Multivariate statistic for surface water quality assessment: case study of Zohreh river, Hendijan, Iran

Jalal Valiallahi 1\* , Saideh Khaffaf Roudy2

1 Associate professor of Environmental Science Department, Shahid Rajae Teacher Training University,

Lavizan, Tehran, Iran. Postal code:1678815811

Dean of Scientific Association of Environmental Education and Sustainable Development, (EESD) [iraneesd@yahoo.com](mailto:iraneesd@yahoo.com)

2 Shahid Rajae Teacher Training University Faculty, University. Lvizan Tehran

1. Coresponding authour Email: [jvaliallahi@yahoo.com](mailto:jvaliallahi@yahoo.com)

*Abstract*—Evaluation of spatial variations and interpretation of Zohrehh River water quality data were made by using multivariate analytical techniques including factor analysis and cluster analysis also the Arc GIS® software. The research method was formulated to achieve objectives herein, including field observation, numerical modeling, and laboratory analyses. The results showed that dataset consisted of 11,250 observations of seven-year monitoring program (measurement of 15 variables at 3 main stations from April 2010 to March 2017). Factor analysis with principal component analysis extraction of the dataset yielded seven varactors contributing to 82% of total variance and evaluated the incidence of each varactor on the total variance. The results of cluster analysis became complete with t-test and made water quality comparison between two clusters possible. Results of factor analysis were employed to facilitate t-test analysis. The t-test revealed the significant difference in a confidence interval of 95% between the mean of calculated varactors 1, 2, 6 and 7 between two clusters, but there was no significant difference in the mean of other varactors 3, 4 and 5 between two groups. The result shows the effect of agricultural fertilizers on stations located at downstream of the ASK dam.

Keywords— Surface water quality · Multivariate statistical techniques · Factor analysis · Cluster analysis · Zohreh river

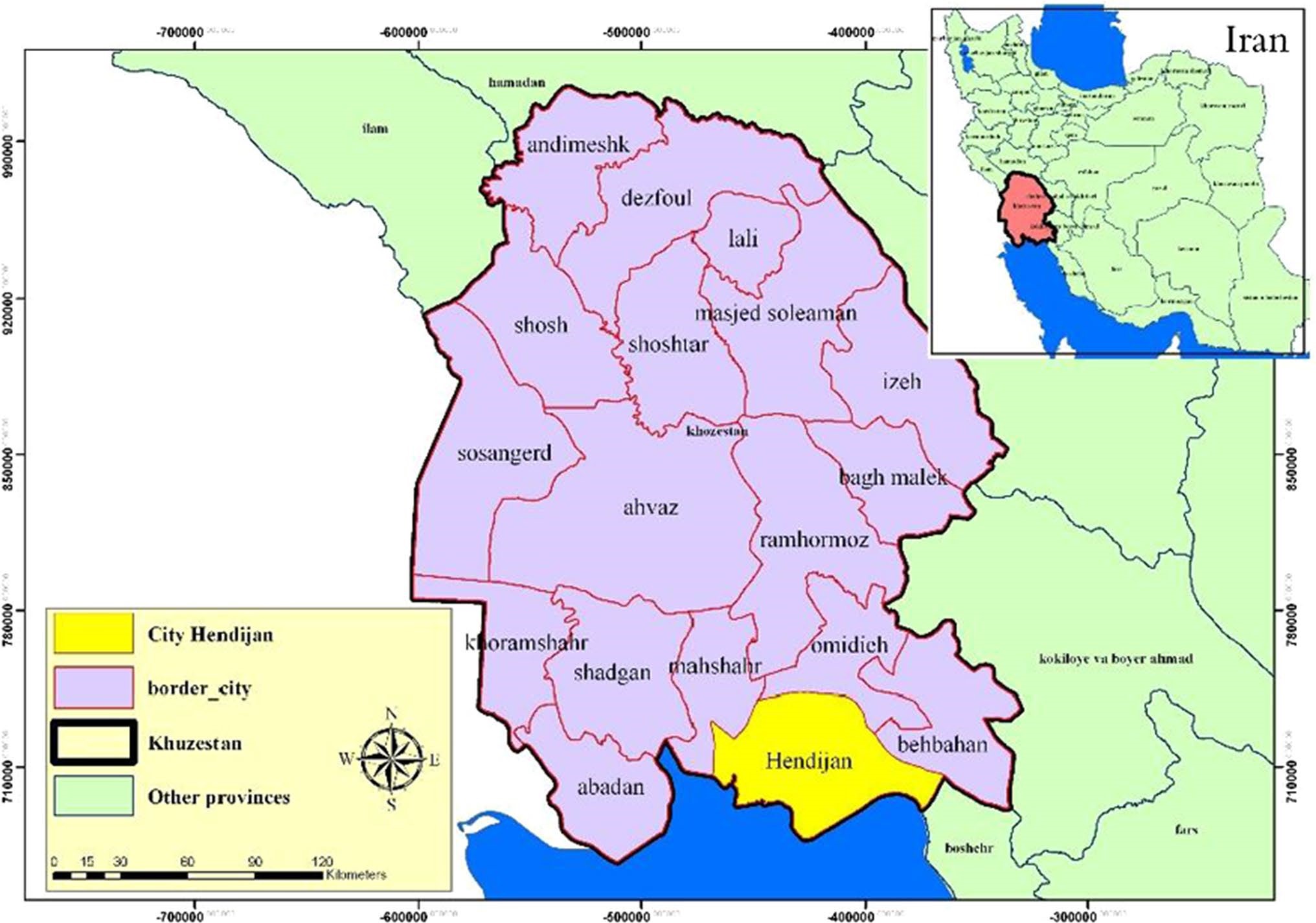
1. **Introduction**

Freshwater resources are vulnerable and have the potential to be strongly impacted by climate change, with wide-ranging consequences for human societies and ecosystems (Bates et al. 2008). The Zohreh river basin is one of the important arias that nowadays is affected by anthropogenic and climate change.

“Classification, modeling, and interpretation of monitored data are the most important steps in the assessment of water quality” (Iscen et al. 2008). “Water quality monitoring has one of the highest priorities in surface water protection policy. Multivariate statistical methods have been used successfully in hydrochemistry for many years. Using multivariate statistical methods help to better understand water quality in the interpretation of complex datasets (Zhang et al. 2009). "Multivariate statistical methods are widely used in various environmental measurements.

They have been presented as appropriate tools in water quality assessment, identification of pollution sources/factors and understanding temporal/spatial variations in water quality for effective river water quality management" (Singh et al. 2005). They may be particularly useful when there is a large volume of experimental results and sometimes they provide insight into the multidimensional patterns in the data that would be overlooked with uni variety analyses" (Shtangeeva et al. 2009). The multivariate methods applied in water quality analyses are cluster analysis (CA), discriminant analysis (DA), factor analysis (FA), principal component analysis (PCA), and exploratory factor analysis (EFA).

"Factor analysis was particularly useful for considering several related random environmental variables simultaneously, and so identification of a new, smaller set of uncorrelated variables that accounted for a large proportion of the total variance in the original variables"(Wang et al. 2007). Varimax factor rotation method rotates the axis such that the two vertices remain 90 degrees (perpendicular) to each other and assumes uncorrelated factors (also referred to as orthogonal rotation). “Cluster analysis (first used by Tryon 1939) classifies objects so that each object can be similar to the others in the cluster with respect to a predetermined selection criterion. The clusters of objects must be homogeneous (within the cluster) and not very solid (between clusters) "(Varol and Sen 2008).

The goal is to compare the data of a similar (or related) group to one another and different from (or unrelated) data in other groups, the greater the similarity (or homogeneity) within a group, and the greater the difference between groups, the better or more distinct the clustering. The component analysis and exploratory factor analysis are a multivariate statistical technique used to identify the important components or factors that explain more of the variance of the system. To reduce the number of variables to a small number of indicators (that is, the main components or factors are designed to reduce the relationships existing in the original data "(Ouyang 2005). PCA is a declining variable technique and is used when the variables overlap strongly. This method reduces the number of observed variables to a number of major components, which is more than the variance of observed variables. The total value of the variance in the PCA is equal to the number of variables examined. In PCA, the observed variables are standardized, for example, the mean = 0, standard deviation = 1, the matrix diameter is 1. The value of the described variance is equal to the matrix trace (sum of the matrices decomposed correlation diameters). The number of extracted components is equal to the number of variables observed in the analysis. The first major component identified is the most variance between data. The second component is identified for the second level of variance in the data and does not match the first major component, and so on. The accounting components are kept for the maximum variance, while other accounting components are not kept for the importance of the variance value. "Specific values represent the amount of variance explained by each component. The special features of the weights are used to calculate the components of the components." (Diana 1999).

“Surface water is influenced by natural processes (precipitation, erosion, weathering of crustal materials) and anthropogenic effects (urban, industrial, and agricultural activities and increasing consumption of water resources)” (Zhang et al. 2008). “Anthropogenic effects have the vast affinity to hasten natural phenomes that affect water quality” (Yidanaa et al. 2008). “Water quality denotes to the physical, biological and chemical aspects of water” (Wanga et al. 2005).

“Discharges from municipals and industries are considered as a point and constant polluting source, while surface runoff is a seasonal phenomenon and non-point source due to its characteristics that are highly influenced by climate and seasonal changes” (Han et al. 2009; Shrestha and Kazama 2007).

Vol.:(0123456789)1 3

High temperatures and extreme rainfall, including floods and droughts, affect the quality of water and change many forms of water pollution, including sediment, nutrients, and organic soluble carbon in water (Bates et al.

2008).

1. **Materials and methods**

The purpose of this study is to accurately assess water quality status in the Zohreh river (Figs. 1, 2). To archiving this purpose, a multivariate statistical technique, PCA and hierarchical cluster analysis, factor analysis (FA) have been used to assess the spatial variability of water quality parameters.

In order to obtain the information required about Zohreh river, three main sampling stations were selected. Selected stations are Chambostan, Dehmolla, Shiri, respectively.

The data for water quality of these stations from 2010 to 2017 for the various parameters of water quality dataset were obtained from Khuzestan Province Regional Water Agency. This agency collects surface and groundwater samples monthly and analyzes them using standard methods for electrical conductivity (EC), total dissolved solids (TDS), pH, total hardness (TH), major cations as calcium (Ca2+), magnesium (Mg2+), sodium (Na+) and potassium(K+) and major anions as chloride (Cl–) and sulphate (SO42–), bicarbonate (HCO3–) temperature (T) in addition to sampling locations latitude and longitude. The data were extracted from the monitoring program conducted by Khuzestan water and wastewater company and include samplings along the Zohreh river during a seven years’ period (from April 2010 to March 2017); this program provided a dataset of 3 main stations (Figs. 1, 2) in which the samplings were carried out monthly.

Dataset consisted of 11,250 observations of seven-year monitoring program (measurement of 15 variables at 3 main stations from April 2010 to March 2017). Among these data, descriptive statistics analyzing carried out on 6 water quality parameter TDS, EC, pH, Hco3, Ca and Mg, among a total number of 729 selected sample from three main stations in Zohreh river.

Data from laboratory results were interred to the statistical software SPSS version 23 and also were transferred to Arc GIS ver. 10. Information provided including lab results and statistical calculation was transferred to the map layer.

|  |
| --- |
| **Fig. 2** Dem of Hendijan city and three main stations, from right to left, Cham Bostan, Shiri and Dehmola along the Zohreh river |

1. **Tables and graphs**

The present study was conducted over a relatively long range of Zohreh river about 275 km over a seven-year long period of time and using qualitative parameters values at 3 stations.

1. **Factor analysis**

The results of Table 1. indicate the highest value of EC (2570) and lowest values (101.80) also descriptive statistics analyzing of 6 water quality parameter TDS, EC, pH, Hco3, Ca and Mg, from a total number of 729 sample at three main stations in Zohreh river (Tables 2, 3).

For involving PCA method, a combination matrix was offered (Table 4). Components that covered a small part of the changes in dataset were excluded. Finally, 2 main factors, the main variables in each component were extracted. Also, share obtained was presented in the Table 5.

At the next step, the correlation coefficients between the factors specified in Table 1 were applied to all data in all station. Correlation matrix variables were examined and showed a strong correlation between the variables. In this

****

FIGURE 3 – Dendogram using average linkage between group.

Table 1. Descriptive Statistics of total number of 729 water sample analyzing of 6 water quality parameter TDS, EC, pH, Hco3, Ca and Mg, in Zohreh river.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | N | Minimum | Maximum | Mean | Std. Deviation |
| TDS | 729 | 2.09 | 1670.50 | 215.1837 | 92.90130 |
| EC | 729 | 101.80 | 2570.00 | 337.3466 | 143.75352 |
| PH | 729 | 5.51 | 8.60 | 7.8391 | .31378 |
| Hco3 | 729 | .50 | 5.77 | 2.4611 | .52871 |
| Ca | 729 | 3.06 | 30.00 | 8.2146 | 3.13451 |
| Mg | 729 | .80 | 43.24 | 3.6202 | 2.41219 |
| Valid N (listwise) | 729 |  |  |  |  |

Table 2. Correlation matrix between 6 main water quality parameter in Zohereh river, high relationships between factors are shown by two stars.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Correlations** | | TDS | Ec | PH | Hco3 | Ca | mg |
| TDS | Pearson Correlation | 1 |  |  |  |  |  |
| Sig. (2-tailed) |  |  |  |  |  |  |
| N | 729 |  |  |  |  |  |
| EC | Pearson Correlation | .979\*\* | 1 |  |  |  |  |
| Sig. (2-tailed) | .000 |  |  |  |  |  |
| N | 729 | 729 |  |  |  |  |
| pH | Pearson Correlation | -.149\*\* | -.134\*\* | 1 |  |  |  |
| Sig. (2-tailed) | .000 | .000 |  |  |  |  |
| N | 729 | 729 | 729 |  |  |  |
| Hco3 | Pearson Correlation | .268\*\* | .261\*\* | -.249\*\* | 1 |  |  |
| Sig. (2-tailed) | .000 | .000 | .000 |  |  |  |
| N | 729 | 729 | 729 | 729 |  |  |
| Ca | Pearson Correlation | .469\*\* | .418\*\* | -.284\*\* | .226\*\* | 1 |  |
| Sig. (2-tailed) | .000 | .000 | .000 | .000 |  |  |
| N | 729 | 729 | 729 | 729 | 729 |  |
| Mg | Pearson Correlation | .777\*\* | .783\*\* | -.138\*\* | .274\*\* | .433\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 | .000 | .000 | .000 |  |
| N | 729 | 729 | 729 | 729 | 729 | 729 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | | | | |

Table 3 KMO test result.

|  |  |  |
| --- | --- | --- |
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .724 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 3412.427 |
| df | 15 |
| Sig. | .000 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4. Component Matrix **a** Extraction Method: Principal Component Analysis. a two component extracted | | |  | Table 5. Extraction Method: Principal Component Analysis | | |  | Table 6 Rotated Component Matrix a | | |
|  | Component | | Communalities | | |  | Component | |
| 1 | 2 |  | Initial | Extraction |  | 1 | 2 |
| TDS | .937 | .238 | TDS | 1.000 | .934 | TDS | .956 | .144 |
| EC | .926 | .265 | Ec | 1.000 | .928 | EC | .956 | .114 |
| pH | -.309 | .792 | pH | 1.000 | .723 | pH | .022 | -.850 |
| Hco3 | .437 | -.521 | Hco3 | 1.000 | .462 | Hco3 | .201 | .650 |
| Ca | .637 | -.262 | Ca | 1.000 | .474 | Ca | .486 | .488 |
| Mg | .868 | .197 | mg | 1.000 | .791 | Mg | .876 | .154 |

| Table 7. Proximity Matrix | | | | | | |  | Table 8 - The parameters in the two clusters Membership. | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Case | Squared Euclidean Distance | | | | | | Case | 2 clusters |
| 1: TDS | 2: EC | 3: pH | 4: Hco3 | 5: Ca | 6: Mg | 1: TDS  2: EC  3: pH  4: HCO3  5: Ca  6: Mg | 1  1  2  2  2  2 |
| 1: TDS | .000 | 279880.240 | 773868.980 | 799599.543 | 756388.214 | 782119.235 |
| 2: EC | 279880.240 | .000 | 1981669.860 | 2022720.617 | 1953738.782 | 1994818.333 |
| 3: pH | 773868.980 | 1981669.860 | .000 | 223.059 | 157.552 | 42.811 |
| 4: Hco3 | 799599.543 | 2022720.617 | 223.059 | .000 | 641.462 | 120.667 |
| 5: Ca | 756388.214 | 1953738.782 | 157.552 | 641.462 | .000 | 271.445 |
| 6: Mg | 782119.235 | 1994818.333 | 42.811 | 120.667 | 271.445 | .000 |
| This is a dissimilarity matrix | | | | | | |

way, Table 2 was generated. The correlation that results are summarized in Table 2 (mathematical and statistical computations). As showed in Table 2, high relationships between factors are indicated by two stars, so we can say that correlation between these parameters has correlated at 99% confidence.

In the third step, rotated component matrix was applied, and the results are shown in Table 6. To extract these components from within the main parameters of water quality, from time to time varimax rotated factor was used. Accordingly, each component parameter which the highest load factor (positive or negative) that was the best representative of that component is described. Based on PCA, the data were classified into 2 groups.

The result of Table 6 shows all the variables fall into two groups and variables TDS, EC and mg in one group of variables pH, Hco3 and Ca are other groups.

1. **Cluster analysis**

The hierarchical cluster was used to implement this method. Initially, a proximity matrix was obtained according to Table 7.

The parameters in the two clusters are shown in Table 8.

According to this table and graph 4, TDS and EC at the bottom are in the first cluster parameters and parameters of PH, Hco3, Ca and Mg were placed in the second cluster.

And finally, dendrogram charts presented in Fig. 3.

1. **Results**

Factor analysis with principal component analysis extraction (PCA) method was investigated. PCA of the dataset resulted in seven principal components (PCs) with eigenvalues > 1 accounting for almost 82% of total variance; the higher the eigenvalue, the more significant the corresponding component (Tables 2, 3).

Based on the results, all measured values for the Zohreh river indicate that the river’s water at upstream can be used for drinking only by applying advanced refining.

In fact, over the course of this seven-year period, the decline in water quality, especially at the Shiri Station, has been highlighted.

In general, by evaluating of water quality trends in time series at the measuring stations, we find that river water and its pollutants are somehow balanced. In other words, the volume of input pollutants has been proportional to river self-purification capacity.

Meanwhile, the water quality of the river varies in the various watery and various seasons.

The EC has gradually reached its maximum levels in different years at all three stations.

Which is higher at Shiri station compared to other stations. The reason for the high rate of this variable at this station can be attributed to the ASC dam construction, which is due to trapping of water during summer and autumn and entering of wastewater into Zohreh river.

The reason for the increase in pollution at the Shiri Station during all studied periods was the increase in consumption of pesticides and fertilizers in the area and accordingly increasing in releasing of a combination of agricultural and industrial wastewater and urban sewage, as a consequence the nutrients, carrying from surface runoff and depositing materials downstream will be increased. Due to frequent droughts and reducing of water flow from water resources, at downstream of this station volume of pollution will be increased and overflow into the station, so the rate of pollution in this area is more intense.

In addition, due to the high temperature at this station in all years and due to the inverse relationship between dissolved oxygen and temperature, the Shiri Station has the lowest amount of dissolved oxygen in all studied years in this research. Obviously, this increase in temperature and the reduction of oxygen will trigger changes in natural and living organisms, start changes in the color of the water, reduce reproduction in animal groups and reduce the ecosystem balance in this area. In addition, the reference to this issue is necessary, unfortunately, the construction of two dam (ASC and Kowsar) will damage the health of the aquatic environment in threat.

According to the results obtained during the 7 years and at most stations, winter season water has better quality due to more precipitation, thus reducing pollutants vice versa at summer season water quality is poorer due to reducing of rainfall.

1. **Discussion**

In this case study, multivariate techniques factor analysis, cluster analysis, and t-test were investigated to determine sources of pollution, the correlation between variables, and the effect of Zohreh dam (Ask Dam) on the water quality.

Major contamination of the Zohreh river is related to agricultural wastes entering the Zohreh river. The river is also of great importance to agriculture, which encounters a variety of natural and artificial sources of pollution along its pathway.

Other sources of water pollution of the Zohreh river can be noted urban wastewater near the sampling sources. Increasing urbanization and industrial and technological development have not only increased urban water consumption, but also increased environmental pollution. In particular, the problem of water pollution has become more complex during the forthcoming time and the increasing use of various chemicals such as detergents, insecticides, herbicides, various chemical fertilizers, etc., and the drainage of wastewater from local factories mixing with irrigation water.

The health, local welfare and social issues associated with the passageway of this important waterway, for example, the air and water health in the city of Hendijan significantly depends on the clarity and health of Zoherh river.

In the past, because the extension and density of incoming waste were not exceeded from the self-restorative and self-purifying proses of water; these issues were not problematic, but nowadays, due to decrease in water flow, the river self-restoring has fewer effects on water quality so river is become contaminated for about at last 6 month of year. The droughts and dehydration have plunged the banks of the Zohreh river. The waters of the Zohre river have already collapsed sharply so that in some places it has been inhabited by coarse fish and insects (Valiallahi Jalal, Khaffaf Roudy Saideh, 2021).

1. **Conclusion**

This conclusion showed that many of the water quality parameters at the studied stations in the summer are due to evaporation, rainfall shortage, and the total decrease in discharge the river has a worse condition. Also showed that seasons and weather conditions have a great impact on water quality.

We concluded that damming and climate changes have drastic effects on vegetation cover, the forests located on the margins of the main rivers including Zohreh river and extreme effects on waterway system and wetlands.

It turns out that rising water crisis is due to climate change in the Zohreh river is more than twice as high as water consumption. In years when runoff is reduced, there is a significant reduction in the quality of water resources. It seems that changes in quantity and quality of water due to climate change will reduce food security and increase the vulnerability of poor rural farmers, especially in the arid areas of the Zohre River Basin.

Iranians have a brilliant history in water systems management, no doubt, they will pay more attention to the environment and will find numerous solution for this problem with help of adjacent country and all of the authorities and scientist in water planning sector. Water can be managed with changes in our attitude with water, and living style, our cultivation procedures, such as reforming the crop pattern, farmland quota, and supplying water to the industry and agriculture from recycling and reusing of water and genuine in retention of water from precipitation.

1. **Financial Disclosure statement**

This paper is part of Saideh khaffaf Roudy theses for revising M.Sc. in Environmental Science, in Environmental Science Department at Shahied Rajaee Teacher Training University, and is part of plan entitled “Water crises in Iran” is carrying out by first author: Jalal Valiallahi in this University and the authors received no specific funding for above mentioned theses.

1. **Compliance with ethical standards**

**Conflict of interest** For this article, there is no conflict of ethical standards.

1. **References**

Bates BC, Kundzewicz ZW, Wu S, Palutikof JP (eds) (2008) Climate change and water. Technical paper of the intergovernmental panel on climate change, IPCC secretariat, Geneva, 210 pp. 2008, Intergovernmental panel on climate change, ISBN: 978–92–9169–123–4.

Diana S (1999) Principal component analysis vs. exploratory factor analysis*.* University of Northern Colorado, SUGI 30, Paper 203. [http:// www2. sas. com/ proce edings/ sugi30/ 203- 30. pdf](http://www2.sas.com/proceedings/sugi30/203-30.pdf)

Han S, Kim E, Kim S (2009) The water quality management in the Nakdong river watershed using multivariate statistical techniques. KSCE J Civ Eng 13(2):97–105. [https:// doi. org/ 10. 1007/](https://doi.org/10.1007/s12205-009-0097-5) [s12205- 009- 0097-5](https://doi.org/10.1007/s12205-009-0097-5)

Iscen CF, Emiroglu Ö, Ilhan S et al (2008) Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environ Monit Assess 144:269–276

Manshouri M, Ahmadisharaf E, Alamdari N (2011) Qualitative study of Zohreh river water using QUAL2K model. Conference paper · August 2011, Doi: [https:// doi. org/ 10. 13140/ RG.2. 1. 2032. 3043](https://doi.org/10.13140/RG.2.1.2032.3043), [https:// www. resea rchga te. net/ publi cation/ 23709 8981](https://www.researchgate.net/publication/237098981)

Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. Water Res 39(12):2621–2635

Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin. Japan. Environ Model Softw 22(4):464–75.

Shtangeeva I, Alber D, Bukalis G et al (2009) Multivariate statistical analysis of nutrients and trace elements in plants and soil from northwestern Russia. Plant Soil 322:219–228

Singh K, Malika A, Sinha S (2005) Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. Anal Chim Acta

538:355–374

Varol M, Sen B (2009) Assessment of surface water quality using multivariate statistical techniques: a case study of Behrimaz Stream. Turkey. Environ Monit Assess 159(1–4):543-53. [https:// doi. org/](https://doi.org/10.1007/s10661-008-0650-6) [10. 1007/ s10661- 008- 0650-6](https://doi.org/10.1007/s10661-008-0650-6)

Wang X, Lu Y, He G et al (2007) Multivariate analysis of interactions between phytoplankton biomass and environmental variablesin

Taihu Lake, China. Environ Monit Assess 133:243–253

Wanga X, Homerb M, Dyerb SD et al (2005) A river water qualitymodel integrated with a web-based geographic information system. J Environ Manag 75(3):219–228

Valiallahi Jalal · Khaffaf Roudy Saideh, 2021, Application of multivariate statistical techniques for investigating climate change and anthropogenic effects on surface water quality assessment: case study of Zohreh river, Hendijan, Iran. Applied Water Science (2021) 11:99 https://doi.org/10.1007/s13201-021-01399-0

Yidanaa SM, Ophoria D, Banoeng-Yakubo B (2008) A multivariatestatistical analysis of surface water chemistry data—the Ankobra Basin, Ghana. J Environ Manag 86(1):80–87

Zhang Q, Li Z, Zeng G et al (2009) Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang watershed. China. Environ Monit Assess. 152:123. [https:// doi. org/ 10. 1007/ s10661- 008- 0301-y](https://doi.org/10.1007/s10661-008-0301-y)

Zhang Y, Guo F, Meng W et al (2009) Water quality assessment and source identification of Daliao river basin using multivariate statistical methods. Environ Monit Assess 152:105–121

آمار چند متغیره برای ارزیابی کیفیت آب

مطالعه موردی: رودخانه زهره هندیجان خوزستان

جلال ولی الهی \*1، سعیده خفاف رودی 2

1 دانشیار گروه علوم محیط زیست دانشگاه تربیت دبیر شهید رجایی ، پست الکترونیکی :

رئیس انجمن علمی آموزش محیط زیست و توسعه پایدار ایران

2**فارغ التحصیل رشته علوم محیط زیست دانشگاه تربیت دبیر شهید رجایی لوزان تهران** [saideh\_khaffaf@yhoo.com](mailto:saideh_khaffaf@yhoo.com)

\*نویسنده مسئول: [iraneesd@yahoo.com](mailto:iraneesd@yahoo.com) [jvaliallahi@srttu.edu](mailto:jvaliallahi@srttu.edu)

چکیده:

ارزیابی تغییرات مکانی و تفسیر داده‌های کیفی آب رودخانه زهره با استفاده از تکنیک‌های تحلیلی چند متغیره شامل تحلیل عاملی و تحلیل خوشه‌ای و همچنین نرم‌افزار Arc GIS® انجام شد. روش تحقیق برای دستیابی به اهدافی از جمله مشاهده میدانی، مدلسازی عددی و تجزیه و تحلیل آزمایشگاهی فرموله شده است. نتایج نشان داد که مجموعه داده شامل 11250 مشاهده برنامه پایش هفت ساله (اندازه گیری 15 متغیر در 3 ایستگاه اصلی از آوریل 2010 تا مارس 2017) است. تجزیه و تحلیل عاملی با استخراج تجزیه و تحلیل مؤلفه های اصلی مجموعه داده، هفت وارکتور را به دست آورد که در 82 درصد از واریانس کل مشارکت داشتند و میزان بروز هر وراکتور را بر واریانس کل ارزیابی کرد. نتایج تجزیه خوشه ای با آزمون t کامل شد و مقایسه کیفیت آب بین دو خوشه را امکان پذیر کرد. از نتایج تحلیل عاملی برای تسهیل تحلیل t-test استفاده شد. آزمون t تفاوت معنی داری را در فاصله اطمینان 95 درصدی بین میانگین واراکتورهای 1، 2، 6 و 7 محاسبه شده بین دو خوشه نشان داد، اما در میانگین سایر وارکتورهای 3، 4 و 5 تفاوت معنی داری وجود نداشت. دو گروه نتیجه نشان دهنده تأثیر کودهای کشاورزی بر ایستگاه های واقع در پایین دست سد ASK است. 

کلمات کلیدی— کیفیت آب سطحی · تکنیک های آماری چند متغیره · تحلیل عاملی · تحلیل خوشه ای · رودخانه زهره