Egypt. J. Chem. **Vol. 64,** No. 6 pp. 3069 - 3075 (2021)

Egyptian Journal of Chemistry

http://ejchem.journals.ekb.eg/

Drinking water quality assessment using Principal Component Analysis: Case study of the town of Souk Ahras, Algeria

Bensoltane M.A.[a], Zeghadnia L[a], Bordji N[b], Gagan M.[c], Bouranen S[d]

[a] Dept. of Civil Engineering, Faculty of Science and Technology, Univ. of Mohamed Cherif Messaadia, Souk Ahras, Algeria.
[b] Dept. of Management, Faculty of Economic, Commerce and Management Sciences, Univ. of Mohamed Cherif Messaadia, Souk Ahras, Algeria.
[c]Dept. of Zoology and Environmental Science, Gurukula Kangri Vishwavidyalaya, Haridwar, India
[d] Dept. of Process Engineering, Faculty of Science and Technology, Univ. of Mohamed Cherif Messaadia, Souk Ahras, Algeria

Abstract

Water quality can be considered a key contributor to both health and disease for humans. To obtain more advanced information about water quality the principal component analysis (PCA) was applied using the drinking water supply network data during four seasons in 2018 in Souk Ahras city/Algeria in order to extract principal factors that can explain the quality of drinking water. These factors aim to bring together the main explanatory elements of the water quality and their order of responsibility on the quality of the water through physicochemical parameters measured in situ. During the course of study 14 parameters were used for the characterization of the quality of drinking water in Souk Ahras city, these parameters were replaced by six (6) principal components explaining approximately 65.33% of the variance of the data. The study highlights two main problems, the strong signs of the presence of harmful germs and probably malfunction in part of the equipment of the processing and distribution facilities; and second, amount of the residual chlorine in water.

Key words: Drinking water quality; Principal component analysis (PCA); Parameter reduction; Souk Ahras city; Algeria

1. Introduction

Quality of water is an important factor in survival and development of all living forms on earth. In course of time, the quality of drinking water is assessed based on quality criteria across the world. Indeed, the use of poor water for the production of drinking water can increase the risks for the consumer, if the water contains products harmful to health as well as the environment. But while focusing on water as food product, it has gradually become the most monitored food product, and subject to the strictest quality standards [1, 2, 3].Water quality monitoring corresponds to the performance of observations, analyses, tests and results of certain parameters in different locations of the drinking water supply network. The main objective of this operation is to verify that the water distributed obeys the drinking water criteria for protecting the public health [4].A significant proportion of recognized diseases linked to drinking water are related to a deterioration in the quality of distribution[5-7,8]. However, most of the world's water supply is unreliable due to intermittency; in general, the contamination that occurs is local. There are many significant evidences that quality changes in such circumstances can be extreme [9] .Most domestic and industrial drinking water supplies deteriorate in quality before reaching the consumer.

Regular monitoring of drinking water is essential to provide people safe and high quality water that meets all requirements and standards. Distribution systems are generally monitored at many sampling points where samples are regularly taken and then analyzed on site as well as in laboratories according

---

*Corresponding author e-mail: lotfi.zeghadnia@univ-soukahras.dz.;.

Receive Date: 15 December 2020, Revise Date: 03 February 2021, Accept Date: 25 February 2021
DOI: 10.21608/EJCHEM.2021.53654.3112

to a monitoring plan. The sampling frequency and the number of parameters examined must comply with government requirements and regulations, while the assessment of water quality must be applied using statistical analysis of the physical, chemical and biological results. The current study explains the quality of drinking water while assessing water quality parameters based on factors which aim to bring together the main explanatory elements of the water quality. The value of the variation explained can informs us about the cluster by which we must begin to improve the quality otherwise its order of responsibility on the quality of the water and this through the physicochemical parameters measured in situ using the principal component analysis to identify the most important elements affecting the water quality.

2. Materials And Methods

2.1. Collection and analysis of water quality data

The results of the 14 parameters were obtained from 201 samples of drinking water taken from the water supply network of the city for one year (2018) according to a monitoring plan. Water taps located mainly in private and commercial buildings were selected as sampling points, while sampling twice on each point during the study period. Water analyses, including sample collection and storage, were performed in accordance with methods standardized by the laboratories health agency and the Algerian authority for treatment and water supply (Algerienne Des Eaux: ADE) of Souk Ahras. The parameters measured are: pH (measured with pH Meter, type (WTW pH 330i/SET)), ammonium, nitrite, nitrate and sulfates (determined by a colorimetric method using a UV/visible spectrometer type (JASCO V-530)), turbidity (measured by Nephelometer using 0.02 NTU standards), temperature (measured on the site using mercury thermometer), conductivity (using pre-calibrated conductivity meter model 611), permanganate Index(ultraviolet-visible spectrophotometry), residual chlorine (visual spectrometry), phosphates (using colourimetric technique on photometer type palintest), dissolved O_2 (using an oximeter Type HACH), magnesium and calcium (were measured by EDTA-(Ethylene diamine tetra acetic-acid)- titration method).

2.2. Collection and analysis of water quality data

The original data matrix (201x14) was prepared and processed in MS Excel 2010. The columns were constructed from analyzed parameters of drinking water. Missing values in the data set are estimated by 08 measures. Principal component analysis and other statistical calculations were performed using IBM SPSS 24 (Statistical Package for the Social Sciences) software. Before calculation, the data are normalized in order to avoid classification errors resulting from the different orders of magnitude of the variables tested.

2.3. Principal component analysis (PCA)

PCA is the dimension reduction technique, which transforms data in high dimensional space to lower dimensional space by keeping most of the useful information while reducing unwanted information [10-13, 14]. Third, it provides a way to understand and visualize the structure of the complex data set while helping to identify new significant underlying variables [15].The PCA is based on the decomposition of a covariance/correlation matrix by the decomposition of eigenvalues or by the decomposition of singular values of matrices of real data. Principal component analysis was used for grouping the data and looking for hidden relationships between them. Unlike other statistical methods (e.g. discriminant analysis), PCA is a robust technique that does not require normally distributed and uncorrelated variables [16].

3. Results And Discussion

The drinking water samples (n = 201), taken for the purpose of regular monitoring of the quality of drinking water in the supply system of the city of Souk Ahras, were characterized by 14 physicochemical variables (Table.1).

From these data, 6 main or principal components, explaining approximately 63% of the total variance, were estimated using Kaiser criterion [17], eigenvalues greater than or equal to 1 using Cattel plot [18].

Bartlett's sphericity test, allows us to judge the inequality of the latent roots, which means the significant absence of sphericity of the model mentioned. Bartlett's test is a hypothesis test, an approximate form of chi-square. The observed value should be 0,05 or less. In the case of the Souk Ahras city (Table 3) the meaning is equal to: 0,000; in other words we could therefore continue to study the main components of water quality. Hence positive loading indicates that the contribution of the variables increases with the increasing loading in dimension; and negative loading indicates a decrease [19].

For the quality of the representation table 04 is established, table 04 shown the components loadings, six (06) components were obtained:

- The first component expresses 15.668% of the variance explained;
- The second component, 12.651% of this variance;
- The third component, 10.89% of the variance;
- The fourth explains, 9.675%;
- The fifth explains, 8.819%;
- The last explains, 7.636% of the total variance.

Table. 1. Summary statistics for drinking water samples.

	N		Mean	Median	Std. Deviation	Minimum	Maximum	Standard
	Valid	**Missing**						
pH	201	3	7.56	7.59	0.36	6.00	8.40	6.5 - 9
Cl_2(mg/l)	200	4	0.16	0.10	0.19	0.00	1.20	0.2 - 0.5
Cond(µS/cm)	201	3	525.48	487.00	167.91	2.48	1777.00	2800.00
T(°C)	201	3	14.59	14.00	4.56	6.00	29.00	25.00
Mg (mg/l)	201	3	41.89	43.00	10.82	17.49	68.00	50.00
SO_4^{2-}(mg/l)	194	10	160.44	159.50	74.39	27.74	386.71	400.00
O_2(mg/l)	201	3	0.16	0.15	0.04	0.05	0.29	5.00
Turb (NTU)	201	3	2.08	1.34	1.99	0.33	12.00	5.00
PI(mg/l)	201	3	1.29	0.96	0.89	0.30	4.80	5.00
NO_3^- (mg/l)	201	3	0.55	0.00	4.70	0.00	52.00	50.00
NH_4(mg/l)	201	3	0.02	0.00	0.08	0.00	0.47	0.50
PO_4(mg/l)	201	3	0.01	0.00	0.07	0.00	0.48	0.40
Ca (mg/l)	199	5	99.72	87.00	33.16	72.00	200.00	200.00
NO_2^- (mg/l)	198	6	0.01	0.00	0.03	0.00	0.12	0.20

Table 02: Correlation matrix of parameters

	Correlation													
	pH	**Cl_2 (mg/l)**	**Cond (µS/cm)**	**T (C°)**	**Mg (mg/l)**	**SO_4^{2-} (mg/l)**	**O_2 (mg/l)**	**Turb (NTU)**	**PI (mg/l)**	**NO_2^- (mg/l)**	**NH_4 (mg/l)**	**PO_4 (mg/l)**	**Ca (mg/l)**	**NO_3^- (mg/l)**
pH	1.00													
Cl_2 (mg/l)	0.08	1.00												
Cond (µS/cm)	-0.29	0.02	1.00											
T(°C)	-0.22	0.10	0.20	1.00										
Mg (mg/l)	-0.04	-0.08	-0.03	0.02	1.00									
SO_4^{2-} (mg/l)	-0.08	0.04	0.04	0.03	-0.01	1.00								
O_2 (mg/l)	0.08	-0.07	0.03	-0.03	0.14	-0.08	1.00							
Turb (NTU)	-0.11	-0.06	-0.22	-0.08	0.12	0.02	-0.02	1.00						
PI (mg/l)	-0.02	-0.15	-0.24	-0.19	-0.02	0.17	-0.08	0.32	1.00					
NO_3^- (mg/l)	-0.15	-0.10	-0.22	-0.04	0.05	0.07	-0.14	0.77	0.62	1.00				
NH_4 (mg/l)	-0.10	-0.22	-0.16	-0.10	0.04	0.07	-0.13	0.40	0.44	0.56	1.00			
PO_4 (mg/l)	0.02	0.01	0.00	-0.11	0.14	0.30	0.04	0.06	0.16	0.02	0.04	1.00		
Ca (mg/l)	-0.08	0.04	-0.03	0.03	-0.06	-0.01	0.00	-0.05	0.04	-0.07	0.07	-0.05	1.00	
NO_2^- (mg/l)	-0.06	-0.05	-0.04	-0.04	-0.10	-0.08	0.06	0.00	-0.03	0.00	0.06	-0.04	-0.01	1.00

***Determinant = 0.13**

Table 03: KMO and Bartletts Test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.65
Bartlett's Test of Sphericity	Approx. Chi-Square.	312.917
	Df	91
	Signification	0.000

Table.4. Quality of the representation.

Variables	Initial	Extraction
pH	1.000	0.600
Cl_2(mg/l)	1.000	0.478
Cond(µS/cm)	1.000	0.497
T(°C)	1.000	0.529
Mg (mg/l)	1.000	0.671
SO_4^{2-}(mg/l)	1.000	0.568
O_2(mg/l)	1.000	0.578
Turb (NTU)	1.000	0.615
PI(mg/l)	1.000	0.582
NO_3-(mg/l)	1.000	0.716
NH_4(mg/l)	1.000	0.733
PO_4(mg/l)	1.000	0.674
Ca (mg/l)	1.000	0.837
NO_2- (mg/l)	1.000	0.705

This factor model explains 65.33% of all the variables measured for the quality of drinking water in the city of Souk Ahras.

The figure 1 shows the eigenvalues sorted from largest to smallest according to the number of principal components. After the sixth main component the components can be omitted.

The loadings of the components, their eigenvalues and their variances are summarized in table 5. The factors are maximized using Varimax option and after saturation the results were reported in table 06:

The first component is formed by 05 elements. The high turbidity is a sign of the presence of microorganisms [20, 21, 22] especially with an increases in temperature, in summer the temperature reaches up to 29 ° C and may reflect malfunction in the processing and distribution facilities; the presence of the permanganate index increases this hypothesis - always in the summer period- when it is close to the upper limit. The conductivity varies between average mineralization (<300 µs /cm) and high (> 1000 µs /cm) [23]; this variation explains the change in water supply sources, in fact, the city of Souk Ahras some neighborhoods are supplied from spring water. For the PH the variation range is an indicator of the variation of the sources of water supply. Acid pH (6.0 and even 5.0) characterizes peaty land, forest regions (this description complies with the case of Souk Ahras).

The second component consists of two elements namely ammonium (NH_4) and nitrate (NO_3-); we can say that these waters are distributed to consumers without complete oxidation by chlorine. Generally, if chlorination is done properly, the ammonium concentration (ammonium ion) is zero [24, 25, 26].

The presence of ammonium is an indication of pollution caused by the discharges of human or industrial origin [27].Nitrates are plant nutrients; the main reason for the presence of nitrates in water is due to the agricultural practice with excessive use of industrial fertilizers which is already confirmed by the presence of NH_4 [25]. Its excessive presence can cause a significant public health problem.

The third component is formed by phosphates and sulfates. These two parameters can help to explain the unpleasant taste that the consumer can find; both parameters are close to the higher values in certain places compared to the standards.

The fourth component is consist of 03 parameters, the level of free chlorine and that of oxygen, they are very connected and the relationship is reversed, the districts which recorded a drop in free chlorine are exposed to an increase in microorganisms which favors the hypothesis of the presence of pathogenic germs [26]. Magnesium also confirms that the supply is made from several sources of different geological forms.

The fifth and sixth components consist mainly of nitrate and calcium respectively. Unlike nitrate the concentration of nitrite is very low, it is often around their detection limits. The fifth principal component characterizes traces of inorganic nitrogen. Calcium has reached the upper limit in certain areas; the control of this parameter proves to be a necessity to avoid the harmful consequences of its exceeding of the standards.

Table.5. Total Variance Explained by the principal components

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	2.19	15.66	15.66	2.19	15.658	15.658	1.95	13.913	13.91
2	1.77	12.65	28.31	1.77	12.651	28.309	1.78	12.99	26.9
3	1.41	10.09	38.40	1.41	10.890	39.199	1.50	11.745	38.645
4	1.21	8.68	47.08	1.21	9.675	48.875	1.34	9.952	48.597
5	1.12	8.02	55.10	1.12	8.819	57.693	1.12	8.931	57.528
6	1.07	7.64	62.73	1.07	7.636	65.329	1.09	7.801	65.329
7	0.95	6.79	69.52						
8	0.85	6.04	75.56						
9	0.77	5.52	81.08						
10	0.68	4.84	85.92						
11	0.61	4.33	90.25						
12	0.53	3.77	94.02						
13	0.49	3.49	97.51						
14	0.35	2.49	100.00						

Table 06: Principal component loadings/ After Rotation

	Component					
	1	**2**	**3**	**4**	**5**	**6**
Turb (NTU)	**0.74**	-0.04	0.19	-0.06	-0.08	0.12
T(°C)	**-0.70**	-0.06	0.00	-0.10	0.15	0.06
Cond(µS/cm)	**-0.58**	-0.33	0.14	0.07	-0.16	0.07
pH	**0.53**	-0.17	-0.41	-0.04	0.23	-0.26
PI(mg/l)	**0.50**	0.00	0.40	0.34	0.04	0.26
NH_4(mg/l)	0.08	**0.83**	0.08	0.06	-0.05	0.17
NO_3^-(mg/l)	0.04	**0.83**	0.14	-0.07	0.02	-0.10
PO_4(mg/l)	0.10	0.16	**0.76**	0.21	-0.08	-0.13
SO_4^{2-}(mg/l)	-0.03	0.02	**0.70**	-0.22	0.15	-0.04
O_2(mg/l)	0.12	-0.33	0.01	**0.66**	-0.07	0.11
Mg (mg/l)	-0.15	0.13	0.07	**0.63**	0.44	-0.19
Cl_2(mg/l)	0.01	-0.33	0.09	**-0.52**	0.29	0.06
NO_2^-(mg/l)	0.02	0.03	-0.03	0.04	**-0.83**	-0.08
Ca (mg/l)	-0.02	0.03	-0.11	-0.03	0.07	**0.90**

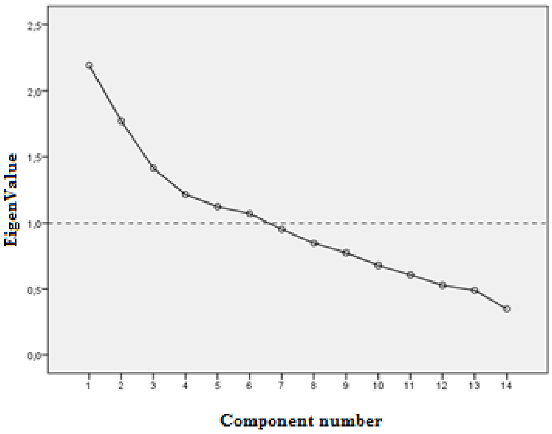

Fig. 1 Scree plot of the eigenvalues.

4. Conclusions

Based on the outputs of the PCA method, the 14 parameters used for the characterization of the quality of drinking water in Souk Ahras city can be replaced by six (6) principal components explaining approximately 65.33% of the variance of the data. The first (1st) component clearly shows that the main problem is the presence of harmful germs and perhaps it is a witness to malfunction in the processing and distribution facilities. The second (2nd) problem is the presence of ammonium and nitrate, which means that the chlorination was not done properly. The unpleasant taste problem is temporary and can be treated, but the priority should be for the first and the second issues. The small residual amount of the free chlorine at the ends points of use is remarkable and need to be enhanced. Attention must be paid to the calcium concentration to keep it within the standards values. The principal component analysis method has proven to be suitable for reducing water quality parameters and determining their relationships by gathering them into principal factors in order to identify the main water quality problems in Souk Ahras city.

5. Conflicts of interest

There are no conflicts to declare.

6. References

[1] Defranceschi M : L’eau dans tous ses états. Ellipses, Paris, 1996.

[2] Ministère du Développement Durable, De l’Environnement de la faune et des Parks : Critères de qualité de l’eau de surface, 3éme Edition, Québec, Direction du suivi de l’état de l’environnement, 2013.

[3] Zenati N, Belahcen N, Rezak A, Zeghadnia L, Djelloul M, Belkadi O : origine de la degradation de la qualite des eaux d’oued medjerda en aval. region de souk ahras. Intern J Of The Environ and water. 7(2) ,62-72, 2018.

[4] Muriel H : Suivi de la qualité de l’eau produite et distribuée : Elaborer et mettre en œuvre un plan des sécurités sanitaire des eaux, Direction des affaires sanitaires et sociales de la nouvelle Caldonie, Santé et environnement, NOUMEA cedex. P 02, 2010.

[5] Guebail A, Zeghadnia L, Djebbar Y: How Does the Rainwater Harvesting Protect the Urban Areas Against Flooding? In: Kallel A., Ksibi M., Ben Dhia H., Khélifi N. (eds) Recent Advances in Environmental Science from the Euro-Mediterranean and Surrounding Regions. EMCEI 2017. Advances in Science, Technology & Innovation (IEREK Interdisciplinary Series for Sustainable Development). Springer, Cham, 2018.

[6] Guebail A, Zeghadnia L, Djebbar Y, BOURANENE S: Rainwater harvesting in Algeria: utilization and assessment of the physico-chemical quality Case study of Souk-Ahras region. Courrier du Savoir, 23,85-94, 2017.

[7] Ainsworth RA : Water quality changes in piped distribution systems. World health organization, 2002.

[8] Craun GF, Calderon RL: Waterborne disease outbreaks caused by distribution system deficiencies. J American water works associ 93,64-75, 2001.

[9] Quick RE, Venczel LV, Mintz ED, Soleto L, Aparicio J, Gironaz M, Hutwagner L, Greene K, Bopp C, Maloney K, Chavez D, Sobsey M, Tauxe RV: Diarrhoea prevention in bolivia Through point of use water treatment and safe storage: a promising new strategy. Epidemiology and infection 122, 83-90, 1999.

[10] Stafford J, Bodson P: Analyse multivariiée avec SPSS. Presses de l’Université du Québec, Canada, 2006.

[11] Alver A, Zeynep K : Prediction of full-scale filtration plant performance using artificial neural networks based on principal component analysis Author links open overlay panel. Separation and Purification Technology. 230, 1-12, 2020.

[12] Subba Rao N, Sunitha B, Adimalla N: Quality criteria for groundwater use from a rural part of Wanaparthy District, Telangana State, India, through ionic spatial distribution (ISD), entropy water quality index (EWQI) and principal component analysis (PCA). Environ Geochem Health 42, 579–599, 2020.

[13] Jolliffe IT, Cadima : Principal component analysis: a review and recent developments Philos. Trans. Roy. Soc. A: Math. Phys. Eng. Sci., 374 (2065),20150202, 2016.

[14] Abdi H, Williams LJ: Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat., 2 ,433-459, 2010.

[15] Sanguansat P: Principal component analysis. InTech, Rijeka, Croatia, 2012.

[16] Parus P: Urban water quality evaluation using multivariate analysis. Acta Montanistica Slovaca. (12)2, 150-158, 2007.

[17] Kaiser HF: The Application of Electronic Computers to Factor Analysis. Education and psychological measurement 20(1), 141-151, 1960.

[18] Cattell R B : The scree test for the number of factors. Multivariate Behavioral Research 1(2), 245–276, 1966.

[19] Mishra A: Assessment of water quality using principal component analysis: A case study of the river Ganges. J Water Chem Technol 32(04), 227–234, 2010.

[20] World Health Organization: water quality and health - review of turbidity: Information for regulators and water suppliers. Technical brief. WHO/FWC/WSH/17.01, 2017.

[21] Patel H: Use of Naturally Prepared Coagulants for the Treatment of Wastewater from Dyeing Mills. In: Characterization and Treatment of Textile Wastewater. Elsevier, pp 147-158, 2015.

[22] Rwooa MA, Juahira H, Roslanb NM, Torimana ME, Enduta A, Azida A, Hasnama CNC, Saudia ASM, Kamarudina MKA, Rwooc SA, Mustafaa AD: The assessment of the variation of physico-chemical sources for drinking water quality using chemometrics: a Case study at water treatment plants in klang valley. Jurnal Teknologi (Sciences & Engineering) 78(11),1–9, 2016.

[23] Rodier J : L'analyse de l'eau, 9e édition. DUNOD (éditeur), Paris, France, 2009.

[24] Bensoltane M A, Zeghadnia L, Djemili L, Gheid A, Djebbar Y: Enhancement of the free residual chlorine concentration at the ends of the water supply network: Case study of Souk Ahras city–Algeria. J W and Land Develop, 38(1), 3-9, 2018.

[25] Guergazi S, ACHOUR S : Caracteristiques physico-chimiques des eaux d'alimentation de la ville de biskra. pratique de la chloration, Rev. Larhyss, 4,119-127, 2005.

[26] Doré M : Chimie des oxydants- traitement des eaux, Ed. Lavoisier, Paris, 1989.

[27] Bensoltane M A, Zeghadnia L, Guebail A, Araibia AS, Djemili L: Assessment Of The Bacterial Pollution In The Distribution Network/ Case Study Of Souk Ahras Town, Algeria. 2nd Euro-Mediterranean Conference for Environment Integration, Sousse/Tunisia, 2019.