Environmental Education and Sustainable Development

Available online at www.ap.ardabil.iau.ir

**Islamic Azad University, Ardabil Branch**

Anthropogenic Pollution Journal, Vol 6 (1), 2022: 1-9

ISSN: 2783-1736- E-ISSN: 2588-4646

I.A.U Islamic Azad University Ardabil Branch

ORIGINAL RESEARCH PAPER

# The use of multivariate statistical methods for the classification of groundwater quality: a case study of aqueducts in the east of Tehran, Iran

**Mohammad Bagher Salarian[1], Mina Mackialeagha[*2], Azita Behbahaninia[3]**

[1] PhD student in Environmental Management, Department of Environmental Sciences and Engineering, Roudehen Branch, Islamic Azad University, Roudehen, Iran

[2] Assistant Professor, Department of Environmental Sciences, Department of Environmental Sciences and Engineering, Roudehen Branch, Islamic Azad University, Roudehen, Iran

[3] Assistant Professor, Department of Environmental Sciences, Department of Environmental Sciences and Engineering, Roudehen Branch, Islamic Azad University, Roudehen, Iran

## ARTICLE INFORMATION

**Received:** 2022.08.15
**Revised:** 2022.09.13
**Accepted:** 2022.10.22
**Published online:** 2022.12.17

**DOI:** 10.22034/AP.2022.1965587.1134

## KEYWORDS

cluster analysis
principal component analysis
water quality classification
pollutant sources
Qanat
Tehran

## ABSTRACT

Natural and human factors have always threatened the health of Qanats, valuable water sources for arid and semi-arid regions. The present study decided to qualitatively classify eight selected Qanats of East Tehran, Iran, using two multivariate statistical methods, cluster analysis (CA) and principal component analysis (PCA) based on parameters including pH, TDS, EC, $Na^{2+}$, $Ca^{2+}$, $Mg^{2+}$, $K^{+}$, $Cl^{-}$, $NO_3$ and $SO4^{2-}$ according to standard methods during the summer of 2020. Data were analyzed by CA and PCA methods, the results of which based on the degree of pollution divided the studied stations into three groups, high pollution (anthropogenic origin), moderate pollution (natural and anthropogenic origin) and low pollution (natural origin). The stations close to each other for quality status were placed in the same group. The eigenvalues obtained from PCA based on the evaluated parameters showed that the first and second components explained more than 58% of changes between the stations. Analyzing the coefficients of each parameter (eigenvectors) for the first and second components revealed that the main causes for the difference between the stations were $Cl^{-}$, $Na^{2+}$, $Mg^{2+}$ and $SO4^{2}$ , TDS and $NO_3$. The two-dimensional display of the stations based on the first two main components confirmed the grouping resulting from the cluster analysis and was able to separate the investigated stations from each other like cluster analysis. The findings of this research highlighted the usefulness and efficiency of two multivariate statistical techniques, CA and PCA, to effectively manage Qanat water quality.

**How to Site:** Salarian MB., Mackialeagha M., Behbahaninia A., The use of multivariate statistical methods for the classification of groundwater quality: a case study of aqueducts in the east of Tehran, Iran, Anthropogenic Pollution Journal, Vol 6 (2), 2021: 1-9, DOI: 10.22034/AP.2022.1965587.1134.

**Corresponding author:** mackialeagha@yahoo.com

## 1. Background

Water pollution caused by chemical, physical and biological pollutants is globally recognized as one of the serious challenges of water resources management (Farhadi et al., 2020; Abbaspour et al., 2013: Hajjabbari and Fataei, 2016). The role of underground water sources in supplying water for consumption in arid and semi-arid parts of the world, including Iran, necessitates the investigation of the risk of groundwater pollution (Majidi et al, 2019; Babaei et al., 2017; Zare Aghbolagh and Fataei, 2016). Qanats have been built as one of the greatest hydrological achievements of the ancient world (Hamidiana et al., 2015) and a water transportation system in arid and semi-arid regions (Radaei et al., 2020). This hydraulic system acts as a long underground network for the exploitation, transmission, maintenance, distribution and supply of drinking water, irrigation and flood control. In addition to natural factors, industrial and domestic sewage pollutants, animals falling into the canals, chemical fertilizers and pesticides also threaten the health of the Qanat (Motiee et al., 2006). The term Qanat is derived from the Semitic word (Semitic race) meaning exploration, digging or water channel, which is generally used in Iran. Qanats have been built thousands of years ago by humans in desert and waterless areas. Qanat is one of the old and at the same time suitable methods of water supply in desert and arid areas. The largest center where Qanat is known as the main source of water supply is located around the deserts of Iran, so that the civilizations related to Qanat were formed mostly in these areas. Except for the narrow strip of the Caspian Sea coast, the territory of Iran has the characteristics of arid and semi-arid areas. Due to the fact that there are no permanent rivers in many areas such as desert ecosystems, and that water sources are often seen seasonally, thus the main water sources in these areas are Qanats. The area studied in the present work, that is, the city of Tehran (the capital of Iran), is one of the areas where there are many Qanats due to its hot and dry climate. Therefore, due to the importance of the Qanat in supplying water to arid and semi-arid regions like Iran, it is necessary to check its quality. Collecting field information for environmental purposes is a complex, costly and time-consuming task (Fataei et al, 2012); accordingly, it should be possible to make maximum use of the available information and data to determine the required parameters (Fataei and Kourandeh, 2013). On the other hand, many experimental sciences face a kind of uncertainty in the scope of their investigation, so nowadays statistics and probability have found a key role in most sciences (Fataei et al., 2010). A branch of statistics called multivariate statistical methods has been developed in recent years, which has wide application fields in different sciences, climate and meteorology, hydrology, and environment (Fataei et al., 2011).

One of the key points in monitoring the quality of water resources systems is the identification and separation of physical and chemical pollutant parameters and their relationship with each other (Kahvaei et al., 2021; Fataei et al., 2012). The importance of this issue stems from the fact that physical pollutants are usually caused by the nature and climate of the basin, but chemical pollutants originate from man-made pollutant sources (Rasi Nezami et al., 2013). Therefore, if the correlation between these two categories of pollutants is high, they can naturally be similar and originate from the same source. On the other hand, finding these connections with laboratory methods involves heavy costs. One of the strategies to reduce costs is to use a multivariate statistical method to identify the main qualitative parameters affecting the quality of water resources (Mirzade Ahari et al., 2019; Fataei and Shiralipoor, 2011).

Principal component analysis (PCA) is widely used technique in the determination of surface and underground water quality (Siyue, 2009). Gurunathan and Ravichandran (1994) applied PCA method to determine the quality of open aquifers in Italy, and found evaporation, irrigation cycle and bedrock type as key variables. Sánchez-Martos et al. (2001) used the PCA technique to determine the effective variables in the analysis of groundwater quality in a number of aquifers in Almería province in Spain. Mohammadzadeh and Heydarizad (2011), by studying the hydrochemistry of underground water resources in Andarkh Karstic region (north of Mashhad, Iran) using PCA, showed that the quality of underground and surface water resources is widely dependent on the lithology of the region. Babaei et al. (2017) by identifying the effective wells in determining the depth of underground water in Urmia plain using PCA, concluded that this method could reduce sampling points and summarize data and that the cost, time and manpower required to measurements and analysis process cut into quarters.

Gangopadhyay et al. (2001) used the PCA technique to identify the importance of monitoring wells for predicting dynamic variables related to the pneumatic level in an area in Bangkok, Thailand. Bhuiyan et al. (2016) assessed the groundwater quality of Lakshimpur district of Bangladesh using 70 groundwater samples collected from shallow to deep wells using water quality indicators, geological methods, and multivariate analysis (PCA).

Ranjan Naik et al. (2021) in evaluating the geochemical quality of groundwater using multivariate water quality index and risk potential in the industrial belt of central Odisha region of India, a total of 106 groundwater samples were analyzed for 14 different parameters. In approximately 30%, 34.9% and 4.7% of the groundwater samples, the concentration of F-, NO3- and uranium exceeded the limit set by WHO, respectively. In addition, they used the groundwater quality index (GWQI), entropy weighted water quality index (EWQI), principal component analysis (PCA), loading based water quality index (PCWQI) and human health risk assessment (HHRA). Depending on the models used, about 19.1 ± 0.9, 70.5 ± 1.9, and 10.38 ± 1.9 percent of water samples were classified as "excellent", "good" and "moderate" using the water quality indicators used in qualitative classification, respectively.

Qanats in the eastern regions of Tehran are among the most important ones of this city, whose quality has been affected by urban development, so that they have been exposed to the threat of environmental pollution in recent years due to neglect in preserving this ancestral achievement and the expansion of urban constructions. Therefore, the management of these Qanats is currently facing serious challenges. Due to the fact that the cleaning of the polluted aquifer requires a very long time and also a huge cost, dealing with the negative environmental consequences of urban Qanats requires qualitative monitoring and classification and determination of the sources of pollutants, which has been investigated in this study through multivariate statistical analysis.

## 2. Materials and Methods

### *2.1. The area under study*

The specifications of the eight studied Qanats located in the eastern area of Tehran, including municipal districts of 1, 4, 8, 13, 14, 15 and 20, are presented in Table 1.

**Table 1.** Specifications of Qanats studied in East Tehran, Iran

| No. of Qanats | Names of Qanats | coordinate (x) | coordinate (Y) | Municipal districts | Purposes of uses |
|---|---|---|---|---|---|
| Qanat 1 | Mobarakabad | 543889 | 3958306 | 4 | Irrigation of green space of Imam Ali |
| Qanat 2 | Majidieh | 542744 | 3955806 | 4 | Irrigation of green space of Majidieh Park and sidewalks around Majidieh Park - Irrigation of small green spaces of houses |
| Qanat 3 | Azizieh | 544189 | 3947963 | 13 | Green space irrigation |
| Qanat 4 | Qeytarieh | 540439 | 3961195 | 1 | Irrigation of Green space of Qeytarieh Park |
| Qanat 5 | Elimon | 538000 | 3941000 | 20 | Green space irrigation |
| Qanat 6 | Mehdiabad | 547868 | 3953764 | 8 | Green space irrigation |
| Qanat 7 | Sulaymaniyah | 543068 | 3948193 | 8 | Green space irrigation |
| Qanat 8 | Aminabad | 543156 | 3931283 | 15 | Green space irrigation |

Tehran is one of the hot and dry regions of Iran. Currently, according to the statistics published by the Environment Unit of Tehran Municipality, out of 500 Qanats in Tehran, about 155 are active, about 150 are semi-active and 195 are abandoned, which have been dug years ago to provide drinking water for humans and livestock and to irrigate agricultural and garden lands (Jafari Aval et al., 2017). Figure 1 shows the location of

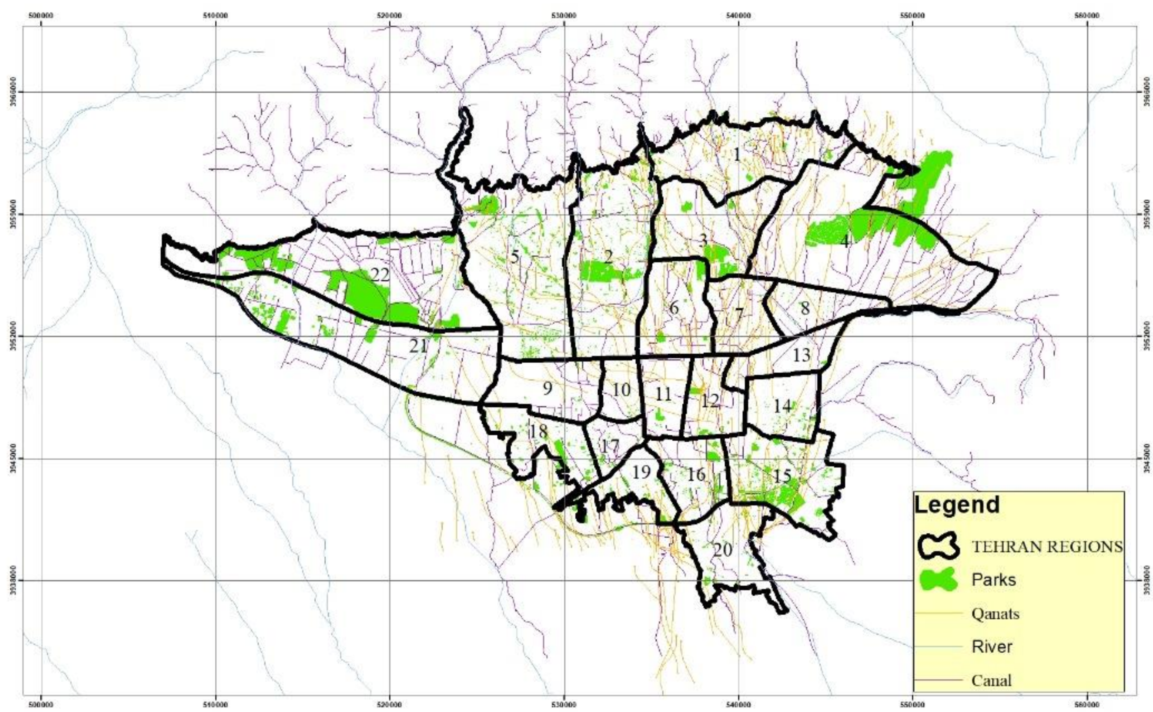


**Figure 1.** Geographical situation of Larestan

the studied stations in the Qanats of the eastern region of Tehran, which were selected based on the location of the pollutant sources.

### 2.2. Sampling method and sample analysis

In this research, samples were taken from eight Qanats in the east of Tehran at three sampling stations along the route of each Qanat (mother well, one of the vertical access shafts and outlet) once in the summer of 2020. The samples were transported to the laboratory under standard conditions and subsequently tested according to the 2017 Standard Methods for the Examination of Water and Wastewater.

Sampling was carried out through glass bottles. The sampling containers were first washed with detergent and tap water, then with nitric acid and finally with distilled water. Before transferring the samples to the laboratory, we performed the preliminary requirements such as fixing the samples with the required chemicals, determining the air and water temperature, and preparing the sample adhesive (including station specifications, sampling time, atmospheric conditions). Following the required preparation, the samples were transferred to the laboratory under suitable temperature conditions and in the shortest possible time and kept in the refrigerator until the subsequent testing.

The samples were analyzed for 10 physical and chemical parameters, including pH, TDS, EC, Na2+, Ca2+, $Mg^{2+}$, $K^+$, $Cl^-$, $NO_3$ and $SO4^{2-}$. Some physicochemical properties of water such as temperature, EC, TDS and pH were measured at the sampling site using a HACK portable device with an accuracy of 0.01 µS/cm for EC and 0.01 for pH. To determine the concentration of the main cations, first the water samples were filtered by a 0.45-µm filter and adjusted to pH below 2 by nitric acid and then measured by ICP-MS method. The main anions of water along with nitrate were measured in the relevant laboratory.

### 2.3. Statistical analysis of data

Data were analyzed by two multivariate statistical methods, cluster analysis (CA) and principal component analysis (PCA), using SPSS 22 and MINITAB 16 software. Kolmogorov-Smirnov (K-S) test was used for normal distribution and goodness of fit of the data. Bartlett's test was performed to test the normality of data in the evaluation of principal components (PCs) and factor analysis. Multivariate statistical methods have been developed in recent years, which aim to reduce the volume of data so that the reduced data includes almost all the informative aspects of the original data. PCA and CA are multivariate statistical methods that can be used to reduce the complexity of data obtained from quality monitoring stations in cases where we are faced with a large amount of information and also for better interpretation of information (Camdevyren, 2005). Therefore, the current research also used these two methods of statistical analysis, as follows:

#### 2.3. 1. Statistical analysis of data

CA is one of the multivariate analysis methods whose main purpose is to group stations based on their characteristics. Thus, the stations with similar characteristics are placed in the same group. According to the predetermined criteria, the grouping will be suitable when the intra-group stations have more homogeneity and the inter-group stations have more heterogeneity. Hierarchical clustering is usually represented by a dendrogram.

Considering that the qualitative grouping of the eight studied qanats based on 10 measured parameters requires the analysis of the correlation matrix of 8*10, and on the other hand, it is necessary to check the quality of the studied qanats by considering all the parameters, so the CA was performed on the basis of distance matrix due to the large volume of data through Minitab16 software. To this end, the analysis was carried out by Ward's minimum-variance method and based on the least-squared Euclidean distance after data standardization.

#### 2.3. 2. Principal component analysis

PCA is one of the multivariate statistical methods that can be used to reduce the complexity of analyzing the original variables of the problem in cases where we are faced with a large amount of information and their better interpretation. Applying this method turns original variables into new and independent components (with a correlation coefficient of zero for both components). The newly created components are linear combinations of the original variables (Manly, 1986). Using these techniques, combinations of p original variables, x4, ..., X2, X1, are constructed to create P independent components (equal to the number of initial variables used), Zp, ..., Z1, Z2. No correlation between these components is a useful feature, because non-correlation means that the components represent different aspects of original variables (Noori et al., 2009). In this method, instead of directly using original variables, first they are converted into new components, and then these components are used instead of original variables. Due to the use of all variables in the formation of components, the information of the original variables is provided by the resulting components with the least error

## 3. Results

### 3.1. Results of cluster analysis

The CA was used to classify the studied Qanats based on qualitative characteristics, so that the Qanats with similar qualitative characteristics were placed in the same group. Cutting the dendrogram based on the furthest Euclidean distance (Laurie et al., 2005) according to the n/2 radical relationship divided the studied stations into three groups. Figure 1 shows the dendogram of the cluster analysis of the studied Qanats based on the measured parameters.

The cluster analysis dendogram (Figure 1) classified the studied Qanats into three groups based on physicochemical characteristics as follows:

The first group (first cluster) contained only Qanat No. 1 (Mobarakabad). The average values obtained for the K, Ca, EC and pH parameters in this cluster were higher than the average of the total measured values.

The second group (the second cluster) included Qanats No. 2, 3, 4 and 5, and the average values obtained for the parameters K, NO3, Cl, TDS and EC in this cluster were higher than the average of the total measured values.

The third group (third cluster) included Qanats No. 6, 7 and 8 and the average values obtained for Na, Mg, SO4 and Cl parameters in this cluster were higher than the average of total measured values, and the average values of the rest of the parameters in this cluster were lower than the total average (Table 2).

The results of one-way ANOVA test confirmed a significant difference between the resulting groups in terms of nitrate and calcium parameters at the level of 1% probability and in terms of chlorine and sodium at the level of 5% probability (Table 2). Examining the difference between the resulting groups indicated that there was no significant difference within each cluster in terms of the evaluated parameters.

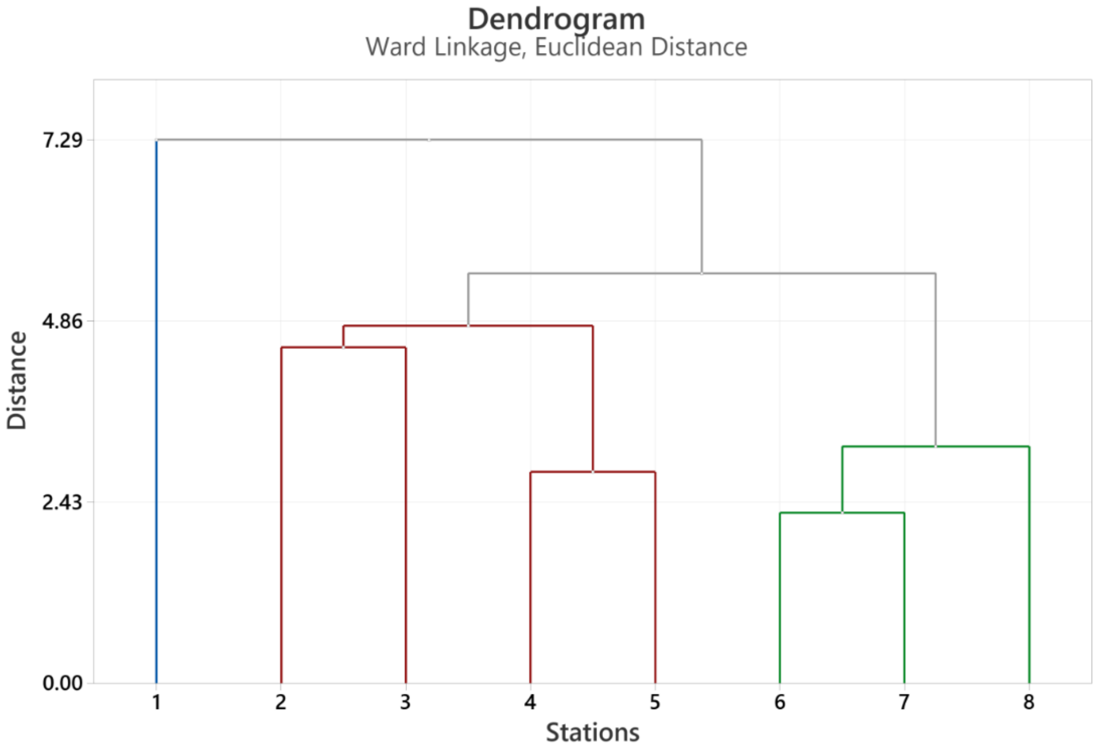


**Figure 1.** Dendrogram derived from the cluster analysis of studied Qanats based on PC1 and PC2 values for the parameters measured in the eastern zone of Tehran, Iran

**Table 2.** Average, deviation from the total average and standard deviation of the averages in three clusters resulting from cluster analysis for the evaluated parameters

| Clusters | Statistical parameters | Na | K | Ca | Mg | SO4 | NO3 | Cl | TDS | Ec | pH |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $\bar{x}$ | 40.00 | 7.50 | 5.00 | 1.80 | 130.00 | 28.00 | 12.00 | 470.00 | 948.00 | 7.76 |
| 2 | $\bar{x}$ | 82.50 | 6.13 | 1.04 | 2.33 | 188.25 | 50.00 | 117.00 | 599.00 | 955.00 | 6.86 |
| 3 | $\bar{x}$ | 187.07 | 5.00 | .66 | 2.67 | 240.00 | 31.33 | 104.67 | 556.00 | 877.33 | 6.93 |
| Mean | Total | 116.63 | 5.88 | 1.39 | 2.39 | 200.38 | 40.25 | 99.25 | 566.75 | 925.00 | 7.00 |
| | One-way ANOVA test results | * | - | ** | - | - | ** | * | - | - | - |

*** and **** = Significance between clusters at the level of 5% and 1% probability, respectively

### *3.2. Results of principal component analysis*

Table 3 presents the results of PCA including the values of eigenvectors, eigenvalues, relative and cumulative variance of the main components of the investigated parameters. The PCA results based on the evaluated parameters showed that the first two components with high eigenvalues explained more than 58% of the changes (Table 3).

Based on the data in Table 2 and Figure 2, $Cl^-$, $Na^{2+}$, $Mg^{2+}$, and $SO4^{2-}$ parameters had the highest positive impact and Ca2+, K+, and pH parameters had the least impact with a negative coefficient on the formation of the first component. In other words, the quality of Qanats No. 5, 7 and 8 were more affected by the concentrations of $Cl^-$, $Na^{2+}$, $Mg^{2+}$and $SO4^{2-}$ parameters. In the formation of the second component, TDS and $NO_3$ parameters had the highest positive coefficient. Therefore, it can be said that the quality of Qanats No. 2 and 4 respectively had the greatest effect from these parameters. The rest of the Qanats (1, 3 and 6) had the same susceptibility in terms of the measured parameters and components, and had an average level of susceptibility in terms of the studied parameters.

**Table 3.** The values of the eigenvectors of the parameters in principal component analysis and rotated factor matrix along with their eigenvalues and their percentage of variance explanation

| Variables | Principal component analysis | | |
|---|---|---|---|
| | PC1 | PC2 | PC3 |
| Na2+ | 0.374 | -0.362 | -0.205 |
| K+ | -0.372 | 0.040 | 0.012 |
| Ca2+ | -0.425 | -0.153 | 0.138 |
| Mg2+ | 0.326 | -0.380 | 0.147 |
| SO42- | 0.266 | -0.146 | -0.364 |
| NO3 | 0.166 | 0.452 | 0.521 |
| Cl- | 0.452 | 0.080 | 0.316 |
| TDS | 0.105 | 0.458 | -0.111 |
| EC | 0.017 | -0.393 | 0.632 |
| pH | -0.368 | -0.321 | 0.039 |
| Eigenvalues | 3.890 | 1.966 | 1.716 |
| Variance percentage | 0.389 | 0.197 | 0.172 |
| Cumulative percentage | 0.389 | 0.586 | 0.757 |

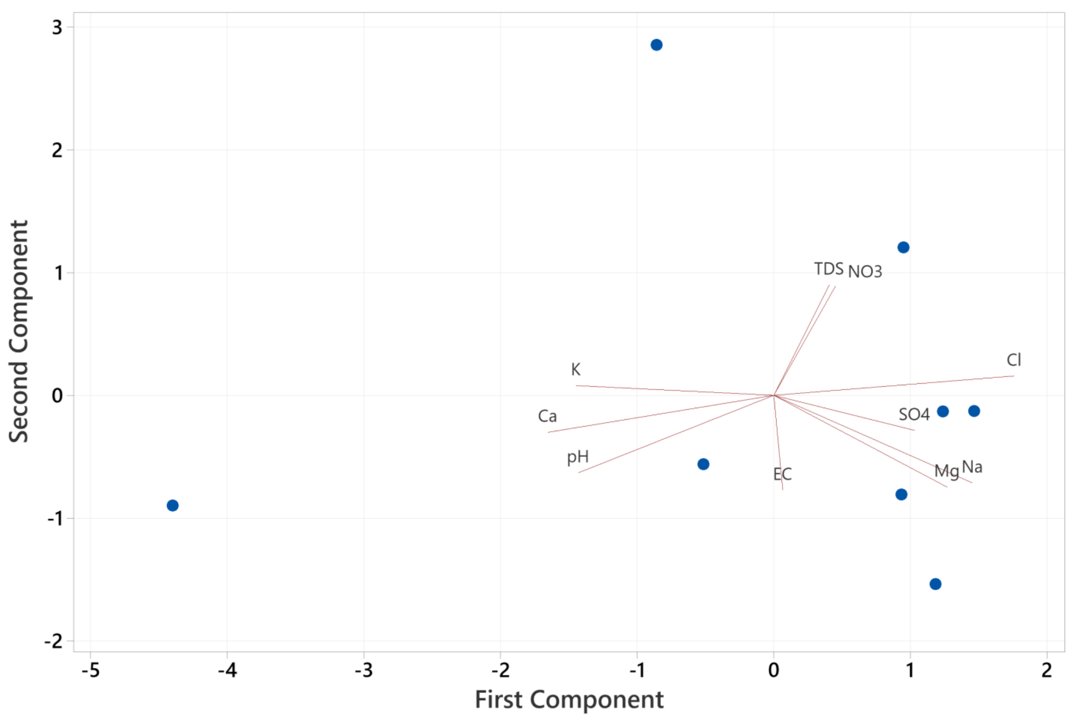


**Figure 2.** Distribution of the effects of different investigated parameters (Biplot) based on the first and second components of principal component analysis on the water quality of eight Qanats in the east of Tehran, Iran

Considering that the Qanats No. 5, 7 and 8 are present in the downstream of the studied area, where the sources of pollutants are concentrated, the highest susceptibility of the quality of these stations was to urban pollutant sources. Accordingly, PC1 can be introduced as the pollution component of anthropogenic activities, that is, any station that has more PC1, its pollution is mainly affected by anthropogenic origin. Figure 3 shows the Biplot of PC1 and PC2 values along with the cluster analysis classification of the stations based on the PC1 and PC2 values in terms of the measured qualitative parameters.

As shown in Figure 3, the results of PCA, like the results of CA (Figure 2), classified the sampling stations into three qualitatively similar groups

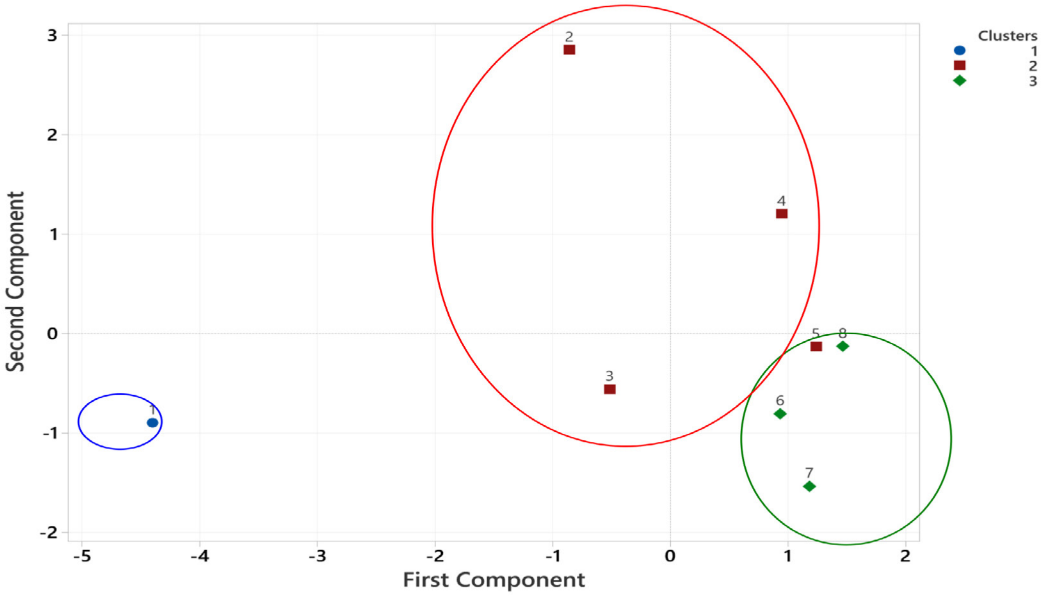


**Figure 3.** Distribution of different investigated stations (Biplot) based on the first and second components of the principal component analysis

## 4. Discussion

The results of the cluster analysis classified the eight studied stations in terms of 10 measured parameters into three qualitative groups, low pollution (cluster 1, natural origin), moderate pollution (cluster 2, natural and anthropogenic origin) and high pollution (cluster 3, anthropogenic origin). In evaluating the water quality of Garasou River in northwestern Iran using cluster analysis, Fataei et al. (2010) classified 11 studied stations using 18 parameters based on the quality status of the stations, dividing them into three groups with high, moderate and low pollution. They mentioned the origin of qualitative changes mainly due to anthropogenic activities, while the cause of qualitative changes between the studied stations in the current study was both natural and anthropogenic sources.

The results of the qualitative classification of the Qanats studied in the eastern area of Tehran showed that the presence of calcium and potassium cations in the Qanat water of the first cluster (Mobarakabad) is of natural origin, so that the significant concentration of calcium in this Qanat could be related to the weathering of the common minerals in most of the rocks of this Qanat and the geological structure. Mohammadzadeh and Heydarizad (2011) determined that the quality of groundwater and surface water sources was predominantly dependent on the lithology of the region. In the second cluster, high levels of NO3 and TDS were mainly related to Qanat No. 2 (Majidieh) due to anthropogenic origin as well as the direction of pollution caused by the effluent of absorbing wells and drainage of fertilizer used in the green space of the area. On the other hand, the presence of potassium and chlorine in the Qanats water of this cluster had a natural origin and geological structure and was due to the weathering of the common minerals in most of the rocks in the region. Sánchez-Martos et al. (2001) investigated the water quality of aquifers in Almería province in Spain and showed that the three variables of sulfate, nitrate and iron had the greatest effect on changing the quality of groundwater with anthropogenic origin, confirming the anthropogenic origin of nitrate in the present study. The reason for the high levels of Cl-, Na2+, Mg2+, and SO42- in Qanat water of the third cluster was of anthropogenic origin, which was related to the introduction of sewage from residential areas, surface runoff, worn texture in the areas, non-observance of proper distance between water wells and sewage wells and urban gardens, and fertilizing agricultural lands with Qanats in these areas.

The PCA results showed that Cl-, Na2+, Mg2+, SO42- , TDS and NO3 had the highest effect on the quality changes of the studied stations, which were caused by natural and anthropogenic sources. Using PCA, Ouyang

(2005) investigated the critical influencing parameters in 22 stations on the lower St. Johns River in Florida, the USA. They found that TOC, DOC, TN, NO3, NO2, orthophosphate, alkalinity, salinity, magnesium and calcium were the most important parameters in the evaluation of river water quality variables, with anthropogenic origin. Bhuiyan et al. (2016) in assessing the groundwater quality of Lakshimpur region of Bangladesh using PCA showed that the groundwater quality in the study area was mainly of geogenic origin followed by anthropogenic source (agrogenic, domestic sewage, etc.). Ranjan Naik et al. (2022) assessed the geochemical quality of groundwater using multivariate water quality index and risk potential in the industrial belt of central Odisha region of India. They found that the groundwater in the study area had different types of water quality issues, including fluoride, nitrate, and uranium pollution of natural and anthropogenic origin.

Babaei et al. (2017) identified effective wells in determining the depth of underground water in Urmia Plain (Iran) using PCA, and announced that by reducing the number of wells based on their importance, through this method, by determining the important wells and reducing the number of sampling wells, it is possible to save a lot on the cost and time of studies. On the other hand, Fataei et al. (2011) in the qualitative evaluation of the Garasou River using multivariate statistical methods showed that important parameters influencing the quality of water resources can be used to reduce the number of sampling stations, without losing much information. In the present study, by specifying six influential parameters, out of the 10 parameters measured in determining the quality of studied Qanats, the measurement of only the parameters of Cl-, Na2+, Mg2+, SO42-, TDS and NO3 can provide information on the quality of the underground water while saving time and cost to monitor the qualitative changes of the Qanats in in the eastern zone of Tehran in the future.

## 5. Conclusions

Water quality management of qanats in the eastern zone of Tehran is of particular importance as one of the groundwater resources. Based on the results of Qanat water analysis and using multivariate statistical methods (cluster analysis and principal component analysis), the studied stations were classified in terms of qualitative status in three groups of low pollution (cluster 1, natural origin), moderate pollution (cluster 2, natural and anthropogenic origin) and high pollution (cluster 3, anthropogenic origin).

Principal component analysis of water parameters based on the correlation coefficient table of 10 × 8 from the average periodic data indicated that the eigenvalues of the parameters with high coefficient in the first three components justified the 75.72% of the changes between the stations. The first and second components, with the eigenvalues of 3.89 and 1.96 respectively, justified a total of 58.6 % of the qualitative changes between the stations. The share of each of the first and second components was 38.9% and 19.6%, respectively. The rest of the components had a minor contribution in justifying the water quality changes among the sampling stations.

In the formation of the first component, Cl-, Na2+, Mg2+ and SO42- parameters had the most positive effect, but in the formation of the second component, TDS and NO3 parameters had the highest positive coefficient. Therefore, the quality of stations 5, 7 and 8 was more affected by the levels of Cl-, Na2+, Mg2+ and SO42- parameters and the quality of stations 1, 3 and 6 was affected by TDS and NO3. The rest of the stations (2 and 4) were affected by the same and on average of the parameters evaluated. Therefore, the anthropogenic activities such as the excessive increase in urban population in the eastern area of Tehran, industrial activities, excessive consumption of chemical fertilizers and pesticides, and wastewater discharge can be introduced as the main factors of water pollution in these Qanats. In addition, the development of garden areas, for example, in forest areas, the decrease of rainfall in recent years and the excessive consumption of water from Qanats for industry and irrigation of parks and gardens can be introduced as other influencing factors in the quality of these Qanats.

The results showed that multivariate statistical techniques (cluster analysis and principal component analysis) can be a suitable method in qualitative assessment and classification and determination of the main factors of Qanats water pollutant. Therefore, the findings of the present research show the usefulness of multivariate statistical techniques in qualitative assessment of underground water resources (Qanat), identifying the sources and polluting factors for effective management of the quality of water resources.

## References

1. Abbaspour M, Javid AH, Jalilzadeh Yengjeh R, Hassani AH, Mostafavi P G, (2013) The biodegradation of methyl tert-butyl ether (MTBE) by indigenous Bacillus cereus strain RJ1 isolated from soil. Petroleum Science and Technology, 31(18): 1835-1841.
2. Babaei A. A, Ghanbari F, Yengejeh, R. J, (2017). Simultaneous use of iron and copper anodes in photoelectro-Fenton process: concurrent removals of dye and cadmium. Water Science and Technology, 75(7), 1732-1742.
3. Babaei S, Hamdami G, Ghasemieh H, (2017) Identify the Effective Wells in Determination of Groundwater Depth in Urmia Plain Using Principle Component Analysis. Journal of Water and Soil, 31(1): 40-50.
4. Bhuiyan MAH, Bodrud-Doza M, Islam ARMT et al (2016) Assessment of groundwater quality of Lakshimpur district of Bangladesh using water quality indices, geostatistical methods, and multivariate analysis. Environ Earth Sci 75, 1020. https://doi.org/10.1007/s12665-016-5823-y
5. Camdevyren H, Demyr N, Kanik A, Keskyn S, (2005) Use of principal component scores in multiple linear regression models for prediction of chlorophyll-a in reservoirs. Ecological Modeling, 181: 581-589.
6. Farhadi H, Fataei E, Kharrat Sadeghi M, (2020) The Relationship Between Nitrate Distribution in Groundwater and Agricultural Landuse (Case study: Ardabil Plain, Iran). Anthropogenic Pollution, 4(1): 50-56.
7. Fataei E, Mosavi S, Imani AA, (2012) Identification of anthropogenic influences on water quality of Aras river by multivariate statistical techniques, 2nd International Conference on Biotechnology and Environment Management IPCBEE, Pukit, Thiland.
8. Fataei E, Monavari SM, Hasani AH, Karbasi AR, Mirbagheri SA, (2010)

Heavy metal and agricultural toxics monitoring in Garasou River in Iran for water quality assessment. Asian Journal of Chemistry, 22 (4): 2991.

9. Fataei E, Monavari SM, Hasani H, Mirbagheri A, Karbasi A, (2011) Surface Water Quality Assessment Using Cluster Analysis. Environmental sciences, 8 (4): 137-146.
10. Fataei E, Shiralipoor S, (2011) Evaluation of surface water quality using cluster analysis: a case study. World Journal of Fish and Marine Sciences, 3 (5): 366-370.
11. Fataei E, Tolou I, Nasehi F, Imani A, (2013) Qualitative classification and determination of Karaj River pollutant sources using multivariate statistical methods. Adv Environ Biol, 7 (11): 3517-3521.
12. Gangopadhyay S, Gupta A, Nachabe MH, (2001) Evaluation of ground water monitoring network by principal component analysis. Ground Water, 39(2): 181-191.
13. Gurunathan K, Ravichandran S, (1994) Analysis of water quality data using a multivariate statistical technique - a case study. IAHS Pub, 219.
14. Hajjabbari S, Fataei E, (2016) Determination cadmium and lead pollution resources of Ardabil Plain underground waters. Open Journal of Ecology, 6(9): 554-561.
15. Hamidian A, Ghorbani M, Abdolshahnejad M, Abdolshahnejad A, (2015) Qanat, Traditional Eco-Technology for Irrigation and Water Management. Agriculture and Agricultural Science Procedia, 4: 119-125. doi: 10.1016/j.aaspro.2015.03.014
16. Jafari Aval Y, Ebadti N, Yousefi H, Kalantari B, Mirzaei M, (2017) Distribution and monitoring of aqueducts water quality for water resources management. Iranian Journal of Ecohydrology, 4(1): 39-52.
17. Kahvaei S, Takdastan A, Jalilzadeh Yengejeh R, (2021) Evaluating the Efficiency of Clarifier Returned Sludge with Poly-aluminum Chloride Coagulant (PAC) for Improving the Removal of Turbidity, COD and PVC. Journal of Health Sciences & Surveillance System, 9(4), 298-304.
18. Majidi H, Mahmoudi Qaraei MH, Minaei M, Malekzadeh Shafaroudi A, (2019) Investigating groundwater nitrate pollution in Qanat hydraulic structure: Jalmabadan and Ramshin villages, north of Sabzevar, 19th Iranian Hydraulic Conference in date 2021-02-14 by Ferdowsi University of Mashhad, Iranian Hydraulic Association in City Mashhad.
19. Manly BF (1986) Multivariate Statistical Methods: A Primer. London, UK: Chapman and Hall.
20. Mirzade Ahari S, Jalilzadeh Yangejeh R, Mahvi A.H, Dadban Shahamat Y, Takdastan A, (2019) A new method for the removal of ammonium from drinking water using hybrid method of modified zeolites/ catalytic ozonation, 170 : 148–157 . doi: 10.5004/dwt.2019.24619
21. Mohammadzadeh H, Heydarizad M, (2011) Hydrochemical and stable isotopes study (O18 and H2 surface and groundwater resources) Andarkh Karstic region (north of Mashhad). Earth science research, 2(5): 59-69. (in Persian).
22. Motiee H, Mcbean E, Semsar A, Gharabaghi B, Ghomashchi V, (2006) Assessment of the contributions of traditional qanats in sustainable water resources management. International Journal of Water Resources Development, 22(4): 575-588.
23. Noori R, Farokhnia A, Riahi Medvar H, (2009) Effects of preprocessing the inlet parameters to artificial neural network (ANN) for predicting the monthly flow by PCA and wavelet. Journal of Water & Wastewater, 20(69): 13-22. [In Persian].
24. Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. Water Res, 39(12): 2621-2635.
25. Radaei M, Salehi E, Faryadi Sh, Masnavi MR, Zebardast L, (2020) Ecological wisdom, a social–ecological approach to environmental planning with an emphasis on water resources: the case of Qanat Hydraulic Structure (QHS) in a desert city of Iran. Environment, Development and Sustainability, 23(7): 10490-10511.
26. Ranjan Naik M, Barik M, Jha V, Sahoo SK, Sahoo NK, (2021) Spatial distribution and probabilistic health risk assessment of fluoride in groundwater of Angul district, Odisha, India. Groundwater for Sustainable Development, 14: 100604.
27. Rasi Nezami S, Nazariha M, Baghvand A, Moridi A, (2013) Karkheh River Water Quality Using Multivariate Statistical Analysis and Qualitative Data Variations. J Health Syst Res, 8(7): 1280-1292.
28. Sánchez-Martos F, Jiménez-Espinosa R, Pulido A, (2001) Mapping groundwater quality variables using PCA and geostatistics: A case study of Bajo Andarax, Southeastern Spain. Environmental Management, 30(5): 716-726.
29. Siyue L, (2009) Water quality in the upper Han River, China: The impacts of land use/land cover in riparian buffer zone. Hazardous Materials, 165(1): 317-324.
30. Toossab Consulting Engineers Company (2011) Study of Qantas water quality of Tehran, Report Number 430752-5139, Tehran, 221.
31. Zare Aghbolagh J, Fataei E, (2016) The study of changes in Ardabil plain groundwater level using GIS , Advances in Science and Technology Research Journal, 10(29): 109-115.
32. Duan, F., Song, F., Chen, S., Khayatnezhad, M. & Ghadimi, N. 2022. Model parameters identification of the PEMFCs using an improved design of Crow Search Algorithm. International Journal of Hydrogen Energy, 47, 33839-33849.
33. Elsayied Abdein, A. A. 2022. The efficiency of Nitrogen utilization and root nodules' life cycle in Alfalfa after various mineral fertilization and cultivation of soil. Water and Environmental Sustainability, 2, 13-20.
34. Guo, H., Gu, W., Khayatnezhad, M. & Ghadimi, N. 2022. Parameter extraction of the SOFC mathematical model based on fractional order version of dragonfly algorithm. International Journal of Hydrogen Energy, 47, 24059-24068.
35. Lin, H. 2022. Levafix blue color's visible light degradation utilizing Fenton and photo-Fenton procedures. Water and Environmental Sustainability, 2, 1-8.
36. Mobar, S. & Bhatnagar, P. 2022. ling women by Greenhouse plan as illustrated in the Post-Feminist Tamil Film 36 Vayadhinile. Water and Environmental Sustainability, 2, 9-12.
37. Zhang, J., Khayatnezhad, M. & Ghadimi, N. 2022. Optimal model evaluation of the proton-exchange membrane fuel cells based on deep learning and modified African Vulture Optimization Algorithm. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44, 287-305.