Identifying key processes in the hydrochemistry of a basin through the combined use of factor and regression models

Sandow Mark Yidana*, Bruce Banoeng-Yakubo
and Patrick Asamoah Sakyi

Department of Earth Science, University of Ghana, Legon, Accra, Ghana.
Corresponding author. e-mail: smyidana@ug.edu.gh

An innovative technique of measuring the intensities of major sources of variation in the hydrochemistry of (ground) water in a basin has been developed. This technique, which is based on the combination of R-mode factor and multiple regression analyses, can be used to measure the degrees of influence of the major sources of variation in the hydrochemistry without measuring the concentrations of the entire set of physico-chemical parameters which are often used to characterize water systems. R-mode factor analysis was applied to the data of 13 physico-chemical parameters and 50 samples in order to determine the major sources of variation in the hydrochemistry of some aquifers in the western region of Ghana. In this study, three sources of variation in the hydrochemistry were distinguished: the dissolution of chlorides and sulfates of the major cations, carbonate mineral dissolution, and silicate mineral weathering. Two key parameters were identified with each of the processes and multiple regression models were developed for each process. These models were tested and found to predict these processes quite accurately, and can be applied anywhere within the terrain. This technique can be reliably applied in areas where logistical constraints limit water sampling for whole basin hydrochemical characterization. Q-mode hierarchical cluster analysis (HCA) applied to the data revealed three major groundwater associations distinguished on the basis of the major causes of variation in the hydrochemistry. The three groundwater types represent Na–HCO₃, Ca–HCO₃, and Na–Cl groundwater types. Silicate stability diagrams suggest that all these groundwater types are mainly stable in the kaolinite and montmorillonite fields suggesting moderately restricted flow conditions.

1. Introduction

Natural groundwater is generally free from contamination and is usually the better alternative to surface water, which is mostly contaminated by anthropogenic activities. However, in some cases, the quality of groundwater can deteriorate to such intolerable levels that interventions are needed prior to usage. Several factors work together to influence groundwater hydrochemistry and determine its suitability for most purposes. Principal among these factors is the interaction between the water and the host rock. Mineral weathering processes contribute significantly to the hydrochemistries of surface and groundwater reservoirs throughout the world, and have been the subject of most hydrochemical analysis in the literature (e.g., Apambire *et al* 1997; Abu-Jaber 2001; Wen *et al* 2005; Coetsiers and Walravens 2006; Cloutier *et al* 2008; Yidana 2008; Banoeng-Yakubo *et al* 2009). The degree to which rock/mineral weathering influences groundwater hydrochemistry is dependent on several factors including the residence time of groundwater in the host rock, the ambient

temperature and pH among others. For instance, silicate mineral weathering is a slow process which is facilitated by low pH and temperatures higher than 50°C. The process is slow and can take several years to register significant effects (Drever 1988; Appelo and Postma 2005), even under the desired conditions. Where the residence time is high, the interaction between the rocks and water endures for a long time, leading to a greater degree of interaction which leaches ions into groundwater.

The literature is also rich in instances where groundwater is contaminated by seawater effects and anthropogenic activities (e.g., Helstrup *et al* 2007; Trabelsi *et al* 2007; Antonellini *et al* 2008; Yidana *et al* 2010). Shallow unconfined aquifers are much more prone to contamination from surface activities than deep confined aquifers. This is because the shallow unconfined aquifer systems serve as open receptors of contaminants, and where surface activities are not regulated, the degree of contamination can be severe and remediation can cost a fortune. Seawater effects are common in areas where excessive groundwater withdrawal lowers the hydraulic head below sea level to encourage seawater upcoming. A typical case of seawater upcoming occurs in the shallow unconfined aquifer system in the Keta Basin, where excessive groundwater extraction for irrigation almost always encourages saline water intrusion (Jorgensen and Banoeng-Yakubo 2001; Helstrup *et al* 2007; Yidana *et al* 2010). Every useful hydrochemical study must consider the sum total of all these sources of contamination or ionic constitution of groundwater. In a bid to accurately unravel these common sources of variation in the hydrochemistry, various authors have adopted several innovative techniques, which have been adequately documented in the literature.

Factor and cluster analyses have been used widely with conventional graphical techniques to adequately characterize hydrochemical systems. Their uses have assisted in the resolution of serious environmental problems, and also provided clues to the understanding of groundwater flow regimes. Classical applications of these techniques in the past decade are contained in Meng and Maynard (2001); Güler *et al* (2002); Güler and Thyne (2004); Thyne *et al* (2004); Helstrup *et al* (2007); Yidana (2008); and Yidana *et al* (2008a, 2008b). In addition to facilitating an understanding of the processes influencing the quality of water in basins, the application of multivariate statistical methods assists in ranking the various processes influencing groundwater hydrochemistry in order of importance/significance. One of the most widely used multivariate statistical methods in water quality studies is factor analysis which is a data dimension reduction technique being used to assist in the

resolution of many problems in earth sciences and other fields. Factor analysis reduces a large number of data of so many variables into a smaller number of uncorrelated factors which are easier to handle and interpret. The method derives a subset of uncorrelated factors that explain the variance in hydrochemical datasets (Anazawa and Ohmori 2005). It is used to uncover the latent structure in the datasets of a set of hydrochemical variables (Yidana *et al* 2008a). In technical terms, common factor analysis represents the common variance of variables, excluding unique variance, and is thus a correlation-focused approach seeking to reproduce the inter-correlation among the variables. On the other hand, components (from principal components analysis, PCA) reflect both common and unique variance of the variables and may be seen as a variance-focused approach that reproduces both the total variable variance with all components as well as the correlations. PCA is far more commonly used than principal factor analysis (PFA) (Yidana *et al* 2008a).

Factor analysis makes it possible to rank hydrochemical processes in order of importance and also identify areas most impacted by a particular process in a basin. These advantages of factor analysis over the conventional graphical methods make these techniques very attractive to most environmental earth scientists and explain why these methods are widely used. Most often, factor analysis is not used in isolation, but is combined with several graphical techniques to provide meaning to hydrochemical analyses. Factor analysis, like any other statistical method, does not produce cause-and-effect relationships, but rather results in associations from which such cause-and-effect relationships can be inferred. The tool makes sense when used with a proper understanding of the system being studied. For instance, the interpretations of the hydrochemistry of a groundwater basin are only logical if the scientist understands the prevailing hydrogeology and the geochemistry of the aquifers.

Factor and principal component analyses are often used with cluster analyses in hydrochemical data analyses and in general groundwater studies. Cluster analysis groups variables into clusters or associations based on perceived similarities or dissimilarities in the variation of the datasets of these variables in space and/or in time. Hierarchical cluster analysis (HCA) is a particular type of cluster analysis whereby the parameters or cases (samples) are grouped into hierarchical classes based on similarities/dissimilarities discerned from the distribution of the datasets. Parameters in the same hierarchical cluster are more similar to each other than those in other clusters. When factor and cluster analyses are applied simultaneously to a

dataset, the result is a robust statistical model which assists in interpreting the datasets and provides leads to discerning the likely causes of the observed trends in the variables. Most of the applications of multivariate statistical tools in geo-environmental data analyses have involved the joint application of factor and cluster analyses.

Although these techniques have been widely used and universally accepted as reliable tools of hydrochemical analysis, their use have been largely centered on working on large sets of data to provide useful insights into the flow regime and sources of variation in the quality. Where spatial distribution of the factors responsible for the observed phenomenon is required, linear interpolation is often performed using the factor scores of each factor. In most cases, the product, an interpolation surface, produces unbelievably strange factor scores because the variograms do not fit well or the sampling was not dense enough. This leads to misleading interpretations. In this study, a new technique to circumvent this problem is being introduced. It relies on the combined use of factor and regression analyses to relate key hydrochemical parameters to the factors in the final factor model. The resulting models can be applied with data of a limited set of variables in order to determine the severity or otherwise of the major processes influencing groundwater hydrochemistry at any location within the basin. After the development of the factor model, key parameters were identified under each factor. Linear regression models were produced from the relationship between the factor scores and the data for the key parameters. These regression relationships were then used as key representatives of the severity or otherwise of the processes represented by each of the factors in the factor model. The methodology has been applied to a hydrochemical dataset from some Proterozoic aquifers in some parts of the western region of Ghana and finds use in areas where logistical constraints make it difficult to sample and analyze for all the major physico-chemical parameters everywhere.

2. The study area

The area lies between latitudes 4.8°–6.5°N and longitudes 1.75°–2.5°W and occupies approximately 8400 km² (figure 1). The study area is an important mining hub in the country as the rocks host gold and some other economic minerals whose exploitation sometimes lead to the contamination of surface water bodies. The need for sustainable supply of water of sufficient quality to meet growing needs has led to the development of groundwater resources at various depths to supply increasing populations with diverse water needs.

2.1 *Geography*

The detailed geography of the area is described by Dickson and Benneh (1995). The topography is characterized by ridges separated by valleys whose altitudes range between 46 and 76 m above mean sea level. Ground elevations as high as 200 m above mean sea level are common in the area especially among the ridges (Dickson and Benneh 1995). Due to the heavy rainfall and forest vegetation that prevent sheet erosion, the whole area is highly dissected and reduced to uniformly moderate relief with a gentle slope to the south (Dickson and Benneh 1980; Kortatsi 2007).

The climate is characterized by the wet semi-equatorial climatic conditions in the north and equatorial climatic conditions in the south, with double maxima in the rainfall distribution (Dickson and Benneh 1995). The annual rainfall is higher in the south (about 2000 mm) than the north (about 1600 mm) due to the differences in climatic zones. Throughout the study area, the annual peaks in rainfall distribution occur in June and October. The highest mean monthly temperature in the area is about 30°C and occurs during the months of March–April, whilst the lowest monthly temperature of 26°C occurs in August (Dickson and Benneh 1980; Kortatsi 2007).

2.2 *Geology and hydrogeology of the study area*

The geology of the area is copiously described in reports of Junner *et al* (1942) and Kesse (1985). It is underlained by rocks of Lower Proterozoic origin and comprises rocks of the Birimian and Tarkwaian basins which host much of the gold mineralization in Ghana (figure 2). The Birimian is subdivided into a Lower Birimian and an Upper Birimian with some intrusives. Over 60% of the area is underlained by rocks of the Lower Birimian, which consists of thick sequences of shale, phyllites, and greywacke, with tuffs and lavas. The Upper Birimian consists of rocks of volcanic and pyroclastic origin and underlies about 8% of the study area (Kortatsi 2007). Rock types in the Upper Birimian are greenstones, tuffs and sediments with minor bands of phyllite that include a zone of manganiferous phyllites bearing manganese ore (Junner *et al* 1942; Kesse 1985). The Birimian rocks have been intruded by sills and dykes of gabbro, norite, and meta-dolerite.

The Tarkwaian rocks unconformably overlie the Birimian rocks and consist of thick clastic sequence of arenaceous and argillaceous sediments with two well-defined locations of pebbly beds and conglomerates (Kesse 1985). Rocks of the Tarkwaian basin occupy about 15% of the study area (Kortatsi 2007), forming an elongated, narrow geosyncline
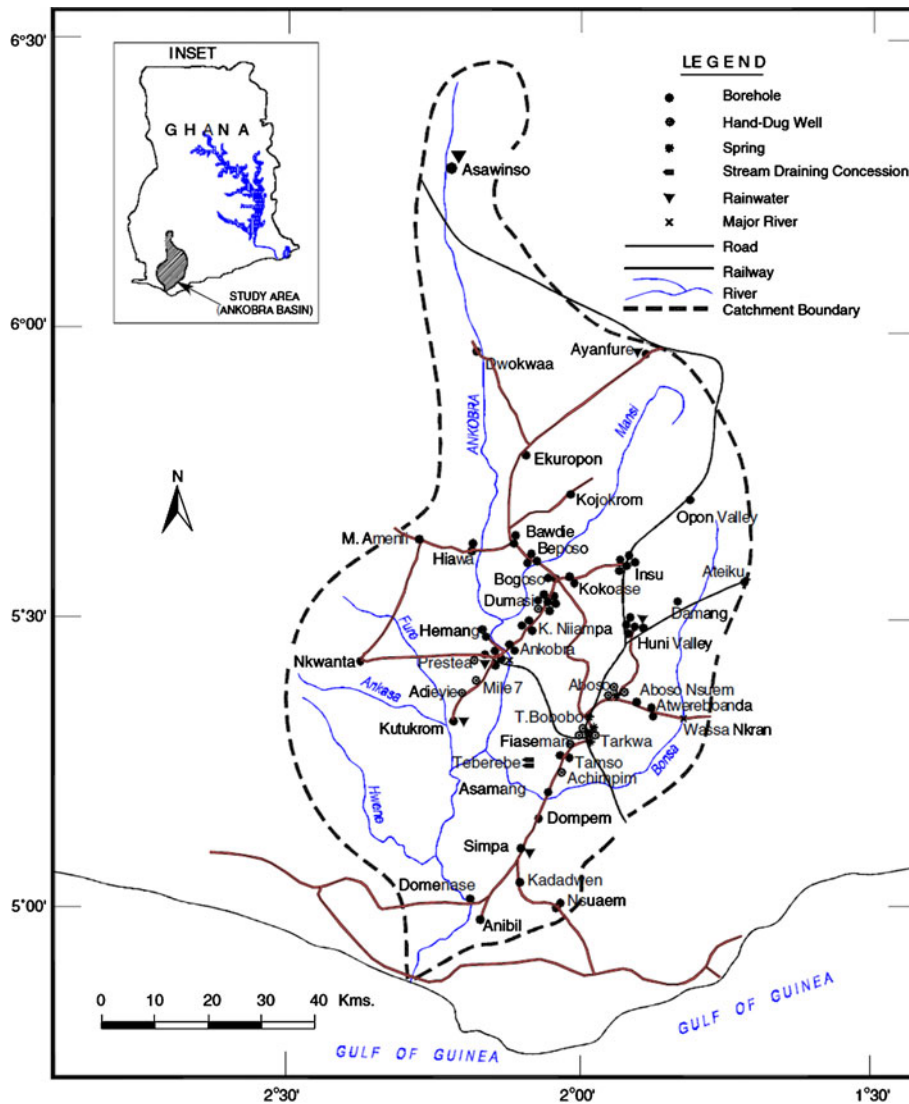
Figure 1. Map of the study area showing the major drainage networks in the area (modified from Kortatsi 2007).

which is about 16 km wide and stretches in a north-eastern to southwestern direction. There is a granitoid complex which occupies close to 10% of the study area. The remaining close to 3% of the study area is underlained by Eocene deposits which occur along the coast.

The hydrogeology of the area is controlled by secondary permeabilities created in the wake of fracturing and weathering of the rocks which have enhanced the storage and transmissive properties of the rocks. Tectonic lines of weaknesses created in the wake of several orogenic events have exposed the rocks to deep seated weathering whose degree varies with the lithology. The extent of weathering is highest in the Birimian where thickness of the weathered zones can range from 90 to 120 m (Banoeng-Yakubo *et al* 2009). This thickness of weathered material serves as prolific aquifers, providing groundwater to some communities living

within the neighbourhood of the basin. In areas where there is a significant proportion of the clay fraction, the transmissivity of the material is significantly reduced. The weathered zone in the granitoids and the other intrusive rocks is not as thick but ranges between 20 and 80 m (Kortatsi 2007). In the Tarkwaian basin, the weathered zone is even thinner and rarely exceeds 20 m (Junner *et al* 1942; Kesse 1985). Due to the high clay content of the top material, aquifers in the area are either confined or semi-confined.

3. Methodology

The data for this research was obtained through funding from the Danish International Development Agency (DANIDA) in the Department of Geology (now Department of Earth Science), and
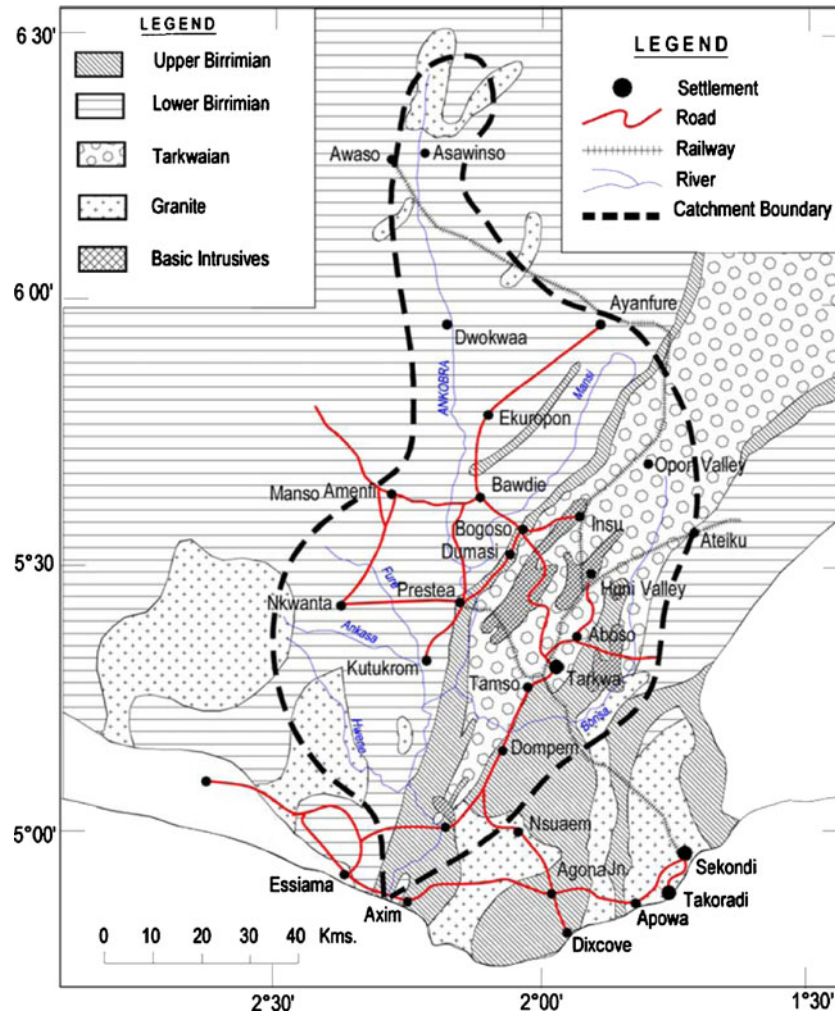
Figure 2. Geological map of the study area (modified from Kortatsi 2007).

was used by graduate students to complete their various theses researches. The sampling and analytical procedures that were strictly followed to obtain the data are described as follows. A total of 50 groundwater samples, comprising 12 samples from hand dug wells and 38 from boreholes were analyzed for the concentrations of the major hydrochemical parameters. Strict standard sampling and analytical protocols were adhered to, in order to obtain representative data from each of the sampling points. The samples were collected in 100 mL clean, acid-rinsed high density polyethylene bottles and preserved in clean ice chests for onward transmission to the laboratory for analyses. The samples for the major ions were acidified to pH < 2 in order to discourage bacterial action. On-site analyses were necessary for electrical conductivity (EC), pH, Eh, temperature, and alkalinity. This is because these parameters are likely to change on transit to the laboratory. pH, Eh, temperature, and EC were measured using a WTW multi-line P4

universal meter. For each of these parameters, final readings were taken when the wells and boreholes were pumped until the values were stabilized. At this point, it was safe to assume or conclude that the effects of drilling fluids and other influences on the parameters were almost negligible and that the readings represented the *in situ* groundwater conditions. Alkalinity values were obtained through titration on site. This was done using a HACH Digital Titrator Model 16900. Samples for the major ions were transported to the laboratory within 24 hours for the concentrations of the major ions to be determined. The major ions (Na^+, Ca^{2+}, Mg^{2+}, K^+, HCO_3^-, SO_4^{2-}, Cl^-) as well as NO_3^-, were analyzed using a double column Dionex DX-120 ion chromatograph.

The resulting data for the major ions were subjected to internal consistency tests. The concentrations of the major anions and cations were compared to ensure that these concentrations were within ±5% of each other, using the charge

balance error formula (Appelo and Postma 2005) equation (1).

$$\frac{\sum \text{cations} - \sum \text{anions}}{\sum \text{cations} + \sum \text{anions}} \times 100\% \qquad (1)$$

where the concentrations of the anions and cations were expressed in meq/L.

All the 50 samples passed the internal consistency test and were used for the next stage of the research.

Optimal results in multivariate statistical modelling require normal distribution and homoscedasticity. The variances of the parameters need to be compatible so that the interpretation of the final results makes logical sense. Where there are vast differences in the variances of the data of the parameters, the results will be influenced more by the parameters with the highest variances in the distribution of their dataset. In this respect, the data for each of the parameters were tested for normal distribution using normal probability plots from SPSS 17.0. Apart from pH whose values were close to normal distribution, all the other parameters strongly deviated from the requirement of normal distribution. There are various methods of transforming data to take the semblance of normal distribution and thus fit the requirements of optimal multivariate analyses. These include differencing, log-transformation, amongst others. In this study, the parameters were log-transformed to approach normal distribution, and the log-transformed data were standardized to their corresponding z-scores using equation (2). Data standardization does not affect its distribution but scales and shifts the data so that the variances are brought closer. This will ensure that the data for all the parameters are assigned equal weight in the ensuing analyses.

$$z = \frac{x - \bar{x}}{s}, \qquad (2)$$

where x, $\bar{x}$ and s are respectively the raw data, the mean, and standard deviation of the data for each parameter.

Where the data is highly variable in space, the variance of the data of a parameter can be significantly large. Standardization reduces the variances so that the parameters look similar in the distribution of their data. Factor analysis makes sense only when the distributions of the data of all the parameters look similar so that no single parameter gains significant advantage over all others due to its high variance. Where the variances are significantly different, the result of factor analysis may not reflect a fair contribution of all the parameters in the resulting model. This usually leads to faulty and unreliable interpretations.

The standard z-scores were then subjected to R-mode factor analysis with 'principal components' as the extraction method and Kaiser (1960) normalization. The factor matrix was varimax rotated in order to achieve maxima in the differentiation of the factors resulting from the factor analysis (Güler *et al* 2002; Cloutier *et al* 2008). Varimax rotation is the process of applying an orthogonal matrix to the factor matrix in order to enhance the differences among the resulting factors for easy interpretation of the results. The factor analysis was carried out repeatedly, each time taking into consideration the contribution of the individual parameters as well as the contribution of each of the factors to the final model. The degree to which each parameter contributed to the factor model was examined from their communalities on the communality table, which is one of the results of factor analysis. A cut-off was placed at a communality of 0.50, below which a parameter would be regarded as insignificant or inconsequential in the analysis. Such a variable (parameter) has to be dropped so that it does not cloud the results of the analysis. The communality of a variable in factor analysis is the sum of the squares of its loadings on all the factors in the final factor model. In the complete factor model, the communality of each variable is expected to be equal to 1. However, most often, the final factor model is not the same as the complete factor model since some of the factors in the complete model are redundant and do not usually represent unique processes controlling the distribution of the data of the parameter. In that case, it is possible that the communalities of some parameters in the selected final factor model may not be equal to 1. Although no single parameter has a communality of exactly 1.0 in a final factor model of less factors than the complete factor model, it is expected that each parameter in the model contributes significantly to the factor model, and those with low communalities need to be weeded out of the factor model in order to ensure the latter is not unduly distorted by such parameters. The communality of a parameter therefore represents its significance in the factor model.

In reality, it is possible to produce as many factors as there are variables in the model. The complete factor model should therefore have factors equal to the number of variables (parameters). A factor model produced in this way would have partially defeated the essence of factor analysis as a data dimension reduction methodology. Kaiser (1960) developed a methodology for cutting the number of factors in the final model to size based on merit. The Kaiser criterion requires that a worthy factor in the final model has an eigenvalue of at least 1.0. The significance of this is that a unique factor should be able to account for the variance

Table 1. *Statistical summary of the hydrochemical parameters used for this study.*

Parameter	Mean	Standard deviation	Minimum	Maximum
Temp	26.83	0.58	25.70	28.00
pH	5.68	0.54	4.00	6.80
EC(μS/cm)	246.37	145.84	37.00	670.00
TDS(mg/L)	147.86	87.49	22.20	402.00
Na^+(mg/L)	18.26	11.97	4.70	71.20
K^+(mg/L)	2.31	3.46	0.40	24.10
Ca^{2+}(mg/L)	18.18	13.85	0.30	67.90
Mg^{2+}(mg/L)	6.52	4.99	1.10	24.80
Cl^-(mg/L)	25.95	25.32	1.00	105.00
SO_4^{2-}(mg/L)	15.59	11.47	0.50	52.80
HCO_3^-(mg/L)	71.65	52.50	7.30	330.60
NO_3^-(mg/L)	1.89	2.05	0.00	8.50
SiO_2(mg/L)	22.57	17.85	0.00	90.60

of at least one variable (parameter). This criterion effectively reduces the number of factors in the final model to factors with eigenvalues greater than or equal to 1.0. Other methods including predetermining the total number of factors to be expected in the final factor model have commonly been in use.

In this study, the Kaiser (1960) criterion was used to select the factors to be included in the final factor model. The K^+ ion was eliminated from the analysis because it registered a low communality (<0.50) in the first run. Factor analysis was carried out using SPSS 17.0. When the final factor model was determined, factor scores were calculated for each of the factors in the resulting factor model. The factor scores were computed by the method of regression. In addition to the factor scores, key parameters were selected from the factor model to represent each of the factors. The selection of these key parameters was based on their loadings under each factor in the final factor model. The concentrations of these parameters were expressed in mmol/L except in the case of SiO_2 where it was expedient to use concentrations in mg/L. Regression relationships were then developed between these key parameters on one hand

Table 2(a). *Factor loadings matrix from R-mode factor analysis.*

Parameter	Factor 1	Factor 2	Factor 3
pH	0.010	0.769	0.245
EC	0.913	0.337	0.036
Na^+	0.933	0.086	−0.156
Ca^{2+}	0.194	0.872	0.016
Mg^{2+}	0.679	0.194	0.405
Cl^-	0.929	−0.086	0.068
SO_4^{2-}	0.705	−0.019	−0.168
HCO_3^-	0.046	0.904	0.219
NO_3^-	0.109	−0.298	−0.788
SiO_2	0.019	0.092	0.822

Table 2(b). *Variances explained by the various factors.*

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	3.909	39.088	39.088	3.909	39.088	39.088	3.578	35.779	35.779
2	2.615	26.145	65.233	2.615	26.145	65.233	2.433	24.326	60.105
3	1.115	11.145	76.378	1.115	11.145	76.378	1.627	16.273	76.378
4	0.776	7.759	84.137						
5	0.562	5.617	89.754						
6	0.471	4.712	94.466						
7	0.319	3.194	97.660						
8	0.148	1.484	99.143						
9	0.054	0.541	99.685						
10	0.032	0.315	100.000						

Extraction method: Principal component analysis.

and the factors they represent in the factor model on the other.

Q-mode HCA was applied to the z-scores of the parameters in order to determine the spatial relationships amongst the various samples or sampled locations. Squared Euclidean distances were used to measure the degree of similarity/dissimilarity amongst the parameters whilst the Ward's agglomeration technique was used to link the initial clusters. Unique clusters were sorted out of the dendrogram, which is the graphical result of the HCA. Statistical summaries of the measures of the parameters under each cluster were computed to assist in the description of the general groundwater flow regime and the general hydrochemical trends in the study area. Factor scores of the members of these clusters were used to plot scatter diagrams against the concentrations of the key parameters identified in the final factor model in order to identify any useful trends in the distribution of the data.

4. Results and discussions

4.1 *General hydrochemistry*

The statistical summaries of the hydrochemical data used for this study are presented in table 1. The temperature of the groundwater sampled is generally room temperature and ranges between $25.70°$ and $28.00°C$, with a mean of $26.83°C$. This means that groundwater in the study area has not been affected by geothermal heat which is insignificant at those depths. As a result, the weathering of silicate minerals, which is usually facilitated by higher temperatures and low pH, will be slower. The same applies to all reactions that usually require higher temperatures than room temperature conditions to be spontaneous. Groundwater in the area is acidic, as the pH ranges between 4.00 and 6.80 with an average of 5.68. The low pH is probably the result of the oxidation of pyrite which is present in high quantities in association with the gold mineralization. The EC ranges between 37.00 and 670.00 $\mu S/cm$, which corresponds with TDS in the range of 22.20 and 402.00 mg/L. This means that groundwater in the study area is fresh and pristine. All the parameters are generally low in concentrations and fall within the ranges prescribed by the WHO (2003) for domestic and other purposes.

4.2 *The factor model*

The final factor model has three factors which account for 76% of the total variance in the hydrochemistry. Table 2(a and b) presents the factor
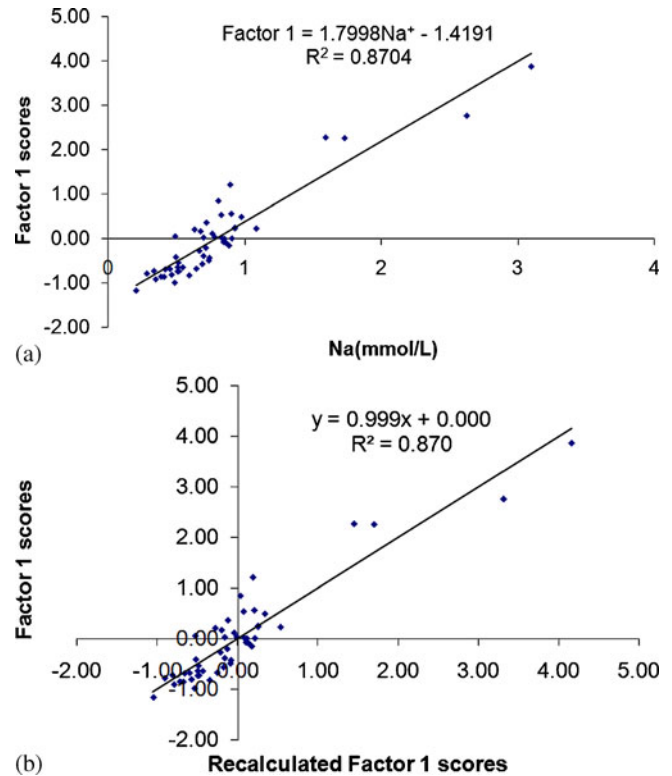

Figure 3. (**a**) A plot of recalculated factor scores against the concentrations of the key parameter and (**b**) a plot of recalculated factor 1 scores against the real factor 1 scores obtained directly from the factor analysis.

loadings matrix. Factor 1 explains 39% of the total variance and has high positive loadings for most of the major ions (Na^+, Mg^{2+}, Cl^- and SO_4^{2-}). The loadings of Na^+ and Cl^- under factor 1 are compatible and suggest a similar source for them. The underlying geology is not known to contain appreciable quantities of halite, but the phyllite and shale will certainly contain the sulfates and chlorides of the major cations whose dissolution could influence groundwater hydrochemistry. Factor 1 certainly does not correlate significantly with SiO_2, and will not be associated with the weathering of silicate minerals. Factor 1 is a mixed factor representing the contribution of the dissolution of chlorides and sulfates of the major cations. Although the geochemistry of rocks of the Lower Birimian and Tarwaian were not part of the objectives of this study, the general mineralogy of the rocks contained in Kesse (1985) and Junner (1940) suggest a wide range of compositional variations of the rocks within the Lower Birimian. The Mg^{2+} ion probably resulted from the dissolution of gypsum and associated minerals in the aquifers. The Na^+ ion has the highest loading with factor 1 and was identified as the key and representative parameter for the latter. The regression relationship between concentrations of Na^+ in mmol/L

against the scores of factor 1 is presented in equation 3. The high R^2 value suggests a strong relationship. This regression equation can therefore be used to predict the impacts of factor 1 anywhere, the concentrations of the Na^+ ion are known in groundwater in the study area. The regression model is easier to handle than recalculating factor scores from all the parameters in table 2(a and b). What is more, it is only the concentration of the key parameter(s) which is (are) required to determine the pervasiveness of factor 1 at the site in question. Factor scores usually indicate the degree to which the processes represented by the factor affect groundwater quality in the area. Where the factor score is higher than 0, the contribution of factor 1 at the location is high (Jiang *et al* 2009; Kim *et al* 2009). The converse is true for negative factor scores. The more positive the factor score, the higher the effect of the factor in question at the location.

$$\text{Factor } 1 = 1.8Na^+ - 1.42$$
$$R^2 = 0.87. \tag{3}$$

The relationship between the concentrations of the Na^+ ion at the various sampled locations and the factor scores, is presented in figure 3(a), together with a plot of factor scores recalculated from equation (3) (figure 3b). It is obvious in figure 3(b) that the factor scores have been adequately predicted and equation (3) can be relied upon at 87% level of accuracy. The degree of accuracy can be improved upon by adding more parameters that load well with factor 1. In this respect, since both Na^+ and Cl^- ions both load highly under factor 1 at loadings higher than 90%, a multiple regression model was developed with factor 1 scores as the dependent variable and the two key parameters as independent variables. This resulted in equation (4) with

a significantly improved efficiency and accuracy in predicting the factor scores.

$$\text{Factor } 1 = 0.998\ Na^+ + 0.692\ Cl^- - 1.288$$
$$R^2 = 0.94 \tag{4}$$

where the concentrations of the ions are in mmol/L.

The strength of the relationship between the original factor 1 scores and the recalculated factor 1 scores is vividly presented in figure 4. It is obvious from figure 4 that the predicted and observed factor scores are almost the same and equation (4) can be used with a high degree of confidence to predict the intensity of the effects of factor 1 everywhere in the study area. The effects of factor 1 are highest in 20 locations where the factor scores are higher than zero. All the locations with the highest effects of factor 1 (positive factor scores) have the lowest depths in the sample space, and range in depth between 20 and 28 m below ground surface. The low depths make them more susceptible to contamination from surface activities than the deeper ones, which have low factor scores. Majority of the low depth wells/boreholes have also been completed in the weathered zones which are semi-confined to unconfined. Such conditions predispose them to the effects of surface activities. Factor 1 can therefore be used to indicate the degree of confinement of the aquifers being tapped by the wells/boreholes in the study area. The lower the factor score, the higher the degree of confinement and therefore, the less likely it is for the aquifer to be affected by surface activities such as waste disposal. Cloutier *et al* (2008) used a similar approach to distinguish groundwater types in the Quebec area. Perhaps the singularity of the current work is the development of the regression model from
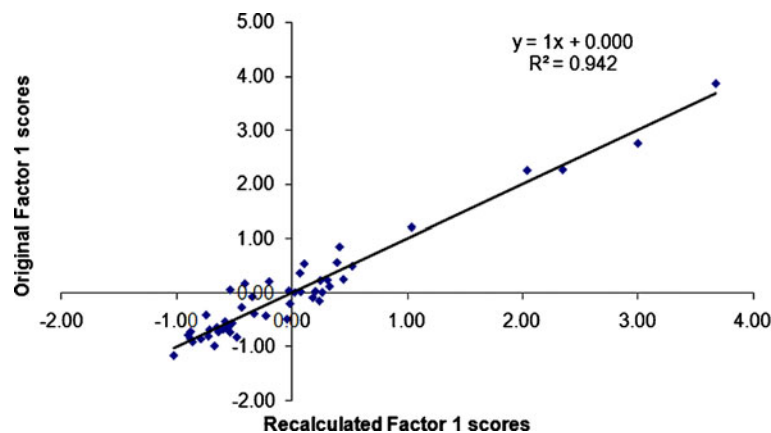

Figure 4. A plot of recalculated factor 1 scores (when more than one key parameter is selected) against the real factor 1 scores obtained from the factor analysis.

which the effects of the factor can be directly computed from the concentrations of two ions in the hydrochemistry. The resulting factor score will in turn be used to determine the effects of factor 1 or anthropogenic contamination on groundwater resources in the area.

Factor 2 in the final factor model has high positive loadings for pH, Ca^{2+} and HCO_3^- and represents the effects of carbonate mineral weathering on the hydrochemistry of groundwater in the study area. It accounts for 26% of the total variance in the hydrochemistry (table 2b) and therefore represents the second most important process in the hydrochemistry of groundwater in the study area. Amponsah (pers. comm.) reveals high calcite content (over 15% in some cases) in the phyllite and metagreywackes of the Lower Birimian in the Brong Ahafo region, northeast of the study area. Some of the manganese ores associated with phyllites in the area are carbonates. Akabzaa and Yidana (2012) reported kutnahorite in sediments and soils in the area. Incongruent silicate mineral weathering, facilitated by carbonic acid, leads to the generation of the bicarbonate ion (HCO_3^-) and cations. The process, after a long period, can lead to the precipitation of the carbonates of cations which were originally present in the minerals. In the case of Ca-rich silicate minerals (such as anorthite), this can result in the precipitation of calcite or aragonite in the resulting weathered zone. It is the dissolution of such carbonate minerals that results in the loadings observed under factor 2. Since carbonate mineral weathering leads to increases in the pH of the medium, factor 2 also loads positively with the latter. Factor scores of factor 2 were similarly computed as in the case of factor 1 and both Ca^{2+} and HCO_3^- were used as the key parameters in the regression modelling. The resulting model is presented in equation (5).

$$\text{Factor } 2 = 1.255Ca^{2+} + 0.662HCO_3^- - 1.33$$
$$R^2 = 0.90 \tag{5}$$

where concentrations of the ions are in mmol/L.

Since factor 2 represents the weathering of carbonate minerals, the effects are highest in areas where carbonate minerals are present in high concentrations to be weathered. As such the high positive scores of factor 2 are localized. Factor 2 increases with increasing alkalinity due to the enrichment of the HCO_3^- ion in the groundwater system. Factor 2 can therefore be linked to increasing alkalinity in the study area.

The last factor in the model accounts for about 11% of the total variance in the hydrochemistry and has a high positive loading for SiO_2 and a high negative loading for NO_3^-. The negative loading of factor 3 with NO_3^- suggests that the processes it represent go contrary to the processes that generate NO_3^- in the system. Factor 3 loads positively with both the HCO_3^- ion and pH although these loadings are weak at 0.219 and 0.245, respectively (table 2a). Factor 3 represents the weathering of silicate minerals which require acid conditions and high temperatures to be effective and contribute substantially to the hydrochemistry of groundwater in the area. The positive correlation of pH with factor 3 suggests increases in the latter with increasing silicate mineral weathering. The negative correlation of the NO_3^- ion which represents acidic medium is consistent with this explanation. The acidity of the medium is being consumed in silicate mineral weathering, and the result is increasing pH or alkalinity. Factor 3 also loads positively with the Mg^{2+} ion at 0.405, which suggests the weathering of Mg^{2+} containing silicate minerals such as the micas in the sediments, and some olivines in the metavolcanics. High biotite content has been recorded in some of the rocks in the Birimian. As shown in table 1, an average of 22.6 mg/L and standard deviation of 17.9 mg/L of silica (SiO_2) are concentrations generally found in natural waters probably reflecting the low solubility of clay minerals, feldspars, quartz and other common silicate minerals. Maximum value of 90 mg/L in table 1 is rarely found, except in water from volcanic rocks and sediments derived from volcanic or granitic rocks as it is the case in the study area (figure 2). For instance, the solubility of quartz is low and ranges from 5 to 15 mg/L SiO_2 in groundwater. Other silicate minerals have solubilities intermediate between quartz and amorphous silica. In this instance, according to the theory pH of groundwater will normally have little effect on the amount of silica which will go into solution (Krauskopf 1956). However, above a pH of 9.0 (which is not the case for this study), the solubility of silicates increases rapidly. The fact that pH is low (acidic conditions) only might explain the low solubilities of SiO_2.

Again, the intensity of the processes of silicate mineral weathering is limited and restricted to locations where the prevailing conditions favour mineral weathering processes. The concentrations of SiO_2 and NO_3^- were used to generate a multiple regression model for factor 3. This model is presented in equation (6).

$$\text{Factor } 3 = 0.033\,SiO_2 - 0.281\,NO_3^- - 0.247$$
$$R^2 = 0.89 \tag{6}$$

where the concentrations are in mg/L.

Results of the Q-mode HCA are presented in figure 5 in which three spatial groundwater associations have been distinguished. The phenon line was drawn across the dendrogram at a linkage

HIERARCHICAL CLUSTER ANALYSIS
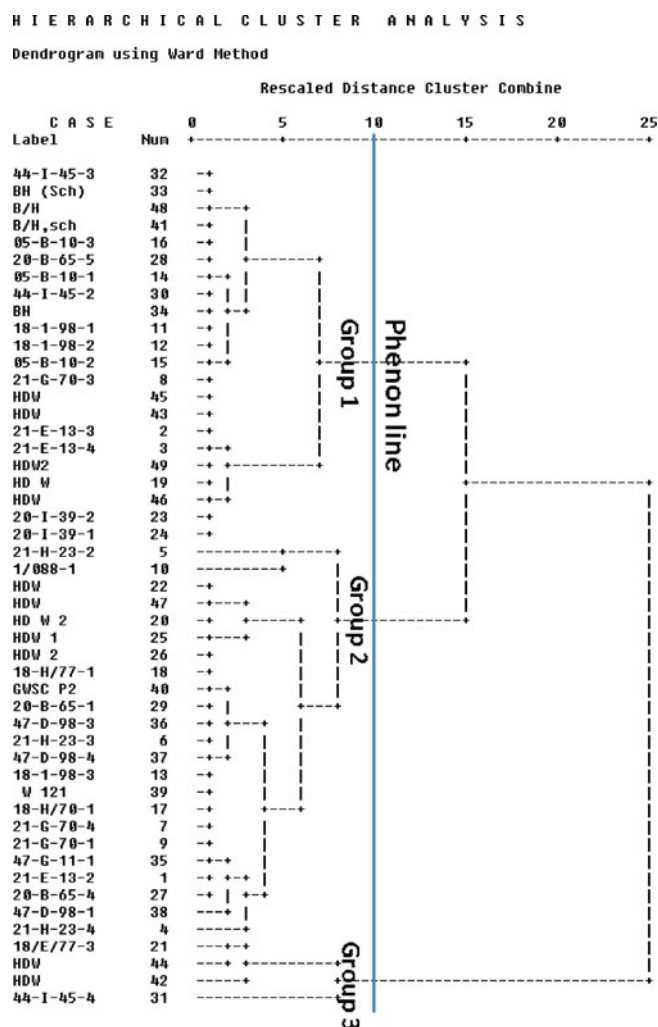
Dendrogram using Ward Method

Figure 5. Dendrogram from the Q-mode HCA.

distance of 10. This resulted in three spatial groundwater associations represented by the three clusters in figure 5. Stiff diagrams were generated for the arithmetic averages of the measures of the physico-chemical parameters used in this study. Figure 6 presents the stiff diagrams (figure 6a) and the spatial distribution (figure 6b) of all the three clusters (groups) in the study area. Each cluster (group) represents the position of its members in the groundwater flow regime and/or the key processes influencing groundwater hydrochemistry in the areas where the membership of these clusters are presented. For instance group 1 represents intermediate type, Na–HCO$_3$ fresh groundwater types in the flow regime; group 2 represents fresh Ca–HCO$_3$ groundwater types which are characteristic of groundwater at or close to recharge areas in the general flow regime; group 3 represents Na–Cl groundwater type which is characteristic of groundwater in discharge areas of the flow regime. Although clusters 1 and 2 represent intermediate and recharge type groundwaters, they may have

resulted purely from the interaction between the rock and groundwater and other intervening processes in the groundwater hydrochemistry. There is no detailed map of the potentiometric surfaces of groundwater in the study area for immediate comparison with the spatial location of the membership of these clusters. On account of this, it is apposite to associate the groundwater types to mineral weathering processes. The different groundwater types therefore resulted from the variability in the composition of the rocks. The group 1 Na–HCO$_3$ groundwaters are not confined to any particular geology in the study area (figure 6b). Group 2 members are largely within the Tarkwaian metasediments, and some of the granitoid intrusions. The mineralogy of these rocks support the Ca–HCO$_3$ groundwater types observed. In the Tarkwaian, this is most likely due to the dissolution of carbonate minerals whereas silicate mineral dissolution is likely attributable to the observed groundwater type in the ganitoids.

Group 3 members are largely confined to the Upper Birimian rocks. The Ca–HCO$_3$ groundwater types result from the weathering of carbonate minerals as suggested by factor 2 and may not necessarily represent groundwater in recharge areas of the flow regime. The Na–HCO$_3$ groundwater type results from the interaction or mixing of groundwaters influenced by the two processes of mineral dissolution. In all three cases, the concentrations of the ions are within acceptable ranges for most domestic uses as the values are within acceptable WHO (2003) and national standards.

Factor scores were calculated using the regression relationships above, for the three groundwater associations distinguished by the Q-mode HCA. Figure 7(a and b) respectively show scatter plots of factor 1 against factor 2 and factor 2 against factor 3. In the scatter plot, the samples have distinguished themselves into distinct associations. The group 3 members display score lower than zero for factor 2 and more positive scores for factor 1 (figure 3). Group 2 members plot within the positive range of the factor scores for factor 2, suggesting higher influence of carbonate mineral weathering.

Ca–HCO$_3$ water types are often associated with carbonate mineral and/or incongruent silicate mineral weathering (Drever 1988; Appelo and Postma 2005). The average TDS values of the three groundwater groups suggests the highest value of 362 mg/L for group 3, which may have resulted from the relative ease of dissolution of the chlorides of the major cations, compared to the other processes contributing to the hydrochemistry. In figure 7(b), group 2 members show as much association with factor 3 as they do with factor 2. This suggests that silicate mineral weathering plays a
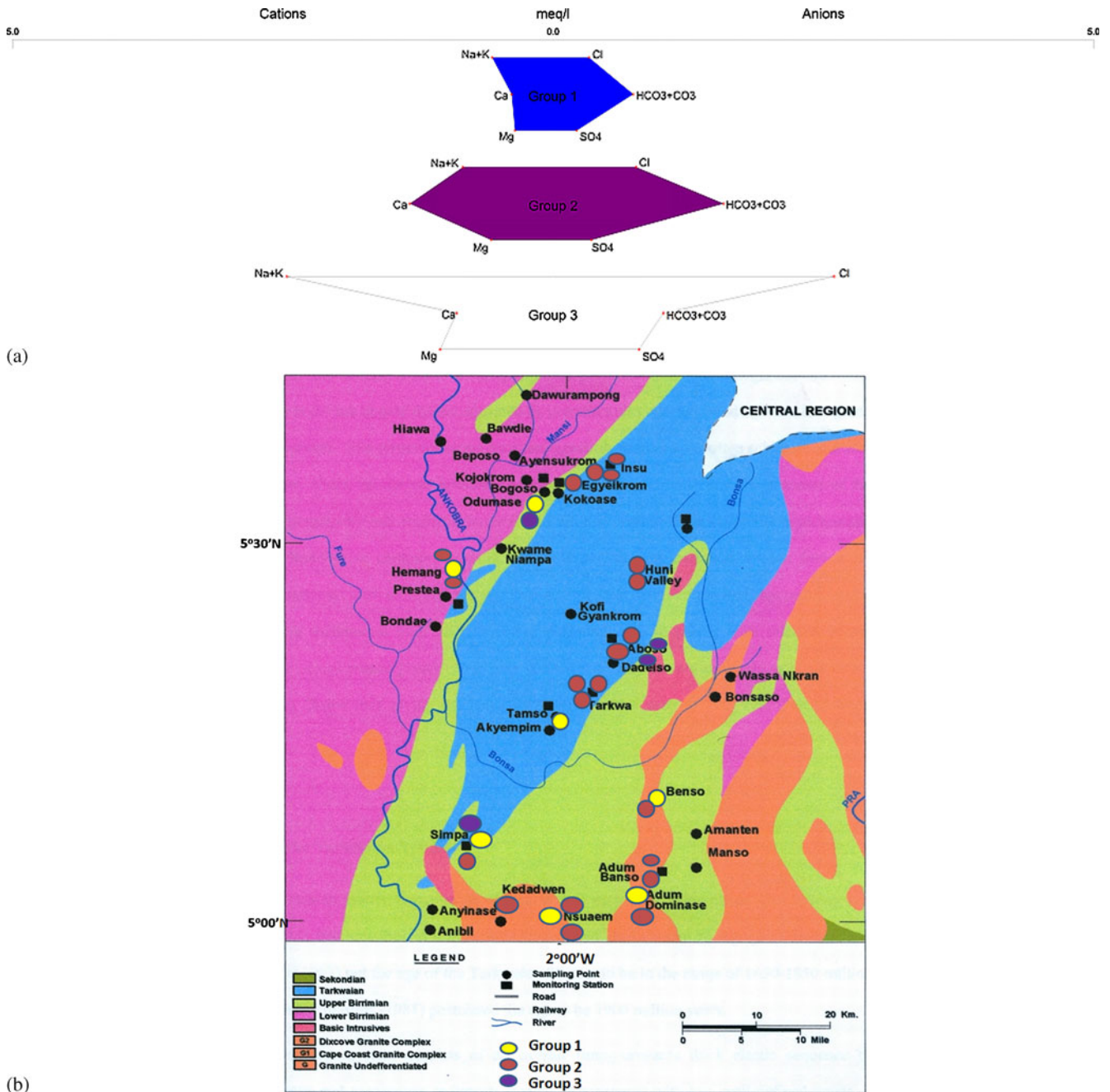
Figure 6. (**a**) Stiff diagrams from the statistical summaries of the major ions in the three clusters from the Q-mode HCA. (**b**) Map of the study area showing the spatial distribution of the membership of the three clusters (modified from Kortatsi 2004).

role in the generation of the Ca–HCO$_3$ groundwater types in the study area. It is also obvious that about half of the group 1 members plot within the positive range for factor 3, an indication that silicate mineral weathering probably contributed to the Na–HCO$_3$ groundwater types in the study area. Incongruent weathering of albite and related plagioclase feldspars has the potential to result in Na–HCO$_3$ compositional groundwater types. Thus silicate mineral weathering is the process of significant effect on the members of groups 1 and 2. In

figure 7(c), the hydrochemistries of groups 1 and 2 appear to associate more closely with factor 3 than factor 1, whereas group 3 members quite distinctly display a strong bias towards factor 1.

4.3 *Gibbs and silicate stability diagrams*

Gibbs (1970) diagram has been used extensively in water resource studies to help understand hydrochemical trends. The diagram is divided into
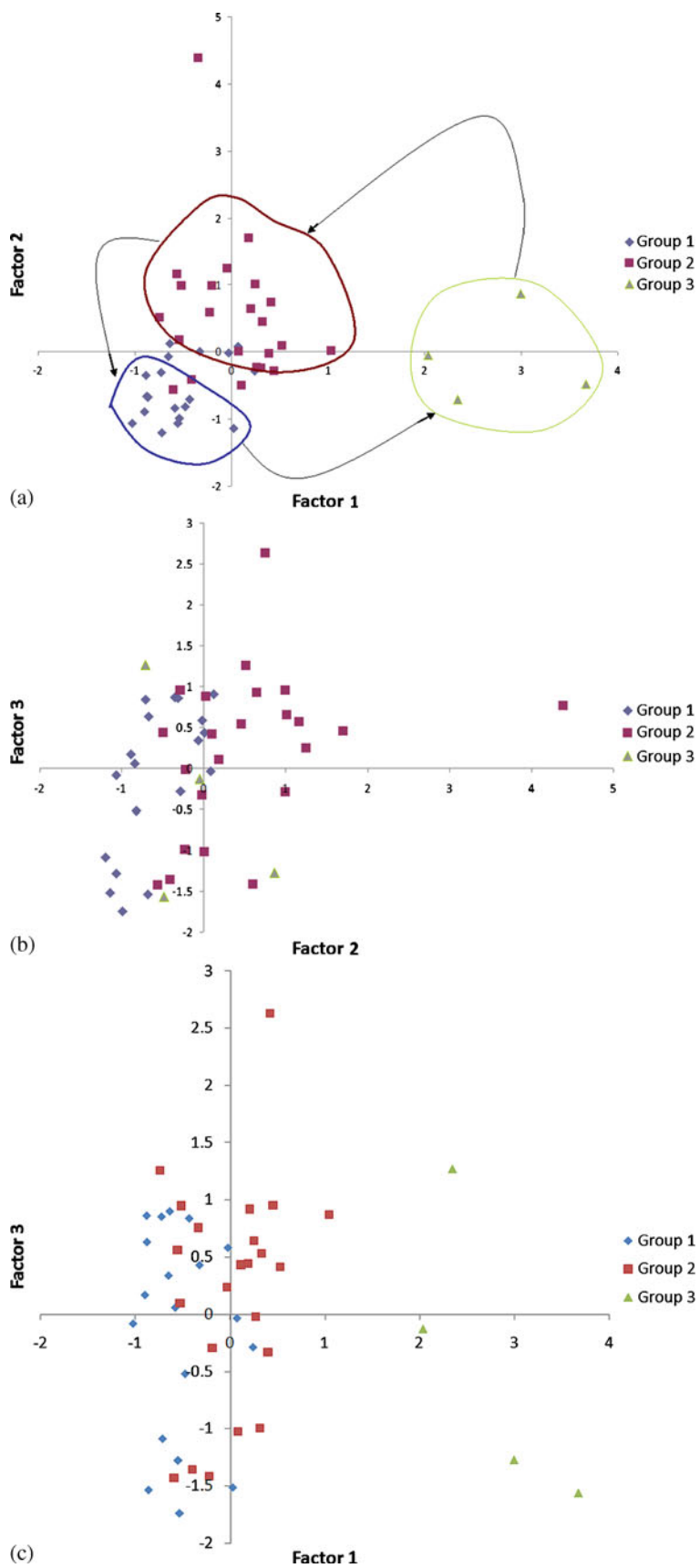
Figure 7. A plot of (**a**) factor 1 against factor 2, (**b**) factor 2 against factor 3, and (**c**) factor 1 against factor 3.
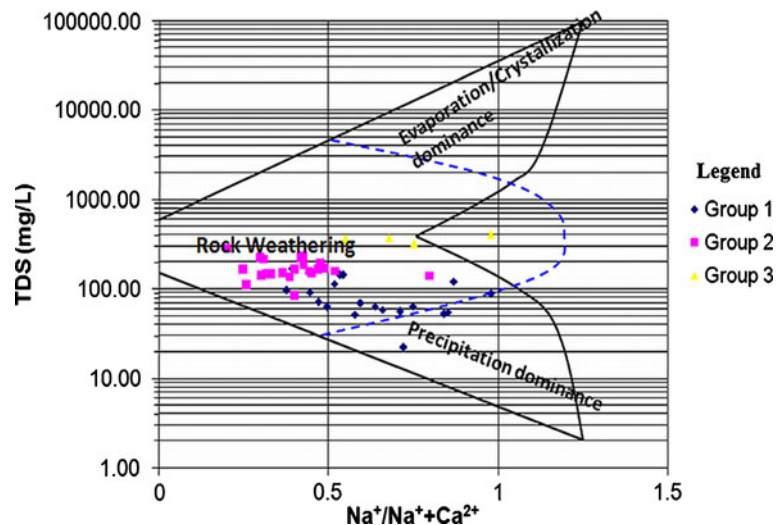
Figure 8. Gibbs diagram showing the positions of all three groundwater groups in the study area.

zones based on the contribution of recharging pre-
cipitation, rock/mineral weathering and evapora-
tion/crystallization on the hydrochemistry. The
discretization is based on the concentration of the
Na^+ ion relative to the sum of the concentrations
of the Na^+ and Ca^{2+} ions in the hydrochemistry
and the TDS content of the water. Waters affected
by high evaporative effects are generally expected
to present high TDS and Na/Ca+Na ratios. Where
the hydrochemistry is controlled mainly by min-
eral weathering processes, the TDS is moderate,
and the Na/Ca+Na is also moderate. For waters
influenced mostly by the hydrochemistry of pre-
cipitation, the TDS is low but the Na/Ca+Na
ratio is moderate to high. One of the major short
comings with the Gibb's (1970) diagram is that
there is no place for waters which have been
affected by domestic contamination or other
sources of hydrochemical enrichment outside of the
three sources outlined above. It is probably on
account of this reason that the diagram is often
used together with other hydrochemical assessment
methods such as multivariate statistical analyses.

In this study, the Gibb's diagram was used to
corroborate the major findings from the HCA and
factor analysis, and also bring to light some of its
inherent weaknesses. The Gibb's diagram, contain-
ing the members of all three groundwater associ-
ations revealed by the Q-mode HCA, is presented
in figure 8. It is apparent that much of the data
plots within the rock dissolution category of the
diagram. Although group 3 members have been
noted to have significant anthropogenic contribu-
tions, they plot within the rock dissolution por-
tion of the Gibb's diagram due to the moderate
TDS and Na/Ca+Na ratio. This highlights one of
the weaknesses discussed above in association with
the diagram. Group 1 members plot within and

close to the 'precipitation dominance' region. This
is in agreement with the dominant water type in
group 1 and the interpretation of the factor models
above. It is the high precipitation influence that is
responsible for the moderate to high Na^+ content
compared to the other cations.

All group 2 members plot within the rock dom-
inance portion of the diagram as expected. There
appears to be a gradual evolution of groundwa-
ter from group 1 through group 2 to group 3 as
the TDS increases. The high Na/Ca+Na ratio in
group 3 is associated with high contribution of the
Na^+ ion from anthropogenic activities. As shown in
figure 7, group 3 members are less influenced by
general mineral weathering compared to the degree
of unconfinement or closeness of the aquifer to the
surface. It is these factors that determine the mag-
nitude of the anthropogenic activities on ground-
water quality in the study area. This is not in any
way discounting any contribution of mineral weath-
ering processes on the hydrochemistry of group 3
members. In so far as groundwater remains in con-
tact with the host rock, interactions between the
latter and former will result in the leaching of some
ions into the groundwater system. In this respect,
all the major groundwater associations in the area
have been influenced variously by mineral weath-
ering processes. In the foregoing analysis, however,
emphasis has been placed on the most significant
process or set of processes of significance in the
hydrochemistry of each group.

Mineral stability diagrams have been used
extensively in hydrochemical research to constrain
silicate mineral weathering processes and high-
light the most stable silicate mineral phases in the
hydrochemistry of surface and groundwater basins.
Figure 9 shows the stability diagrams for the
$NaO–H_2O–Al_2O_3–SiO_2$ and $CaO–H_2O–Al_2O_3–SiO_2$

Figure 9. A mineral stability diagram for (**a**) the NaO–Al$_2$O$_3$–SiO$_2$–H$_2$O system, and (**b**) the CaO–Al$_2$O$_3$–SiO$_2$–H$_2$O system.

time is short. Stability within the gibbsite field in a silicate terrain is revealing of water that is relatively fresh. It is the low level of contact between the groundwater and host rock that leads to the formation of gibbsite. This occurs in tropical environments where the flushing of rainwater is rapid (Hiscock 2005). In areas of arid to semi-arid climate conditions, and where the flow is restricted, the interaction between silicate minerals and groundwater leads to the formation of montmorillonite, a higher silica clay mineral end member. Restricted flow conditions retain the water in contact with the rock/aquifer material for a long time, thus exposing the minerals to weathering. Flow conditions intermediate between those that lead to the formation of gibbsite and montmorillonite lead to the formation of kaolinite. The silicate stability diagram can therefore aid in the understanding of the flow system, and the prevailing climatic conditions in the area.

In this study, figure 9(a and b) shows just a few points within the gibbsite field due to relatively low silica contents in these samples. In the case of the two samples of group 3 that plot within the gibbsite region, the logical inference is the absence of silica or insignificant contribution of silicate mineral weathering to the hydrochemistry of these waters. The group 1 and 2 members in the gibbsite region are relatively younger groundwaters, close to recharge areas in the flow regime. It is the combinations of the tropical climatic conditions and generally slow or restricted flow conditions that have led to the observed stability within the kaolinite field in the NaO–H$_2$O–Al$_2$O$_3$–SiO$_2$ (figure 9a). On the stability diagram of the relatively higher temperature Ca-rich end member plagioclase (anorthite) in figure 9(b), it is apparent that the data is split between kaolinite and montmorillonite fields. Almost all the samples plotting within the Ca-montmorillonite field belong to the group 2 groundwater type. The observed high Ca^{2+} and HCO$_3^-$ is therefore attributed in part to the contribution of the incongruent silicate mineral weathering which occurred to result in the weathered zone. It is probably these processes that occurred to concentrate Ca^{2+} and HCO$_3^-$ ions enough to result in the precipitation of carbonate minerals in the weathered zone. The Goldich weathering sequence (Hiscock 2005) is the converse of the Bowen's reaction series and suggests that the high temperature minerals weather much more readily whilst the low temperature ones are much more resistant to chemical weathering. It is in the light of this that the CaO–H$_2$O–Al$_2$O$_3$–SiO$_2$ stability field shows more points in the montmorillonite field than that of the NaO–H$_2$O–Al$_2$O$_3$–SiO$_2$. The mineral stability diagrams as presented in figure 9 represent minerals with ideal chemical

systems in the study area. Both stability diagrams indicate stability in the kaolinite field. In figure 9(a), most of the group 2 members plot within the Ca-montmorillonite field due to high Ca^{2+} and SiO$_2$ contents in the samples. In groundwater hosted by silicate terrain, the interaction between the rocks and groundwater leads to the leaching of ions into the groundwater system. As a result of such interactions, different products result, depending on the residence time of the water, the climatic conditions and the general flow regime. Incongruent weathering leads to gibbsite, a silica deficient end member in the silicate stability diagram, where the flow is rapid and the residence

compositions, and may not accurately depict the prevailing conditions in most aquifers (Hiscock 2005). Nevertheless, such diagrams provide useful information that will lead to the understanding of the flow regimes in most groundwater systems.

Perhaps the most useful inference from figure 9 is the fact that groundwater flow in the study area is moderately-to-highly restricted, lending credence to the fact that groundwater flow and storage in the area is facilitated by secondary permeabilities resulting from the fissuring and/or weathering of the rock whose primary permeability is negligible. As a result, the flow is restricted to the weathered and/or fractured zones and impeded by the relatively fresh rock. Silicate mineral weathering processes are facilitated in an environment where the pH is low (acidic conditions) as explained above. In the study area, these conditions result from the dissolution of CO_2 in the groundwater system from the respiration of plant root cells (Appelo and Postma 2005; Hiscock 2005), and CO_2 dissolved in recharging precipitation water. The dissolution of CO_2 provides the required acid conditions for silicate mineral weathering. The oxidation of organic matter by any oxidizing agent such as the NO_3^- ion can create the desired pH conditions for silicate mineral weathering. In addition, the gold mineralization in the study area is associated with sulfide minerals, whose oxidation creates the desired pH conditions. It is for this reason that the pH range of the sampled locations is within the acid range. Thus silicate mineral weathering is an active process in the groundwater flow system in the study area.

5. Conclusions

A methodology has been developed to facilitate the proper characterization of the hydrochemistry of surface and groundwater basins. The methodology is based on the combined use of factor and cluster analyses. This study also finds that groundwater hydrochemistry in the aquifers of the Tarkwaian and Birimian basins in the western region, is controlled by three major factors: dissolution of the chlorides and sulfates of the major cations, carbonate mineral dissolution, and silicate mineral weathering. Sodium and Cl^- ions, Ca^{2+} and HCO_3^- ions, and SiO_2 and NO_3^- ions are the key parameters used to trace the effects of the dissolution of chlorides and sulphates of major ions, carbonate mineral dissolution and silicate mineral weathering, respectively. Multiple regression models have been developed to measure the intensity of the above-mentioned processes using the concentrations of the key parameters mentioned. Using these regression models, one will be able to assess the intensities of the major processes of significance in the hydrochemistry of groundwater in the study area without measuring the concentrations of all the physico-chemical parameters. Stability diagrams for the study area suggests stability in the kaolinite and montmorillonite fields suggesting restricted groundwater flow leading to longer residence times in some cases.

Acknowledgements

References

Abu-Jaber N 2001 Geochemical evolution and recharge of the shallow aquifers at Tulul al Ashaqif, NE Jordan; *Environ. Geol.* **41** 372–383.

Akabzaa T M and Yidana S M 2012 An integrated approach to environmental risk assessment of cumulatively impacted drainage basin from mining activities in southwestern Ghana; *J. Environ. Earth Sci.* **65** 291–312.

Anazawa K and Ohmori H 2005 The hydrochemistry of surface waters in Andesitic Volcanic area, Norikura volcano, central Japan; *Chemosphere* **59** 605–615.

Antonellini M, Mollema P, Giambastiani B, Bishop K, Caruso L, Minchio A, Pellegrini L, Sabia M, Ulazzi E and Gabbianelli G 2008 Salt water intrusion in the coastal aquifer of the southern Po Plain, Italy; *Hydrogeol.* **16** 1541–1556.

Apambire W B, Boyle D R and Michel F A 1997 Geochemistry, genesis, and health implications of fluoriferous groundwaters in the upper regions of Ghana; *Environ. Geol.* **33** 13–24.

Appelo C A J and Postma D 2005 *Geochemistry, Groundwater and Pollution*; 2nd edn, Balkerma, Netherlands.

Banoeng-Yakubo B, Yidana S M and Nti E 2009 An evaluation of the genesis and suitability of groundwater for irrigation in the Volta region, Ghana; *Environ. Geol.* **57** 1005–1010.

Cloutier V, Lefebvre R, Therrien R and Savard M M 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.

Coetsiers M and Walravens K 2006 Chemical characterization of the Noegene Aquifer, Belgium; *Hydrogeol.* **14** 1556–1568.

Dickson K B and Benneh G 1980 *A new geography of Ghana*; Longmans Group Limited, London.

Dickson K B and Benneh G 1995 *A new geography of Ghana*, 3rd edn, Longman, Malaysia.

Drever J I 1988 *The Geochemistry of Natural Waters*; 2nd edn, Prentice Hall, New Jersey.

Gibbs R J 1970 Mechanisms controlling world water chemistry; *Science* **17** 1088–1090.

Güler C and Thyne G D 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 D, McCray J E and Turner A K 2002 Evaluation of graphical and multivariate statistical methods for classification of water chemistry data; *Hydrogeol. J.* **10** 455–474.

Helstrup T, Jorgensen N O and Banoeng-Yakubo B 2007 Investigation of hydrochemical characteristics of groundwater from the Cretaceous–Eocene limestone in southern Ghana and southern Togo using hierarchical cluster analysis; *Hydrogeology* **15** 977–989.

Hiscock K M 2005 *Hydrogeology, Principles and Practice*; Blackwell Publishing, Malden Massachusset, USA.

Jiang J, Wu Y, Groves C, Yuan D and Kambesis P 2009 Natural and anthropogenic factors affecting the groundwater quality in the Nandong karst underground river system in Yunan, China; *J. Contam. Hydrol.* **109** 49–61.

Jorgensen N O and Banoeng-Yakubo B K 2001 Environmental isotopes (^{18}O, ^{2}H, and ^{87}Sr/^{86}Sr) as a tool in groundwater investigations in the Keta Basin, Ghana; *Hydrogeology* **9** 190–201.

Junner N R 1940 Geology of the Gold Coast and Western Togo; Bulletin No. 11.

Junner N R, Hirst T and Service H 1942 Tarkwa Goldfield; *Memoir*, No. 6, *Gold Coast Geological Survey*.

Kaiser H F 1960 The application of electronic computers to factor analysis; *Educ. Psychol. Meas.* **20** 141–151.

Kesse G O 1985 *The mineral and rock resources of Ghana*; AA Balkema, Rotterdam, Boston.

Kim K O, Yun S-T, Choi B-Y, Chae G-T, Joo Y, Kim K and Kim H-S 2009 Hydrochemical and multivariate statistical interpretations of spatial controls of nitrate concentrations in a shallow alluvial aquifer around oxbow lakes (Osong area, central Korea). *J. Contam. Hydrol.* **107** 114–127.

Kortatsi B K 2004 Hydrochemistry of groundwater in the mining area of Tarkwa–Prestea, Ghana. Unpublished Ph. D thesis. 103 p.

Kortatsi B K 2007 Hydrochemical framework of groundwater in the Ankobra Basin, Ghana; *Aquat. Geochem.* **13** 41–74.

Krauskopf K B 1956 Dissolution and precipitation of silica at low temperatures; *Geochim. Cosmochim. Acta* **10** 1–26.

Meng S X and Maynard J B 2001 Use of statistical analysis to formulate conceptual models of geochemical behavior: Water chemical data from Butucatu aquifer in Sao Paulo State, Brazil; *J. Hydrol.* **250** 78–97.

Thyne G, Guler C and Poeter E 2004 Sequential analysis of hydrochemical data for watershed characterization; *Ground Water* **42** 711–723, doi: 10D1111/j.1745–6584D2004.tb02855.x.

Trabelsi R, Zairi M and Dhia B H 2007 Groundwater salinization of the Sfax superficial aquifer, Tunisia; *Hydrogeol. J.* **15** 1341–1355.

Wen X, Wu Y, Su J, Zhang Y and Liu F 2005 Hydrochemical characteristics and salinity of groundwater in the Ejina Basin, Northwestern China; *Environ. Geol.* **48** 665–675.

World Health Organization (WHO) 2003 *Guidelines for drinking water quality*; Revision of the 1993 Guidelines. Final Task Group Meeting, Geneva, 21–25 September 1992.

Yidana S M 2008 Management of groundwater for productive uses in the Afram Plains area, Ghana; PhD dissertation, Montclair State University, Montclair, New Jersey, USA.

Yidana S M, Banoeng-Yakubo B and Akabzaa T 2010 Analysis of groundwater quality using multivariate and spatial analyses in the Keta basin, Ghana; *J. African Earth Sci.* **58** 220–234.

Yidana S M, Ophori D and Banoeng-Yakubo B 2008a A multivariate statistical analysis of surface water chemistry – the Ankobra Basin, Ghana; *Environ. Manag.* **86** 80–87.

Yidana S M, Ophori D and Banoeng-Yakubo B 2008b Hydrochemical evaluation of the Volta Basin: The Afram Plains area; *J. Environ. Manag.* **88** 697–707.