Alharbawee and Mohammed *Iraqi Journal of Science, 2024, Vol. 65, No.3, pp: 1266-1275*
DOI: 10.24996/ijs.2024.65.3.8

ISSN: 0067-2904

Water Quality Assessment of Tigris River Using Multivariate Statistical Techniques

Noor A.H. Alharbawee *, Ahmed Jasim Mohammed
Department of Biology, College of Science, University of Baghdad, Baghdad, Iraq

---

Received: 5/12/2022 Accepted: 23/3/2023 Published: 30/3/2024

ABSTRACT

Factor and cluster analysis, two multivariate statistical approaches, have been used to analyze temporal and spatial variations and pollutants sources effecting the water quality of Tigris River from the north to the south of Baghdad city. From October 2021 to September 2022, 15 water quality parameters (electrical conductivity (EC), pH, water temperature (WT), turbidity (Tur), total hardness (TH), biochemical Ooxygen Ddemand (BOD), dissolved oxygen (DO), Mg+, Ca+, SO4, dissolved solids (TDS), total suspended solids (TSS), total chlorine (TC), Cr, and Al) were tested at five monitoring sites (Balad, Al-Ghrai'at, Al-Shuhada'a Bridge, Al-Za'franiya, and Jisr Diyala), along the study area of the river. Results revealed that the major factor that affects the water quality in the study area is the anthropogenic factor. Spatial cluster analysis of the monitoring sites revealed three different groups that were compared to one another where each site suggested a different degree of contamination in the Tigris river water quality.

Keywords: Multivariate analysis techniques, Tigris River, Factor Analysis, Cluster Analysis, Water Quality.

تقييم جودة مياه نهر دجلة باستخدام تقنيات التحليل الاحصائي متعددالمتغيرات

نور عبدالرزاق حسن الحرباوي*, احمد جاسم محمد

قسم علوم الحياة , كلية العلوم, جامعة بغداد, بغداد , العراق

الخلاصة

تم استخدام تحليل العوامل والتحليل العنقودي ، وهما نهجان إحصائيان متعددا المتغيرات لتحليل التباين الزماني والمكاني وتأثير مصادر الملوثات على جودة مياه نهر دجلة من شمال إلى جنوب مدينة بغداد. من أكتوبر 2021 إلى سبتمبر 2022 ، 15 معلمة لجودة المياه (التوصيلية الكهربائية ، درجة الحموضة ، درجة حرارة الماء ، التعكر ، العسرة الكلية ، المتطلب الحيوي للاوكسجين ، الاوكسجين المذاب ، ايونات المغنسيوم ، ايونات الكالسيوم ، الكبريتات ، إجمالي المواد المذابة ، مجموع المواد الصلبة العالقة ، إجمالي الكلور ، تركيز الكروم ، تركيز الالمنيوم) في خمسة مواقع للرصد (بلد ، الكَريعات ، جسر الشهداء ، الزعفرانية ، وجسر ديالى) (على امتداد منطقة الدراسة على النهر). كشفت النتائج أن العامل الرئيسي الذي يؤثر على جودة المياه في منطقة الدراسة هو العوامل البشرية ، وقد كشف التحليل العنقودي المكاني لمواقع المراقبة عن ثلاث مجموعات مختلفة كانت قابلة للمقارنة مع بعضها البعض ، كل منها يشير إلى درجة مختلفة من التلوث في جودة مياه نهر دجلة.

---

*Email: nalharbawee@gmail.com

1. Introduction

The problem of pollution of the various components of the environment has aggravated as a result of neglect, corruption and wars that have spanned for decades, resulting in various types of solid, liquid and gaseous pollutants of chemical, biological and nuclear nature by getting accumulated in the Iraqi environment. Most of them have entered various ecosystems threatening their sustainability and the safety of their resources [1].The water resources in Iraq suffer from many different factors resulting due to external factors in the water policies and the sources of the Tigris and Euphrates rivers and the dams on them and their tributaries without taking into account Iraq's water needs as well as the internal factors that are present by throwing industrial, agricultural, and municipal waste into the rivers [2], [3]. In addition to the lack of application of laws and legislations, preventing water pollution and preserving the ecosystem of the rivers healthy have also contributed to the problem [4]. Tigris River flows through the major cities of the country with Baghdad being one of the cities that splits it into sides (Al-Ressaffa and Al-Karkh) [5] which has the highest population and industrial density in the country. Due to the lack of environmental monitoring and majority of the pollutants thrown into the river, there has been continuous deterioration in the environmental ecosystem. The assessment of Tigris River water quality became necessary to be investigated to ensure its suitability for human uses and the aquatic environment. Assessment of water quality is done by using valuable techniques such as water quality indexes where water quality data is summarized for reporting to the public in a consistent such as the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI), Biological Water Quality Index (BWQI) etc. [6], [7] and multivariate statistical techniques [8], [9].

Using various multivariate statistical techniques like factor (FA) and cluster analysis (CA) implemented to facilitate the explanation of complex data matrices to best assessment for ground and surface water quality and ecosystem status studies, as well as, these techniques enable the identification of potential factors or sources that affect the quality of water systems, and provide useful tools for the reliable management of water resources and a quick resolution of pollution issues [10].

In hydrochemistry, Factor Analysis (FA) and Correspondence Analysis (CA) have been successfully utilized for assessing quality of surface water and environmental studies. Using these methods is well-presented in the literature. Group of Iranian researchers studied the water quality of Gorgan Rod River. From 1984 through 2008, water samples collected from 18 sampling locations along the river were evaluated by using these techniques. The results of the findings showed that PCA (principal component analysis) linked anthropogenic and natural changes in the various stations; and that the locations of the various stations under study were compatible with their outward appearances. In addition, due to the anthropogenic effects, water quality in some stations was among the worst. However, no changes were recorded in other stations because there were no changes to the land uses [11]. Ismail and Robescu in 2019 adopted multivariate statistical methods to discover the main factors causing pollution in the water quality in Danube River. Fourteen parameters were assessed to monitor spatial and temporal similarities and differences among the sampling locations for one year. The results showed three significant distinctions found were industrial waste discharge, home wastewater and agricultural activity in agricultural areas. As well as the sampling sites' differences were revealed by hierarchical spatial CA as two distinct groups, each indicating a different degree of river water contamination [12]. In a study at Damascus University, a five-year data, from 2015 to 2019, examined the Orontes River's water quality using multivariate statistical approaches. Cluster analysis was utilized in order to recognize the pollution areas along the river. Five main factors were found to have been responsible for explaining 92.86% of the total variance. It included 7 basic parameters: Electrical conductivity (EC), turbidity (TUR), NO_3, Na, pH, NH_4, chemical oxygen demand(COD) that were recorded by using factor analysis (13). A separate study presented validity of utilizing of FA and CA methods to assess the Tigris River's water quality within Baghdad City from February 2017 to February 2018. The findings demonstrated that changes in all parameters throughout both wet and dry seasons resulted in a decline in water quality [14].

The current study aimed to employ same multivariate analysis techniques for determining the main parameters that affect the water quality and the most significant variables that cause the difference in water quality of the Tigris River in the extension of the study area.

2. Materials and Methods:

2.1 Study Area

One of the longest rivers in the Middle East, the Tigris River, has a catchment area of 235,000 km^2 and a length of about 1,900 km, passing through Iraqi land for nearly 1415 km. The river shares this role with the Euphrates River as a primary source of water for human consumption, particularly for water supply since they pass through the main cities in the region [4]. The Tigris River flows from the Toros Mountains in southeast Turkey through Syria, Iraq and Turkey. This river is the principal source of water supply for Baghdad, the capital of Iraq. It has numerous tributaries, including the Botmanse, Kessora, A1-Khabur, the Greater Zabs, Lesser Zabs, the Al-Adhaim and Diyala Rivers (14). From Al-Tarmiyahm District in the north to Al-Madaain District in the south, the Baghdad stretch of the Tigris River is located in the Mesopotamian alluvial plain with geographical coordinates 33°14'-33°25' N, 44°31'-44°17' E and an altitude between 30.5 and 34.85 m above sea level (Figure 1). From the north to the south flow, the river splits the city of Baghdad into two sections: the right side (Al-Karkh) and the left side (Al-Risafa). The region has a dry, hot and semi-arid environment with cold winters and a mean annual rainfall of about 151.8 mm [15]. More than eight million residents live in Baghdad which is regarded as the most populous and developed metropolis in Iraq. Most industrial and municipal waste is dumped into the stream without being properly treated

2.2 Sampling and Analysis

Five locations were selected for sampling on the reach of the Tigris river, namely: Balad (S1) which is in the north of Baghdad; Al-Ghrai'at (S2) which is 116 kilometers south of S1 along the river's length; Al-Shuhada'a Bridge (S3) which is 13.5 km to the south from S2; Za'franiya (S4) which is about 27 km south from S3; and Jisr Diyala (S5) which is nearly 7 kilometers south of S4 along the Tigris reach in southern region of Baghdad (Figure 2). Data for five water quality monitoring locations for fifteen water quality parameters, was collected by using polypropylene bottles at seasonal intervals from sampling locations from October 2021 to September 2022. For several water quality parameters analysis, sampling was used following the recommended standard methods [18]. Bottles were utilized to collect and examine water quality parameters, and for biochemical oxygen demand (BOD) analysis of the samples to prevent unanticipated changes. Water samples analysis begun as soon as was feasible following collection. Samples of microorganisms were collected in sterile, opaque glass containers. Within around 24 hours, the bottles were examined while being stored at 4°C. The analysis of the collected samples were done in the biological and chemical laboratory , Department of Water Resource Techniques, biological and chemical laboratory, Institute of Technology-Baghdad Middle Technical University, and in thecentral laboratory, Department of Biology College of Science, University of Baghdad.

The tested parameters included were EC, water (pH), water temperature(WT), turbidity, total hardness (TH), (BOD, dissolved oxygen (DO), magnesium ion (Mg+), calcium ion (Ca+), sulphate (SO_4), total dissolved solids (TDS), total suspended solids (TSS), total chlorine (TC), chromium (Cr) and aluminum (Al).

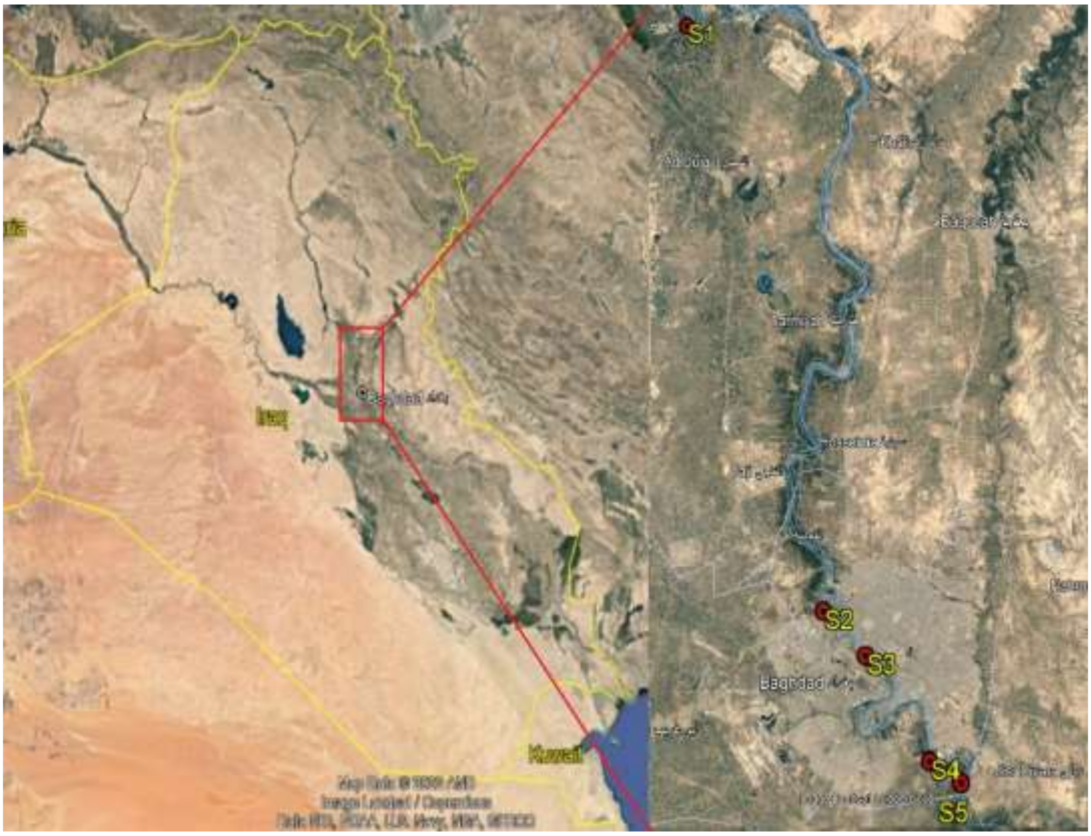

Figure 1: Map of the study area and sampling locations [32].

2.3 Data Treatment and Multivariate Analysis

Prior to adopting multivariate statistical techniques, Shapiro-Wilk (W) test was used to verify that each variable distribution was normal [17]. Water quality variables must fit the normal distribution in order to be used in factor and cluster analysis. All of the variables, with the exception of WT, turbidity, DO, and TSS, were found to be normally distributed using the Shapiro-Wilk (W) test. As a result, the initial non-normal distribution variable data was transformed using the formula $x' = \log 10(x)$ [18]. After log transformation and the Shapiro-Wilk (W) test, the extra variables DO, WT and TSS were taken out, leaving only the normalized turbidity.

The water quality Tigris River was studied using FA that is used to define the primary contributing factors and the most significant variables. All variables were passed in three main stages, first was correlation matrices produced, then an initial set of factors was extracted using the PCA approach, and lastly the factors that were extracted were subsequently rotated using the Varimax rotation [14]. The fundamental concept behind factor analysis can be stated as follows:

$$z_j = a_{j1} f_1 + a_{j2} f_2 + \ldots + a_j f_m + e_{ij};\ j = 1,2,\ldots,p \qquad (1)$$

Where Z: value that is being measured, f: the score of factor, a: the loading of factor, e: the residual term accounting for errors or other sources of variation, i: number of samples, j: number of variables, m: total number of factors.

CA was similarly utilized to categorize objects with comparable qualities, which divides large amounts of objects into lesser number of homogenous clusters depending on the high similarity [19]. The most popular method used in this study was hierarchical agglomerate grouping using Ward's algorithm [20] which establishes similarity correlations between any one sample and the complete dataset and is often depicted by a dendrogram (tree diagram). One of the most widely used measures of distance is the Euclidean distance approach. The multidimensional space's geometric separation was calculated using the formula [21].

$$\text{Distance } (X, Y) = \left\{ \sum i (Xi - Yi)^2 \right\}^{1/2} \qquad (2)$$

IBM SPSS25 software was used to analyze the statistical data, FA and CA.

4. Results and Discussion

Five stations in the research regions were used to monitor the water quality over the course of the period (Oct 2021 to Sep 2022) (Figure 1). pH, EC, BOD5, TH, TBR, SO_4, Mg+, Ca+, TDS, TSS, TC, Cr and Al) parameters were chosen to assess surface water quality characteristics. Table 1 provides a descriptive statistical summary of the Tigris River's measured water quality data during the period of study.

4.1 Factor Analysis (FA)

Using IBM SPSS 25, factor analysis was conducted on all 15 parameters of water quality at all monitoring stations through the body of Tigris River during the study period [24], [25]. Centroid approach was used to create the correlation matrix of variables, and factors were then removed and rotated using Varimax rotation [26]. The factor with the highest eigenvalue is the most significant as it provides a measure of the significance for the component. Significant eigenvalues are those with values of 1.0 or above [27]. As shown in the Figure 2, first four factors were retrieved and the the rest of the factors were deleted after interpreting the factor loadings. This indicated that the first four factors accounted for most of the total variation of the original data. Then, factor rotation (Varimax) was utilized to create factor loadings that were simple to understand [28]. Table 2 displays the percentage of total variance that can be explained by the first four factors.

Table 1: Summary of the descriptive statistical data on Tigris River water quality in the area of study and Iraqi and world standards [22], [23]

Parameters	Iraqi Standard (2009)	WHO Standards (2011)	Minimum	Maximum	Mean	Std. Deviation
pH	6.5-8.5	6.5-8.5	6.54	8.85	7.82	0.73
WT	-	-	12.60	32.50	20.49	6.96
EC	1500	1500	407.00	1012.00	755.50	183.98
Turbidity	5	5	6.35	96.20	34.48	28.53
TH	500	500	210.00	490.00	374.50	81.52
DO	>5	4	2.00	8.50	6.31	1.90
BOD	5	<5	1.00	6.00	3.00	1.62
Mg+	50	150	7.20	81.60	33.12	16.85
Ca+	50	200	44.00	136.00	94.60	24.42
SO4	400	250	13.44	65.28	31.58	12.03
TDS	500	1000	137.00	384.00	258.90	71.21
TSS	120	25	200.00	1000.00	490.00	255.26
T. chloride	250	250	177.00	540.00	336.04	91.57
Cr	0.05	0.05	0.02	0.04	0.02	0.01
Al	0.2	0.2	2.75	7.17	4.78	1.44

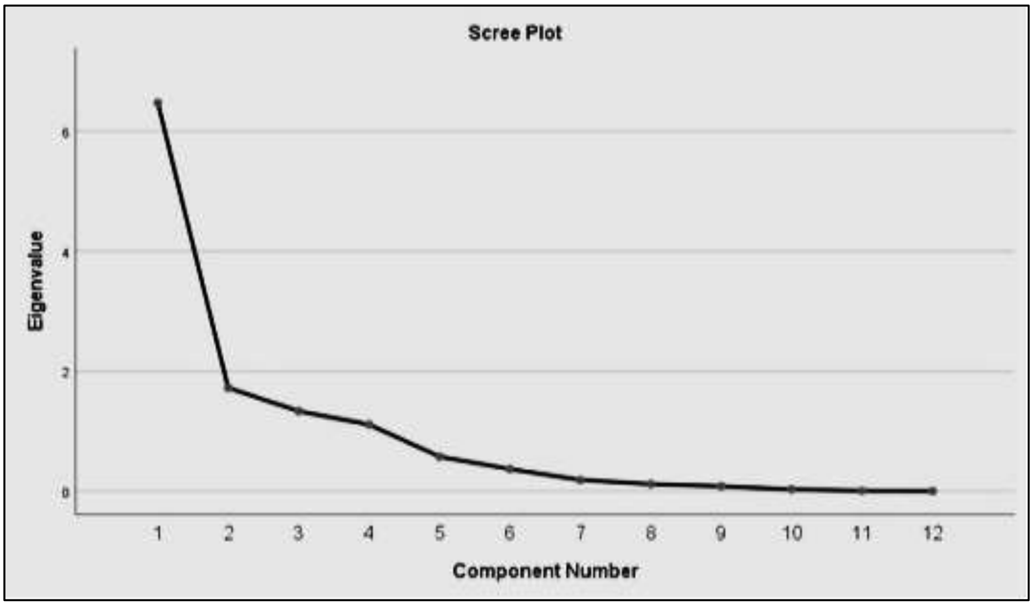

Figure 2: Scree plot of eigenvalues versus components for the observed water quality

Table 2: Explanation of total variance before and after Varimax rotation

Total Variance Explained

Component	**Initial Eigenvalues**			**Extraction Sums of Squared Loadings**			**Rotation Sums of Squared Loadings**		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	6.475	53.960	53.960	6.475	53.960	53.960	4.827	40.228	40.228
2	1.721	14.341	68.301	1.721	14.341	68.301	2.253	18.776	59.005
3	1.333	11.111	79.412	1.333	11.111	79.412	2.112	17.600	76.605
4	1.109	9.245	88.657	1.109	9.245	88.657	1.446	12.052	88.657
5	.574	4.784	93.441						
6	.367	3.061	96.502						
7	.185	1.544	98.046						
8	.116	.968	99.014						
9	.083	.688	99.701						
10	.030	.251	99.952						
11	.006	.048	100.000						
12	-4.510E-17	-3.759E-16	100.000						

Extraction Method: Principal Component Analysis.

It is apparent that the first, second, third and fourth components, in that sequence, represented 40.228%, 18.776%, 17.600%, and 12.052% of the total variance of the data on the observed water quality. The first four components represented about 88.656% of the overall variation, but the following eight only account for 11.344%. Table 3 shows the factor loadings for the first four components of the factor analysis of the observed water quality data.

The first factor (F1) had positively significant loads for EC, TH, BOD, Ca+, SO_4, TDS, and Cr; low positive loading for Mg+ and Al; and negative loading for turbidity. It also explained 40.228% of the total variance. Anthropogenic pressures included domestic wastewater discharge and industrial activities that transport households and industrial effluents of the city, which might have contributed to many factors. The latter can be connected to distinct sources based on the existence of various ingredients in the factors extracted. The first factor (F1) was influenced by sources that included point-source pollutants from household and industrial trash. The presence of heavy metals in home and industrial waste can be seen from the greater Ca+ loadings in factor 1. This element also consistently demonstrated that the TDS inside this river was primarily caused by the TH. For obvious reasons, the connection to EC is well known. T. chloride had strongly positive loads of F1 (0.828). According to Arzu *et. al.*, the rubbish from cities raises the amount of feces inside the water, suggesting that F1 may have originated from a city.

The second factor, F2, had modest positive loading from other factors and substantial negative loading from turbidity which accounted for 18.776% of the total variation. This factor represented the total TSS in the river water. The third factor (F3) showed high positive loadings on Mg+ and Cr gradually and explained 17.600% of the overall variation. The fourth factor (F4) exhibited high positive loadings on Al and SO_4 gradually, accounting for the least amount of the total variance (12.052%). Agriculture, municipal wastewater, the erosion effect of soil cultivation and high rains from upland areas are all related to this component [16].

The results of the factor analysis led to the conclusion that there were four components, each of which represented a different process: anthropogenic, erosional, rainfall, and TSS factors.

Table 3: Factor loading matrix and total variance explained after Varimax rotation.

Parameters	Factor 1	Factor 2	Factor 3	Factor 4
EC	0.858	0.363	0.235	0.139
TH	0.734	0.186	0.595	0.202
BOD	0.728	-0.380	0.210	0.256
Mg+	0.066	0.058	0.966	0.180
Ca+	0.904	0.182	-0.316	0.062
SO4	0.607	0.254	0.201	0.565
TDS	0.871	0.372	0.249	0.083
T. chloride	0.828	0.270	0.398	0.161
Cr	0.532	0.049	0.595	-0.349
Al	0.141	0.132	0.060	0.882
Turbidity	-0.249	-0.875	-0.068	-0.160

4.2 Cluster Analysis (CA)

In this research, to identify the similarities between the monitoring sites, sampling site categorization was carried out using cluster analysis (squared Euclidean distance as measure of similarity and Ward's method of linking). Dendrogram below displays the cluster analysis (CA) results. In cluster analysis, dendograms are a helpful graphical tool to calculate the number of clusters that describe the underlying mechanisms causing regional differences [29].

CA outputs have been revealed as dendograms for every season below. Figure 4 shows the spatial clustering of monitoring sites during autumn season. It can be seen that there are two main groups agglomerated (Group A and Group B). Group A includes stations Balad, Al-Ghrai'at and Al-Shuhada'a Bridge whereas Group B includes Al-Za'franiya and Jisr Diyala stations. In this season, 2 levels of pollution were identified in which stations located south of Baghdad such as Al-Za'franiya and Jisr Diyala reflected the highly polluted sites. Other stations reflected the moderate pollution.

In winter season, 2 groups were agglomerated (Figure 5). Group A included 4 stations namely Al-Ghrai'at and Al-Shuhada'a Bridge, Al-Za'franiya and Jisr Diyala. While group B included Balad station only. Figure 6 shows the spatial clustering of monitoring sites during spring season. In this season, 2 agglomerated groups were A and B. Group A further included 2 sub-groups: A1 and A2. CA outputs of summer and winter seasons were very close (Figure 7). Balad station was agglomerated in group B reflecting relatively good water quality whereas group A agglomerating sites were polluted.

Generally, it can be concluded that three levels of pollution were seen from CA. Balad referred to a less polluted site, Al-Ghrai'at and Al-Shuhada'a Bridge revealed to be moderately polluted sites, and Al-Za'franiya and Jisr Diyala represented highly polluted sites. These stations are located south of Baghdad and the river in this area receives pollution load from wastewater treatment plant, as well as a group of industries also disposing off their wastewater without adequate treatment [30], [31].

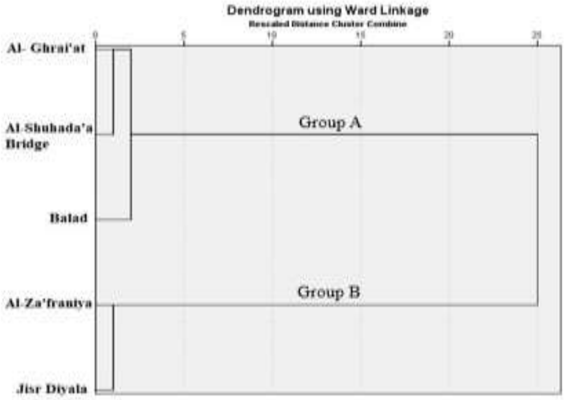

Figure 4: Dendrogram showing spatial clustering of monitoring sites during autumn season**.**

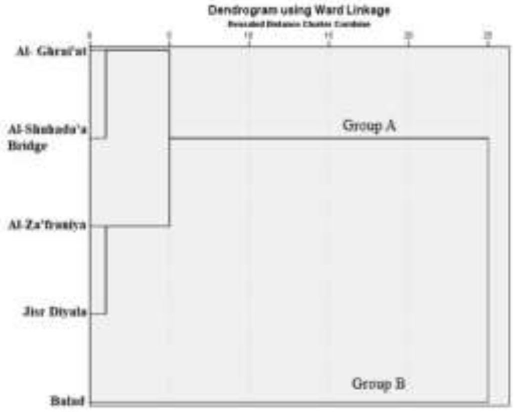

Figure 5: Dendrogram showing spatial clustering of monitoring sites during winter season.

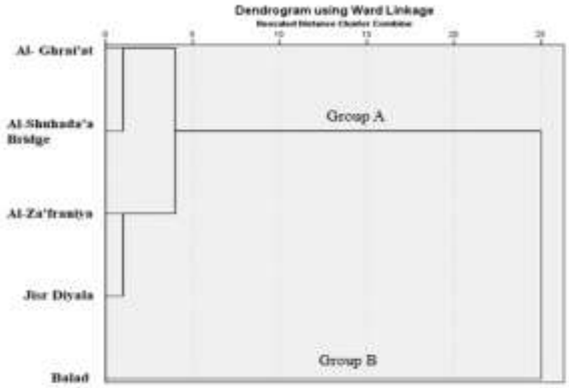

Figure 6: Dendogram showing spatial clustering of monitoring sites during spring season.

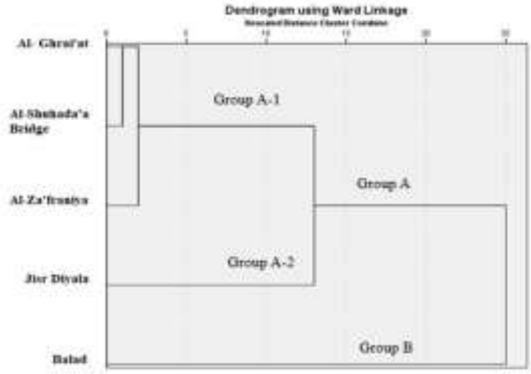

Figure 7: Dendrogram showing spatial clustering of monitoring sites during summer season.

5. Conclusion

Factor and cluster analysis were applied on a water quality dataset of Tigris River for a period of 1 year. Four variables were found to be responsible for the data structure by FA accounting for 88.656% of the dataset's overall variation. Nearly 40%, 19%, 18% and 12% of the total variations were explained by the first, second, third and fourth factors respectively. Anthropogenic, erosion, rainfall and TSS variables were found to be the primary determinants of the river water quality. Spatial CA identified three statistically significant clusters among the five riverside monitoring locations. Al-Ghrai'at and Al-Shuhada'a bridges were examples of generally moderately contaminated places, whereas Balad Station turned out to be an example of a comparatively less polluted site. The third group (Za'franiya and Jisr Diyala) relates to locations that were found to be comparatively heavily polluted. The use of FA and CA approaches are useful tools for managing water resources effectively and explained complicated datasets of the studied river water quality.

References

[1] M. Al-Taay, M. S. A., Al-Assie, A. H. A., & Rasheed, R. O., “Impact of bazian cement factory on air, water, soil, and some green plants in Sulaimani City-Iraq”, *The Iraqi Journal of Agricultural Science*, vol.49, no. 3, pp. 463-477, 2018.

[2] N. Al-Ansari, "Water Resources of Iraq," *Journal of Earth Sciences and Geotechnical Engineering*, vol. 11, pp. 15-34, 2021.

[3] M. Ibrahim, M. Mohammed-Ridha, H. Hussein, and A. Faisal, "Artificial neural network modeling of the water quality index for the Euphrates River in Iraq," *The Iraqi Journal of Agricultural Science*, vol. 51, pp. 1572-1580, 2020.

[4] Price, R.A., “Environmental risks in Iraq, K4D Helpdesk Report”, Brighton, UK: Institute of Development Studies, 2018.

[5] J. Rzóska, *Euphrates and Tigris, Mesopotamian ecology and destiny* vol. 38: Springer Science & Business Media, 2012.

[6] K. Yürekli, M. Erdoğan, A. H. Ismail, and M. A. Shareef, "How far can hydrochemical characteristics of surface water meet drinking and irrigation criteria: a pragmatic study for the Euphrates River basin, Turkey," *Sustainable Water Resources Management*, vol. 7, pp. 1-11, 2021.

[7] Q. Y. Al-Kubaisi and S. A. Khorsheed, "hydrochemistry Evaluation of Groundwater Suitability for consumption in Yaychi area (Southwest Kirkuk city, North Iraq)," *Iraqi Journal of Science*, pp. 119-134, 2018.

[8] A. H. Ismail and D. Robescu, "Assessment of water quality of the danube river using water quality indices technique," *Environmental Engineering and Management Journal*, vol. 18, pp. 1727-1737, 2019.

[9] M. Alhadithi, "Evaluation of groundwater quality using water quality index (WQI) and GIS techniques," *Iraqi Journal of Agricultural Sciences*, vol. 49, 2018.

[10] M. Varol and B. Şen, "Assessment of surface water quality using multivariate statistical techniques: a case study of Behrimaz Stream, Turkey," *Environmental monitoring and assessment*, vol. 159, pp. 543-553, 2009.

[11] Y. Khaledian, S. Ebrahimi, U. Natesan, N. Basatnia, B. B. Nejad, H. Bagmohammadi, *et al.*, "Assessment of water quality using multivariate statistical analysis in the Gharaso River, Northern Iran," in *Urban ecology, water quality and climate change*, ed: Springer, 2018, pp. 227-253.

[12] A. H. Ismail and D. Robescu, "Application of multivariate statistical techniques in water quality assessment of Danube River, Romania," *Environ. Eng. Manag. J*, vol. 18, pp. 719-726, 2019.

[13] L. Khouri and M. B. Al-Mufti, "Assessment of surface water quality using statistical analysis methods: Orontes River (Case study)," *Baghdad Science Journal*, 2022.

[14] R. Al-Ani, A. Al Obaidy, and F. Hassan, "Multivariate analysis for evaluation the water quality of Tigris River within Baghdad City in Iraq," *The Iraqi Journal of Agricultural Science*, vol. 50, pp. 331-342, 2019.

[15] A. S. Al-Adili and S. M. Ali, "Hydrochemical Evolution of Shallow Ground Water System Within Baghdad City–Iraq," *The Iraqi Geological Journal*, pp. 134-139, 2005.

[16] A. H. Ismail, B. S. Abed, and S. Abdul-Qader, "Application of multivariate statistical techniques in the surface water quality assessment of Tigris River at Baghdad stretch," *Journal of Babylon University,* vol. 22, pp. 450-462, 2014.

[17] Al-Adili, A. S., “Geotechnical evaluation of Baghdad soil subsidence and their treatments” (Doctoral dissertation, University of Baghdad),1998.

[18] W. E. Federation and A. Association, "Standard methods for the examination of water and wastewater," *American Public Health Association (APHA): Washington, DC, USA,* vol. 21, 2005.

[19] S. K. Chapagain, V. P. Pandey, S. Shrestha, T. Nakamura, and F. Kazama, "Assessment of deep groundwater quality in Kathmandu Valley using multivariate statistical techniques," *Water, Air, & Soil Pollution,* vol. 210, pp. 277-288, 2010.

[20] K. L. de Oliveira, R. L. Ramos, S. C. Oliveira, and C. Christofaro, "Spatial variability of surface water quality in a large Brazilian semiarid reservoir and its main tributaries," *Environmental Monitoring and Assessment,* vol. 193, pp. 1-15, 2021.

[21] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to data mining*: Pearson Education India, 2016.

[22] I. S. IQS, "Iraqi standard of drinking water No. 417," *Second modification,* 2009.

[23] WHO, "Guidelines for drinking-water quality," *World health organization,* vol. 216, pp. 303-304, 2011.

[24] F. Murtagh and P. Legendre, "Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion", *Journal of classification*, vol. 31, pp. 274-295, 2014.

[25] U. K. Singh, M. Kumar, R. Chauhan, P. K. Jha, A. Ramanathan, and V. Subramanian, "Assessment of the impact of landfill on groundwater quality: a case study of the Pirana site in western India," *Environmental monitoring and assessment,* vol. 141, pp. 309-321, 2008.

[26] U. C. Panda, S. K. Sundaray, P. Rath, B. B. Nayak, and D. Bhatta, "Application of factor and cluster analysis for characterization of river and estuarine water systems–a case study: Mahanadi River (India)," *Journal of hydrology,* vol. 331, pp. 434-445, 2006.

[27] S. Shrestha and F. Kazama, "Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji River basin, Japan," *Environmental Modelling & Software,* vol. 22, pp. 464-475, 2007.

[28] S. M. Ahmed, M. Hussain, and W. Abderrahman, "Using multivariate factor analysis to assess surface/logged water quality and source of contamination at a large irrigation project at Al-Fadhli, Eastern Province, Saudi Arabia," *Bulletin of Engineering Geology and the Environment,* vol. 64, pp. 315-323, 2005.

[29] J.-O. Kim, O. Ahtola, P. E. Spector, J.-O. Kim, and C. W. Mueller, *Introduction to factor analysis: What it is and how to do it*: Sage, 1978.

[30] A. H. Ismail and A. Muntasir, "Estimation of river Tigris dispersivities using a steady-state numerical model," *Applied Water Science,* vol. 8, pp. 1-9, 2018.

[31] B. Abed, M. H. Daham, and A. H. Ismail, "Water quality modeling and management of Diyala river and its impact on Tigris River," *Journal of Engineering Science and Technology,* vol. 16, pp. 122-135, 2021.

[32] Google Map , https://www.google.com/maps/@33.6604661,44.5507459,9.53z, 2022.