

Use of Multivariate Statistical Approaches for Decoding Chemical Evolution of Groundwater near Underground Storage Caverns

Jeonghoon Lee*

Department of Science Education, Ewha Womans University, Seoul 120-750, Korea

다변량통계기법을 이용한 지하저장시설 주변의 지하수질 변동에 관한 연구

이 정 훈*

이화여자대학교 과학교육과, 120-750, 서울특별시 서대문구 이화여대길 52

Abstract: Multivariate statistical analyses have been extensively applied to hydrochemical measurements to analyze and interpret the data. This study examines anthropogenic factors obtained from applications of correspondence analysis (CA) and principal component analysis (PCA) to a hydrogeochemical data set. The goal was to synthesize the hydrogeochemical information using these multivariate statistical techniques by incorporating hydrogeochemical speciation results calculated by the program, commonly used, WATEQ4F included in the NETPATH. The selected case study was LPG underground storage caverns, which is located in the southeastern Korea. The highly alkaline groundwaters at this study area are an analogue for the repository system. High pH, speciation of Al and possible precipitation of calcite characterize these groundwaters. Available groundwater quality monitoring data were used to confirm these statistical models. The present study focused on understanding the hydrogeochemical attributes and establishing the changes of phase when two anthropogenic effects (i.e., disinfection activity and cement pore water) in the study area have been introduced. Comparisons made between two statistical results presented and the findings of previous investigations highlight the descriptive capabilities of PCA using calculated saturation index and CA as exploratory tools in hydrogeochemical research.

Keywords: multivariate statistical analysis, LPG storage caverns, principal component analysis

요약: 다변량통계기법은 수리지구화학 자료의 분석 및 해석에 많이 이용되어 왔다. 본 연구에서 대응분석과 주성분분석을 동시에 사용하여 인위적인 활동에 의한 지하수의 특징을 살펴보았다. 본 연구의 목적은 NETPATH 프로그램 속의 WATEQ4F를 이용하여 지하수 화학성분의 분화를 계산하고 이를 다변량통계기법을 이용하여 지구화학적 정보를 추출하는 것이다. 연구지역은 한반도의 남동쪽에 위치한 울산의 LPG 저장시설이다. 본 연구지역에서는 다른 저장시설에서 관찰되는 초염기성의 조성을 가지는 지하수가 관찰되었다. 이러한 인위적인 영향에 의한 높은 pH를 가지는 지하수로 인해 Al의 분화특징과 탄산염의 침전을 유발할 수 있다. 본 연구에서는 연구지역에 지하수에 영향을 주는 두 인위적인 요소(세정작용과 시멘트영향)에 의해서 수리지구화학적 특징과 상이 어떻게 변하는가에 초점을 두었다. 이전 연구결과와 두 통계분석을 통해 제시된 결과를 비교하여 지구화학적 정보를 이용한 주성분분석과 대응분석인 수리지구화학 연구에서 기초연구로 활용될 수 있음을 알 수 있다.

주요어: 다변량분석, LPG 저장시설, 주성분 분석

*Corresponding author: jeonghoon.d.lee@gmail.com

Tel: +82-2-3277-3794

Fax: +82-2-3277-2684

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Due to the complexity of the chemical evolution of groundwaters and the substantially large amount of data available, researchers often include as many variables as possible to ensure that no relevant variables will be omitted in exploratory studies. These analyses will indicate that the data matrix in the study area have a variety of hydrochemical characteristics, which cannot be revealed using standard graphical techniques, such as Piper diagrams. One may wish to reduce the dimensionality of the problem, i.e., reduction of the number of variables without losing much information. In this case, statistical methods can be a useful tool to confirm and discriminate the influence of each phenomenon on the final chemical composition of groundwater. Therefore, they have been widely used for the hydrochemical interpretations of commonly collected groundwater data. The analysis of multivariate data is used for classifying the variables and observation units and defining the relationship between them (David, 1973; Join et al., 1997).

Multiple regression, principal components and correspondence analysis, have been used for establishing relations of inter-dependence between variables (Kwon and Kim, 1997; Lee et al., 2008; Ko et al., 2010; Jeong, 2012). The aim of these techniques is to establish the relationship between variables and define the contribution of each of them, or each combination, to the structure of hydrochemical data. Principal components analysis (PCA) is a method complementary to a classical approach for hydrogeochemical research. In the case of hydrochemical studies, these techniques can be used for distinguishing different processes of pollution (Giménez and Morell, 1997), calculating the effect of the groundwater mixing on the obtained groundwater composition (Laaksoharju et al., 1999) and tracing groundwater circulation (Join et al., 1997).

Correspondence analysis (CA), on the other hand, can be viewed as simultaneous linear regression scheme with dual scaling, which allows the interpretation of both sample numbers and variables in the same vector space (Jobson, 1992). Such a unique feature is strictly

related to the scaling and weighing procedure used in CA. CA is a technique for displaying the rows and columns of a data matrix as points in low-dimensional vector spaces. Three advantages are involved (David, 1973): (1) groups of sample points belong to a specific family; (2) the closer the variables, the greater the correlation between them; and (3) the closeness of variable(s) and sample(s) might be an indication that the same underlying process are determining such proximity in the plotting positions. The use of CA in hydrogeochemistry is somewhat limited because the mathematical formulation for recomputation of the observed data is less suitable for the hydrogeochemical interpretation of the factors than the formula used by factor analysis (Valenchon, 1982).

All of the minerals in the study area react to some degree with groundwater. The complexity generated by this mineralogical interactions between minerals and water makes it impossible to construct a unique reaction model describing the groundwater chemistry. However, acquisition of relatively large sets of calculated saturation index (SI) has become routine in a variety of hydrogeochemical modeling techniques (Plummer et al., 1976). These quantities of data are particularly valuable for tracing water-rock interactions by using SI as variables. I attempt here to expand the usefulness of PCA by dealing with the specific case of groundwater evolution. This approach may be vulnerable to subjective bias and is inadequate in situations involving several or more samples requiring comparison.

Following the earlier studies on the groundwater flow and hydrochemical transport systems of the study areas (Ko et al., 2002; Lee et al., 2008; Ko et al., 2010), I explore the use of multivariate statistical methods in this investigation as a means of rapidly searching for chemical similarities between different groundwater samples collected from the study area. The hydrogeochemistry and hydrogeology of the study area have been studied quite extensively in earlier works (Kim et al., 2000; Ko et al., 2002; Lee et al., 2008) and in order to further increase the understanding of this it was necessary to obtain more information

about the groundwater. Here, I employed a large data set consisting of calculated SI. The purpose of this study is to discuss the use of principal component analysis using the SI and correspondence analysis, multivariate statistical procedures, to interpret the major factors contributing to chemical variability of the groundwater system in the study area.

Site Descriptions

Study Area

Topography of the study area is composed of a hilly mountain, Bongdae-San, in the range of 0-70 m with a 132 m peak elevation above mean sea level (a.m.s.l) (Fig. 1). The average annual precipitation is 1,250 mm (Lee et al., 2008). Park and Yoon (1968) have investigated the geology of some other hills around the Bongdae-San area. Geostock (1985, 1992) has reported on the general local geology and geologic discontinuities. Two sedimentary units of the Ulsan Formation occupy the major part of the map area. One of the units is composed of purple shale, grayish green shale, sandstone and conglomerate and the other, younger one is composed of hornfels and minor sandstone and conglomerate, representing the metamorphosed ring of a large granite stock located in the south. Locally, andesite, granite, granodiorite and quartz porphyry intrude the sedimentary and metasedimentary formations.

As shown in Fig. 1, three main tectonic faults are present in the study area. They have been considered to play an important role as natural factors for groundwater flow and hydrochemical transport. Fault 1 separates the propane and butane storage facilities, Fault 2 lies in the eastern part of the water curtain and Fault 3 is located in the western part of the propane caverns. Investigations during construction of the LPG facility showed that the fault crossing the propane and butane caverns had strike/dip of N30°E/70°SE, width 90-100 m, and formed a very permeable water-containing zone (Geostock, 1985). Groundwater flow and chemistry in this study area have been investigated by several researchers (Kim et al., 2000;

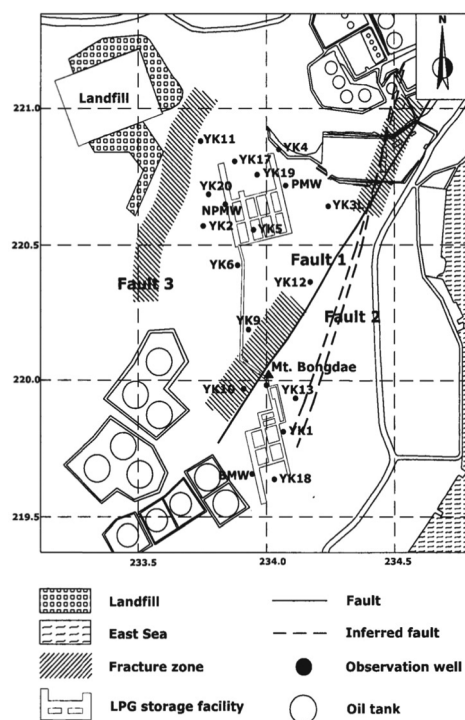


Fig. 1. Plan view of LPG storage caverns and locations of groundwater samples taken for this work. The coordinates are from the local Korean grid.

Ko et al., 2002; Lee et al., 2008; Ko et al., 2010; Lee et al., 2011). In terms of hydraulic properties, the rock system can be divided into two main types of formations. The medium permeability formation, usually encountered between surface and tentatively about 60 m below the ground level, has average hydraulic conductivity values in the range of 1.61×10^{-7} to $8.06 \times 10^{-6} \text{ m min}^{-1}$. The wells installed in this fault (YK9U/L) showed high hydraulic conductivities (6.77×10^{-6} - $8.06 \times 10^{-6} \text{ m min}^{-1}$). The high hydraulic conductivities of fault zones in the study area suggest that they could be main conduits for the natural groundwater flow.

LPG storage facilities

A construction of the LPG storage facilities was completed in 1987, having taken about two and half years. The LPG storage facilities are mainly composed of three sub-facilities. One sub-facility is a pipeline

system from ship unloading pier to the underground storage cavern. Another sub-facility is an underground cavern system and the third sub-facility is a gas extraction system. Basically, a principle of LPG storage in rock caverns is to suppress gases using groundwater in rock fractures that surround caverns. The underground cavern system has an access tunnel, water curtain, water curtain tunnel and storage caverns, which is conceptualized in Fig. 2. The access tunnel was used as an access route from the surface to the storage caverns at the time of construction, and is now used as the water pathway from surface pond to water curtain tunnels.

To assure that the groundwater pressure is higher than the LPG gas pressure at the cavern boundary, water curtain are made of multiple tributary tunnels extending from the access tunnel (Lindblom, 1989). Fan-like bundles of small holes (so-called 'water curtain holes') connected from the water curtain tunnels to the water curtains so that water in the access tunnel and water curtain tunnels can be supplied to the water curtains. The water supplied to the water curtains flow into storage caverns to satisfy the gas containment condition (Fig. 2). When the rate of inflow from the water curtains to caverns exceeds the rate of water inflow from ambient groundwater to the access tunnel and water curtain tunnels, water can

be artificially supplied to a surface pond or to the access tunnel. For the system studied here the water supply to the access and water curtain tunnels occurs naturally by groundwater discharge into the tunnels from ambient formations.

Methods

Field and Analytical Methods

Groundwater samples from the study area were taken in May 1999 for the analysis of hydrochemical parameters. A Watterra[®] tube (1.5 inches, 3.8 cm) and a bailer (3.0 inches, 7.6 cm) lowered into the well were connected directly to a flow cell, where pH, dissolved oxygen (DO), redox potential (Eh), and chemical conductivity (EC) were measured by electrodes, according to the methods described by Eaton et al. (1995) until these parameters had stabilized. A dissolved oxygen (DO) measurement was then made by continually filling a bottle holding a DO probe and samples for DO were fixed in the field by Winkler method and analyzed by titration for comparison. Samples were then collected through a 0.45-m filter and acidified with concentrated HNO₃ for cations and appropriately preserved at 4°C for transport back to analytical laboratories.

Chemical analyses were performed using standard

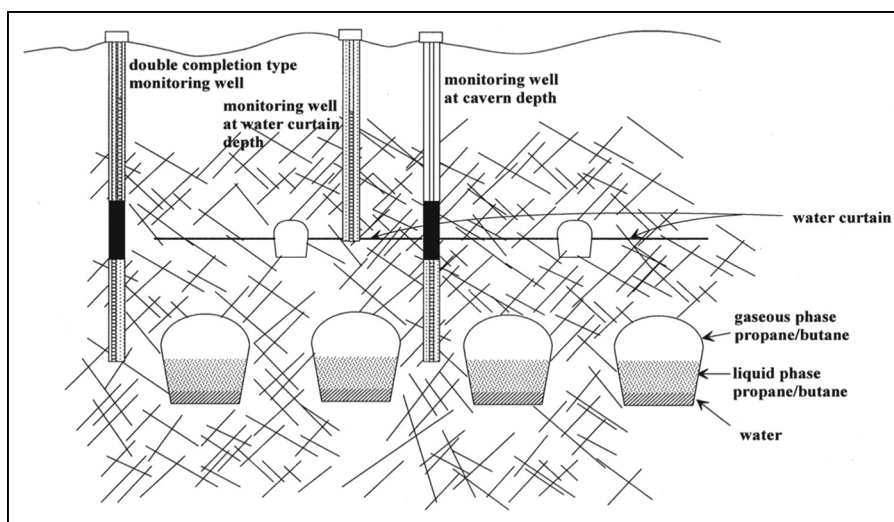


Fig. 2. Vertical schematic profile of LPG storage caverns (Modified from Kim et al., 2000).

methods. Anions were analyzed by ion chromatography (Dionex DX-500 equipped with AS40 autosampler), cations and trace elements by inductively coupled plasma (ICP-AES, Jobin Yvon JY138), alkalinity by gran titration (Stumm and Morgan, 1996). Standards were run every 10 samples for cations and anions. The relative standard deviations for these chemical analyses are less than 2%.

Statistical Methods

Two types of multivariate analysis, called PCA and CA were used on the data set. For this study, a statistical package developed by the SAS 8 package was used to perform the analyses. Values for elements below the method detection limits (<DL) were substituted with DL/2 prior to statistical analyses. When the number of '<DL' is small, replacing them with a constant (i.e., DL/2) is generally satisfactory, as suggested by the previous investigation (Stetzenbach et al., 1999).

CA is described in many textbooks (Jobson, 1992), so the following discussion of correspondence analysis is summarized. As discussed in the previous section, CA is a multivariate statistical technique for plotting data from the rows and columns of a two-way contingency table as points in low-dimensional vectors spaces. In performing CA, it is important to determine the intervariable relationship between attributes and other attributes, individuals and other individuals, and between individuals and attributes. It is also crucial to examine these intervariable relationships graphically by acquiring orthogonal information extracted from the variance-covariance data that can subsequently be used to form orthogonal axes from which the correlations between attributes and attributes, individuals and individuals, and attributes and individuals can be graphically displayed. Attributes and/or individuals that plot close to each other on these plots are related or correlated to each other. Consequently, if two individuals, e.g., two groundwater samples, plot close to each other, then these two individuals are strongly correlated. For this study, this means that the two

groundwater samples have very similar hydrochemical compositions.

The starting point of PCA is the description of the variability of n points in a p -dimensional space (n number of samples, p number of variables) by processing the variance-covariance matrix. The unit eigenvector coordinates are PCA loadings and describe a linearly independent combination of parameters that accounts for the maximum amount of variance in the data set. Principal component analysis scores represent projections of the data onto the eigenvectors. In this study, PCA was applied to a correlation matrix because the variables considered in this study vary by different orders of magnitude. In the correlation matrix, each variable is normalized to unit variance and contributes equally. The summarized information presented in the PCA plot is utilized for finding relationships, patterns, groundwaters that have extreme chemical composition and for further statistical modeling (Laaksoharju et al., 1999; Mumford et al., 2007).

Results and Discussions

Piper diagram has been shown to be very useful for depicting the cation exchange process, which is a major process that takes place in groundwater contamination events involving brines or other highly mineralized waters from the petroleum industry, the agricultural industry, manufacturing industries and from natural sources (Cates et al., 1996). Major chemical compositions of the groundwater indicate that three chemically distinct zones are present in the study area (Fig. 3). Fig. 3 is an extended Piper diagram of the samples taken from the study area. A considerable compositional variation was observed in the groundwater of study area. Hydrochemical composition defines a continuous trend from Ca/Na-HCO₃ groundwaters to Na-[(HCO₃+CO₃)/Cl] type after an input of sodium hypochlorite (NaOCl). Sodium hypochlorite is known as a disinfectant, which had high pH, high sodium and chloride concentration (Ko et al., 2002; Lee et al., 2008). The groundwater

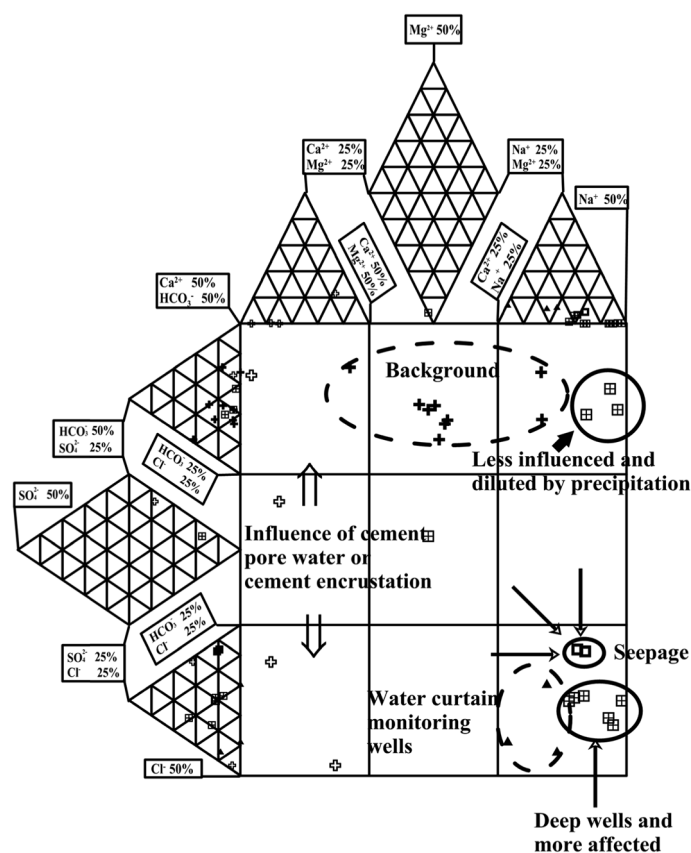


Fig. 3. Expanded Durov diagram from major cation and anion concentrations (units are milliequivalents per liter).

samples collected in the study area fairly show the effect of disinfectant and are classified by the influence of that. The seepage (PSP and BSP), water curtain monitoring wells (PMW, NPMW and BMW) and the groundwaters affected by the disinfection activity are plotted in the area of $\text{Na}-[(\text{HCO}_3+\text{CO}_3)/\text{Cl}]$ type in the Piper diagram. The chemical evolution of groundwater terminates at the seepage of caverns, which play a role as an integrated groundwater discharge.

A plugging of wells (YK5U, YK11U and YK14U) to the deposition of cementing encrustation by groundwater are illustrated in the area of $\text{Ca}-[(\text{HCO}_3+\text{CO}_3)/\text{Cl}]$ type in the Piper diagram (Smellie, 1998; Perkins and Palmer, 2000; Lee et al., 2008). The hydrochemistry indicates evolution of calcium-chloride-type from calcium-bicarbonate-type water, as a result of cement pollution, with no evidence of

reverse ion exchange (Chadha, 1999). However, water curtain monitoring wells (PMW, NPMW and BMW) contaminated by the dual effects, addition of the NaOCl and cement pollution, show different pattern of the previous explanations. These wells define less pH, sodium and calcium concentrations than both polluted disinfectant and cement wells because of dilution of $\text{Ca}-\text{HCO}_3$ groundwater from the overlying and nearby rock matrix as time goes. Chloride concentration (Cl) is, however, still high. These wells can be seen in the Piper diagram under cation exchange.

Multivariate statistics were used in this study to determine the relationships between chemical parameters and to make inferences, regarding the processes that account for these relationships. To conduct the study, I have closely followed the guidelines given by key publications on the subject (Ko et al., 2002; Lee et al., 2008; Ko et al., 2010). Their terminology and symbols

Table 1. Classification of groundwater samples by hydrogeochemical characteristics

Group		Sample	pH	Ca	Na	Cl	Characteristics
I	P ^a	YK2L, YK6, YK11L	<8.5	6–26	<35	<30	Wells with background composition
	B ^a	YK1, YK13U, YK13L					
	AT ^a	YK9U, YK9L, YK12U					
II	P	YK2U, YK3L, YK4, YK17, YK19, YK20	>9.0	3–21	>60	>30	Wells affected by disinfection
	B	YK14L, YK18					
	AT	YK12L					
III	P	PMW, NPMW	8–10	15–65	50–150	80–330	Water curtain monitoring wells
	B	BMW					
IV	P	PSP	8.5–9.5	11–21	70–130	75–180	Seepage
	B	BSP					
V	P	YK5U, YK5L, YK11U	10–12	75–520	10–50	10–220	Cement pollution
	B	YK10, YK14U					

^aP, B and AT denote propane, butane and access tunnel area, respectively.

Units of Na, Ca and Cl are mg L⁻¹.

Table 2. Summary of results by correspondence analysis (inertia and chi-square decomposition)

Singular values	Principal inertias	Chi-squares	Percents
0.615	0.379	2678.02	33.5%
0.545	0.297	2102.22	26.3%
0.410	0.168	1191.25	14.9%
	
	1.130	7991.55	

Degrees of freedom=405

were also adopted. In addition, wells polluted by cement pollution (YK5U, YK11U and YK 14U) were defined into Group V (Table 1).

Prior to a systematic hydrogeochemical evaluation, the groundwater data was first explored using a correspondence analysis (CA) to determine any broad regional grouping of data that might facilitate further interpretations. In the CA, a factor is represented by the eigenvector of the normalized covariance and correlation matrix of the data. From the CA, three factors were estimated to represent about 75% (33.5+26.3+14.9) of the variance of the original data (Table 2). The table consists of 16 attributes for individual wells and seepage and, consequently, represents a 28×16 matrix. The analysis, therefore, is made on the basis of 2-D plot to make easier the visualization of the variables grouping (Fig. 4). In interpreting the diagrams, one should keep in mind the spatial

importance attached to the factor order. A visual inspection of Fig. 4 indicates that three wells (YK5U, YK11U and YK14U), namely Group V (Polluted by cement contamination), situate close to calcium (Ca²⁺) and carbonate (CO₃²⁻). Nine wells located to the northeast are related to high concentration of HCO₃⁻ and NO₃⁻. Groundwaters forming Group I (Background) are typical direct recharge groundwaters of the cavern system. Most of wells (Group II and III, effect of disinfection activity and enhanced due to contamination in the figure, respectively) situated in the south area were affected disinfection activity and increased sodium and chloride concentrations. And then, the evolution of groundwater terminates at the seepage of caverns, which play a role as an integrated groundwater discharge. 3-D plot (see Fig. 5) makes it possible to distinguish YK11U from Group I. The variation is caused by NO₃⁻ and SO₄²⁻. The highest values of linkage in YK11U are mainly attributed to the high concentrations of NO₃⁻ (31.3 mg L⁻¹).

Saturation indices (SI) of various mineral phases were calculated using groundwater chemistry and the geochemical speciation program, WATEQ4F, included in the NETPATH (Plummer et al., 1976). Thermodynamic data used for these calculations were provided in the WATEQ database. The saturation index is defined as the log (activity product of the ions in solution: solubility product). Positive values of SI indicate that

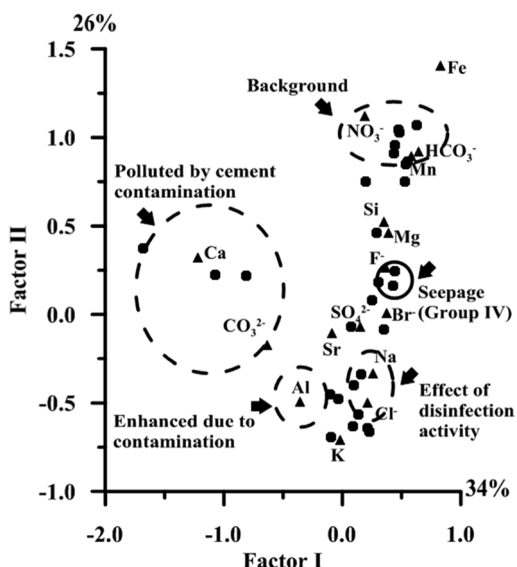


Fig. 4. Distribution of variables (▲) and samples (●) as calculated by correspondence analysis.

the solution is predicted to be oversaturated with respect to a particular mineral phase and precipitation would be expected. Negative values indicate that a solution in equilibrium with a given solid would show a higher concentrations of its ions than is measured (Langmuir, 1997). The approach taken in this paper is to assume that the fast-reacting minerals, such as

carbonates, have a much greater impact on groundwater chemistry than minerals that react relatively slowly. Additionally, possible equilibria between Al^{3+} and some minerals and pH changes in these equilibria, are discussed. If this assumption results in an acceptable approximation, then the reaction models that are constructed will be nonunique but nevertheless useful.

PCA was selected for this study because an unbiased and efficient tool was sought that would facilitate analysis of the saturation index that were calculated from the WATEQ4F. It was anticipated that this type of information would be useful in understanding the response of the groundwater system to natural or anthropogenic processes (Khan et al., 2008). As mentioned in the section of Methods, two major assemblages of data can be visually noted on the PCA graph (Fig. 6). Principal component 1 is plotted along the horizontal axis and principal component 2 is plotted along the vertical axis. The PCA was carried out by a diagonalization of the correlation matrix, so the problems arising from different measurement scales and numerical ranges of the original variables are avoided, since all variables are automatically to mean zero and variance unit. Tables 3 and 4 summarize the PCA results including

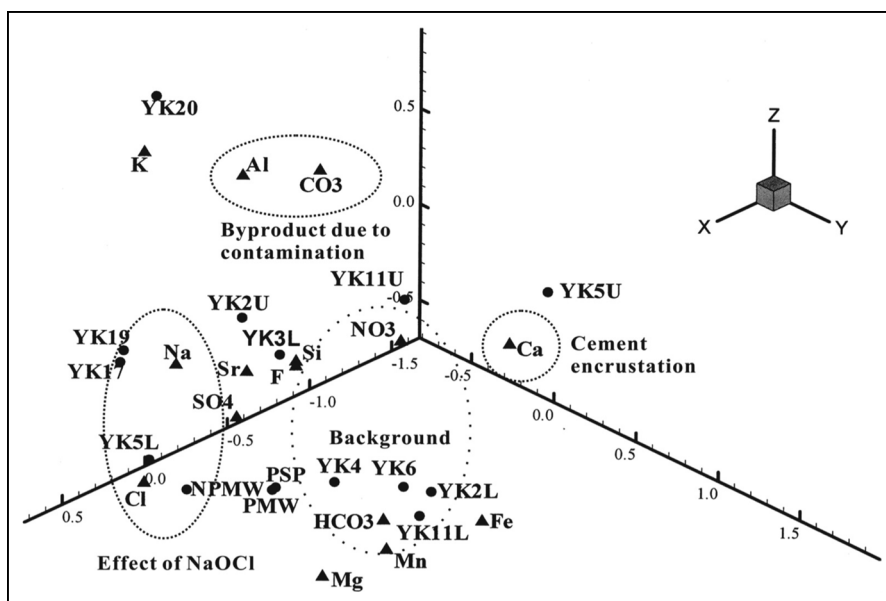


Fig. 5. 3-D plot of distributions of variables and samples as calculated by correspondence analysis in the propane area.

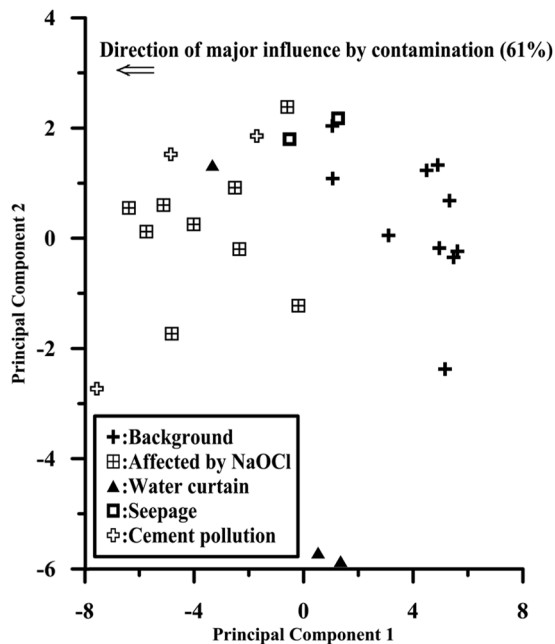


Fig. 6. Principal component analysis of calculated saturation index (SI) affecting on groundwater quality.

the loadings (participation of the original variables in the new ones) and the eigenvalue of each PC. PC 1 (horizontal) accounts for 61% of the total variance, while PC 2 accounts for 15% of the variation (A third axis accounts for 12% of the variation and the remaining is 12%). As shown in Fig. 6, PC 1 variation relative to PC 2 means a major factor

Table 3. Summary of results by principal component analysis combined geochemical speciation modeling (eigenvalue of the correlation matrix)

	Eigenvalue	Difference	Proportion	Cumulative
#PC1	17.619	-	0.608	0.608
#PC2	4.336	13.283	0.150	0.757
#PC3	3.600	0.736	0.124	0.881

#PC: Principal components

affecting groundwater quality relying on the horizontal axis. Fig. 7 shows the loadings on the PC 1. The interesting variables on PC 1 are the saturation index of calcite, dolomite and fractionation of Al. The more contaminated, the calcite saturation index may be increased. It means possible chemical clogging along the fractures can be suggested when calcite precipitation may be occurred.

Acidification of freshwater and groundwater with harmful effects on the aquatic biota is a widespread phenomenon (Baker and Schofield, 1982). The key role of Al in understanding damage to various aquatic ecosystems has been realized for some time. The toxicity of aqueous Al varies with the speciation. The simple hydroxides ($\text{Al}(\text{OH})_2^+$ and $\text{Al}(\text{OH})_3^+$) are regarded as the most dangerous forms while organically bound and polymeric forms are less toxic or essentially harmless. Knowledge of concentrations of various species of aqueous Al is therefore necessary to evaluate groundwater quality. Such information is also

Table 4. Summary of results by principal component analysis combined geochemical speciation modeling (eigenvectors)

Phase	PC1	PC2	Phase	PC1	PC2
calcite	-0.22	0.05	Ca-Montmorillonite	0.22	0.18
dolomite	-0.19	0.08	talc	-0.19	0.23
siderite	0.20	-0.07	illite	0.2	0.24
rhodochroite	-0.16	-0.02	hematite	0.1	0.03
strontinite	-0.22	0.05	goethite	0.11	0.02
gypsum	-0.07	0.4	$\text{Fe}(\text{OH})_3$ (a)	0.11	0.02
anhydrite	-0.07	0.4	pyrolusite	-0.18	0.11
celesite	-0.07	0.42	hausmannite	-0.23	0.06
fluorite	-0.11	0.05	manganite	-0.21	0.08
SiO_2 (a)	0.17	0.31	pyrochroite	-0.23	0.01
chalcedony	0.17	0.31	P_{CO_2}	0.23	-0.04
quartz	0.17	0.31	melanterite	0.22	0.00
gibbsite (c)	0.22	-0.04	alunite	0.23	0.03
$\text{Al}(\text{OH})_3$ (a)	0.22	-0.04	K-jarosite	0.21	0.09
chlorite	-0.21	0.13			

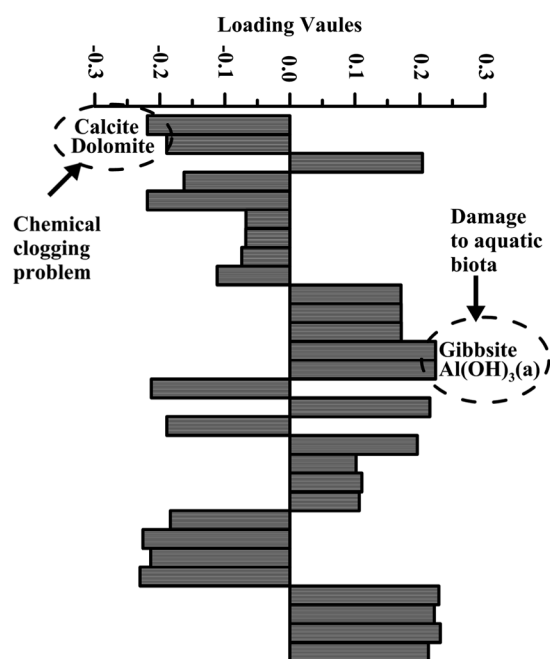


Fig. 7. Bar graphs of eigenvectors for 29 statistically determined variables, indicating relative importance of variables on principal component axis.

important for understanding the water-rock interactions in a study area. Methods for determining various forms of Al separately have been described by several authors (Seip et al., 1984). However, the present work is related to examine the hydrogeochemical processes controlling the groundwater chemistry. Generally, high Al concentration can be easily observed in the environment with low pH such as acid mine drainage. In the study area, the diagnostic high Al concentration in the samples with high pH are observed. Few studies except for environmental study have not reported the high Al in high pH natural environment (Ko et al., 2002). This can be explained in the activity diagram of dissolved aluminum species with gibbsite, kaolinite, and pyrophyllite (Fig. 8). Al in pristine groundwater of neutral pH (6-8) oversaturated and understaturated in high pH with gibbsite. The results discussed above demonstrate that high pH by contamination for disinfection or cement pollution occurs aluminum dissolution or phase change of Al species and oversaturation of calcite.

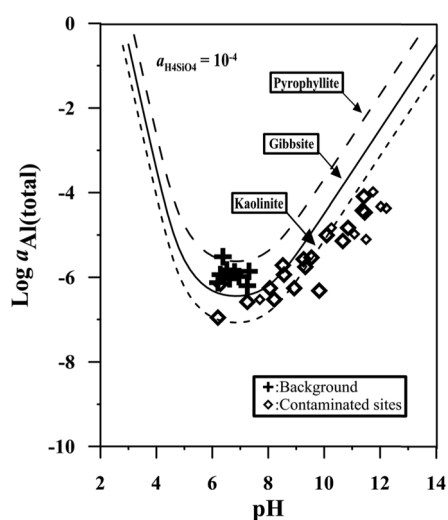


Fig. 8. Sum of activities of dissolved aluminum species in equilibrium with gibbsite, kaolinite, and pyrophyllite at 10^{-4} M of H_4SiO_4 .

Summary

Multivariate statistical analyses, correspondence analysis (CA) and principal component analysis (PCA), have been utilized for hydrochemical measurements to analyze and interpret the data from LPG underground storage caverns, which is located in the southeastern Korea. Using both PCA and CA, this work investigates two main anthropogenic factors due to the operations of the underground storage caverns, in which the highly alkaline groundwater are an analogue for the repository system. Based on the two contaminants, sodium hypochlorite (NaOCl) and cement pollution, wells were classified into five groups in the study area. Using the CA, the visualizations of the grouping between variables and samples facilitate the explanations of the processes that account for the associations. High pH, speciation of Al and possible precipitation of calcite characterize these groundwaters by the PCA with hydrogeochemical speciation, in particular, saturation index. Comparisons made between the two statistical results and the classifications from the previous studies underscore the descriptive capabilities of PCA and CA as an exploratory tool in hydrogeochemical studies.

Acknowledgments

This work was supported by the Basic Research Project (14-3211-2) of the Korea Institute of Geoscience and Mineral Resources (KIGAM) funded by the Ministry of Science, ICT and Future Planning. I appreciate Dr. Dong-Chan Koh and an anonymous reviewer, whose comments led to improvements.

References

- Baker, J.P. and Schofield, C.L., 1982, Aluminum toxicity to fish in acidic waters. *Water, Air and Soil Pollution*, 18, 289-309.
- Cates, D.A., Knox, R.C., and Sabatini, D.A., 1996, The impact of ion exchange processes on subsurface brine transport as observed on piper diagrams. *Ground Water*, 34, 532-544.
- Chadha, D.K., 1999, A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. *Hydrogeology Journal*, 7, 431-439.
- David, J.C., 1973, *Statistics and Data Analysis in Geology*. John Wiley and Sons, Incorporated, New York, USA, 550 p.
- Eaton, A.D., Clesceri, L.S., and Greenberg, A.E., 1995, *Standard methods for the examination of water and wastewater*. American Public Health Association, Washington, USA, 1100 p.
- Geostock, 1985, Yugong LPG storage project in Ulsan final reconnaissance report. Geostock, Paris, France, 55 p.
- Geostock, 1992, Ulsan LPG storage cavern Yugong gas hydrogeological monitoring. Geostock, Paris, France, 68 p.
- Giménez, E. and Morell, I., 1997, Hydrogeochemical analysis of salinization processes in the coastal aquifer of Oropesa (Castellón, Spain). *Environmental Geology*, 29, 118-131.
- Jeong, Y.-K., 2012, Cluster analysis of the 1000-hPa height field around the Korean Peninsula. *Journal of Korean Earth Science Society*, 33, 337-349.
- Jobson, J.D., 1992, *Applied multivariate data analysis*. Springer-Verlag, New York, USA, 731 p.
- Join, J.L., Coudray, J., and Longworth, K., 1997, Using principal components analysis and Na/Cl ratios to trace groundwater circulation in a volcanic islands: The example of Reunion. *Journal of Hydrology*, 190, 1-18.
- Khan, S., Chen, H.F., and Rana, T., 2008, Optimizing ground water observation networks in irrigation areas using principal component analysis. *Ground Water Monitoring and Remediation*, 28, 93-100.
- Kim, T., Lee, K.-K., Ko, K.-K., and Chang, H.W., 2000, Groundwater flow system inferred from hydraulic stresses and heads at an underground LPG storage cavern site. *Journal of Hydrology*, 236, 165-184.
- Ko, K.-S., Lee, J., Lee, K.-K., and Chang, H.-W., 2010, Multivariate statistical analysis for groundwater mixing ratios around underground storage caverns in Korea. *Carbonates and Evaporites*, 25, 35-42.
- Ko, K.-S., Chang, H.W., Lee, K.-K., and Kim, T., 2002, Factors affecting the groundwater system around an underground LPG storage cavern. *Quarterly Journal of Engineering Geology and Hydrogeology*, 35, 279-290.
- Kwon, B.D. and Kim, C., 1997, Application of multivariate statistical analysis technique in landfill investigation. *Journal of Korean Earth Science Society*, 18, 515-521.
- Laaksoharju, M., Skårman, C., and Skårman, E., 1999, Multivariate mixing and mass balance (M3) calculations, a new tool for decoding hydrogeochemical information. *Applied Geochemistry*, 14, 861-871.
- Langmuir, D., 1997, *Aqueous environmental geochemistry*. Prentice Hall, NJ, USA, 600 p.
- Lee, J., Ko, K.-S., Kim, J.-M., and Chang, H.-W., 2008, Multivariate statistical analysis of underground gas storage caverns on groundwater chemistry in Korea. *Hydrological Processes*, 22, 3410-3417.
- Lee, J., Jung, B., Kim, J.-M., Ko, K.-S., and Chang, H.-W., 2011, Determination of groundwater flow regimes in underground storage caverns using tritium and helium isotopes. *Environmental Earth Sciences*, 63, 763-770.
- Lindblom, U., 1989, The performance of water curtains surroundings rock caverns used for gas storage. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 26, 85-97.
- Mumford, K., MacGregor, J.F., Dickson, S.E., and Frappa, R.H., 2007, Multivariate analysis of groundwater and soil data from a waste disposal site. *Ground Water Monitoring and Remediation*, 27, 92-102.
- Park, Y.D. and Yoon, H.D., 1968, Geologic map of Korea (Bangeojin sheet 1:50,000). Geological Survey of Korea.
- Perkins, R.B. and Palmer, C.D., 2000, Solubility of $\text{Ca}_6[\text{Al}(\text{OH})_6]_2(\text{CrO}_4)_3 \cdot 26\text{H}_2\text{O}$, the chromate analog of ettringite; 5-75°C. *Applied Geochemistry*, 15, 1203-1218.
- Plummer, L.N., Jones, B.F., and Truesdell, A.H., 1976, WATEQF; A FORTRAN IV Version of WATEQ: A computer program for calculating chemical equilibrium of natural waters. U. S. Geological Survey Water-Resources Investigation Report, 76-13, 66 p.

- Seip, H.M., Mller, L., and Naas, A., 1984, Aluminum speciation: Comparison of two spectrophotometric analytical methods and observed concentrations in some acidic aquatic systems in southern Norway. *Water, Air and Soil Pollution*, 23, 81-95.
- Smellie, J.A.T., 1998, MAQARIN natural analogue study: Phase III. Technical Report TR-98-04, Stockholm, Sweden, 350 p.
- Stetzenbach, K.J., Farnham, I.M., Hodge, V.F., and Johannesson, K.H., 1999, Using multivariate statistical analysis of groundwater major cation and trace element concentrations to evaluate groundwater flow in a regional aquifer. *Hydrological Processes*, 13, 2655-2673.
- Stumm, W. and Morgan, J.J., 1996, *Aquatic chemistry*. Wiley, NY, USA, 1040 p.
- Valenchon, F., 1982, The use of correspondence analysis in geochemistry. *Mathematical Geology*, 14, 331-342.

Manuscript received: June 16, 2014

Revised manuscript received: July 8, 2014

Manuscript accepted: July 25, 2014