Research Journal of Biology Sciences
Biyoloji Bilimleri Araştırma Dergisi
E-ISSN: 1308-0261, 11(1): 1-4, 2018, www.nobel.gen.tr

Use of Principle Component Analysis (PCA) to Evaluate the Groundwater Resources of Havsa District (Edirne, TURKEY)

Cem TOKATLI[1], Arzu ÇİÇEK[2], Esengül KÖSE[3]*, Alper UĞURLUOĞLU[4], Hayri DAYIOĞLU[5], Özgür EMİROĞLU[6]

[1]Trakya University, Ipsala Vocational School, Department of Laboratory Technology Program, İpsala/Edirne, TURKEY
[2]Anadolu University, Applied Environmental Research Centre, Eskişehir, TURKEY
[3]Eskişehir Osmangazi University, Eskişehir Vocational School, Department of Environmental Protection and Control, Eskişehir, TURKEY
[4]Ministry of Forestry and Water Affairs, General Directorate of Water Management, Ankara, TURKEY
[5]Dumlupınar University, Faculty of Science, Department of Biology, Kütahya, TURKEY
[6]Eskişehir Osmangazi University, Faculty of Science, Department of Biology, Eskişehir, TURKEY

*Sorumlu Yazar
E-posta: ekose@ogu.edu.tr

Geliş Tarihi: 24 Mayıs 2018
Kabul Tarihi: 31 Temmuz 2018

Abstract
Havsa District is located in the Edirne Province in Thrace Region of Turkey and provides drinking water to about 30.000 people. This study was carried out to evaluate the groundwater quality of Havsa District by using some multivariate statistical techniques. For this purpose, groundwater samples were collected from 15 stations from the villages of Havsa District in winter season of 2016. Some water quality parameters including dissolved oxygen, % oxygen saturation, total dissolved solids (TDS), salinity, electrical conductivity (EC), pH, sulphate (SO_4), nitrate (NO_3), phosphate (PO_4) and chlorine (CI) in groundwater of region were determined and Factor Analysis (FA), Cluster Analysis (CA) and Pearson Correlation Index (PCI) were applied to detected data in order to evaluate the groundwater quality. According to the results of PCI, significance positive and negative relations were determined among the investigated parameters at the 0.05 and 0.01 levels. According to results of FA, 2 factors named as "Agricultural Factor" and "Oxygenation Factor" explained 87.16% of the total variance. According to CA, 2 statistically significant clusters were formed, which were corresponded to polluted and unpolluted stations.
Keywords: Havsa District, Groundwater Quality, Factor Analysis, Cluster Analysis, Pearson Correlation Index.

INTRODUCTION

One of the most contaminated parts of environment is of course freshwater resources and pollution caused by especially agricultural and industrial activities reduces the quality of the limited freshwater [1]. Only about 2.8% of water is suitable and fresh for human consumption and about 30.1% of these freshwaters are underground [2]. It is known that groundwater pollution is one of the most serious problems relating to human health in the twenty-first century. Groundwater is source of drinking water, irrigation and industrial processes for numbers of villages, districts, provinces etc.. But unfortunately many organic and inorganic pollutants sourced from anthropogenic activities have been identified as strong contaminants found in groundwater [3-7]. Therefore, monitoring groundwater quality has an importance both for human health and ecosystem health especially in rural areas.

Havsa District contains very large and productive agricultural lands and economic life is largely based on agriculture. The majority of the population including the district center and the villagers is engaged in agriculture and animal husbandry. The most important agricultural products are wheat, sunflower, rice, corn, sugar beet, vineyard, fruit and other plants and the most important farming products are milk and fattening cattle and sheep [8]. The aim of this study was to evaluate the groundwater quality of Havsa District by a statistical approach.

MATERIALS and METHODS

2.1. Study Area and Collection of Samples

The Havsa District with a surface area of 49.635 square meters is 27 km away from the Edirne City center and its altitude is about 30 meters. The land is usually in the form of plain and has a young age because of the mass formation of geological periods. The area has no mountains, highlands, rivers and forests, so it is dominated by agricultural land. Terrestrial climate prevails throughout the district during all the year [8].

In this study, groundwater samples were collected in winter season of 2016 from 15 stations from the drill fountains of the villages located in the Havsa District. Water samples were then collected at the outflow of drill pump in polyethylene bottles. Coordinate information and localities of selected stations are given in Table 1 and the topographic map of Havsa District with selected stations is given in Figure 1.

Table 1. Location properties of selected stations

Station Number	Locality	Coordinates	
		North	South
S1	Bakışlar	41.458108	26.831325
S2	Tahal	41.42818	26.856707
S3	Naipyusuf	41.483062	26.904915
S4	Yolageldi	41.518485	26.949749
S5	Bostanlı	41.610994	26.97024
S6	Köseömer	41.58754	26.888348
S7	Havsa	41.549663	26.819859
S8	Osmalı	41.587715	26.835672
S9	Hasköy	41.639331	26.858687
S10	Arpaç	41.690341	26.880257
S11	Habiller	41.669214	26.795204
S12	Oğulpaşa	41.605067	26.752064
S13	Abalar	41.546367	26.740826
S14	Azatlı	41.502259	26.701552
S15	Şerbettar	41.461724	26.759727

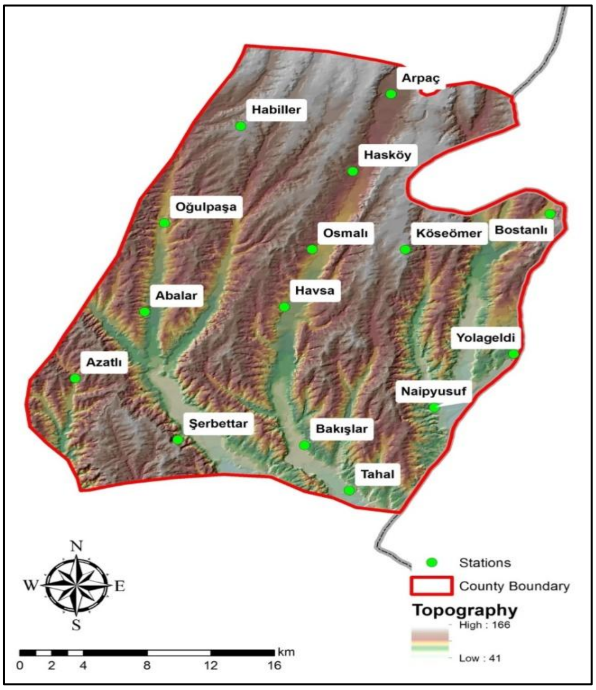

Figure 1. Topographic map of Havsa District

2.2. Physical and Chemical Analysis

Dissolved oxygen, % oxygen saturation, total dissolved solids (TDS), salinity, electrical conductivity (EC) and pH parameters were determined by using "Hach Lange HQ40D Multiparameter" device during the field studies and sulphate (SO_4), nitrate (NO_3), phosphate (PO_4) and chlorine (CI) were determined by using "Hach Lange DR3900 Spectrophotometer" device during the laboratory studies.

2.3. Statistical Analysis

Pearson Correlation Index (PCI) was applied to the results in order to determine the relations between the investigated parameters by using the SPSS statistical package program (version 17). Multivariate techniques such as cluster analysis(CA) and factor analysis (FA) have been used increasingly in studies conducted to measure and monitor groundwater because these techniques useful for identifying relationships between variables [9-11]. Factor Analysis (FA) was applied to the results in order to determine the effective varifactors on groundwater of Havsa District according to correlated variables by using the SPSS statistical package program (version 17). Cluster Analysis (CA) was applied to the results in order to classify the investigated villages according to contamination status by using the Past statistical package program.

RESULTS

3.1. Pearson Correlation Index (PCI)

The relations among the investigated physical and chemical water quality parameters in groundwater of Havsa District were determined by using all the data (n = 15 for all parameters) and all the detected relations are given in Table 2.

It was found that, the relations between dissolved oxygen – % oxygen saturation (+), pH (+) and PO_4 (-); between % oxygen saturation – pH (+) and PO_4 (-); between TDS – salinity (+), EC (+), SO_4 (+), NO_3 (+), PO_4 (+) and Cl (+); between salinity – EC (+), SO_4 (+), NO_3 (+), PO_4 (+) and Cl (+); between EC – SO_4 (+), NO_3 (+), PO_4 (+) and Cl (+); between pH – PO_4 (-); between SO_4 – PO_4 (+) and Cl (+); between NO_3 – Cl (+); between PO_4 – Cl (+) levels were directly proportional at the 0.01 significance level. It was also found that the relations between dissolved oxygen – SO_4 (-); between % oxygen saturation – SO_4 (-); between SO_4 – NO_3 (+); between NO_3 – PO_4 (+) levels were directly proportional at the 0.05 significance level.

Table 2. PCI coefficients of parameters

Parameters	DO	O_2 sat	TDS	Sal.	EC	pH	SO_4	NO_3	PO_4	Cl
DO	1									
O_2 sat	.995**	1								
TDS	-.408	-.446	1							
Sal.	-.404	-.444	1.000**	1						
EC	-.413	-.451	1.000**	1.000**	1					
pH	.708**	.678**	-.393	-.387	-.400	1				
SO_4	-.621*	-.635*	.707**	.702**	.712**	-.529*	1			
NO_3	-.367	-.413	.867**	.868**	.864**	-.383	.528*	1		
PO_4	-.737**	-.734**	.669**	.662**	.674**	-.727**	.831**	.591*	1	
Cl	-.383	-.413	.920**	.919**	.923**	-.492	.710**	.688**	.709**	1

DO: Dissolved oxygen; O_2 sat: % oxygen saturation; Sal.: Salinity
*: Correlation is significant at the 0.05 level ($p<0.05$);
**: Correlation is significant at the 0.01 level ($p<0.01$)

3.2. Factor Analysis (FA)

In the present application, a total of 15 variables (all the investigated parameters) were used to detect the effective varifactors (n = 15 for all parameters). Result of KMO (Kaiser – Meyer – Olkin) test that presents the measure of sampling adequacy was 0.58 and this value means that the sampling adequacy was in a good level for the this application (>0.5) [9].

Eigenvalues higher than one were taken as criterion for evaluate the principal components that required to explain the sources of variance in the data (Figure 2). According to rotated cumulative percentage variance, 2 factors explained 87.16% of the total variance (Figure 2).

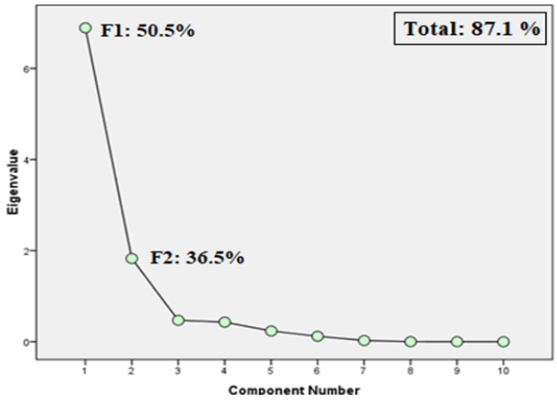

Figure 2. Scree plot of FA

The parameter loadings higher than 0.5 calculated after rotation for 2 components are given in Figure 3. Also component plot in rotated space that shows the related variables of detected 2 factors is given in Figure 4. Liu et al. 2003 [9], classified the factor loadings according to loading values as strong (>0.75), moderate (0.75 – 0.50) and weak (0.50 – 0.30).

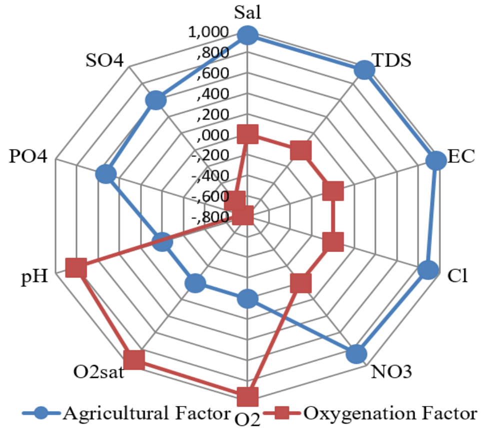

Figure 3. Rotated component matrix

First factor (F1), named as "Agricultural Factor" explained 50.50% of total variance and it was related to the variables of total dissolved solids, salinity, electrical conductivity, sulphate, nitrate, phosphate and chlorine parameters. Total dissolved solids, salinity, electrical conductivity, nitrate, and chlorine parameters were strong positively and sulphate and phosphate parameters were moderate positively loaded with this factor (Figure 3 and 4).

Second factor (F2), named as "Oxygenation Factor" explained 36.65% of total variance and it was related to the variables of dissolved oxygen, % oxygen saturation, pH, sulphate and phosphate parameters. Dissolved oxygen, % oxygen saturation and pH parameters were strong positively, phosphate parameter was strong negatively and sulphate parameter was moderate negatively loaded with this factor (Figure 3 and 4).

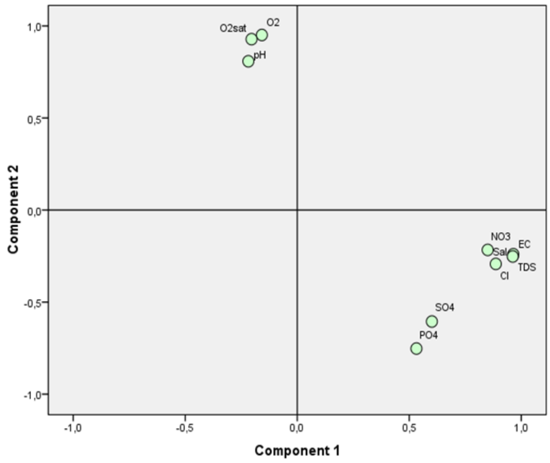

Figure 4. Component plot in rotated space

3.2. Cluster Analysis (CA)

The diagram of CA calculated by all the investigated parameters was given in Figure 5. According to results of CA, 2 statistically significant clusters were formed: Cluster 1 corresponded to Naipyusuf, Yolageldi, Hasköy, Habiller and Şerbettar villages, where were determined as the polluted parts of the investigated area; Cluster 2 corresponded to Bakışlar, Tahal, Bostanlı, Köseömer, Havsa, Osmalı, Arpaç, Oğulpaşa, Abalar and Azatlı villages, where were determined as the uppolluted parts of the investigated area. Maximum similarity was observed between Oğulpaşa and Bakışlar villages (99%) and minimum similarity was observed between Yolageldi and Arpaç villages (61%).

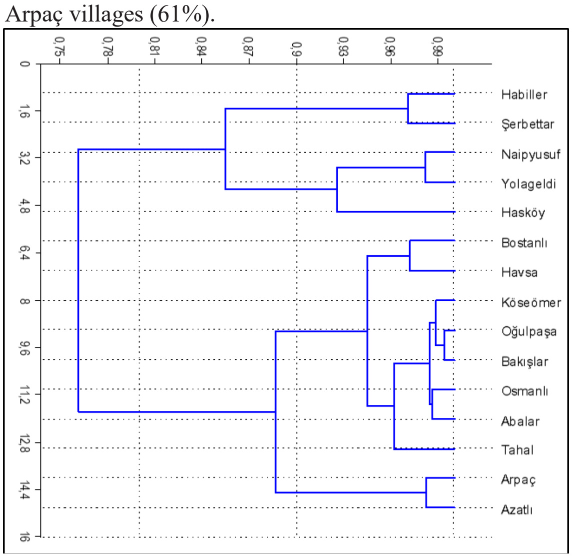

Figure 5. Diagram of CA

DISCUSSION

Factor and Cluster Analysis that provide valuable and easy explaining data are two of the most widely used multivariate statistical techniques. These are being used in many countries in order to evaluate many different aquatic habitats [12-18].

In a study performed in Edirne Province of Turkey, groundwater quality of İpsala District was evaluated by using CA and FA. According to results of CA, 8 statistically significant clusters were formed in terms of groundwater quality characteristics of investigated villages in İpsala District and according to results of FA, 3 factors explained 74.867% of the total variance. As similar to this study "Agricultural Factor" that explained 22.9% of total variance was an effective component on groundwater quality of İpsala District and nitrate parameter was strong positively loaded with this factor. Also dissolved oxygen and conductivity parameters had opposite factor loadings as in freshwater resources of Balkan Arboretum Area [5].

In another study performed in Aegean – Central Anatolia Regions of Turkey, groundwater quality of Türkmen Mountain was evaluated by using CA and FA. According to the results of CA determined by using physiochemical parameters, maximum similarity was recorded as 99% and minimum similarity was recorded as 57% among the investigated villages on the mountain. According to results of FA, 4 factors explained 79.25% of the total variance. As similar to this study "Agricultural Factor" that explained 14.06% of total variance was an effective component on groundwater quality of Türkmen Mountain and phosphate parameter was moderate positively loaded with this factor [10].

In a macroscopic point of view as a result of this study, multistatistical techniques like Factor and Cluster Analysis are necessary for a sophisticated evaluation in freshwater quality assessment studies, because of determined very large numbers of data and also difficulty of the interpretations of these data.

ACKNOWLEDGEMENT

The author would like to thank for the financial and technical supports supplied by Trakya University, Turkey. This investigation has been supported by the project numbered as 2016/247 accepted by Trakya University, Commission of Scientific Research Projects.

REFERENCES

[1] Tokatlı, C. 2015. Assessment of the Water Quality in the Meriç River: As an Element of the Ecosystem in the Thrace Region of Turkey. Polish Journal of Environmental Studies. 24 (5): 2205-2211.

[2] Gupta, A. D. 1997. Importance of groundwater as water resource. Proceedings of Seminar and Training on Groundwater Contaminated by Hazardous Substances, Bangkok, January 20–21.

[3] Canter, L. W. 1987. Groundwater Quality Protection. Lewis Publications, Inc, Chelsea, MI.

[4] Hudak, P. F. 1999. Chloride and Nitrate Distributions in the Hickory Aquifer, Central Texas, USA. Environment International 25 (4), 393–401.

[5] Tokatlı, C. 2014. Drinking Water Quality of a Rice Land in Turkey by a Statistical and GIS Perspective: İpsala District. Polish Journal of Environmental Studies, 23 (6): 2247-2258.

[6] Li, P., Wu, H. Q., Lyu, X., Liu, H. 2014. Origin and assessment of groundwater pollution and associated health risk: a cese study in an industrial park, northwest China. Environmental Geochemistry and Health, 36:693-712

[7] Neshat, A., Pradhan, B., Javadi, S. 2015. Risk assesment of groundwater pollution using Monte Carlo approach in an agricultural region: An example from Kerman Plain, İran. Computers, Environment and Urban Systems. 50:66-73

[8] http://www.havsa.gov.tr/

[9] Liu, C. W., Lin, K. H., Kuo, Y. M. 2003. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Science of the Total Environment, 313, 77–89.

[10] Tokatlı, C., Çiçek, A., Köse, E. 2013. Groundwater Quality of Türkmen Mountain (Turkey). Polish Journal of Environmental Studies, 22 (4), 1197-1208.

[11] Arslan, H., Turan, A. N. 2015. Estimation of spatial distribution of heavy metals in groundwater using interpolation methods and multivariate statistical techniques; its suitability for drinking and irrigation purposes in the Middle Black Sea Region of Turkey. Envinronmental Monitoring and Assesment, 187:516.

[12] Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A., Shah, A. Q. 2009. Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotoxicol Environ Saf 72:301–309.

[13] Akin, B. S., Atıcı, T., Katircioglu, H., Keskin, F. 2011. Investigation of water quality on Gökçekaya Dam Lake using multivariate statistical analysis, in Eskişehir, Turkey. Environ Earth Sci 63:1251–1261.

[14] Çiçek, A., Bakış, R., Uğurluoğlu, A., Köse, E., Tokatlı, C. 2013. The Effects of Large Borate Deposits on Groundwater Quality of Seydisuyu Basin (Turkey). Polish Journal of Environmental Studies, 22 (4): 1031-1037.

[15] Tokatlı, C. 2013. Use of Statistical Methods in Water Quality Assessment: A Case Study of Balkan Arboretum Area in Trakya University (Edirne, Turkey). Journal of Applied Biological Sciences, 7 (3): 79-83.

[16] Köse, E., Tokatlı, C., Çiçek, A. 2014. Monitoring Stream Water Quality: A Statistical Evaluation. Polish Journal of Environmental Studies, 23 (5): 1637-1647.

[17] Tokatlı, C., Çiçek A., Emiroğlu Ö., Arslan N., Köse E., Dayioğlu H. 2014a. Statistical Approaches to Evaluate the Aquatic Ecosystem Qualities of a Significant Mining Area: Emet Stream Basin (Turkey). Environmental Earth Sciences, 71(5): 2185-2197.

[18] Tokatlı, C., Köse, E., Çiçek, A. 2014b. Assessment of the Effects Of Large Borate Deposits on Surface Water Quality by Multi Statistical Approaches: A Case Study of the Seydisuyu Stream (Turkey). Polish Journal of Environmental Studies, 23 (5): 1741-1751.