


Application of multivariate statistical analysis concepts for assessment of hydrogeochemistry of groundwater—a study in Suri I and II blocks of Birbhum District, West Bengal, India

Shreya Das¹ · S. K. Nag¹ 

Received: 29 January 2015 / Accepted: 3 June 2015 / Published online: 20 June 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Multivariate statistical techniques, cluster and principal component analysis were applied to the data on groundwater quality of Suri I and II Blocks of Birbhum District, West Bengal, India, to extract principal factors corresponding to the different sources of variation in the hydrochemistry as well as the main controls on the hydrochemistry. For this, bore well water samples have been collected in two phases, during Post-monsoon (November 2012) and Pre-monsoon (April 2013) from 26 sampling locations spread homogeneously over the two blocks. Excess fluoride in groundwater has been reported at two locations both in post- and in pre-monsoon sessions, with a rise observed in pre-monsoon. Localized presence of excess iron has also been observed during both sessions. The water is found to be mildly alkaline in post-monsoon but slightly acidic at some locations during pre-monsoon. Correlation and cluster analysis studies demonstrate that fluoride shares a moderately positive correlation with pH in post-monsoon and a very strong one with carbonate in pre-monsoon indicating dominance of rock water interaction and ion exchange activity in the study area. Certain locations in the study area have been reported with less than 0.6 mg/l fluoride in groundwater, leading to possibility of occurrence of severe dental caries especially in children. Low values of sulfate and phosphate in water indicate a meager chance of contamination of groundwater due to anthropogenic factors.

Keywords Groundwater quality · Fluoride · Correlation analysis · Hierarchical cluster analysis · Principal component and factor analysis · Suri I and II

Introduction

Assessment of the hydrochemical characteristics of water and aquifer hydraulic properties is important for groundwater planning and management in the study area. Generally, the motion of groundwater along its flow paths below the ground surface increases the concentration of the chemical species (Domenico and Schwartz 1990; Freeze and Cherry 1979; Kortasi 2007). Groundwater is the purest form of water sourced from natural resources and is usually clear, colorless and remains relatively at constant temperature. Therefore, it is normally superior to surface water in terms of sanitary consideration. But groundwater has higher salt contents than surface water because slowly moving water remains in contact with sub-strata for longer periods. Hence, the groundwater chemistry could reveal important information on the geological history of the aquifers as well as the suitability of groundwater for domestic and agricultural purposes. Quality of groundwater is equally important to its quantity owing to the suitability of water for various purposes (Schiavo et al. 2006; Subramani et al. 2005). Hydrochemical evaluation of groundwater systems is usually based on the availability of a large amount of information concerning groundwater chemistry (Aghazadeh and Mogaddam 2010; Hossein 2004). Groundwater chemistry, in turn, depends on a number of factors, such as general geology, degree of chemical weathering of the various rock types, quality of recharge water and inputs from sources other than water rock interaction. Such factors and their interactions result

✉ S. K. Nag
nag_sk@yahoo.com

¹ Department of Geological Sciences, Jadavpur University, Kolkata 700032, India

in a complex groundwater quality (Domenico and Schwartz 1990; Guler and Thyne 2004; Sunne et al. 2005). Various studies on the hydrogeological and hydrochemical properties of the aquifers underlying the country are currently being pursued. Recent studies have been focused on the quality of groundwater from specific hydrogeologic terrains in the country. Researchers (e.g. Kortatsi et al. 2008; Yidana et al. 2008a, b, c, d; Yidana 2008; Banoeng-Yakubo et al. 2009; Nag and Ray 2015; Nag 2014; Nag and Ghosh 2013) have applied different methodologies to understand the sources of variation in the quality of surface and groundwater basins in India.

The competition for water resources has gained importance in recent years, not only in India but also in many places of the world. The development of human societies and industry result in bioenvironmental problems; pollution puts the water, air and soil resources at risk (Milovanovic 2007). Groundwater has become the major source of water supply for domestic, industrial and agricultural sectors of many countries and groundwater chemistry depends on the quality of recharged water, atmospheric precipitation, inland surface water and sub-surface geochemical processes. In recent years, many cities of developing countries are experiencing rapid demographic growth due to rural exodus. Urbanization and the unregulated growth of the population have altered the local topography and drainage system directly which affect both quality and quantity of the groundwater (Vasanthavigar et al. 2010). Temporal changes in the origin and constitution of the recharged water, hydrological and human factors frequently cause periodic changes in groundwater chemistry and quality (Milovanovic 2007; Sreedevi 2004).

The present study has been carried out in two blocks—Suri I and II, of Birbhum district, West Bengal, India. Groundwater is an important water resource for drinking, agriculture and industrial use in the present study area. In this study, physical and hydrochemical data from the groundwater system has been used to determine the main factors and mechanisms controlling the chemistry of groundwater in the area. Geochemical studies of groundwater provide a better understanding of possible changes in quality as development progress. Besides anthropogenic factors, natural factors also play a role in contaminating groundwater, rendering it unfit for drinking. Fluoride contamination in groundwater is one such factor which leads to irreparable physiological damage to the human body in form of dental or skeletal fluorosis. As per World Health Organization guidelines, the permissible limit of fluoride in drinking water is 1.5 mg/l (WHO 2011). Intense agricultural and urban development has caused a high rise in demand of groundwater resources in arid and semi-arid regions of West Bengal and thereby the resource has been put at a greater risk to contamination (Moghaddam and

Najib 2006; Jalali 2006; Khazaei et al. 2006). In the present study area, people depend solely on groundwater supply for all domestic chores as well as for consumption. Excess fluoride in groundwater has been reported previously in Junidpur, Rampurhat and Nalhati blocks of Birbhum district (Gupta et al. 2006; Gupta and Banerjee 2011). The present study reports the presence of excess fluoride in groundwater and intends to identify the hydrogeochemical and anthropogenic contamination factors responsible for deterioration of water quality using statistical concepts.

Inter-relationships amongst the analyzed water quality parameters, with special emphasis on relation of fluoride with the rest of the parameters have been studied with the use of multivariate statistical analysis tools in this study. In the analysis of hydrochemical data, the application of multivariate statistical analyses has been profusely taken up by various researchers. These analyses assist us in systematically studying and interpreting hydrogeochemical patterns and relationships between the various parameters determining water quality. Some examples of studies were multivariate analyses methods have been effectively employed in comprehending hydrochemical processes are studies by Güler et al. (2002), Güler and Thyne (2004), Helsrup et al. (2007), Yidana et al. (2008a), Yidana and Yidana (2009), Cloutier et al. (2008), Machiwal and Jha (2015), Mrazovac and Miloradov (2011), Gummadi et al. 2014, Bajpayee et al. 2012, Jothivenkatachalam et al. 2010 and others. Multivariate statistical methods do not directly present hydrogeochemical processes being carried out but gives an idea about the factors that control these processes.

Study area

The present study has been carried out in Suri (comprising of two blocks—Suri I and Suri II), the district headquarter of Birbhum district, West Bengal, India. The two blocks are located between latitudes 23.76°N–23.99°N and longitudes 87.42°E–87.64°E, and covers an area of 220 km² (approx.). The area usually experiences an arid climate. Summer temperatures soar to a maximum of 40 °C or above whereas in winter temperatures dip to around 10 °C. Majority of the rainfall is limited to the monsoon season from June to October and hovers around an average of 1100 mm. The area is characterized by rural setting and major occupation of the people is agriculture. The study area is largely comprised of alternating layers of sand and clay, which are soft sediments and part of the Ganga-Kosi formation and are of the tertiary age. The clays interbedded with sand partings are poorly sorted with occasional layers of grit. Granite gneiss which date back to the Archean age, are hard and foliated type rocks belonging to the Chotanagpur Gneissic complex and, constitute the north

western part of the study area (Mukherjee et al. 1969). These poly-metamorphosed rocks contain streaks of pegmatite, alkali feldspars and mica. Hard clays dominate small areas in the extreme upper and lower parts and a stretch in the central part—all lying in the eastern region of the study area. Presence of laterites and lateritic soil, which render a reddish color to the top soil in many regions of the study area, is observed to have a scattered distribution spanning both Suri I and II blocks. Figure 1 presents the study area map marked with the 26 sampling location points from where water samples have been collected along with the geological map of the study area.

Materials and methods

Groundwater sample collection from the study area has been done in two phases: post-monsoon and pre-monsoon sessions. Water samples were collected from bore to wells at 26 sampling locations spread homogeneously over the two blocks, Suri I and II. Depth of the water table ranged widely in the study area. During post-monsoon the Depth to Water Level (DTWL) values ranged from 1.69 to 13.83 m and during pre-monsoon it ranged from 3.03 to 21.24 m. The samples were collected in 500 ml sterilized polythene bottles. For all samples collected, parameters

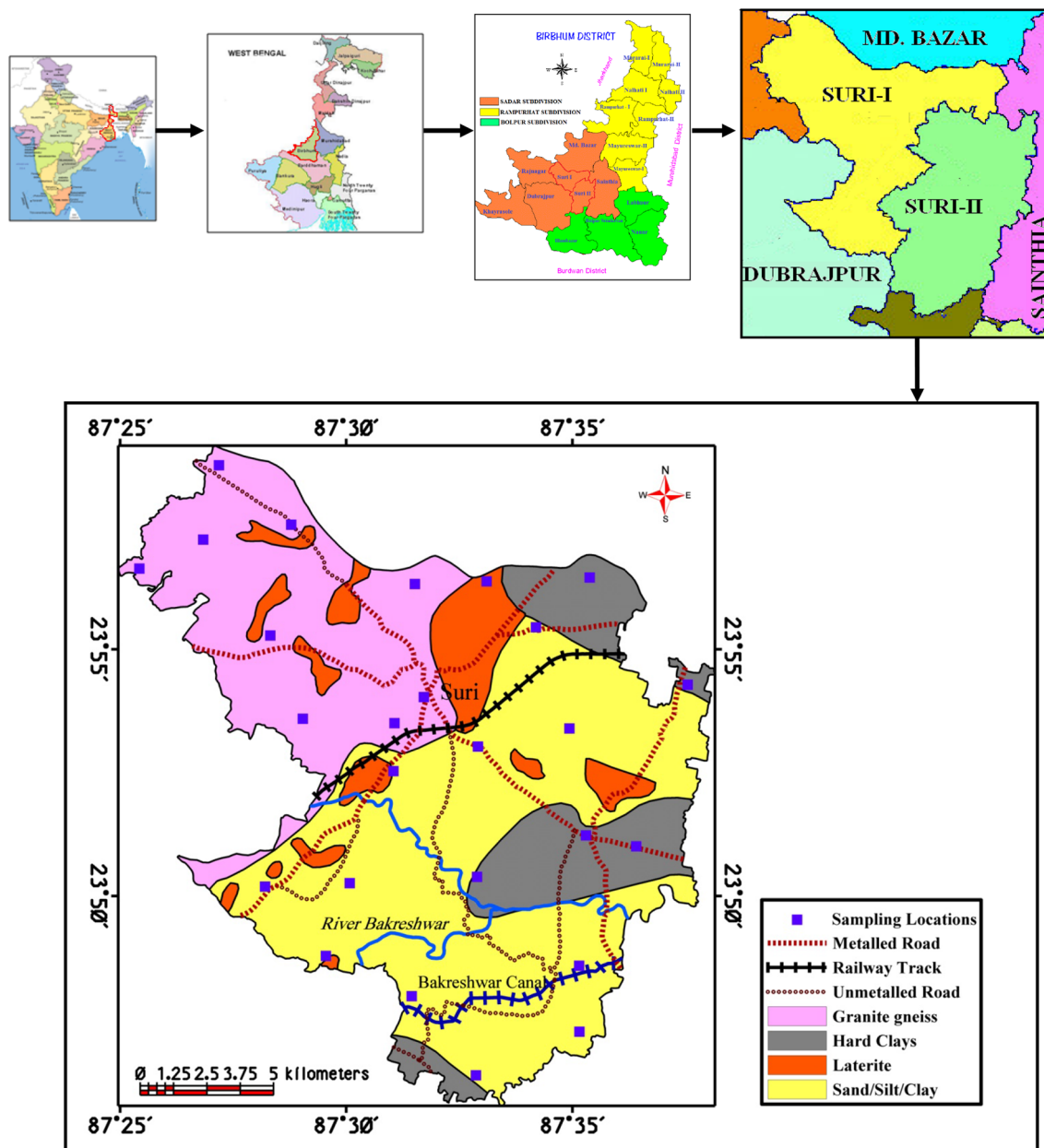


Fig. 1 Map of the study area presenting the sampling location points and geology of the area

like pH, electrical conductivity (EC) and total dissolved solids (TDS) were measured in the field. This is necessary as they are likely to change during transport. The water samples were tested for a total of sixteen parameters. The quantitative analysis for major cations and anions in water including fluoride was carried out in the laboratory employing standard methods (APHA 1995). Various statistical analysis methods have been used in the past to determine water quality and hydrogeochemistry (Nelson and Ward 1981; Briz-Kishore and Murali 1992; Ramesh Kumar and Riyazuddin 2008; Routroy et al. 2013). Based on the quantitative chemical analysis data generated, such statistical analyses to the likes of correlation analysis, hierarchical cluster analysis, principal component and factor analysis were performed using software program StatistiXL.

Results and discussions

Major ion chemistry and spatial distribution

The pH values of the groundwater varies from 7.00 to 8.40 (in post-monsoon) with an average of 7.50, and 6.4–8.4 (in pre-monsoon) with an average of 7.1, which indicates that water is almost neutral in nature. The average concentration of Total Dissolved Solids (TDS) ranged from 244.3 (post-monsoon) to 249.0 (pre-monsoon) mg/l in the study area. Normally TDS in water may originate from natural sources and sewage discharges. The electrical conductivity (EC) in the study area varies from 90.00 to 300.00 (in post-monsoon) and 55.00 to 590.00 (in pre-monsoon) with an average of 212.7 (post-monsoon) and 556.2 (pre-monsoon) $\mu\text{S}/\text{cm}$ at 25 °C. The Total hardness in water is derived from the solution of carbon dioxide released by bacterial action in the soil, in percolating rain water. Low pH conditions develop and lead to the dissolution of insoluble carbonates in the soil and in limestone formations to convert them into soluble bicarbonates. Impurities in limestone, such as sulfates, chlorides and silicates, become exposed to the solvent action of water as the carbonates are dissolved so that they also pass into solution. The total hardness (TH) of water is a measure of mainly calcium carbonate and magnesium carbonate dissolved in groundwater. The general acceptance level of hardness is 300 mg/l, although WHO has set an allowable limit of 600 mg/l. The total hardness in the study area ranges between 55 and 365 mg/l in post-monsoon while in pre-monsoon it ranges between 48 and 384 mg/l.

Calcium concentration ranged from 12.6 to 109.2 mg/l in post-monsoon and 6.7–95.8 mg/l in pre-monsoon periods. The magnesium concentration ranges between 3.2–42.5 mg/l in post-monsoon and 7.6–35.3 mg/l. in pre-

monsoon. Sodium concentration in water varies from 7.00 to 58.3 mg/l with an average of 26.04 mg/l in post-monsoon and 8.0–69.0 mg/l with an average of 30.5 mg/l in pre-monsoon period. Iron is an essential element in the human body and is required physiologically on various aspects (Moore 1973). Although iron has little concern as a health hazard, it is still considered as a nuisance in excessive quantities (Dart 1974). It causes staining of clothes and utensils. It is also not suitable for processing of food, beverages, dyeing, bleaching, etc. The concentration limits of iron in drinking water ranges between 0.3 mg/l (maximum acceptable) and 1.0 mg/l (maximum allowable). Iron concentrations of Suri I and II blocks range between 0.0 and 1.8 mg/l in post-monsoon and 0.0 to 6.1 mg/l in pre-monsoon. At Bhagabanbati Primary School (Location no. 13 located in the eastern part of the study area) the iron concentration is above the desirable limit (0.30 mg/l) during both post and pre-monsoon. High iron concentration affects the taste of water, has adverse effects on domestic uses and promotes growth of iron bacteria. Measures should be taken before consumption by installation of iron removing plants.

Bicarbonate ion varies from 48.80 to 1073.60 and 48.8 to 292.8 mg/l in post and pre-monsoon respectively. Chloride concentration in groundwater samples in the study area ranged from 15.0 to 124.96 and 15.0 to 274.9 mg/l in post and pre-monsoon respectively. WHO has set standards of 200–500 mg/l for chloride in drinking water. Too much of chloride leads to bad taste in water and also chloride ion combines with the Na (that is being derived from the weathering of granitic terrains) and forms NaCl, whose excess presence in water makes it saline and unfit for drinking and irrigation purposes. Here too, as exhibited by mean values in Table 1, the chloride value decreases during post-monsoon.

The sulfate ion causes no particular harmful effects on soils or plants; however, it contributes in increasing the salinity in the soil solution. Sulfate ion varied from 0.43 to 48.76 mg/l during post-monsoon and 0.1 to 56.7 mg/l in pre-monsoon periods.

The range of presence of each of the sixteen parameters analyzed, along with their average and standard deviation values for both the sampling sessions have been presented in Table 2. Based on the analyses performed on the quantitative data, the following results have been reported.

Correlation analysis

The correlation analysis study involving statistical calculations was devised by Pearson (1896). Based on the value of correlation coefficient ‘*r*’, the correlation between two parameters plotted on a XY scatter diagram can be termed as positive or negative. Correlation analysis is a common

Table 1 Quantitative chemical analysis results of post and pre-monsoon period

Parameter	Post-monsoon						Pre-monsoon					
	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
pH	7.0	8.4	7.5	0.34	0.37	0.09	6.4	8.4	7.1	0.43	0.87	1.52
TDS	74.0	622.0	244.3	109.2	1.41	4.81	55.0	590.0	249.0	125.9	0.76	1.04
EC	90.0	300.0	212.7	48.65	-0.54	1.01	150.0	1200.0	556.2	256.6	1.05	1.80
TA	50.0	880.0	530.0	204.24	-0.54	0.15	40.0	320.0	158.8	66.7	0.42	0.10
TH	55.0	365.0	206.9	70.5	-0.07	0.17	48.0	384.0	148.3	68.7	1.66	4.52
Ca ²⁺	12.6	109.2	46.8	21.1	0.97	1.81	6.7	95.8	29.3	18.4	2.11	5.87
Mg ²⁺	3.2	42.5	21.9	8.0	0.06	1.12	7.6	35.3	18.3	7.9	0.36	-0.76
Na ⁺	7.0	58.3	26.04	15.4	0.80	-0.38	8.0	69.0	30.5	17.8	0.81	-0.20
K ⁺	0.7	10.0	2.4	2.1	2.22	6.49	0.2	2.5	0.6	0.54	2.00	4.69
Fe ²⁺	0.0	1.8	0.12	0.35	4.95	24.94	0.0	6.1	0.47	1.21	4.39	20.33
CO ₃ ²⁻	0.0	204.0	56.8	75.5	1.07	-0.48	0.0	36.0	6.0	9.75	1.70	2.52
HCO ₃ ⁻	48.8	1073.6	531.2	222.6	-0.36	1.22	48.8	390.4	181.6	76.2	0.75	1.10
Cl ⁻	15.0	129.9	39.4	31.15	1.99	3.52	15.0	274.9	49.0	58.65	2.93	9.16
SO ₄ ²⁻	0.43	48.8	8.4	13.43	-0.36	1.22	0.1	56.7	8.7	15.03	2.43	4.94
PO ₄ ³⁻	0.0	0.04	0.01	0.01	1.23	1.59	0.0	4.8	0.3	0.93	4.80	23.80
F ⁻	0.21	2.38	0.6	0.50	2.36	6.34	0.24	2.84	0.73	0.59	2.34	6.44

Table 2 Correlation coefficient matrix for post-monsoon session

Parameters	PH	TDS	EC	TA	TH	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	Fe ²⁺	CO ₃ ²⁻	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻	PO ₄ ³⁻	F ⁻
pH	1.00															
TDS	-0.42	1.00														
EC	-0.36	0.53	1.00													
TA	-	0.55	0.59	1.00												
TH	-0.66	0.84	0.61	0.49	1.00											
Ca ²⁺	-0.71	0.85	0.45	0.43	0.92	1.00										
Mg ²⁺	-0.28	0.45	0.58	0.36	0.69	0.33	1.00									
Na ⁺	-	0.61	0.58	0.57	0.38	0.31	0.19	1.00								
K ⁺	-0.08	0.43	0.48	0.40	0.25	0.26	0.17	0.46	1.00							
Fe ²⁺	-0.15	0.11	0.13	-0.15	0.19	0.20	0.03	-	0.13	1.00						
CO ₃ ²⁻	0.19	0.04	0.43	0.47	-0.05	0.02	-0.15	0.47	0.38	-0.08	1.00					
HCO ₃ ⁻	-0.13	0.58	0.37	0.79	0.66	0.50	0.51	0.31	0.19	-0.11	-0.16	1.00				
Cl ⁻	-0.48	0.70	0.35	0.00	0.52	0.61	0.20	0.42	0.32	-0.02	-0.13	0.09	1.00			
SO ₄ ²⁻	-0.47	0.75	0.42	0.13	0.51	0.59	0.22	0.56	0.58	-0.08	0.01	0.14	0.85	1.00		
PO ₄ ³⁻	0.14	-0.16	-0.31	-0.24	-0.11	-0.14	0.11	-0.08	-0.09	0.07	-0.15	-0.17	-0.12	-0.02	1.00	
F ⁻	0.41	0.03	0.14	0.38	-0.14	-0.21	0.04	0.46	0.03	-0.10	0.22	0.27	-0.07	-0.16	-0.17	1.00

and useful statistical tool for water quality studies indicating which ions control the water chemistry (Box et al. 1978; Chapman 1996; Shrivastava and Patil 2002; Zeng et al. 2005). It is simply a measure to exhibit how well one variable predicts the other (Kurumbein and Graybill 1965). In the present study relationship of fluoride has been worked out with rest of the analyzed water quality parameters and the following observations have been made from the trend analysis graphs for both post-monsoon

(Fig. 2) and pre-monsoon (Fig. 3) sessions. Tables 3 and 4 present the correlation coefficient matrices prepared for the post-monsoon and pre-monsoon sessions, respectively.

In the post-monsoon session, fluoride is observed to share a positive correlation with pH ($r = 0.41$), total alkalinity ($r = 0.38$), sodium ($r = 0.46$), carbonate and bicarbonate ($r > 0.2$). In case of magnesium, potassium and TDS the 'r' value is positive but very close to zero indicating a not so strong correlation of these parameters

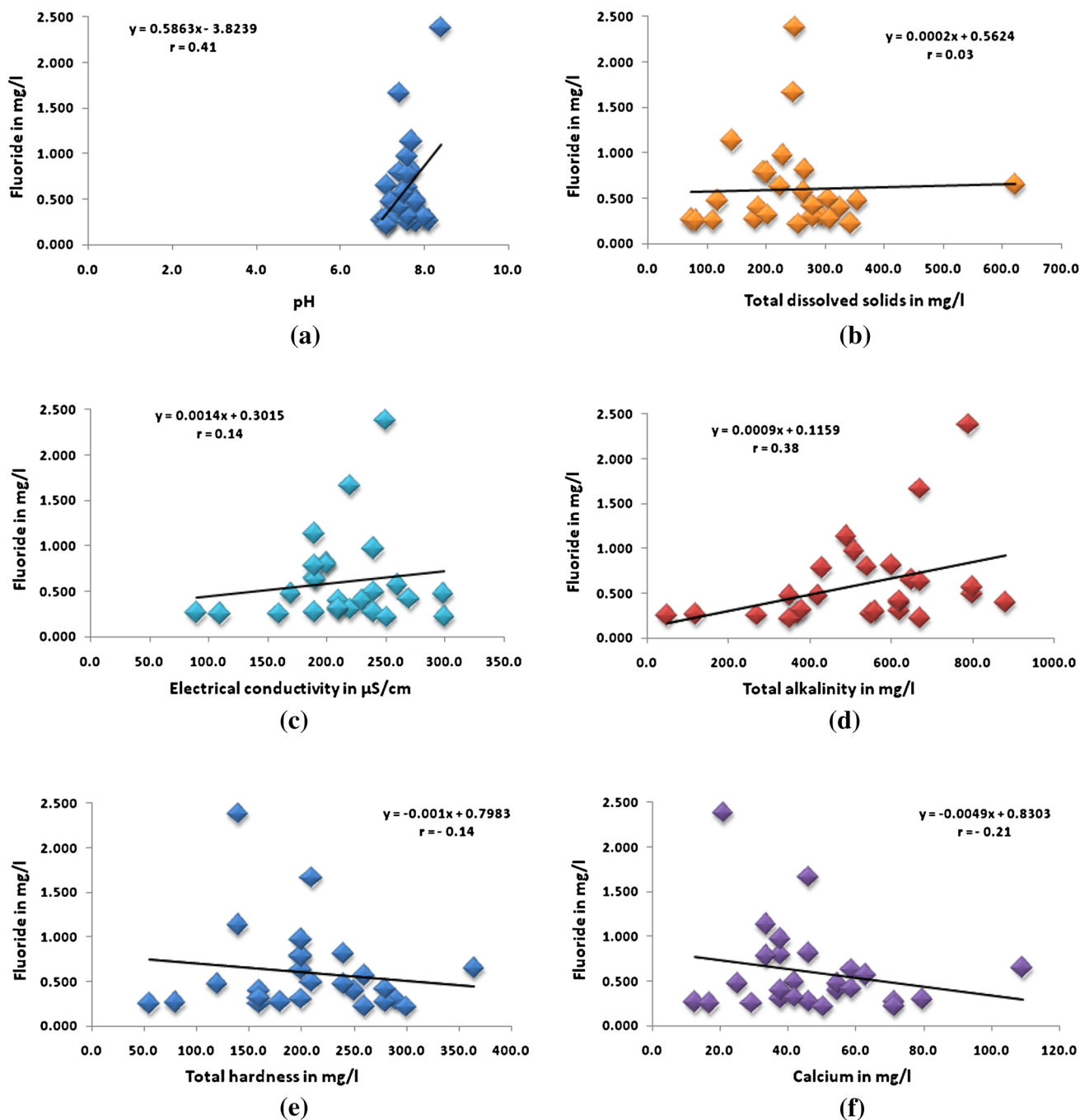


Fig. 2 Correlation Analysis graphs presenting correlation trends of fluoride with the other analyzed water quality parameters in post-monsoon session

with fluoride. Correlation of fluoride with calcium, total hardness, sulfate and phosphate is found to be negative. Bicarbonate shares positive correlations with alkalinity ($r = 0.79$), hardness ($r = 0.66$), calcium and magnesium ($r > 0.5$). The statistical significance of these correlations has been tested at p value = 0.05; and all have been found to be significant.

In the pre-monsoon session, fluoride shares fairly positive correlation with total alkalinity, sodium ($r = 0.39$),

carbonate ($r = 0.67$) and bicarbonate. In case of correlation of fluoride with pH, although parameters are positively related, but the ' r ' value is very close to zero, a trend opposite than that observed during post-monsoon. With almost all the other analyzed parameters, fluoride shares a negative relationship. The correlation of fluoride with potassium, iron and phosphate is found to be statistically insignificant during the pre-monsoon session. In the pre-monsoon session bicarbonate shares a strong but

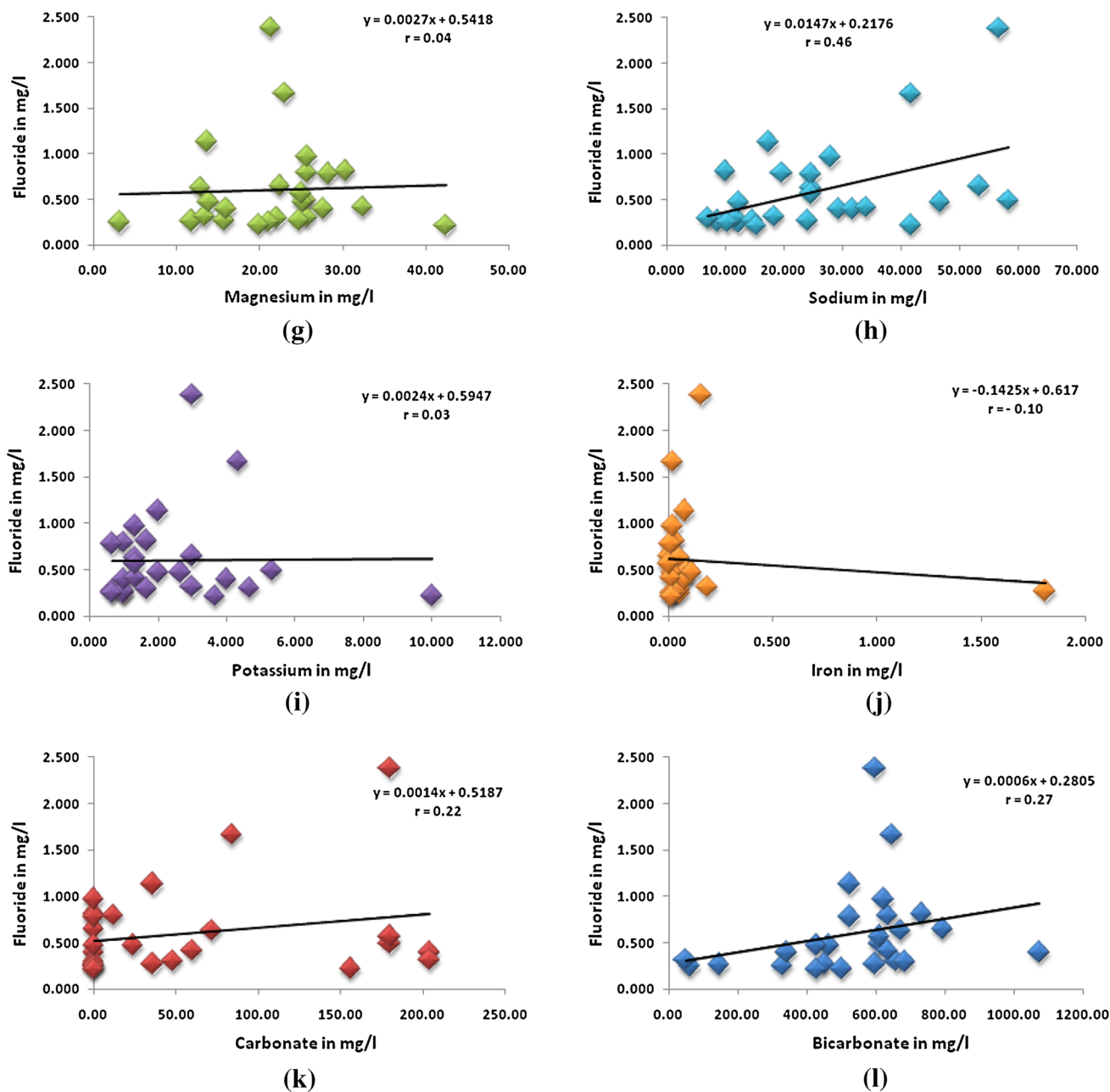


Fig. 2 continued

statistically insignificant correlation ($r = 0.97$) with alkalinity whereas with hardness it is the opposite ($r = 0.16$). Both calcium and magnesium show strong correlations with total hardness ($r = 0.91$ and $r = 0.82$ respectively) indicating that Ca and Mg regulate the alkalinity character of groundwater. Iron shares a strong correlation with hardness ($r = 0.65$) as opposed to that during post-monsoon.

The box and whisker plots (Fig. 4a, b), representing a comparison of distribution of the analyzed water quality parameters based on maximum, minimum and quartile

values, primarily indicate that groundwater in the study area is comparatively more alkaline with much lower conductivity during post-monsoon than that during pre-monsoon.

Hierarchical cluster analysis

Multivariate statistical analysis methods have been put to use in various water quality and hydrogeochemical studies. With the help of these analytical tools various forms of grouping studies can be performed in case of water samples

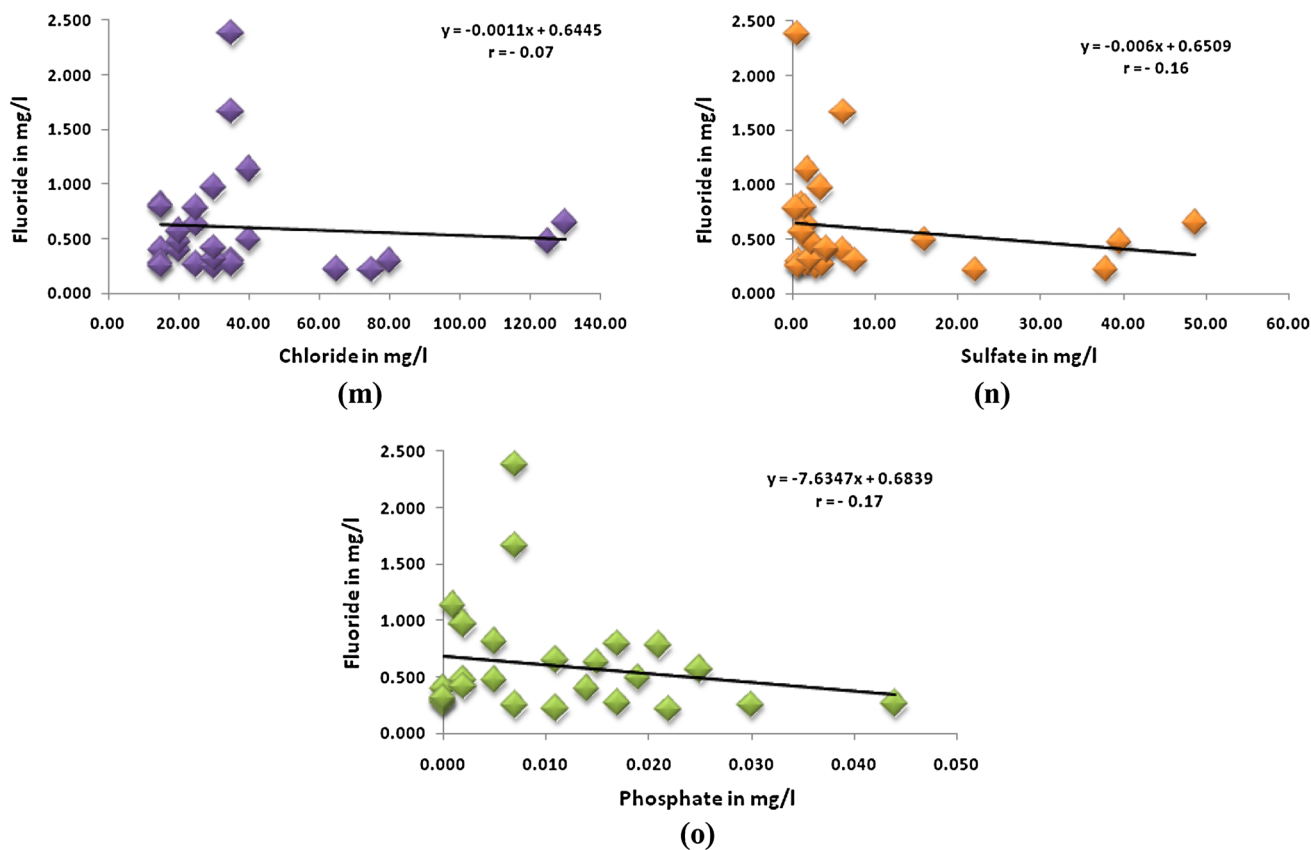


Fig. 2 continued

collected from various sources. The multiple water quality parameters analyzed can be clustered based on their interdependency (Routroy et al. 2013) or sampling locations can be clustered based on chemical similarity in samples collected from these locations (Singh et al. 2005; Shrestha and Kazama 2006). In the present study, the StatistiXL software has been used to run the Hierarchical cluster analysis function to group the multiple water quality parameters in clusters and generate dendrograms for the post-monsoon (Fig. 5a) and pre-monsoon (Fig. 5b) sessions. The dendrograms generated demonstrate high homogeneity within the clusters and high heterogeneity between the clusters (Sarbu and Pop 2005). The analyzed values of water quality parameters tested have been standardized (using the mean and standard deviation values) to lessen the disparity in range of the quantitative values of each parameter. The linkage distance amongst the parameters was measured in terms of squared euclidean distance (and the clustering method chosen was Ward's method).

In the post-monsoon session, the closest clusters generated were that of calcium-total hardness and sulfate-chloride. The alkalinity or salinity character of water is

depicted by the sodium-carbonate and bicarbonate-total alkalinity clusters which further group with the magnesium—conductivity cluster. Fluoride forms cluster with pH and is primarily linked to the iron-phosphate cluster. On a larger scale three major clusters can be observed from top—the first one comprising seven parameters (Mg^{2+} , EC, HCO_3^- , TA, CO_3^{2-} , Na^+ , K^+), second one comprising five parameters (Ca^{2+} , TH, TDS, SO_4^{2-} , Cl^-) and the third one of four parameters (PO_4^{3-} , Fe^{2+} , F^- , pH).

In the pre-monsoon session, the hardness factors like calcium, magnesium, total hardness are clubbed with TDS and conductivity. Fluoride is clustered with the alkaline parameters such as pH, carbonate, bicarbonate and total alkalinity. Sulfate and phosphate form another distinct cluster in the pre-monsoon session demonstrating the factor of anthropogenic contamination of groundwater. In form of larger clusters three major ones can be observed from top—comprising of four (SO_4^{2-} , K^+ , Na^+ , PO_4^{3-}), seven (Ca^{2+} , TH, Cl^- , Fe^{2+} , TDS, EC, Mg^{2+}) and five parameters (F^- , CO_3^{2-} , pH, HCO_3^- , TA).

The major variation observed in clustering patterns of both sessions is mainly related to fluoride. In post-monsoon

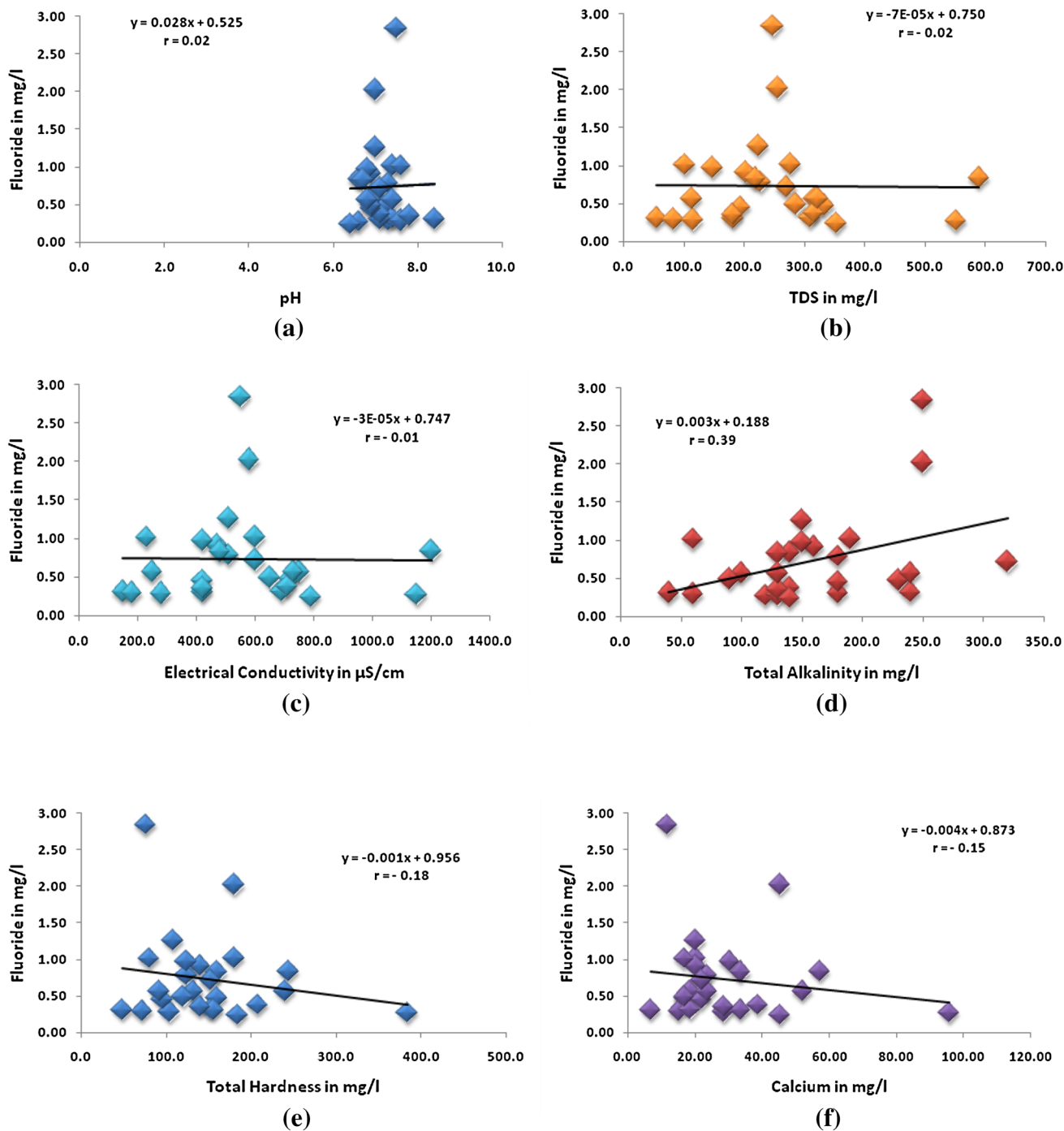


Fig. 3 Correlation Analysis graphs presenting correlation trends of fluoride with the other analyzed water quality parameters in pre-monsoon session

session, fluoride contamination seems to be regulated by iron concentration and pH of groundwater whereas in pre-monsoon fluoride contamination in groundwater seems to be regulated primarily by the alkaline factors. Iron in the pre-monsoon session groups with hardness factors.

Lateritic soil and clay soils dominate parts of the present study area as depicted in the geological map. Hence localized presence of iron in excess amounts in groundwater has been observed both during post and pre-monsoon sessions. Being a divalent ion, Fe^{2+} also has potential to

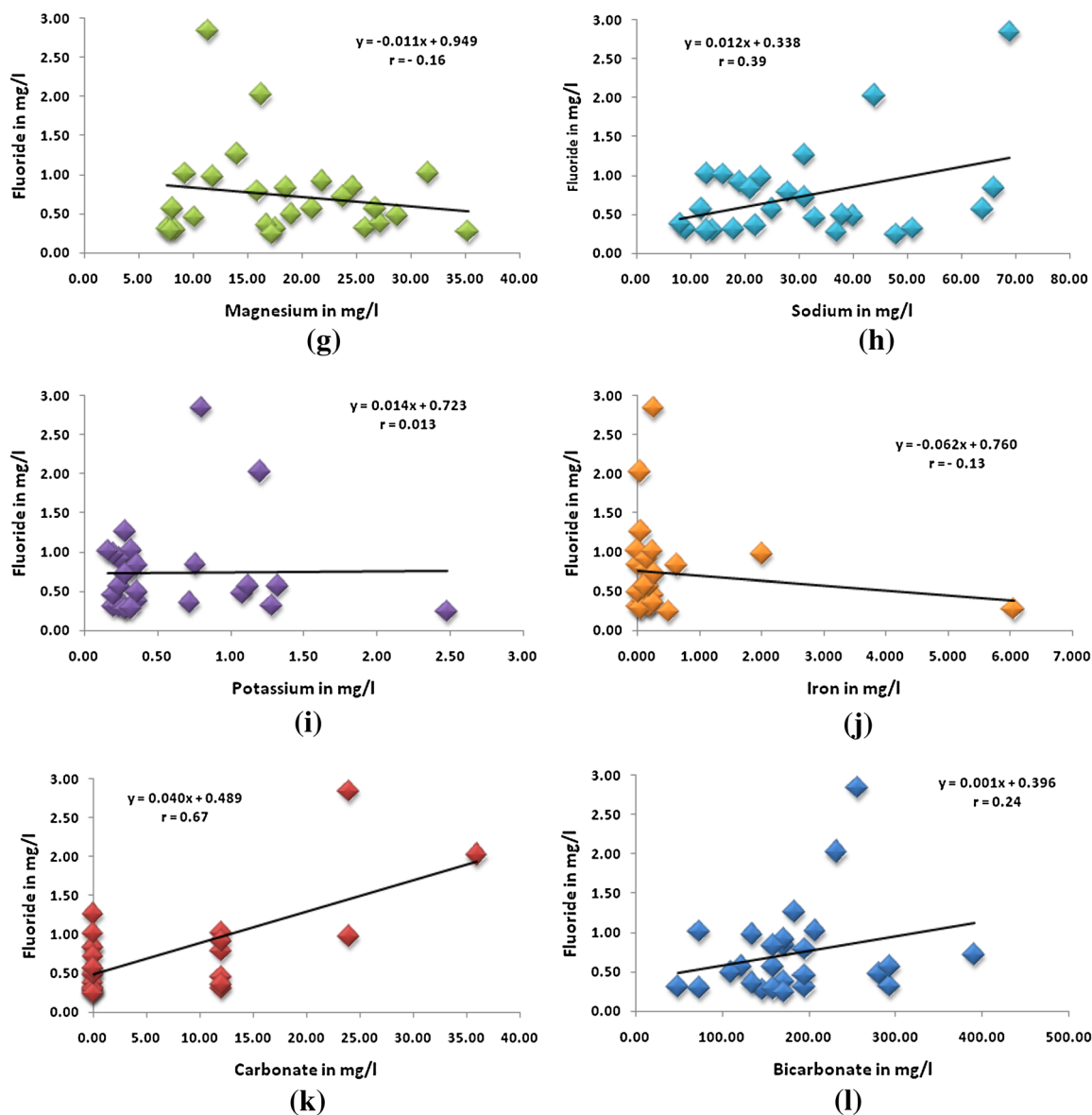


Fig. 3 continued

contribute to total hardness of water. Such a pattern is observed in the study area during pre-monsoon session, when iron shares a strong correlation with total hardness and calcium which is reflected in the pre-monsoon dendrogram as well, where iron is clustered with parameters defining the hardness character of groundwater.

Principal component and factor analysis

Principal component analysis (PCA) is a useful statistical tool for identifying patterns in a data set comprising of numerous parameters each with wide data range. Using

principal component analysis method in factor analysis we can compress the data by reducing the number of dimensions without much loss of information. Factor analysis (FA) related to principal component analysis, generates an unobserved or latent variable which mainly compiles variations in three or more observed variables. Thus factor analysis as well as principal component analysis methods are put to use in water quality evaluation studies, where various parameters are tested in search for particular causes leading to change in quality of water (Helena et al. 2000; Reghunath et al. 2002; Yidana 2010). Prior to factor analysis on the study area data, the skewness and kurtosis

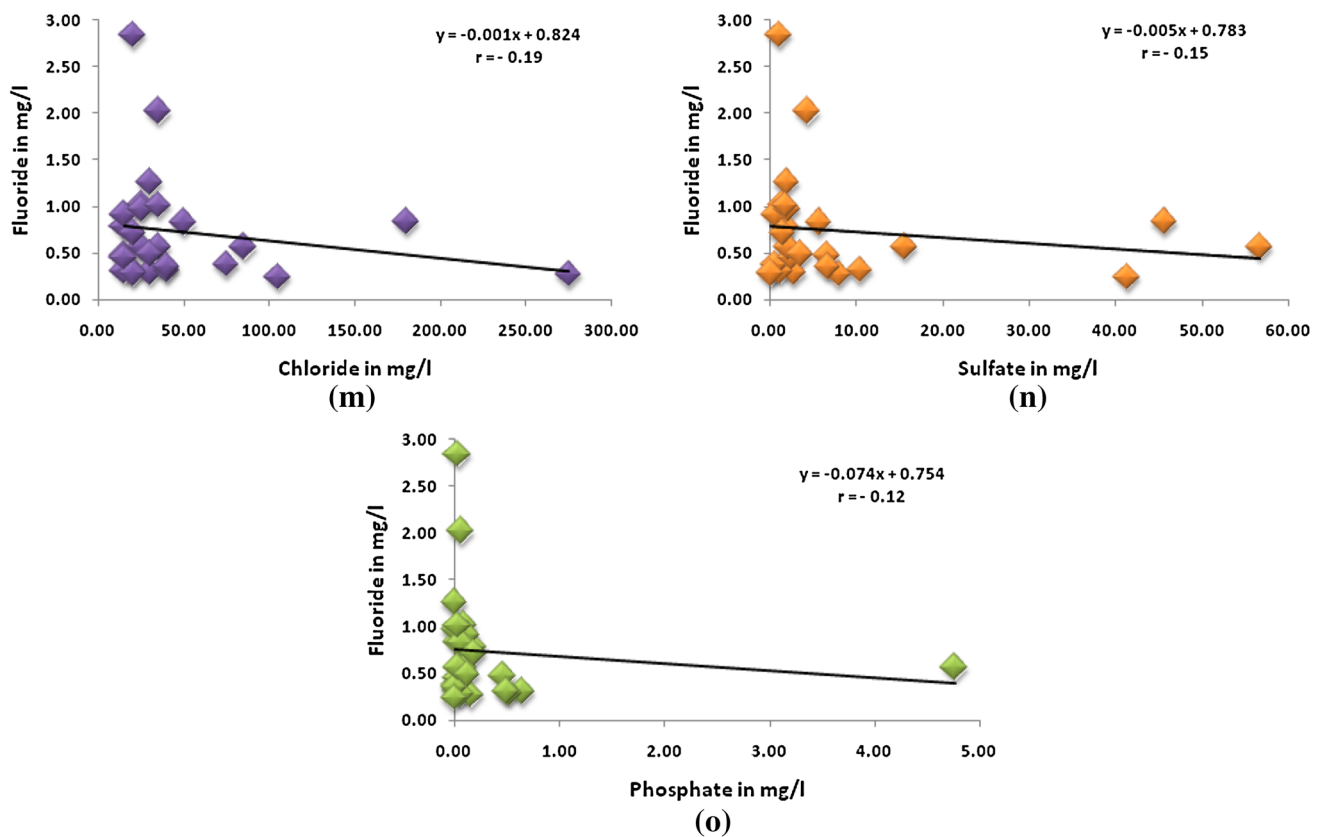


Fig. 3 continued

Table 3 Correlation coefficient matrix for pre-monsoon session

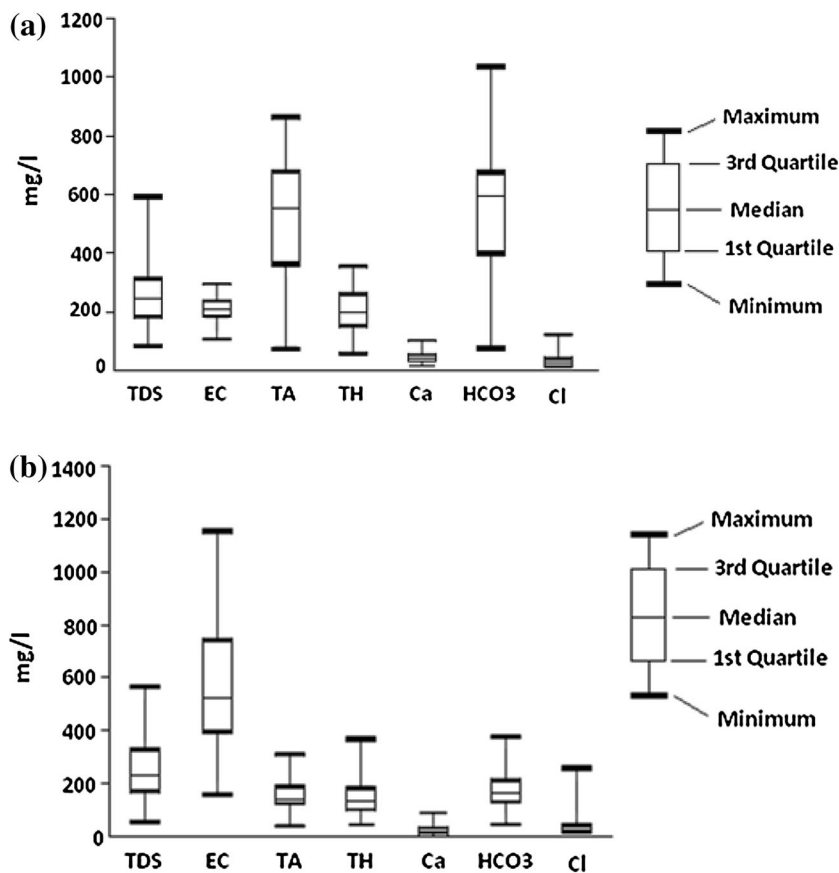
Parameters	PH	TDS	EC	TA	TH	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	Fe ²⁺	CO ₃ ²⁻	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻	PO ₄ ³⁻	F ⁻
pH	1.00															
TDS	-0.56	1.00														
EC	-0.56	0.99	1.00													
TA	-0.10	0.32	0.29	1.00												
TH	-0.57	0.84	0.84	0.12	1.00											
Ca ²⁺	-0.60	0.69	0.70	-0.07	0.91	1.00										
Mg ²⁺	-0.35	0.79	0.78	0.36	0.82	0.51	1.00									
Na ⁺	-0.26	0.64	0.62	0.53	0.21	0.16	0.23	1.00								
K ⁺	-0.24	0.44	0.40	0.31	0.23	0.18	0.23	0.60	1.00							
Fe ²⁺	-0.33	0.42	0.42	-0.11	0.65	0.72	0.36	0.05	-0.12	1.00						
CO ₃ ²⁻	0.02	-0.14	-0.16	0.37	-0.11	-0.03	-0.18	0.15	-0.01	-0.01	1.00					
HCO ₃ ⁻	-0.12	0.38	0.36	0.97	0.16	-0.07	0.43	0.53	0.33	-0.11	0.14	1.00				
Cl ⁻	-0.49	0.76	0.79	-0.18	0.86	0.90	0.53	0.29	0.18	0.73	-0.26	-0.13	1.00			
SO ₄ ²⁻	-0.44	0.57	0.56	-0.06	0.47	0.47	0.32	0.40	0.64	-0.03	-0.26	0.01	0.49	1.00		
PO ₄ ³⁻	-0.08	0.08	0.06	-0.12	0.23	0.20	0.20	-0.12	0.16	-0.07	-0.16	-0.09	0.09	0.61	1.00	
F ⁻	0.02	-0.01	-0.02	0.39	-0.18	-0.15	-0.16	0.39	0.01	-0.13	0.67	0.24	-0.19	-0.15	-0.12	1.00

of each parameter was calculated for post and pre-monsoon sessions. In the post-monsoon session, conductivity, total alkalinity, total hardness, bicarbonate and sulfate were found to be negatively skewed. All the other parameters

were positively skewed. The kurtosis coefficients were greater than zero (95 % confidence) for all parameters except sodium and carbonate. In case of pre-monsoon session, all the analyzed parameters were positively

Table 4 Varimax rotated factor loadings for post-monsoon session

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
pH	0.478	0.065	-0.294	-0.209	0.237	-0.611
TDS	-0.745	0.151	0.555	0.101	-0.097	-0.085
EC	-0.272	0.632	0.513	0.201	-0.178	0.040
TA	-0.014	0.478	0.714	-0.133	-0.227	-0.334
TH	-0.517	0.054	0.779	0.190	-0.035	0.240
Ca ²⁺	-0.629	0.039	0.547	0.221	-0.223	0.258
Mg ²⁺	-0.101	0.054	0.799	0.053	0.284	0.102
Na ⁺	-0.524	0.521	0.194	0.081	0.029	-0.546
K ⁺	-0.401	0.663	0.130	-0.278	0.088	0.043
Fe ²⁺	-0.007	-0.030	0.004	0.948	0.055	0.046
CO ₃ ²⁻	0.116	0.899	-0.113	0.028	-0.125	-0.161
HCO ₃ ⁻	-0.096	-0.085	0.877	-0.169	-0.168	-0.263
Cl ⁻	-0.945	-0.029	0.049	-0.002	-0.058	0.047
SO ₄ ²⁻	-0.934	0.213	0.088	-0.119	0.090	0.103
PO ₄ ³⁻	0.053	-0.118	-0.026	0.050	0.932	0.048
F ⁻	0.074	0.076	0.113	-0.003	-0.132	-0.889
Eigen value	6.110	2.677	1.785	1.203	1.109	1.001
% Variance	38.19	16.73	11.16	7.52	6.93	6.26
Cumulative % variance	38.19	54.92	66.08	73.59	80.52	86.78

Fig. 4 a Box and Whisker Plot for Post-monsoon session, **b** box and Whisker Plot for Pre-monsoon session

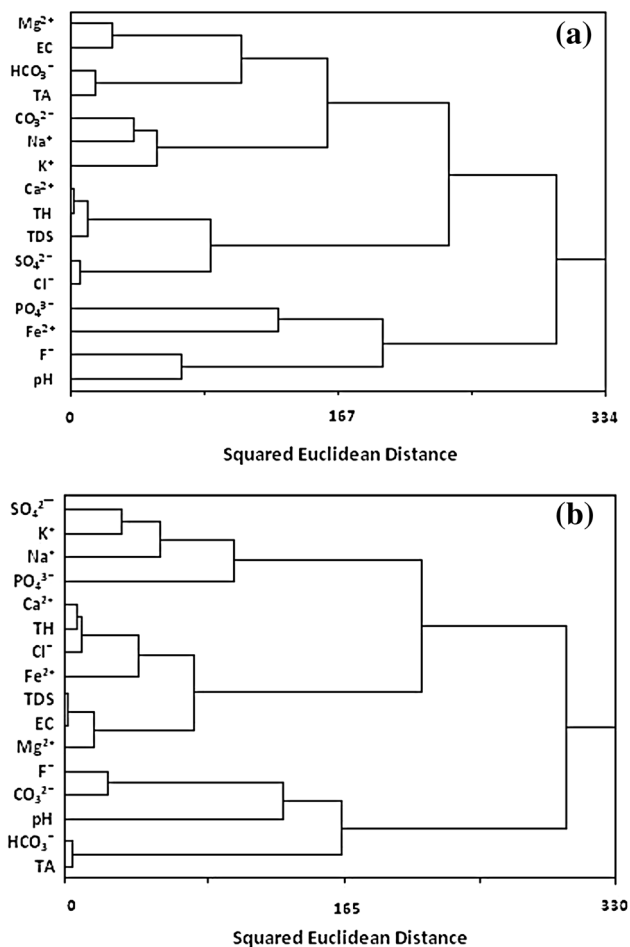


Fig. 5 **a** Dendrogram generated through cluster analysis for post-monsoon session, **b** dendrogram generated through cluster analysis for pre-monsoon session

skewed with kurtosis coefficients greater than zero (95 % confidence) except for magnesium and sodium. Principal component and factor analysis were performed for both sessions using the software Statistix following the Kaiser Normalization and Varimax Rotation procedures (Kaiser 1958).

The Varimax rotated factor loadings for the post-monsoon session have been presented in Table 5. Six factors having Eigen values greater than one were generated in the post-monsoon session and rotations were completed in nine iterations. Figure 6a presents the scree plot generated based on the Eigen values of the components. Factor 1 explains 38.19 % of the total variance and besides fluoride and carbonate which have very low positive loadings, only pH contributes a significant factor loading which possibly indicates that and not the other ions analyzed do not contribute significantly towards pH of the groundwater. In that

case, pH could possibly be controlled by biogenic factors (Helena et al. 2000). pH is negatively correlated to most of the analyzed parameters which is reflected by the negative factor loadings of the same. Factor 2 explains 16.73 % of the total variance and high positive loadings of total alkalinity, conductivity, sodium, potassium and carbonate are shown on this factor. High positive loading of potassium and carbonate and moderate loading value for sulfate demonstrates incidence of anthropogenic contamination sources. The study area is geologically composed of granite gneiss type rocks which are composed of alkali feldspars. Thus high positive loadings of sodium and potassium can also be attributed to dissolution of alkali feldspars. Factor 3 which explains 11.16 % of the total variance can be termed as the hardness factor as it shows high positive loadings of total hardness, calcium, magnesium and bicarbonate. High positive loading of alkalinity coupled with that of bicarbonate also indicates the salinity condition of the groundwater reserves. Moderately high scores of electrical conductivity and total hardness shown by factor 3 also imply chemical leaching of secondary minerals from the aquifer lithology (Routroy et al. 2013). Weathering of parent rock bodies and leaching of minerals leading to rock water interaction seem to be the factors controlling geochemistry of the groundwater. Factor 4 explains 7.52 % of the total variance and has very high positive factor loading for iron which can be attributed to the acidic character of the groundwater system. Factor 5 shows a high (>0.9) positive loading for phosphate and contributes 6.93 % to the total variance and can be associated to anthropogenic contamination factors. Factor 6 explains 6.26 % of the total variance and demonstrates moderate (>0.2) loadings for total hardness and calcium. Most of the other parameters have low or negative loadings thus indicating that this factor also partially defines the hardness quotient of groundwater.

The Varimax rotated factor loadings for the pre-monsoon session have been presented in Table 5. Five factors having Eigen values greater than one were generated in the pre-monsoon session and rotations were completed in seven iterations. Figure 6b presents the scree plot generated based on the Eigen values of the components. Factor 1 explains 40.74 % of the total variance. Similar to post-monsoon session, only pH has a high positive loading for factor 1. The relationship of pH with rest of the analyzed parameters is negative or close to zero in the pre-monsoon session as well, which has been reflected through the loading values. Factor 2 explains 19.85 % of the total variance and shows high positive loadings for total alkalinity, magnesium and bicarbonate, an average loading

Table 5 Varimax rotated factor loadings for pre-monsoon session

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
pH	0.610	-0.015	0.272	-0.084	-0.107
TDS	-0.776	0.363	-0.454	-0.103	0.011
EC	-0.786	0.339	-0.442	-0.118	-0.007
TA	0.020	0.906	-0.173	0.319	-0.069
TH	-0.936	0.206	-0.056	-0.113	0.205
Ca ²⁺	-0.937	-0.108	-0.074	0.014	0.163
Mg ²⁺	-0.649	0.588	-0.014	-0.257	0.203
Na ⁺	-0.217	0.342	-0.782	0.247	-0.236
K ⁺	-0.069	0.183	-0.826	-0.021	0.161
Fe ²⁺	-0.816	-0.118	0.256	0.029	-0.206
CO ₃ ²⁻	0.047	0.102	0.102	0.914	-0.028
HCO ₃ ⁻	0.010	0.941	-0.211	0.103	-0.067
Cl ⁻	-0.923	-0.167	-0.203	-0.170	-0.028
SO ₄ ²⁻	-0.339	-0.096	-0.697	-0.169	0.568
PO ₄ ³⁻	-0.066	-0.050	-0.060	-0.068	0.951
F ⁻	0.098	0.145	-0.141	0.865	-0.088
Eigen value	6.518	3.176	1.886	1.355	1.027
% Variance	40.74	19.85	11.79	8.47	6.42
Cumulative % variance	40.74	60.59	72.38	80.85	87.26

value for sodium and a negative loading score for calcium. This factor can thus be assumed to demonstrate majorly the salinity character of groundwater. Factor 3 explains 11.79 % and shows low factor loadings for pH and iron. Based on the pH values at sampling locations of the study area, factor 3 can be assumed to define the acidic nature of groundwater. Factor 4 shows high loadings for carbonate and fluoride values, moderate factor loadings for alkalinity and it explains 8.47 % of the total variance. Fluoride shows a positive correlation with carbonate in the pre-monsoon session which is reflected in the factor loading values as well. The particularly strong interrelationship of fluoride with carbonate can be attributed to the calcite–fluorite geochemistry in groundwater where fluoride bears a negative correlation with calcium (as reported in the study) and a positive one with carbonate and bicarbonate. A moderate score for alkalinity and positive correlation of fluoride with alkalinity also demonstrates that unlike post-monsoon, rise in fluoride level in pre-monsoon might be associated with the alkalinity character of groundwater. Factor 5 explains 6.42 % of the total variance. Very high factor loadings are demonstrated by sulfate and phosphate in this case. A bulk of the study area is used for agricultural purposes. Thus factor 5 can be associated with anthropogenic contamination of groundwater occurring due to continuous use of fertilizers and pesticides in this area dominated by agricultural practices.

Conclusion

Excess fluoride has been observed in two locations of the study area, both in the post-monsoon and pre-monsoon sessions and the concentration of fluoride is seen to rise in the pre-monsoon session. Just like excess presence of excess fluoride in drinking can lead to dental and skeletal damages in the human body in form of fluorosis, similarly consuming water containing less than 0.6 mg/l fluoride can also lead to dental caries, especially in children in the age bracket of 6–10. In the present study, during both sampling sessions, groundwater in many locations have been found to contain less than 0.6 mg/l of fluoride thus making those areas prone to dental caries. Total alkalinity, the maximum permissible limit for which is 600 mg/l, is found to be in excess in 42 % of the samples in post-monsoon, whereas in pre-monsoon alkalinity is within desirable limit at all locations. Iron is present in excess at one location during post-monsoon, where during pre-monsoon iron content further rises up to more than 6 mg/l. Rest of parameters analyzed were all found to be present within the desirable limit in groundwater. Cluster analysis groups the parameters mainly into hardness, salinity, and dissolved solids factors in both sessions. Correlation analysis and Hierarchical cluster analysis results in post-monsoon do not reveal a particularly strong correlation or dependence of fluoride with any other parameter besides pH. In the pre-

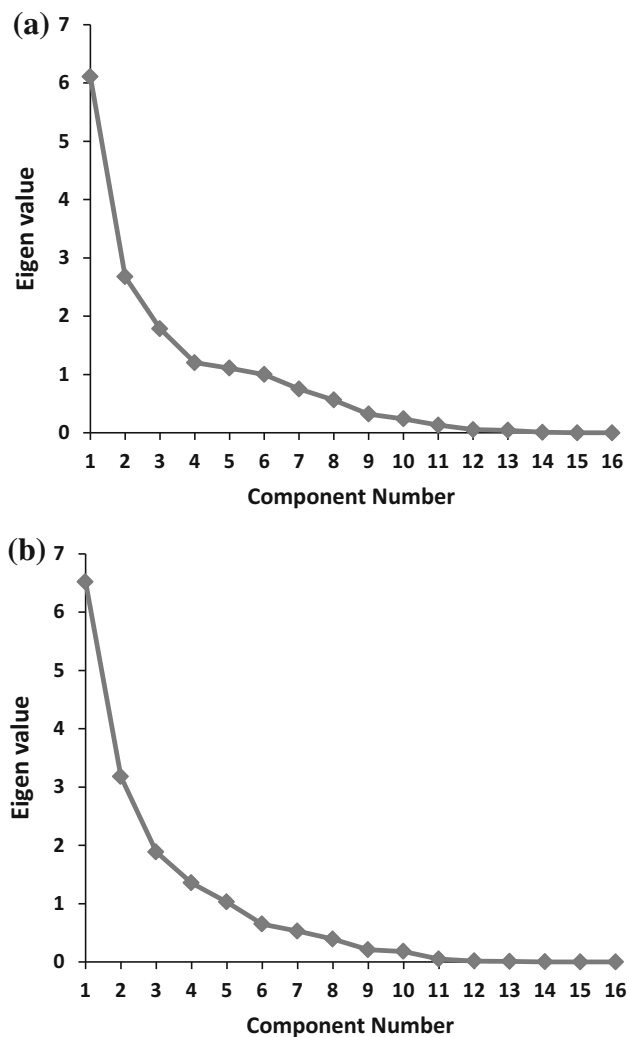


Fig. 6 **a** Scree plot generated through factor analysis for the post-monsoon session, **b** scree plot generated through factor analysis for the pre-monsoon session

monsoon session fluoride is well clustered with the alkalinity parameters of water and seems to be regulated by them. The factor analysis results in pre-monsoon demonstrate strong affinity of fluoride for carbonate and also indicate anthropogenic factors playing a role in contaminating groundwater. Not all parameters follow the same trend in both sampling sessions regarding increase or decrease in their concentration in groundwater. The inter-relationships amongst them are also observed to change on a seasonal basis depending on recharge and discharge of groundwater reserves.

Acknowledgments The author, S. Das is thankful to UGC, New Delhi and Jadavpur University, for providing her the UGC Research Fellowship in Science for Meritorious Student 2012–2013. The other author (S.K. Nag) acknowledges the financial support received from the Centre of Advanced Study (CAS-Phase V), Department of Geological Sciences, Jadavpur University in conducting the field work.

The authors gratefully acknowledge the support received from Dr. S. Gupta in analyzing the water quality parameters.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aghazadeh N, Mogaddam AA (2010) Assessment of groundwater quality and its suitability for drinking and agricultural uses in the Oshnavieh area, Northwest of Iran. *J Environ Prot* 1:30–40
- APHA (American Public Health Association) (1995) Standard methods for examination of water and waste water. American Public Health Association, American Water Works Association and Water Pollution Control Federation, Washington
- Bajpayee S, Das R, Ruj B, Adhikari K, Chatterjee PK (2012) Assessment by multivariate statistical analysis of ground water geochemical data of Bankura, India. *Int J Environ Sci* 3(2):870–880
- Banoeng-Yakubo B, Yidana SM, Nti E (2009) Hydrochemical modeling of groundwater using multivariate statistical methods—the Volta Region, Ghana. *KSCE J Civil Eng* 13(1):55–63
- Box GFP, Hunter WG, Hunter JS (1978) Statistics for experiments. An introduction to design data analysis and model building. Wiley, Toronto, p 653
- Briz-Kishor BH, Murali G (1992) Factor analysis for revealing hydrochemical characteristics of a watershed. *Environ Geol Water Sci* 19(1):3–9
- Chapman D (Ed) (1996) On the behalf of UNESCO, WHO. UXEP. Water quality assessments—a guide to use biota, sediments and water in environmental monitoring. F & F Spoil, London, Chapter 9
- Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353:294–313
- Dart FJ (1974) The hazard of iron. Water and Pollution Control, Ottawa
- Domenico PA, Schwartz FW (1990) Physical and chemical hydrogeology. Wiley, New York, p 824
- Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs
- Guler C, Thyne GD (2004) Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA. *J Hydrol* 285:177–198
- Güler C, Thyne G, McCray JE, Turner AK (2002) Evaluation and graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology* 10:455–474
- Gummadi S, Swarnalatha G, Vishnuvardhan Z, Harika D (2014) Statistical analysis of the groundwater samples from Bapatla Mandal, Guntur District, Andhra Pradesh, India. *IOSR J Environ Sci Toxicol Food Technol* 8(1):27–32
- Gupta S, Banerjee S (2011) Fluoride in crops and dietary intake in a fluoride-endemic area of West Bengal. Research report, July–September 2011. *Fluoride* 44(3):153–157
- Gupta S, Banerjee S, Saha R, Dutta JK, Mondal N (2006) Fluoride geochemistry of groundwater in Nalhati-1 Block of the Birbhum District, West Bengal, India. *Res Rep Fluoride* 39(4):318–320

- Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000) Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res* 34(3):807–816
- Helsrup T, Jorgensen NO, Banoeng-Yakubo B (2007) Investigation of hydrochemical characteristics of groundwater from Cretaceous-Eocene limestone aquifers in southern Ghana and Togo using hierarchical cluster analysis. *Hydrogeology* 15:977–989
- Hossien MT (2004) Hydrochemical evaluation of groundwater in the Blue Nile Basin, eastern Sudan, using conventional and multivariate techniques. *Hydrogeol J* 12:144–158
- Jalali M (2006) Chemical characteristics of groundwater in parts of mountainous region, Alvand, Hamadan, Iran. *Environ Geol* 51:433–446
- Jothivenkatachalam K, Nithya A, Chandra Mohan S (2010) Correlation analysis of drinking water quality in and around Perur block of Coimbatore District, Tamil Nadu, India. *Rasayan J Chem* 3(4):649–654
- Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23:187–200
- Khazaei E, Stednick JD, Sanford WE, Warner JW (2006) Hydrochemical changes over time in the Zahedan aquifer, Iran. *Environmental Monitoring and Assessment*
- Kortatsi BK (2007) Hydrochemical framework of groundwater in the Ankobra Basin, Ghana. *Aquat Geochem* 13(1):41–74
- Kortatsi BK, Tay CK, Anornu G, Hayford E, Dartey G (2008) Hydrogeochemical evaluation of groundwater in the lower Offin basin, Ghana. *Environ Geol* 53:1651–1662
- Kurumbain WC, Graybill FA (1965) An introduction to statistical models in geology. McGraw-Hill, New York
- Machiwal D, Jha MK (2015) Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. *J Hydrol Reg Stud*. doi:10.1016/j.ejrh.2014.11.005
- Milovanovic M (2007) Water quality assessment and determination of pollution sources along the Axios/Vardar River, Southeast Europe. *Desalination* 213:159–173
- Moghaddam A, Najib M (2006) Hydrogeologic characteristics of the alluvial tuff aquifer of northern Sahand Mountain slopes, Tabriz, Iran. *Hydrogeol J* 14:1319–1329
- Moore CV (1973) Iron in: modern nutrition in health and disease. Philadelphia, Lea and Febiger, pp 297
- Mrazovac S, Miloradov MV (2011) Correlation of main physico-chemical parameters of some groundwater in northern Siberia. *J Geochem Explor* 108:176–182
- Mukherjee B, Rao MG, Karunakaran C (1969) Genesis of kaolin deposits of Birbhum, West Bengal, India. *Clay Miner* 8:161–170
- Nag SK (2014) Evaluation of hydrochemical parameters and quality assessment of the groundwater in Gangajalghati Block, Bankura District, West Bengal, India. *Arabian J Sci Eng* 39:5715–5727
- Nag SK, Ghosh P (2013) Variation in groundwater levels and water quality in Chhatna Block, Bankura District, West Bengal, a GIS approach. *J Geol Soc India* 81(2):261–280
- Nag SK, Ray S (2015) Hydrochemical evaluation of groundwater quality of Bankura I and II Blocks, Bankura District, West Bengal, India: emphasis on Irrigation and Domestic Utility. *Arabian J Sci Eng* 40:205–214
- Nelson JD, Ward RD (1981) Statistical considerations and sampling techniques for ground-water quality monitoring. *Groundwater* 19(6):617–625
- Pearson K (1896) Mathematical contributions to the theory of evolution, III. regression, heredity and panmixia. *Philos Trans R Soc Lond* 187:253–318
- Ramesh Kumar A, Riyazuddin P (2008) Application of chemometric techniques in the assessment of groundwater pollution in a suburban area of Chennai city, India. *Curr Sci* 94(8):235–242
- Reghunath R, Sreedhara Murthy TR, Raghavan BR (2002) The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Res* 36:2437–2442
- Routroy S, Harichandran R, Mohanty JK, Panda CR (2013) A statistical appraisal to hydrogeochemistry of fluoride contaminated ground water in Nayagarh District, Odisha. *J Geol Soc India* 81:350–360
- Sarbu C, Pop HF (2005) Principal component analysis versus fuzzy principal component analysis, a case study: the quality of Danube water (1985–1996). *Talanta* 6:1215–1220
- Schiavo MA, Havser S, Gusimano G, Gatto L (2006) Geochemical characterization of groundwater and sub-marine discharge in the southeastern Sicily. *Cont Shelf Res* 26(7):826–834
- Shrestha S, Kazama F (2006) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ Model Softw* 22:464–475
- Shrivastava VS, Patil PR (2002) Tapti river water pollution by industrial wastes: a statistical approach. *Nat Environ Pollut Technol* 1(3):279–283
- Singh PK, Malik A, Sinha S (2005) Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. *Anal Chim Acta* 538:355–374
- Sreedevi PD (2004) Groundwater quality of Pageru river basin, Cuddapah district, Andhra Pradesh, India. *J Geol Sci* 64:619–636
- Subramani T, Elango L, Dhamodarasamy SR (2005) Groundwater quality and its suitability for drinking and agricultural use in Chithar River Basin, Tamilnadu, India. *Environ Geo* 47:1099–1110
- Sunne Vazquez E, Sanchez Vila X, Carrera J (2005) Introductory review of specific factors influencing urban groundwater, an emerging branch of hydrogeology, with reference to Barcelona, Spain. *Hydrogeol J* 13:522–533
- Vasanthavignar M, Srinivasamoorthy K, Vijayaragavan K, Rajiv Ganthi R, Chidambaram S, Sarama VS, Anandhan P, Manivannan R, Vasudevan S (2010) Application of water quality index for groundwater quality assessment: Thirumanimuttar Sub-Basin, Tamilnadu, India. *Environ Monit Assess* 171(1–4):595–609
- WHO (2011) Guideline for drinking water quality, 4th edn. World Health Organization, Geneva
- Yidana SM (2008) Groundwater resources management for productive uses in the Afram Plains area, Ghana. Doctoral Dissertation, Montclair State University, USA
- Yidana SM (2010) Groundwater classification using multivariate statistical methods: Southern Ghana. *J Afr Earth Sci* 57:455–469
- Yidana SM, Yidana A (2009) Assessing groundwater quality using water quality index and multivariate statistical analysis—the Voltaian basin, Ghana. *J Environ Earth Sci*. doi:10.1007/s12665-009-0132-3
- Yidana SM, Ophori D, Banoeng-Yakubo B (2008a) Hydrogeological and hydrochemical characterization of the Voltaian Basin: the Afram Plains area, Ghana. *Environ Geol* 53:1213–1223
- Yidana SM, Ophori D, Banoeng-Yakubo B (2008b) Hydrochemical evaluation of the Voltaian system—the Afram Plains area, Ghana. *Environ Manage* 88:697–707
- Yidana SM, Ophori D, Banoeng-Yakubo B (2008c) Groundwater availability in the shallow aquifers of the southern Voltaian system: a simulation and chemical analysis. *Environ Geol* 55:1647–1657
- Yidana SM, Ophori D, Banoeng-Yakubo B (2008d) Groundwater quality evaluation for productive uses—the Afram Plains area, Ghana. *J Irrig Drain Eng* 134(2):222–227
- Zeng X, Rasmussen TC (2005) Multivariate statistical characterization of water quality in Lake Lanier, Georgia, USA. *J Environ Qual* 34:1980–1991