Characterizing shallow groundwater contamination depending on different land use types

Shallow groundwater is vulnerable to major pollutants, such as nitrate-nitrogen (NO$_3$-N). The pollution characteristics of NO$_3$-N need to be understood to protect the quality of shallow groundwater. As already known, shallow groundwater contamination by NO$_3$-N is closely linked to agricultural fertilizers, domestic animals, and sewage. Correspondingly, nonpoint pollution sources such as fertilizers and sewage should be characterized by land use type. In this study, shallow groundwater contamination was examined according to three land use types (forest-dominant, agricultural-dominant, and mixed areas). Principal component analysis was adopted to extract two chemical components (Cl and SO$_4$) related to NO$_3$-N among a total of 11 components. Cluster analysis was performed through a self-organizing map by applying three water quality clusters with a ratio of NO$_3$-N/Cl. As a result, the forest-dominant area was characterized by a natural origin of NO$_3$-N with a [NO$_3$-N]/[Cl] ratio of 0.45; the agricultural-dominant area was indicated by NO$_3$-N from nitrogen fertilizer with a [NO$_3$-N]/[Cl] ratio of 0.56; and the mixed area was designated by NO$_3$-N from sewage with a [NO$_3$-N]/[Cl] ratio of 0.34.

Introduction

Shallow groundwater is closely related to the health of residents in areas where public water is not supplied (WHO and UNICEF, 2000). In developing countries, shallow groundwater is used as an important source of drinking water (Howard et al., 1999). In South Korea, the ratio of the water supply to the total population increased from 97.7% in 2010 to 99.3% in 2019 (Ministry of Environment, 2019). The rate of municipal water supply (99.3%) in South Korea is similar to that in other OECD countries, such as Australia (99.7%), Germany (99.3%), Japan (97.5%), and Denmark (97.2%) (IWA, 2016a), while the rate of groundwater use in South Korea is very low compared to that in other OECD countries. In 2016, groundwater among the sources for water supply was 15.64% in Australia, 62.26% in Germany, 22.92% in Japan, 35.94% in the United States, and 6.02% in Canada; however, the value was only 2.18% in Korea (IWA, 2016b). In addition, the rate of public water supply is lower in rural areas than in urban areas. In the circumstance of high public water supply percentages to private water use, a low ratio of groundwater to surface source water causes the groundwater management vulnerability in South Korea. In fact, contrary to the 99.3% of total municipal water supply in 2019, in rural county areas, only 78.6% of the population has a benefit of a public water supply (Ministry of Environment, 2019) because in those areas, the public water supply system has not been well established, e.g., the groundwater facilities of small-scale village waterworks and private wells (NIER, 2016). Therefore, the water quality of shallow groundwater should be managed very carefully since the shallow groundwater used for drinking and living water is closely related to the health safety of the local people.

Groundwater quality is influenced by many factors, such as microbiology, water-rock interactions, geochemical processes, land use patterns, fertilizer, and discharge sewage (Stumm and Morgan, 1996; Gelinas et al., 1996; Andrade et al., 2008; Mondal et al., 2008; Jiang et al., 2008; Compton and Boone, 2000; Kouras et al., 2007; Mondal and Singh, 2011). Most shallow groundwater is easily contaminated by nitrate–nitrogen (NO$_3$-N) (Kaown et al., 2007; Wick et al., 2012). In particular, agricultural activity has a great effect on the pollution of shallow groundwater (Liu et al., 2017; Lamsal et al., 2009). According to Scanlon et al. (2007), the groundwater recharge in crop areas is twice as high as that in forest areas, and groundwater recharge occurs in the soil and accelerates the NO$_3$-N contamination in shallow groundwater. Several studies (Chen et al., 2010; Bawa and Dwivedi, 2019) have reported a reverse relationship between the area of agricultural land and the pollution degree of shallow groundwater. Therefore, it is essential to ponder land use information to manage the water quality of shallow groundwater. Shallow groundwater pollution is characterized by the concentration and source of NO$_3$-N. Additionally, in agricultural activity areas, NO$_3$-N pollution in groundwater is closely related to chloride ions (Cl$^-$) (Rodvang et al., 2004; Menció et al., 2016). Accordingly, the [NO$_3$-N]/[Cl$^-$] ratio is widely used to confirm groundwater contamination by nitrogen fertilizers (Bourke et al., 2019). Denitrification is indicated by a decrease in the [NO$_3$-N]/[Cl$^-$] ratio since NO$_3$-N (or NO$_2$) is a highly reactive
ion, whereas Cl is a nonreactive ion (Weil et al., 1990; McCallum et al., 2008).

Principal component analysis (PCA) and self-organizing map (SOM) are the commonly used statistical methods for analyzing water quality data (Rao et al., 2010; Mohapatra et al., 2011; Selvakumar et al., 2017; Walley et al., 2000; Aguilera et al., 2001; Sánchez-Martos et al., 2002). PCA is suitable for effectively analyzing complex, large data (Mishra, 2010), reducing the dataset from high to low dimensions improves K-means clustering (Marín Celestino et al., 2018), and grouping datasets has the advantage of identifying characteristics (Pan et al., 2019; Loh et al., 2020). SOM, a type of cluster analysis, is an unsupervised learning method of deriving clusters through dimensional reduction from high- to low-dimensional data (Kohonen, 1982). Recently, SOM has been frequently used for clustering water quality data along with factor analysis and PCA, e.g., cluster analysis of groundwater quality data (Nguyen et al., 2015; Li et al., 2020), development of cluster models for the groundwater quality index (Gholami et al., 2020), cluster analysis for distinguishing freshwater and saltwater (Belkhiri et al., 2018), and regionalization of hydrological model parameters (Wallner et al., 2013).

This study aimed to reveal the shallow groundwater quality for the three land use types of forest, agriculture, and mixed areas, considering the pollution tendency of shallow groundwater in the different land use characteristics. For this study, chemical constituent analysis was conducted on a total of eleven components (pH, nitrate-N (NO$_3$-N), fluoride (F), chlorine (Cl), sulfate (SO$_4$), aluminum (Al), manganese (Mn), zinc (Zn), iron (Fe), copper (Cu), and turbidity). PCA and SOM were used to characterize the eleven components.

**Land Use and Geology of the Study Area**

The National Geographic Information Institute of Korea classifies the territory into four large categories (forest, agricultural, residential, and water areas), 14 intermediate categories, and 38 small categories on a map at a 1:25,000 scale. This study area is classified into three land use types according to the four large categories. The forest-dominant area, which is bounded by 36°10’ to 36°38’N and 127°30’ to 127°55’E, with the administrative districts of Boeun-gun of 583 km$^2$ and Okcheon-gun of 542 km$^2$ in Chungcheongbuk-do Province, is largely composed of mountains; of the total area, 69.45% is forest area, 21.90% is agricultural area, 3.48% is residential area, and 5.17% is water area (Fig. 1(A) and Table 1).

The agricultural-dominant area that is bounded by 36°04’ to 36°34’N and 126°41’ to 127°04’E, with the administrative districts of Buyeo-gun of 624 km$^2$ and Cheongyang-gun of 480 km$^2$ in Chungcheongnam-do Province (Fig. 1(B) and Table 1), accounts for 54.03% of the agricultural domain and 35.05% of the forest, with a larger agricultural domain than that of the forest-dominant area. The mixed area occupies 36°37’~36°57’N and 126°50’-127°25’E, with a total of 1,180 km$^2$ and consists of 39.52% agricultural area, 30.10% forest area, 29.38% residential area, and 1.00% water area (Fig. 1(C) and Table 1).
of Seobuk-gu in Cheonan city (198 km²), Dongnam-gu in Cheonan city (439 km²), and Ansan city (543 km²) in Chungcheongnam-do Province (Fig. 1(C) and Table 1). Unlike the forest and agricultural areas, this mixed area contains 11.52% of residential urban, 44.77% of forest, and 38.57% of agricultural domains.

These forest-dominant, agriculture-dominant, and mixed areas are similar to those by domestic statistical survey of Korea statistical information service (KOSIS, 2021); the forest-dominant area (Boeun-gun and Okcheon-gun) accounts for 191% of agricultural domain, 66.2% of forest domain, and 8.9% of residential urban domain; the agriculture-dominant area (Buyeo-gun and Cheongyang-gun) contains 26.9% of agricultural domain, 57.2% of forest domain, and 7.2% of residential urban domain; and the mixed area (Cheonan and Ansan cities) consists of 29.7% of agricultural domain, 43.9% of forest domain, and 23.2% of residential urban domain. This classification of the three types based on land use characteristics and regional statistics is reasonable within an acceptable accuracy even though Korea shows various land uses in its small territory with a difficulty of land use classification.

According to the geology of the study area based on the 1:250,000

| Table 1. Areas of land use map in each of the three types |
|-----------------|-------|-------|--------|-------|-------|
|                 | Forest | Agriculture | Residential area | Water | Total |
| **Forest-dominant area** |       |       |       |       |       |
| Area (km²)      | 781.57 | 246.44 | 39.22 | 58.21 | 1125  |
| Percent of area (%) | 69.45 | 21.90 | 3.48 | 5.17 | 100   |
| **Agriculture-dominant area** |       |       |       |       |       |
| Area (km²)      | 387.07 | 596.61 | 57.10 | 63.44 | 1104  |
| Percent of area (%) | 35.05 | 54.03 | 5.17 | 5.75 | 100   |
| **Mixed area** |       |       |       |       |       |
| Area (km²)      | 528.23 | 455.05 | 135.93 | 60.73 | 1180  |
| Percent of area (%) | 44.77 | 38.57 | 11.52 | 5.14 | 100   |

Figure 2. Geological maps of (A) forest-dominant area, (B) agriculture-dominant area, and (C) mixed area.
scale geological map (Kim et al., 2020), the forest-dominant area consists of the unknown-aged Okcheon formation along with Cenozoic volcanic rocks and Bulguksa intrusive rocks. In the unknown-aged Okcheon formation, many limestone cavities exist along with the limestone belt of Cambrian period. Groundwater abundantly takes place in such limestone areas (Bakalowicz, 2005; Papadopoulou-Vrynioti et al., 2013). The 80-84% of total groundwater quantity in the limestone area belongs to Ca-HCO₃ type due to the influence of the limestone (Moon, 2017). In the agricultural-dominant area, Gyeonggi gneiss complex of Pre-Cambrian is widely distributed and Nampo group of the Jurassic is distributed in the west along north-south direction. Also, Jurassic Daebo granite intrudes these rocks with a wide distribution. In the mixed area, Gyeonggi gneiss complex of Pre-Cambrian occupies most of the area with intrusion of Jurassic Daebo granite and is almost similar to the agricultural-dominant area.

Materials and Methods

Groundwater Quality Analysis

In total, 1,108, 1,074, and 2,088 groundwater samples were collected from the forest-dominant area, agriculture-dominant area, and mixed area, respectively (Fig. 1). Among the eleven components, NO₃-N, pH, and turbidity were measured in the field, and NO₂-N, F, Cl, SO₄, Al, Mn, Zn, Fe, and Cu were analyzed in the laboratory. Field NO₂-N analysis was conducted using a HACH Co. paper test kit. Any sample that exceeded 9.0 mg/L NO₂-N was analyzed for NO₂-N, fluorine, chloride, and sulfate ions using ion chromatography (861 compact IC, Metrohm Co.) and for aluminum, manganese, lead, zinc, iron, and arsenic via inductively coupled plasma-optical emission spectrometry (ICP-OES) (710-ES, Varian Co.) (Hanwool Life Science Co., Daejeon, South Korea). Turbidity was measured with a 2100p Hach Co. turbidity meter, and pH was measured on a Star A121 Orion CO. The water quality analysis followed Article 5 (Water Quality Management for Drinking Water, etc.) of the ‘Drinking Water Management Act’ (Ministry of Environment, 2017). Table 2 shows the water quality standards for the 13 chemical variables used in this study as well as their instrument detection limits.

| Item    | Korea’s drinking water standard | Detection limit | Instrument                  |
|---------|--------------------------------|-----------------|-----------------------------|
| NO₃-N   | 10.0 mg/L                      | 0.1 mg/L        | IC (861 compact IC)         |
| pH      | 4.5-9.5                         |                 | Star A121                   |
| F       | 1.50 mg/L                      | 0.15 mg/L       |                             |
| Cl      | 250 mg/L                       | 0.4 mg/L        | IC (861 compact IC)         |
| SO₄     | 250 mg/L                       | 2 mg/L          |                             |
| Al      | 0.20 mg/L                      | 0.02 mg/L       |                             |
| Mn      | 0.300 mg/L                     | 0.004 mg/L      |                             |
| Zn      | 3.000 mg/L                     | 0.002 mg/L      | ICP-OES (710-ES)            |
| Fe      | 0.30 mg/L                      | 0.05 mg/L       |                             |
| Cu      | 1.000 mg/L                     | 0.008 mg/L      |                             |
| Turbidity | 1,000 NTU[a]               | 0.02 NTU        | Turbidity meter (2100p)     |

[a]NTU: Nephelometric turbidity units

Principal Component Analysis and Self-organizing Map

PCA solves the raw data variable interdependence issue through dimension reduction. It is a combination of a covariance matrix and an eigenvector from normalized data. If the structure of the variables is not simple and the amount is large, the covariance matrix has a very complex structure. Therefore, the purpose of PCA is to summarize the characteristics of raw data and more easily express the covariance matrix with preserving the information of the raw data as much as possible. First, raw data is normalized so that the reprocessed dataset has a mean of 0 and a standard deviation of 1; then, a covariance matrix between variables is estimated using the normalized data. Second is to analyze the correlation between the variables. And then, eigenvalues and eigenvectors are calculated from the covariance and correlation matrix. Finally, the number of eigenvalues is determined using appropriate methods (e.g., Bartlett’s test (Bartlett, 1950), scree test (Cattell, 1966), Kaiser test (Kaiser, 1960), and parallel analysis (Horn, 1965)).

The SOM algorithm is a kind of density-based cluster analysis and presents a heuristic model to visualize and search linear and nonlinear relationships in high-dimensional datasets. In general, it is more difficult to search a trend with a large number of data sets and variables, especially in case of shallow groundwater that is easily affected by land use/surface cover and numerous factors (fertilizer, waste inflow into the well, well aging, etc.). SOM is a cluster analysis using unsupervised learning and subjectively determines the optimal number of clusters, with elbow and scree plot that relies on personal subjective judgment. By contrast, silhouette index, an evaluation of clustering validity, can be objectively judged when searching for the optimal number of clusters and then can be used to determine an appropriate number of clusters (Rousseeuw, 1987). The index indicates the distance between clusters with efficient separation (a range from -1 to 1). A value closer to -1 indicates that a cluster is close to other clusters, whereas a value closer to 1 means that a cluster is far from other clusters.

Classifying into clusters in SOM has an advantage of easier visualization than other cluster analyses (HCA and K-means) and can check the proportion of each cluster to the total data. In contrast, other cluster analyses with a large number of data or a high density data in a narrow area, possess a difficulty to distinguish each cluster when plotting total data on 2-dimensional or 3-dimensional graph.
SOM divides the entire data into clusters to overcome this difficulty and then re-groups the clusters. Therefore, SOM is more time- and cost-efficient than other analyses, with easier visualization (Vesanto, 1999; Qian et al., 2019).

SOM was first proposed by Kohonen (1982) and is processed: neurons received from the input space are compared with simple discriminant functions to select the largest function value unit; and next, the sample vector constituting the input layer is determined through competitive learning with the nearest neighbors to the selected value. In SOM analysis, an input layer is generated from the raw data, and an output layer of one or two dimensions is derived from the learning process. A sample vector that constitutes the input layer is determined through competitive learning based on Euclidean distance between any selected vector and the nearest neighbor vector. Input vectors with similar Euclidean distances are placed in the same cell or adjacent cells and belong to the same cluster (Kohonen, 1982). In this study, PCA was executed by using the psych package (Revelle, 2020), and SOM was performed by the SOMbrero package (Vialaneix et al., 2020) in open source R and RStudio (R Core Team, 2020).

**Results**

**Groundwater Quality According to Land Use Types**

The pollution pattern of NO$_3$-N in shallow groundwater was examined in the study area. The second chemical analysis was performed for the shallow groundwater that had a concentration higher than 9.0 mg/L at the first chemical analysis. The median concentrations of NO$_3$-N, pH, and Cl in the forest-dominant area were 5.54, 6.80, and 11.77 mg/L, respectively. In the agricultural-dominant area, the median concentrations of NO$_3$-N, pH, and Cl were 10.51, 6.50, and 16.99 mg/L, respectively. In the mixed area, the median NO$_3$-N, pH, and Cl concentrations were determined to be 9.50, 6.89, and 25.60 mg/L, respectively (Fig. 3). The concentrations of all items are presented in Table 3. Under natural circumstances, the average NO$_3$-N concentration in shallow groundwater is typically less than 2 mg/L (Mueller et al., 1995) and generally originates from industry, agriculture, waste disposal, and sewage (Ogrinc et al., 2019).

The ratio of NO$_3$-N and Cl can be used as an index to identify nitrate mixing or biological processes (nitrification, denitrification, assimilation, and remineralization), with Cl ions and conservative tracers, playing a role as an indicator of contamination in shallow groundwater (Yue et al., 2017). In this study, the ratios of [NO$_3$-N] and [Cl] in the forest- and agricultural-dominant areas were relatively high, with values of 0.48 and 0.51, respectively, and the ratio in the mixed area was low, with a value of 0.37, indicating contamination in the forest-dominant and agricultural-dominant areas by chemical fertilizers and in the mixed area by sewage (Fig. 4). It has been reported that chemical fertilizers contain a high nitrate concentration with a high [NO$_3$-N]/[Cl] value, whereas sewage contains a high content of [Cl] with a low [NO$_3$-N]/[Cl] value (Liu et al., 2006; Widory et al., 2005). Chen (2009) proposed a high value (0.56) of [NO$_3$-N]/[Cl] ratio in groundwater due to input of chemical fertilizers in Beijiang watershed in China. The source of pollution can be identified using [NO$_3$-N]/[Cl] ratio with isotope data. However, the pollution source is distinguishable based on [NO$_3$-N]/[Cl] ratios, provided there is an understanding of each site, local hydrogeology, and potential sources (Bourke et al., 2019). Therefore, the analysis of potential sources (fertilizers, sewage, and nature) in this study was interpreted with [NO$_3$-N]/[Cl] ratio as well as using mean values of NO$_3$-N concentration.

**Classification of Chemical Components by PCA and SOM**

In this study, at first stage, PCA was conducted for screening the components related to NO$_3$-N among eleven water quality components through dimensionality reduction, using shallow groundwater data from a total of 4,270 wells. At second stage, SOM was executed for clustering and visualization. The result of SOM was validated by Figure 3. Box-whisker plots of (A) NO$_3$-N concentration, (B) pH, and (C) Cl concentration for the land use types.
The items Cl and SO$_4$ were classified to belong to factor 2 based on reflecting their contribution to NO$_3$ contamination in the shallow groundwater of the study area. Hence, with three clusters, the basic input data of the study area was determined to be three, maintaining the unity of the study area. Among the numerous techniques used to determine an appropriate number of clusters, in this study, hierarchical cluster analysis was used, with Euclidean distance calculating the distance between clusters and Ward’s method determining the connection between clusters (Ward, 1963). Finally, SOM was performed to characterize NO$_3$-N in all three land use types, with Euclidean distance calculating the distance between clusters and Ward’s method determining the connection between clusters (Ward, 1963). Finally, SOM was performed to characterize NO$_3$-N contamination in the shallow groundwater of the three areas. In the SOM analysis, the hexagonal plot of 5 × 5 superclusters was divided into three clusters. The conditions of the SOM model were equally applied to all three land use areas, with 2000 trials, random choice for the initial neuron location, and Euclidean distance for the distance between clusters. Although the maximum number of trials was 2000, the model was stabilized at ~200 trials with a reduced number of water samples through PCA (Fig. 6(A)). The appropriate number of clusters was derived from the dendrogram (Fig. 6(B)). Forest-dominant and agriculture-dominant areas could be classified into three clusters, but mixed areas could be classified into three or five detailed clusters. Finally, the appropriate number of clusters in all three land use areas was determined to be three, maintaining the unity of the study area. Hence, with three clusters, the basic input data of SOM were determined to be NO$_3$-N, Cl, and SO$_4$.

### Regional Characterization According to Land Use Type

In Fig. 7, the three hexagonal grid plots are composed of red cluster 1, green cluster 2, and blue cluster 3. In the forest-dominant area, clus-

| Number of samples | Land-use type | NO$_3$-N (mg/L) | pH | F (mg/L) | Cl (mg/L) | SO$_4$ (mg/L) | Al (mg/L) | Mn (mg/L) | Zn (mg/L) | Fe (mg/L) | Cu (mg/L) | Turb. (NTU) |
|-------------------|--------------|-----------------|----|---------|----------|-------------|--------|--------|--------|--------|---------|-----------|
| Forest-dominant area | 326 | Forest | 6.73 | 6.96 | 0.23 | 13.89 | 16.57 | 0.00 | 0.19 | 0.04 | 0.00 | 0.40 |
| | 93 | Agriculture | 5.99 | 6.95 | 0.19 | 12.31 | 17.06 | 0.00 | 0.13 | 0.02 | 0.05 | 0.23 |
| | 685 | Residential | 7.65 | 6.85 | 0.21 | 16.74 | 17.88 | 0.01 | 0.12 | 0.05 | 0.00 | 0.42 |
| | 4 | Water | 3.55 | 7.00 | 0.38 | 18.95 | 14.24 | 0.00 | 0.07 | 0.00 | 0.00 | 0.11 |
| Agricultural-dominant area | 363 | Forest | 10.87 | 6.60 | 0.16 | 20.49 | 11.89 | 0.01 | 0.14 | 0.02 | 0.01 | 0.51 |
| | 93 | Agriculture | 9.05 | 6.59 | 0.17 | 21.07 | 13.12 | 0.00 | 0.05 | 0.01 | 0.00 | 0.43 |
| | 614 | Residential | 13.14 | 6.54 | 0.17 | 24.63 | 12.52 | 0.02 | 0.19 | 0.03 | 0.00 | 0.57 |
| | 4 | Water | 12.26 | 6.65 | 0.20 | 35.29 | 3.10 | 0.00 | 0.04 | 0.00 | 0.00 | 0.54 |
| Mixed area | 721 | Forest | 8.82 | 7.04 | 0.15 | 27.88 | 14.83 | 0.01 | 0.13 | 0.00 | 0.01 | 0.35 |
| | 146 | Agriculture | 5.86 | 7.15 | 0.27 | 19.25 | 14.76 | 0.01 | 0.14 | 0.01 | 0.01 | 0.50 |
| | 1213 | Residential | 8.90 | 6.96 | 0.16 | 26.58 | 16.00 | 0.00 | 0.02 | 0.13 | 0.00 | 0.35 |
| | 8 | Water | 10.71 | 6.95 | 0.21 | 40.62 | 15.63 | 0.00 | 0.09 | 0.00 | 0.01 | 0.36 |

**Table 3. Mean concentration of chemical constituents by the land use types**

**Figure 4. Relationship between land use type and [NO$_3$-N]/[Cl] ratio.**

using silhouette index. When classifying chemical items by PCA, it is very important to determine the appropriate number of factors. In this study, PCA was used to classify items related to NO$_3$-N, and then the factors containing NO$_3$-N were applied to SOM. In this study, silhouette index was estimated as 0.47, 0.48, and 0.44 for the agriculture-dominant, forest-dominant, and mixed areas, respectively (Fig. 5(A)). The calculated values of silhouette index seems to be reasonable for clustering since the silhouette index of 0.5 is the criteria of reasonable clustering (Everitt et al., 2001).

In the biplot visualization, NO$_3$-N, Cl, and SO$_4$ belong to factor 1 in the forest-dominant area, and factor 2 in both the agriculture-dominant area and the mixed area (Table 4). In the mixed area, Cl and SO$_4$ were classified to belong to factor 2 based on reflecting their contribution to NO$_3$-N even though they showed approximately equivalent level of loading to factor 2 and factor 3. Thus, Cl and SO$_4$ displayed a positive relationship with NO$_3$-N, while pH showed a negative relationship with NO$_3$-N in the three land use types (Fig. 5(B)).
Cluster 3 was predominant (Fig. 7(A)). In particular, as shown in '3.3 Water quality results', cluster 3 in the forest-dominant area represented an average NO$_3$-N concentration of 4.10 mg/L in the natural state, with a [NO$_3$-N]/[Cl$^-$] ratio of 0.45. In contrast, cluster 1 was predominant in the agriculture-dominant area (Fig. 7(B)). The average NO$_3$-N concentration in cluster 1 was 6.51 mg/L, which was higher than that of the forest-dominant area, and a higher [NO$_3$-N]/[Cl$^-$] value of 0.56. The pollution of these two areas had slightly higher NO$_3$-N concentrations than the natural background concentration (below 4 mg/L; APEC Water, 2021; USEPA, 1982) and was characterized by pollution mainly caused by nitrogen fertilizers. In the mixed area, clusters 2 and 3 were almost equally important, with cluster 2 illustrating an average NO$_3$-N concentration of 25.92 mg/L and a [NO$_3$-N]/[Cl$^-$] value of 0.47 and cluster 3 demonstrating an average NO$_3$-N concentration of 5.91 mg/L and a [NO$_3$-N]/[Cl$^-$] value of 0.34 (Fig. 7(C)). Cluster 2 was judged to be polluted by nitrogen fertilizers, similar to the forest-dominant and agriculture-dominant areas.
nant areas, while cluster 3 was judged to be polluted by sewage. The pollution sources of each cluster can be inferred from the location of the wells, the areal distribution of NO$_3$-N and Cl concentrations, and the [NO$_3$-N]/[Cl] ratios, as presented in Table 5. The [NO$_3$-N]/[Cl] ratio was the highest in the agriculture-dominant area, followed by the forest-dominant area and then the mixed area (Table 5). The first pollution source comes from agricultural activities. Cluster 1 in the forest-dominant area and Cluster 2 in the agriculture-dominant area represent agricultural activities with high concentration of NO$_3$-N and Cl, and accordingly, high [NO$_3$-N]/[Cl] ratios. In particular, the wells in agricultural area mostly belong to these clusters. The second pollution source is sewage. The effect of sewage augments the [NO$_3$-N]/[Cl] ratios with high concentration of NO$_3$-N and low concentration of Cl. This effect is recognized as cluster 2 in the forest-dominant area and clusters 1 and 2 in the mixed area. The wells belonging to these clusters are distributed in residential area which supports the origin of sewage. Hence, it was confirmed that the pollution patterns of shallow groundwater were governed by the land use characteristics through PCA and SOM, which could clearly distinguish clusters according to land uses.

Table 4. PCA results using varimax rotation of three land use types

| Variables | Forest-dominant area | Agriculture-dominant area | Mixed area |
|-----------|----------------------|---------------------------|------------|
|           | Factor 1 | Factor 2 | Factor 3 | Communality | Factor 1 | Factor 2 | Factor 3 | Communality | Factor 1 | Factor 2 | Factor 3 | Communality |
| NO$_3$-N  | 0.75     | -0.16   | -0.26   | 1.3        | 0.04     | 0.84     | -0.08   | 1.0        | -0.07     | 0.80     | 0.16     | 1.1        |
| pH        | -0.42    | -0.02   | 0.62    | 1.7        | 0.00     | -0.59    | 0.03    | 1.0        | -0.03     | -0.77    | 0.01     | 1.0        |
| F         | 0.03     | 0.01    | 0.66    | 1.0        | 0.03     | -0.06    | 0.79    | 1.0        | -0.04     | -0.46    | 0.29     | 1.7        |
| Cl        | 0.83     | -0.02   | 0.01    | 1.0        | 0.09     | 0.84     | 0.04    | 1.0        | -0.07     | 0.51     | 0.56     | 2.0        |
| SO$_4$    | 0.76     | 0.05    | 0.32    | 1.4        | -0.05    | 0.42     | -0.05   | 1.1        | -0.08     | 0.24     | 0.59     | 1.4        |
| Al        | -0.06    | 0.61    | -0.11   | 1.1        | 0.24     | -0.01    | 0.80    | 1.2        | 0.93      | 0.00     | -0.01    | 1.0        |
| Mn        | 0.12     | 0.08    | -0.11   | 2.7        | -0.04    | 0.03     | 0.02    | 2.2        | 0.04      | -0.07    | 0.49     | 1.1        |
| Zn        | -0.01    | 0.06    | -0.27   | 1.1        | 0.61     | 0.08     | -0.31   | 1.5        | 0.02      | -0.01    | 0.41     | 1.0        |
| Fe        | 0.07     | 0.83    | -0.02   | 1.0        | 0.88     | 0.02     | 0.28    | 1.2        | 0.93      | -0.01    | 0.05     | 1.0        |
| Cu        | -0.02    | 0.02    | -0.29   | 1.0        | 0.14     | 0.02     | -0.17   | 1.9        | 0.07      | -0.07    | 0.45     | 1.1        |
| Turbidity | -0.05    | 0.84    | -0.04   | 1.0        | 0.80     | 0.02     | 0.26    | 1.2        | 0.87      | -0.01    | 0.05     | 1.0        |
| Eigen values | 2.03     | 1.81   | 1.18    | 1.96      | 1.87     | 1.55    | 2.52    | 1.76       | 1.39      |         |         |            |
| Variance explained | 0.40     | 0.36   | 0.24    | 0.36      | 0.35     | 0.29    | 0.44    | 0.31       | 0.25      |         |         |            |
| Cumulative of variance | 0.40     | 0.76   | 1.00    | 0.36      | 0.71     | 1.00    | 0.44    | 0.75       | 1.00      |         |         |            |

Figure 6. (A) Iteration plot of the SOM model and (B) dendrogram of clustering in the forest-dominant area.

Discussion

Shallow groundwater quality is seriously affected by land use change in relation to an increase in population, an acceleration of urbanization and climate change. In this study, forest-dominant, agriculture-domi-
nant, and mixed areas could be classified according to the proportion of land use in Chungcheongnam-do Province, South Korea. There was accelerated urbanization in the mixed area compared with that in the forest- and agriculture-dominant areas. The average cultivated area in the world was ~12% of the total global land area from 2014 to 2016 (Kim et al., 2019). In contrast, in South Korea, agricultural land decreased from 19.2% to 15.9% from 1998 to 2018 due to urbanization (Kim et al., 2019). In the end, it is inferred that the total agricultural land in South Korea has decreased mainly in the mixed area due to urbanization while the forest-dominant and agriculture-dominant areas are being maintained as they were. As a consequence, the mixed area is significantly affected by urban or residential sewage, which is different from the major pollution sources of the other two areas.

The pollution by nitrogen fertilizers has been noted as representative pollution in shallow groundwater not only in Korea but also worldwide (Kaown et al., 2007; Wick et al., 2012). In this study, using the SOM model, the water quality clusters of the forest-dominant and agriculture-dominant areas were distinctly distinguished from the water quality clusters of the mixed area. Panno et al. (2005) reported that wastewater discharged from sewage treatment facilities contains relatively high concentrations of Na, K, NO₃, NH₄, Cl, and PO₄. In the [NO₃-N]/[Cl] ratio, Cl is nonreactive, but NO₃-N is reactive with denitrification (Kimble et al., 1972; Liu et al. 2006). In general, high NO₃-N concentrations and high [NO₃-N]/[Cl] ratios indicate chemical fertilizer, while high Cl concentrations and relatively low [NO₃-N]/[Cl] ratios indicate sewage origin (Widory et al., 2005). In addition, areas heavily affected by livestock burial sites and/or livestock manure distinctly show NO₃-N concentration close to 30 mg/L (Kim et al., 2017).

In this study, cluster 2 of the forest-dominant area, cluster 1 of the agriculture-dominant area, and cluster 2 of the mixed area were considered to have very high NO₃-N concentrations (exceeding 10.0 mg/L) and high [NO₃-N]/[Cl] ratios. In contrast, clusters 1 and 3 of the mixed area showed very low [NO₃-N]/[Cl] ratios compared to the other clusters. Hence, based on land use change and the [NO₃-N]/[Cl] ratio, clusters 1 of the mixed area specified the effect of sewage. However, for a more accurate interpretation, it can be necessary to acquire the value of δ¹⁵NNO₃. Instead, in this study, according to the three clusters by the SOM model and the [NO₃-N]/[Cl] ratio, the forest-dominant and agriculture-dominant areas were characterized by a similar pollutant (nitrogen fertilizer), and the mixed area was characterized by sewage. This result is consistent with the urban district distribution and land use change in the study area.

**Figure 7. Hexagonal grid plots of three clusters depending on the land use types.**

**Table 5. NO₃-N, Cl, and SO₄ concentrations and the [NO₃-N]/[Cl] ratio of each cluster**

| Cluster       | Number of samples | NO₃-N (mg/L) | Cl (mg/L) | SO₄ (mg/L) | [NO₃-N]/[Cl] ratio |
|---------------|------------------|-------------|-----------|-----------|-------------------|
| Forest-dominant area |                   |             |           |           |                   |
| 1             | 285              | 11.20       | 21.21     | 22.70     | 0.53              |
| 2             | 117              | 16.70       | 41.70     | 38.06     | 0.40              |
| 3             | 709              | 4.10        | 9.05      | 11.97     | 0.45              |
| 1             | 646              | 6.51        | 11.52     | 8.22      | 0.57              |
| Agriculture-dominant area |             |             |           |           |                   |
| 2             | 232              | 16.98       | 33.98     | 28.91     | 0.50              |
| 3             | 196              | 24.27       | 47.62     | 6.20      | 0.51              |
| 1             | 318              | 13.03       | 40.79     | 42.89     | 0.32              |
| Mixed area    |                   |             |           |           |                   |
| 2             | 612              | 25.92       | 55.08     | 15.69     | 0.47              |
| 3             | 1158             | 5.91        | 17.54     | 11.22     | 0.34              |
As of 2011, the world average arable land per capita was 0.6 ha, with 77.4 ha in Canada, 55.3 ha in Australia, 32.3 ha in the USA, 14.0 ha in France, and 11.3 ha in Argentina (Shim et al., 2013). Countries with large territories, such as the USA, Canada, and Australia, possess a single land use of a large area that is characterized by a distinct groundwater quality. Unlike large countries with large-scale agricultural activities, in a small country such as South Korea where complex land use takes place in a small area with rapid land use change due to urbanization as well as agricultural types, shallow groundwater quality is not easy to characterize due to the complex land uses (Lee et al., 2021). Hence, the cluster analysis in this study can help characterize shallow groundwater in small areas with complex land use.

**Conclusion**

This study discriminated the contamination characteristics of shallow groundwater in forest-dominant, agriculture-dominant, and mixed areas in Chungcheongnam-do Province in South Korea. By chemical analysis, the median concentrations of $\text{NO}_3^-$, pH, and Cl in the shallow groundwater were 5.54, 6.80, and 11.77 mg/L in the forest-dominant area, 10.51 mg/L, 6.50, and 16.99 mg/L in the agriculture-dominant area, and 9.50, 6.89, and 25.60 mg/L in the mixed area, respectively, showing the lowest values in the forest-dominant area followed by those in the mixed area and the agriculture-dominant area. Additionally, the Cl ions showed an increasing order of forest-dominant areas, agriculture-dominant areas, and mixed areas, with a similar pH value in the three areas. This trend indicates a relatively higher influence of shallow groundwater pollution by agricultural activities and domestic wastes. In addition, $[\text{NO}_3^-]/[\text{Cl}]$ ratios for the wells in the forest- and agricultural-dominant areas, are higher as 0.48 and 0.51, respectively, than 0.37 for the wells in the mixed area and represent the effect of chemical fertilizers in the forest-dominant and agricultural-dominant areas but sewage effect in the mixed area.

The chemical constituents (Cl and SO$_4$) connected to NO$_3$ among the eleven components were extracted by using PCA. Then, three common clusters from water sample data were derived in the forest-dominant, agriculture-dominant, and mixed areas by using HCA. Through SOM analysis, three similar clusters were determined for all three land use areas. Cluster 1, 2, and 3 in the forest-dominant area reflect natural effect, sewage effect, and natural effect, respectively. Cluster 1, 2, and 3 in the agriculture-dominant area correspond to cluster 2 in the forest-dominant area and cluster 1 in the mixed area. The wells belonging to these clusters are distributed in residential area which supports the origin of sewage.

Hence, through PCA, HCA, and SOM analysis of the chemical data of shallow groundwater as well as the $[\text{NO}_3^-]/[\text{Cl}]$ ratio of the three clusters, it was confirmed that shallow groundwater in the study area was obviously characterized based on the different land use types. This result indicates that the land use effect is a critical factor of shallow groundwater contamination and shallow groundwater management associated with land use and climate change in the future.

**Acknowledgments**

This research was supported by the National Research Foundation of Korea (NRF) under the Ministry of Science and ICT (No. NRF-2020R1A2B5B02002198), Basic Research Project (21-3411) of the Korea Institute of Geoscience and Mineral Resources, and the research project ‘A study on Optimal Management Measures for Securing Safe Groundwater’ of the National Institute of Environmental Research of Korea (NIER-SP2015-387).

**Funding Information and Conflicts of Interest**

The authors declare no competing interests.

**References**

Aguilera, P.A., Frenich, A.G., Torres, J.A., Castro, H., Vidal, J.L.M., and Canton, M., 2001, Application of the Kohonen neural network in coastal water management: methodological development for the assessment and prediction of water quality. Water Research, v. 35, pp. 4053–4062. doi:10.1016/S0043-1354(01)00151-8

APEC Water, 2021, Nitrate (nitrate nitrogen) in drinking water. https://www.freedrinkingwater.com/water-education2/79-nitrate-nitrogen.htm [accessed 1st April 2021].

Andrade, E.M., Palácio, H.A.O., Souza, I.H., de Oliveira Leão, R.A., and Guerreiro, M.J., 2008, Land use effects in groundwater composition of an alluvial aquifer (Trussu River, Brazil) by multivariate techniques. Environmental Research, v. 106, pp. 170–177. doi:10.1016/j.envres.2007.10.008

Bakalowicz, M., 2005, Karst groundwater: A challenge for new resources. Hydrogeology Journal, v. 13, pp. 148–160. doi:10.1007/s10040-004-0412-9

Bartlett, M.S., 1950, Tests of significance in factor analysis, British Journal of Psychology, v. 3, pp. 77–85. doi:10.1111/j.2044-8317.1950.tb00285.x

Belkhirli, L., Mouni, L., Tiri, A., Sheikhy Narany, T., and Nouibet, R., 2018, Spatial analysis of groundwater quality using self-organizing maps. Groundwater for Sustainable Development, v. 7, pp. 121–132. doi:10.1016/j.gsd.2018.04.001

Bourke, S.A., Ivanshyn, M., Kohn, J., and Hendry, M.J., 2019, Sources and fate of nitrate in groundwater at agricultural operations overlying glacial sediments. Hydrology and Earth System Sciences, v. 23, pp. 1355–1373. doi:10.5194/hees-23-1355-2019

Cattell, R.B., 1966, The scree tests for the number of factors. Multivariate Behavioral Research, v. 1, pp. 245–276. doi:10.1207/s15327906mbr0102_10

Chen, F., Jia, G., and Chen, J., 2009, Nitrate sources and watershed denitrification inferred from nitrate dual isotope in the Beijiang River, south China. Biogeochemistry, v. 94, pp. 163–174. doi:10.1007/s10533-009-9316-x

Chen, S., Wu, W., Hu, K., and Li, W., 2010, The effects of land use change and irrigation water resource on nitrate contamination in shallow groundwater at county scale. Ecological Complexity, v. 7, pp. 131–138. doi:10.1016/j.ecocom.2010.03.003

Compton, J.E., and Boone, R.D., 2000, Long-term impacts of agriculture on soil carbon and nitrogen in New England forests. Ecology, v. 81, pp. 2314–2330. doi:10.1890/0012-6688(2000)81[2314:LTIOAOJ:CO;2]Everitt, B.S., Landau, S., and Leese, M., 2001, Cluster analysis, McAllister, I. De. Fourth Edition, London, Arnold, ISBN: 0-340-76119-9.

Gelinas, Y., Randall, H., Robidoux, L., and Schnit, J.-P., 1996, Well water survey in two districts of Conakry (Republic of Guinea) and compari-
son with the piped city water. Water Research, v. 9, pp. 2017–2026. doi:10.1016/j.watres.2016.04.013

Gholami, V., Khaleghi, M.R., and Taghvaee Salimi, E., 2020, Ground-water quality modeling using self-organizing map (SOM) and geophysi-
c3c0n information system (GIS) on the Caspian southern coasts. Journal of Mountain Science, v. 17, pp. 1724–1734. doi:10.1007/s11629-019-
5483-y

Horn, J.L., 1965, A rationale and test for the number of factors in factor analysis, Psychometrika, v. 30, pp. 179–185. doi:10.1007/BF0289447

Howard, G., Bartram, J.K., and Luyima, P.G., 1999, Small water supplies in urban areas of developing countries. In: Covuto JA, Crauf GR, Hearne N, editors. Providing safe drinking water in small system: technology, operations and economics. Lewis Publishers, Washington, DC, USA, pp. 83–93.

IWA (International Water Association), 2016a, IWA Statistics & Econom-
cies, Water delivered supply rate for drinking water. http://waterstatistics.iwa-network.org/ [accessed 18th March 2021].

IWA (International Water Association), 2016b, IWA Statistics & Econom-
cies, Abstraction source for water supply in %. http://waterstatistics.iwa-net-
cwork.org/ [accessed 18th March 2021].

Jiang, Y., Zhang, C., Yuan, D., Zhang, G., and He, R., 2008, Impact of land use change on groundwater quality in a typical karst watershed of southwest China: a case study of the Xiaojiang watershed, Yunnan Province. Hydrogeology Journal, v. 16, pp. 727–735. doi:10.1007/s10040-007-0259-9

Kaiser, H.F., 1960, The application of electronic computers to factor analysis, Educational and Psychological Measurement, v. 20, pp. 141–151. doi:10.1177/001316446002000116

Kim, J.-C, Koh, H., Lee, S.R., Lee, C.-B., Choi, S.-J., Park, K., Kim, D.-
H., Hwang, J.-H., Song, K.-Y., Lee, B.-J., Kim, Y.B., Cho, D.L., Choi, H., Jeon, H., Kim, B.C., Kee, W.-S., Kang, P. Jin, M.-S., Park, D.H., Choi, B.Y., and Choi, Y.S., 2020, Geological Map of Korea (on a scale of 1:250,000). Korea institute of Geoscience and mineral resources, Daejeon, South Korea. doi:10.2274/data.20201202.1141

Kim, M., Kim, C., and Park, D., 2019, The Position of Korean Agriculture in the World. Korea Rural Economic Institute, Naju, South Korea, pp. 1–269.

Kim, H., Park, S., Hwang, J., Kim, M., Jo, H., Jeon, S., Lee, K.-K., and Jeon, S.-W., 2017, Groundwater qualities of wells around carcass burial areas. Journal of the Geological Society of Korea, v. 53, pp. 433–446. doi:10.14477/jgs.2017.53.3.433

Kimble, J.M., Bartlett, R.J., McIntosh, J.L., and Varney, K.E., 1972, Fate of nitrate from manure and inorganic nitrogen in a clay soil cropped to continuous corn. Journal of Environmental Quality, v. 1, pp. 413–415. doi:10.2134/jeq1972.0047242500010004017x

Kohenen, T., 1982, Self-organized formation of topologically correct feature maps. Biological Cybernetics, v. 43, pp. 59–69. doi:10.1007/BF00337288

KOSIS (Korea Statistical Information Service), 2021, Domestic statistics, sta-
tus by land category. https://kosis.kr/statisticsList/statisticsListIndex. do?svwcd=M_TITITLE&enumdnl=M_01_01 [accessed 25th September 2021].

Kouras, A., Katsoyiannis, I., and Voutsas, D., 2007, Distribution of arsenic in groundwater in the area of Chalkidiki, Northern Greece. Journal of Hazardous Materials, v. 147(3), pp. 890–899. doi:10.1016/j.jhazmat.2007.01.124

Lamsal, S., Bliss, C.M., and Graetz, D.A., 2009, Geospatial mapping of soil nitrate-nitrogen distribution under a mixed land use system. Pedicsphere, v. 19, pp. 434–445. doi:10.1016/S1002-0160(09)60136-3

Lee, C.-M., Choi, H., Kim, Y., Kim, M., Kim, H., and Hamm, S.-Y., 2021, Characterizing land use effect on shallow groundwater contamination by using self-organizing map and buffer zone, Science of the Total Environment, v. 800, 149632. doi:10.1016/j.scitotenv.2021.149632

Loh, Y.S.A, Akurugua, B., Manu, E., and Aliou, A., 2020, Assessment of groundwater quality and the main controls on its hydrochemistry in some Voltaian and basement aquifers, northern Ghana, Groundwater for Sustainable Development, v. 10, 100296. doi:10.1016/j.gsf.2019.100296

Li, J., Shi, Z., Wang, G., and Liu, F., 2020, Evaluating Spatiotemporal Variations of Groundwater Quality in Northeast Beijing by Self-Organizing Map. Water, v. 12, pp. 1382. doi:10.3390/w12051382

Liu, C.Q., Li, S.L., Lang, Y.C., and Xiao, H.Y., 2006, Using δ15N and δ18O values to identify nitrate sources in karst ground water, Guiyang, Southwest China. Environmental Science & Technology, v. 40, pp. 6928–6933. doi:10.1021/es0610129

Liu, J., Jiang, L.H., Zhang, C.J., Li, P., and Zhao, T.K., 2017, Nitrate-nitro-
gen contamination in groundwater: Spatiotemporal variation and driving factors under cropland in Shandong Province, China. IOP Conference Series: Earth and Environmental Science 82. doi:10.1088/1755-1315/82/1/012059

Marín Celestino, A.E., Martínez Cruz, D.A., Otazo Sánchez, E.M., Gavi

Reyes, F., and Vásquez Soto, D., 2018, Groundwater Quality Assessment: An Improved Approach to K-Means Clustering. Principal Component Analysis and Spatial Analysis: A Case Study. Water, v. 10, pp. 437. doi:10.3390/w10040437

McCullum, J.E., Ryan, M.C., Mayer, B., and Rodvang, S.J., 2008, Mixing-
induced groundwater denitrification beneath a manured field in southern Alberta, Canada. Applied Geochemistry, v. 23, pp. 2146–2155. doi:10.1016/j.apgeoche.2008.03.018

Mencio, A., Mas-Pla, J., Otero, N., Regas, O., Boy-Roura, M., Puig, R., Bach, J., Domenech, C., Zamorano, M., Brusi, D., and Folch, A., 2016, Nitrate pollution of groundwater; all right ..., but nothing else!. Science of the Total Environment, v. 539, pp. 241–251. doi:10.1016/j.scitotenv.2015.08.151

Ministry of Environment, 2019, 2019 Statistics of waterworks; Ministry of Environment: Sejong, Korea. http://law.go.kr/engLsSc.do?menuId=

Ministry of Environment, 2017, Drinking Water Management Act; Ministry of Environment: Sejong, Korea. http://law.go.kr/engLsSc.do?menuId=

Ministry of Environment: Sejong, Korea. http://law.go.kr/engLsSc.do?menuId=

Niño, editors. Providing safe drinking water in small system: technology, Water delivered supply rate for drinking water. http://waterstatistics.iwa-network.org/ [accessed 18th March 2021].

Mondal, N.C., Saxena, V.K., and Singh, V.S., 2008, Occurrence of eleva-
ted nitrate in groundwaters of Krishna delta, India. African Journal of Environmental Science and Technology, v. 2, pp. 265–271.

Mondal, N.C., and Singh, V.P., 2011, Hydrochemical analysis of saliniza-
tion for a tannery belt in Southern India. Journal of Hydrology, v. 405, pp. 235–247. doi:10.1016/j.jhydrol.2011.05.058

Moon, S.H., 2017, Hydrogeochemistry of Groundwater Occurring in Complex Geological Environment of Yeongdong Area, Chungbuk, Korea. Economic and Environmental Geology, v. 50, pp. 445–466. doi:10.9719/EEG2017.50.6.445

Mueller, D.K., Hamilton, P.A., Helsel, D.R., Hitt, K.J., and Barbara, C.R., 1995, Nutrients in Ground Water and Surface Water of the United States-An Analysis of Data Through 1992. U.S. Geological Survey Water Resources Investigations Report 95-4031.

Nguyen, T.T., Kawamura, A., Tong, T.N., Nakagawa, N., Amauchi, H., and Gilbuena, J.R., 2015, Clustering spatio-seasonal hydrochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. Journal of Hydrology, v. 522, pp. 661–673. doi:10.1016/j.jhydrol.2015.01.023

NIER (National Institute of Environmental Research), 2016, A Study on Safe Groundwater Supply Project for Municipal Water Unsupplied Area in 2016. NIER, Incheon, South Korea, 236p.
Ogrinc, N., Tamšek, S., Zavladlav, S., Vrzel, J., and Jin, L., 2019, Evaluation of geochemical processes and nitrate pollution sources at the Ljubljansko polje aquifer (Slovenia): A stable isotope perspective. Science of the Total Environment, v. 646, pp. 1588–1600. doi:10.1016/j.scitotenv.2018.07.245

Pan, C., Ng, K.T.W., and Richter, A., 2019, An integrated multivariate statistical approach for the evaluation of spatial variations in groundwater quality near an unplanned landfill. Environmental Science and Pollution Research, v. 26, pp. 5724–5737. doi:10.1007/s11356-018-3967-x

Panno, S.V., Hackley, K.C., Hwang, H.H., Greenberg, S., Krapac, I.G., Landsberger, and O’Kelly, D.J., 2005, Characterization and Identification of Na–Cl Sources in Ground Water. Groundwater, v. 44, pp. 176–187. doi:10.1111/j.1745-6584.2005.00127.x

Papadopoulou-Vrynioti, K., Bathrellos, G.D., Skilodimou, H.D., Kaviris, G., and Makropoulos, K., 2013, Kaest collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area. Engineering Geology, v. 158, pp. 77–88. doi:10.1016/j.enggeo.2013.02.009

Qian, J., Nguyen, N.P., Oya, Y., Kikugawa, G., Okabe, T., Huang, Y., Ohuchi, F.S., 2019, Introducing self-organized maps (SOM) as a visualization tool for materials research and education. Results in Materials, v. 4, 100020. doi:10.1016/j.rinma.2019.100020

R Core Team, 2020, R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, https://www.R-project.org/

Rao, Y.R.S., Keshari, A.K., and Gosai, A.K., 2010, Evaluation of regional groundwater quality using PCA and geostatistics in the urban coastal aquifer, East Coast of India. International Journal of Environment and Waste Management, v. 5, pp. 163–180. doi:10.1504/IJEWM.2010.029700

Revelle, W., 2020, psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston, Illinois, USA, https://cran.r-project.org/web/packages/psych/index.html

Rodvang, S., Mikelson, D., and Ryan, M., 2004, Changes in ground water quality in an irrigated area of southern Alberta. Journal of Environmental Quality, v. 33, pp. 476–487. doi:10.2134/jeq2004.4760

Rousseeuw, P.J., 1987, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, v. 20, pp. 53–65. doi:10.1016/0377-0427(87)90125-7

Sánchez-Martos, F., Aguilar, P.A., Garrido-Frenich, A., Torres, J.A., and Pulido-Bosch, A., 2002, Assessment of groundwater quality by means of self-organizing maps: application in a semiarid area. Environmental Management, v. 30, pp. 716–726. doi:10.1007/s00267-002-2746-2

Scarlat, B.R., Jolly, I., Sophocleous, M., and Zhang, L., 2007, Global impacts of conversion from natural to agricultural ecosystems on water resources: Quantity versus quality. Water Resources Research, v. 43, W03437. doi:10.1029/2006WR005486

Selvakumar, S., Chandrasekar, N., and Kumar, G., 2017, Hydrogeochemical characteristics and groundwater contamination in the rapid urban development areas of Coimbatore, India. Water Resources and Industry, v. 17, pp. 26–33. doi:10.1016/j.wri.2017.02.002

Shim, M.H., Chai, C.H., and Jang, D.H., 2013, The Position of Korean Agriculture in the World, Korea Rural Economic Institute, Naju, South Korea, p.269.

Stumm, W., and Morgan, J.J., 1996, Aquatic chemistry: chemical equilibria and rates in natural waters, 3rd ed. Wiley: New jersey, USA; A277.

USEPA, 1982, Nitrate and Nitrite in Drinking-water. Publications of the World Health Organization can be obtained from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland, https://www.who.int/water_sanitation_health/dwq/chemicals/nitrate-nitrite2ndadd.pdf [accessed 1st April 2021].

Vesanto, J., 1999, SOM-based data visualization methods. Intelligent Data Analysis, v. 3, pp. 111–126. doi:10.1016/S1088-467X(99)00013-X

Vialaneix, N., Maigne, E., Mariette, J., Oltewa, M., Rossi, F., Bendhaiba, L., and Bolaert, J., 2020, SOMbrero: SOM Bound to Realize Euclidean and Relational Outputs. R package version 1.3-1. https://cran.r-project.org/web/packages/SOMbrero/index.html

Ward, J.H., 1963, Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, v. 58, pp. 236–244.

Walley, W.J., Martin, R.W., and O’Connor, MA, 2000, Self-organizing maps for classification of river quality from biological and environmental data. In: Denzer, R., Swayne, D.A., Purvis, M., and Schimak, G., Environmental Software Systems, ISESS 1999. IFIP — The International Federation for Information Processing: 39. Springer, Boston, MA, doi:10.1007/978-0-387-35503-0_4

Wallner, M., Haberlandt, U., and Dietrich, J., 2013, A one-step similarity approach for the regionalization of hydrological model parameters based on Self-Organizing Maps. Journal of Hydrology, v. 494, pp. 59–71. doi:10.1016/j.jhydrol.2013.04.022

Weil, R.R., Weismiller, R.A., and Turner, R.S., 1990, Nitrate contamination of groundwater under irrigated coastal plain soils. Journal of Environmental Quality, v. 19, pp. 441–448. doi:10.2134/jeq1990.00472425001900030015x

WHO, and UNICEF, 2000, Global Water Supply and Sanitation Assessment 2000 Report United Nations. New York, 87p

Wick, K., Heumesser, C., and Schmid, E., 2012, Groundwater nitrate contamination: Factors and indicators. Journal of Environmental Management, v. 111, pp. 178–186. doi:10.1016/j.jenvman.2012.06.030

Widory, D., Petetlet-Giraud, E., Negrel, P., and Ladouceur, B., 2005, Tracking the sources of nitrate in groundwater using coupled nitrogen and boron isotopes: A synthesis. Environmental Science & Technology, v. 39, pp. 539–548. doi:10.1021/es0493897

Yue, F.-J., Li, S.-L., Liu, C.-Q., Zhao, Z.-Q., and Ding, H., 2017, Tracing nitrate sources with dual isotopes and long term monitoring of nitrogen species in the Yellow River, China. Scientific Reports, v. 7, 8537. doi:10.1038/s41598-017-08756-7

Chung-Mo Lee is a Postdoctoral researcher of the department of groundwater research center in Korea Institute of Geoscience and Mineral Resources (KIGAM). He received a PhD in Department of Geological Sciences at Pusan National University, South Korea. He does research in Hydrogeology, Geostatistics and Geological Information System in earth sciences.

Yongcheol Kim is a principal researcher and head of the department of groundwater research center in Korea Institute of Geoscience and Mineral Resources (KIGAM). His research interests include managed aquifer recharge, coastal aquifer management, flow and transport in saturated/unsaturated porous/fractured media, DNA-PIL. problems, tracer tests and hydraulic tests for site characterization.
MoonSu Kim currently working as the Groundwater Team leader of Soil and Groundwater Division in National Institute of Environmental Research, Incheon, South Korea. He obtained his M. S. degree and Candidate of Phil degree in the Department of Geology, from Pusan National University.

Se-Yeong Hamm is professor of Department of Geological Sciences at Pusan National University, South Korea. He obtained his Ph.D. in Hydrogeology from Montpellier University. His research interests are groundwater flow characteristic in porous and fractured media, groundwater flow and mass transport modeling, and groundwater response to earthquake.

Hyun-Koo Kim currently working as the Director of Soil and Groundwater Division in National Institute of Environmental Research, Incheon, South Korea. He obtained his Phil degree in Environmental Engineer, from Hokkaido University.