A Multivariate Analysis of Groundwater Chemistry Data

Ahmed Garba, Muhammad Mukhtar, and Aishatu Hussein Santuraki

Abstract

Groundwater serves as the main source of water for drinking and other domestic purposes, and often it is over exploited. Nowadays, it is facing threats due to anthropogenic activities. In this study, a total of 20 groundwater samples from boreholes used for drinking, and other domestic purposes were collected from Hadejia town northeastern Jigawa State, Nigeria, and analyzed for their physicochemical characteristics and trace metal levels using standard methods aimed at assessing the groundwater quality. The well-founded techniques of principal component analysis (PCA) and cluster analysis (CA) were jointly used. PCA was used to investigate the origin of each water quality parameters and yielded five varimax factors/components with 78.69% total variance in the groundwater quality data. First two factors reveal that 47.69% of the total variance dominated by total hardness, TDS, Cl⁻, NO₃⁻, Cu, EC, Fe, and Cr indicating the major variations are related to anthropogenic activities and natural processes. Cluster analysis results grouped the 15 parameters and 20 sampling sites into two and four statistically significant clusters, respectively, based on their similarities. The results showed that concentrations of Mn, Cr, and Pb are slightly higher above the maximum permissible limit of Nigerian standard for drinking water quality (NSDWQ, Nigerian Industrial Standard. NIS 554, Nigerian Standard for Drinking Water Quality (NSDWQ). (2007). Nigerian Industrial Standard (NIS) 554:2007 ICs 13.060.20.) and hence the need for the replacement of the affected wells.

1 Introduction

Groundwater constitutes the largest readily available fresh water reserve on earth. It is the most vital natural resource, which forms the core of the ecological system. It has become the major source of water supply for drinking, domestic, household, agricultural, industrial, recreational, and environmental activities, etc. The usefulness of water for particular purpose is determined by its quality. Groundwater quality is mainly controlled by the range and type of human influence as well as geochemical, physical, and biological processes occurring in the ground.

Groundwater supply schemes consist typically of a large number of hand-pumped boreholes, providing untreated, unmonitored, and often unconnected supplies (Hiscock, 2009). A poor quality of water threatens human health and plant growth (Olajire & Imeokparia, 2001), thereby adversely affecting economic development and social prosperity (Milovanovic, 2007). It is therefore crucial to establish the current status of groundwater quality in the study area and thus its appropriateness for use in drinking and other purposes.

The determination of groundwater quality for human consumption is important for the well-being of the ever-increasing population. Groundwater quality depends, to some extent, on its chemical composition (Wadie & Abduljalil, 2010) which may be affected by natural and anthropogenic factors. Variation in groundwater quality in

A. Garba (✉) · M. Mukhtar
Department of Science Laboratory Technology, Binyaminu Usman Polytechnic, Hadejia, Nigeria

A. H. Santuraki
Department of Chemical Sciences, Federal University Kashere, Gombe, Nigeria

an area is a function of physical and chemical parameters that are greatly influenced by natural processes such as geological formations and anthropogenic activities.

Multivariate techniques such as factor analysis, principal component analysis, and cluster analysis are powerful techniques to identify the underlying processes that control groundwater chemistry, grouping of samples of similar composition and origin, and to predict the variables that differentiate the sampling stations temporally and spatially. These techniques have widely been used as unbiased methods in the analysis of groundwater quality data to characterize groundwater composition influenced by natural and anthropogenic factors (Lambrakis et al., 2004; Dragon, 2006).

2 Study Area

This study is carried out in Hadejia Local Government Area, and it lies in the northeastern corner of Jigawa State between latitude 12° 13′–13° 60′ N and longitude 90° 22′–11′ 00E (Orisakwe, 2013). The topography of the area is dominated by the river Hadejia which it lies on its Northern bank which drained into Lake Chad, which at some points is 1–2 m above the surrounding plains including the town. The climate of the area is semi-arid. It is characterized by a long dry season and a short wet season from June to September. The total annual rainfall ranges from 600 mm in the northern part to 762 mm in the southern areas (Orisakwe, 2013). The regional vegetation falls within the Sudan Savannah type.
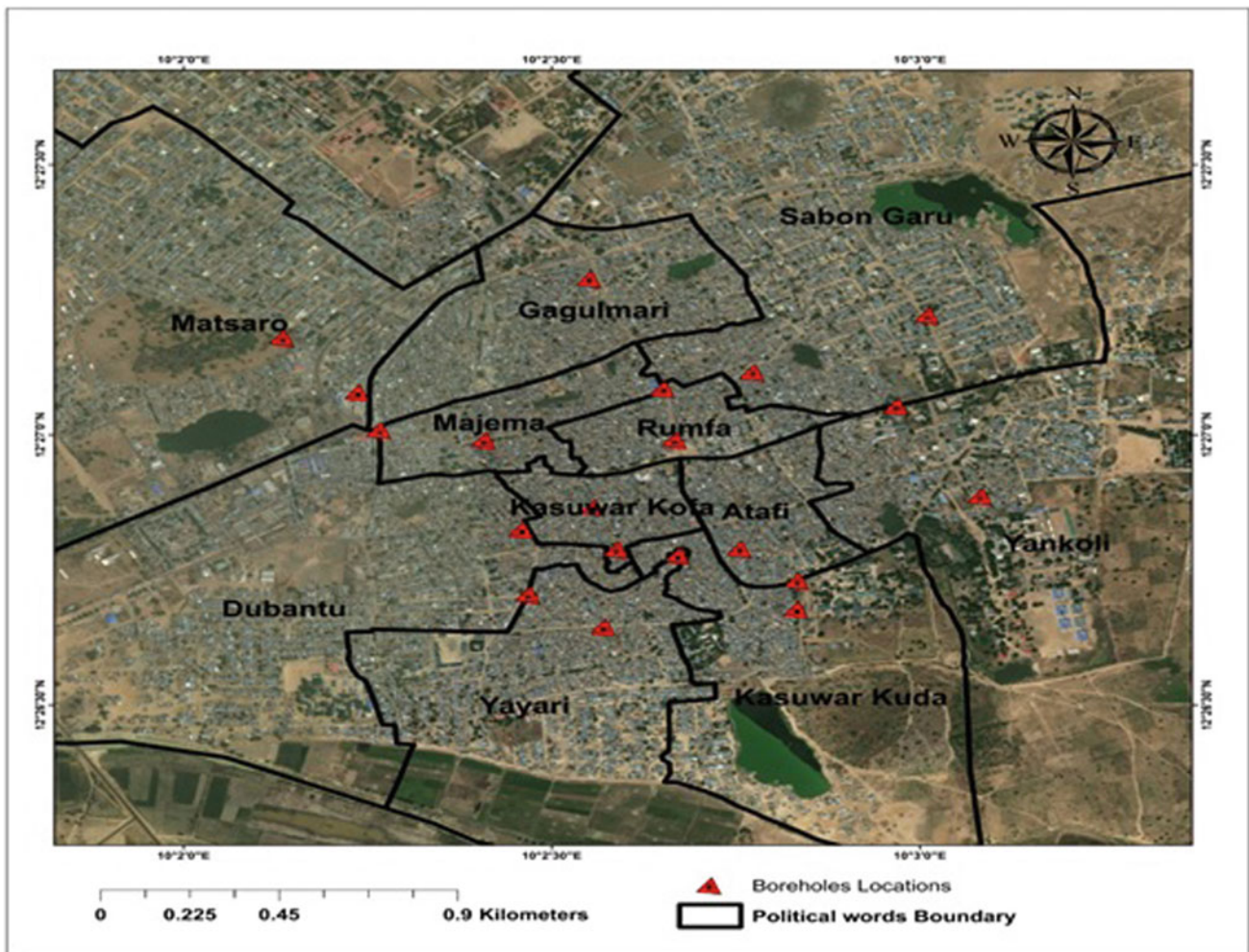
3 Sample Collection and Analysis

Groundwater samples for the analysis were collected from 20 different sites monthly in February, March, and April, 2019 during the dry season. All the samples were collected and treated in accordance with the (American Society for Testing and Material, 2004; American public health association [APHA], 2005). The boreholes were allowed to flow for about 3 min before the water is collected. Temperature, pH, electrical conductivity, and total dissolved solids were measured in situ using Hanna pH 210 model microprocessor, EQ-660 digital conductivity meter, and Hatch 446,600 model TDS meter, respectively. T/H, Ca, and Mg were determined using titrimetric method (American Society for Testing and Material, 2004). NO_3^-, Cl^-, Fe, Cu, Zn, Mn, Cr, and Pb were determined using CHEMetrics Vacu-Visual self-filling ampoules water testing kits with CHEMetrics Photometric Multi-Element Analyzer model V-2000. The samples were chemically analyzed at the Center for Energy, Research and Training, Ahmadu Bello Univeristy Zaria, Nigeria.

4 Results and Discussion

The values for physicochemical variables were presented in Table 1. Parameters values such as EC, pH, Temp., TDS, TH, Ca, Mg, Cl^-, and NO_3^- were found to be below or within the permissible limit set by the (NSDWQ, 2007) in majority of the boreholes. The results similarly showed that concentrations of Mn, Cr, and Pb are slightly higher above the maximum permissible limit of Nigerian standard for drinking water quality (NSDWQ, 2007) and hence the need for the replacement of the affected wells (Table 2).

The rotated factor loading, eigenvalues, percentages of variance, and cumulative percentages of variance associated with each factor for PCA are summarized in Table 3. The results show that five (5) principal components were extracted and rotated using the varimax normalization (Liu et al., 2003) which accounts for more than 78% of the total variance, which is quite good and can be relied upon to identify the main source of variation in the hydrochemistry. The factor loading was sorted according to the criteria of (Bartram & Balance, 1996) by which strong, moderate, and weak loadings corresponds to absolute loading values of >0.75, 0.75–0.50, and 0.50–0.30, respectively.

In PCA, eigenvalues are normally used to determine the number of components (PCs) that are retained for further analysis. An eigenvalue gives a measure of the significance of the factor. The factors with highest eigenvalue are the most significant. Eigenvalues of 1.0 or greater are considered significant (Shrestha & Kazama, 2007)). Number of components is equal to the number of variables in PCA.

The first component loading accounted for 30.27% of the total variance showed higher loading for TH, TDS, NO_3^-, Cu, and Cl^-, with significant contribution from Ca and Mn. It is reasonable to observe a strong positive loading of TH, TDS, NO_3^-, with Cl^-, which may be probably due to leachate from domestic waste discharge in some part of the area. The higher Cl^- and TDS is an indication that there is a possibility that the groundwater is polluted by sewage or waste from refuge dumping site (Amadi, 2010). This factor can be labeled as the hardness and anthropogenic.

Table 1 Value of physicochemical parameter in groundwater of Hadejia

Parameter	Results range	NSDWQ (2007)	WHO (2011)
pH (Range)	5.96–7.58	6.5–8.50	6.5–8.50
Temperature (°C)	27.60–28.00	Ambient	Ambient
EC (µS/cm)	98.00–1306	1000	1000
TH (mg/L)	62.00–374	150	150
TDS (mg/L)	53.90–723	500	500
Fe (mg/L)	0.14–0.88	0.30	0.30
Cu (mg/L)	0.02–5.57	1.0	1.0
Zn (mg/L)	0.01–2.37	3.0	3.0
Mn (mg/L)	0.28–19.50	0.20	0.20
Cr (mg/L)	0.04–0.25	0.05	0.05
Pb (mg/L)	0.34–2.95	0.01	0.01
Ca (mg/L)	9.61–76.95	NS	75
Mg (mg/L)	6.32–56.42	50–150	50–150
Cl^- (mg/L)	0.69–12.64	250	250
NO_3^- (mg/L)	1.35–34.80	50	50

Table 2 PCA rotated component matrix with Kaiser normalization

	VR1	VR2	VR3	VR4	VR5
pH	−0.103	−0.266	−0.153	−0.037	−0.58
Temp	0.066	−0.057	0.005	**0.877**	0.181
EC	−0.292	**−0.729**	0.112	0.046	−0.176
TH	**0.871**	−0.187	−0.180	−0.234	0.153
TDS	**0.869**	0.169	0.341	0.237	0.034
Ca	0.362	0.201	**−0.661**	0.069	0.395
Mg	−0.015	0.497	**0.665**	0.005	0.288
Cl⁻	**0.549**	−0.048	**0.621**	0.305	0.204
NO₃	**0.891**	0.244	−0.003	0.063	−0.179
Fe	−0.242	**0.607**	0.109	−0.483	−0.178
Cu	**0.811**	0.285	0.317	0.203	−0.045
Zn	−0.144	−0.027	−0.08	−0.043	**0.906**
Mn	0.346	0.037	**0.780**	0.050	0.001
Cr	−0.154	**−0.859**	−0.093	0.116	−0.127
Pb	−0.03	0.168	−0.128	**−0.782**	0.289
Eigenvalues	4.54	2.615	1.974	1.624	1.051
% of variance	30.267	17.431	13.161	10.826	7.006
Cumulative %	30.267	47.698	60.858	71.684	78.690

Table 3 Mean values of geochemistry data

Cluster	pH	Temp	EC (μS/cm)	TH (mg/L)	TDS (mg/L)	Ca (mg/L)	Mg (mg/L)	Cl⁻ (mg/L)	NO₃⁻ (mg/L)	Fe (mg/L)	Cu (mg/L)	Zn (mg/L)	Mn (mg/L)	Cr (mg/L)	Pb (mg/L)
1 (n = 9)	6.53	27.8	*235*	154	207	**53.8**	30.4	1.98	7.35	0.49	0.48	0.34	*0.75*	0.11	1.57
2 (n = 3)	6.25	27.9	597	151	210	28	28	4.57	*1.25*	**0.53**	0.20	**0.87**	0.98	0.17	1.70
3 (n = 4)	6.38	27.8	286	**301**	**588**	44.8	32.5	6.79	**29.7**	0.29	3.53	0.05	**5.56**	*0.08*	1.39
4 (n = 4	6.76	27.8	**1116**	*145*	91	*19.7*	*20.1*	0.28	2.78	*0.26*	*0.01*	*0.03*	0.93	**0.18**	*0.99*

Factor loading (VR2) accounted for 17.43% of total variance represents higher loadings for Fe and higher negative loading for Cr and Electrical conductivity, with significant contribution from Mg. These could be conceived to mainly originate from the ionic dissolution in the course of groundwater migration. This factor can be labeled as a natural process leading to heavy metal pollution. VR3 explains 13.16% of total variance and has a higher loading on Mn, Mg, Cl, and moderate negative loading on Ca with significance contribution from TDS and Cu. The major variables constituting VR3 (Ca, Mg, Cl, and Mn) are related to the hydrochemical variables originating from mineralization of groundwater.

5 Cluster Analysis

Cluster analysis is applied to detect the similarity among the different sampling sites. The clustering procedure generates either cluster or groups on the basis of similarity or dissimilarity of variables. Cluster analysis classifies a set of observations into two or more unknown groups based on combination of internal variables. Therefore, the cluster analysis is to discover a system of organized observations where a number of groups/variables share properties in common, and it is cognitively easier to predict mutual properties based on an overall group membership. This helps
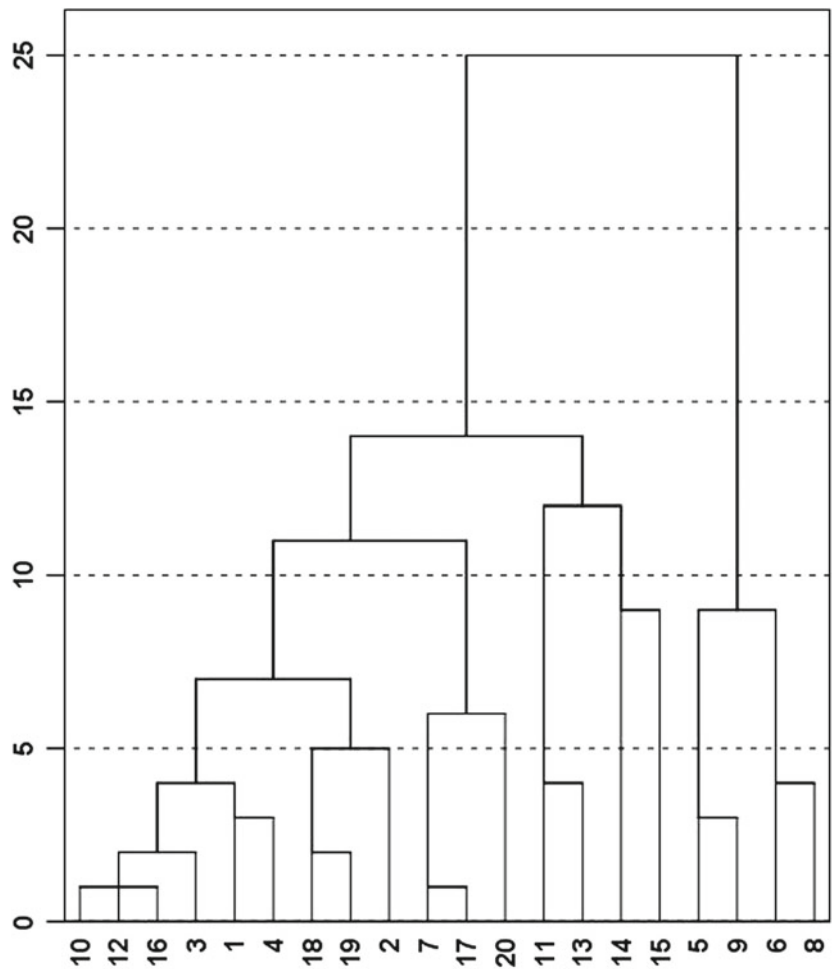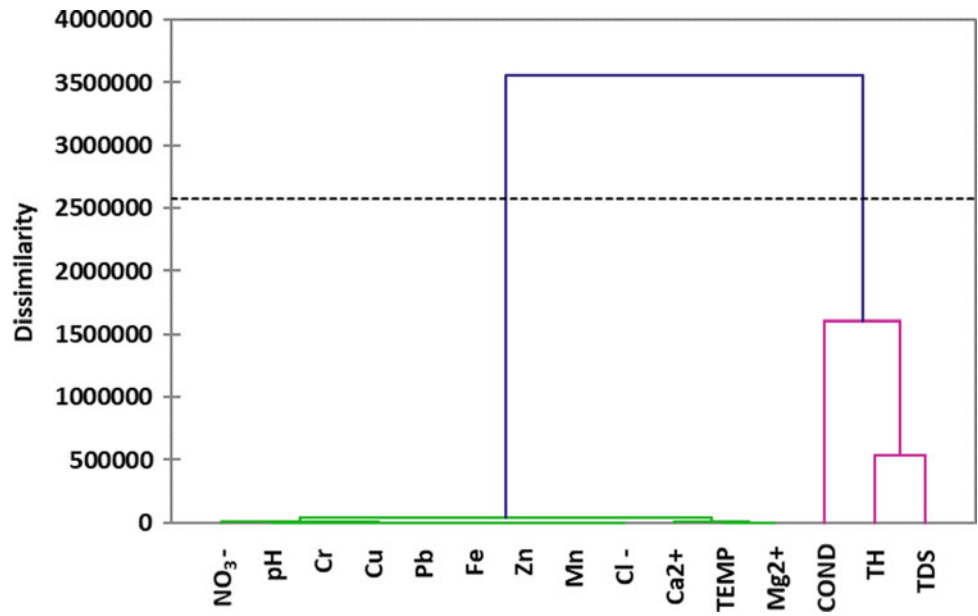
Fig. 1 Dendrogram of water sampling sites

Fig. 2 Dendrogram of hydrochemical parameters

define source profiles of variables and their interpretation in terms of possible sources. The clustering procedure generates either cluster or groups on the basis of similarity or dissimilarity of variables. From Figs. 1 and 2, the cluster analysis resulted in the grouping of 20 monitoring boreholes 15 parameters into four (4) and two (2) statistically significant clusters, respectively, based on their similarities.

Cluster 1: This includes nine (9) boreholes no. 10, 12, 16, 3, 1, 4, 18, 19, and 2(45%)

Cluster 2: It consists of three (3) boreholes which include no. 7, 17, and 20(15%)

Cluster 3: It consists of four (4) boreholes which include no. 11, 13, 14, and 15(20%)

Cluster 4: It consists of four (4) boreholes which include no. 5, 9, 6, and 8(20%).

Samples from cluster 1 are composed of nine (9) boreholes 1, 2, 3, 4, 10, 12, 16, 18, and 19, constituted 45% of the water samples, and are characterized by high concentration of Ca, and low concentration of Mn and EC in all clusters. Sample from cluster 2 is composed of three (3) boreholes 7, 17, and 20, constitutes 15% of the water samples, and is characterized by high concentration of Fe, Zn and Pb and low concentration of NO_3^- in all clusters. Samples from cluster 3 are composed of four (4) boreholes 11, 13, 14, and 15, constituted 20% of the water samples, and are characterized by high concentration of T/Hardness, TDS, Mg, Cl^-, NO_3^-, Cu, Mn, and low concentration of Cr in all clusters. Samples from cluster 4 are composed of four (4) boreholes 5, 6, 8, and 9, constituted of 20% of the water samples, and are characterized by high concentration of EC and Cr and low concentration of T/Hardness, TDS, Ca, Mg, Cl^-, Fe, Cu, Zn, and Pb in all clusters.

6 Conclusion

The quality evaluation of groundwater sources in Hadejia Local Government Area of Jigawa State revealed that the results showed that concentrations of Mn, Cr, and Pb are slightly higher above the maximum permissible limit of Nigerian standard for drinking water quality (NSDWQ, 2007). The values for physicochemical variables EC, pH, Temp., TDS, TH, Ca, Mg, Cl^-, and NO_3^- were found to be below or within the permissible limit set by the (NSDWQ, 2007) in majority of the boreholes. PCA was used to investigate the origin of each water quality parameters and yielded five varimax factors/components with 78.69% total

variance, indicating that the major variations are related to anthropogenic activities and natural processes. Cluster analysis results grouped the 20 sampling points and 15 parameters into four (4) and two (2) statistically significant cluster based on their similarities.

References

Amadi, A. N. (2010). Effects on urbanization on groundwater quality: A case study of Port-Harcort, southern Nigeria. *Natural and Applied Sciences Journal, 11*(2), 143–152.

American Public Health Association (APHA). (2005). *Standard methods for the examination of water and waste water* (20th ed., pp. 65–68) Washington, D.C.

American Society for Testing And Material. (2004). *Annual book of ASTM standards, water and environmental technology.* (Vol. 11.01, pp. pp 79–170). West Conshohocken, PA: Water (i) ASTM International.

Bartram, J., & Balance, R. (1996). *Water quality monitoring: A practical guide to the design of fresh water quality and monitoring programmes*(pp. 34–39). London: Chapman and Hall.

Dragon, K. (2006). Application of factor analysis to study contamination of semi-confined aquifer (Wickopolska buried valley aquifer, Poland). *Journal of Hydrology, 331*, 12–279.

Hiscock, K. M. (2009). *Hydrology: Principles and practice* (pp. 66–75). Oxford: Blackwell Science (2009).

Lambrakis, N., Antonakos, A., & Panagoponlous, G. (2004). The use of multi- component. statistical analysis in hydrological environmental research. *Water Resource, 38*, 1862–1872.

Liu, C. W., Lin, K., & Kuo, Y. M. (2003). Application of factor analysis in the assignment of groundwater Quality in Blackfoot disease area in Taiwan. *The Science of the Total Environment, 313* (1–3), 77–89.

Milovanovic, M. (2007). Water quality assessment and determination of pollution sources along the Axios/Vardar River, southestern Europe. *Desalination, 213,* 159–173.

Nigerian Standard for Drinking Water Quality (NSDWQ). (2007). Nigerian Industrial Standard (NIS) 554:2007 ICs 13.060.20.

Olajire, A. A., & Imeokparia, F. E. (2001). Water quality assessment of Osun River: Studies on inorganic nutrients. *Environmental Monitoring and Accessment, 69*(1), 17–28.

Orisakwe, K. U. (2013). Challenge detection analysis of land uses in Hadejia township of Jigawa State of Nigeria. *International Journal of Applied Sciences and Technology, 3*(3), 60–168.

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Model, Software, 22,* 464–475.

Wadie, A. S. T., & Abduljalil, G. A. D. S. (2010). Assessment of hydrochemical quality of groundwater under some urban areas within Sana'a Secretariat. *Ecletica Quimica, 35*(1), 77–84.