural patches increases for each basin landscape. The neural network pattern shows that high PLADJ (low dispersion) values were associated with lower nitrate levels in groundwater. In the case of Dobrogea, the highest levels of nitrate were associated with basins where agricultural patches were spread throughout the basin.

Area weighted proximity index (PROX_AM) predicted of variation in nitrate levels in mid-level ranges (Fig. 6C). PROX_AM is a measure of landscape configuration that deals explicitly with the spatial arrangement of patches[15,19] and is calculated as the sum of agricultural patch area divided by the nearest edge-to-edge distance squared of a corresponding agricultural patch within a specified search radius. AM (area-weighted mean), in our case, equals the total area of all agricultural patches divided by the total area of each watershed landscape. PROX_AM is dimensionless and should be used as a comparative index for complexity and configuration. For basins in Dobrogea, PROX_AM was positively associated with middle-range nitrate loadings in groundwater. Generally speaking, the higher the density of large agricultural patches, the greater the potential for increased levels of nitrate in groundwater. The relationship is similar at that for CONTIG_AM, except that CONTIG_AM predicts changes in nitrate at low concentrations and PROX_AM better predicts middle-range nitrate variation.

CONCLUSIONS

In conclusion, land cover and landscape reflect a complex mosaic of historical artifact, human legacy, and present-day activity. Changes to land-cover through the appropriation of natural landscapes to provide for human need is one of the most pervasive alterations resulting from human activity. As such, the relationships between each individual landscape metric and nitrate in groundwater can be interpreted in many ways. Researchers in landscape ecology have been continuously developing metrics for quantifying such patterns on different response variables[12,14,16,19–21], but there remains a demand for understanding the linkages between landscape metrics and ecological process and functions[22–25].

This study suggests that the presence of forests, in combination with the spatial patterns of agricultural patches can be an effective predictor of groundwater contamination risk for nitrates. In spite of the relatively small sample size of the study (55 wells across 20 sub-basins), the statistical significance and the directionality of trends detected by the neural network model suggest that such studies have the potential to develop landscape models for predicting sub-surface water quality risk. These models can assist in the efficient use and allocation of monitoring resources

Acknowledgements. This research was made possible by support from the US National Science Foundation Grant IRES 0623474 to TJE.

REFERENCES

1. M. ALBERTI:. The Effects of Urban Patterns on Ecosystem Function. International Regional Science Review, **28** (2), 168 (2005); K. ARMS: Environmental Science. 2nd ed. Saunders College Publishing, Fort Worth, TX. 1994.
2. Health Canada. 1992. Nitrate/nitrite. Available at: http://www.hc-sc.gc.ca/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/water-eau/nitrate_nitrite/nitrate_nitrite-eng.pdf. (verified 29 April 2009). Health Canada,Ottawa, ON.
3. M. G. TURNER: Landscape Ecology: The Effect of Pattern on Process. Annual Review of Ecology and Systematics, **20**, 171 (1989).
4. NRC: Ground Water Vulnerability Assessment: Contamination Potential under Conditions of Uncertainty. National Academy Press, Washington, DC, 1993.
5. J. C. LEHRTER: Effects of Land Use and Land Cover, Stream Discharge, and the Interannual Climate on the Magnitude and Timing of Nitrogen, Phosphorus, and Organic Carbon Concentrations in Three Coastal Plain Watersheds. Water Environment Research, **78** (12), 2356 (2006).

6. P. M. VITOUSEK: Beyond Global Warming: Ecology and Global Change. Ecology, **75**, 1861 (1994).
7. G. GULIS, M. CZOMPOLYOVA, J. R. CERHAN: An Ecologic Study of Nitrate in Municipal Drinking Water and Cancer Incidence in Tranava District, Slovakia. Environmental Resources, **88**, 182 (2002).
8. World Health Organization: Rolling Revision of the WHO Guidelines for Drinking-water Quality: Nitrates and Nitrites in Drinking-water. 2004. http://www.who.int/water_sanitation_health/dwq/chemicals/nitratesfull.pdf. (verified 29 April 2009). WHO, Geneva.
9. M. VIDAL, J. MELGAR, A. LOPEZ, M. SANTOALLA: Spatial and Temporal Hydrochemical Changes in Groundwater under the Contaminating Effects of Fertilizers and Wastewater. J. of Environmental Management, **60**, 215 (2000).
10. R. V. O'NEILL, J. R. KRUMMEL, R. H. GARDNER, G. SUGIHARA, B. JACKSON, D. L. DeANGELIS, B. T. MILNE, M. G. TURNER, B. ZYGMUNT, S. W. CHRISTENSON, V. H. DALE, R. L. GRAHAM: Indices of Landscape Pattern. Landscape Ecology, **1**, 153 (1988).
11. A. LEITAO, J. MILLER, J. AHERN, K. McGARIGAL: Measuring Landscapes: A Planner's Handbook. Island Press, Washington DC, USA, 2006.
12. K. McGARIGAL, B. J. MARKS: FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. General Technical Report PNW-GTR-351, USDA Forest Service, Pacific Northwest Research Station, Portland, OR, 1995.
13. M. ALBERTI: Advances in Urban Ecology: Integrating Humans and Ecological Processes in Urban Ecosystems. Springer, New York, NY, 2008.
14. M. ALBERTI, D. BOOTH, K. HILL, B. COBURN, C. AVOLIO, S. COE, D. SPIRANDELLI: The Impact of Urban Patterns on Aquatic Ecosystems: An Empirical Analysis in Puget Lowland Sub-basis. Landscape and Urban Planning, **80**, 345 (2007).
15. J. LaGRO, Jr.: Assessing Patch Shape in Landscape Mosaics. Photogrammetric Eng. Remote Sens., **57**, 285 (1991).
16. W. B. SOLLEY, R. R. PIERCE, H. A. PERLMAN : Estimated Use of Water in the United States in 1995. US Geological Survey Circular 1200, 1998. 77 p.
17. M. SERBANESCU: Dobrogea. Romart Design Publishing House, Constanta, 2005.
18. K. McGARIGAL, S. CUSHMAN, M. NEEL, E. ENE: FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Comp. Software Program. Amherst, MA, University of Massachusetts, 2002. (http://www.umass.edu/landeco/research/fragstats/fragstats.html.)
19. C. D. A. McLAY, R. DRAGTEN, G. SPARLING, N. SELVARAJAH: Predicting Groundwater Nitrate Concentrations in a Region of Mixed Agricultural Land Use: A Comparison of Three Approaches. Environmental Pollution, **115**, 191 (2001).
20. S. T. A. PICKETT, M. L. CADENASSO, J. M. GROVE, C. H. NILON, R. V. POUYAT, W. C. ZIPPERER, R. COSTANZA: Urban Ecological Systems: Linking Terrestrial Ecology, Physical, and Socioeconomic Components of Metropolitan Areas. Annual Review of Ecology and Systematics, **32**, 127 (2001).
21. N. K. C. TWARAKAVI, J. J. KALUARACHCHI: Sustainability of Ground Water Quality Considering Land Use Changes and Public Health Risks. J. of Environmental Management, **81**, 405 (2006).
22. M. ALBERTI, J. MARZLUFF, E. SHULENBERGER, G. BRADLEY, C. RYAN, C. SUMBRUNNEN: Integrating Humans into Ecology: Opportunities and Challenges for Studying Urban Ecosystems. Bioscience, **53** (12), 1169 (2003).
23. J. D. ALLAN: Landscape and River Spaces: The Influence of the Land Use on Stream Ecosystems. Annu. Rev. Ecol. Syst., **35**, 257 (2004).
24. SAS Institute Incorporated. 2006. JMP ™ System for Statistics. Cary, NC. USA.
25. United States Environmental Protection Agency (USEPA). 2007. Drinking Water Contaminants. Available at http://www.epa.gov/safewater/contaminants/index.html. United States Environmental Protection Agency, Washington DC, USA.

Received 4 July 2009
Revised 18 September 2009