Global NEST Journal, Vol 20, No 2, pp 355-362

Statistical analysis tools for the assessment of groundwater chemical variations in Wadi Bani Malik area, Saudi Arabia

Rehman F.[1,3,*], Cheema T.[2], Lisa M.[1], Azeem T.[1], Ali Naseem A.[1], Khan Z.[4], Rehman F.[5] and Rehman S.U.[5]

[1]Department of Earth Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

[2]Department of Hydrogeology, Faculty of Earth Sciences, King Abdulaziz University, KSA.

[3]Department of Geophysics, Faculty of Earth Sciences, King Abdulaziz University, KSA.

[4]Schlumberger (SIS), Pakistan

[5]Department of Earth Sciences University of Sargodha, Pakistan

Received: 31/12/2016, Accepted: 17/02/2018, Available online: 21/07/2018

*to whom all correspondence should be addressed: e-mail: faisalrehman_atd@yahoo.com

Abstract

Multivariate statistical analysis including cluster analysis, factor analysis and correlation coefficient was applied to the groundwater samples to understand the groundwater condition in Wadi Bani Malik area. Water samples collected from the study area were analysed for physical parameters, major cations and major anions. The groundwater quality is not good in this area and the physicochemical parameters go beyond the acceptable limits of the Kingdom of Saudi Arabia and WHO drinking water standards. The water type in the study area was primarily NaCl. Correlation coefficients were determined among different parameters and then regression analysis was performed to realize the linear relationship between the best correlated parameters. Results of factor analysis specify that factor 1 is dominant and is responsible for 45.5% of the total variance and is dominated by Cl^-, Na^+, Mg^{++} and K^+. Ground water samples can be divided into four clusters.

Keywords: statistical, sample, analysis, groundwater, parameter

1. Introduction

Groundwater remains the main source of water for drinking, industrial and irrigation purposes in many parts of the world. During the last few decades, the disposal of waste without any treatment on the earth makes groundwater vulnerable to contamination. Groundwater contamination problems have started to be discussed in countless studies due to the development in industries all over the world. The groundwater contamination is a significant risk for the population who uses contaminated water for any purpose (Nalbantcilar and Pinarkara, 2015; Tank and Chandel, 2010).

The application of different multivariate statistical techniques, such as cluster analysis (CA), correlation analysis and factor analysis helps in the interpretation of complex data matrices to better understand the water quality and ecological status of the study area. Moreover, these methods help to identify the possible sources/factors that affect the water systems (Kalaivani and Krishnaveni, 2015). Statistical analysis, such as the regression model can be applied to evaluate the quality of water and to determine the main controls and influential variables on the groundwater quality (Guezgouz *et al.*, 2017). A study of correlation decreases the range of ambiguity related to decision-making. In order to analyse the quality of groundwater, and also in the future to evaluate its quality, qualitative analysis has been conducted followed by a correlation–regression study. Also, with least parameters, future forecast of the groundwater quality can be made. Cluster analysis has been performed to acquire the dendrogram displaying the groupings of different sampling locations (Priya and Arulraj, 2011).

The goal of the present investigation is to study the assessment of groundwater chemical variations in Wadi Bani Malik area, Saudi Arabia and change in physicochemical parameters due to the impact of discharge of untreated sewage in the study area.

2. Materials and Methods

2.1 Location and geological setting of study area:

The study area is positioned around 40km east side of Jeddah; one of the major urban centers in Saudi Arabia. It represents more than three million residents and population is increasing by 2.35% per year. The water depletion is approximately 200 liters per capita per day (Rehman and Cheema, 2016). About 70% of Jeddah region is not connected to sewerage pipelines; waste water collected in underground cesspools and then by truck tankers wastewater is transported to the Al Misk lake for more than 10 years. It was estimated that more than 800 tankers vacant 40,000 cubic meters of waste water into the lake every day and the amount has expanded significantly (Elfeki *et al.*, 2010; Ewea, 2010; Rehman *et al.*, 2016a; Rehman *et al.*, 2016b).

There are seven main rock units in the study area and most of the rock units of the study area are cross-cutted by basic dykes. Main rock units are sm (Madrakah formation), xgb (unassigned meta-gabbro and gabbro), kddi, kdqd (Dighbij complex of Kamil suite, diorite and quartz diorite), khtt (Hafnah complex of Kamil suite), gs (syenogranite of unassigned plutons), Quaternary deposits and mafic dykes as shown in Figure 1 (Moore and Al-Rehaili, 1989; Rehman *et al.*, 2016c).

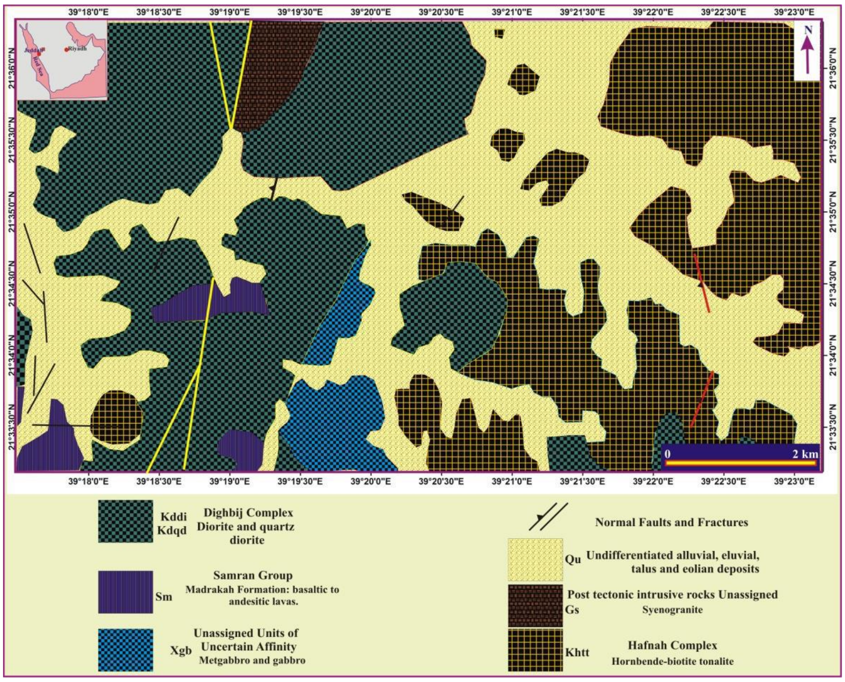


Figure 1. Geological map of Al Misk Lake study area, Eastern Jeddah, Saudi Arabia

A groundwater well survey was directed in order to collect the field hydrogeological parameters and to identify the groundwater wells for geochemical analysis. Based on the survey, nineteen water wells were identified in the area. As there were no other wells, all the water wells were sampled for major ion analysis from Wadi Bani Malik area. Proper sampling techniques were applied for representative sample and data collection. The geochemical analysis was performed by following standard routines. The major anions include sulfate, carbonate, bicarbonate, chloride (SO_4^{--}, CO_3^{--}, HCO_3^- and Cl^-) and cations contain calcium, magnesium, sodium, potassium (Ca^{++}, Mg^{++}, Na^+, K^+). The analysis also includes pH, electrical conductivity, total dissolved solids and total hardness.

2.2. Statistical Analysis

2.2.1 Cluster Analysis

The main goal for cluster analysis is to find subgroups within larger group and to create profile for further action such as marketing and medical intervention (Guezgouz *et al.*, 2017). Clustering is one of the main data mining technique (examining for patterns of concern in a specific representation form) and targets at combining the data objects into significant clusters or classes such that the resemblance of objects within classes is maximized, and the resemblance of objects from different classes is minimized (Zimek, 2009). It leads off with regard to each element as a separate cluster and then combines them into larger clusters. Precisely, in each specific step of clustering, it determines the most faithful pair of clusters and then chains them into a new parent cluster. The outcomes of clustering can be presented with a tree structure design called dendrogram. The dendrogram generated in cluster analysis delivers a beneficial graphical tool in deciding the number of clusters which describe the fundamental process that contributes to spatial variation (Kalaivani and Krishnaveni, 2015; Liu *et al.*, 2012).

2.2.2 Correlation Coefficient Analysis

Correlation coefficient analysis is a useful technique to find out the connection between numerical variables. It can be obtained from the correlation analysis of the ionic concentration of water between different physiochemical parameters. It can be applied to understand the geochemical processed in the groundwater. Correlation analysis measures the closeness and the point of linear association between independent and dependent variables. The closeness and degree of linear relationship between independent and dependent variables can be measured from correlation analysis (Jabal *et al.*, 2014; Kalaivani and Krishnaveni, 2015; Ketata *et al.*, 2011; Pazand *et al.*, 2012; Singh *et al.*, 2012; Srinivasamoorthy *et al.*, 2012). The measure of interdependence of two or more variables is called the correlation coefficient (r) and is usually used to measure and find a relation amongst the variables. Between the two variables the relationship will be considered as well and perfectly associated with a positive or negative linear correlation respectively, when the value of $r \cong +1$ or -1. There will be no correlation between the two variables when $r \cong 0$. Therefore, variables with the positive correlation mean that they have a common source, whereas variables with negative correlation specify different sources. Variables will be strongly correlated if the value of r >0.7 or -0.7 and shows moderate correlation when the values ranges from 0.5 to 0.7 or -0.5 to -0.7 (Jabal *et al.*, 2014; Pazand *et al.*, 2012). The correlation between sodium and chloride is normally used to recognise the mechanism of salinity in arid and semiarid regions (Jabal *et al.*, 2014).

2.2.3 Regression Analysis

It is a statistical tool used to examine the relationships between variables. Generally, the researcher seeks to determine the causal effect of one variable upon another (Sykes, 1993). Regression analysis can be efficiently used for stating the association of different parameters of ground water. Once the regression model for a special area is generated, then in the future the model can be applied for the qualitative analysis of the water (Joarder *et al.*, 2008; Priya and Arulraj, 2011).

2.2.4 Factor Analysis:

Factor analysis technique employed to ease the huge data sets off high complexity by defining a small number of variables that explain the ultimate variance in all of the original variables (Lin *et al.*, 2012). Factor analysis is a beneficial tool for revealing the prospective information buried in the data set comprising hydrochemical processes controlling the groundwater chemistry and environmental data. It permits both the classification of groups of data set and hydrochemical facies examination and also the interpretation of their source (Dragon, 2008; Kalaivani and Krishnaveni, 2015; Sun and Gui, 2015). In factor analysis technique components extraction has been acquired with eigenvalue greater than one by using varimax rotation method. Eigenvalue is a measure of the related variance of the principal component (Dehghanzadeh *et al.*, 2014).

3. Results and discussion:

3.1 Groundwater chemistry and water quality

Qualitative analysis for physicochemical parameters of the groundwater samples are presented in Table 1, which explain the different ions with their units, minimum, maximum, average and standard deviation (SD) values. The results show that most of the samples are high in total dissolved solids, chloride and sodium. The values were also compared with the maximum allowed concentration of different ions in Kingdom of Saudi Arabia and standards of the World Health Organization (Environment, 2011; Jabal *et al.*, 2014; WHO, 2011). According to the maximum allowed concentration standard limits, groundwater quality in the study area is not suitable for drinking as well as agriculture purposes.

Table 1. Summary statistics for concentration of physicochemical parameters, with comparison to WHO (2011) standards.

Parameters	Minimum	Maximum	Ave	SD	WHO (2011)
Ca^{++} (mg/l)	160.3	3767.5	1556.8	847.6	200
Mg^{++}(mg/l)	0.5	1605.1	588.9	331	150
Na^{+}(mg/l)	3000	17900	7349	3535.8	200
K^{+}(mg/l)	6.1	320	38.79	69.0	20
SO_4^{--}(mg/l)	724.3	8645.5	3665.2	1793.8	250-400
CO_3^{--}(mg/l)	10	1128	218.9	256.0	-
HCO_3^{-}(mg/l)	24.4	1040.7	283.8	259.5	350
Cl^{-}(mg/l)	3195	29820	12303.5	6616.1	250
TDS(mg/l)	10008	57240	26130.21	10923	1,000
E.C(mS/cm)	12.49	92.5	36.29	18.7	1.5
Total Hardness(mg/l)	1404.79	13008.38	6306.47	2729.5	500
pH	7.35	8.34	7.84	0.3	6.5–9.2

A thorough examination of major ions exhibited that the area mainly comprised of NaCl type water. The major ions percentage in all water wells is represented in Table 2. From examination, it is clear that chloride and sodium are most abundant ions in the study area. The chloride ranges from 31 to 59% whereas the sodium ranges from 21 to 32%. Sulfate is the third most abundant ion that ranges from 3 to 31%. Whereas calcium, magnesium, bicarbonate and carbonate ions are the less abundant and potassium is least abundant ion in the study area.

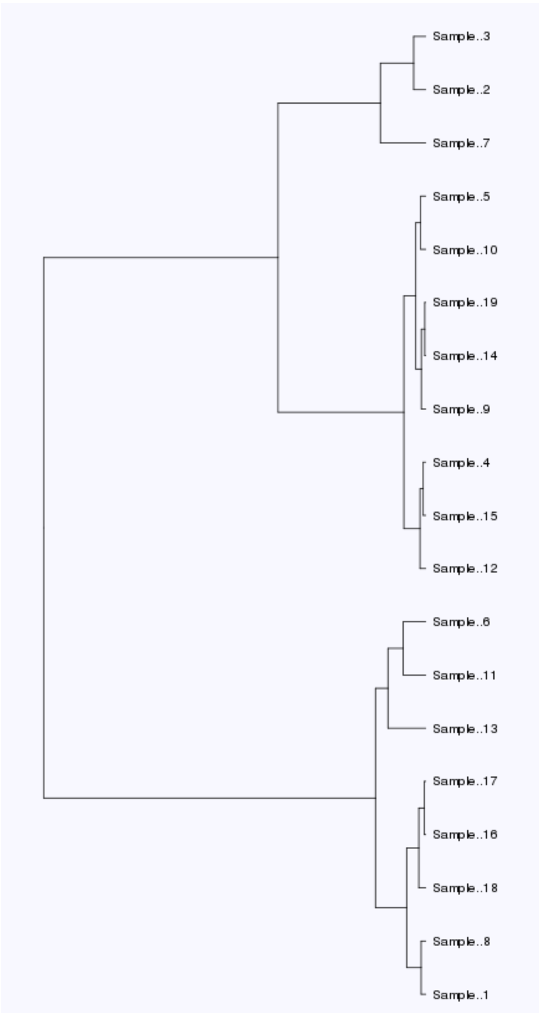
3.2 Cluster Analysis of groundwater samples:

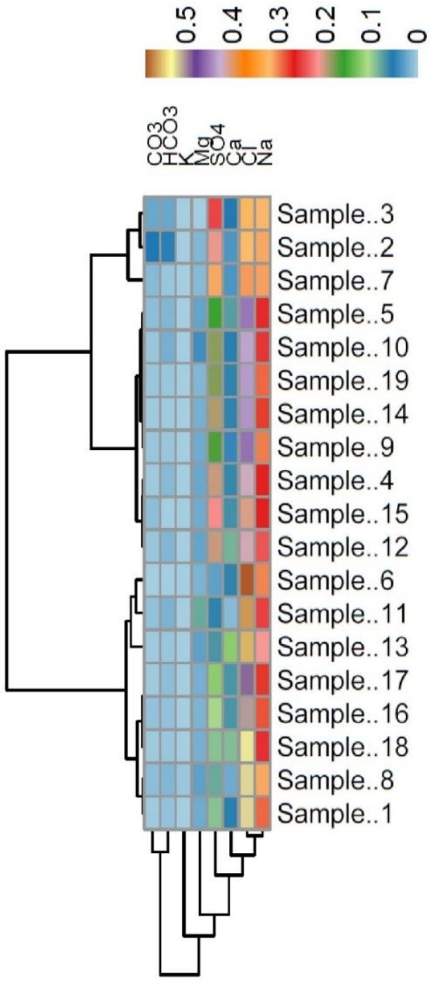
For cluster analysis, we used the "Dendrogram" resulted from cluster analysis based on correlation coefficient of similarity (the paired group averaged method) and heat map cluster analysis.

The groundwater samples in the study area can be divided into four main groups. Group 1 consists of three water samples that are sample 2, 3 and 7. Group 2 consists of eight water samples (5, 10, 19, 14, 9, 4, 15 and 12). Three samples that are sample 6, 11 and 13 are included in group 3. And the last group no 4 consists of five water samples that are 17, 16, 18, 8 and 1. Group 1 is further divided into two subgroups, 1st subgroup consists of sample 3 and 2, whereas the 2nd consists of only sample 7. Similarly, the group 2, 3 and 4 can be subdivided into their subgroups (Figure 2a and 2b). Color scale in Figure 2b shows the concentration of different ions in each sample. It is clear that all the groups have high concentrations of chloride, sodium and sulfate whereas potassium, carbonate, bicarbonate, magnesium and calcium are less in concentration in most of the groups. Similar conclusions were found by Guezgous *et al.,* (2017) for the shallow aquifers located in Algeria.

Table 2. Percentage of major ions in each water sample.

Sample No	Ca^{++} (%)	Mg^{++} (%)	Na^{+} (%)	K^{+} (%)	SO_4^{--} (%)	CO_3^{--} (%)	HCO_3^{-} (%)	Cl^{-} (%)	Water Type
1	6	2	29	0.1	9	0.2	0.9	52	NaCl
2	3	1	31	0.2	21	5	5	32	NaCl
3	6	0	32	0.07	25	2	2	33	NaCl
4	6	2	27	0.2	20	0.3	1	42	NaCl
5	7	3	26	0.2	16	1	2	45	NaCl
6	6	2	30	0.1	3	0.02	0.2	59	NaCl
7	3	1	31	0.08	31	1	0.7	31	NaCl
8	2	3	31	0.6	8	1	2	52	NaCl
9	4	2	30	0.05	17	0.6	0.5	46	NaCl
10	6	4	26	0.1	18	0.7	2	43	NaCl
11	1	8	25	0.09	6	0.8	2	57	NaCl
12	9	3	24	0.1	20	0.6	2	42	NaCl
13	12	3	21	0.07	7	0.4	0.5	56	NaCl
14	6	2	28	0.08	19	0.3	0.4	44	NaCl
15	7	2	27	0.04	22	1	0.1	41	NaCl
16	7	2	29	0.1	11	0.8	0.8	51	NaCl
17	7	2	28	0.1	12	0.4	1	49	NaCl
18	9	2	26	0.07	9	0.4	0.6	53	NaCl
19	6	1	29	0.03	18	0.6	0.6	44	NaCl





Groundwater Sampling and Analysis

Figure 2a. Dendrogram resulted from cluster analysis based on correlation coefficient of similarity

Figure 2b. Heat map showing the cluster of samples in the study area

3.3 Correlation coefficient analysis of chemical parameters

The Pearson correlation coefficients analysis (r) among twelve water quality parameters, namely electrical conductivity (EC), pH, total dissolved solids (TDS), total hardness, calcium, magnesium, sodium, potassium, sulfate, carbonate, bicarbonate and chloride were calculated for correlation analysis. Interpretation of correlation analysis provides an indication of quick water quality monitoring method. Correlation analysis of the parameters of the groundwater samples in Wadi Bani Malik was accomplished, and the values of r were presented in a correlation matrix (Table 3).

Table 3. Correlation coefficient matrix between major ions of groundwater in Almisk Lake area.

	EC	pH	TDS	Total hardness	Ca^{++}	Mg^{++}	Na^{+}	K^{+}	SO_4^{--}	CO_3^{--}	HCO_3^{-}	Cl^{-}
EC	1											
pH	0.38	1										
TDS	0.40	0.003	1									
Total hardness	0.22	-0.34	<u>0.71</u>	1								
Ca^{++}	0.04	-0.52	0.48	<u>0.87</u>	1							
Mg^{++}	0.38	0.13	<u>0.68</u>	<u>0.64</u>	0.19	1						
Na^{+}	0.40	0.09	<u>0.98</u>	<u>0.55</u>	0.31	<u>0.63</u>	1					
K^{+}	<u>0.72</u>	0.42	<u>0.75</u>	0.35	-0.02	<u>0.75</u>	<u>0.78</u>	1				
SO_4^{--}	0.06	0.38	0.21	-0.18	-0.17	-0.11	0.28	0.12	1			
CO_3^{--}	0.43	0.19	0.14	-0.25	-0.33	0.01	0.22	0.37	0.33	1		
HCO_3^{-}	<u>0.60</u>	0.40	0.27	-0.08	-0.29	0.30	0.33	<u>0.62</u>	0.13	<u>0.86</u>	1	
Cl^{-}	0.36	-0.12	<u>0.96</u>	<u>0.80</u>	<u>0.56</u>	<u>0.73</u>	<u>0.92</u>	<u>0.70</u>	-0.04	-0.01	0.17	1

The underline prominences on when the values of r were greater than +0.5 are significantly correlated.

From the calculated Pearson correlation coefficient analysis, it is clear that TDS and sodium show the highest correlation coeffiecient, while the pH, SO_4 and CO_3 exhibit low and weak correlation with all other examined variables. The electrical conductivity is strongly correlated with potassium (r = +0.72) and bicarbonate (r = +0.60), whereas the pH is negatively correlated with total hardness, calcium and chloride. Relationships between TDS and the examined major and minor ions is presented in Figure 3 and Table 3. The positive high correlation of total dissolved solids was observed with a total hardness (r = +0.71), sodium (r = +0.98), potassium (r = +0.75) and chloride (r = +0.96). Also, TDS is strongly correlated with magnesium (r = +0.68). Total hardness shows a strong correlation with calcium and chloride that is +0.87 and +0.8, respectively. It shows a moderate correlation with magnesium (r = +0.64) and sodium (r = +0.55). Calcium has only a moderate correlation with chloride that is +0.56. Magnesium shows a moderate correlation with sodium (r = +0.63) whereas the strong correlation with potassium and chloride that is +0.75 and +0.73 respectively. The correlation coefficient of sodium with potassium and chloride is +0.78 and +0.92, which means they are strongly correlated. Strong correlation between sodium and chloride indicates that the main portion of salinity in the groundwater is due to sodium chloride salt. The correlation of the potassium with bicarbonate and chloride is strong, that is +0.62 and +0.70 respectively. Carbonate and bicarbonate are strongly correlated with each other and correlation coefficient is +0.86.

3.4 Regression model

For the most correlated parameters, the regression models were generated. A thorough analysis shows that there exhibits a positive correlation between sodium and chloride, between sodium and total dissolved solids, between chloride and total dissolved solids and between calcium and total hardness. For these parameters, a linear relationship was established for the study area:

For Normality of residuals, Jarque-Bera test has been utilized. For Homogenity of variance of residuals, White test of heteroskedasticity has been utilized. All the regression equations satisfy the both assumptions.

TDS = 3.02*Na + 3943.34	**or**	Na = 0.316 **TDS – 917.48
(0.00)		**(0.00)**
R^2 = 0.96		**R^2 = 0.96**
TDS = 1.59* Cl + 6566.17	**or**	Cl = 0.583* TDS – 2940.87
(0.00)		**(0.00)**
R^2 = 0.93		**R^2 = 0.93**
Na = 0.49 *Cl + 1314.56	**or**	Cl = 1.717 *Na – 316.61
(0.00)		**(0.00)**
R^2 = 0.84		**R^2 = 0.84**
Total Hardness = 2.811* Ca+1929.70	**or**	**Ca** = 0.271*TH – 152.98
(0.00)		**(0.00)**
R^2 = 0.76		**R^2 = 0.76**

The value in parenthesis is p-value of the test for significance of β_1. *β_1 is the regression coefficient (slope) and R^2 is the coefficient of determination.* And R^2 of all regression equations is also mentioned.

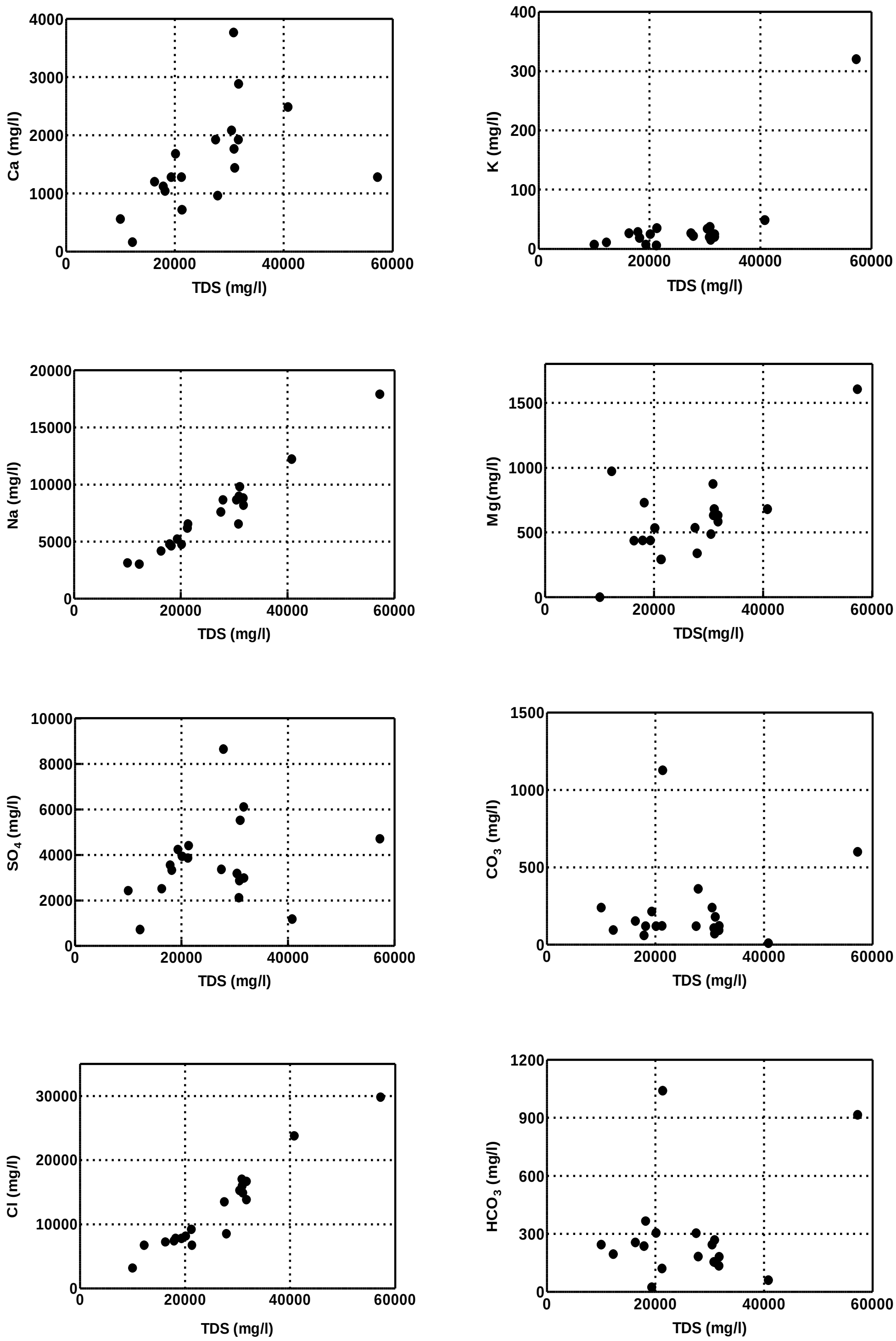


Figure 3. Relationships between TDS and the examined major and minor ions

3.5 Factor Analysis

We performed the factor analysis on the data set of all the samples that are Ca^{++}, Mg^{++}, Na^{+}, K^{+}, SO_4^{--}, CO_3^{--}, HCO_3^{-}, Cl^{-}. Three principal components were extracted that explained about 86.09% of total sample variance. Component extraction has been acquired with eigenvalue greater than one by using the varimax rotation method as shown in the Table 4.

Factor 1 is responsible for 45.5% of the total variance and is dominated by Cl^{-}, Na^{+}, Mg^{++} and K^{+}. The second factor is responsible for 27.76% of the total variance and is mainly contributed to HCO_3^{-} and CO_3^{--}, whereas the explanation of the third factor is only 12.82% and is predominantly by SO_4^{--}.

Table 4. Results of factor analysis

	Factors		
	1	2	3
Ca^{++}	0.452	-0.667	0.008
Mg^{++}	0.839	0.101	-0.252
Na^{+}	0.911	0.061	0.32
K^{+}	0.829	0.448	0.011
SO_4^{--}	0.006	0.161	0.954
CO_3^{--}	0.096	0.863	0.275
HCO_3^{-}	0.326	0.892	0.014
Cl^{-}	0.97	-0.168	0.021
Eigenvalues	3.64	2.22	1.02
Variance	45.50%	27.76%	12.82%

4. Conclusions

Groundwater chemistry data of Wadi Bani Malik area shows that the concentration of sodium, chloride, sulfate, total dissolved solids and total hardness in the groundwater was found to be very high in the study area. The groundwater quality was poor and the water type was predominated by sodium chloride. Groundwater in the study area can be divided into four main groups according to cluster analysis. Correlation coefficient analysis indicates that the highest correlation coefficient in the study area is between total dissolved solids and sodium. In conclusion, the result of this research shows that the integration of the multivariate statistical techniques helps to delineate the groundwater chemical variations and to find the major controls on the groundwater quality.

References

Dehghanzadeh R., Hir N.S., Sis J.S. and Taghipour H. (2014), Integrated assessment of spatial and temporal variations of groundwater quality in the eastern area of Urmia Salt Lake Basin using multivariate statistical analysis, *Water Resources Management*, **29** (4), 1351-1364.

Dragon K. (2008), The influence of anthropogenic contamination on the groundwater chemistry of a semi-confined aquifer (the Wielkopolska Buried Valley Aquifer, Poland), *Water Resources Management*, **(22)**3, 343-355.

Elfeki A., Ewea H. and Al-Amri N. (2010), Linking Groundwater Flow and Transport Models, GIS Technology, Satellite Images and Uncertainty Quantification for Decision Making: Buraiman Lake Case Study Jeddah, Saudi Arabia, *4 International Conference on Water Resources and Arid Environments: Riyadh, Saudi Arabia,* p. 122-130.

Environment, P. o. M. a., 2011, National Ambient Water Quality Standards for KSA, 1409-01.

Ewea H. (2010), Hydrological Analysis Of Flooding Wastewater Lake In Jeddah: Saudi Arabia, *Journal of King Abdel Aziz Univ: Met., Env. & Arid Land Agric faculty*, **125-142**(9).

Guezgouz N., Boutoutaou D., Zeggane H. and Chefrour A. (2017), Multivariate statistical analysis of the groundwater flow in shallow aquifers: a case of the basins of northern Algeria, *Arabian Journal of Geosciences*, **10**(15), p. 336.

Jabal M.S.A., Abustan I., Rozaimy M.R. and El Najar H. (2014), Groundwater beneath the urban area of Khan Younis City, southern Gaza Strip (Palestine): Hydrochemistry and water quality, *Arabian Journal of Geosciences*, 1-13.

Joarder M., Raihan F., Alam J. and Hasanuzzaman S. (2008), Regression analysis of ground water quality data of Sunamganj District, Bangladesh, *International Journal Of Environmental Research*, **(2)**3, p. 291-296.

Kalaivani K. and Krishnaveni M. (2015), Multivariate Statistical analysis of pollutants in Ennore Creek, South-East Coast of India, *Global NEST Journal*, **17**(3), 618-627.

Ketata M., Hamzaoui F., Gueddari M., Bouhlila R. and Ribeiro L. (2011), Hydrochemical and statistical study of groundwaters in Gabes-south deep aquifer (south-eastern Tunisia), *Physics and Chemistry of the Earth, Parts A/B/C,* **36**(5), 187-196.

Lin C.Y., Abdullah M.H., Praveena S.M., Yahaya A.H.B. and Musta B. (2012), Delineation of temporal variability and governing factors influencing the spatial variability of shallow groundwater chemistry in a tropical sedimentary island, *Journal of Hydrology*, **(432)**, 26-42.

Liu X., Zhu X.-H., Qiu P. and Chen W. (2012), A correlation-matrix-based hierarchical clustering method for functional connectivity analysis, *Journal of Neuroscience Methods*, **211**(1), 94-102.

Moore T. and Al-Rehaili M., 1989, Geologic map of the Makkah Quadrangle: Sheet 21D, Kingdom of Saudi Arabia, Ministry of Petroleum and Mineral Resources, Deputy Ministry For Mineral Resources Publication, Jeddah, SA.

Nalbantcilar M. and Pinarkara D. (2015), Impact of industry on ground water contamination: A case study in Konya City, Turkey, *Global Nest Journal*, **(17)**4, 796-815.

Pazand K., Hezarkhani A., Ghanbari Y. and Aghavali N. (2012), Groundwater geochemistry in the Meshkinshahr basin of Ardabil province in Iran, *Environmental Earth Sciences*, **(65)**3, 871-879.

Priya K. and Arulraj G.P. (2011), A correlation–regression model for the physicochemical parameters of the groundwater in Coimbatore city, India, *Environmental Technology*, **(32)**7, 731-738.

Rehman F., Abuelnaga H.S.O., Harbi H.M., Tariq C. and Atef A.H. (2016a), Using a combined Electrical Resistivity Imaging and Induced Polarization Techniques with the chemical analysis in determining of groundwater pollution at Al Misk Lake, Eastern Jeddah, Saudi Arabia, *Arabian Journal of Geosciences,* **9**: 286. https://doi.org/10.1007/s12517-016-2423-9.

Rehman F. and Cheema T. (2016), Effects of sewage waste disposal on the groundwater quality and agricultural potential of a floodplain near Jeddah, Saudi Arabia, *Arabian Journal of Geosciences,* 9: 307. https://doi.org/10.1007/s12517-016-2340-y.

Rehman F., Cheema T., Abuelnaga H., Harbi H., Atef A. and Lisa M. (2016b), interpretation of groundwater chemistry data using multivariate statistical techniques, *Global NEST Journal*, **(18)**3, 665-673.

Rehman F., Helmy S.O.A., Hussein M.H. and Ali H.A. (2016c), Application of Vertical Electrical Sounding and Water Analysis for Study the Contaminated Area at Al Misk Lake, Eastern Jeddah, Saudi Arabia, *Nature Environment and Pollution Technology*,**(15)**1, 311-316.

Singh V.K., Bikundia D.S., Sarswat A. and Mohan D. (2012), Groundwater quality assessment in the village of Lutfullapur Nawada, Loni, District Ghaziabad, Uttar Pradesh, India, *Environmental Monitoring and Assessment*, **(184)**7, 4473-4488.

Srinivasamoorthy K., Vijayaraghavan K., Vasanthavigar M., Sarma S., Chidambaram S., Anandhan P. and Manivannan R. (2012), Assessment of groundwater quality with special emphasis on fluoride contamination in crystalline bed rock aquifers of Mettur region, Tamilnadu, India, *Arabian Journal of Geosciences,* **(5)**1, 83-94.

Sun L. and Gui H. (2015), Statistical Analysis for Understanding Groundwater Chemical Variations, *Nature Environment and Pollution Technology*, **(14)**1, 41.

Sykes A.O., 1993, An introduction to regression analysis.

Tank D.K. and Chandel C.S. (2010), Analysis of the major ion constituents in groundwater of Jaipur City, *Nature and Science*,**(8)**10, 1-7.

WHO, 2011, Guidelines for drinking-water quality, World Health Organization.

Zimek A. (2009), Correlation clustering: ACM SIGKDD, *Explorations Newsletter*, **(11)**1, 53-54.