

지리 공간 자료의 다중회귀분석을 이용한 제주도 남측사면 용천수의 시기별 질산성 질소 농도 예측

정윤영^{1,2} · 고동찬^{1*} · 강봉래³ · 고경석¹ · 유용재²

¹한국지질자원연구원 지하수연구실, ²충남대학교 지구환경과학부, ³제주특별자치도 환경자원연구원

Prediction of Seasonal Nitrate Concentration in Springs on the Southern Slope of Jeju Island using Multiple Linear Regression of Geographic Spatial Data

Youn-Young Jung^{1,2}, Dong-Chan Koh^{1*}, Bong-Rae Kang³, Kyung-Suk Ko¹ and Yong-Jae Yu²

¹Groundwater department, Korea Institute of Geoscience and Mineral Resources (KIGAM), Daejeon 305-350, Korea

²Department of Geology and Earth Environmental Sciences, Chungnam National University Daejeon 305-764, Korea

³Jeju Special Self-Governing Province Institute of Environmental Resources Research, Jeju 690-700, Korea

Nitrate concentrations in springs at the southern slope of Jeju Island were predicted using multiple linear regression (MLR) of spatial variables including hydrogeological parameters and land use characteristics. Springs showed wide range of nitrate concentrations from <0.02 to 86 mg/L with a mean of 20 mg/L. Spatial variables were generated for the circular buffer when the optimal buffer radius was assigned as 400 m. Selected regression models were tested using the p values and Durbin-Watson statistics. Explanatory variables were selected using the adjusted R², Cp (total squared error) and AIC (Akaike's Information Criterion), and significance. In addition, mutual linear relations between variables were also considered. Small portion of springs, usually <10% of total samples, were identified as outliers indicating limitations of MLR using circular buffers. Adjusted R² of the proposed models was improved from 0.75 to 0.87 when outliers were eliminated. In particular, the areal proportion of natural area had the greatest influence on the nitrate concentrations in springs. Among anthropogenic land uses, the influence of nitrate contamination is diminishing in the following order of orchard, residential area, and dry farmland. It is apparent quality of springs in the study area is likely to be controlled by land uses instead of hydrogeological parameters. Most of all, it is worth highlighting that the contamination susceptibility of springs is highly sensitive to nearby land uses, in particular, orchard.

Key words : spring, nitrate contamination, multiple linear regression, volcanic aquifer, land use

제주도 남측사면에서 산악지역부터 해안지역에 걸쳐 분포하는 용천수에 대해 풍수기와 갈수기의 2회에 걸쳐 측정된 NO₃ 농도를 수리지질학적 인자 및 토지 이용 특성 인자를 포함하는 공간 변수들의 다중선형 회귀모형으로 예측하였다. 용천수의 NO₃ 농도는 평균 20 mg/L이며, <0.02~86 mg/L의 범위를 보여 인위적 오염의 정도가 매우 다양하다. 공간 변수는 용천수를 중심으로 원형 버퍼를 설정하여 추출하였으며, 수정결정계수 증가율과 원형 버퍼의 제한점을 고려하여 반경 400 m를 최적 범위로 설정하였다. 선택된 회귀 모형들은 p-값과 더빈-왓슨 통계치에 근거하여 모두 통계적으로 유의하였다. 설명변수는 수정결정계수, Cp (total squared error), AIC (Akaike's Information Criterion)등을 기준으로 선택하였으며 변수들의 유의성과 다중공선성을 확인하여 최적 회귀 모형을 제시하였다. 일부 용천수들은 이상치로 확인되었으나 전체 시료의 10%이내였으며, 이들은 원형 버퍼를 사용하는 다중회귀분석의 한계를 지시한다고 할 수 있다. 변수의 유의성 기준으로 선정된 최적 회귀 모형의 결정계수는 이상치 제거 전이 0.74-0.79, 제거 후가 0.86-0.87의 범위로 높은 설명력을 보여주었으며, 자연지역 면적 비율이 용천수의 NO₃ 농도에 가장 큰 영향력을 가지는 것으로 나타났다. 용천수 NO₃ 농도에 대한 인위적 토지이용의 영향력은 최적 버퍼 반경에서 두 조사 시기 모두

*Corresponding author: chankoh@kigam.re.kr

과수원 > 주거지역 > 밭의 순으로 나타났다. 이러한 결과는 제주도 남측사면 용천수의 수질이 수리지질학적 인자보다는 토지 이용 특성에 크게 좌우됨을 지시하며, 용천수의 오염 취약성이 주변의 지표 오염원, 특히 과수원 분포에 민감함을 보여준다.

주요어 : 용천수, 질산성 질소 오염, 다중 선형 회귀, 화산암 대수층, 토지 이용

1. 서 론

제주도 용천수의 수질특성은 크게 자연적인 광물화에 의한 수질진화와 인위적 오염에 의해 좌우된다 (Lee *et al.*, 2002; Jung *et al.*, 2010). 인위적 오염에 의해 영향을 받은 용천수는 상대적으로 NO_3 및 Cl 의 농도가 높아 지표오염에 의한 용질 증가경향을 보인다. 제주도의 용천수는 화산암 대수층의 지하수 유동구조를 파악하는데 있어 중요한 연구대상으로 인위적 요소에 의한 지하수내 오염물질의 유입에 취약하다. 이는 제주도가 육지에 비해 높은 투수성의 화산암체로 구성되어 있기 때문에 토지이용과 같은 공간변수는 지하수 오염과 관련하여 매우 중요하다 (Youn and Park, 1998). 관정 지하수에서도 해안지역에서 배경치 이상으로 검출되는 NO_3 농도는 농업활동과 주거지역 분포 등 토지 이용과 밀접한 관계된 것으로 알려져 있다 (Koh *et al.*, 2007). 질소동위원소조성을 이용한 오염원별 기여도 평가에서 NO_3 가 주변지역뿐만 아니라 광역적인 범위에서 대수층으로 유입됨이 보고되었다 (Song *et al.*, 1999).

토양 및 대수층 매질 특성, 지형 등의 수리지질학적 공간변수는 전통적으로 DRASTIC과 같은 오염 취약성 평가 (Hamm *et al.*, 2004; Panagopoulos *et al.*, 2005; Cheong *et al.*, 2008)에서 지수화되어 활용되어 왔다. 지하수의 질산성 질소 오염 예측에서는 다중선형회귀 (multiple linear regression; McLay *et al.*, 2001; Kim *et al.*, 2008), 토빗 회귀 (Tobit regression; Gardner and Vogel, 2005), 로지스틱 회귀 분석 (logistic regression; Nolan, 2001) 등의 여러 회귀분석방법이 주로 적용되는데, 수리지질학적 변수 외에 토지이용과 오염 부하량 등의 자료가 주요 독립 변수로 이용된다. 이 중 다중회귀방법은 지하수보다는 오염원의 영향이 더 직접적으로 나타나는 지표수의 질산성 질소 오염 예측에 널리 활용되었다 (Thornton *et al.*, 1998; Basnyat *et al.*, 1999; Kloiber, 2006). 지하수의 질산성 질소 오염 예측에서 다중회귀법의 대표적인 사례는 Nolan and Hitt (2006)으로 이들은 세부적인 공간 자료를 바탕으로 질산성 질소와 관련된 인자들의 기작적인 구조를 비선형 회귀로 구현하였다. 따라서 높은 투수성과 빠른 지하

수의 함양특성이 나타나는 제주도의 경우 지하수의 오염특성을 파악하는데 있어 다중회귀방법의 적용이 가능할 것으로 판단된다.

이번 연구에서는 고도에 따라 고르게 분포하는 제주도 남측사면의 용천수를 대상으로 지하수내 NO_3 농도를 종속변수로 하고, 토양형태, 지형구배, 토양심도, 토지이용 등과 같은 수리지질학적 인자와 오염원을 반영하는 공간변수들을 독립 변수로 하는 다중회귀분석을 실시하였다. 공간변수는 다양한 공간 범위에서 추출하고, 모형 예측력을 기준으로 최적 공간 규모를 설정하였다. 이러한 과정을 통해 구축된 회귀모형을 이용하여 질산성 질소 농도 예측의 적용성을 평가하고, 공간변수와 질산성 질소 농도의 상관관계와 개별 공간변수가 질산성 질소 농도에 미치는 영향력을 분석하여 주요 영향 인자를 규명하고자 하였다. 이번 연구는 인위적 토지 이용 증가에 따라 점차 지하수의 오염이 증가하는 연구지역의 지하수 수질 보전 대책 수립에 대한 기초자료로 활용될 수 있을 것으로 기대된다.

2. 연구지역

2.1. 지형 및 지질

연구지역의 지형은 북쪽의 한라산 정상을 기점으로 남측으로 갈수록 고도가 감소하며, 해발고도 600 m를 축으로 상부는 급경사면을 보이는데 반해 하부는 경사가 완만해지는 지형적 특성을 보인다. 한라산일대는 화산활동의 후기에 속하는 조면암질 암상특성에 기인한 것으로 관입상을 보이는 조면암류에 의해 하부지역의 지형특성과는 차이를 나타낸다 (Park *et al.*, 2000).

제주도는 제3기 플라이오세에서부터 제4기에 이르는 수 십 차례의 화산활동에 의해 형성된 화산섬으로 최하부의 용결응회암 및 화강암을 비롯하여 시대미상의 미고결 퇴적층, 서귀포층, 화산쇄설성 퇴적층, 화산분출암 등의 순서로 층서를 이룬다 (Won *et al.*, 2006). 연구지역의 지표의 대부분은 화산암으로 현무암, 조면현무암, 현무암질조면안산암, 조면안산암 및 조면암류로 분류되며 육지지역에 비해 높은 투수성을 보인다 (Koh *et al.*, 2005). 제주도 남측사면은 다른 지역보다 하부

층서에 서귀포층이란 불투수층의 발달되어 지하수의 유동을 규제하고 해수의 침투를 막는 차수벽 역할을 하는 것으로 보고되었다 (Koh, 1997).

2.2. 기상 및 수리지질

지난 30년간 (1971-2000년) 기상청 관측자료 (<http://www.kma.go.kr>)에 의하면 연구지역에 해당하는 서귀포시의 평균 기온은 16.2°C, 연평균 강수량은 1850.7 mm로 최근 4년 (2006-2009년)에는 평균 기온 17.2°C, 평균 강수량은 1898.1 mm로 온도상승과 강수량의 증가를 보여준다. 연구지역은 행정구역상 서귀포시에 해당하며 제주, 성산, 고산에서 측정된 기상 관측자료와 비교해보았을 때 가장 높은 기온과 강수량을 보인다. 연구지역이 가장 높은 강수량이 기록되는 이유는 풍수기에 해당하는 6~9월경 북서풍의 북태평양 고기압의 영향이 나타나기 때문이다 (Kim *et al.*, 2002). 연구지역의 고도에 따른 강수량은 고도효과에 의한 지형성강우의 영향으로 해발고도 100 m 증가 시 약 241 mm의 강수량이 증가하는 것으로 나타난다 (Jeju-do and K-water, 2003).

분출된 용암의 상대적인 냉각수축 (cooling contraction)에 의한 수직의 주상절리 발달 및 용암관 (lava tube), 용암터널 (lava tunnel) 등 투수성 수직구조가 발달된 제주도의 경우 지하수의 함양이 매우 원활하다고 할 수 있다 (Koh, 1997). 함양된 지하수는 제주도의 용암류 사이에 발달한 크링커 (clinker)층과 스큐리아 (scoria)층과 같이 높은 투수성을 보이는 층서구조를 따라 유동하게 되며, 유동하는 지하수는 유동과정에서 주변에 발달한 파쇄대나 절리를 따라 용출하게 된다. 제주도의 경우 파쇄대 및 절리 등의 발달은 암상의 경계주변에 발달되어 있으며 실제로 연구지역에 해당하는 서귀포시 일대에 분포한 용천수 168개소 중 146개소가 암상분포의 차이가 나타나는 용암류의 경계에 위치한다 (Jeju-do, 1999).

2.3. 토양 특성 및 토지 이용

연구지역의 토양은 토성 (Soil texture)에 의해 미사질양토 (SiL), 미사질식양토 (SiCL), 둥근 바위가 있는 미사질양토 (R-Cob-SiL), 바위가 있는 미사질양토 (Cob-SiL), 자갈이 있는 미사질양토 (Peb-SiL), 모래 (S), 둥근바위 (R-Cob)의 총 7개의 토양형태로 구분된다. 이들 대부분은 전형적인 화산회토로 주된 모래는 현무암이며 일부는 조면암 및 조면암질 안산암에서 유래된

다 (Jeju-do, 1997). 토양심도의 경우 연구지역의 지형 경사에 높은 연관성을 갖는데 경사의 변화가 나타나는 중산간 (해발고도 600~200 m) 지역의 경우 평균 유효토심이 약 60 cm로 나타난다 (Jeju-do, 1997). 대부분의 경작지가 분포하는 해안지역 (해발고도 200 m)의 경우 0.5~1 m의 유효토심이 차지하는 면적은 30.4%, 1 m 이상의 유효토심이 차지하는 면적은 36%로 경작지 및 인위시설이 많은 해안지역을 중심으로 깊은 유효토심을 보인다 (www.asis.rda.go.kr).

연구지역이 포함되는 서귀포시의 토지지목별 면적은 임야가 43,802 ha, 논과 밭을 포함한 경지면적이 26,562 ha, 대지 및 공장용지 등의 인위시설이 6,059 ha 순이다 (Jeju Provincial Government, 2009). 경지면적은 전체면적의 약 30.5%로 과수원을 포함한 밭이 대부분이며 논의 비율은 0.6% 미만이다. 산림 및 초지를 포함하는 임야는 경작지 및 인위시설과 해발고도 약 300 m를 기준으로 구분되며 이는 연구지역의 지형특성을 반영한다.

3. 연구방법

3.1. 용천수의 NO₃ 농도 자료

회귀분석에 사용된 용천수의 NO₃ 농도자료는 제주도 남측 사면에 분포하는 총 39개소의 용천수에서 갈수기와 풍수기에 대해 측정되었다. 총 39개소의 용천수는 고도에 따라 고지대 용천수 (해발고도 600 m 이상) 8개소, 중산간 용천수 (해발고도 600~200 m) 6개소, 해안지역 용천수 (해발고도 200 m 이하) 25개소로 구분된다 (Fig. 1). 시료채취 시기는 갈수기에는 2009년 4~6월, 풍수기에는 9월이다. 자세한 수리지구화학적 조사방법과 농도자료는 Jung *et al.* (2010)에 제시되어 있다.

3.2. 공간변수 설정

연구지역의 용천수는 고도가 감소함에 따라 지하수 유동과정에서 발생하는 물-암석 반응 및 지표오염물질의 유입에 의해 용존 광물질 함량의 증가가 나타난다 (Jung *et al.*, 2010). 지표오염원의 유입에 따른 NO₃ 및 Cl 농도 증가는 빠른 지하수의 함양이 이루어지는 연구지역에서 지표특성에 따라 영향이 나타난다. 따라서 용천수의 NO₃ 농도 예측에서 공간변수는 연구지역의 지표특성과 연관되어야 하며 고도별 용천수의 구분되는 수질특성을 반영해야 한다. 다중회귀분석을 위한 공간변수의 선정은 기존의 지표오염원 예측과 관련된

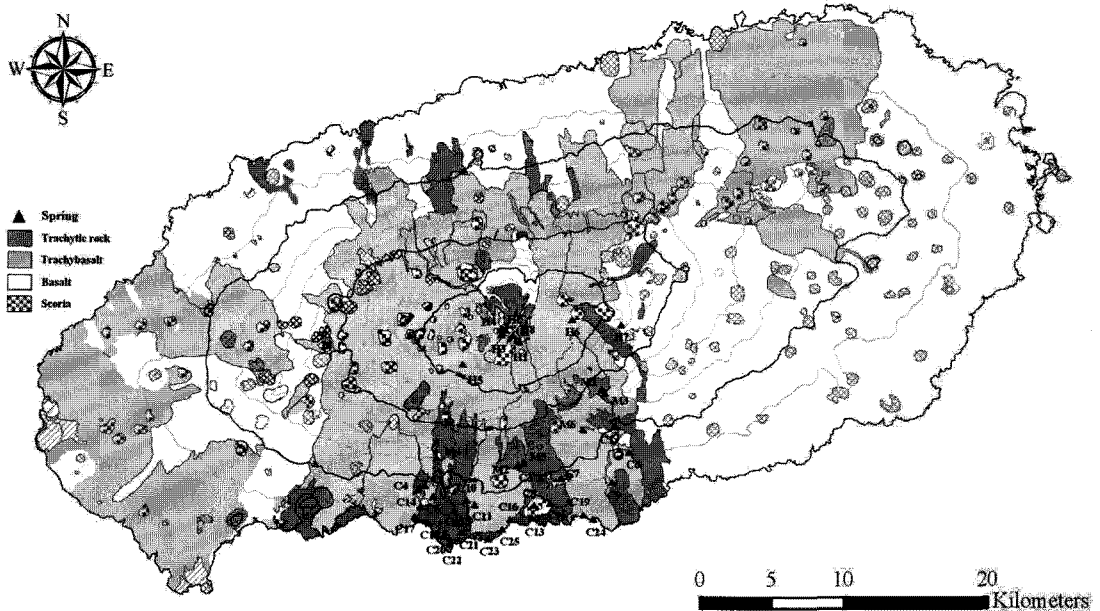


Fig. 1. Sample location map.

연구사례를 고려하여 토지이용 (Land use), 토양형태 (Soil type), 지형구배 (Topography) 토양심도 (Soil depth), 용출량 (Flow rate), 고도 (Elevation)를 선정하였다 (Gardner and Vogel, 2005; Cheong *et al.*, 2008; Kim *et al.*, 2008).

공간변수를 이용한 용천수의 NO_3 농도 예측을 위한 최적의 공간적 영향력이 있는 범위를 파악하기 위해 용천수의 위치를 중심으로 다양한 반경을 갖는 원형 버퍼 (buffer)를 적용하였으며 최적 공간 범위를 결정하기 위해 버퍼의 반경을 변화시키며 다중회귀분석을 실시하였다 (Eckhardt and Stackelberg, 1995; Kolpin, 1997; McLay *et al.*, 2001; Kim *et al.*, 2008). 공간변수와 관련된 자료는 제주특별자치도 환경자연연구원에서 제공하는 토지이용도 및 토양도를 이용하였으며 모든 공간변수에 대한 좌표정보는 TM좌표계의 중부원점을 기준으로 설정하였다.

3.3. GIS 분석을 통한 공간변수 생성

Won *et al.* (2004)은 도시화가 진행되는 소규모 지역에서 지하수의 NO_3 분포양상이 토지이용 변화와 높은 상관성을 나타낸다고 보고한 바 있다. 이는 토지이용을 하나의 공간변수로 사용할 때 토지이용은 연구지역의 특성을 고려해야한다는 것을 지시한다. 따라서 연구지역의 토지이용 변수는 연구지역 용천수의 수질특

성 및 고도에 따라 구분되는 토지이용 특성을 고려하여 크게 과수원(O), 논/밭(F), 인위시설(R), 자연지역(N)으로 분류하였다. 과수원은 일반적으로 밭에 포함되거나 해안지역의 토지이용 특성을 고려했을 때 가장 많은 토지이용 면적에 해당하기 때문에 이를 별도의 토지이용 변수로 구분하였다. 논은 제주도에서 면적이 매우 작기 때문에 과수원을 제외한 밭과 하나로 통합하여 하나의 토지이용의 변수로 활용하였다. 인위시설에는 주거지, 시설용지, 도로 등의 인위적 시설을 자연지역의 경우 산림 및 초지를 하나로 통합하여 토지이용 변수로 이용하였다. 토성에 따른 토양형태는 연구지역의 토양형태의 특성 및 분포한 위치상의 관계에 입각하여 미사질양토와 미사질식양토를 S1, 둥근바위가 있는 미사질양토와 바위가 있는 미사질양토를 S2, 자갈이 있는 미사질양토를 S3, 모래와 둥근바위를 S4로 구분하였다. 토지이용 및 토양형태에 관한 공간변수는 GIS를 이용해 버퍼 반경별 지오메트리계산을 통해 얻어진 개별 변수의 면적에 대해 총 버퍼 면적에 대한 비율 통계분석에 이용하였다 (Fig. 2).

지형구배는 낮을수록 지하수 산출능력이 좋고 지형 고도가 높아질수록 산출능력이 저하되며, 광역 지형경사가 급해짐에 따라 투수량계수와 비양수량이 감소하는 음의 상관관계를 보인다 (Lee *et al.*, 2008). 지표유출량과는 정의 상관성을 보이며 이는 오염물질의 유입

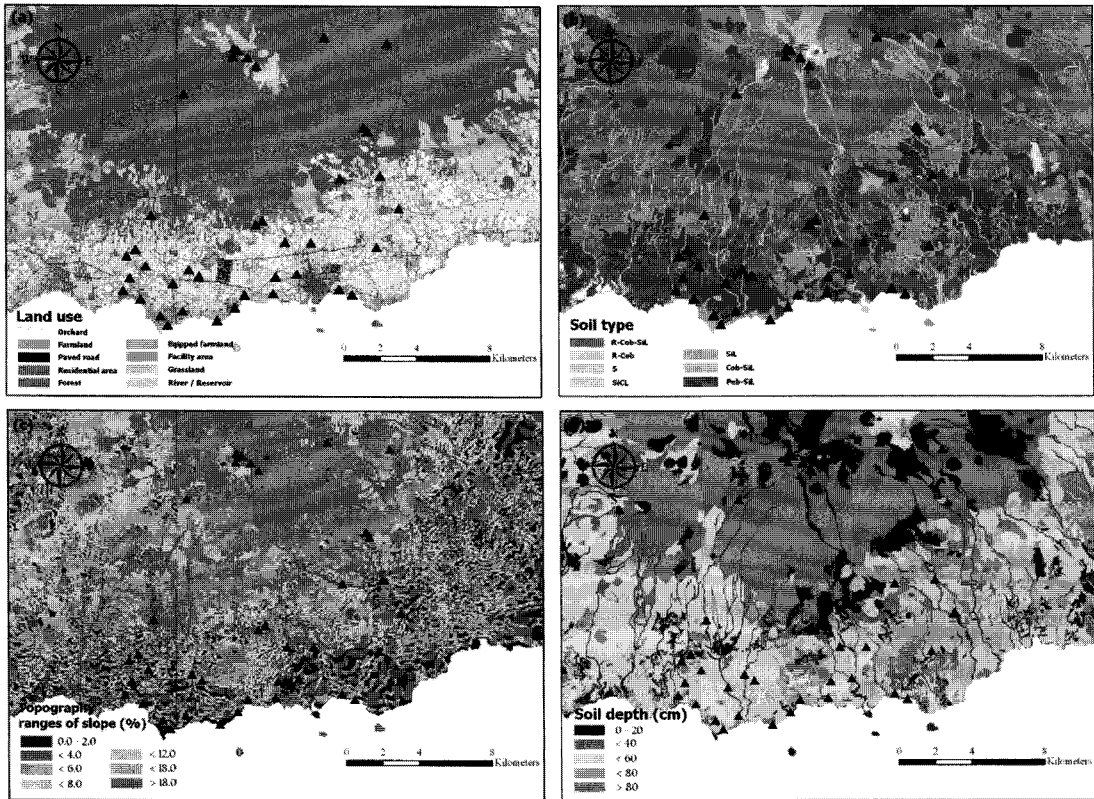


Fig. 2. GIS maps of spatial variables used for multiple linear regression. (a) Land use, (b) Soil type, (c) Topographic relief, (d) Soil depth.

과는 음의 상관성을 나타낸다 (Hamm *et al.*, 2004). 따라서 지형구배는 오염물의 거동양상에 영향을 주는 요인으로 하나의 공간변수로 다중회귀분석에 활용하였다. 연구지역의 지형구배분석을 위해 10 m 간격의 등고선도를 이용하여 GIS상에서 TIN생성을 통해 연구지역의 고도자료를 수치고도자료 (DEM)로 변환하였다. 수치고도자료는 다시 30 m × 30 m의 격자로 변환한 후 버퍼별 포함되는 격자의 평균값을 해당 버퍼 반경의 지형구배 변수로 사용하였다 (Fig. 2).

지표오염물질은 토양층을 통과하는 과정에서 중력과 표면장력에 의한 모세관현상에 의해 수직 및 수평 방향으로 확산되게 된다. 확산과정이 진행되면서 토양층은 점차 포화상태에 이르게 되고 모세관현상에 의한 지표오염물질의 이동은 점차 약해진다 (KIGAM, 2008). 이는 토양심도와 연관되어 토양심도가 깊을수록 지표오염물질의 확산이 더더진다는 특성을 보이므로 토양심도는 질소와 같은 지표오염물의 거동양상을 지시하는 공간변수로 다중회귀분석에 응용할 수 있을 것이

라 판단하였다. 토양심도는 GIS상에서 각각의 버퍼내 토양심도가 차지하는 면적비를 이용하여 해당 버퍼별 평균 토양심도의 계산을 통해 다중회귀분석에 사용하였다 (Fig. 2).

3.4. 다중선형회귀분석

3.4.1. 분석방법 개요

다중선형회귀분석이란 두 개 이상의 독립변수가 종속변수와 선형관계를 갖고 있는 경우의 회귀분석을 뜻하는 것으로 서로 상관성을 갖고 있는 변수들 간의 관계를 규명하고 모형화하기 위한 통계적 기법이다. 이때 영향을 주는 변수를 독립변수 (independent variable) 또는 설명변수 (explanatory variable)라 하고 x로 나타내며, 독립변수에 의해 영향을 받는 변수를 종속변수 (dependent variable) 또는 반응변수 (response variable)라 하고 y로 표현한다. 종속변수 y에 영향을 미치는 독립변수가 k개인 경우 다중선형회귀모형은 관계식 (1)로 나타낼 수 있다.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

$$i = 1, 2, \dots, n \quad (1)$$

y_i : i -th measured value of the dependent variable.
 x_{ki} : i -th measured value of the response variable.
 β_k : Regression coefficient of the population.
 ε_j : A error value of the j -th measured.

다중선형회귀분석은 예측과 설명에 있어 서로 배타적이지 않으며, 두 가지 연구 문제 모두에 중점을 두고 분석 가능한 방법으로, 사면결합지수에 영향을 주는 사면 붕괴 요인간의 상관성을 평가하는 방법 및 지질조건에 따른 투수계수에 대한 추정을 가능하게 하는 등 다양한 지구과학분야에 적용되고 있다 (Kang and Um, 2007; Jun *et al.*, 2010).

이번 연구에서는 지하수 내 지표오염물의 유입양상이 어떠한 요소에 의해 지배되는지를 알아보기 위하여 종속변수는 지하수 내 지표오염물질의 유입을 반영하는 개별 용천수의 NO₃ 농도(lnNO₃)를 설정하였으며, 독립변수는 공간변수인 토지이용(O, F, R, N), 토양형태(S1, S2, S3, S4), 지형구배(lnTopo), 토양심도(lnSD), 용출량(lnFlow), 용천수가 분포한 고도(lnElev)를 설정하여 다중회귀분석을 실시하였다. 이때 종속변수인 NO₃ 농도 및 독립변수인 고도, 용출량, 지형구배, 토양심도 총 5개 변수에 관하여 자료 분포의 치우침을 감소시키고자 자연로그(ln)를 취하여 다중회귀를 실시하였다. 토지이용 및 토양형태에 대해서는 용천수별 변화가 독립적이며 각각의 값에 대한 변동폭이 크지 않아 자연로그(ln)를 취하지 않았다. 다중회귀분석을 위해 SAS 9.2를 사용하였고 용천수를 중심으로 버퍼의 반경변화를 통해 공간적 영향력이 있는 범위와 각각의 공간변수가 갖는 영향력에 대해 알아보려 하였다.

3.4.2. 모형 검정

모든 독립변수에 의해 추정된 모형이 적합한지 알아보기 위해 F-검정을 이용하였으며, 모형의 유의확률(significant probability)이 0.05를 기준으로 0에 근접한 값을 보일수록 해당하는 모형은 유의하다고 판단하였다. 이때 회귀모형의 유의확률이 0.05 이상인 경우 회귀모형은 적합하지 않다고 가정하고 해당 모형에 관하여는 기각하였다 (Kim *et al.*, 2008). 유의하다고 판단된 모형에 관해서는 잔차항들이 서로 독립적이며 정규분포 한다는 다중선형회귀분석의 가정을 알아보기 위해 더빈-왓슨 (Durbin-Watson) 통계량을 이용한 잔차분석을 시행하였다. 잔차들간에 상관관계는 더빈-왓슨

의 기준값인 2로부터 개별 버퍼 모형의 더빈-왓슨값이 양의 상관관계(0)나 음의 상관관계(4)로 치우치는지에 대한 파악을 통해 모형의 적합성을 파악하였다(Kim and Lee, 2008).

3.4.3. 독립 변수 선택 방법

모형의 유의확률 및 잔차분석을 통해 적합하다 판단되는 모형의 경우 회귀모형으로 유의함을 갖는다는 것을 의미할 뿐 각 변수의 유의성을 검정한 것은 아니다. 따라서 유의한 모형에 관하여 개별 독립변수가 유의성 검증을 통해 해당하는 종속변수에 대한 최적의 모형을 선정하여야 한다. 이를 위하여 수정결정계수 (Adj. R² : Adjusted coefficient of determination)를 사용하는 최대결정계수 선택법을 통해 해당하는 모형에 의미를 갖는 개별 변수의 선택을 수행하였다. Liao and McGee (2003)가 제안한 수정결정계수를 사용하는 최대결정계수 선택법은 고전적인 변수선택방법인 전진선택법 (forward selection), 후진제거법 (backward elimination), 단계적 선택법 (stepwise selection) 등에 비해 집합한 최적의 모형을 결정하는데 있어 가장 효율적인 변수선택방법으로 알려져 있다 (Hong *et al.*, 2005). 수정결정계수에 관한 관계식은 (2)로 나타낼 수 있다.

$$Adj. R^2 = 1 - \frac{SSE(n-p-1)}{SST(n-1)} \quad (2)$$

이때 n 은 표본의 크기, p 는 독립변수의 수, SST는 총자승합, SSE는 오차자승합을 의미한다. 수정결정계수를 사용하는 최대결정계수 선택법은 모든 가능한 회귀모형에 대해 수정결정계수를 계산하는 방법으로 결정계수 (R² : coefficient of determination)를 사용하는 최대결정계수 선택법에서 자유도를 고려하여 수정된 방법이다. 이는 결정계수를 사용하는 최대결정계수 선택법에 비해 부적절한 설명변수의 추가 또는 표본크기의 변화에 민감하지 않아 더 나은 효율성을 가지고 있다. 따라서 독립변수의 증가에 따른 변수들간의 연관성 정도를 과대 추정할 수 있는 것을 방지할 수 있다 (Liao and McGee, 2003).

각 버퍼별 유의한 독립변수의 선택에 있어 신뢰성을 높이기 위해 수정결정계수 이외에도 모형 선택의 기준이 되는 통계량인 총오차자승합 (Cp=total squared error) 과 간명적합지수 중 하나인 AIC (Akaike's Information Criterion)를 함께 고려하였다 (Mallows, 1973; Akaike, 1978). 이때 Cp와 AIC 통계량은 서로 다른 독립변수로

이루어진 모형에 대해 전체 독립변수간의 적합성 비교 시 이용하였으며, 해당 모형의 수정결정계수는 높을수록, Cp와 AIC 값은 작을수록 적절한 독립변수로 구성된 모형으로 평가하였다 (Helsel and Hirsch, 2002).

3.4.4. 개별 독립변수의 유의성 검정 및 설명력 파악
수정결정계수와 Cp, AIC를 통해 선택된 독립변수라 할지라도 상관계수가 높은 경우 변수들간의 상관성에 의해 모형이 적절한 예측값을 보이는 것처럼 높은 결정계수를 갖는다. 이러한 경향은 독립변수들 사이에 다중공선성 (multicollinearity)이 존재하기 때문이다. 수정 결정계수에 의해 선택된 최적모형에 진입한 독립변수를 대상으로 분산팽창인자 (VIF=Variance Inflation Factor)가 >10, 고유값 (Eigenvalue)가 <0.01, 조건지표 (Condition Number)가 >100인 경우 다중공선성이 존재하는 것으로 판단하였으며 해당 변수는 모형에서 제외시켰다 (Kleinbaum *et al.*, 1988; Kim and Lee, 2008). 이와 더불어 모형에 진입된 독립변수의 유의확률이 $p < 0.1$ 인 경우에 통계적으로 유의하다 판단하였다 (Bauder *et al.*, 1993). 위 조건을 모두 만족하는 변수에 대하여 표준화 회귀계수 (STB=Standardized Estimate)를 구하는 방식으로 최적모형에 진입한 개별 독립변수가 종속변수인 질소의 농도에 얼마만큼 중요도를 갖는지 파악하였다.

3.4.5. 이상치 진단

용천수 주변의 공간변수를 이용한 NO₃의 농도예측에 있어 개별 용천수는 하나의 종속변수와 여러 독립변수로 이루어진 하나의 관찰치라고 할 수 있다. 제주도에서 용천수의 경우 다양한 지하층서 구조 및 구조대의 발달양상 등에 따라 지하수의 유동경로가 개별 용천수별로 다를 수 있다. 특히 연구지역의 경우 하부에 서귀포층과 같은 불투수층이 불규칙적으로 발달해

있을 뿐만 아니라 지형구배의 변화가 심해, 하부와는 연관성이 없는 지형적 특성을 보인다는 점에서 특정 용천수의 경우 다른 용천수들과 비교했을 때 전혀 다른 지하수 유동구조가 나타날 수 있다. 이러한 상이한 지하수 유동구조를 갖는 용천수는 이번 연구에서 선정된 공간변수에 의한 NO₃ 농도 예측에 상이한 결과를 나타낼 수 있다는 점에서 Cook's D 통계량과 영향치 진단을 이용하여 전체 모형에 심각한 차이를 발생시키는 이상치를 선정하였다. Cook's D의 통계량을 통해 이상치일 가능성이 있는 관찰치를 구분하였고, 구분되어진 관찰치에 대하여 영향치 진단을 함으로써 각 개별 이상치를 선정하였다. 영향진단을 통한 이상치의 선정기준은 표준화 제외 잔차 (Rstudent : studentized deleted residual)가 ±2를 초과한 관찰치로 표준화 제외 잔차가 ±2를 초과를 초과하더라도 영향치 진단 (Hat Diag H)값이 1 미만, 예측차이 (Dffits : Difference in FITS)값이 2 미만, 회귀계수 차이 (Dbetas : Difference in BETAS)값이 2 미만일 시 이는 영향력을 갖는 이상치가 아니라 판단하였다 (Kim, 2000).

4. 결과 및 토의

4.1. 용천수의 NO₃ 오염 특성

연구지역 용천수의 갈수기 평균 NO₃ 농도는 20 mg/L 이고, 고지대, 중산간, 해안지역별로 1.0, 3.1, 29 mg/L 이다. 풍수기 평균 NO₃ 농도는 19 mg/L이며 고지대, 중산간, 해안지역 용천수별로 0.5, 0.7, 29 mg/L이다 (Table 1). 두시기 고도별 용천수의 측정된 NO₃ 농도는 누적확률분포도에서 전구간 대비 특정 구간으로 치우치는 경향이 나타나지만 자연로그를 취했을 때 좀 더 균등하게 분산되는 효과가 나타난다 (Fig. 3).

Koh *et al.* (2009)에서 제시한 제주도 지하수의 지표오염물질에 대한 영향을 지시하는 NO₃ 농도 기준인

Table 1. Distribution of nitrate concentration in springs (mg/L as NO₃)

Sampling period	Summary statistics	total	high-altitude	mid-mountain	coastal
Dry season	min	<0.02	<0.02	<0.02	3.88
	max	60.4	2.30	15.5	60.4
	mean	19.7	1.01	3.09	28.9
	median	13.8	0.78	0.10	24.5
Rainy season	min	<0.02	<0.02	<0.02	0.42
	max	86.2	1.76	3.51	86.2
	mean	19.0	0.49	0.66	29.3
	median	14.6	0.20	0.05	24.9

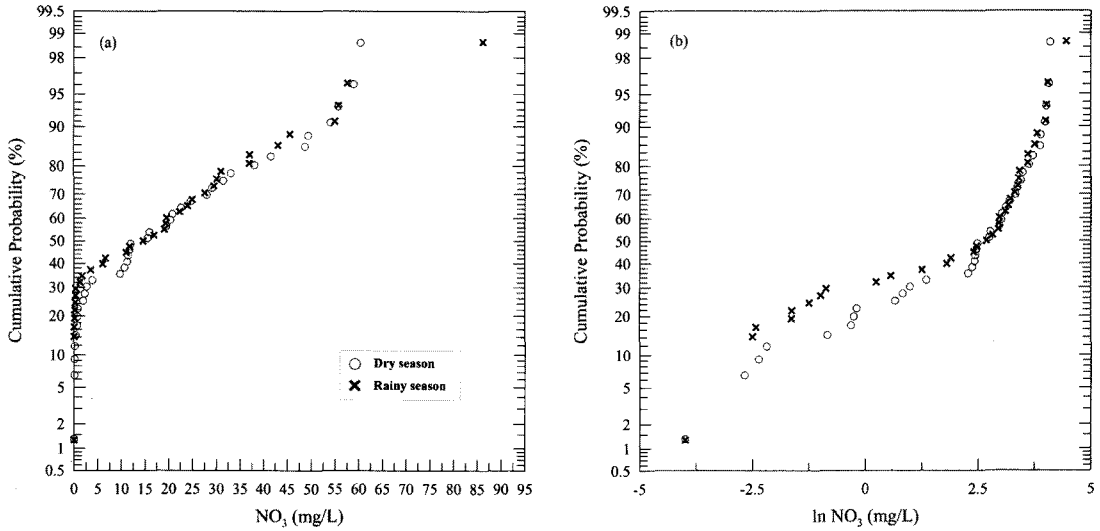


Fig. 3. Cumulative probability of measured nitrate concentrations in springs. (a) the measured nitrate, (b) taking the natural logarithm of the measured nitrate.

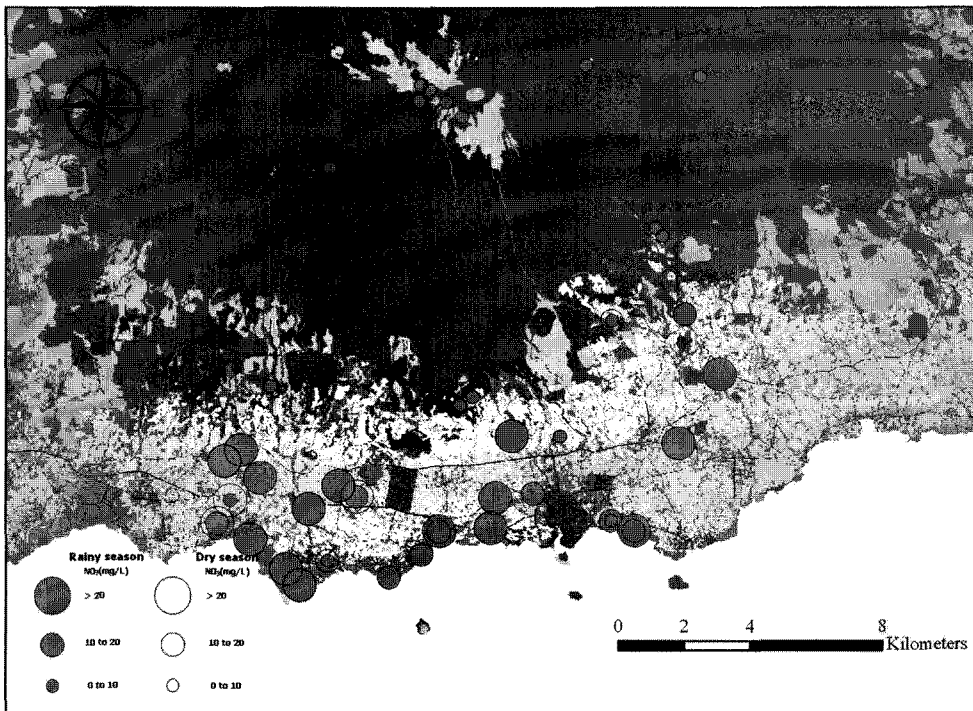


Fig. 4. Distribution of nitrate concentration in spring waters on the land-use map.

6 mg/L과 비교하면 대부분의 해안지역 용천수에서 오염의 영향이 상당히 높게 나타남을 알 수 있다. NO₃의 경우 우리나라 지하수에 나타나는 대표적인 지표오염물질로 연구지역에서 NO₃의 농도분포는 고도에 따

라 구분되는 토지이용과 높은 상관성을 갖는다는 것을 알 수 있다 (Fig. 4). 이러한 경향은 농경지와 주거시설이 대부분을 차지하는 해안지역의 특성을 반영하는 것으로 화학비료, 유기질 비료 및 생활하수 등에 의해

지하수의 오염이 진행되었음을 지시한다.

해안지역 용천수군의 경우 두 시기별로 유사한 NO₃ 농도를 보이는 용천수도 있지만 두 시기별로 차이를 보이는 용천수들도 존재한다. 이러한 차이는 용천수 별로 지표오염원을 반영하는 지하수 유동경로가 다양하며 함양되는 지점에 따라 지하수의 체류시간이 서로 다르다는 데 있다. 그럼에도 불구하고 해안지역 용천수군의 전반적인 지표오염원의 유입은 강수량의 영향을 받아 갈수기보다 풍수기에 더 강한 특성을 보인다 (Jung *et al.*, 2010). 이러한 특성은 음이온인 NO₃가 일반적으로 음전하를 띠는 토양입자에 흡착이 이루어지지 않기 때문에 투수성이 좋은 연구지역에서 시기별 비료 살포시기와 강수량의 차이가 직접적으로 지표오염물질의 지하수 내 유입을 좌우한다고 할 수 있다 (Brady and Weil, 2002). 최근에 제주도에서 화학비료의 사용량은 감소하고 있지만, 유기질 비료 사용량은 해마다 증가하는 추세로 토양내 NO₃의 지속성이 증가함에 따라 NO₃의 지하수 내 유입은 비료 살포시기보다는 강수량에 의한 영향이 커질 것으로 사려 된다 (Jeju Provincial Government, 2010).

4.2. 회귀분석 결과

분석에 이용된 모든 버퍼 반경에 대해 모든 독립변수를 고려한 모형의 유의확률 (p-value)은 0.05 이하로 통

계적으로 유의하며, 더빈-왓슨 검정치 (Durbin-Watson statistic)는 3보다 작아 잔차간 자기상관성이 작다고 할 수 있다 (Table 2). 따라서 NO₃ 농도 예측에 설정한 공간변수를 이용한 다중회귀분석이 적합하다고 할 수 있다. 수정결정계수 선택방법과 Cp, AIC 통계량을 이용해 분석에 유의한 공간변수를 선택하였으며, 변수의 유의확률 (p-value)과 다중공선성 (VIF)을 이용하여 통계적으로 유의하면서 변수간의 상관성이 존재하지 않는 최적 회귀모형을 제시하였다 (Table 3).

두 시기의 회귀모형은 버퍼 반경이 커질수록 수정결정계수 (Adj. R²)가 증가하는 특징을 보인다 (Fig. 5). 그러나 원형 버퍼의 경우 버퍼 반경이 커질수록 실제 지하수 유동 특성에 대한 반영이 어렵다. 특히 조사 지역과 같이 경사도가 큰 지역에서는 용천수의 하부 지역은 용천수 수질에 큰 영향을 준다고 보기 어렵다. 따라서 이번 연구에서는 수정결정계수 증가율이 감소하고 그 값이 0.7 이상이 되는 최소반경을 최적 반경으로 결정하였다. 이러한 기준으로 공간변수를 이용한 NO₃ 농도 예측에서 두 시기에 공통적으로 적용 가능한 최적의 버퍼 반경은 400 m라 판단된다. 두 시기의 버퍼 반경별 Cook's D 통계량과 이상치 진단을 이용해 나타난 이상치에 해당하는 용천수는 고지대 1개, 중산간 5개, 해안지역 4개로 버퍼 반경별로 차이를 보인다 (Table 4). 최적의 버퍼 반경인 400 m에서 이상치

Table 2. Results of regression analysis using all variables

Sampling period	Buffer radius (m)	R ²	Adj. R ²	F	p	DW
Dry season (n=38)	50	0.597	0.427	3.50	0.0042	1.821
	100	0.703	0.560	4.93	0.0004	1.831
	150	0.710	0.570	5.09	0.0003	1.660
	200	0.749	0.628	6.21	<0.0001	1.933
	250	0.750	0.630	6.24	<0.0001	2.028
	300	0.832	0.752	10.33	<0.0001	1.952
	400	0.851	0.780	11.92	<0.0001	2.141
	500	0.845	0.771	11.38	<0.0001	1.981
	750	0.860	0.793	12.82	<0.0001	2.102
	1000	0.829	0.748	10.13	<0.0001	2.487
Rainy season (n=39)	50	0.647	0.503	4.50	0.0007	1.668
	100	0.687	0.543	4.76	0.0004	1.710
	150	0.668	0.514	4.35	0.0008	1.754
	200	0.701	0.564	5.09	0.0003	1.947
	250	0.737	0.616	6.08	<0.0001	1.940
	300	0.784	0.685	7.87	<0.0001	2.085
	400	0.834	0.758	10.90	<0.0001	2.213
	500	0.816	0.730	9.58	<0.0001	2.038
	750	0.818	0.734	9.73	<0.0001	2.492
	1000	0.810	0.722	9.24	<0.0001	2.586

Table 3. Regression analysis of the various circular buffers using significant variables only

Sampling period	Buffer radius (m)	R ²	Adj. R ²	F	p	DW	Significant variables ¹⁾
Dry season (n=38)	50	0.541	0.500	13.34	<.0001	1.700	N, S4, lnSD
	100	0.683	0.645	17.80	<.0001	1.794	N, S1, S4, lnSD
	150	0.635	0.615	30.49	<.0001	2.055	N, lnSD
	200	0.713	0.678	20.48	<.0001	2.017	N, S1, S3, lnSD
	250	0.730	0.697	22.27	<.0001	2.078	N, S1, S3, lnSD
	300	0.755	0.726	25.45	<.0001	2.100	N, S1, S3, lnSD
	400	0.794	0.769	31.78	<.0001	2.302	N, S1, S3, lnElev
	500	0.821	0.799	37.81	<.0001	2.220	N, S1, S3, lnElev
	750	0.852	0.834	47.33	<.0001	2.071	N, S1, S3, lnElev
1000	0.813	0.791	35.96	<.0001	2.287	N, S1, S3, lnElev	
Rainy season (n=39)	50	0.598	0.563	17.35	<.0001	1.638	S2, S4, lnSD
	100	0.616	0.583	18.72	<.0001	2.029	O, F, R
	150	0.630	0.620	62.86	<.0001	2.183	N
	200	0.655	0.646	70.36	<.0001	2.197	N
	250	0.671	0.662	75.39	<.0001	2.194	N
	300	0.698	0.690	85.43	<.0001	2.191	N
	400	0.741	0.726	51.42	<.0001	2.287	N, S2
	500	0.748	0.768	38.61	<.0001	2.191	N, S3, lnElev
	750	0.792	0.774	44.38	<.0001	2.054	N, S3, lnElev
1000	0.776	0.757	40.49	<.0001	2.035	N, S3, lnElev	

¹⁾Explanatory variables were selected for those with p<0.1 and VIF<10

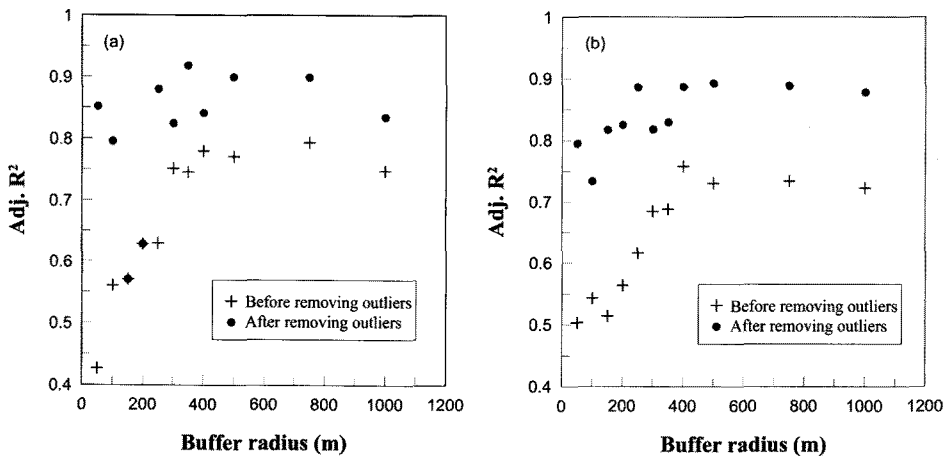


Fig. 5. Coefficients of determination for various radius of circular buffers. (a) dry season, (b) rainy season.

로 나타난 용천수는 갈수기 C7, 풍수기 C7, C10으로 400 m 이외의 버퍼 반경에서 선정된 두 시기의 이상치에 비해 매우 적게 나타남을 알 수 있다. 이는 용천수를 중심으로 하는 버퍼 반경 400 m가 연구지역을 대표할 수 있는 최적의 공간범위임을 지시한다.

최적 버퍼 반경에서 회귀모형의 수정결정계수는 이상치를 제거하여 재분석된 모형에서 더욱 높게 나타나 공간변수를 이용한 NO₃의 농도예측의 신뢰성이 증가

함을 알 수 있다 (Table 5). 두 시기의 이상치가 제거된 최종 모형에 진입한 공간변수가 설명할 수 있는 설명력은 갈수기 약 85%, 풍수기 약 84%로 용천수의 NO₃ 농도 예측에 있어 상당히 높은 수준임을 알 수 있으며 두 시기의 회귀식은 (3), (4)와 같다.

Dry season (objective)

$$\ln\text{NO}_3 = -3.822 - 8.727(N) + 4.812(S2) + 1.389(\ln\text{SD})$$

Table 4. Removed outliers to various radius of circular buffers of dry season and rainy season

Buffer radius (m)	Removed outliers	
	Dry season	Rainy season
50	H6 M2 M5 C22	M2 M5 C22
100	M2 M5	M2 M5 C7
150		M4 M5 C7 C24
200		H6 M5 C7 C24
250	M3 M5 M3 M6 C7	H6 M3 M5 M6 C7
300	C7	C7
350	H6 M6 C7 C10	C7
400	C7	C7 C10
500	M1 M6 C7	M1 M6 C7
750	M6 C7	M6 C7
1000	M4 M6	M4 M6 C7

$$+ 0.461(\ln\text{Elev}), R^2=0.848 \quad (3)$$

Rainy season (objective)

$$\ln\text{NO}_3=4.892 - 10.316(N) - 2.263(S1) - 3.981(S3) + 0.561(\ln\text{Elev}), R^2=0.843 \quad (4)$$

이상치의 제거 전 회귀모형과 제거 후 회귀모형에서 NO₃의 농도분포에 영향력을 갖는 공간변수는 자연지역으로 음(-)의 가장 높은 표준화 회귀계수 값을 보인다. 이는 자연지역이 차지하는 면적이 클수록 NO₃의 농도는 감소한다는 것을 지시하는 것으로 연구지역이 높은 투수성의 화산암으로 이루어져있으며 고도에 따라 구분되는 토지이용 특성을 보인다는 점을 고려했을 때 용천수 주변의 인위적 토양이용에 따른 지표오염물질의 지하수내 유입이 이루어짐을 파악할 수 있다. 자연지역 다음으로 높은 영향력을 미치는 공간변수는 토양형태로 S1, S2, S3가 두 시기와 이상치 제거 전후의 회귀모형에서 차이를 보이며 나타난다. 표준화 회귀계수가 음(-)의 부호를 띠는 S2가 분포한 위치는 주로 고지대에 해당하며 표준화 회귀계수가 양(+)의 부호를 띠는 S1과 S3가 분포한 위치는 주로 해안지역이다. 이는 토양형태가 고도 차이에 기인한 영향보다는 지형학적 특징에 영향을 받는다는 것을 의미하는 것으로 연구지역의 토성 특성이 경사가 급하고 고도가 높은 지역일수록 모래함량이 많고, 구릉지와 평탄한 지역에서 미사와 점토의 함량이 높게 나는 점이 반영됨을 알 수 있다 (Jeju-do, 1997). 고도의 표준화 회귀계수는 양(+)의 부호를 띠어 고도가 높을수록 NO₃의 농도가 증가함을 지시한다. 용천수의 분포가 200 m 이하의 해안지역에 편중되어 있고, 중산간지역에는 용천

수가 거의 존재하지 않는 공간적인 편향성을 고려할 때 이는 조사지역 전체보다는 용천수가 밀집한 고도 200 m 이하의 해안지역의 특성을 지시한다고 볼 수 있다. Jung *et al.* (2010)은 동일 조사지역에서 고도 25 m 이하의 해변 용천수가 상위 해안지역 용천수에 비해 지표오염물질 유입특성이 작게 나타나고, 이는 자연지역에서 기원한 기저지하수의 비중 증가에 따른 희석효과를 반영하는 것으로 해석한 바 있다. 이번 연구에서 제시된 회귀모형은 자료수의 부족으로 기저지하수의 유입 효과는 고려하지 못하였다. 따라서, 저지대에서 토양 유형이나 토지 이용 등의 지표 특성과 질산성 질소 오염간의 상관관계를 더 명확히 구분하기 위해서는 해안에 인접한 용천수는 제외하는 것이 좋을 것으로 판단된다. 그러나, 해안지역의 토지 이용이 거의 전적으로 인위적이란 점에서 저지대 시료만으로는 자연지역을 포함하는 다양한 공간변수의 영향을 평가하기 어렵다. 따라서 이번 연구에서는 고지대에서 해안지역에 이르는 전체 시료를 활용하였다.

토양심도의 표준화 회귀계수는 양(+)의 부호로 깊이에 따른 표면장력에 의해 지표오염물질의 이동이 느려진다는 일반적인 특성을 반영하지 않는다. 연구지역의 경우 토양의 대부분은 화산회토로 구성되어 있어 전용적밀도가 낮아 풍식에 대한 저항성이 낮으며 공극률과 투수성이 높아 하향침투가 빠른 특성을 보인다는 점을 고려했을 때 작물재배를 위해 개간된 경작지의 특성을 지시한다고 할 수 있다.

4.3. 이상치에 대한 고찰

제주도의 지하수 유동구조는 화산층서와 관련되어 용천수별로 서로 다른 다양한 지하수 유동특성을 보인다는 점에서 모든 용천수를 포함하는 회귀모형으로 NO₃ 농도를 예측하는 것은 모형의 신뢰성을 낮출 수 있다. 특히, 지하수 유동이나 공간변수 특성이 다른 용천수들과 크게 다른 용천수는 이러한 문제를 더욱 크게 할 수 있다. 따라서 버퍼 반경별 Cook's D 통계량과 이상치 진단을 이용해 이상치에 해당하는 용천수에 대해 파악하였다 (Fig. 6).

고지대 용천수인 H6의 NO₃ 농도는 갈수와 풍수기 모두 가장 낮은 수치로 분석상의 검출한계 (0.02 mg/L)의 값과 매우 유사한 값을 갖는다 (Jung *et al.*, 2010). 이번 연구에서는 고지대 및 중산간 용천수의 낮은 NO₃ 농도는 고지대와 중산간의 지형 및 수리질학적 요소를 반영하는 수질특성으로 판단하였다. 따라서 NO₃ 농도가 검출한계에 가까운 용천수에 대해 제외하지 않

Table 5. Comparison of regression models before and after removal of outliers in the optimal buffer radius based on significance of variables

Sampling period	Outlier	n ¹⁾	R ²	Adj. R ²	F	p	Variables	Parameter Estimate	Standard Error	p	VIF	t	STB
Dry season							Intercept	5.943	1.268	<0001		4.69	
	Included	38	0.794	0.769	31.78	<0001	N	-8.715	1.049	<0001	5.289	-8.31	-1.510
							S1	-3.596	1.309	0.0096	2.626	-2.75	-0.352
							S3	-4.353	1.079	0.0003	3.021	-4.03	-0.554
		InElev					0.383	0.205	0.0709	3.658	1.87	0.282	
		Intercept					-3.822	3.383	0.2671		-1.13		
	Removed	37	0.865	0.848	51.13	<0001	N	-8.727	0.908	<0001	5.733	-9.61	-1.496
							S2	4.812	0.852	<0001	3.192	5.65	0.656
							InSD	1.389	0.781	0.0847	2.766	1.78	0.192
		InElev					0.461	0.153	0.0049	2.979	3.02	0.339	
Rainy season							Intercept	3.366	0.321	<0001		10.48	
	Included	39	0.741	0.726	51.42	<0001	N	-7.265	1.031	<0001	3.156	-7.04	-1.062
							S2	2.197	1.282	0.0951	3.156	1.71	0.258
							Intercept	4.892	1.278	0.0006		3.83	
	Removed	37	0.860	0.843	49.24	<0001	N	-10.316	1.106	<0001	6.028	-9.33	-1.513
							S1	-2.263	1.332	0.10	2.698	-1.70	-0.185
							S3	-3.981	1.087	0.00	3.107	-3.66	-0.427
		InElev					0.561	0.208	0.0112	3.977	2.69	0.355	

1) number of springs used for the regression.

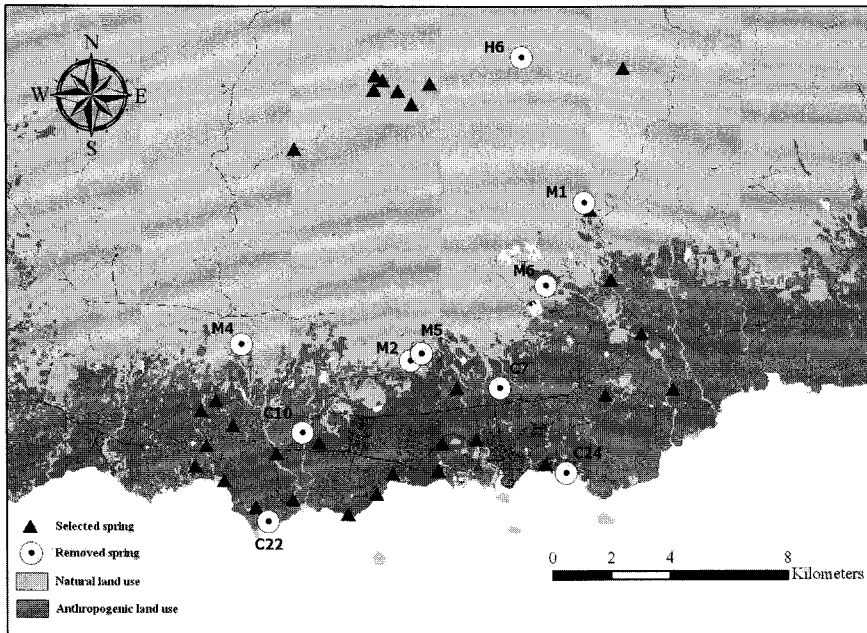


Fig. 6. Distribution of outliers removed from regression analysis.

은 점이 회귀분석 상에서 이상치로 선정되었다고 파악된다. 일반적으로 검출한계 이하의 값을 갖는 시료의 경우 정규분포에서 벗어난 특성을 나타내는데 이 경우 Tobit regression을 이용하여 회귀분석을 실행할 수 있다 (Gardner and Vogel, 2005).

해발고도 300 m 인근의 중산간 용천수의 경우 대부분의 버퍼 반경에서 이상치로 나타난다. 이러한 경향은 연구지역에서 중산간 용천수가 분포한 위치와 높은 상관성을 갖는다고 사려된다. 중산간의 경우 고지대의 자연적 특성이 지배적인 토지이용 특성과 해안지역의 인위적 토지이용 특성이 구분되는 경계부에 해당하기 때문이다 (Fig. 4). 또한 연구지역의 지하수 수질특성은 지형경사와 낮아짐에 따라 물-암석반응에 의한 자연적 수질진화특성을 보인다 (Jung *et al.*, 2010). 따라서 중산간 용천수가 분포한 위치를 기준으로 지형구배에 따른 지하수 유동 특성상 불필요한 하부의 공간변수가 포함되는 원형버퍼가 갖는 한계를 보여준다 할 수 있다.

이상치로 선정된 해안지역 용천수 중 C7은 대부분의 버퍼 반경에서 이상치로 선정된다. C7의 경우 수질 특성에 있어 다른 해안지역 용천수들과 달리 갈수기와 풍수기에 NO₃ 농도가 6 mg/L 이하로 지표오염원의 영향이 나타나지 않은 해안지역 용천수이다. C7은 지표암상의 변화가 관찰되는 곳에 위치한 지층경계형 용천수로 시기별 온도는 14.6°C로 일정하며 HCO₃ 및

SiO₂의 농도는 해안지역 용천수군의 평균값에 인접하다 (Jung *et al.*, 2010). 이러한 수질 특성은 C7이 해안지역 용천수 중에서도 물-암석반응에 의한 자연적인 수질진화특성이 지배적으로 나타나는 용천수임을 지시한다. 이와 같이 C7이 다른 해안지역 용천수와는 구분되는 수질특성을 나타내는 원인으로는 지형적 특성을 반영하는 것으로 C7의 용출지점을 기점으로 상부의 지형구배 변화가 급격히 변하는 특성이 나타난다는데 있다. 이는 C7의 지하수 유동은 고지대 및 중산간 지역과 같이 지표오염원이 거의 없는 곳에서 함양된 지하수가 용출지점까지의 유동과정에서 지표오염원의 영향을 받은 천부지하수와와의 혼합이 이루어지지 않았음을 지시한다. 그 밖의 이상치로 나타난 해안지역 용천수의 경우 인위적 오염의 영향이 상대적으로 감소하는 특성을 보인다. 이는 불투수층서의 발달과 기저지하수에 의한 지표오염원의 희석효과에 의해 다른 해안지역 용천수와 구분되어 이상치로 선정된 것으로 보인다. 이처럼 다중회귀분석에서 이상치는 단순히 이상 시료에 대한 파악뿐만 아니라 이상치로 선정된 개별 용천수와 다른 용천수들 간의 구분되는 지하수 유동체계의 차이를 제시한다고 사료된다.

4.4. 인위적 토지이용별 영향력 평가

Song *et al.* (1999)은 질소동위원소를 이용하여

Table 6. Comparison of regression models before and after removal of outliers in the optimal buffer radius based on anthropogenic land use variables

Sampling period	Outlier	n ¹⁾	R ²	Adj. R ²	F	p	Variables	Parameter Estimate	Standard Error	p	VIF	t	STB	
Dry season							Intercept	-0.088	0.477	0.8556		-0.18		
							O	7.845	0.959	<0001	2.792	8.18	1.140	
	Included	38	0.778	0.743	22.38	<0001	F	8.405	2.437	0.0016	2.764	3.45	0.478	
							R	6.405	1.528	0.0002	1.988	4.19	0.493	
							S1	-4.276	1.486	0.0071	3.044	-2.88	-0.419	
							S3	-4.357	1.154	0.0007	3.106	-3.78	-0.555	
							Intercept	-4.494	0.704	<0001		-6.39		
							O	8.474	0.829	<0001	2.647	10.23	1.201	
	Removed	37	0.833	0.813	40.00	<0001	F	8.853	1.895	<0001	2.227	4.67	0.503	
							R	6.474	1.308	<0001	1.940	4.95	0.497	
						S2	4.817	0.952	<0001	3.230	5.06	0.656		
Rainy season							Intercept	-2.407	0.458	<0001		-5.26		
							O	5.751	0.752	<0001	1.033	7.65	0.695	
	Included	39	0.720	0.696	30.10	<0001	F	5.105	2.560	0.054	1.787	1.99	0.238	
							R	6.344	1.863	0.0017	1.747	3.40	0.402	
							Intercept	-5.568	2.007	0.0093		-2.77		
							O	8.790	1.055	<0001	2.838	8.33	1.030	
	Removed	37	0.833	0.806	30.95	<0001	F	7.471	2.445	0.0046	2.497	3.06	0.354	
							R	9.390	2.027	<0001	3.157	4.63	0.604	
							S3	-2.896	1.020	0.0079	2.220	-2.84	-0.310	
							lnTopo	1.166	0.628	0.0728	4.481	1.86	0.288	

1) number of springs used for the regression.

NO₃ 오염원을 추정하는 연구를 수행하여 화학비료, 혼합오염원, 동물성 폐기물과 폐수에 의해 주로 영향을 받은 시료가 각각 52, 24, 19%임을 보여주었다. 이는 지하수내 NO₃에 대해 화학비료의 기여가 높다는 것을 지시하는 사례로 용천수의 NO₃ 농도분포에 영향을 주는 공간변수는 자연지역보다는 인위적 토지이용임을 알 수 있다. 이에 용천수의 NO₃ 농도에 양의 관계로 기여하는 공간변수 중 가장 중요한 인위적 토지이용간의 상대적 영향력 차이의 평가를 위해 토지이용 중 자연지역을 제외하고 최적의 버퍼 반경 400 m를 대상으로 과수원(O), 논/밭(F), 주거 및 인위시설(R)을 변수로 이용하여 이상치 제거 이전과 이후의 모형을 재차 회귀 분석 하였다 (Table 6).

재분석된 회귀모형의 유의확률은 0.05 이하, 선택된 변수의 유의확률 0.1 이하, 다중공선성 10 이하로 통계적으로 유의함을 알 수 있다. 재분석된 회귀모형은 객관적 기준에 의한 회귀분석방법에 비해 0.041~0.03의

범위에서 모든 모형의 수정결정계수가 낮게 나타난다. 인위적 토지이용 공간변수를 기준으로 최적의 버퍼 반경인 400 m에서 표준화 회귀계수 (STB)는 갈수기 O > S3 > R > F > S1, 풍수기 O > R > F로 두 시기 모두 과수원의 영향력이 가장 큰 것으로 나타났다. 인위적 토지이용에 해당하는 공간변수의 영향력은 해당하는 변수가 연구지역에서 차지하는 면적에 비례하는 것으로 높은 투수성의 화산암체로 구성된 연구지역의 경우 지표오염물의 유입이 직접적으로 인위적 토지이용과 연관됨을 지시한다. 400 m 반경에 대해 인위적 토지이용 공간변수를 기준으로 하며 이상치가 제거된 최종 모형에 진입한 공간변수가 설명할 수 있는 설명력은 갈수기 약 81%, 풍수기 약 80%로 용천수의 NO₃ 농도 예측에 있어 상당히 높은 수준임을 알 수 있으며 두 시기의 회귀식은 (5), (6)와 같다.

Dry season (Land use)

$$\ln\text{NO}_3 = -4.494 + 8.474(O) + 8.853(F) + 6.474(R)$$

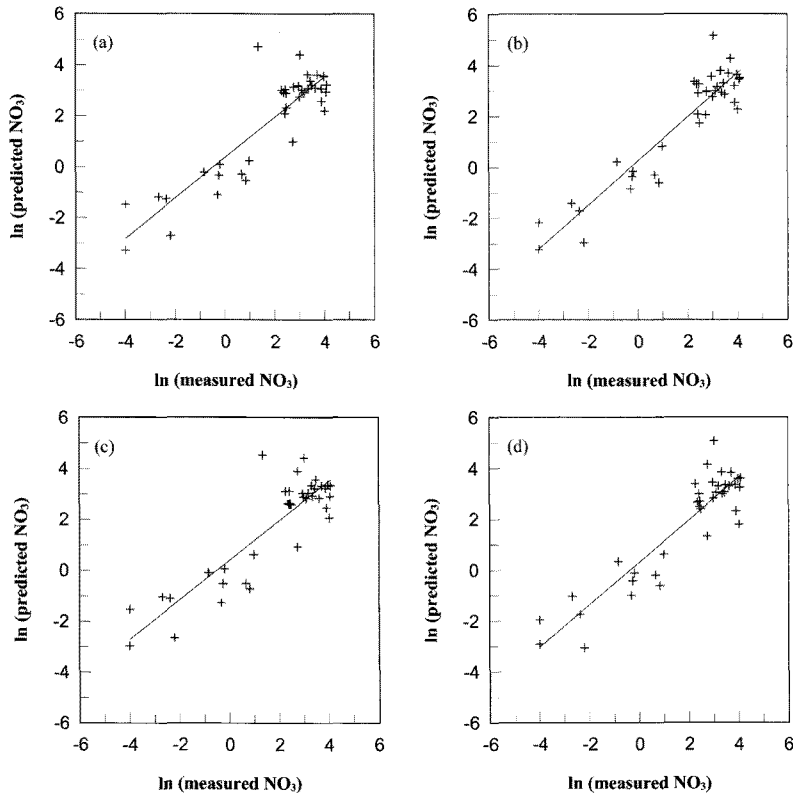


Fig. 7. Comparison of predicted and observed nitrate concentrations in dry season spring waters. (a) prediction using the best regression model before removing outliers, (b) prediction using the best regression model after removing outliers, (c) prediction using the anthropogenic land use variables before removing outliers, (d) prediction using the anthropogenic land use variables after removing outliers.

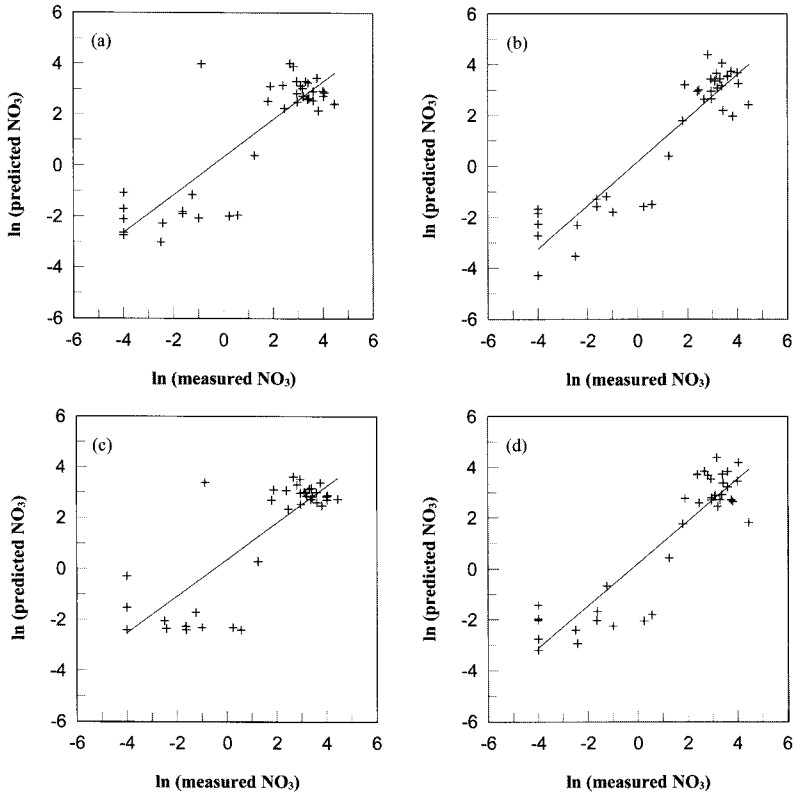


Fig. 8. Comparison of predicted and observed nitrate concentrations in rainy season spring waters. (a) prediction using the best regression model before removing outliers, (b) prediction using the best regression model after removing outliers, (c) prediction using the anthropogenic land use variables before removing outliers, (d) prediction using the anthropogenic land use variables after removing outliers.

$$+ 4.817(S2), R^2=0.813 \tag{5}$$

Rainy season (Land use)

$$\ln NO_3 = -5.568 + 8.79(O) + 7.471(F) + 9.39(R) - 2.896(S3) + 1.166(\ln Topo), R^2=0.806 \tag{6}$$

인위적 토지이용 공간변수를 기준으로 하는 회귀분석방법은 객관적인 기준에 의한 회귀분석방법에 비해 수정결정계수의 변화는 큰 차이를 나타내지 않아 연구 지역의 지표오염원의 분포에 중점을 둔 회귀분석은 용천수의 NO₃ 농도 예측에 있어 연구지역의 특성을 반영하는 적절한 회귀분석방법이라 판단된다. 또한 이상치의 제거 이전과 이후의 차이를 비교하였을 때 이상치의 제거를 통해 회귀식에 대한 신뢰성이 증가함을 알 수 있다 (Figs. 7 and 8). 따라서 의미를 갖는 변수를 통해 얻어진 최적 회귀모형이 연구지역의 전반적인 모형을 제시할 수 있으며 이상치의 제거를 통해 지표오염물질의 지하수 내 유입양상이 다른 유동구조를 갖는 용천수의 구별이 가능하다고 판단된다.

본 연구에서 제시한 추정치는 연구지역 내 모든 용천수의 지표오염과 관련한 수질특성에 차이를 보일 수는 있지만, 전반적인 지하수 내 지표오염원의 유입양상에 대한 유동특성을 파악하는데 있어 일차적인 예비자료로 활용할 수 있다는데 의미를 가질 수 있다.

5. 결 론

제주도 남측사면에 위치한 용천수의 질산성 질소 농도를 규제하는 요인을 분석하기 위해 토지이용, 지형, 토양 유형과 심도 등의 지리공간변수를 설명변수로 이용하는 다중회귀분석을 실시하였다.

1) 공간변수 추출을 위해 설정된 원형 버퍼의 반경이 커질수록 회귀모형의 결정계수가 증가하나, 원형 버퍼의 제한점과 계수 증가율을 고려할 때 풍수기와 같수기 모두 반경 400 m를 최적의 변수 추출 범위로 제시하였다.

2) 최적 회귀 모형에서 이상치로 확인되는 지점들은

대체로 전체의 10% 이내이며, 최적 버퍼 환경에서는 1-2개소로 이상치 제거로 인해 전체 시료에 대한 설명력 저하는 크지 않았다.

3) 이상치로 확인되는 용천수는 NO₃ 농도가 검출한계 이하인 경우, 기저 지하수 비율이 높은 해안 지역 용천수, 자연적 토지 이용과 인위적 토지 이용의 경계부에 위치한 용천수 등으로 나타나, 원형버퍼를 사용하는 다중선형회귀모형의 한계점을 보여주었다.

4) 수정결정계수, Cp, AIC를 기준으로 선택된 최적 회귀 모형의 결정계수는 이상치 제거 전이 0.74-0.79, 제거 후가 0.86-0.87의 범위로 높은 설명력을 가진다.

5) 수리지질학적 인자 및 토지 이용 특성 인자 중 용천수의 NO₃ 농도에 대한 영향력은 변수의 유의성 기준으로 자연지역의 면적 비율이 가장 크며, 토양 형태가 그 다음이다.

6) 용천수 NO₃ 농도에 대한 인위적 토지이용의 영향력은 최적 버퍼 환경에서 두 조사 시기 모두 과수원 > 주거지역 > 밭의 순으로 나타나 과수원이 용천수의 NO₃ 오염에 가장 높은 기여를 하는 것으로 나타났다.

참고문헌

Akaike, H. (1973) A new look at the statistical model identification, *IEEE Trans. Automat. Control*, v.19, p.716-723.

Basnyat, P., Teeter, L.D., Flynn, K.M. and Lockaby, B.G. (1999) Relationships between landscape characteristics and nonpoint source pollution inputs to coastal estuaries. *Environmental Management*, v.23, p.539-549.

Bauder, J.W., Sinclair, K.N., and Lund, R.E., 1993, Physiographic and land use characteristics associated with nitrate-nitrogen in montana groundwater, *Journal of Environmental Quality*, v.22, p.255-262.

Brady, N.C. and Weil, R.R. (2002) *The nature and properties of soils*. 13th ed. Prentice Hall, Upper Saddle River, NY.

Cheong, B.K., Chae, G.T., Koh, D.C., Ko, K.S. and Koo, M.H. (2008) A Study of Improvement for the Prediction of Groundwater Pollution in Rural Area: Application in Keumsan, Korea. *Journal of KoSSGE.*, v.13, p.40-53.

Eckhardt, D.A.V. and Stackelberg, P.E. (1995) Relation of ground-water quality to land use on Long Island New York. *Ground water*, v.33, p.1019-1033.

Gardner, K.K. and Vogel, R.M. (2005) Predicting Ground Water Nitrate Concentration from Land use. *Ground Water*, v.43, p.343-352.

Hamm, S.Y., Cheong, J.Y., Kim, M.J., Kim, I.S. and Hwang, H.S. (2004) Assessing Groundwater Vulnerability Using DRASTIC Method and Groundwater Quality in Changwon City. *Korea Soc. Econ. Environ. Geol.*, v.37, p.631-645.

Helsel, D.R. and Hirsch, R.M. (2002) *Statistical Methods in Water Resources*. USGS, Elsevier Science publishers B.V., The Netherlands, 522p.

Hong, C.S., Han, J.H. and Kim, H.I. (2005) Variable Selection for Logistic Regression Model Using Adjusted Coefficients of Determination. *The Korean Journal of Applied Statistics*, v.18, p.435-443.

Jeju Provincial Government (2009) *Statistical Yearbook of Jeju*.

Jeju Provincial Government (2010) *Report on present condition of farming and stockbreeding*. Jeju Provincial Agricultural Policy Department, 163p. (title translated).

Jeju-do (1997) *Comprehensive survey in middle mountain area of Jeju Island*. Korea Research Institute for Human Settlements, 344p. (title translated).

Jeju-do (1999) *Water of Jeju*, Spring. 392p. (title translated).

Jeju-do and K-water (2003) *Report on general survey of hydrogeology and groundwater resource(III)*. 425p. (title translated).

Jun, D.C., Song, Y.S. and Han, S.I. (2010) Proposal of Models to Estimate the Coefficient of Permeability of Soils on the Natural Terrain considering Geological Conditions. *Korean Society of Engineering Geology*, v.20, p.35-45.

Jung, Y.Y., Koh, D.C., Yu, Y.J. and Ko, K.S. (2010) Analysis of groundwater flow systems for springs in the southern slope of Jeju Island using hydrogeochemical parameters. *Journal of the Geological Society of Korea*, v.46, p.253-273.

Kang, T.S. and Um, J.G. (2007) Risk Assessment of the Road Cut Slopes in Gyeongnam based on Multiple Regression Analysis. *Korean Society of Engineering Geology*, v.17, p.393-404.

Kim, C.R. (2000) *The SAS statistical box, Data-Plus press*, Korea, 592p. (title translated).

Kim, E.Y., Koh, D.C., Ko, K.S. and Yeo, I.W. (2008) Prediction of Nitrate Contamination of Groundwater in the Northern Nonsan area Using Multiple Regression Analysis. *Journal of KoSSGE.*, v.13, p.57-73.

Kim, K.H., Lee, M.H., Lee, G.W., Kim, Y.P., Youn, Y.H. and Oh, J.M. (2002) Observations of aerosol-bound ionic compositions at Cheju Island, Korea. *Chemosphere*, v.48, p.317-327.

Kim, S.U. and Lee, K.S. (2008) Regional Low Flow Frequency Analysis Using Bayesian Multiple Regression. *Jour. Korea Resources Association*, v.41, p.325-340.

Kleinbaum, D.G., Kupper, L.L. and Muller, K.E. (1988) *Applied regression analysis and other multivariate methods* (2nd ed.). PWS Publishing Co., Boston, 366p.

Kloiber, S.M. (2006) Estimating nonpoint source pollution for the twin cities metropolotan area using landscape variables. *Water, Air, and Soil Pollution*, v.172, p.313-335.

Koh, D.C., Chae, G.T., Yoon, Y.Y., Kang, B.R. Koh, K.W. and Park, K.H. (2009) Baseline geochemical characteristics of groundwater in the mountainous area of Jeju Island, South Korea: Implications for degree of mineralization and nitrate contamination. *Journal of Hydrology*, v.376, p.81-93.

Koh, D.C., Chang, H.W., Lee, K.S., Ko, K.S., Kim, Y. and

- Park, W.B. (2005) Hydrogeochemistry and environmental isotopes of groundwater in Jeju volcanic island, Korea: implications for nitrate contamination. *Hydro. Process*, v.19, p.2225-2245.
- Koh, D.C., Ko, K.S., Kim, Y.J., Lee, S.G. and Chang, H.W. (2007) Effect of agricultural land use on the chemistry of groundwater from basaltic aquifers, Jeju Island, South Korea. *Hydrogeology Journal*, v.15, p.727-743.
- Koh, K.W. (1997) Characteristics of the groundwater and hydrogeologic implications of the Seoguipo Formation in Cheju Island. Doctoral thesis of Pusan National University, 325p. (title translated).
- Kolpin, D.W. (1997) Agricultural chemicals in groundwater of the Midwestern United States: relations to land use. *Journal of Environmental Quality*, v.26, p.1025-1037.
- Korea Institute of Geoscience and Mineral Resources (KIGAM) (2008) Integrated Analysis of Groundwater Occurrence in Jeju. 365p.
- Lee, B.J., Moon, S.H., Park, K.H., Koh, D.C. and Koh, K.W. (2002) Hydrogeochemical characteristics of the spring waters in Jeju. *Journal of the Geological Society of Korea*, v.38, p.421-439.
- Lee, S.R., Kim, Y.S., Kim, N.J. and Ahn, K.H. (2008) Analysis of Relationships Between Topography/Geology and Groundwater Yield Properties at Pohang using GIS. *Korea Soc. Econ. Environ. Geol.*, v.41, p.115-131.
- Liao, J.G. and McGee, D. (2003) Adjusted Coefficients of Determination for Logistic Regression. *The American Statistician*, v.57, p.161-165.
- Mallows, C.L. (1973) Some Comments on Cp. *Technometrics*, v.15, p.661-675.
- McLay, C.D.A., Dragten, R., Sparling, G. and Selvarajah, N. (2001) Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches. *Environmental Pollution*, v.115, p.191-204.
- Nolan, B.T. (2001) Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States. *Ground Water*, v.39, p.290-299.
- Nolan, B.T., Hitt, K.J. (2006) Vulnerability of shallow groundwater and drinking-water wells to nitrate in the United States. *Environmental Science and Technology*, 40 (24), p.7834-7840.
- Panagopoulos, G.P., Antonakos, A.K. and Lambrakis, N.J. (2005) Optimization of the DRASTIC method for groundwater vulnerability assessment via the use of simple statistical method and GIS. *Hydrogeology Journal*, v.14, p.894-911.
- Park, K.H., Cho, D.L., Kim, Y.B., Kim, J.C., Cho, B.W., Jang, Y.N., Lee, B.J., Lee, S.R., Son, B.K., Cheon, H.Y., Lee, H.Y. and Kim, Y.U. (2000) Geologic report of the SeogwipoHahyori Sheet (1:50,000). Jeju Provincial Government, 163p.
- Song, Y.C., Koh, Y.K. and U, J.G. (1999) Estimation of Nitrate Sources in Cheju Island Groundwater using $\delta^{15}\text{N}$. *Jour of the Korean Society of Groundwater Environment*, v.6, p.107-110.
- Thornton, G.J.P. and Dise, N.B. (1998) The influence of catchment characteristics, agricultural activities and atmospheric deposition on the chemistry of small streams in the English Lake District. *Science of the Total Environment*, v.216, p.63-75.
- Won, J.H., Lee, J.Y. and Kim, J.W. (2006) Groundwater occurrence on Jeju Island, Korea. *Hydrogeology Journal*, v.14, p.532-547.
- Won, J.S., Woo, N.C. and Kim, Y.J. (2004) Analysis Influential Factors on Nitrate Distribution in Ground Water in an Urbanizing Area using GIS. *Korea Soc. Econ. Environ. Geol.*, v.37, p.647-655.
- Youn, J.S. and Park, S.W. (1998) Hydrochemical Characteristics of Spring Water in Cheju Island. *Journal of the Korean Society of Groundwater Environment*, v.5, p.66-79.