Contribution of Statistical Analysis Methods to Assessment of the Physicochemical Quality of the Dam

Aissam Gaagai*, Hani Amir Aouissi, Salah Eddine Maalam, Mostefa Ababsa

Scientific and Technical Research Center on Arid Regions (CRSTRA), Biskra 07000, Algeria

ABSTRACT

We used statistical analysis to try to evaluate the temporal variations concerning surface water located in Baber watershed. So, we collected twenty-one samples, for each city, seven ones were collected. In these samples, we determined eleven physicochemical parameters concerning the eastern part (S1): The salinity of samples proves high TDs values, while low TDs levels were found in the west (S2). In the dam (S3) TDs values are considered as intermediate on average. Regarding the water type: It was SO_4-Na in S1, HCO_3-Ca-Mg in S2 and finally SO_4-Ca-Mg in the dam (S3) which characterizes mixed water. In R-mode, Cluster analysis was used; we classified seven variables into two clusters which are controlled by SO_4. About Q-mode, 21 samples were grouped into three groups corresponding to our three stations: S1, S2 and S3. Factor analysis is indicating to us two factors: These factors explain 87% of the total variance related with data set about water quality. We have two factors (Salinization and sulfates) that explain respectively 73% and 14% of the total variance. We noticed by relying on time series analysis that Cl, Na, Mg, Ca and SO_4 have very similar trend with TDs, and we noticed similar pattern between K and NO_3. It suggests that five components are controlling the evolution of TDs. Negative correlations between K & NO_3 with TDs because of human activities, like domestic sewages (thrown into the tributaries of Wadi Arab) and fertilizers for example. We could notice, by establishing this study, the real benefit of the used technique (statistical analysis) to interpret complex data sets related to the analysis of spatio-temporal variations in the water surface's quality.

Keywords: Statistical analysis; Salinization; Surface water

INTRODUCTION

Most people in the world do not have access to safe drinking water, which has led to an increase in water-borne diseases [1]. Water used for consumption must be potable. Drinking water is water that is free of diseases that produce micro-organisms and chemicals that are dangerous to health [2]. Water consists of different soluble salts and the quality of drinking water is affected by the presence of these different soluble salts [3]. Surface water quality is mainly influenced by the nature and anthropogenic processes especially in urban areas and agricultural activities around rural areas [4]. These anthropogenic processes lead to the contamination of water. Contaminated water not only affects the health of the public, but also the consumption of polluted water can cause various waterborne diseases such as diarrhea, dysentery and complaints of skin, teeth, and other abdominal diseases [5]. Transmission of the disease by drinking water is therefore one of the main concerns for drinking water supply [6]. World demand for water remains mainly growing due to population growth, economic growth, rapid urbanization and increasing demand for food and energy [7]. Therefore, assessing the availability of water resources at the relevant spatial and temporal scales is important [8], as well as the ability to assess the availability of freshwater resources has been a question of importance in most countries for several decades [9].

The main objective of this present work is the characterization of the chemistry of surface water, and the determination of the origin of chemical elements founded by using two techniques: time series analysis and the multivariate statistical techniques. Our investigation area is in Algerian South-East. Most inhabitants (more than 20,000) are concentered in the town of Baber.

MATERIALS AND METHODS

Study area

The study area is located at the downstream of the Wadi (River) Arab. The area of the watershed is 567 km²; Babar dam with a capacity of 54 mm³ is built to impound the Wadi waters (Figure 1).

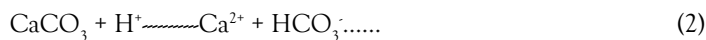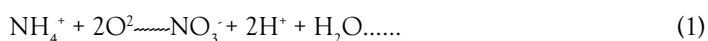**Figure 1:** Map of situation of Babar area.

Methodology

The application of multi-variate environmental statistical techniques such as factor analysis (FA), hierarchical ascending classification (CAH) and time series analysis (EN) has increased considerably in recent years for data analysis and draw useful information [10-12]. Thus, the application of different multivariate statistical techniques can facilitate the interpretation of complex data matrices and the understanding of how to increase the water quality and ecological status of a freshwater system. It also helps to identify possible factors influencing water systems, while being a valuable tool for reliable management of water resources in the area. The objective of this work is to characterize the chemical composition of surface waters and to determine the origin of the chemical elements present using multivariate statistical techniques and the time series. All the statistical computations were made using Excel 2010 (Microsoft Office ®) and STATISTICA 6 (StatSoft, Inc. ®).
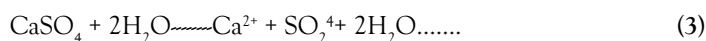
Physicochemical analysis

Monthly surface water was sampled from Wadi Arab and Babar dam, from October 2007 to April 2008 at three stations. The Station 1 (S1) and Station 2 (S2) are located downstream of urban and industrial discharges and Station 3 (S3) is located at the lake of the dam (Figure 1). Total 21 samples (7 from each station) were collected. Samples were analyzed for the major ions. The physico-chemical data that were analyzed are pH, temperature of water (T), Total dissolved solids (TDS), calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^{+}), potassium (K^{+}), chloride (Cl), sulfate (SO_4^{2-}), bicarbonates (HCO_3^{-}), nitrate (NO_3^{-}) (Table 1). This was achieved using standard methods suggested by the American Public Health Association [13]. The respective ionic balance is generally around 5%.

RESULTS AND DISCUSSION

We determined eleven physicochemical parameters during this study (Table 1). To determine the chemical facies of this surface water, the application to diagram of Schoeller Berkaloff demonstrated that the sulphate-calcium is dominate for this all echantillons (Figure 2). The source of the calcium can derive from fertilizer nitrogen who participate to the dissolution of carbonates according to reaction 1 and 2 [14].

$$NH_4^+ + 2O^{2-} \text{------} NO_3^- + 2H^+ + H_2O...... \qquad (1)$$

$$CaCO_3 + H^+ \text{------} Ca^{2+} + HCO_3^-...... \qquad (2)$$

For the sulfate, he can derive from the dissolution of gypsum, anhydrite and halite occur according to reaction 3 [15].

$$CaSO_4 + 2H_2O \text{------} Ca^{2+} + SO_2^4+ 2H_2O....... \qquad (3)$$

During the study period, the change in water temperature is similar in the three stations and is largely influenced by the general climate of the region. The low temperatures coincide with the cold period and high temperatures coincide with the warm periods. High temperatures correspond to the warm periods while the low temperatures correspond to the cold ones. pH value of water samples varied from 7.4 to 8.9 indicating that water is slightly alkaline in nature. In station S2, pH values are higher than the pH values of the two other stations, this being because of the presence of carbonate formations that we found in this part of the field.
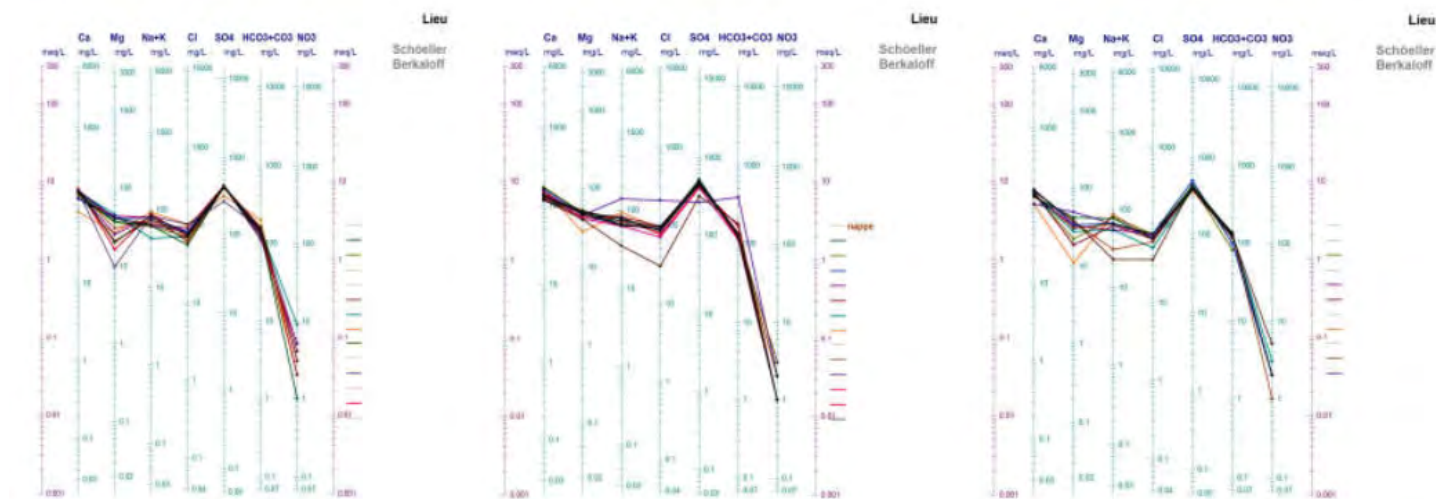
Conventionally, the process of buffering calcite and dolomite, are dominant for the pH range 6.5 to 7.5 [16]. For the salinity, the S1 with high TDS (2837.1 ± 285.3 mg/l), S2 with low TDS (917 ± 201.5 mg/l) and S3 with intermediate and average TDS (1557.1 ± 295 mg/l) therefore, the salinity can be classified in this order S1>S3>S2. Higher concentration of TDS observed in S1 and S3, due to the impact of solubility of evaporate element as gypsum, anhydrite, and halite [17,18]. High concentrations of evaporite elements (Na, K, Cl and SO_4) are recorded at S1 with 908, 7.7, 293.2 and 966.1, respectively. This is mostly due because of the presence of evaporate formations in the concerned area. Another reason for the increase of Na, Cl and SO_4 concentration is related with the effluents input from the industrial and urban sector [19]. It would also be wise to note that the S3 waters are strongly influenced by those of S1. The Ca and Mg content are greater in S1; this is in connection with the dissolution of gypsum and epsomite. As against the HCO_3 concentration which is higher in S2 due to dissolution of carbonates formations. The variation of the concentration of NO_3 appears to be similar in the three stations. However, we note that the high values are recorded in periods of drought while low concentrations recorded during floods [20]. All values are in mg/l except pH and T (°C).

STATISTICAL ANALYSIS

In this case study, environmental techniques and time series analysis were used to evaluate temporal variations in surface water in the Baber watershed. The sources of water pollution in the Baber watershed could more probably be derived from industrial and urban wastewater, irrigation activity and mineral weathering. The application of Piper diagram for water surface in the region can show three chemical facies. The water type in S1 is SO_4-Na for S1, in S2 is HCO_3-Ca-Mg and is SO_4-Ca-Mg water in the dam. These water types are, in fact, a reflection of the predominant influence of gypsum, anhydrite and halite in the Eastern part of study area. The predominant influence of carbonate material is in the western part of study area and mixed water in the dam. In the cluster analysis (R-mode) seven variables were classified into two clusters controlled by SO_4. These groups are: cluster 1 with evaporate elements: Na, Cl and K, cluster 2 with carbonate elements: Ca, Mg and HCO_3. In Q-mode, 21 samples were classified into three groups. The first group is made up by the samples belonging to the station S1. Second group 2 is represented by the samples belonging to the station S2. The third group is made up by the samples belonging to the station S3. In PCA, the two main principal components explain 87.2%

Table 1: Range, mean, standard deviation (SD), of water quality parameters of the Babar dam during 2007-2008.

Variables	T	pH	TDS	Ca	Mg	Na	K	CL	SO$_4$	HCO$_3$	NO$_3$
Station						S1 (n=7)					
Min	3.7	7.6	2320	224.4	148.9	811.9	6.7	255.6	880	190.3	50
Mean	10.9	7.9	2837	261.5	168.3	908	7.7	293.2	966	216.6	81.4
Max	20.5	8.1	3200	284.6	194.6	1019	9.6	308.9	998	262.3	120
SD	5,5	0.2	285.3	20.3	17.2	73	0.9	16.9	41.3	23.5	27
Cv (%)	51	3	10	7	10	8	11	6	4	10	33
Station						S2 (n=7)					
Min	3	7.5	700	92.1	55.2	53.3	3.8	35.3	32.5	229.4	5
Mean	10.7	7.9	917.1	110.7	72.5	66.2	5.7	57.8	38.9	262.3	10.1
Max	20.6	8.3	1250	149.7	104.4	80.8	7	71	46	298.9	16
SD	6	0.2	210.5	18.6	15.1	11.4	1.1	12.2	4.4	26.3	3.9
Cv (%)	56	3	22	17	21	17	18	21	11	10	38
Station						S3 (n=7)					
Min	4	7.4	1000	120.2	71.9	490.6	6.9	113.6	650	222	10
Mean	10.9	7.8	1557	147.4	93.2	574	7.6	158.4	726	303.1	42.9
Max	20.6	8.1	1890	176.4	117.1	679.8	8.1	205.9	770	384.3	90
SD	57	0.3	295	18.1	16.8	60.5	0.5	30.8	40.7	47.2	24.9
Cv (%)	52	3	19	12	18	11	7	19	6	16	58
						Total (n=21)					
Min	3	7.4	700	92.1	55.2	53.3	3.8	35.3	32.5	190.3	5
Mean	10.8	7.8	1770	173.2	111.3	516.1	6.9	169.8	577	260.6	44.8
Max	20.6	8.3	3200	284.6	194.6	1019	9.6	308.9	998	384.3	120
SD	5.9	0.2	862.2	68.6	45.4	359.1	1.3	101.2	404	50.2	36.9
Cv (%)	54	3	48	39	40	69	18	59	70	19	82

Figure 2: Diagram of Schoeller Berkaloff applied to Wadi Arab water.

of the total variance. Time series analysis also provided the same results. The autocorrelation shows that TDS, Ca, Mg, Na, K, Cl and HCO$_3$ have a strong linear interrelationship, and they are subjected to periodically changing sources. In fact, NO$_3$ and SO$_4$ are affected by anthropogenic sources, urban action, and geological condition. Spectral density functions of Ca, Mg, Na, Cl and HCO$_3$ has almost similar trend with TDS. It suggests that TDS was affected by these elements and that TDS plays a more vital role on the surface water quality. But the multiple peaks of the spectral density functions of pH, K, SO$_4$ and NO$_3$ resulted from the human activity, fertilizer, and domestic sewage. The cross-correlations of Ca, Mg, Na, Cl and

SO$_4$ have very similar trend with TDS and K have similar pattern as NO$_3$. It suggests that five components are the controlling factors of the TDS. K and NO3 represent a negative correlation with TDS, it resulted from the human activity, fertilizer, and domestic sewage inputs from the tributaries of Wadi Arab. The results of this study clearly demonstrate the usefulness of the multivariate statistical analysis in hydrochemistry.

CONCLUSION

To quantify pollution, environmental techniques and time series analyzes were used to assess temporal variations in surface water

in the Babar watershed. Sources of Water Pollution in the Babar Watershed Could Be Further Derived from Industrial and Urban Wastewater, Irrigation Activities and Mineral Storms. In the group analysis (R mode), seven variables were classified into two groups controlled by SO_4. These groups are group 1 with evaporated elements: Na, Cl and K, group 2 with carbonate elements: Ca, Mg and HCO_3. In Q mode, 21 samples were classified into three groups. The first group consists of the samples belonging to the S1 station. The second group 2 is represented by the samples belonging to the station S2. The third group consists of the samples belonging to the S3 station. In the ACP, the two main components explain 87.2% of the total variance. Time series analysis also yielded the same results. The autocorrelation shows that TDS, Ca, Mg, Na, K, Cl and HCO_3 have a strong linear interrelation, and they are subjected to periodically changing sources. Indeed, NO_3 and SO_4 are affected by anthropogenic sources, urban action, and geological status. The spectral density functions of Ca, Mg, Na, Cl and HCO_3 have an almost similar trend with TDS. He suggests that TDS has been affected by these elements and that the TDS plays a more vital role in the quality of surface water. But the multiple peaks of the spectral density functions of pH, K, SO_4 and NO_3 result from human activity, fertilizers, and domestic wastewater. The cross-correlations of Ca, Mg, Na, Cl and SO_4 have a very similar tendency with TDS. By cons K has a model like NO_3. He suggests that five components are the controlling factors of TDS. K and NO_3 are negatively correlated with TDS, resulting from human activity, fertilizers, and domestic sewage from tributaries of Oued El Arab. The results of this study clearly demonstrate the utility of multivariate statistical analysis in hydrochemistry.

DISCLOSURE

There is no conflict of interests.

ACKNOWLEDGMENTS

No disclosures or acknowledgments.

REFERENCES

1. TWAS. Safe drinking water, the need, the problem, solutions, and action plan. 3rd World Academic of sciences, Italy. 2002;8-12.

2. Lamikanra A, Essential Microbiology for students and Practitioners of Pharmacy, Medicine and Microbiology. (2nd edn). Amkra. 1999;406.

3. Sonawane VY, Khole AM, Water quality of some drinking waters in Parbhani city: A case study. J Chem Pharmaceut Res. 2010;104-107.

4. Ayeni AO, Balogun II, Soneye ASO. Seasonal assessment of physico-chemical concentration of polluted urban River: A case of Ala River in south western Nigeria. Res J Environ Sci. 2011;21-35.

5. Bhardwaj RM. Water quality monitoring in India achievements and constraints. International Work Session on Water Statistics. Vienna. 2005;2:20-22.

6. Ahmed TK, Tahir SS, Rauf N. Bacteriological analysis of water collected fro m different dams of Rawalpindi/Islamabad region in Pakistan. J Biol Sci. 2004;7:662-666.

7. GWP (Global Water Partnership), A handbook for Integrated Water Resources Management in Basins. 2009.

8. Yang H, Zehnder A. Virtual Water: An unfolding concept in integrated water resources management. Water Resour Res. 2007;43:12301.

9. World Meteorological Organization (WMO) Technical Material for Water Resources Assessment. Geneva. 2012.

10. Shihab AS, Abdul Baqi YT. Multivariate analysis of ground water quality of Makhmor plain/North Iraq. J Damascus Univer. 2010;26:19-26.

11. Batayneh A, Zumlot T. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk Basin of north Jordan. Res Environ Earth Sci. 2012;4:756-768.

12. Laaksoharju M, Skarman C, Skarman E. Multivariate mixing and mass balance (M3) calculations, a new tool for decoding hydrogeochemical information. J Appl Geochem. 1999;14:861-871.

13. American Public Health Association, American Water Works Association, Water Pollution Control Federation, & Water Environment Federation. Standard methods for the examination of water and wastewater. American Public Health Association. 1915.

14. Boudoukha A, Bouguerne A, Mebarkia A. Impact of climate change and human activity on the water quality of the Ain Zada dam. Algeria. Int J Environ Global Clim Change. 2014;2:74-85.

15. Stumm W, Morgan JJ. Aquatic chemistry: Chemical equilibria and rates in natural waters. John Wiley & Sons, NY, USA. 2012;126.

16. Geller W, Friese K, Herzsprung P, Kringel R, Schultz M. Limnology of sulphuracidic mining lakes. II Chemical properties: The main constituents and buffering systems. Verha Inter Ver Limnol. 2000;27:2475-2479.

17. Nas B, Berktay A. Groundwater quality mapping in urban groundwater using GIS. Environ Monit Assess. 2010;8:215-227.

18. Pacini N, Donabaum K, De Villeneuve PH, Konecny R, Pineschi G, Pochon Y, et al. Water-quality management in a vulnerable large river: the Nile in Egypt. Int J River Basin Manag. 2013;11:205-219.

19. Kura NU, Ramli MF, Ibrahim S, Sulaiman WNA, Aris AZ. An integrated assessment of seawater intrusion in a small tropical island using geophysical, geochemical, and geostatistical techniques. Environ Sci Pollut Res. 2014;12:7047-7064.

20. Saebo HV. Statistical analysis of effects of measures against agricultural pollution. Environ Monit Assess. 1991;17:137-146.