International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2017; 5(3): 80-85

Received: 13-03-2017
Accepted: 14-04-2017

G Satyanarayana
School of Chemistry, Andhra University, Visakhapatnam, Andhra Pradesh, India

P Ramadas
Department of Geophysics, Andhra University, Visakhapatnam, Andhra Pradesh, India

P Padmavathi Devi
Department of Geophysics, Andhra University, Visakhapatnam, Andhra Pradesh, India

NVBSS Prasad
Department of Geophysics, Andhra University, Visakhapatnam, Andhra Pradesh, India

G Nageswara Rao
School of Chemistry, Andhra University, Visakhapatnam, Andhra Pradesh, India

Correspondence
G Nageswara Rao
School of Chemistry, Andhra University, Visakhapatnam, Andhra Pradesh, India

Ground water quality interpretation in coastal Srikakulam, Andhra Pradesh by using principal component analysis

G Satyanarayana, P Ramadas, P Padmavathi Devi, NVBSS Prasad and G Nageswara Rao

Abstract
An attempt was made to study the ground water quality. Water samples were collected from 51locations in Uddanam region and analyzed for 10 water quality variables. The principal components analysis (PCA) applied to define the parameters responsible for the main variability in water quality variance for Uddanam ground water. The PCA produce three significant main components and explain more than 77% of the variance. Namely, seasonal effects, agricultural effects and geological effect that represent 54.8%, 12.53%, and 11.01% respectively, of the total variance of water quality in Uddanam area. The results of PCA reflected a good look on the water quality monitoring and interpretation of the groundwater.

Keywords: Variance, PCA, groundwater, water quality

1. Introduction

The study of the hydro chemical evolution in complex aquifers requires manipulation of a wide range of diverse origin. The physicochemical parameters indicate the diversity of the groundwater and orientation of the possible processes that take place through the aquifer. A series of geological variables must also be considered which determine the chemical evolution of the water. The understanding and the particularities of the hydro geochemistry of aquifer systems report a great opportunity to apply multivariate statistical analysis, a quantitative approach allowing classify ground water samples, to study correlations between the sampling sites (Cloutier *et al*., 2008) [26]. The data sets contain rich information about the behaviour of the water body. The classification, modelling and interpretations of monitoring data are the most important steps in the assessment of water quality. Water quality parameters interact with each other to define the resource water quality. Many re-searchers treated water quality parameters individually by describing the seasonal variability and their causes. While it is difficult to interpret all the parameters patterns in combinations (Shihab, 1993) [7]. In recent years, many studies have been done using PCA in the interpretation of water quality parameters. Lohani and Todino (1984) [20] utilized principal components (PC) technique to provide a quick analytical method for the water quality of Chao Phraya river in Thailand. Shihab (1993) [7] used this technique in order to describe the variation in water quality in Al-Mosul dam reservoir and the regulating lake to reduce the number of water quality parameters needed for monitoring the lake water. Mazlum *et al*. (1999) [14] determined factors that caused variations in water quality at the Porsuk Tributary in the Sakarya river basin by using PCA. Researchers referred that PCA is more reliable than factor analysis and it is a pure mathematical technique without any assumption. Also, PCA has been successfully applied to sort out hydro geological and hydro geochemical processes from commonly collected groundwater quality data (see Jayakumar and Siraz, 1997; Salman and Abu Ruka'h 1998; Praus, 2005; Thilagavathi, *et al*., 2012; Olobaniyi and Owoyeni, 2006) [3, 18, 2, 4, 5]. Iyer *et al*. (2003) [21] developed a statistical model which is based on the PCA for coastal water quality data from the Cochin coast in south west India, which explained the relationships between the various physicochemical variables that have been monitored and environmental conditions effect on the coastal water quality. PCA is employed in this study; this technique can be used to obtain relationship between parameters, to identify the factors and sources influencing groundwater quality and to suggest useful tools for both management of water resources and monitoring of groundwater quality in Uddanam area.

2. Study area
Uddanam region is a part of south eastern part of Srikakulam ghat terrain located 200 km from the city of Visakhapatnam. The Uddanam region, covering an area of about 290 sq km, is imposed between "18.56219– 19.16563"N latitude and "84.30934– 84.76844"E longitude (Fig 1.1). The area is covered by survey of India toposheets 65O/1, O/2 & O/3 on 1:50,000 scale. The area is fairly connected to NH-5 in addition to local highways.

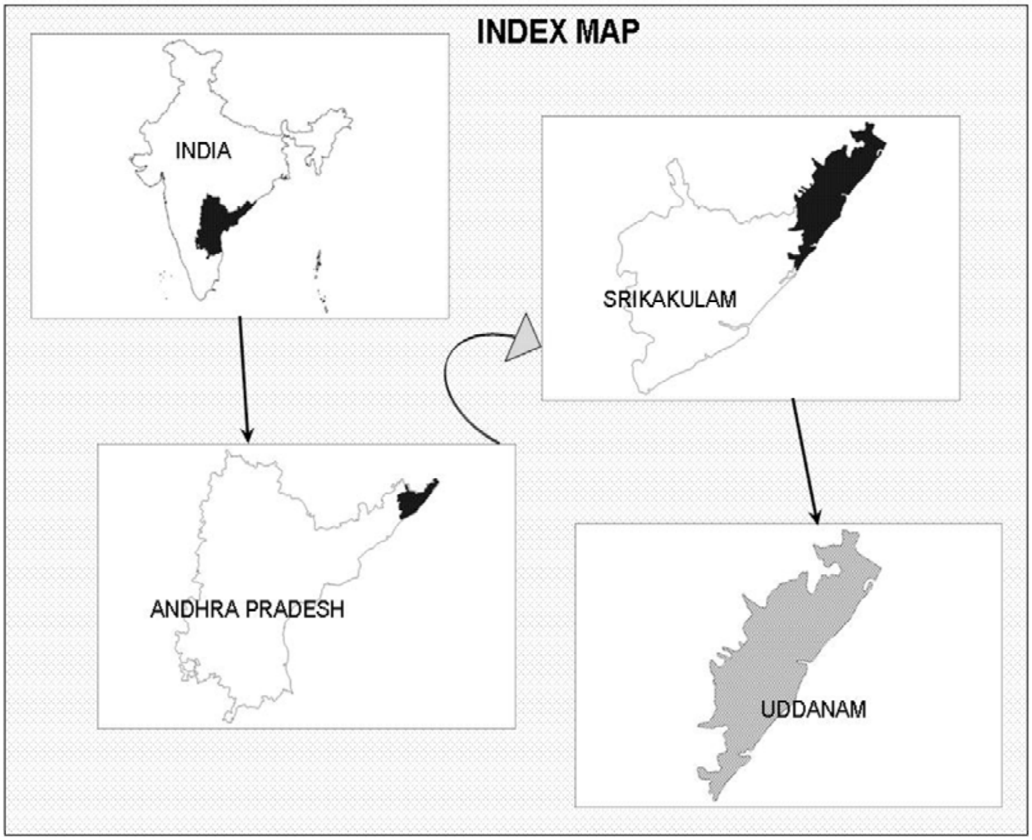

Fig 1: Location and administrative map of Uddanam region

2.1 Communication
The study area, being nearer to the city of Srikakulam, is well connected through roads and rail network. A number of local highways are connected to different locations of nearby NH-5. Tekkali, is another town on NH-5 is also connected to the study area on the western side, while Palasa town provides the major connectivity and approach to the interior villages of problematic. However, Somepeta, Vajrapukotturu, Kanchili located in central part of the affected areas provide connectivity to almost all the villages in that area.

2.2 Drainage
The main rivers drain the district are Vamsadhara and Nagavali. The other important rivers flowing in the district are Suvarnamukhi, Vegavati, Mahendra Tanya and Bahuda. Among the rivers Vamsadhara, Nagavali and Suvarnamukhi are perennial. The general drainage pattern is dendritic to sub-dendritic and occasionally parallel at places. The drainage in western part of the district resembles dendritic type, where as in the central part it is parallel to sub-dendritic. The overall drainage is of medium to coarse textured towards west and north of the district, whereas in central and southern parts it is very coarse. The drainage density varies from less than 0.2 to 1 km/km2. In Palakonda area the density is of 0.6 to 1 km/km2. While areas with a density of more than 0.2 km/sq.km are observed in Ichapuram, Sompeta, Narasannapeta and Srikakulam. The density is less than 0.2 km/sq.km in plain area of Tekkali

3. Sample collection and treatment
A total of 51 locations were selected for the characterization of the physicochemical properties of the ground water. All water samples were collected in a single sampling compaign during monsoon period. All the wells chosen for ground water sampling were commonly utilized for irrigation water supply and/or domestic purposes. Geographic coordinates of each sampling site were recorded with a handhold GPS unit using UTM coordinate system and WGS84 datum. To determine the general quality of the water in the study area, general precautions like keeping the container clean, rinsing with water sample, avoiding re-aeration through trickling water along the sides, sealing the container etc., described by (Hem, J. D., 1991; WHO/GEMS Water quality operational guide, 1978; APHA, 1995) were strictly adopted. The physical parameters like p^H, temperature, electrical conductivity, total dissolved solids, dissolved oxygen, salinity of the water samples were determined in the field using the water analyzer kit. The water samples were analysed for major cations (Na^+, TH, K^+) and anions ($Cl^–$, SO_4^{2-}, HCO_3^-) and minor elements like Silica, Phosphate and Nitrate. The chemical analysis was carried out as per the procedure given in APHA, (1995). The major, minor and trace elements were analyzed in the laboratory using the following methods. Chloride, Bicarbonate, Total Hardness (Titrimetric methods) Nitrate, Silicate (UV Visible Spectrophotometer (190nm-900nm) Sulphate (Nephelometer) Sodium and Potassium (Flame photometer)

3.1 Data Processing

The 10 parameters obtained from the laboratory analysis were used as variables inputs for PCA. Since water quality parameters had different magnitudes and scales of measurements, so the data were standardized to produce a normally distribution of all variables (Davis, 1973) [10]. In the standardization, the raw data were converted to unit of zero mean and a variance of one, by subtracting from each variable the mean of data set and dividing by standard deviation. From the standardized covariance or correlation matrix of the data the initial factor solution were extracted by the multivariate PC extraction, then a number of PC were selected from the initial according to their eigenvalues and scree diagram. Orthogonal rotation of the selected initial components to terminal factor solutions was done by Kaiser's Verimax schemes that attempt to achieve simple structure with respect to both the rows and columns (Kleinbaum *et al*., 1988) [23]. In this study, the eigenvalue one criterion was used to determine the numbers of PCs based on the assumption that only eigenvalues greater than one were considered important and the higher eigenvalues are more significant. Varimax normalization was then applied as the rotation method in the analysis on the PCs for better interpretation of results. Varimax factor loadings of 77% were considered strong, although the terms 'strong', 'moderate' and 'weak' as applied to loadings, refer to absolute loading values of >0.75, 0.5-0.75 and 0.30-0.50, respectively.

4. Discussion

PCA was performed on standardized matrix of the raw data in which a water samples are described by 10 physical and chemical parameters. This technique aims to transform the observed variables to a new set of variables (PC) which are uncorrelated and arranged in decreasing order of importance in order to simplify the problem. Table 5a, represented the determined initial PC and its eigenvalues and percent of variance contributed in each PC. Figure 2, show that the scree plot of the eigenvalues for each component. Eigenvalues accounts and scree plot showed that the first four PC is the most significant components which represent 77% of the variance in water quality of ground water in uddanam area: 54.866% by PC1, 12.530% by PC2, and, 11.011% by PC3; in addition, it have eigenvalues of more than one. Components loading and communalities for each variable in four selected components before: Varimax rotation was explained in Table 5b; and after Varimax rotation in Table 5c. Communalities provide an index to the efficiency of the reduced set of components and degree of contribution of each variable in the selected three components. In ourcase, communalities showed that all the variables have been described to acceptable levels in the selected components with more than 50%. Table 6 shows the correlation matrix of components of Varimax rotated three PC. It is noted that there are no correlation between components each components represent a discrete unit from others. Components loading (correlation coefficients), which measure the degree of closeness between the variables and the PC. The largest loading either positive or negative, suggests the meaning of the dimensions, positive loading indicates that the contribution of the variables increases with the increasing loading in dimension; and negative loading indicates a decrease (Jayakumar and Siraz, 1997) [3]. In general, component's loading larger than 0.6 may be taken into consideration in the interpretation, in other words, the most significant variables in the components represented by high loadings have been taken into consideration in evaluation the components (Mazlum *et al*.,1999) [14].

4.1 Descriptive Statistics for All the Parameters in ppm except (EC& pH)

Table 1: Descriptive statistics

	N	Min	Max	WHO Standards	BIS Standards
EC	50	111	4870	-	-
TDS	50	71	3116.8	500	500
pH	50	6.1	9	6.5-8.5	6.5-8.5
Na	50	6.1	417.4	200	-
K	50	1	172	-	20
TH	50	30.3	676.7	300	200
Ca	50	11.9	142.9	75	75
Mg	50	0.4	125	50	30
HCO_3	50	48.8	1049.2	-	-
Cl	50	13.5	803.1	250	-
SO_4	50	0.4	100.2	500	200
NO_3	50	0.4	11.7	50	45
SiO_2	50	37.8	267.3	-	-

The concentrations of the parameters in groundwater in the study area are all within the WHO (WHO, 2004) allowable limit for drinking water. TDS higher than the limits in some points. In total hardness (TH) classification the water in the area is moderated.

4.2 Principal Components Interpretations

Principal Component I; has a high loading of TDS, Na, K, Cl, SO4, and TH which explain 54.866% of the total variance (Table 5a, 5b) This component reflex the Aquifer effect. The geological structure of the uddanam area is constituted cyclic deposition of chornkites, limestone and gypsum with sandstone at its upper part that represent the major sources (see Davis and Dewiest, 1966; Al-Rawi, *et al*., 1990 and Al-Tamir, 2005) [17, 16, 13]. Principal Components II has a high loading of DO and Sio_2 explain 12.53% of the total variance. Principal Component III has 9.241% of the total variance in shallow aqua fires and deep aqua fires represent the variation in natural atmospheric conditions, from temperature and rainfall that represent the seasonal effects upon water quality In summary, the two extracted PC representing four different processing are reflecting the effects of seasonal, agricultural drainage and storm water, geological and wastewater pollution from domestic.

4.3 Correlation

Pearson correlation (r) matrix was used to determine the relationship between variables. The classification was based on Guildford's rule (Guildford, 1973) [11] of thumb for interpreting the Pearson product moment correlation The correlation matrix (Table 2), describes the interrelationship between variables and the results for 10 hydro chemical parameters which show that the very high positive correlation (p) exist between Cl-(TDS, Na). High positive correlation (S) exist between TDS-(Na, k, HCO_3, SO_4), Na-SO_4 and Cl-HCO_3. A moderate correlation (M) exists between Na-(K, HCO_3), K-(HCO_3, SO_4, Cl), TH-(HCO_3, SO_4, Cl, TDS). The high positive loding indicated strong linear correlation between the component and parameter

Table 2: Guildford s rule of thumb for interpreting correlation coefficient.

R Value	Interpretation
0.0 to 0.29	Negligible or little correlation (N)
0.3 to 0.49	Low correlation (weak) (W)
0.5 to 0.69	Moderate correlation (M)
0.7 to 0.89	High correlation (Strong) (S)
0.9 to 1	Very high correlation (Perfect) (p)

Table 3: correlation matrix

Correlation		TDS	DO	Na	K	TH	HCO3	Cl	SO42
	TDS	1.000	-.099	.841	.758	.656	.723	.907	.854
	DO	-.099	1.000	-.336	.095	.005	-.235	-.175	-.139
	Na	.841	-.336	1.000	.566	.488	.757	.908	.810
	K	.758	.095	.566	1.000	.465	.585	.699	.599
	TH	.656	.005	.488	.465	1.000	.559	.634	.665
	HCO3	.723	-.235	.757	.585	.559	1.000	.739	.624
	Cl	.907	-.175	.908	.699	.634	.739	1.000	.840
	SO42	.854	-.139	.810	.599	.665	.624	.840	1.000
	NO3	-.412	.032	-.295	-.316	-.423	-.220	-.319	-.397
	SiO2	-.172	-.200	-.278	-.141	.083	-.169	-.193	-.238
Sig. (1-tailed)	TDS		.244	.000	.000	.000	.000	.000	.000
	DO	.244		.008	.253	.487	.049	.110	.165

Table 4: Communalities

	Initial	Extraction
TDS	1.000	.910
DO	1.000	.883
Na	1.000	.917
K	1.000	.650
TH	1.000	.732
HCO3	1.000	.709
Cl	1.000	.907
SO42	1.000	.813
NO3	1.000	.495
SiO2	1.000	.824

Table 5a: Explain initial components

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	5.487	54.866	54.866	5.487	54.866	54.866	5.452
2	1.253	12.530	67.396	1.253	12.530	67.396	1.224
3	1.101	11.011	78.407	1.101	11.011	78.407	1.165
4	.777	7.765	86.172				
5	.499	4.987	91.159				
6	.376	3.761	94.920				
7	.235	2.351	97.271				
8	.139	1.391	98.662				
9	.084	.844	99.506				
10	.049	.494	100.000				

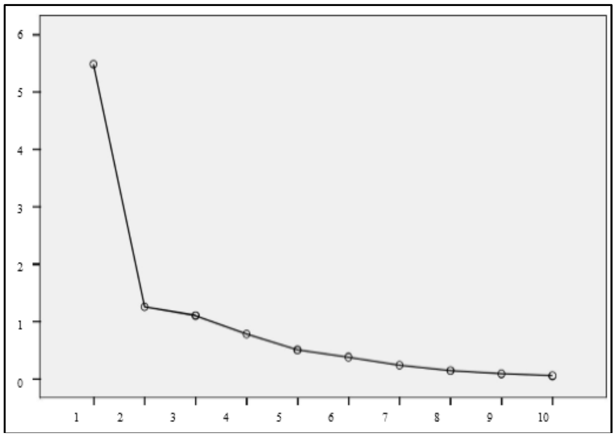

Fig 2: Scree plot for component with its Eigen values

Component Number

Table 5b: Component Matrix

Parameter	Component		
	1	2	3
TDS	.952	.043	.034
DO	-.173	.851	.358
Na	.902	-.118	-.300
K	.760	.238	.127
TH	.724	-.055	.452
HCO3	.816	-.114	-.172
Cl	.948	-.023	-.094
SO42	.901	.032	-.013
NO3	-.459	-.031	-.532
SiO2	-.218	-.662	.582

Table 5c: Rotated component matrix

		Component	
	1	2	3
TDS	.949	.032	.093
DO	-.146	.925	-.081
Na	.911	-.262	-.135
K	.762	.252	.071
TH	.688	.140	.489
HCO3	.818	-.198	-.031
Cl	.949	-.085	.009
SO42	.900	.002	.053
NO3	-.423	-.260	-.498
SiO2	-.295	-.319	.797

Table 6: Component Transformation Matrix

C.o.	1	2	3
1	.996	-.023	.087
2	.059	.890	-.451
3	-.067	.455	.888

5. Conclusion

From the 10 components in Table 5a, the first four components are sufficient to explain the monitoring area. These components explain 77% of the total variance of the original data set in study area. Moreover, the first four selected PC explained more than 54.8% of the variance of each quality variable; see communalities in Table 5a and 5b. PCA of water quality data for the Uddanam region ground water showed that seasonal effects, agricultural wastes and storm water effects and geological effects are the main variation in water quality. The results of PCA reflected a good look on the water quality monitoring and interpretation of the groundwater. Performing multivariate analysis for different seasons and years would help to assess temporal variation of groundwater quality. The results obtained through above mentioned tools will be helpful for the decision makers to adopt suitable remedial measures to protect the ground water sources.

6. References

1. Al-Kharabsheh, Ta'any R. Influence of urbanization on water quality deterioration during drought periods at south Jordan. Journal of Arid Environments. 2003; 53:619-630.
2. Praus P. Water Quality Assessment Using SVD-based Principal Component Analysis of Hydrological Data, Water SA, 2005; 31(4):417-422.
3. Jayakumar R, Siraz L. Factor analysis in hydrogeochemistry Of Coastal aquifers-preliminary study, Environmental Geology. 1997; (3/4)31:174-177.
4. Thilagavathi R, Chidambaram S, Prasanna MV, Thivya C, Singaraja CA. A study of groundwater geochemistry and water quality in layered aquifers system of Pondicherry region, southeast India. Appl. Water Sci. 2012; 2:253-269.
5. Olobaniyi SB, Owoyemi FB. Chemical Facies of groundwater in the Deltaic Plain Sands Aquifer of Warri, Western Niger Delta, Nigeria, AJST, 2006; 7(1):73-81.
6. Batayneh A, Zumlot T. Multivariate Statistical Approach to Geochemical Methods in Water Quality Factor Identification; Application to the Shallow Aquifer System of the Yarmouk Basin of North Jordan. Research Journal of Environmental and Earth Sciences. 2012; 4(7):756-768.
7. Shihab S. Application of Multivariate Method in the Interpretation of Water Quality Monitoring Data of Saddam Dam Reservoir Confidential, 1993.
8. Salameh, Water quality degradation in Jordan. Royal Society for the Conservation of Nature and Freidrich Ebert Stiftung, Amman, Jordan, 1996, 179.
9. Bayacioglu H. Surface water quality assessment using factor analysis, Water SA, 2006; 32(3):389-394.
10. Davis JC. Statistics and data analysis in geology, John Wiley and Sons, inc. New York, 1973, 550.
11. Guildford JP. Fundamental Statistics in Psychology and Education; McGraw-Hill: New York, NY, USA, 1973.
12. Nosrati K, Van Den Eeckhaut M. Assessment of Groundwater Quality using Multivariate Statistical Techniques in Hashtgerd Plain, Iran. J. Environmen-Interpretation of groundwater quality parameters for springs 43 tal Earth Science. 2012; 65(1):331-344.
13. Al-Tamir MA. Water quality Index for a group of wells in northwest of Mosul city. Journal of Sci. and Tech. Geology special issue. 2005; 16(2):27-40.
14. Mazlum N, Ozer A, Mazlum S. Interpretation of Water Quality Data by Principal Components Analysis, J. of Engineering and Environmental Science. 1999; 23:19-26.
15. Usman NU, Toriman ME, Juahir H, Abdullahi MG, Rabiu AA, Isiyaka H. Assessment of Groundwater Quality Using Multivariate Statistical Techniques in Terengganu, Science and Technology, 2014; 4(3):42-49.
16. Al-Rawi SM, Al-Azzo SI, Abbawi SA. Hydrogeochemical evaluation of groundwater in some parts of Mosul city and suitability for irrigation, SDRC conference, University of Mosul, Iraq: 1990; 18-20:224-238.

17. Davis SN, Dewiest RJM. Hydrogeology, John Wiley and sons, New York, 1966, 463.
18. Salman SR, Abu Ruka' h YH. Multivariate and principle component Statistical Analysis of contamination in urban and agricultural Soils from north Jordan. Environmental Soils from north Jordan, Environmental Geology. 1998; (3)38:265-270.
19. WHO. Guidelines for drinking water quality, 3rd edition. Geneva, Switzerland, WHO, 2004.
20. Lohani N, Todino G. Water Quality Index for Chao Pharya River. Journal of ASCE, Environmental Engineering, 1984; 110(6):1163-1176.
21. Iyer CS, Sindhu M, Kulkarni SG, Tambe SS, Kulkarni BD. Statistical analysis of the Physico-chemical data on the coastal water of Cochin, J. Environ. Monit. 2003; 5:324-327.
22. Department of Meteorology. Internal report. Amman-Jordan, 2004.
23. Kleinbaum G, Kupper LL, Muller KE. Applied Regression Analysis and Other Multivariable Methods, PWS-KENT Publishing Comp., 2nd ed., USA. 1988, 718.
24. Interpretation of groundwater quality parameters for springs in tafaileh area in south of Jordan using principal component analysis. Environmental Sciences. 2015; 3(1):31-44.
25. Indian meteorological department.
26. Cloutier V, Lefebvre R, Therrien R, Savard MM. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system, J. Hydrol. 2008; 353:294-313.
27. Satyanarayana G, Ramadasu P, Padmavathi Devi P, Prasad NVBSS, Nageswara Rao G. Ground water quality Assessment in Uddanamregion, Costal Srikakulam, Andhra Pradesh, India. IJPDA. 2017; 5(4):116-128.