Bull. Environ. Contam. Toxicol. (2006) 76:1–7

DOI: 10.1007/s00128-005-0882-x

Bulletin of Environmental Contamination and Toxicology

Geostatistical Analysis for Hydrogeochemical Characterization of the Han River, Korea: Identification of Major Factors Governing Water Chemistry

J.-S. Ryu,[1,4] K.-S. Lee,[1] J.-H. Kim,[2] K.-H. Ahn,[3] H.-W. Chang[4]

[1] Environmental Tracer Team, Division of Isotope Geoscience, Korea Basic Science Institute, Yuseong-gu, Daejeon, 305-333, Korea
[2] Korea Institute of Geoscience and Mineral Resources, Yuseong-gu, Daejeon, 305-350, Korea
[3] Korea Institute of Science and Technology, Seongbuk-gu, Seoul, 136-791, Korea
[4] School of Earth and Environmental Sciences, Seoul National University, Kwanak-gu, Seoul, 151-742, Korea

Received: 1 November 2004/Accepted: 12 October 2005

The Han River is the largest river in South Korea in terms of its length, drainage and mean annual discharge. The river is of importance for cultural and economic reasons because it is a unique source of drinking water for more than 20 million people living in Seoul and its surrounding areas. Intensive industrial development in Seoul during the last four decades has greatly increased the risk for pollution of river water and deterioration of water quality.

In this study, multivariate statistical analysis including cluster analysis (CA) and principal component analysis (PCA) have been used to investigate the factors influencing the water quality of the Han River, and to discriminate the effect of human activities on the river hydrochemistry.

MATERIALS AND METHODS

The study area is the Han River basin with a length of 481.7 km and a drainage area of ~26,000 km^2 (Data from the Han River Flood Control Office). The mean annual discharge is estimated from 16.0 to 18.9 km^3 (Vorosmarty et al. 1998). The two major tributaries, the North and the South Han Rivers, join at Yangsu-Ri, forming the main channel.

A total of 32 samples were collected from two main stems and their tributaries during wet and dry seasons of 2000 (Fig. 1b). Sampling, preservation and chemical analysis were conducted based on the standard procedures of APHA (1998). Temperature (T), electrical conductivity (EC) and pH were measured using temperature-compensated electronic meters in the field. Alkalinity was determined by the Gran titration method using 0.1N HCl. For the chemical analysis, water samples were filtered in the field with a hand-held syringe using 0.45μm membrane filters and then acidified to pH<2 with concentrated HNO_3. Fourteen elements (Na, Mg, Si, S, K, Ca, Sr, Al, Mn, Ni, Cu, Zn, Mo and Ba) were analyzed by ICP-AES and ICP-MS at Korea Basic Science Institute (KBSI) and two anions (Cl and NO_3) by an ion chromatograph at Seoul National University.

Sixteen chemical constituents (Na, Mg, Si, S, K, Ca, Sr, Al, Mn, Ni, Cu, Zn, Mo,

Correspondence to: H.-W. Chang

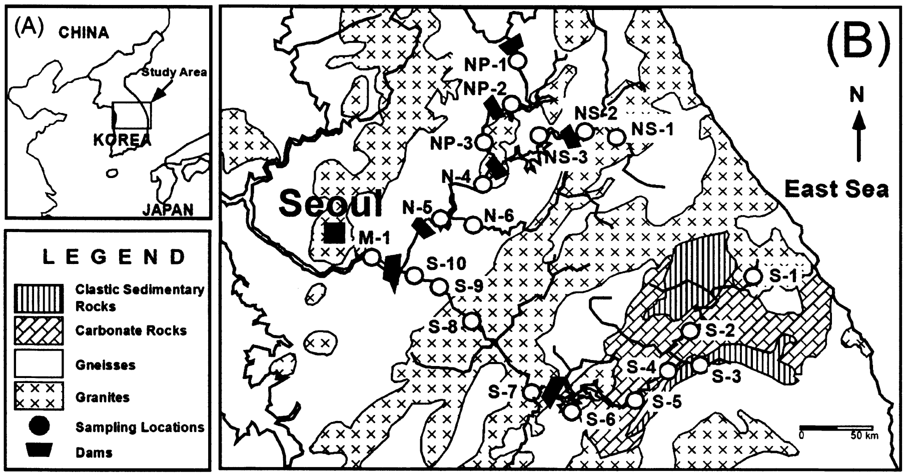

Figure 1. Maps showing the sampling sites and lithological characteristics of the study area.

Ba, Cl and NO_3) were used for the multivariate statistical analysis. Values for constituents lower than the method detection limits (< DL) were substituted with DL/2 prior to statistical analysis (Farnham et al. 1998). Each variable was standardized to the mean 0.0 and variance 1.0 in order to avoid misclassifications arising from the different order of magnitude of the variables. For the multivariate statistical analysis, the SAS 8.0 package was used. Cluster analysis (CA) was performed using Ward's method. This method uses an analysis of variance approach to evaluate the distance between clusters, attempting to minimize the sum of squares of any two (hypothetical) clusters that could be formed at each step. In general, it is regarded as very efficient (Vega et al. 1998; Helena et al. 2000). Based on the standardized data, R-mode factor analysis (FA) including principal component analysis (PCA) was carried out to reveal the interaction among the variables. In general, it attempted to simplify the complex and diverse relationships that exist among a set of observed variables by revealing common and unobserved factors that link together with the seemingly unrelated variables.

RESULTS AND DISCUSSION

Based on the standardized chemical data, cluster analysis was performed to split the water sampling points into a finite number of groups (zones) with similar hydrogeochemical composition. Figure 2 shows the resulting dendrogram of the study area. The groups are mainly discriminated by their location: Group A and Group B are from the South Han River basin except for one sample (N-6), whereas Group C is from the North Han River basin. Therefore, the Han River water can be classified into two major groups according to their geology.

To examine the chemical variation of groups and the difference among the groups, box-and-whisker plots of individual variables were plotted (Figure 3). It shows an

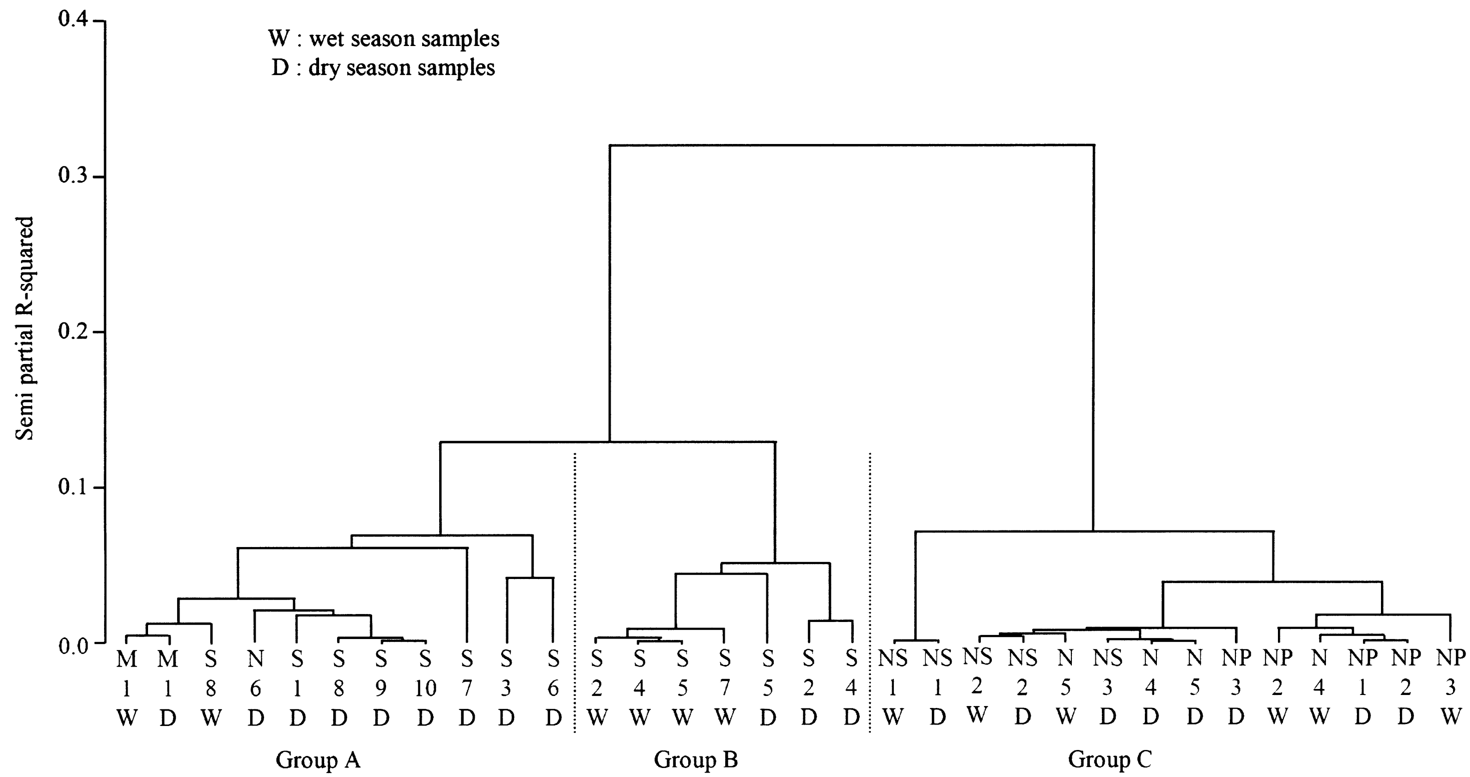

Figure 2. Dendrogram using Ward's method for dry (D) and wet (W) seasons.

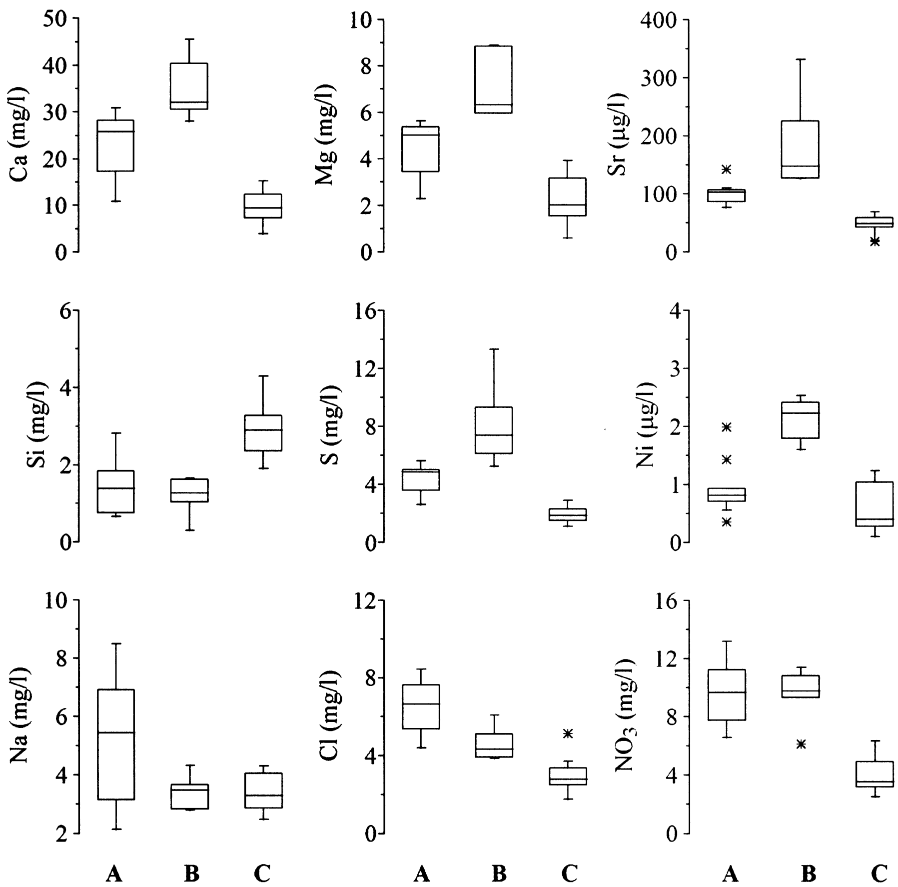

Figure 3. Box plots for Ca, Mg, Sr, Si, S, Ni, Na, Cl and NO_3 in three groups.

example of box plots for some meaningful variables related to the water quality. By inspecting these plots, it is possible to perceive differences among the groups. Ca, Mg, Sr, S and Ni have an increasing trend in the order of groups C, A, B, whereas Na, Cl and NO_3 have a different trend which increases in the order of groups C, B, A. It seems like that Group B having the highest concentration of Ca, Mg, Sr, S and Ni has been affected by the carbonates and clastic sediments as well as metallic deposits, whereas Group A which has a median value of Ca, Mg, Sr, S and Ni and the highest in Na, Cl and NO_3 by both carbonates and human activities.

With the correlation matrix between the 16 variables, significant correlation over 0.70 is found between Mg, Si, S, Ca, Sr, Ni, NO_3 and Cl, suggesting that these variables are strongly influenced by the same factor (Table 1). Based on the eigenvalues (> 1.5) calculated by the principal component analysis, the first three factors are selected to represent the hydrogeochemical processes of the water,

Table 1. Correlation matrix among 16 chemical compositions of water.

	Na	Mg	Si	S	K	Ca	Sr	Al	Mn	Ni	Cu	Zn	Mo	Ba	Cl	NO_3
Na	1.00															
Mg	0.06	1.00														
Si	-0.23	**-0.80**	1.00													
S	0.08	**0.94**	-0.66	1.00												
K	0.20	0.22	-0.37	0.12	1.00											
Ca	0.04	**0.98**	**-0.79**	**0.95**	0.19	1.00										
Sr	0.07	**0.89**	-0.59	**0.97**	0.07	**0.89**	1.00									
Al	-0.30	0.25	-0.14	0.25	-0.01	0.28	0.20	1.00								
Mn	0.07	0.18	-0.12	0.16	0.19	0.20	0.06	0.33	1.00							
Ni	-0.12	**0.73**	-0.49	**0.70**	0.11	**0.72**	0.65	0.51	0.27	1.00						
Cu	0.26	0.47	-0.40	0.36	0.55	0.41	0.32	0.21	0.35	0.40	1.00					
Zn	-0.09	0.12	-0.32	0.06	0.12	0.15	0.05	0.12	0.07	-0.07	0.03	1.00				
Mo	-0.10	0.03	-0.15	-0.02	0.04	0.10	-0.16	-0.20	-0.09	-0.05	-0.23	0.09	1.00			
Ba	0.04	0.14	-0.26	-0.03	0.45	0.03	-0.05	-0.16	0.14	-0.10	0.29	0.10	-0.22	1.00		
Cl	**0.70**	0.48	-0.64	0.46	0.45	0.53	0.42	-0.11	0.12	0.11	0.33	0.09	0.17	0.07	1.00	
NO_3	0.28	**0.76**	**-0.71**	**0.74**	0.21	**0.84**	**0.72**	0.17	0.10	0.42	0.32	0.20	0.13	-0.15	**0.78**	1.00

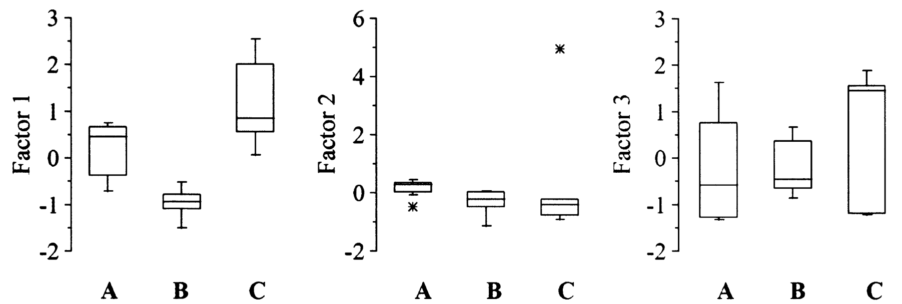

Figure 4. Box plots for the three factor scores in three groups.

Table 2. Rotated factor pattern of three factors after varimax rotation.

	Factor 1	Factor 2	Factor 3
Na	-0.003	0.903	0.077
Mg	0.962	0.062	0.171
Si	-0.730	-0.252	-0.333
S	0.963	0.076	-0.004
K	0.084	0.326	0.702
Ca	0.964	0.100	0.068
Sr	0.950	0.051	-0.016
Al	0.254	-0.319	-0.143
Mn	0.023	0.132	0.153
Ni	0.740	-0.171	-0.047
Cu	0.328	0.281	0.517
Zn	0.071	-0.134	0.239
Mo	-0.002	0.111	-0.256
Ba	-0.021	-0.084	0.907
Cl	0.428	0.835	0.139
NO_3	0.774	0.443	-0.077

without losing significant information. For a simpler and easier interpretation, the rotated factor pattern after factor extraction using varimax rotation is given in Table 2 and box plots of the chosen factor scores are plotted (Figure 4).

The first factor, accounting for 40.4% of the total variance, is characterized by high positive loadings for Mg, S, Ca, Sr, Ni and NO_3, and negative loadings for Si. It tends to increase median factor scores in the following the sequence: Group C < A < B. As above mentioned, this factor reveals that the carbonates and clastic sediments have mainly affected the water quality. The second factor, accounting for 14.0% of the total variance, is mainly associated with very high positive loadings of only Na and Cl. This factor, which differentiates Group A from Group B, seems to be related to human activities because no evaporates have been found within the Han River basin. The third factor, accounting for 11.5% of total variance, is associated with high positive loadings of Ba and K. Its median factor scores in Group A and C are higher than in Group B. Factor 3 seems to be mostly associated with the silicates.

Acknowledgements. We thank the Green Korea 21 Program of the Korea Institute of Science and Technology and the BK21 program, School of Earth and Environmental Sciences, Seoul National University for financial support of this work.

REFERENCES

APHA, AWWA, WEF (1998) Standard methods for the examination of water and wastewater. 20th Ed. APHA Washington, DC

Farnham I, Smiecinski A, Singh, AK (1998) Handling chemical data below detection limits for multivariate analysis of groundwater, First International Conference on Remediation of Chlorinated and Recalcitrant Compounds. Monterey, CA: 99-104

Helena B, Prardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000) Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Wat Res 34:807-816

Vega M, Pardo R, Barrado E, Deban L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Wat Res 32:3581-3592

Vorosmarty CJ, Fekete B, Tucker BA (1998) River Discharge Database, Version 1.1 (RivDIS v1.0 supplement). Available through the Institute fro the Study of Earth, Oceans and Space, University of New Hampshire, Durham NH (USA)