Spatial control of groundwater contamination, using principal component analysis

N Subba Rao

Department of Geology, Andhra University, Visakhapatnam 530 003, India.
e-mail: srnandipati@gmail.com

A study on the geochemistry of groundwater was carried out in a river basin of Andhra Pradesh to probe into the spatial controlling processes of groundwater contamination, using principal component analysis (PCA). The PCA transforms the chemical variables, pH, EC, Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^-, into two orthogonal principal components (PC1 and PC2), accounting for 75% of the total variance of the data matrix. PC1 has high positive loadings of EC, Na^+, Cl^-, SO_4^{2-}, Mg^{2+} and Ca^{2+}, representing a salinity controlled process of geogenic (mineral dissolution, ion exchange, and evaporation), anthropogenic (agricultural activities and domestic wastewaters), and marine (marine clay) origin. The PC2 loadings are highly positive for HCO_3^-, F^-, pH and NO_3^-, attributing to the alkalinity and pollution controlled processes of geogenic and anthropogenic origins. The PC scores reflect the change of groundwater quality of geogenic origin from upstream to downstream area with an increase in concentration of chemical variables, which is due to anthropogenic and marine origins with varying topography, soil type, depth of water levels, and water usage. Thus, the groundwater quality shows a variation of chemical facies from $Na^+ > Ca^{2+} > Mg^{2+} > K^+$: $HCO_3^- > Cl^- > SO_4^{2-} > NO_3^- > F^-$ at high topography to $Na^+ > Mg^{2+} > Ca^{2+} > K^+$: $Cl^- > HCO_3^- > SO_4^{2-} > NO_3^- > F^-$ at low topography. With PCA, an effective tool for the spatial controlling processes of groundwater contamination, a subset of explored wells is indexed for continuous monitoring to optimize the expensive effort.

1. Introduction

Groundwater along with rainwater is a vital factor for social and economic development in any area due to uncertainty in monsoon. However, with the rapid economic growth and fast growing population, the land use activities have become key influencing processes. The continuous deterioration of groundwater quality is due to indiscriminate disposal of industrial effluents, domestic wastes, and uncontrolled usage of agricultural fertilizers, pesticides and soil amendments (Todd 1980; Hem 1991).

During infiltration, recharged water interacts with soils, weathered materials and/or fractured rocks and carries pollutants released from the land use activities. The pollutants are dissolved depending upon their solubility in the water, before reaching the groundwater body. Transgression (movement of seawater onto surface) and incursion (entering of seawater into subsurface) of saline water deteriorates the quality of groundwater in coastal regions. Thus, it results in groundwater contamination in many river basins in the country (Bhatt and Saklani 1996; Elampooranan *et al.* 1999; Umar and Absar 2003; Aravindan *et al.* 2004;

Khurshid and Zaheerudin 2004; Sreedevi 2004; Subba Rao *et al.* 2005; Subramani *et al.* 2005; Rajmohan and Elango 2006; Gopinath and Seralathan 2006; Jeevanandam *et al.* 2006; John Devadas *et al.* 2007; Giridharan *et al.* 2008; Krishna Kumar *et al.* 2009; Naik *et al.* 2009; Raju 2009; Papiya *et al.* 2010; Mithas *et al.* 2011; Mridul *et al.* 2011; Ravikumar *et al.* 2011; Shankar *et al.* 2011; Subba Rao *et al.* 2012a).

Earlier research studies on the Varaha River basin located in Visakhapatnam District, Andhra Pradesh (figure 1) dealt with the (i) search of groundwater potential zones, using geographical information system (Murthy 2000), (ii) hydrogeological framework and prospects of groundwater resource development (Central Ground Water Board, CGWB 2001), (iii) fluoride-bearing groundwater (Subba Rao 2009a), (iv) identification of effective sites for infiltration of recharge water (Subba Rao 2009b), and (v) chemical characteristics and quality of groundwater (Subba Rao and Surya Rao 2010; Subba Rao *et al.* 2012b). However, no attention was paid for the assessment of controlling processes in deteriorating the groundwater quality at a specific site to implement remedial measures.

The focus of the present paper is on principal component analysis (PCA) to identify the spatial controlling processes of groundwater contamination. The numerical values of principal component (PC) loadings and scores are used (i) to interpret the relative importance of the chemical variables and (ii) to probe into the classification of groundwater wells, based on specific field conditions of topography, soil type, depth of water levels and water usage. The output of PCA is utilized to select the index wells for long-term monitoring of groundwater quality. This approach reduces the number of wells and minimizes the expensive efforts.

2. Site description

2.1 *Location*

The present study area (Varaha River basin) is between latitude $17°23'–17°55'$N and longitude $82°30'–82°53'$E (figure 1). The climate is humid to dry-humid. The mean annual temperatures are in the range of 20 to 36°C. The mean annual rainfall is about 1200 mm (Subba Rao 2009b). The river Varaha originates from the Eastern Ghats, flows towards the southeastern side and finally joins the Bay of Bengal.

2.2 *Geology*

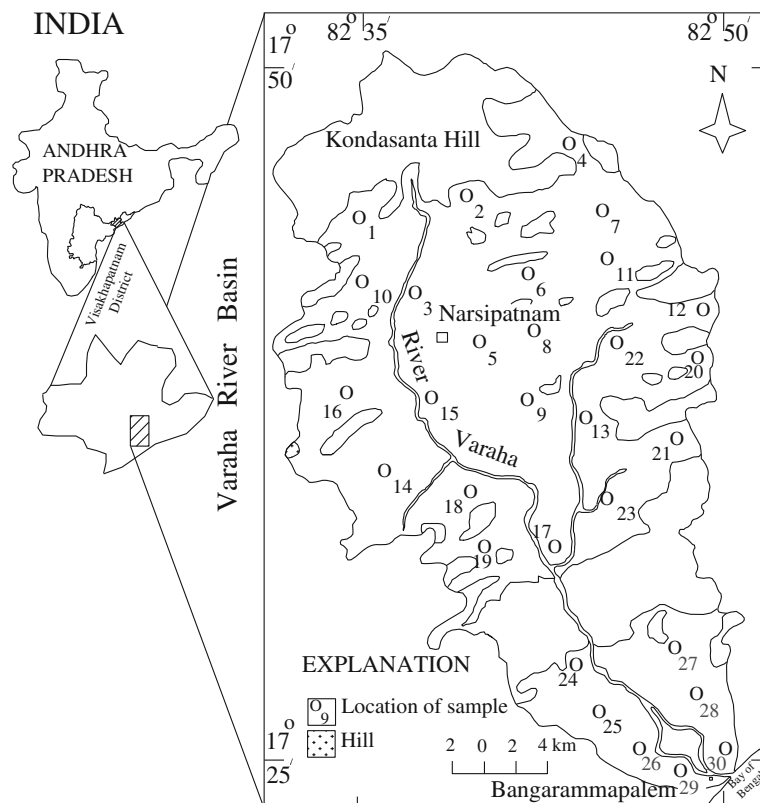The characteristic of the Varaha River basin is its undulating topography, sloping towards the

Figure 1. Groundwater sample locations in the Varaha River basin, Visakhapatnam District, Andhra Pradesh.

southeast. Geologically, the study area is underlain by the Precambrian Eastern Ghats and the Recent Formations (figure 2). The Eastern Ghats include khondalites and charnockites with quartzite, migmatite and granite intrusives. The khondalites are the most dominant rocks, while the charnockites occur in isolated pockets. The former rocks contain quartz, plagioclase and orthoclase feldspars, garnet, sillimanite, apatite, biotite and opaques with medium- to coarse-grained and gneissic texture, while the latter rocks are made up of quartz, potash, and plagioclase feldspars, ortho- and clino-pyroxenes, biotite, apatite, zircon and opaques with fine- to medium-grained and granulitic texture. The foliation trend of the khondalites is towards NE–SW with a deviation of NNW–SSE, N–S and NNE–SSW, and a dip of 70° southeast. The Recent Formations occur over the basement rocks. The formations include silt loam, loamy sand, colluvium, coastal sand and marine clay. They are fine to coarse-grained in nature. The red soil is the dominant type. The $CaCO_3$ concretions, locally known as *kankar*, occur as intercalations in the soil zone.

2.3 *Groundwater conditions*

Groundwater occurs in the weathered and fractured rock portions under water table conditions. The groundwater is drawn through shallow open dug wells and deep drilled bore wells. The water is used for drinking and irrigation purposes. The depth of water level varies from 3 and 15 m below ground level (bgl) at the river basin of Varaha. The shallow depth of water levels at low-lying areas and the deep depth of water levels at elevated grounds are observed.

2.4 *Land use and land cover*

About 65% of the total study area of the Varaha River basin is irrigated land, 21% waste land, 8% forest land, 5% built-up land, and the remaining 1% occupied by water bodies. Here, sanitary conditions are generally poor. Agriculture is the main occupation of the people in the study area. Long-term irrigation is in active practice.

3. Materials and methods

3.1 *Sampling and analytical procedures*

Thirty groundwater samples were collected from the Varaha River basin (figure 1) along with data on topography, soil type, depth of water levels, and water usage during May 2008 (table 1). The physico-chemical parameters: pH, electrical conductivity (EC), calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^+), potassium (K^+), bicarbonate (HCO_3^-), chloride (Cl^-), sulphate (SO_4^{2-}), nitrate (NO_3^-) and fluoride (F^-) were determined, using standard water quality procedures (APHA 1999).

The pH and EC were measured in the field, using portable pH and EC meters. The value of EC [expressed in micro-siemens per centimeter ($\mu S/cm$) at 25°C] was used to calculate the concentration of total dissolved solids (TDS, Hem 1991). The HCO_3^- was analysed volumetrically, using standard HCl. The total hardness (TH) as $CaCO_3$ and Ca^{2+} were determined by titrating, using standard EDTA. The Mg^{2+} content was computed as difference between the numerical values of TH and Ca^{2+}. The flame photometer technique was used for estimation of Na^+ and K^+ ions. The Cl^- was analyzed by titrating with standard $AgNO_3$. The
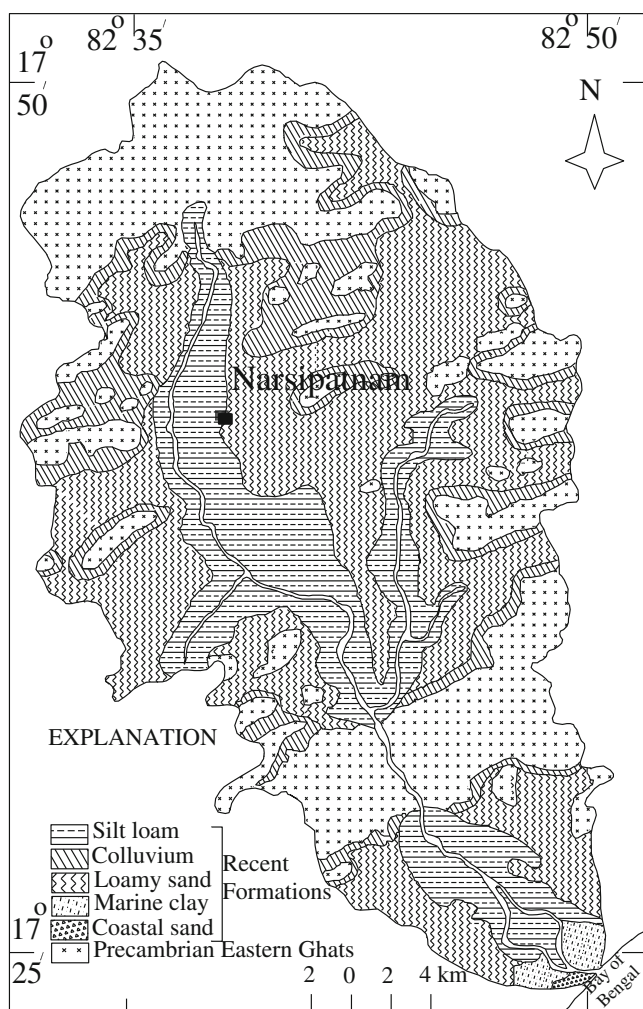

Figure 2. Geology of the Varaha River basin, Visakhapatnam District, Andhra Pradesh.

Table 1. *Chemical composition of groundwater in the Varaha River basin, Visakhapatnam District, Andhra Pradesh.*

Sample no. (figure 1)	Topography	Depth of water level (m bgl)	Soil type	Water use	pH	EC (μS/cm)	Ca^{2+} (mg/L)	Mg^{2+} (mg/L)	Na^+ (mg/L)	K^+ (mg/L)	HCO_3^- (mg/L)	Cl^- (mg/L)	SO_4^{2-} (mg/L)	NO_3^- (mg/L)	F^- (mg/L)
1	High	14.6	Loamy sand	Domestic	7.3	1400	50	30	201	11	490	170	29	33	0.7
2	High	13.5	Silt loam	Domestic	7.4	1430	45	30	210	17	510	170	31	31	0.7
3	High	12.9	Silt loam	Irrigation	7.7	1460	40	47	201	15	535	170	38	36	0.6
4	High	12.4	Loamy sand	Irrigation	7.3	1370	40	35	192	14	450	130	86	38	1.0
5	High	10.9	Loamy sand	Domestic	7.2	1490	55	30	212	18	460	225	21	32	0.8
6	High	11.6	Loamy sand	Domestic	7.6	1460	45	47	194	13	500	195	39	26	0.9
7	High	12.5	Loamy sand	Domestic	7.7	1630	30	47	242	13	530	210	45	26	0.9
8	High	10.1	Loamy sand	Irrigation	7.9	1750	45	40	271	19	560	235	52	37	1.2
9	Medium	8.7	Loamy sand	Domestic	7.9	1675	35	45	254	12	550	210	50	29	0.9
10	High	11.7	Loamy sand	Irrigation	7.5	1660	50	25	263	24	500	205	63	38	1.1
11	Medium	9.8	Loamy sand	Domestic	8.0	1955	50	45	293	21	610	270	65	33	1.2
12	Medium	9.7	Loamy sand	Irrigation	8.1	2155	60	55	330	19	720	315	46	53	1.3
13	Low	3.8	Silt loam	Domestic	7.4	1725	45	55	220	24	530	210	75	29	1.4
14	Medium	8.7	Silt loam	Irrigation	8.0	2015	30	50	311	38	640	310	44	39	1.1
15	Low	4.6	Silt loam	Domestic	7.3	1570	40	30	245	23	510	210	39	31	1.3
16	Medium	7.3	Loamy sand	Irrigation	7.9	1985	45	45	323	10	640	295	46	36	1.1
17	Medium	5.9	Loamy sand	Domestic	8.0	1800	45	35	280	26	540	260	39	29	1.3
18	Medium	6.4	Loamy sand	Irrigation	7.8	2110	60	60	290	12	665	300	49	36	1.2
19	Medium	9.1	Loamy sand	Irrigation	7.9	2140	60	55	312	15	700	285	46	39	1.3
20	Medium	6.9	Loamy sand	Irrigation	8.0	2290	50	55	345	14	700	340	54	39	1.3
21	Medium	5.2	Loamy sand	Domestic	8.0	2250	60	45	348	21	720	320	56	34	1.3
22	Low	2.9	Loamy sand	Domestic	8.0	1710	40	60	230	14	530	230	70	31	1.4
23	Low	3.5	Silt loam	Irrigation	7.8	1785	45	50	250	22	560	250	49	41	1.2
24	Low	4.2	Silt loam	Irrigation	7.9	1800	50	45	260	19	600	235	42	41	1.2
25	Low	5.3	Silt loam	Irrigation	7.8	2510	70	50	358	27	790	310	79	39	1.6
26	Low	4.1	Loamy sand	Domestic	8.1	2710	75	50	425	29	820	385	93	33	1.6
27	Low	3.9	Loamy sand	Irrigation	8.2	2815	70	60	429	19	880	400	85	38	1.7
28	Low	4.7	Loamy sand	Irrigation	8.2	2770	75	60	453	21	890	420	79	37	1.9
29	Low	3.5	Marine clay	Cloth washing	7.1	3490	75	75	588	17	380	940	130	21	0.6
30	Low	2.7	Marine clay	Cloth washing	7.0	3910	70	80	648	22	370	1045	145	19	0.7
Average (1–30)					7.73	2027.33	51.67	47.87	305.93	18.97	596.00	308.33	59.50	34.13	1.15
Standard deviation					0.30	612.75	13.15	12.89	110.34	6.16	134.26	199.38	27.90	6.61	0.33
Coefficient of variation					3.88	30.22	25.45	26.93	36.07	32.47	22.53	64.66	46.89	19.37	28.70
Average (1–28)					7.78	1907.86	50.18	45.75	283.64	18.93	611.79	259.46	53.93	35.14	1.19
Average (29 and 30)					7.05	3700.00	72.15	77.50	618.00	19.50	375.00	992.50	137.50	20.00	0.65

SO_4^{2-} was determined, using turbidimetric procedure, and the NO_3^-, using colorimetric method. The F^- was measured, using spadns method. The chemical variables (except pH) are expressed in milligrams per litre (mg/L).

3.2 *Ionic_balance_error*

Ionic_balance_error (equation 1) between the total concentrations of cations, Ca^{2+}, Mg^{2+}, Na^+ and K^+ (TCC) and the total concentrations of anions, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^- (TCA) for each groundwater sample is within the acceptable range of ±5%.

$$\text{Ionic\_balance\_error} = \left[\frac{(\text{TCC} - \text{TCA})}{(\text{TCC} + \text{TCA})}\right] \times 100 \quad (1)$$

3.3 *Multivariate statistical analysis*

STATISTICA (version 6) is used for the PCA dataset of 11 chemical variables (pH, EC, Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^-) measured for 30 groundwater samples. The varimax rotation option renders it to output principal components (PCs). Kaiser's criterion (Kaiser 1958) of PCs with eigenvalue greater than one is made use of in this exploratory pursuit. Raghunath *et al.* (2002) and Liu *et al.* (2003) reported a detailed PCA for groundwater quality data. The correlation coefficient of the raw data (30 × 11) matrix, its eigenvalue, and eigenvectors are computed. Based on percent of explainability of variance of the data matrix, two PCs are found adequate (figure 3). Table 2 depicts the knowledge
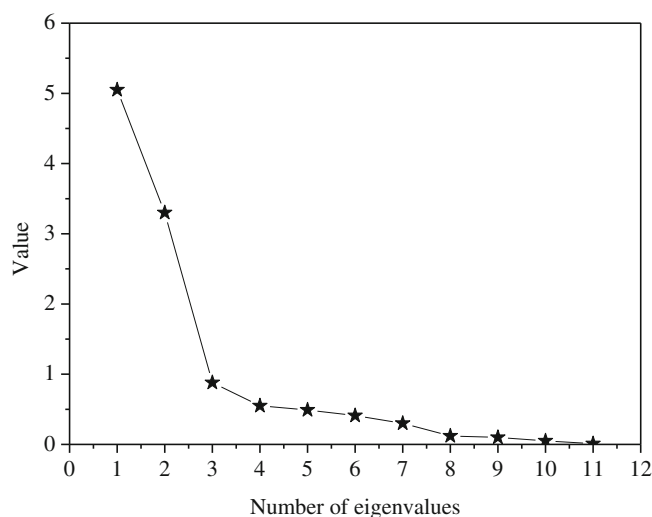

Figure 3. Scree plot.

4. Results and discussion

4.1 *Groundwater chemistry*

The pH of groundwater in the Varaha River basin varies from 7.1 to 8.2, indicating an alkaline nature (table 1). The EC is between 1370 and 3910 (µS/cm), which is a measure of a material's ability to conduct an electric current. The higher the EC, the greater is the enrichment of salts in groundwater. A low enrichment of salts is observed from the upstream area and a high enrichment of salts from the downstream area.

The concentrations (mg/L) of Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^- ions vary from 30 to 75, 25 to 80, 192 to 648, 10 to 38, 450 to 890, 130 to 1,045, 21 to 145, 19 to 53 and 0.60 to 1.90, respectively (table 1). The cations enter groundwater system through the incongruent dissolution of feldspars and ferromagnesium minerals of the host rocks (Drever 1997). The agricultural activities, domestic wastes, and marine source also contribute the Na^+, Mg^{2+} and K^+ ions to the groundwater (Todd 1980; Hem 1991; Subba Rao 2002; Subba Rao *et al.* 2005). The very high concentration of Na^+ is due to its high solubility, while the very low concentration of K^+ is a result of its adsorption on clay minerals.

During the infiltration of recharge, the water adsorbs a large amount of CO_2 released from soil, which is mainly from decay of organic matter and root respiration (equations 2–4). In weathering reactions, it is converted to HCO_3^- salts (Jacks 1973; Berner and Berner 1987). In the present study area, the groundwater samples 1–28 contain HCO_3^- concentration more than that of Cl^- (table 1). The higher HCO_3^- in the groundwater infers a dominance of mineral dissolution (Stumm and Morgan 1996). The source of Cl^- is of non-lithological origin. However, it also contributes from the clay products formed by rock-weathering due to sluggish drainage conditions (Hem 1991). There is no evidence of sulphide-bearing minerals in the study area. The decaying plant and animal sources as well as the uncontrolled usage of fertilizers and soil amendments contribute to the presence of SO_4^{2-} in the groundwater (Subba Rao *et al.* 2012a, b). The source of NO_3^- (>10 mg/L) is mostly poor sanitary and irrigation practices (Cushing *et al.* 1973). The

base and MATLAB functions developed in house for exploratory analysis, which have been in use in computer augmented instruction in geophysical data analysis. These two PC axes are mutually orthogonal and the PC scores are a projection of the data on to the PC axes.

Table 2. *MATLAB functions for exploratory analysis.*

```
function [xmean,                             % Knowledge base
xstd, xcov, xcorrcoef, xmeancenterd, zscore] =
om_scalx(xraw)                               tol = 1.0e-6
                                             if meanx < tol & abs(stdx-1) > tol
[rx,cx] = size(xraw);                          disp(['Raw data itself is mean centered'])
 xmean = sum(xraw)/rx;
 xstd = std(xraw);                           end
 xcov = cov(xraw);                           if meanx < tol & abs(stdx-1) < tol
 xcorrcoef = corrcoef(xraw);                 disp(['Raw data itself is variance scaled'])
 xmeancenterd = xraw-ones(rx,1) * xmean;     end
                                             if meanx > tol
 for i = 1:cx                                  disp(['Data is not mean centered or
  xms(:, i) = xmeancenterd(:, i)./xstd(:, i);   variance scaled'])
 end                                         end

 for j = 1:cx
  c(:, j) = xraw(:, j) - ones(rx, 1)*xmean (j);
  zscore(:, j) = c(:, j)./norm(c(:, j));
 end
end
```

alkaline water mobilizes F^- from the soils/rocks and also releases it from F^--bearing minerals such as apatite, biotite, and clay (equation 5; Jacks *et al.* 1973, 2005; Ramamohana Rao *et al.* 1993; Subba Rao 2003, 2011, 2012b; Madhnure *et al.* 2007). Moreover, the F^- in the groundwater is also enriched by influence of fertilizers under the alkaline environment (Subba Rao 2009a).

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \qquad (2)$$
$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \qquad (3)$$
$$HCO_3^- \rightleftharpoons H^+ + CO_3^{2-} \qquad (4)$$
$$CaF_2 + HCO_3^- \rightleftharpoons CaCO_3 + 2F^- + H^+ \qquad (5)$$

The hydrogeochemical facies changes from $Na^+ > Mg^{2+} > Ca^{2+} > K^+$: $Cl^- > HCO_3^- > SO_4^{2-} > NO_3^- > F^-$ (inferior quality) at low topography to $Na^+ > Ca^{2+} > Mg^{2+} > K^+$: $HCO_3^- > Cl^- > SO_4^{2-} > NO_3^- > F^-$ (relatively better quality) at high topography (table 1).

4.1.1 *Correlation analysis of chemical parameters*

To assess the sources of dissolved salts in the groundwater, a dataset of 11 chemical variables (pH, EC, Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^-) expressed in mg/L, except pH, is used for the correlation analysis (table 3). The pH has a good positive correlation with HCO_3^- ($r = 0.81$) and F^- ($r = 0.69$), and HCO_3^- with F^- ($r = 0.82$), indicating an alkaline environment, in which the dissolved CO_2 played a dominant role. This leads to mineral dissolution, including F^--bearing minerals. A strong positive correlation of EC with Ca^{2+} ($r = 0.77$), Mg^{2+} ($r = 0.78$), Na^+ ($r = 0.99$), Cl^- ($r = 0.91$) and SO_4^{2-} ($r = 0.82$) is observed. It points out that the aquifer chemistry is mainly controlled by Ca^{2+}, Mg^{2+}, Na^+, Cl^- and SO_4^{2-} ions. However, there is a variation in the values of correlation coefficient. This is caused by the difference in the mineral dissolution and solubility, leaching of secondary salts, ion exchange, evaporation, anthropogenic activities and marine sources, following the topographic features and water flow-path conditions (Subba Rao *et al.* 2012b). A high positive correlation coefficient of Ca^{2+} and Mg^{2+} with Na^+ (0.75) reflects the weathering and dissolution of plagioclase feldspars and ferromagnesium minerals. Here, the ion exchange is not evident, as it is masked by the concurrent increase/decrease in the cations due to a result of dissolution/precipitation reactions and concentration effects (Adams *et al.* 2001). The Na^+ has a significant positive correlation with Cl^- ($r = 0.94$) and SO_4^{2-} ($r = 0.82$), indicating the influences of evaporation, domestic wastes, poor drainage conditions and marine source. The Cl^- shows a strong positive correlation with Ca^{2+} ($r = 0.63$) and Mg^{2+} ($r = 0.76$). This is a result of anthropogenic and marine origin. A significant positive correlation of SO_4^{2-} with Mg^{2+} ($r = 0.70$) and Cl^- ($r = 0.82$) indicates the long history of evaporation, leaching of soil salts, anthropogenic activity and marine source (Datta and Tyagi 1996; Park *et al.* 2005; Subba Rao *et al.* 2006, 2012b).

Table 3. *Correlation matrix.*

Chemical variables	pH	EC	Ca^{2+}	Mg^{2+}	Na^+	K^+	HCO_3^-	Cl^-	SO_4^{2-}	NO_3^-	F^-
pH	1.00	0.02	0.05	0.11	−0.01	0.15	**0.81**	−0.28	−0.21	0.56	**0.69**
EC		1.00	**0.77**	**0.78**	**0.99**	0.27	0.26	**0.91**	**0.82**	−0.20	0.21
Ca^{2+}			1.00	0.51	**0.75**	0.14	0.44	**0.63**	0.59	0.01	0.36
Mg^{2+}				1.00	**0.75**	0.01	0.18	**0.76**	**0.70**	−0.20	0.15
Na^+					1.00	0.25	0.20	**0.94**	**0.82**	−0.26	0.15
K^+						1.00	0.23	0.15	0.20	0.12	0.37
HCO_3^-							1.00	−0.12	−0.02	0.57	**0.82**
Cl^-								1.00	**0.82**	−0.44	−0.13
SO_4^{2-}									1.00	−0.37	0.11
NO_3^-										1.00	0.44
F^-											1.00

*Significant at 0.05 level.
Bold denotes correlation coefficient of > 0.60.

4.2 *Exploratory data analysis*

4.2.1 *Scatter plots*

The distribution of raw data for each variable indicates mostly that it is non-Gaussian and emphasizes the trend for above the noise level. Most groundwater sampling points of the study area (1–28) fall below the equiline of Na^+: Cl^-, except two sampling points (29 and 30; figure 4). The excess of Na^+ over Cl^- is a result of rock-weathering (Meyback 1987), while the excess of Cl^- over Na^+ is caused by marine source (Hem 1991). Further, the ascending trend in the plot of Cl^- *vs.* SO_4^{2-} with a correlation of 0.92 (figure 5) indicates a common source of these ions.

The saturation indices of NaCl (average: −2.118) and $CaSO_4$ (average: −1.924) are negative, reflecting their dissolution (figure 6). Further, the groundwater shows a positive saturation index of $CaCO_3$ (average: 2.904) due to evaporation. The occurrence of *kankar* (concretion of $CaCO_3$) in the soil zone supports this observation. On the other hand, the high concentration of Na^+ compared to that of Ca^{2+} in the scatter plot of Na^+ *vs.* Ca^{2+} (figure 7) reflects the evaporation and/or ion exchange.

Langelier and Ludwig (1942) diagram (figure 8) for distribution of Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^- and SO_4^{2-} ions classifies the groundwater samples into two groups, viz., the first is dominated by Na^+–HCO_3^-–Cl^-–SO_4^{2-} ions, and the second by Na^+–Cl^-–SO_4^{2-} ions. Dominances of HCO_3^- and Cl^- ions respectively indicate the persistence of non-saline and saline conditions in the groundwater. All samples in the diagram are far away from the meteoric water zone (Ca^{2+}–Mg^2–HCO_3^-), indicating the pollution activity. It is clear from the diagram that the original chemical characteristics of the meteoric water have been completely obliterated by the entering of non-geogenic
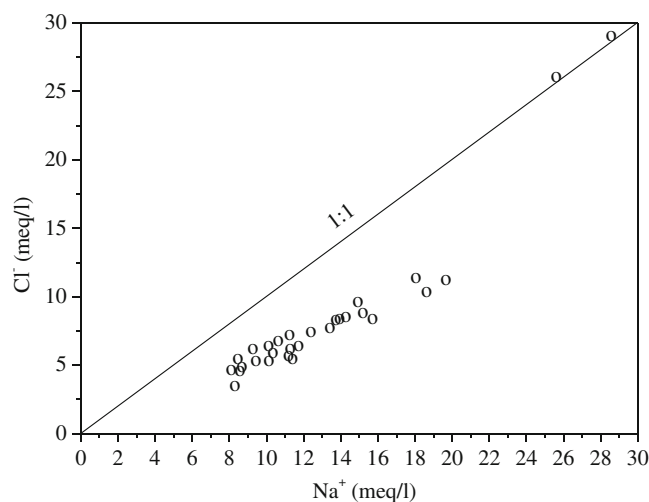

Figure 4. Scatter diagram of variation of Na^+ *vs.* Cl^- in the groundwater samples.
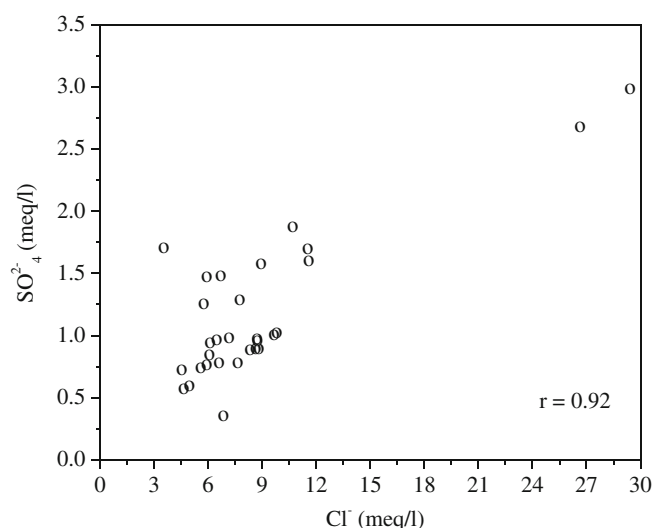

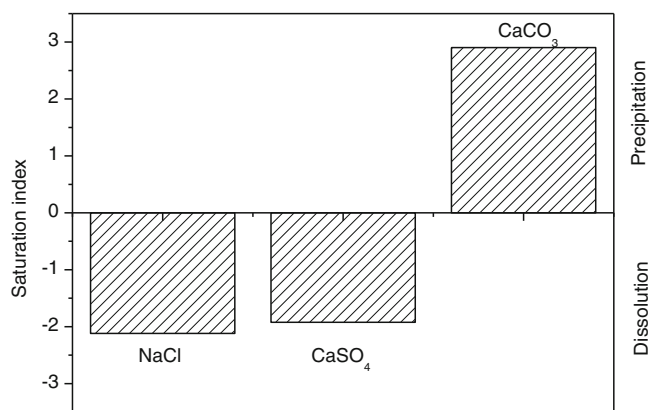Figure 5. Scatter diagram of variation of Cl^- *vs.* SO_4^{2-} in the groundwater samples.

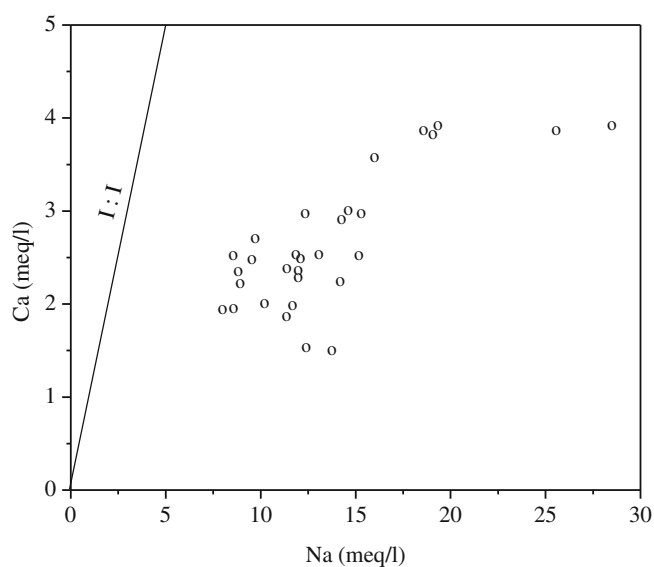Figure 6. Saturation index with respect to CaCO₃, CaSO₄ and NaCl in the groundwater.

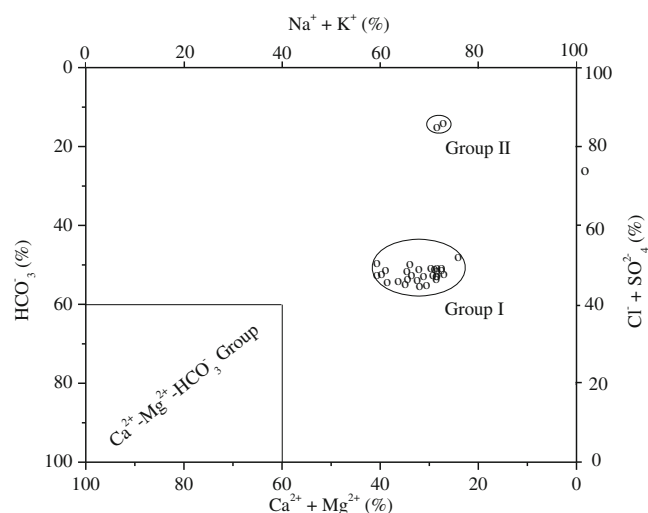Figure 7. Scatter diagram of variation of Na^+ *vs.* Ca^{2+} in the groundwater samples.

Figure 8. Langelier and Ludwig (1942) diagram for discrimination of quality of groundwater.

matter (anthropogenic into marine sources) into the groundwater body. Thus, group I is a result of geogenic (mineral dissolution, ion exchange and evaporation) and anthropogenic (agricultural activities and domestic wastes) origins, while the group II ions are of marine origin (marine clay).

4.3 *Principal component analysis (PCA)*

The PCA reduces the chemical variables (pH, EC, Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^-) into two orthogonal principal components (PC1 and PC2) in the study area of the Varaha River basin (figure 9) based on Scree plot (figure 3). They account for 75.45% of the total variance of the data matrix. The PC1 shows 46.07% of the total variance with an eigenvalue of 5.068. The PC2 with an eigenvalue of 3.231 accounts for 29.38% of the total variance. The communalities of chemical variables vary from 0.518 to 0.999, indicating an involvement of different contributions in changing of quality of groundwater.

4.3.1 *First principal component (PC1)*

The high positive PC loadings of EC (0.975), Na^+ (0.973), Cl^- (0.949), SO_4^{2-} (0.893), Mg^{2+}
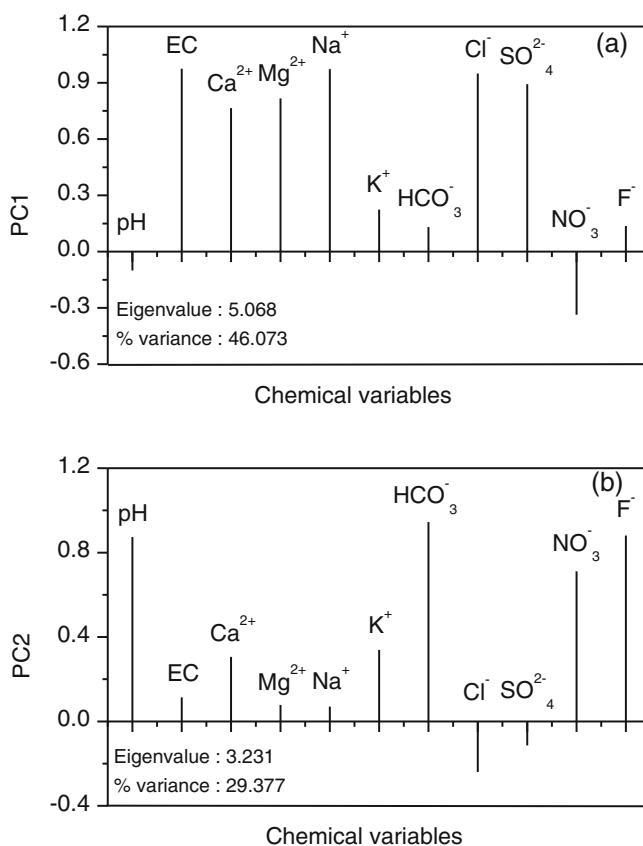

Figure 9. Variation of PC loadings with the communalities of 0.518–0.999.

(0.817) and Ca^{2+} (0.765) are observed for the PC1 (figure 9). The positive PC loadings are heuristically classified into five types to probe into the relative contribution of chemical variables to groundwater quality. These classes based on numerical value of PC loadings are with (a) very low positive (<0.450); (b) low positive (0.450–0.600); (c) medium positive (0.600–0.750); (d) high positive (0.750–0.900); and (e) very high positive (>0.900). Thus, the EC, Na^+ and Cl^- have very high positive PC loadings, while the Mg^{2+}, Ca^{2+} and SO_4^{2-} show high positive PC loadings. Further, the Na^+ and Cl^- ions have higher positive values compared to that of SO_4^{2-} ion due to differences in their solubility.

In PC1 scores, the negative values (−0.028 to −0.982) are observed from the groundwater samples of 1–17 and 22–24, and the positive values (0.051–3.227) from the groundwater samples of 18–21 and 25–30 (figure 10). The positive PC scores are categorized on an adhoc basis into three types to probe into the relative influences of controlling processes on overall quality of groundwater. They are with (a) low positive (<1.000); (b) medium positive (1.000–2.000); and (c) high positive (>2.000) PC scores. Mostly in upstream area (figure 1), 20 groundwater samples (1–17 and

Figure 10. Variation of PC scores with different groundwater quality well samples.

22–24) show negative PC1 scores (−0.028 to −0.982) in the PC1 direction. This reflects a low dissolution of Na^+, Cl^-, SO_4^{2-}, Mg^{2+} and Ca^{2+} ions due to lesser interaction of water with the aquifer material due to high topography and/or recharge area. Thus, a low enrichment of salts, Na^+ (249 mg/L), Cl^- (225 mg/L), SO_4^{2-} (48 mg/L), Ca^{2+} (44 mg/L) and Mg^{2+} (42 mg/L), reflected in EC (1671 μS/cm) are observed from the upstream area (table 4).

The low positive PC1 scores (0.051–0.645) for the groundwater samples (18–21 and 25) show higher average values of EC (2100 μS/cm), Na^+ (305 mg/L), Cl^- (295 mg/L), SO_4^{2-} (55 mg/L), Mg^{2+} (55 mg/L) and Ca^{2+} (54 mg/L) compared to those of the respective average values of the negative PC1 scores (table 4; figure 10). In the region of the medium positive PC1 scores (1.052–1.154), there is a higher rate of dissolution of ions. Thus, the medium positive PC1 scores for the groundwater samples (26–28) show high average values of EC value of 2765 μS/cm, Na^+ (436 mg/L), Cl^- (402 mg/L), SO_4^{2-} (86 mg/L), Ca^{2+} (73 mg/L) and Mg^{2+} (57 mg/L) than those of the average values in the low positive PC1 scores. Progressively in the high positive PC1 scores (2.724 to 3.227) region for the groundwater samples (29 and 30), there is an increase in the average values of EC (3700 μS/cm), Na^+ (618 mg/L), Cl^- (993 mg/L), SO_4^{2-} (138 mg/L) and Mg^{2+} (78 mg/L), except Ca^{2+} content (72 mg/L), compared to those of the average values in the medium positive PC1 scores. This is explained based on low topography and/or discharge area, resulting in higher interaction of water with the aquifer material.

The low positive PC1 scores for the sampling points in the study area (figure 11a) are spread in the eastern, western and southern parts covered with the loamy sand (coarse-grained) of medium topography and medium water levels. The medium positive PC1 scores observed from the southeastern part, which has loamy sand, is low-lying and has shallow water levels. The high positive PC1 scores are located in the extreme southeastern part, where the area is covered with marine clay, is low-lying and has shallow water levels.

Since there is a progressive enrichment of salts from higher to lower topography evident from the spatial distribution of the PC1 scores (figure 11a), depending upon the availability of geogenic, anthropogenic and marine sources in the direction of groundwater flow, the PC1 corresponds to salinity-based phenomena.

4.3.2 *Second principal component (PC2)*

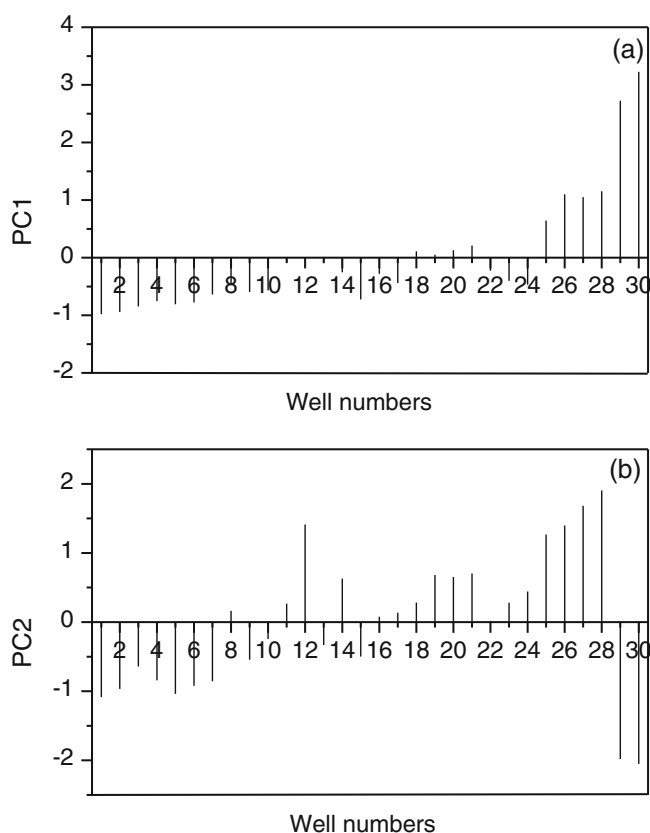The high positive PC2 loadings are observed for HCO_3^- (0.945), F^- (0.881), pH (0.874) and NO_3^-

Table 4. *Average concentrations of influential chemical variables in PC score space.*

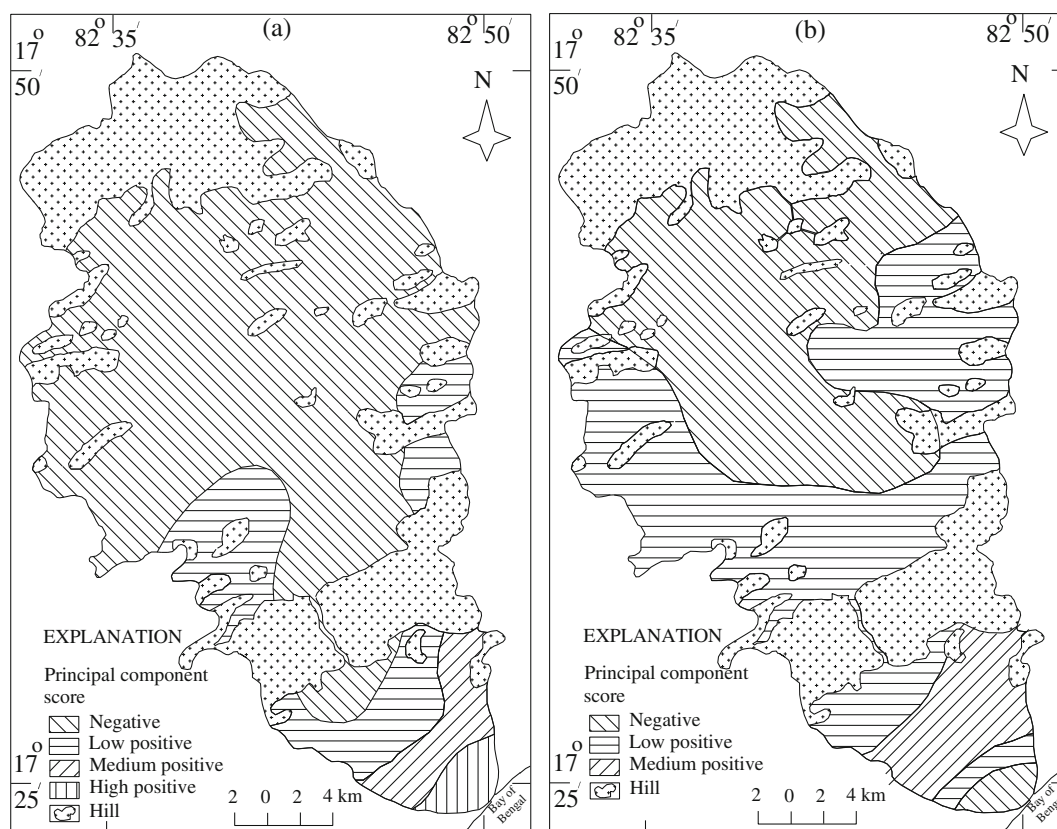|	Average concentrations						
	PC1scores				PC2 scores		
Chemical variables	Negative	Low positive	Medium positive	High positive	Negative	Low positive	Medium positive
pH	–	–	–	–	7.42	7.95	8.01
EC (μS/cm)	1671.00	2100.00	2765.00	3700.00			
Ca^{2+} (mg/L)	44.25	54.00	73.33	72.15	–	–	–
Mg^{2+} (mg/l)	42.30	55.00	56.67	77.50	–	–	–
Na^+ (mg/L)	249.10	305.00	435.67	618.00	–	–	–
HCO_3^- (mg/L)	–	–	–	–	485.77	629.62	845.00
Cl^- (mg/L)	225.25	295.00	401.67	992.50	–	–	–
SO_4^{2-} (mg/L)	48.45	55.00	85.67	137.50	–	–	–
NO_3^- (mg/L)	–	–	–	–	29.92	37.54	36.76
F^- (mg/L)	–	–	–	–	0.89	1.24	1.70

Figure 11. Distribution of **(a)** PC1 scores and **(b)** PC2 scores on the study area map.

(0.711; figure 9). The higher positive loading of NO_3^- is attributed to pollution activity. The hydrochemical parameters, pH, HCO_3^- and F^-, pertain to a set of alkalinity-based phenomenon.

The variations in the concentrations of chemical species from the sign and magnitude of the PC2 loadings are due to differences in the dissolution of ions. Therefore, the observed low average values of pH (7.42), HCO_3^- (486 mg/L), NO_3^- (30 mg/L) and F^- (0.89 mg/L) are reflected as the negative PC2 scores (-0.245 to -0.2051) for the

groundwater samples (1–7, 9, 10, 13, 15, 29 and 30; table 4; figure 10). The medium average values of pH (7.95), HCO_3^- (630 mg/L), NO_3^- (38 mg/L) and F^- (1.24 mg/L) are in the low positive PC2 scores (0.018 to 0.702) for the groundwater samples (8, 11, 14 and 16–24). The high average values of pH (8.01), HCO_3^- (845 mg/L), NO_3^- (37 mg/L) and F^- (1.70 mg/L) are found in the medium positive PC2 scores (1.266 to 1.904) for the groundwater samples (12 and 25–28). The rising of alkalinity is responsible for the increase in the concentration

of F$^-$ almost to double the value (0.89–1.70 mg/L). This is caused by the shifting of PC scores from negative (pH: 7.42 and HCO$_3^-$: 486 mg/L) to medium (pH: 8.01 and HCO$_3^-$: 845 mg/L).

Because of the variations in the concentrations of chemical variables in PC2, the negative PC2 scores are spread mostly in the northern part and to some extent in the extreme southeastern region (figure 11b). The low positive PC2 scores are in the northeastern, southeastern and southwestern zones. In hydrogeological parlance, the study area is mostly loamy sand zone, which is covered with medium topography, and has medium water levels (table 1). On the other hand, the medium positive PC2 scores are observed from the southeastern part, which is a loamy sand zone of low topography and shallow water levels.

4.3.3 *Analysis in loading and score space*

4.3.3.1 *PC loadings*: In the plot of loadings of PC1 *vs.* PC2 (figure 12a), the chemical variables are divided into five clusters based on visual inspection. In the cluster-1, EC, Na$^+$, Mg^{2+} and Ca^{2+},
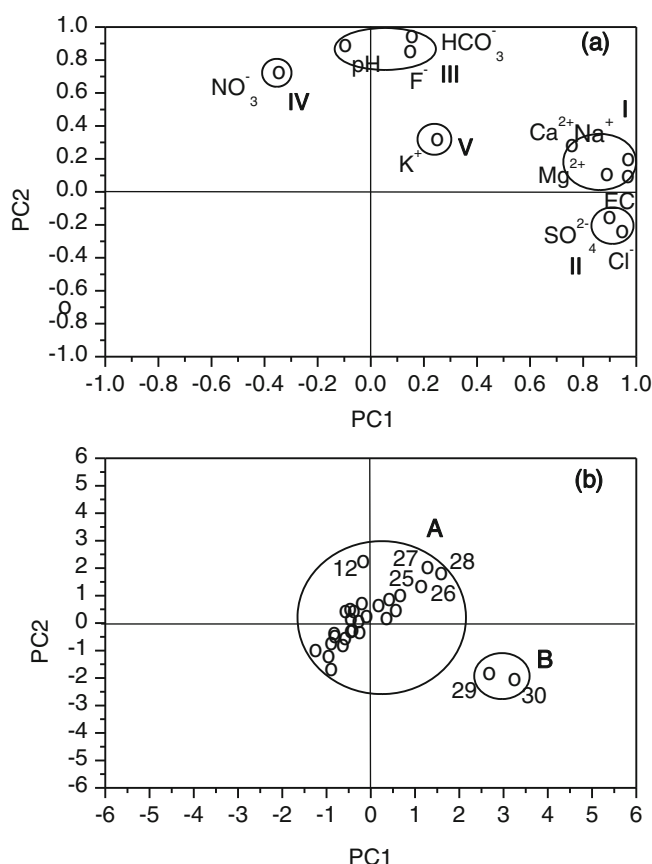
have high to very high positive loadings (0.765–0.975) for the PC1 and very low positive loadings (0.070–0.306) for the PC2. This cluster is attributed to a salinity controlled process based on the geochemical knowledge pertaining to the geogenic, anthropogenic and marine origins. The second cluster (Cl$^-$ and SO$_4^{2-}$) with very high positive loadings (0.893–0.949) for the PC1 and very low negative loadings (-0.114 to -0.240) for the PC2 is assigned to the anthropogenic and marine controlled processes. The cluster-3 (HCO$_3^-$, pH and F$^-$) also has high to very high positive loadings (0.881–0.945) for the PC2 and very low negative loading, of pH (-0.100), as well as very low positive loadings of F$^-$ (0.137) and HCO$_3^-$ (0.131) for the PC1. This cluster represents a set of alkaline controlled process of geogenic origin. The NO$_3^-$ ion forms a singleton cluster-4 with the medium positive loading (0.711) for the PC2 and very low negative loading (-0.336) for the PC1. This explains an anthropogenic activity, leading to pollution. The adsorption of K$^+$ on clay minerals is a geogenic process and results in the second singleton for cluster-5 with low positive loadings (0.224–0.339) on both the PCs.

4.3.3.2 *PC scores*: The PC scores plot of PC1 *vs.* PC2 (figure 12b) show two distinct non-overlapping clusters (A and B). Cluster A with more than 25 sampling points (1–28) reflects the influence of both agricultural (75%) and domestic (25%) activities on the quality of groundwater (table 5). The distribution of two contributions in PC space is overlapping and thus not possible to resolve into component influences with the PCA alone.

The cluster B (29 and 30) with the high positive PC scores (2.724–3.227) on the PC1 and high negative PC scores (-1.982 to -2.051) on the PC2 (figure 12b) corresponds to salinity-based marine pollution rather than a geogenic or anthropogenic based one.

4.4 *Index wells*

Generally, a selection of a large number of wells for continual intensive long-term monitoring of groundwater quality is a welcome feature, but it is infeasible and expensive. A pragmatic way is in selecting a subset of sampling sites from clusters of high positive PC scores, representing the entire study area of the Varaha River basin.

The well 28 is retained from the cluster A, which shows high positive PC scores on both the PC axes (figure 12b). Inclusion of well 30 from cluster B, as well as 27 and 12 from cluster A will increase the information content.

Figure 12. (**a**) Scatter diagram of (**a**) PC1 loadings *vs.* PC2 loadings and (**b**) PC1 scores *vs.* PC2 scores.

Table 5. *Influences of topography, soil type, depth of water levels and water use on the quality of groundwater.*

Hydrogeological characteristics	PC1		PC2		PC1 *vs.* PC2	
	NW (10)	%	NW (17)	%	NW (8)	%
Topography						
Low	6	60	7	41.18	4	50
Medium	4	40	9	52.94	4	50
High	–	–	1	5.88	–	–
Soil type						
Loamy sand	7	70	13	76.47	7	87.5
Silt loam	1	10	4	23.53	1	12.5
Marine clay	2	20	–	–	–	–
Depth of water level						
Shallow	4	40	5	29.41	2	25
Medium	4	40	7	41.18	5	62.5
High	2	20	5	29.41	1	12.5
Water usage						
Irrigation	6	60	12	70.59	6	75
Domestic	2	20	5	29.41	2	25
Cloth-washing	2	20	–	–	–	–

NW = number of wells.

5. Conclusions

The following conclusions are drawn from the present study area:

- Groundwater is characterized by alkaline nature with Na^+ as a dominant ion. HCO_3^- is the abundant ion in most of the groundwater samples. Cl^- is the predominant ion in a few water samples.
- The PCA transforms the physico-chemical variables, pH, EC, Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, Cl^-, SO_4^{2-}, NO_3^- and F^-, into two orthogonal principal components (PC1 and PC2), accounting for 75% of the total variance of the data matrix. The loadings of PC1 are high positive for EC, Na^+, Cl^-, SO_4^{2-}, Mg^{2+} and Ca^{2+}, representing a salinity-based phenomenon with geogenic, anthropogenic and marine origins. The loadings of PC2 are represented by high positive for HCO_3^-, F^-, pH and NO_3^- corresponding to alkalinity and pollution activities. These ions are from geogenic and anthropogenic origins, respectively.
- Mineral dissolution is the dominant geogenic controlling process for the source of ions in the groundwater system. Other influencing processes are ion exchange and evaporation. Agricultural related activities, domestic wastes and marine clays are the additional sources of ions, the former being the most dominant contributor, followed by the second and third categories.
- The spatial distribution of PC scores indicates that the quality of groundwater of geogenic origin is deteriorated progressively from the upstream to the downstream area by anthropogenic and marine sources, depending upon the topography, soil type, depth of water levels and water usage. Thus, the hydrogeochemical facies change from $Na^+ > Ca^{2+} > Mg^{2+} > K^+$: $HCO_3^- > Cl^- > SO_4^{2-} > NO_3^- > F^-$ at high topography to $Na^+ > Mg^{2+} > Ca^{2+} > K^+$: $Cl^- > HCO_3^- > SO_4^{2-} > NO_3^- > F^-$ at low topography.
- The index wells with the highest positive PC scores are selected for the study area to minimize the continuous long-term monitoring of groundwater quality.
- The PCA is an effective tool in exploratory data analysis to probe into the spatial controlling processes. It sheds light on implementing the scientific remedial management measures for controlling the groundwater contamination.

References

Adams S, Tredoux G, Harris C, Titus R and Pietersen K 2001 Hydrochemical characteristics of aquifers near Sutherland in the western Karoo, South Africa; *J. Hydrol.* **241** 91–103.

APHA 1999 *Standard methods for the examination of water and wastewater*; American Public Health Association, Washington DC.

Aravindan S, Manivel M and Chandrasekar S V N 2004 Groundwater quality in the hard rock area of the Gadilam River basin, Tamilnadu; *J. Geol. Soc. India* **63** 625–635.

Berner E K and Berner R A 1987 *The global water cycle, geochemistry and environment*; Prentice-Hall, New Jersey.

Bhatt K B and Saklani S 1996 Hydrogeochemistry of the Upper Ganges River, India; *J. Geol. Soc. India* **48** 171–182.

Cushing E M, Kantrowitz I H and Taylor K R 1973 *Water resources of the Delmarva Peninsula*; US Geological Survey Professional Paper 822, Washington DC, 58p.

CGWB 2001 *Hydrogeological framework and development prospects in Visakhapatnam district, Andhra Pradesh*; Technical Report of Central Ground Water Board, Ministry of Water Resources, Government of India, Southern Region, Hyderabad, India.

Datta P S and Tyagi S K 1996 Major ion chemistry of groundwater in Delhi area – chemical weathering process and groundwater flow regime; *J. Geol. Soc. India* **47** 179–188.

Drever J L 1997 *The geochemistry of natural waters*; 3rd edn, Prentice Hall, Englewood, 436p.

Elampooranan T, Rajmohan N and Abirami L 1999 Hydrochemical studies of Artesian well waters in Cauvery deltaic area, South India; *Indian J. Environ. Health* **41** 107–114.

Giridharan L, Venugopal T and Jayaprakash M 2008 Evaluation of the seasonal variation on the geochemical parameters and quality assessment of the groundwater in the proximity of River Cooum, Chennai, India; *Environ. Monit. Assess.* **143** 161–178.

Gopinath G and Seralathan P 2006 Chemistry of groundwater in the laterite formation of Muvattupuzha river basin, Kerala; *J. Geol. Soc. India* **65** 705–714.

Hem J D 1991 *Study and interpretation of the chemical characteristics of natural water*; 2254, Scientific Publishers, Jodhpur, India, 263p.

Jacks G 1973 Chemistry of groundwater in a district in southern India; *J. Hydrol.* **18** 185–200.

Jacks G, Bhattacharya P, Chaudhary V and Singh K P 2005 Controls on the genesis of some high-fluoride groundwaters in India; *Appl. Geochem.* **20** 221–228.

Jacks G, Rajagopalan K, Alveteg T and Johnsson M 1993 Genesis of high-F groundwaters, southern India; *Appl. Geochem. (Suppl.)* **2** 241–244.

Jeevanandam M, Kannan R, Srinivasalu S and Rammohan V 2006 Hydrogeochemistry and groundwater quality assessment of lower part of the Ponnaiyar River basin, Cuddalore District, south India; *Environ. Monit. Assess.* **132** 263–274.

John Devadas D, Subba Rao N, Thirupathi Rao B, Srinivasa Rao K V and Subrahmanyam A 2007 Hydrogeochemistry of the Sarada River basin, Visakhapatnam District, Andhra Pradesh, India; *Environ. Geol.* **52** 1331–1342.

Kaiser H F 1958 The varimax criterion for analytic rotation in factor analysis; *Psychometrika* **23** 187–200.

Krishna Kumar S, Rammohan V, Dajkumar Sahayam J and Jeevanandam M 2009 Assessment of groundwater quality and hydrogeochemistry of Manimuktha River basin, Tamil Nadu, India; *Environ. Monit. Assess.* **159** 341–351.

Khurshid S and Zaheerudin 2004 Geochemistry of groundwater: An overview of sporadic fluoride and nitrate contamination in parts of Yamuna basin; *J. Appl. Geochem.* **6** 25–35.

Langelier W F and Ludwig H F 1942 Graphic method for indicating the mineral character of natural water; *J. Am. Water Works Association* **34** 335–352.

Liu Ch W, Lin K H and Kuo Y M 2003 Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan; *The Sci. Total Environ.* **313** 77–89.

Madhnure P, Sirsikar D Y, Tiwari A N, Ranjan B and Malpe D B 2007 Occurrence of fluoride in the groundwaters of Pandharkawada area, Yavatmal District, Maharashtra, India; *Curr. Sci.* **92** 675–679.

Meyback M 1987 Global chemical weathering of surficial rocks estimated from river dissolved loads; *Am. J. Sci.* **287** 401–428.

Mithas A D, Sankar K and Imran A D 2011 Major ion chemistry and hydrochemical studies of groundwater of parts of Palar river basin, Tamil Nadu, India; *Environ. Monit. Assess.* **176** 621–636.

Mridul C, Soumya C, Saumen B, Manash J N, Lokendra S, Ravi B, Srivastava H P and Sarma T 2011 Groundwater arsenic contamination in Brahmaputra river basin: A water quality assessment in Golaghat (Assam), India; *Environ. Monit. Assess.* **173** 371–385.

Murthy K S R 2000 Groundwater potential in a semi-arid region of Andhra Pradesh – A geographical information system approach; *Int. J. Remote Sens.* **21** 1867–1884.

Naik P K, Awasthi A K, Anand A V S S and Behera P N 2009 Hydrogeochemistry of the Koyna River basin, India; *Environ. Earth Sci.* **59** 613–629.

Papiya M, Rahul U and Aziz H 2010 Seasonal and spatial variation of Yamuna River water quality in Delhi, India; *Environ. Monit. Assess.* **170** 661–670.

Park S C, Yun S T, Chae G T, Yoo I Y, Shin K S, Heo C H and Lee S H 2005 Regional hydrochemical study on salinization of coastal aquifers, western coastal area of South Korea; *J. Hydrol.* **313** 182–194.

Raghunath R, Murthy T R S and Raghavan B R 2002 The utility of multivariate statistical techniques in hydrochemical studies: An example from Karnataka, India; *Water Res.* **36** 2437–2442.

Rajmohan N and Elango L 2006 Hydrogeochemistry and its relations to groundwater level fluctuations in the Palar and Cheyyar river basins, southern India; *Hydrol. Process.* **20** 2415–2427.

Raju N J 2009 Groundwater quality in the lower Varuna River basin, Varanasi District, Uttar Pradesh; *J. Geol. Soc. India* **73** 78–192.

Ramamohana Rao N V, Suryaprakasa Rao K and Schuiling R D 1993 Fluorine distribution in waters of Nalgonda District, Andhra Pradesh, India; *Environ. Geol.* **21** 84–89.

Ravikumar P, Somashekar R K and Ai Mhasizonuo 2011 Hydrochemistry and evaluation of groundwater suitability for irrigation and drinking purposes in the Markandeya River basin, Belgaum District, Karnataka State, India; *Environ. Monit. Assess.* **173** 459–487.

Shankar K, Aravindan S and Rajendran S 2011 Hydrogeochemistry of the Paravanar River sub-basin, Cuddalore District, Tamilnadu, India; *E-J. Chem.* **8** 835–845.

Sreedevi P D 2004 Groundwater quality of Pageru River basin, Cuddapah District, Andhra Pradesh; *J. Geol. Soc. India* **64** 619–636.

Stumm W and Morgan J J 1996 *Aquatic chemistry*; John Wiley and Sons, New York, 1022p.

Subba Rao N 2002 Geochemistry of groundwater in parts of Guntur District, Andhra Pradesh, India; *Environ. Geol.* **41** 552–562.

Subba Rao N 2003 Groundwater quality – focus on fluoride concentration in rural parts of Guntur District, Andhra Pradesh, India; *Hydrol. Sci. J.* **48** 835–847.

Subba Rao N 2009a Fluoride in groundwater, Varaha River basin, Visakhapatnam District, Andhra Pradesh, India; *Environ. Monit. Assess.* **152** 47–60.

Subba Rao N 2009b A numerical scheme for groundwater development in a watershed basin of basement terrain: A case study from India; *Hydrogeol. J.* **17** 379–396.

Subba Rao N 2011 High-fluoride groundwater; *Environ. Monit. Assess.* **176** 637–645.

Subba Rao N, Devadas John and Srinivasa Rao K V 2006 Interpretation of groundwater quality using principal component analysis from Anantapur District, Andhra Pradesh, India; *Environ. Geosci.* **13** 239–259.

Subba Rao N, Saroja Nirmal I and Suryanarayana K 2005 Groundwater quality in a coastal area – a case study from Andhra Pradesh, India; *Environ. Geol.* **48** 534–550.

Subba Rao N and Surya Rao P 2010 Major ion chemistry of groundwater in a river basin: A study from India; *Environ. Earth Sci.* **61** 757–775.

Subba Rao N, Subrahmanyam A, Ravi Kumar S, Srinivasulu N, Babu Rao G, Surya Rao P and Venkatram Reddy G 2012a Geochemistry and quality of groundwater of Gummanampadu sub-basin, Guntur District, Andhra Pradesh, India; *Environ. Earth Sci.* **67** 1451–1471.

Subba Rao N, Surya Rao P, Vidyasagar G, Nagamani G and Satyanarayana N L V V 2012b Chemical characteristics of groundwater and assessment of groundwater quality in Varaha River basin, Visakhapatnam District, Andhra Pradesh, India; *Environ. Earth Sci.* **184** 5189–5214.

Subramani T, Elango L and Damodarasamy S R 2005 Groundwater quality and its suitability for drinking and agricultural use in Chithar River basin, Tamil Nadu, India; *Environ. Geol.* **47** 1099–1110.

Todd D K 1980 Groundwater Hydrology; John Wiley & Sons, New York, 535p.

Umar R and Absar A 2003 Chemical characteristics of groundwater in parts of the Gambhir River basin, Bharatpur District, Rajasthan, India; *Environ. Geol.* **44** 535–544.