Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

Vol. 9	No. 3	pp. 605-610	2010

Original Research Paper

Quality Assesment of Groundwaters Using Principal Component Analysis in Mianpur Area, Hyderabad, India

Ishrath, G. Prabhakar and M. Raghavender*
Deptt. of Geology, P. G. College of Science, Saifabad, Osmania University, Hyderabad-500 004, A. P., India
*Deptt. of Statistics, P. G. College of Science, Saifabad, Osmania University, Hyderabad-500 004

Nat. Env. Poll. Tech.
ISSN: 0972-6268
www.neptjournal.com

Key Words:
Principal component analysis,
Groundwater assessment
Mianpur urban area

ABSTRACT

Principal component analysis (PCA) of chemical variables of groundwater is used to interpret the relationship between the specific processes that control the quality of water. Groundwater samples were collected from Mianpur urban area, Hyderabad, and analysed for trace elements chemistry. The PCA separates the chemical variables into six principal components by HCO_3, Cl, NO_3, Al, Ba, Ca, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Ti and Zn. The principal component-I shows the significance on cadmium. The principal component-II shows the significance on calcium, chromium and iron. Tin, chromium, copper, manganese and lead show the significance on PC-III, PC-IV, PC-V and PC-VI, which are ignored as these show very less significance on the elements. These three principal components are identified with the processes of salinity, alkalinity and pollution, which are considered as lithologically and non-lithologically controlled factors. Identification of the zones helps to take appropriate management measures to improve quality of groundwater for sustainable development of the area.

INTRODUCTION

With the rapid growth in the industrialization and urbanization in Hyderabad, vast stretches of agricultural lands are converted into residential, industrial and commercial areas. A number of industries have been developed with intensive use of water and thus demand for water has been increased enormously, resulting into severe pressure on groundwater resources as surface water resources are limited here. On the other hand, wastewaters released from industries are left without adequate treatment. They flow mainly through the uncovered channels and finally let into the streams, tanks, ponds, and also mix with the natural waters. Understanding the impacts of land use activities on the quality of groundwater is very important to develop a better policy and legislative or regulatory mechanisms (Subba Rao et al. 2002). Due to the lack of such policies, there is a degradation of water quality, which causes an ill health. Therefore, it is essential to evaluate the relationship among the chemical variables and to identify the local and regional processes, which influence the quality of groundwater resources and management for a sustainable development. Such type of studies are of imperative relevance to the present social needs of the country.

STUDY AREA

The study area is a part of Ranga Reddy district, Andhra Pradesh located in one of the busiest highways (NH 9). It is very fastly developing urban area due to the hitech city which is 4-5 km away from the Mianpur area. It is bounded by 78°15’790” to 78°21’367’ E longitudes and 17°29’295” to

17°32'367" N latitude. The average annual rainfall is 800 mm. The area receives 578 mm south west monsoon and 132 mm in north east. It is located in the SOI Topo Sheet No. 56K/7NW.

MATERIALS AND METHODS

Principal component analysis is widely used for ranking of variables by using SPSS-6 software. In general PCA gives a large set of principal components which explain the relationship among the chemical variables. The first two or three principal component variables are taken on the basis of Eigen values, which explain the reasonable percent of variance. The first three principal components were utilized which have Eigen value greater than unity following Kaiser (1958). The interpretation of the principal components can be simplified using the rotating procedures. In the present study, the varimax rotation of Kaiser (1958) was applied to obtain a simple structure with scores, which shows the intensity of chemical processes described by each principal component (Dalton & Upchurch 1978). Regionally distributed and lithologically controlled variables are extracted first from the principal component scores (Table 5) and the more local pathway/origin controlled variables are then identified (Lawerence & Upchurch 1982).

Principal component analysis (PCA) of the chemical variables of groundwater is employed to interpret the relationship with specific processes that control quality of the water (Subba Rao et al. 2007). Groundwater samples collected from the Mianpur area were analysed for major ion chemistry. The PCA separates the chemical variables into three principal components, which together account for 77.2% of total variance. The five principal components are identified with the process of salinity, alkalinity and pollution, which are considered as lithologically and non-lithologically controlled factors. Litho-aerial distribution of the principal component scores shows that the higher positive scores of PC-I, PC-II, PC-III, PC-IV , PC-V and PC-VI are randomly distributed, but do not coincide with one another, which suggests that the local modifications are due to the regional flow system of groundwater, and are the pathways of recharge zones via soils and anthropogenic activities operated independently, following the natural environmental hydrogeochemical evolution of groundwater. Identification of these zones helps to take appropriate management measures to improve the quality of groundwater for sustainable development of the area.

The percentage of variance can be estimated based on the twelve elements given in Table 1. The highest percentage is shown by Al which is 9.5 and Zn having the least percentage of 0.262. The Cr and Cd show the percentage of 4.1 to 4.9, the Cu and Fe identical percentage of 2.4, and Mn, Ni and Pb the percentages from 1.0 to 1.5 indicating the impact factor on the pollution. Since Al variation is more, it indicates that there is a high impact of Al on water quality whereas Zn having less variation, indicates low impact on pollution with the component but any have rest of other two giving impact on the pollution with respect to above variations.

Table 1: Principal component loadings.

Components	Total	% of variance	Commulative %
Al	1.425	9.502	67.27
Ba	1.130	7.534	74.771
Ca	1.013	6.756	82.527
Cd	0.747	4.978	86.505
Cr	0.617	4.113	90.618
Cu	0.389	2.493	93.212
Fe	0.361	2.408	95.619
Mn	0.219	1.462	97.082
Ni	0.200	1.334	98.416
Pb	0.153	1.017	99.433
Ti	0.046	0.306	99.738
Zn	0.039	0.262	100

Table 2 indicates the principal component factors based on the squared loadings. This method was discovered by Karl Pearson, which is used to yield less efficient estimates than those obtained from the principle of maximum likelihood. Factor I indicates 19 % of total variance, Factor II indicates 18 % of total variance, Factor III indicates 16 %

Table 2: Principal component factors based on the sum of the square loadings.

	Factor – I	Factor-II	Factor-III	Factor –IV	Factor – V	Factor –VI
Total	2.911	2.740	2.456	1.60	1.366	1.157
% of Variance	19.04	18.267	16.376	10.663	9.106	7.711
Cumulative %	19.404	37.671	54.047	64.710	73.816	81.527

of total variance, Factor IV indicates 10 % of total variance, Factor V indicates 9 % of total variance and Factor VI indicates 7 % of total variance.

The significance of *t*-test at level of significant value of *t* for the 2 tailed tests can be obtained by finding the significant value of the different elements (Table 3). These values show the effect of pollution on the groundwater quality. Al (0.116) is less significant, Ba (0.060) moderately significant, Ca (0 .093) highly significant, Cd (0.041) less significant, Cr (0.427) moderate significant, Cu (0.000) not significant, Fe (0.083) high significant, Mn (0.045) moderate significant, Ni (0.186) less significant, Pb (0.062) moderate significant, Ti (0.017) less significant, and Zn (0.145) moderately less significant.

Dispersion: Averages or the measure of the central tendency gives an idea of the concentration of the observations about the central part of the distribution. Literal meaning of dispersion is scatteredness. The study of dispersion is to have an idea about the homogeneity or heterogeneity of the distribution. If it is more homogenous then it is less dispersed than the all. If it is more heterogeneous then it is more scattered or dispersed. The dispersion can be studied by mean and standard deviation.

Mean deviation: If i = 1, 2 , 3 ……n is the frequency distribution, the mean deviation ranges from the average A.

Then mean is:

$$1/N\sum_i f_i\,|X_i - A|, \sum f_i = N, \; for\, i = 1,2,3,....n$$

Standard Deviation: Standard deviation is the positive square root of the arithmetic mean of the square of the deviations of the given values from the Mean. For the frequency distribution,

$$\sigma = \sqrt{1/N\sum fi\,|Xi - \bar{x}|^2} \text{ for I = 1, 2, 3, …n}$$

The standard deviation satisfies almost all the properties laid down for an ideal measure of dispersion except for the general nature of extracting the square root which is not readily comprehensible for a non mathematical person. It may also be pointed out that standard deviation gives greater weight to extreme values and as such has found favour with economists or businessmen who are more interested in the results. The standard deviation in statistical theory is the best and most powerful measure of dispersion.

The standard deviation value of chromium is more (Table 4), which shows that it has less impact on the quality of the groundwater, whereas the standard deviation value of cadmium is less which shows more impact on the quality of groundwater. This is due to the anthropogenic activities going on in the study area and is due to the industries which are discharging the waste dump on road sides which is showing more impact on water quality.

The screen plot (Fig. 1) shows that there is positive correlation in the study area between components and Eigen values. The Table 6 explains the significant correlation between samples and the

Table 3: The independent sample means by using t-test.

		Test for equality of means		
		t	df	sig. (2-tailed)
HCO_3-	Equal variances asumed	4.413	74	0.000
Cl^-	Equal variances asumed	2.362	74	0.021
NO_3^-	Equal variances asumed	17.10	74	0.091
Al (396.153 μg/L)	Equal variances asumed	1.592	74	0.116
Ba (233.527)	Equal variances asumed	1.914	74	0.060
Ca (317.933)	Equal variances asumed	1.700	74	0.093
Cd (228.802)	Equal variances asumed	2.082	74	0.041
Cr (267.716)	Equal variances asumed	0.799	74	0.427
Cu (327.398)	Equal variances asumed	4.169	74	0.000
Fe (238.24)	Equal variances asumed	1.759	74	0.083
Mn (257.610)	Equal variances asumed	2.040	74	0.045
Ni (231.604)	Equal variances asumed	1.334	74	0.186
Pb (220.353)	Equal variances asumed	1.898	74	0.062
Ti (334.940)	Equal variances asumed	2.434	74	0.017
Zn (206.00)	Equal variances asumed	1.474	74	0.145

Table 4: The Descriptive statistics of the samples.

	Mean	Std. Deviation	Analysis N
HCO_3^-	263.584	187.0072	38
Cl^-	1631.54	3087.120	38
NO_3^-	20.47	397.334	38
Al (396.153 μg/L)	73.838	120.3080	38
Ba (233.527)	119.4517	164.45840	38
Ca (317.933)	114.198	157.4509	38
Cd (228.802)	1.45897	3.116349	38
Cr (267.716)	182.2312	386.99312	38
Cu (327.398)	10.30642	7.087427	38
Fe (238.24)	198.268	533.7726	38
Mn (257.610)	213.8096	380.94284	38
Ni (231.604)	11.10239	25.162841	38
Pb (220.353)	25.6224	38.50345	38
Ti (334.940)	3.65384	4.121908	38
Zn (206.00)	220.54308	713.630963	38

principal components where: SPS-Strong Positive Significant, MPS-Moderately Positive Significant, WPS-Weak Positive Significant, SNS-Strong Negative Significant, MNS-Moderately Negative Significant, WNS-Weak Negative Significant and NS-Negative Significant.

RESULTS AND CONCLUSIONS

The interpretation of the principal components can be simplified using certain procedures. In the

Table 5: The rotated component matrix extracted from principal component analysis.

	Component					
	1	2	3	4	5	6
HCO_3^-	0.191	0.898	0.081	-0.002	0.071	-0.032
Cl^-	0.556	0.721	0.266	0.010	-0.052	-0.036
NO_3^-	0.177	0.917	0.009	0.013	-0.009	-0.013
Al (396.153 µg/L)	-0.054	-0.149	0.766	0.158	0.073	-0.062
Ba (233.527)	-0.084	-0.235	-0.181	0.868	-0.025	0.197
Ca (317.933)	0.223	0.208	0.822	-0.125	-0.012	0.015
Cd (228.802)	0.837	0.253	0.009	0.052	-0.69	-0.023
Cr (267.716)	0.058	0.192	0.890	-0.054	0.155	-0.025
Cu (327.398)	0.255	-0.249	0.010	-0.088	-0.630	-0.317
Fe (238.24)	0.167	-0.125	0.177	0.020	0.822	-0.191
Mn (257.610)	0.152	0.260	0.181	0.875	0.101	-0.070
Ni (231.604)	0.910	0.244	-0.064	0.076	-0.125	-0.052
Pb (220.353)	0.727	0.017	0.304	-0.067	0.390	-0.080
Ti (334.940)	0.534	0.376	0.357	-0.091	0.269	0.323
Zn (206.00)	-0.035	-0.072	-0.058	0.089	-0.029	0.925

Extraction method: Principal Component Analysis; Rotation method: Varimax with Kaiser normalization

present study the varimax rotation of Kaiser (1958) was applied to obtain simple structures with scores. The scores obtained are called principal components. In the present study six principal components were obtained, which explain the effects of the elements based on the study. Each element shows the strong positive significance in each component.

1. The samples collected from the Mianpur urban area show the significant correlation between the components and the elements.
2. The PCA of 15 chemical variables in groundwater samples are taken as a complementary solution to hydro-geochemical study, which led to the identification of six principal components having the Eigen value loadings greater than unity and these together account for 87 % of total variance.
3. Graphical representation of principal component loadings brings close relationship of variables of HCO_3, Cl, NO_3, Al, Ba, Ca, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Ti and Zn.
4. The PC-I is identified with salinity, PC-II with alkalinity and PC-III with pollution processes. The PC-IV, PC-V and PC-VI can ignorable as they have highest negative values. They are described as lithologically and non-lithologically controlled factors respectively.
5. The PCA analysis of PC-I, PC-II and PC-III helps to demarcate the samples of higher positive significance which suggests the modification of an aquifer

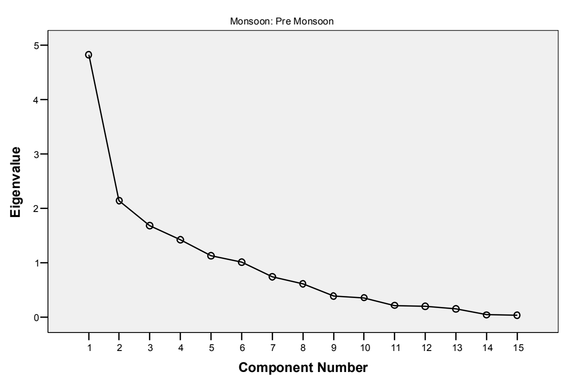

Fig. 1: Scren plot.

Table 6: The significant correlation between the samples and the six principal components.

Elements	Principal Components					
	PC-I	PC-II	PC-III	PC-IV	PC-V	PC-VI
HCO3	WPS	SPS	WPS	WNS	SPS	SNS
CL	SPS	WNS	WNS	MPS	MPS	MNS
NO3	MPS	SNS	WPS	SNS	MPS	SPS
Al	WNS	SPS	NS	WNS	SNS	WPS
Ba	WNS	SNS	WPS	WNS	MNS	SPS
Ca	WPS	SPS	WPS	WNS	MPS	SPS
Cd	SPS	WPS	WNS	WPS	SNS	SNS
Cr	WNS	SPS	SPS	SPS	MPS	WPS
Cu	WPS	WNS	SPS	WPS	MPS	WNS
Fe	WPS	SPS	MPS	WPS	MNS	MNS
Mn	WPS	WPS	SPS	MPS	WPS	WPS
Ni	WPS	SNS	MPS	WPS	WNS	WPS
Pb	WNS	WPS	SPS	WPS	WNS	WNS
Ti	WPS	SPS	MPS	SNS	SPS	WNS
Zn	SNS	WNS	WPS	WNS	SNS	SPS

chemistry occurring at different local levels and the entire area could be taken as a single mass for the regional processes.

6. Based on the results it is understood that PC-I is related to the regional flow system of groundwater, PC-II is related to the pathways of the recharge through soil zones, and PC-III is related to modified recharge through anthropogenic activity of agricultural and urban wastes.
7. The study suggests that the quality of groundwater is mainly controlled by salinity and pollution processes, and hence water softening and rainwater harvesting techniques should be adopted to reduce the salinity in waters, whereas the proper drainage facilities should be arranged to mitigate the pollution of the groundwater system.
8. The methodology can also be applied for the studies of groundwater quality of fast developing urban areas for a proper planning for groundwater resource management for sustainable development.

REFERENCES

Dalton, M.G. and Upchurch, S.B. 1978. Interpretation of hydrogeochemical facies by factor analysis. Ground water, 16: 228-233.

Kaiser, H.F. 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23: 187-200.

Lawrence and Upchurch, S.B. 1982. Identification of recharge areas using geochemical factor analysis. Groundwater, 20: 680-687.

Subba Rao, N., Prakasa Rao, J., John Devadas, D., Srinivasa Rao, K.V., Krishna, C. and Nagamalleshwara Rao, B. 2002. Hydrogeochemistry and groundwater quality in a developing urban environment of a semi arid region, Guntur, Andhra Pradesh. J. Geol. Soc., India, 59: 159-166.

Subba Rao, N., Prakasa Rao, J. and Subramanyam, A. 2007. Prinicpal component analysis in groundwater quality in a developing urban area of Andhra Pradesh. J. Geol. Society of India, 69: 959-969.