

Statistical models for the assessment of nitrate contamination in urban groundwater using GIS

S.M. Lee · K.D. Min · N.C. Woo · Y.J. Kim · C.H. Ahn

Abstract The purpose of this study is to develop statistical models for groundwater quality assessment in urban areas using Geographic Information Systems (GIS). To develop the models, the concentrations of nitrate (expressed as nitrogen, $\text{NO}_3\text{-N}$), which are different according to the type of land use, well depth and distribution of rainfall, were analyzed in the Seoul (the capital of South Korea) area. Data such as land use, location of wells and groundwater quality data for nitrate contamination were collected and a database constructed within GIS. The distribution of $\text{NO}_3\text{-N}$ concentrations is not normal, and the results of the Mann-Whitney U-test analysis show the difference of $\text{NO}_3\text{-N}$ concentration by well depth and by distribution of rainfall. In both the shallow and deep wells, the radius of influence is 200 m in the dry season and 250 m in the rainy season, showing the tendency to increase in the rainy season. The results of correlation and regression analysis indicate that mixed residential and business areas and cropped field areas are likely to be the major contributor of increasing $\text{NO}_3\text{-N}$ concentration. Land uses are better correlated with $\text{NO}_3\text{-N}$ in deep wells than in shallow wells.

Keywords Statistical model · Geographic Information System · Nitrate contamination · Korea · Seoul · Radius of influence · Regression analysis

Introduction

Groundwater systems in urban areas show greater differences in physical and chemical characteristics than natural groundwater systems. A new hydrogeological system has been established in urban areas in Korea because of the artificial water supply (recharge) and discharge as well as a decreased groundwater recharge rate caused by increasing coverage of pavement (Ministry of Science and Technology 1997).

Seoul, the capital of South Korea (Fig. 1), has a population of approximately 10 million in an area of 605.7 km² (Seoul Metropolitan City 2001a). The quantity and quality of groundwater in this metropolitan area, which has expanded rapidly since 1970, have been greatly affected by factories, sewer lines, waste landfills, groundwater exploitation and so on (Kim and others 2001). The depth of wells has increased because of the deterioration of groundwater quality in shallow wells since the late 1980s. Subsequently, groundwater quality in the deep aquifer has been exposed to an increased risk of contamination. The annual precipitation in the Seoul area is 1,360 mm/year. A thin topsoil layer and steep slopes in the watershed causes rapid runoff of rainfall (Seoul Metropolitan City 1996). Groundwater aquifers of the Seoul area consist of thin alluviums and fractured crystalline bedrocks of granites, gneisses and schists. Alluviums are mainly distributed in the nearby Han River and are composed of coarse- and fine-grained sediments that result in high permeability. The geology of the area is mainly biotite granites and banded biotite gneisses, covering nearly 60% of the study area (Fig. 2). Gneiss areas show severe weathering. Groundwater systems are extremely distorted by large buildings, subway systems, water-supply and sewage systems and various underground constructions (Seoul Metropolitan City 1996). About 15,000 wells are used in the Seoul area at the present time and about 4.045×10^7 m³/year of groundwater is pumped from domestic, agricultural and industrial wells (Kim and Lee 2001). Both groundwater depletion and serious groundwater contamination have been reported in some parts of the area (Kim and others 2000). Leakage of water from the sewerage systems has a significant impact on some characteristics of the groundwater. To solve these problems, the factors contributing to groundwater contamination in urban areas have to be analyzed and systematic management plans need to be established (Custodio 1997; Foster 2001).

Received: 8 May 2002 / Accepted: 4 November 2002
Published online: 30 January 2003
© Springer-Verlag 2003

S.M. Lee (✉) · Y.J. Kim
Center for GIS/UIS, Seoul Development Institute,
San 4-5, Yejang-dong, Jung-ku, 120-250 Seoul, Korea
E-mail: lsm@sdi.re.kr
Fax: +82-2726-1187

K.D. Min · N.C. Woo
Department of Earth System Sciences,
Yonsei University, Seoul, Korea

C.H. Ahn
Electronics and Telecommunications Research Institute,
Daejeon, Korea

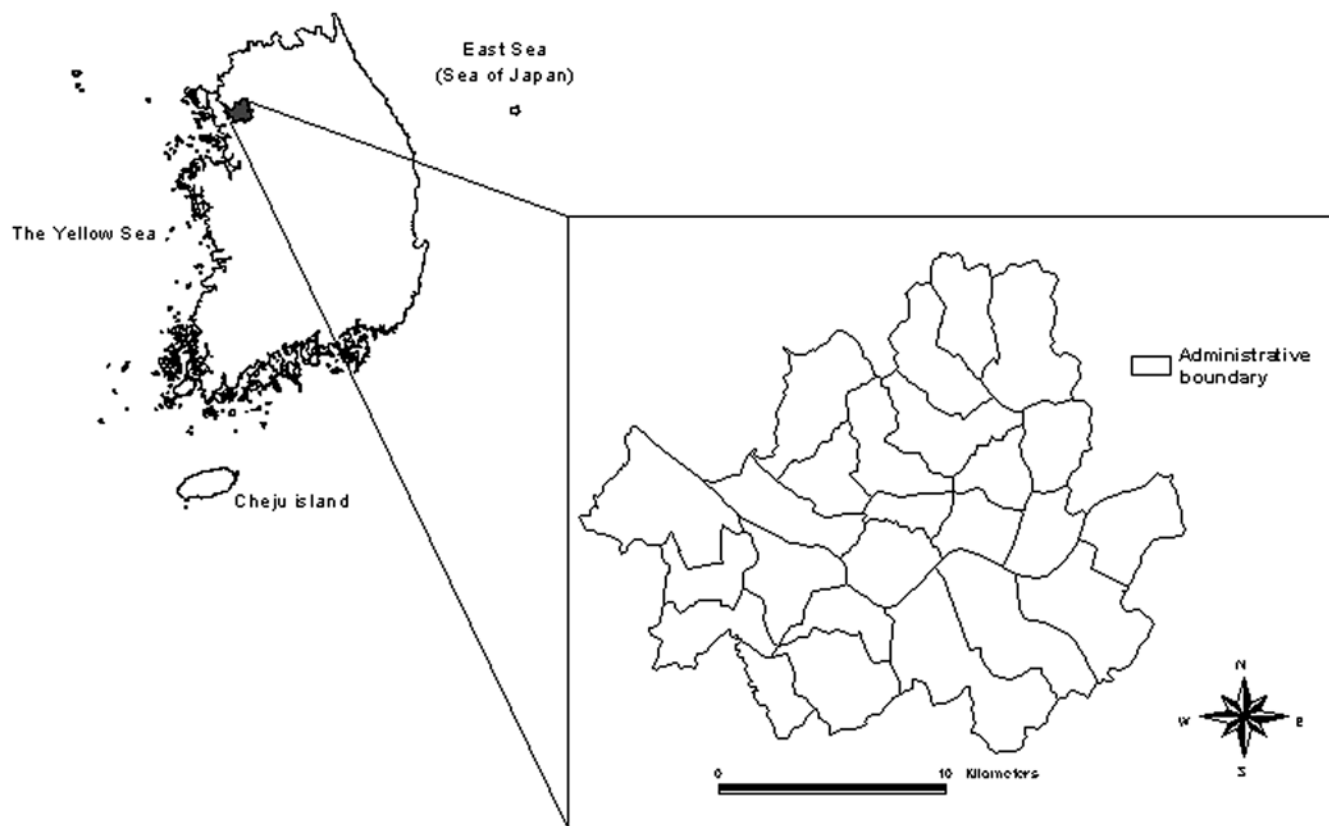


Fig. 1
Location of Seoul, Korea

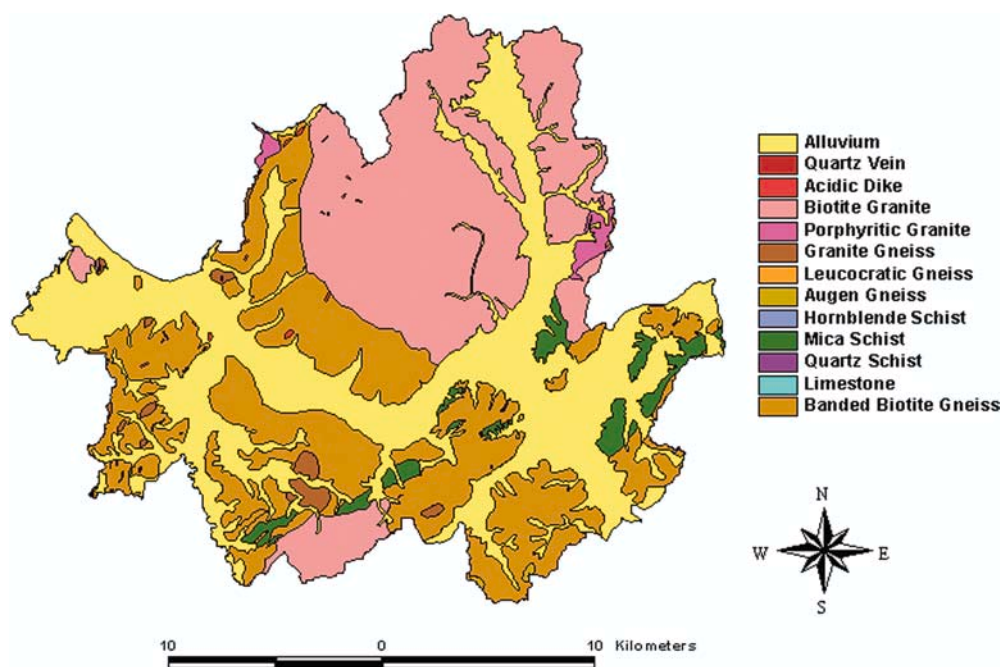


Fig. 2
Geological map of the study area

Recently, many research projects have examined the relationship between specific land-use patterns, corresponding pollutant emissions and the resulting groundwater quality (Trauth and Xanthopoulos 1997; Hong and Rosen 2001; Lasserre and others 1999). The object

of this study is to analyze the relationship between land uses and the well-known contaminant in urban groundwater, $\text{NO}_3\text{-N}$, and subsequently to develop statistical models for the assessment of nitrate contamination.

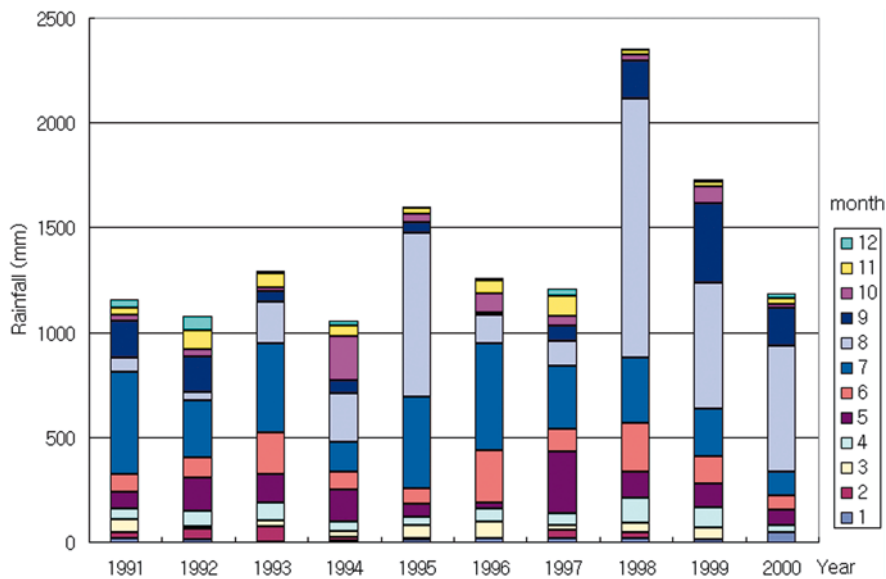


Fig. 3
Rainfall distribution in the Seoul area (1991–2000)

Materials and methods

This study was carried out in two steps: at first, data collection and database construction using a Geographic Information System (GIS), and secondly, statistical analysis to investigate the characteristics of $\text{NO}_3\text{-N}$ according to the well depth, rainfall distribution and land uses. The SPSS statistical package for Windows was used for the statistical analysis. The *P* values for testing the results of statistical analysis were calculated at the 95% significance level. Data collected and used to construct the GIS database include monthly rainfall records, well locations, groundwater quality, land-use data from references and reports and related GIS data.

Figure 3 represents the rainfall distribution in each month from 1991 to 2000. The average annual rainfall is 1,360 mm/year (Seoul Metropolitan City 2001a). Based on the rainfall data for 10 years, the dry and rainy seasons were identified to be from December to February and from July to September, respectively. More than 70% of the rainfall occurs during the 4 months in summer, resulting from seasonal rain fronts in June and July and typhoons in August and September. In this study, the characteristics of $\text{NO}_3\text{-N}$ in groundwater were analyzed according to the rainfall distribution.

The Seoul Metropolitan City (2001b) constructed land-use data, which are based on 1:1,000-scale digital topographic and cadastral maps, and the land uses were classified at three levels (Table 1). As a first-order classification, land uses are classified into urban and open space, which occupy about 58% and about 42% of the total Seoul area, respectively (Seoul Metropolitan City 2001b). In the second-order classification, urban areas are divided into residential areas, commercial and business areas, mixed residential and business areas, industrial areas, public facilities, transportation facilities, urban infrastructure facilities, denuded areas and inaccessible areas. Open spaces include grasslands, field crop areas, forest areas,

managed green areas, and rivers and streams (Fig. 4). In the third-order classification, residential areas are further classified into detached houses, apartment houses, and traditional houses. Transportation facilities are subdivided into railroad, road, airport and related facilities. Urban infrastructure facilities, the potential contamination sources, are classified into sewage treatment facilities, storm-water retention ponds, water supply reservoirs, waste landfills, power plants, waste incineration facilities and waste transfer stations.

A database of well locations and $\text{NO}_3\text{-N}$ concentrations was constructed in 2001. Presently, about 30,000 wells, which include in-use and abandoned wells, constitute the database (Seoul Metropolitan City 2002). Results from a total of 1,988 wells, monitored in January, February, July, August, September and December of the year 2000, were used for statistical analysis.

The well depth was classified based on the maximum depth of weathered bedrocks, which is 40 m, in the Seoul area (Seoul Metropolitan City 1996). If the borehole depth was less than 40 m, the well was considered a “shallow well.” Wells deeper than 40 m were identified as “deep wells.” Figure 5 shows the location of wells used in this study, which are classified according to the well depth in the dry season: (a) 306 shallow wells and (b) 225 deep wells. Figure 6 shows the location of the wells in the rainy season: (a) 849 shallow wells and (b) 608 deep wells.

Figure 7 represents a flowchart for developing the statistical models for contamination assessment. At first, a test for normality is performed for evaluating the distribution of $\text{NO}_3\text{-N}$ concentration grouped according to the well depth and rainfall distribution. Secondly, a test for determining the difference of $\text{NO}_3\text{-N}$ concentration in groundwater is performed in each group according to the well depth and rainfall distribution. Thirdly, the radius of influence (ROI) is evaluated to identify the area over which a certain land use could affect $\text{NO}_3\text{-N}$ concentration in

Table 1
Land uses in the Seoul area (Seoul Metropolitan City 2001b)

First order	Second order	Third order	Area	Area (%)
Urban area	Residential area	Detached houses, apartment houses, traditional houses	163.5 km ²	26.93
	Commercial and business area		35.6 km ²	5.86
	Mixed residential and business area		29.8 km ²	4.91
	Industrial area		8.2 km ²	1.35
	Public facilities	Educational facilities, governmental offices, hospitals or health care centers, research center, large-scale sports facilities	30.6 km ²	5.04
	Transportation facilities	Railroad and related facilities areas, road and related facilities areas, airport area	62.2 km ²	10.24
	Urban infrastructure facilities	Sewage treatment facilities, storm water retention ponds, water supply reservoirs, waste landfill areas, power plants, waste incineration facilities, agricultural and fishery products markets, waste transfer stations	7.9 km ²	1.30
	Denuded areas		14.9 km ²	2.45
	Inaccessible areas	Non-surveyable areas, military areas	12.7 km ²	2.09
	Forest and open space	Grassland areas		17.5 km ²
Field crop areas			29.8 km ²	4.91
Forest areas			156.9 km ²	25.84
Managed green areas			2 km ²	0.33
Rivers, streams, or wetlands			35.6 km ²	5.86

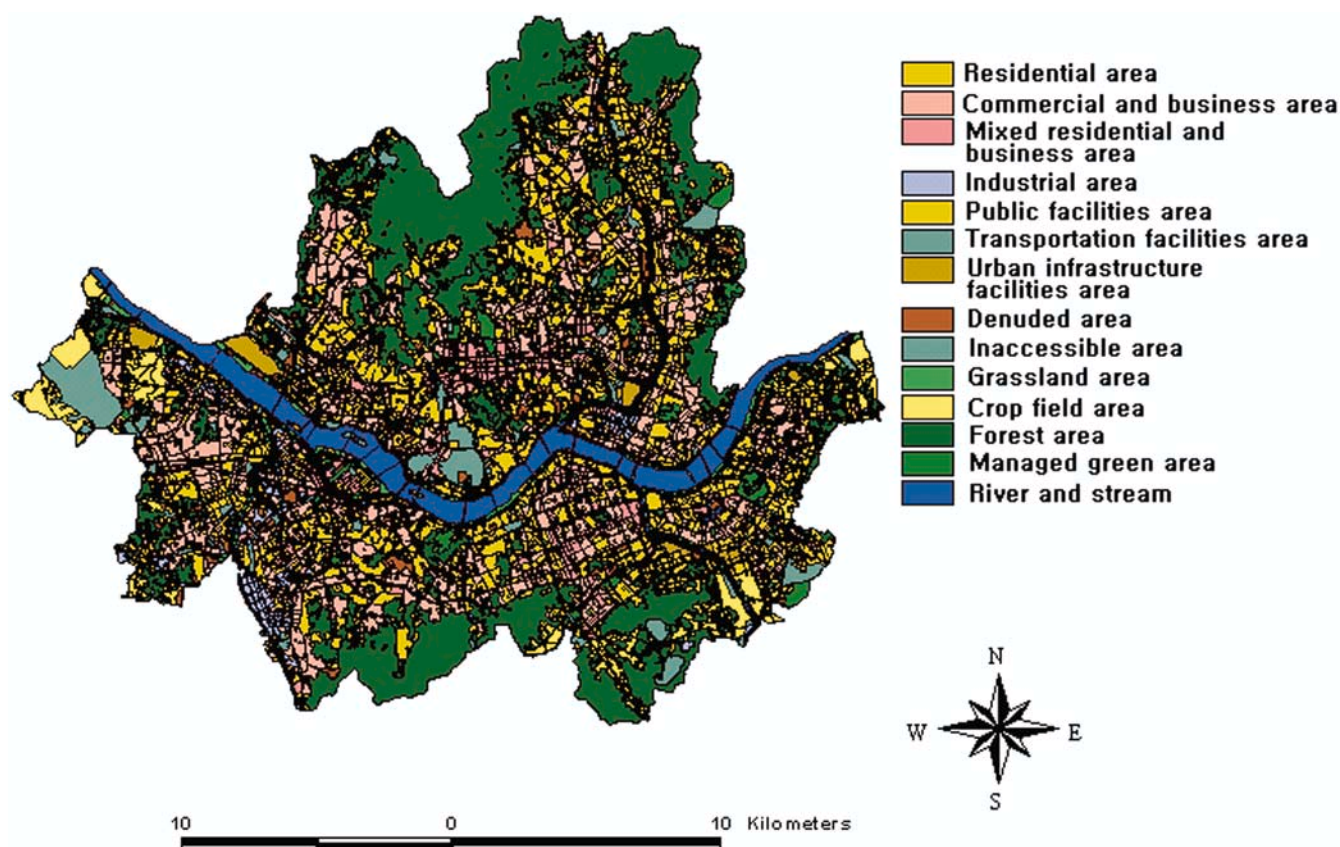


Fig. 4
Land uses in the Seoul area

wells. Finally, correlation and linear regression analyses are performed to analyze the relationship between land uses within specific ROI and NO₃-N concentrations. Before these statistical methods are performed, two GIS processes

are implemented to calculate the area of specific land uses within the radius of influence. Those two processes are buffering at the well center and overlaying the land-use data and buffer layer.

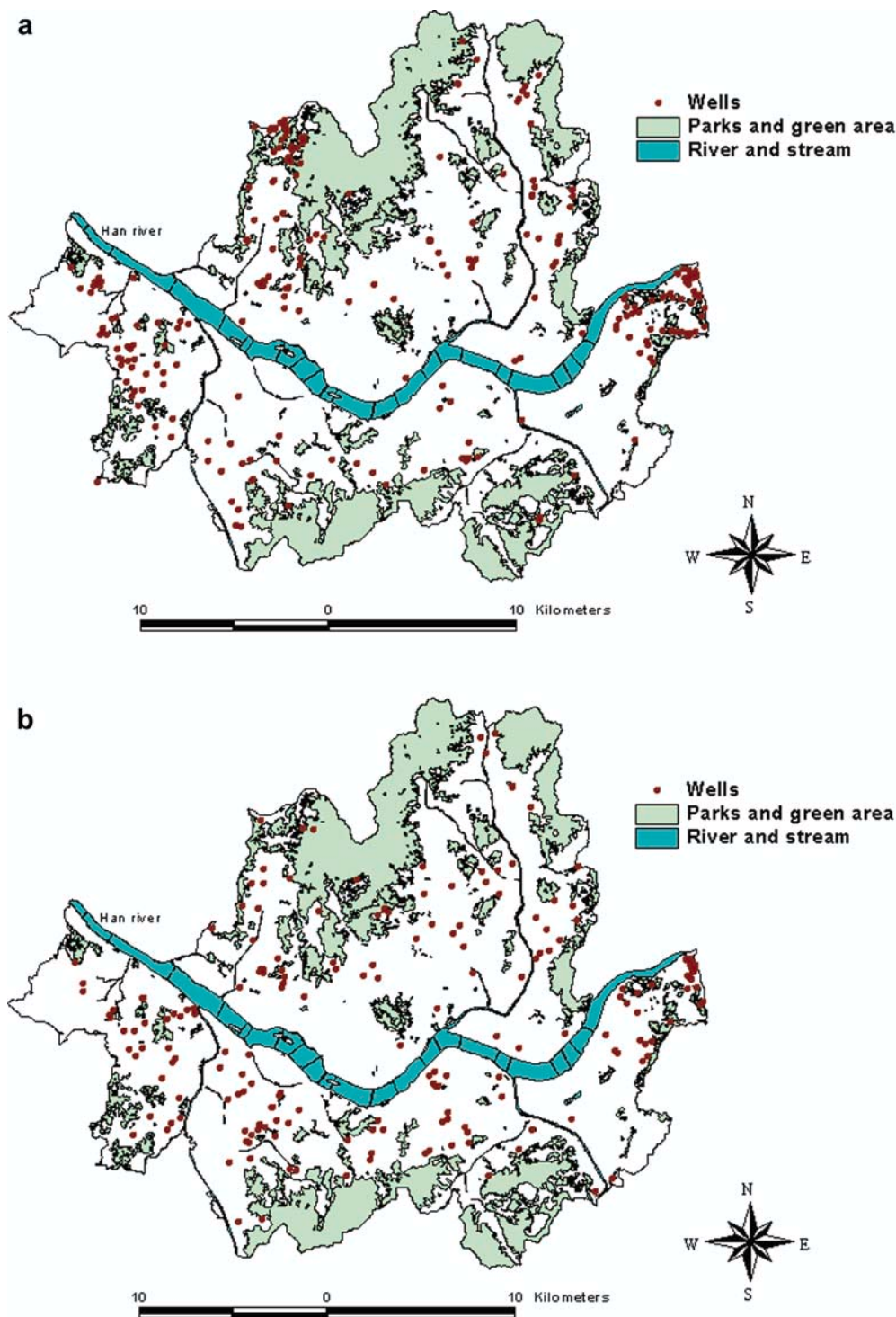


Fig. 5

Distribution of wells monitored in 2000. a Shallow wells in the dry season, b deep wells in the dry season

Results

Present $\text{NO}_3\text{-N}$ distribution

To estimate the extent of contamination from each type of land use, $\text{NO}_3\text{-N}$ concentrations were analyzed based on the number of wells and ratios exceeding the drinking water standard of 10 mg/L (OECD 1986). The mean $\text{NO}_3\text{-N}$ concentration for the 1,988 wells was 6.33 mg/L. Among these wells, 463 wells (23.29%) exceeded the drinking water standard. Table 2 shows the number of wells located

in various second-order land uses and the number of contaminated wells with $\text{NO}_3\text{-N} > 10$ mg/L in each category.

The land use that represents the highest ratio of $\text{NO}_3\text{-N}$ concentrations exceeding 10 mg/L is the field crop area. A total of 70 wells (38.46%) of the total 182 wells exceeded 10 mg/L of $\text{NO}_3\text{-N}$. The next highest is the mixed residential and business areas (28.73%), followed by grass land areas (22.22%), industrial areas (20.37%), residential areas (19.95%), commercial and business areas (19.10%)

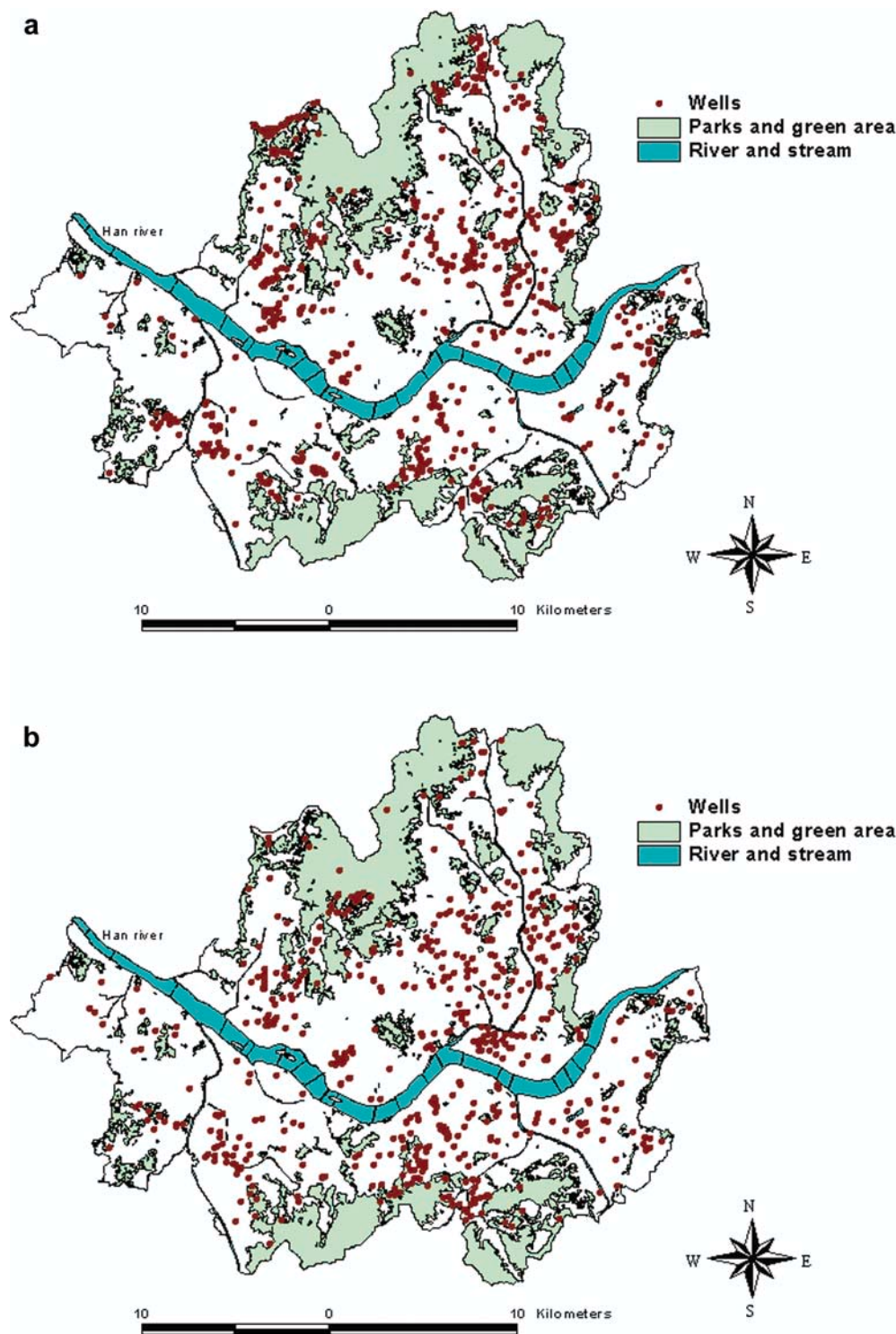


Fig. 6 Distribution of wells monitored in 2000. a Shallow wells in the rainy season, b deep wells in the rainy season

and so on. $\text{NO}_3\text{-N}$ in groundwater generally derived from urban wastewaters, various fertilizers in agriculture and livestock wastes in farming areas (Hem 1992; Huber and others 2000).

In Seoul, field crop areas and mixed residential and business areas show a higher percentage of $\text{NO}_3\text{-N}$ contaminated wells than the average ratio (23.29%) of all the wells exceeding the standard, indicating that general causes of nitrate contamination could be related to the

land uses. The land uses represented by the lowest ratios are inaccessible areas, urban infrastructure facilities, managed green areas and forest (Table 2).

A normality test for distribution of $\text{NO}_3\text{-N}$ concentration is performed to determine if parametric or non-parametric test procedures may be employed. Table 3 shows the results of the normality test for distribution of $\text{NO}_3\text{-N}$ concentration, which present the non-normality in each data group, and as an example, the Kolmogorov-Smirnov value

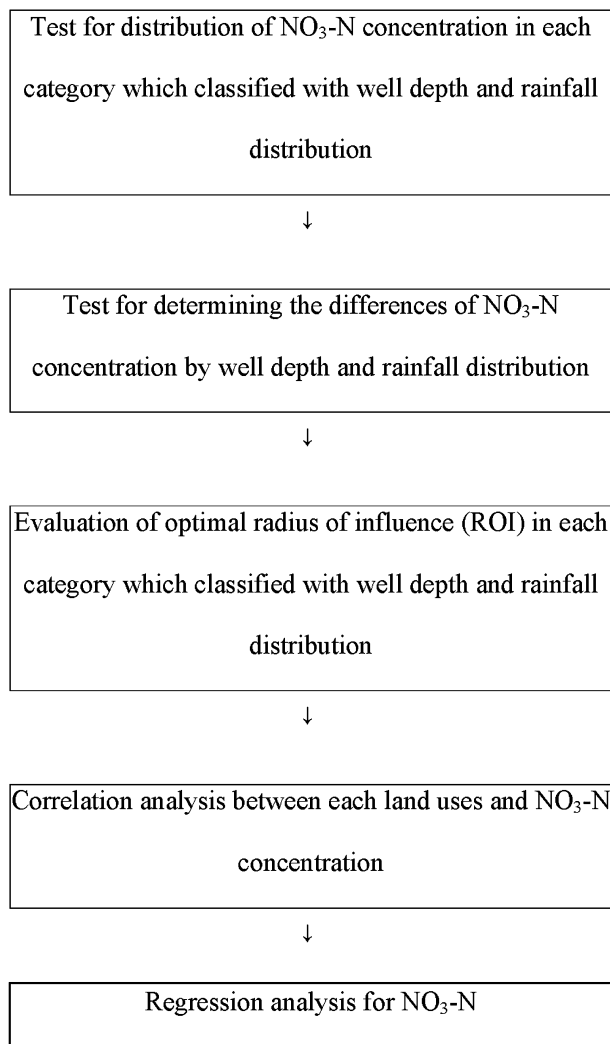


Fig. 7

Developing statistical models for contamination assessment of $\text{NO}_3\text{-N}$

for $\text{NO}_3\text{-N}$ concentration in shallow wells and the rainy season is 3.667, and the P value is <0.001 . This value shows that the P value is smaller than the significance level 0.05, which implies distribution of $\text{NO}_3\text{-N}$ concentration is non-

normal. Other P values are smaller than the significance level of 0.05, demonstrating a non-normal distribution. A Mann-Whitney U-test is performed to verify the population difference of $\text{NO}_3\text{-N}$ concentration in wells. In the comparison of rainy and dry seasons (Table 4a, b), both shallow and deep wells show Z values of -5.068 and -2.888 , respectively. These Z values indicate that there are significant differences of $\text{NO}_3\text{-N}$ concentration in different seasons.

Comparing shallow and deep wells, Z values of the rainy and dry seasons were -6.591 and -4.089 , respectively, showing the statistically significant differences of $\text{NO}_3\text{-N}$ concentrations with depth (Table 4c, d).

During the same season, shallow wells tend to have higher concentrations than deep wells (Table 4a vs. b), indicating that groundwater quality in shallow wells has deteriorated more. For wells of the same depth, the $\text{NO}_3\text{-N}$ values were higher in the dry season than the rainy season (Fig. 8, Table 4c vs. d). Consequently, the results of the Mann-Whitney U-test show that the differences of concentration of $\text{NO}_3\text{-N}$ due to well depth and seasonal rainfall are significant at the 95% confidence level.

Establishment of radius of influence

The radius of influence (ROI) is the range over which a particular land use has the major influence on the groundwater quality. It is very important to establish the optimum ROI when analyzing the relationship between land use and groundwater quality. If the ROI is set at too short a distance, land use characteristics are not reflected properly in the groundwater quality. On the contrary, if the ROI is set too long, unrelated land uses may appear to influence the groundwater quality. Use of a ROI concept ignores the regional and local flows of groundwater or contamination moving into the ROI area of a well. The hydrogeological conditions in the Seoul area show extreme disturbance caused by such influences as well pumping, subway pumping, sewer system leaks, the municipal water-supply system and non-permeable pavement. Regional groundwater flow is therefore difficult to model in the Seoul area (and detailed groundwater flow modeling even more so within the various ROI areas). Thus, although the

Table 2

Range of nitrate-N concentrations in wells according to each land use. (LN1 residential area, LN2 commercial and business area, LN3 mixed residential and business area, LN4 industrial area, LN5 public facilities area, LN6 transportation facilities area, LN7 urban infrastructure facilities area, LN8 denuded area, LN9 inaccessible area, LN10 grassland area, LN11 field crop area, LN12 forest area, LN13 managed green area)

Land use	Area (ha)	Number of wells	Conc. of $\text{NO}_3\text{-N}$ (mg/L)				No. of wells with $\text{NO}_3\text{-N}$ above 10 mg/L	Ratio (%)
			Min	Max.	Median	Mean		
LN1	16,352.2	406	0.001	58.1	4.35	5.87	81	19.95
LN2	3,555.6	487	0.001	27.3	3.70	5.38	93	19.10
LN3	2,975.7	616	0.001	84.0	4.90	6.82	177	28.73
LN4	822.2	54	0.2	28.2	4.40	6.33	11	20.37
LN5	3,057	58	0.1	20.0	3.95	5.07	7	12.07
LN6	6,221.9	50	0.01	20.8	1.75	4.29	9	18.00
LN7	787.3	18	0.001	10.2	2.35	2.47	1	5.56
LN8	1,486.3	40	0.1	18.7	3.25	5.35	7	17.50
LN9	1,271.6	2	5.8	10.0	7.9	7.9	0	0.00
LN10	1,751.2	9	0.3	80.0	4.9	16.77	2	22.22
LN11	2,977.1	182	0.4	47.9	6.40	9.90	70	38.46
LN12	15,692.7	51	0.05	20.0	4.10	4.73	4	7.84
LN13	204.2	15	0.93	22.7	3.10	4.54	1	6.67
Total	57,155	1,988	0.001	84.0	4.40	6.33	463	23.29

Table 3

Results of testing for normality in the distribution of NO₃-N concentration by well depth and season

Types		Number of wells	Kolmogorov-Smirnov value	P value
(a) Shallow wells	Rainy season	849	3.667	<0.0001
	Dry season	306	2.923	<0.0001
(b) Deep wells	Rainy season	608	4.521	<0.0001
	Dry season	225	2.886	<0.0001

Table 4

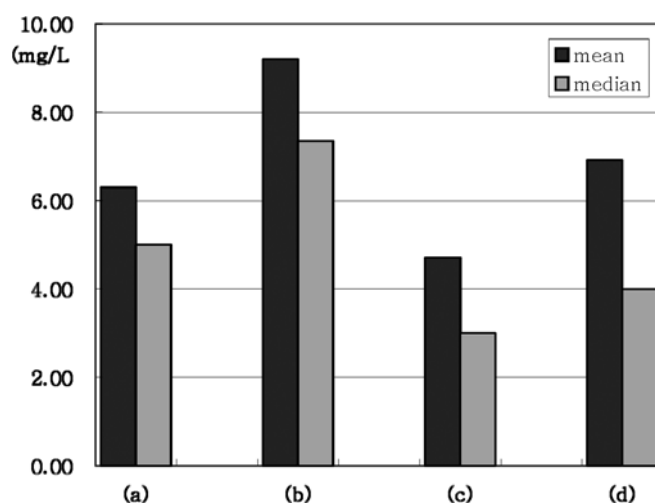
Results of the Mann-Whitney U-test: analyzing the difference in NO₃-N concentration by comparing the well depth and rainfall distribution

Types		Number of wells	Mean	Variance	Average rank	Z value	P value
(a) Shallow wells	Rainy season	849	6.30	30.19	548.14	-5.068	<0.0001
	Dry season	306	9.20	90.83	660.85		
(b) Deep wells	Rainy season	608	4.71	27.26	402.37	-2.888	0.004
	Dry season	225	6.92	63.46	456.54		
(c) Rainy season	Shallow wells	849	6.30	30.19	709.46	-6.591	<0.0001
	Deep wells	608	4.71	27.26	643.17		
(d) Dry season	Shallow wells	306	9.20	90.83	289.34	-4.089	<0.0001
	Deep wells	225	6.92	63.46	234.26		

ROI concept has a few shortcomings, it reflects the effects of contaminant sources, such as specific land uses, in the local area. Figure 9a shows the land uses within specific ROIs from a well located in an urban area, which is composed mainly of residential areas, commercial and business areas and mixed residential and business areas. Figure 9b shows the same information in green and open-space areas, mainly forest areas, field crop areas and grassland areas.

In this study, to establish the optimum ROI, we considered various distances from the well center (50 to 400 m in 50-m increments) and then performed multiple regression analysis between specific land uses and concentrations of NO₃-N at each ROI. As previously mentioned, land uses that correlate with the higher NO₃-N concentrations in wells are mixed residential and business areas and field crop areas. Therefore, these two land uses were used as independent variables for the regression analysis. The optimum ROI was determined based on the highest regression coefficient.

Figure 10 shows the regression coefficients between the two land uses and the NO₃-N concentration according to the radius of the ROI. The regression coefficients for NO₃-N in the dry season increased from 50 to 200 m of radius, then continuously decreased above 250 m of radius in both shallow and deep wells. Therefore, in the dry season, 200 m from a well is considered to be an optimum ROI to capture land-use characteristics affecting the concentration of NO₃-N in groundwater. In the rainy season, an ROI of 250 m has the highest regression coefficient in both shallow and deep wells. Consequently, ROIs of 200 and 250 m for the dry and rainy seasons, respectively, were selected to evaluate the statistical models for contamination assessment. This result confirms Chon and Ahn's (1998) study concluding that the ROI for NO₃-N from potential point sources would be 200 m in the Guro-Ku area, a western part of the Seoul metropolitan area.

**Fig. 8**

Mean and median concentration of NO₃-N according to well depth and rainfall distribution. a Shallow wells in the rainy season, b deep wells in the rainy season, c shallow wells in the dry season, and d deep wells in the dry season

Development of statistical models

Correlation analysis expresses the relationship between two variables. Previous results for the Mann-Whitney U-test show that the NO₃-N concentrations are different according to the well depth and rainfall distribution. Therefore, correlation analysis was performed to see how much each land use could affect the NO₃-N concentration. Data distribution for NO₃-N concentration showed the non-normality in each group, which is classified by well depth and rainfall distribution, so the Spearman's rho test is performed.

Various land uses within the ROI include residential, commercial and business, mixed residential and commercial, industrial, public facilities, transportation facilities, urban

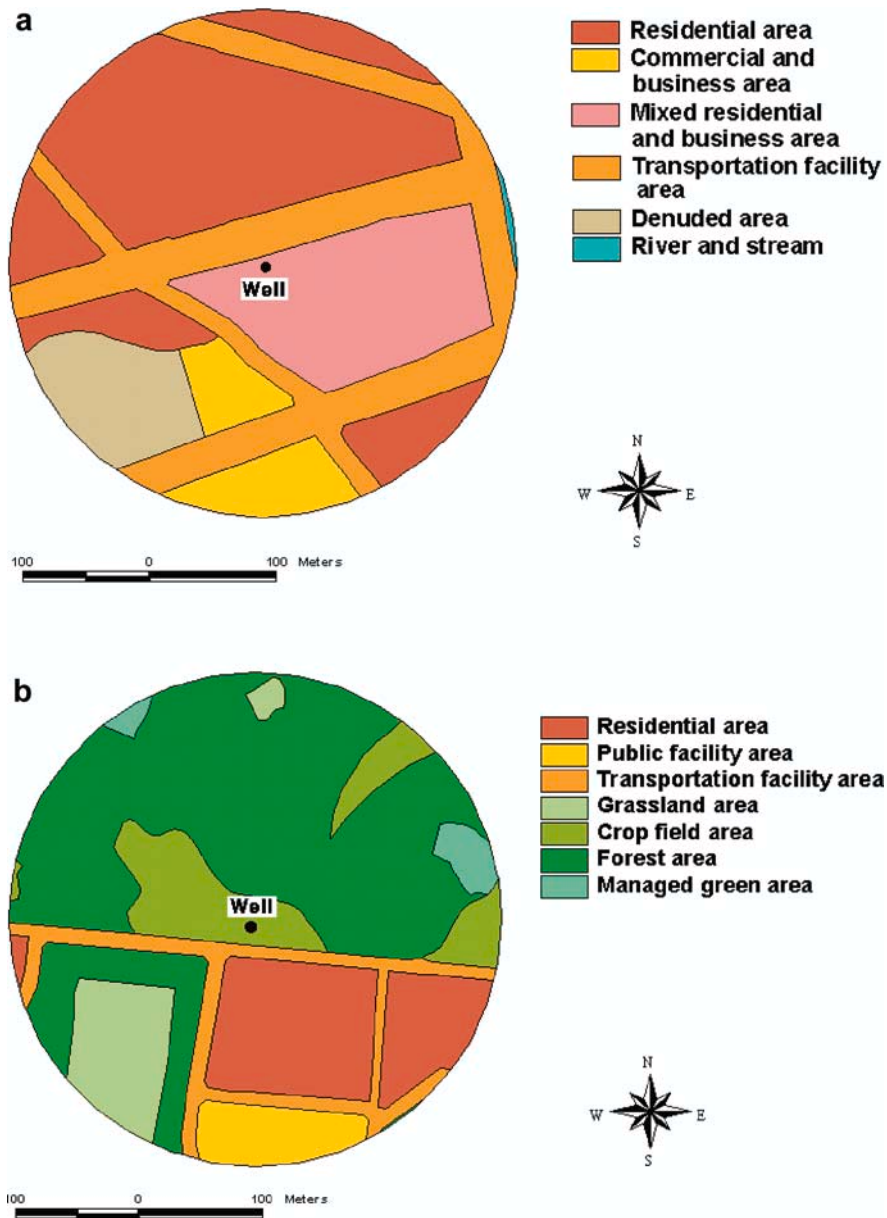


Fig. 9

Land uses within the 200 m ROI. **a** Land uses in the urban area, **b** land uses in the green and open-space area

infrastructure facilities, denuded, inaccessible, grassland, field crop, forest, managed green, and river and stream. Table 5 shows the results of the Spearman's rho test between land uses within the ROI and $\text{NO}_3\text{-N}$ concentration at the center wells. In the dry season, the land uses showing positive coefficients at the significance level of 0.05 are deep wells in field crop areas (LN11). Field crop areas (LN11) show notably higher correlation coefficients than other land uses, implying that the occurrence of $\text{NO}_3\text{-N}$ in wells is related to field crop areas. In the rainy season, residential areas (LN1), mixed residential and business areas (LN3) and public facilities areas (LN5) show positive coefficients in shallow wells at the significance level of 0.05. Mixed residential and business areas (LN3) show positive coefficients with deep wells at the significance level of 0.05. Correlation coefficients are generally low, implying that the groundwater quality was affected by factors other than

the land uses. If the land uses that have negative correlation were increased in some areas, the $\text{NO}_3\text{-N}$ concentration would be decreased.

Multiple regression analysis was used to draw a relationship between the $\text{NO}_3\text{-N}$ concentration and land use. The regression equation is expressed in the following form (Davis 1986):

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (1)$$

where Y_i is the $\text{NO}_3\text{-N}$ concentration as a dependent variable. The independent variables x_1, x_2, \dots, x_p represent the areas of each land use. $\beta_1, \beta_2, \dots, \beta_p$ are regression coefficients and β_0 is a constant. The results of the regression analysis for $\text{NO}_3\text{-N}$ in the dry season are summarized in Table 6: (a) for shallow and (b) for deep wells. Among the land uses used to establish the equation, the

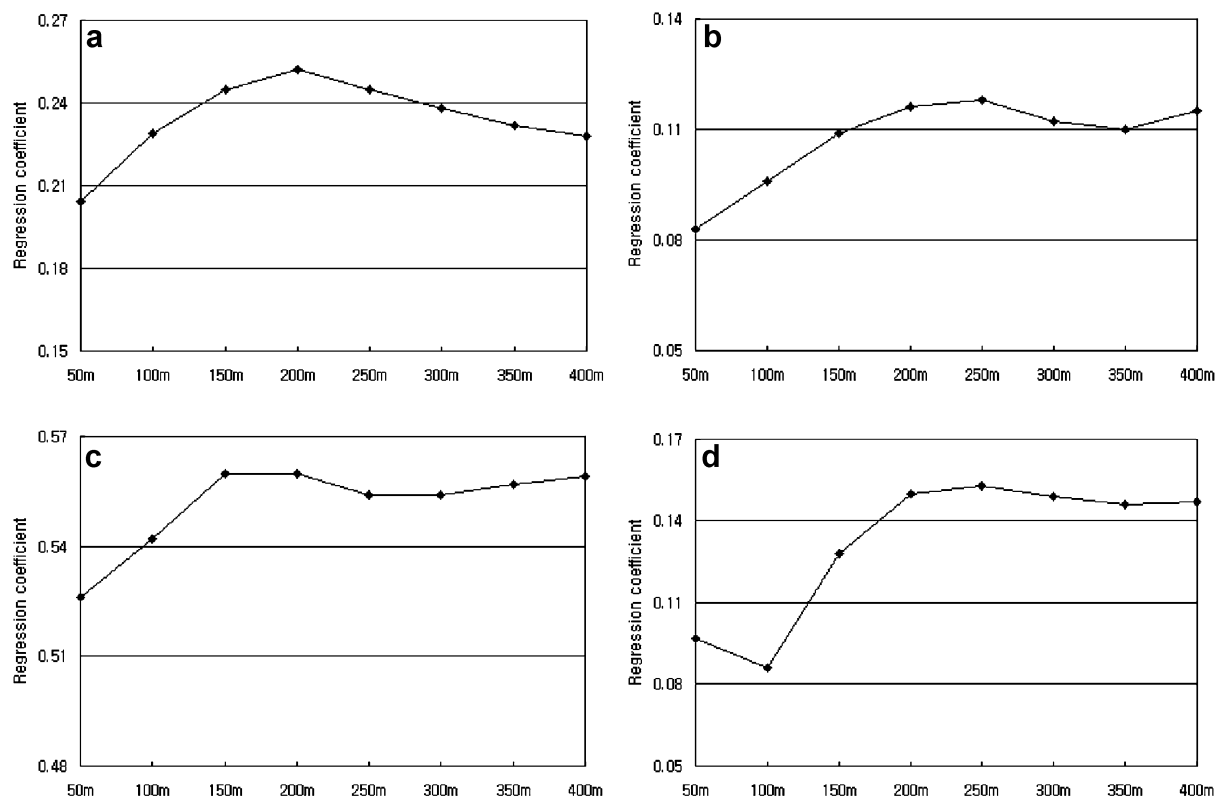


Fig. 10

Regression coefficient variation for $\text{NO}_3\text{-N}$ by increment of radius of influence. **a** Shallow wells in the dry season, **b** shallow wells in the rainy season, **c** deep wells in the dry season, and **d** deep wells in the rainy season

mixed residential and business area (LN3), grassland area (LN10) and field crop area (LN11) have positive coefficients in shallow wells in the significance level of 0.05. In deep wells, grassland area (LN10) and field crop area (LN11) have positive coefficients that denote significance. The land uses with positive coefficients tend to increase the $\text{NO}_3\text{-N}$ concentration. The field crop area, which affects the $\text{NO}_3\text{-N}$ concentration most of all, is represented in both shallow and deep wells. Standardized coefficients for field crop areas are 0.289 and 0.523 in shallow and deep wells, respectively.

Two regression equations for $\text{NO}_3\text{-N}$ in shallow (Eq. 2) and deep (Eq. 3) wells were derived at the significance level of 0.05 by the F-test. Units for $\text{NO}_3\text{-N}$ concentration and land use are mg/L and m^2 , respectively. For example, in Eq. (3), if the field crop area is increased by 1 m^2 , the $\text{NO}_3\text{-N}$ concentration is increased by $1.091\text{E}-04 \text{ mg/L}$. In Eqs. (2) and (3), LN3 is the mixed residential and business area, LN10 is the grassland area and LN11 is the field crop area.

$$\begin{aligned} \text{NO}_3 - \text{N}(\text{mg/L}) = & 5.340 + 5.844\text{E} - 05 \times \text{LN3}(\text{m}^2) \\ & + 2.641\text{E} - 04 \times \text{LN10}(\text{m}^2) \\ & + 8.268\text{E} - 05 \times \text{LN11}(\text{m}^2) \end{aligned} \quad (2)$$

$$\begin{aligned} \text{NO}_3 - \text{N}(\text{mg/L}) = & 4.540 + 2.346\text{E} - 04 \times \text{LN10}(\text{m}^2) \\ & + 1.091\text{E} - 04 \times \text{LN11}(\text{m}^2) \end{aligned} \quad (3)$$

Table 5

Results of the Spearman's rho test between land use and $\text{NO}_3\text{-N}$ concentration. (LN1 residential area, LN2 commercial and business area, LN3 mixed residential and business area, LN4 industrial area, LN5 public facilities area, LN6 transportation facilities area, LN7 urban infrastructure facilities area, LN8 denuded area, LN9 inaccessible area, LN10 grassland area, LN11 field crop area, LN12 forest area, LN13 managed green area, LN14 river and stream)

Land uses	Dry season		Rainy season	
	(a) Shallow	(b) Deep	(c) Shallow	(d) Deep
LN1		-0.232**	0.097**	
LN2		-0.201**		
LN3		-0.187**	0.084*	0.083*
LN4				
LN5		-0.278**	0.120**	
LN6	-0.136*	-0.331**	-0.084*	-0.088*
LN7	-0.134*	-0.143*	-0.086*	
LN8				0.122**
LN9				
LN10		0.430**		
LN11				
LN12	-0.169**			
LN13				
LN14	-0.155**	-0.235**	-0.123**	-0.141**

Significance level: * 0.05; ** 0.01

The results of the regression analysis for $\text{NO}_3\text{-N}$ in the rainy season are summarized in Table 7. In shallow wells, the mixed residential and business area (LN3) have positive coefficients, indicating contributions to the increase of $\text{NO}_3\text{-N}$ concentration. In deep wells, the residential area (LN1), the mixed residential and business area (LN3), the industrial area (LN4), the denuded area

Table 6

Regression analysis results for NO₃-N in the dry season. (LN1 residential area, LN3 mixed residential and business area, LN10 grassland area, LN11 field crop area)

(a) Regression analysis in shallow wells (R ² : 0.094)						
Variable	Parameter estimate	Standard error	Standardized coefficient	T value	P> T	F value (P>F)
Constant	5.340	0.947		5.638	<0.0001	10.453 (<0.0001)
LN3	5.844E-05	0.000	0.205	3.211	0.001	
LN10	2.641E-04	0.000	0.180	3.192	0.002	
LN11	8.268E-05	0.000	0.289	4.629	<0.0001	
(b) Regression analysis in deep wells (R ² : 0.332)						
Constant	4.540	0.494		9.194	<0.0001	55.073 (<0.0001)
LN10	2.346E-04	0.000	0.147	2.599	0.010	
LN11	1.091E-04	0.000	0.523	9.267	<0.0001	

Table 7

Regression analysis results for NO₃-N in the rainy season. (LN1 residential area, LN3 mixed residential and business area, LN4 industrial area, LN6 transportation facilities, LN8 denuded area, LN11 field crop area, LN12 forest area, LN14 river and stream)

(a) Regression analysis in shallow wells (R ² : 0.055)						
Variable	Parameter estimate	Standard error	Standardized coefficient	T value	P> T	F value (P>F)
Constant	7.910	0.519		15.235	<0.0001	12.387 (<0.0001)
LN3	7.953E-06	0.000	0.070	1.975	0.049	
LN6	-5.208E-05	0.000	-0.153	-4.106	<0.0001	
LN12	-2.518E-05	0.000	-0.127	-3.284	0.001	
LN14	-2.009E-04	0.000	-0.181	-5.211	<0.0001	
(b) Regression analysis in deep wells (R ² : 0.048)						
Constant	2.579	.477		5.403	<0.0001	6.104 (<0.0001)
LN1	1.532E-05	0.000	0.129	2.960	0.003	
LN3	2.219E-05	0.000	0.191	4.235	<0.0001	
LN4	2.508E-05	0.000	0.097	2.328	0.020	
LN8	3.855E-05	0.000	0.093	2.288	0.022	
LN11	5.718E-05	0.000	0.164	3.936	<0.0001	

(LN8) and the field crop area (LN11) have positive coefficients. Among these land uses, the mixed residential and business area has the highest regression coefficients, suggesting that the mixed residential and business area (LN3) would cause the most increase of NO₃-N concentration. Unlike the results for the dry season, the R² value in shallow wells in the rainy season (r²=0.055) is higher than that of deep wells (r²=0.048). This is probably a result of infiltration of surface runoff and subsequent recharge processes. The R² values in the rainy season (Table 7: r²=0.055 and 0.048 for shallow and deep wells, respectively) are lower than those in the dry season (Table 6: r²=0.094 and 0.332 for shallow and deep wells, respectively). This result is interpreted as showing that the hydrogeological system in the rainy season becomes more complicated than in the dry season and the effects of land use could be masked by the other processes such as recharge.

The equations in the rainy season were also evaluated by use of the F-test. The equations in shallow and deep wells are presented in Eqs. (4) and (5), respectively. In equations, LN1 is the residential area, LN3 the mixed residential and business area, LN4 the industrial area, LN6 the

transportation facilities, LN8 the denuded area, LN11 the field crop area and LN14 the river and stream.

$$\begin{aligned} \text{NO}_3 - \text{N}(\text{mg/L}) = & 7.910 + 7.953\text{E} - 06 \times \text{LN3}(\text{m}^2) \\ & - 5.208\text{E} - 05 \times \text{LN6}(\text{m}^2) \\ & - 2.518\text{E} - 05 \times \text{LN12}(\text{m}^2) \\ & - 2.009\text{E} - 04 \times \text{LN14}(\text{m}^2) \end{aligned} \quad (4)$$

$$\begin{aligned} \text{NO}_3 - \text{N}(\text{mg/L}) = & 2.579 + 1.532\text{E} - 05 \times \text{LN1}(\text{m}^2) + 2.219\text{E} \\ & - 05 \times \text{LN3}(\text{m}^2) + 2.508\text{E} - 05 \times \text{LN4}(\text{m}^2) \\ & + 3.855\text{E} - 05 \times \text{LN8}(\text{m}^2) \\ & + 5.718\text{E} - 05 \times \text{LN11}(\text{m}^2) \end{aligned} \quad (5)$$

Through the results of the four types of regression analysis, we could identify the grassland area (LN10) and the field crop area (LN11) to be major land uses affecting NO₃-N concentration in the dry season, whereas the mixed residential area (LN3) was a major source in the rainy season. Constants representing the Y intercept are higher in shallow wells than deep wells, also indicating that

shallow wells would be more vulnerable to contamination than deep wells. This confirms that the mean of $\text{NO}_3\text{-N}$ concentration in shallow wells is higher than that of deep wells.

Discussion

In this study, $\text{NO}_3\text{-N}$ concentrations in urban wells were classified according to the rainfall distribution (as the dry and the rainy season) and the depth of the well (as shallow and deep). The $\text{NO}_3\text{-N}$ concentrations were analyzed using GIS techniques and statistical methods to draw correlations with land uses.

The $\text{NO}_3\text{-N}$ concentrations in the rainy season were lower than those in the dry season. This could be attributed to rainfall recharge and resulting dilution effects on the $\text{NO}_3\text{-N}$ concentration. The $\text{NO}_3\text{-N}$ concentrations are higher in shallow wells than in deep wells in both dry and rainy seasons. This indicates that $\text{NO}_3\text{-N}$ sources are mainly located at the surface and these contaminants flow into the groundwater system by infiltration and recharge processes.

The optimum radius of influence (ROI) so that land use could show significant effects on $\text{NO}_3\text{-N}$ concentration in wells is shown to be about 200 m in the dry season and 250 m in the rainy season for both shallow and deep wells. Land uses that affect $\text{NO}_3\text{-N}$ concentration are the grassland area and the field crop area in the dry season, and the mixed residential and business area in the rainy season. Therefore, to manage the groundwater quality for $\text{NO}_3\text{-N}$ concentration, it is necessary to manage land use according to the rainfall distribution.

In the statistical methods, we use land use areas as the non-point sources that affect $\text{NO}_3\text{-N}$ in wells. Correlation results between $\text{NO}_3\text{-N}$ concentration and land uses are generally low and weak even if they are significant at a 0.05 level. The R^2 values from the regression analysis also show the very low correlation. These results suggest that many contaminant sources exist in addition to the land uses, and that hydrogeological conditions are severely disturbed in urban areas. In the next step, to obtain better results, we need to enhance our database to include the number of people, sewer lines and sewage disposal facilities, geological conditions and soil characteristics.

Acknowledgements The authors thank the employees of the Seoul Metropolitan City for providing the land-use data, well location data and groundwater quality data.

References

- Chon HT, Ahn HI (1998) Assessment of groundwater contamination using the Geographic Information System. *J Korean Soc Groundwater Environ* 5:129–140
- Custodio E (1997) Groundwater quantity and quality changes related to land and water management around urban areas: blessings and misfortunes. *Groundwater in the urban environment: problems, processes and management*. Balkema, Rotterdam, pp 11–22
- Davis JC (1986) *Statistics and data analysis in geology*, 2nd edn. Wiley, New York, pp 468–478
- Foster SSD (2001) The interdependence of groundwater and urbanization in rapidly developing cities. *Urban Water* 3:209–216
- Hem JD (1992) *Study and interpretation of the chemical characteristics of natural water*, 3rd edn. United States Government Printing Office, Washington, pp 124–126
- Hong YS, Rosen MR (2001) Intelligent characterization and diagnosis of the groundwater quality in an urban fractured-rock aquifer using an artificial neural network. *Urban Water* 3:217–228
- Huber A, Bach M, Frede HG (2000) Pollution of surface waters with pesticides in Germany: modeling non-point source inputs. *Agric Ecosyst Environ* 80:191–204
- Kim YJ, Lee SM (2001) *The methods and principles of groundwater conservation area designation and its management*. Seoul Development Institute, Seoul, Korea, pp 3–11
- Kim YJ, Won JS, Lee SM (2000) *Guidelines for developing groundwater management plan of Seoul*. Seoul Development Institute, Seoul, Korea
- Kim YY, Lee KK, Sung IH (2001) Urbanization and the groundwater budget, metropolitan Seoul area, Korea. *Hydrogeol J* 9:401–412
- Lasserre F, Razack M, Banton O (1999) A GIS-linked model for the assessment of nitrate contamination in groundwater. *J Hydrol* 224:81–90
- Ministry of Science and Technology (1997) *Study on protection and reclamation for the groundwater resources in Seoul area*. Korea Institute of Geoscience and Mineral Resources, Korea, pp 1–10
- OECD (Organization for Economic Co-operation and Development) (1986) *Water pollution by fertilizers and pesticides*. OECD, Paris, France
- Seoul Metropolitan City (1996) *Report of groundwater basic investigation*. Rural development corporation. Seoul, Korea, pp 19–49
- Seoul Metropolitan City (2001a) *Report of statistics for Seoul Metropolitan City*. Seoul, Korea
- Seoul Metropolitan City (2001b) *Biotope mapping and guideline for establishment of ecopolis in Seoul*. Seoul Development Institute, Seoul, Korea
- Seoul Metropolitan City (2002) *The development of groundwater management system in Seoul area*. Seoul Development Institute, Seoul, Korea
- Trauth R, Xanthopoulos C (1997) Non-point pollution of groundwater in urban areas. *Water Resour Res* 31: 2711–2718