ISSN 2229-5518

Assessment of water quality in Hussainsagar lake and its inlet channels using multivariate statistical techniques

A. Sridhar Kumar*a, K. Shnakaraiahb, P.L.K.M.Raoc and M. Sathyanarayana

a Environmental specialist, Irrigation & CAD, Hyderabad, Telangana State, India
b Research scholar, earth & space sciences, University of Hyderabad
c Senior environmental consultant, Hyderabad
d Managing Director, ClinicaPro Data Research Pvt.Ltd, Hyderabad

*Author email: meetsreedhar1@gmail.com

----------------------------*********----------------------------

Abstract—Application of different multivariate statistical approaches for the interpretation of data obtained during a monitoring programme of Hussainsagar Lake and its catchment area at Hyderabad (India) is presented in this study. The waste water coming from the catchment is being directly discharged in to the lake, forming point and non point sources of contamination of lake water. This study analyzes twenty physico-chemical and heavy metal water quality parameters recorded at five sampling stations during 2012–2013 by using multivariate statistical techniques. Hierarchical clustering analysis (CA) is first applied to distinguish the three general water quality patterns among the stations, followed by the use of principle component analysis (PCA) and factor analysis (FA) to extract and recognize the major underlying factors contributing to the variations among the water quality measures. FA identified seven factors responsible for data structure explaining 73.05% total variance and allowed to group selected parameters according to common features. DO, COD (BOD) and TA were associated and controlled by mixed origin with similar contributions from anthropogenic and natural sources whereas Cd, Ni and Zn were derived from anthropogenic activities. This study indicates the necessity and usefulness of multivariate statistical techniques for evaluation and interpretation of the data with a view to get better information about the water quality and designs some remedial techniques to prevent the pollution caused by toxic elements in future.

Index terms—Heavy metals, Multivariate statistical techniques, Physico-chemical, Principle component analysis, Factor analysis, Inlet channels, lake water.

1 INTRODUCTION

Water quality is the main factor controlling the healthy and diseased states in both humans and animals. Surface water quality is an essential component of the natural environment and a matter of serious concern today. Anthropogenic influences as well as natural processes degrade the surface waters and their use for drinking, industrial, agricultural, recreation or other purposes [1], [2]. The water bodies are suffering because of pollution and are used for disposing of untreated local sewage and industrial effluents. People are becoming the more aware of the complexity of the nature and delicate balance that exist within the global ecosystem [3]. The discharge of effluents and associated toxic compound in to aquatic system represents an ongoing environmental problem due to their possible impact on communities in the receiving aquatic water and a potential effect on human health [4]. Further these materials enter the surface water resulting in pollution of irrigation and drinking water. Therefore, researchers have been paying more attention to the efforts of natural and human activities on water quality, in particular the key contributors of human activities to nutrients and heavy metals.

Many investigations have been conducted on anthropogenic contaminants of ecosystems [5], [6], [7]. Because of the spatial and temporal variation in water quality conditions, a monitoring program which provides a representative and reliable estimation of the quality of surface waters is necessary. The monitoring results produce a large and complicated data matrix that is difficult to interpret to draw meaningful conclusions. Multivariate statistical techniques are powerful tools for analyzing large numbers of samples collected in surveys, classifying assemblages and

ISSN 2229-5518

assessing human impacts on water quality and ecosystem conditions. The application of different multivariate statistical techniques, such as principal component analysis (PCA), factor analysis (FA), and cluster analysis (CA) assists in the interpretation of complex data for a better understanding of water quality and ecological characteristics of a study area. These techniques provide the identification of possible sources that affect water environmental systems and offer a valuable tool for reliable management of water resources as well as rapid solution for pollution issues [8], [9], [10]. Multivariate statistical techniques have been widely adopted to analyze and evaluate surface and freshwater water quality, and are useful to verify temporal and spatial variations caused by natural and anthropogenic factors linked to seasonality [11], [12].

The objective of the present study was to analyze 20 physico-chemical and heavy metal water quality parameters in water samples collected on monthly basis from 2012 to 2013 in Hussainsagar Lake basin in India. The data matrix obtained from field measurement was subjected to the CA, PCA, and FA techniques to define the natural and anthropogenic origin and to estimate the contributions of possible sources on concentrations of determined parameters.

2 MATERIALS AND METHODS

2.1. Study Site and Sample Collection

The Hussainsagar lake is one of the largest lake in India and a large (5.7 sq.km water spread area) and a deep (5.02 m. average depth) lake having 270 sq.km catchment area is situated on the deccan plateau at a height of 1788 feet above the sea level in the southeastern part of India, and is located at 17° 22’ of northern latitude and 78° 29’ of the eastern longitude. The mean maximum and minimum temperature vary from 40 to 14°C and the normal rainfall is 786 mm. Over a decade the lake receiving huge amount of waste water through four major inflow channels from the highly urbanized and industrialized catchment area, therefore, the author taken up the water quality study.
Fig.1. The sampling network including five measured stations was designed to cover a wide range of key locations accounting for inflows and outflow.

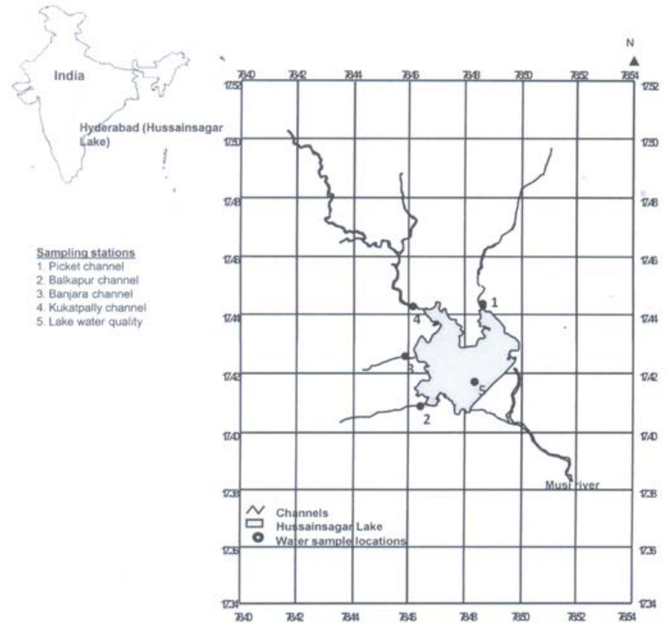

Fig. 1. Map of study area and water quality monitoring stations (listed 1-5) in the Hussainsagar lake basin.

The stations (C1, C2 & C3) namely Picket, Balkapur and Banjara channels are located at north, south and southwest directions of lake and receiving about 30, 17 and 15 million liters per day (MLD) of untreated wastewater respectively. Station (C4) Kukatpally channel located at northwest direction of the lake having four industrial estates and receiving 60 to 70 MLD of domestic & industrial effluents, and all these four stations are inflow sites. Station (L5) is close to the site of lake water outflow.

The selected water quality parameters are presented in table.1, includes pH, temperature, total alkalinity, dissolved oxygen, biochemical oxygen demand, chemical oxygen demand, barium, total nitrogen, total kjeldhal nitrogen, cadmium, copper, lead, nickel, total phosphate, total chromium, total oil & grease, total suspend solids, turbidity, arsenic and zinc. The depth of the sample is subsurface 0.5 m below the water surface.

ISSN 2229-5518

Table. 1. The water quality parameters, their units and methods of analysis.

Parameters	Abbreviations	Units	Analytical methods
pH	-	-	pH -meter
Temperature	WT	°C	Mercury thermometer
Total Alkalinity	T.Alk	(mg/L)	Colourimetric Titration
Dissolved Oxygen	DO	(mg/L)	Oxygen meter
Biochemical Oxygen Demand	BOD	(mg/L)	Five days incubation at 20^0C
Chemical Oxygen Demand	COD	(mg/L)	$K_2Cr_2O_7$ digestion
Barium	Ba	(mg/L)	Atomic Emission Spectrometry - Flame Emission
Total Nitrogen	TN	(mg/L)	Colourimetry
Total Kjeldhal Nitrogen	TKN	(mg/L)	Colourimetry
Cadmium	Cd	(mg/L)	Atomic Absorption Spectrometry - Solvent Extraction
Copper	Cu	(mg/L)	Colourimetry
Lead	Pb	(mg/L)	Atomic Absorption Spectrometry - Solvent Extraction
Nickel	Ni	(mg/L)	Atomic Absorption Spectrometry - Solvent Extraction
Total Phosphate	TP	(mg/L)	Colourimetry
Total Chromium	T.Cr	(mg/L)	Colourimetry
Total Oil & Grease	TO&G	(mg/L)	Petroleum Ether Extraction
Total Suspend Solids	TSS	(mg/L)	Gravimetric method
Turbidity	Turb	NTU	Nephelometric
Arsenic	Ar	(mg/L)	Colourimetry
Zinc	Zn	(mg/L)	Atomic Absorption Spectrometry - Solvent Extraction

3 DATA TREATMENT AND MULTIVARIATE STATISTICAL METHODS

The water quality data sets were subjected to four multivariate techniques: Cluster analysis (CA), principle component analysis (PCA) factor analysis (FA) and discriminate analysis [13], [14], [15]. Descriptive statistics was applied to raw data, whereas PCA, FA and CA were applied to experimental data, standardized through z-scale transformation in order to avoid misclassification arising from the different orders of magnitude of both numerical values and variance of the parameters analyzed [16], [17].

3.1. Cluster Analysis

CA is an unsupervised pattern recognition method that divides a large amount of cases into smaller groups or clusters based on the characteristics of the process. The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between the analytical values form the samples [18]. In the present study, we used Euclidian distance average linkage method (within the group) of cluster analysis. The number of clusters was also decided by practicality of the results as there is ample information available on the study sites. The spatial variability of water quality in the lake was determined from hierarchical CA using the linkage distance [19], [20].

3.2. Factor Analysis / Principal Component Analysis

Factor analysis technique extracts the eigen values and eigen vectors from co-variance matrix of original variables. The principle components (PC) are the uncorrelated (orthogonal) variables obtain by multiplying original correlated variables with eigen vector, which is a list of coefficients (loading or weightings). Thus principal components are weighted linear combinations of original variables. PC provides information on the most meaningful parameters, which describe whole data set affording data reduction with minimum loss of original information [21], [22]. It is a powerful technique for pattern recognition that attempts to explain the variance of large set of inter-correlated variables and transforming in to a smaller set of independent (uncorrelated) variables (principle component). Factor analysis further reduce the contribution of less significant variables obtained from PCA and the new group of variables known as varifactors, are extracted through rotating the axis defined by PCA. A varifactor can include unobservable, hypothetical, latent variables, while a PC is a linear combination of observable water quality variables [23], [24]. PCA of the normalized variables was performed to extract significant PC's and to further reduce the contribution of variables with minor significance. These PC's were subjected to varimax rotation (raw) generating varifactors.

4 RESULTS AND DISCUSSIONS

From the descriptive analysis (table. 2), the average concentrations of pH vary from 7.3 to 7.9 showing neutral to basic and alkaline the water in all the stations. The DO was observed very low average value varies from 0.2 to 2.9 mg/L indicating fragile water quality, and the COD (BOD) average values are very high varies from 116 to 536 mg/L indicating organic pollution. The T.Alkalinity average values are varies between 314 to 559 mg/L, this high alkalinity water may neutralize the acids present in the water. It was observed that there are some high values of TKN, TN, TP, TO&G, Pb, Cd, Ni, T.Cr, Ar and Zn due to point and non-point sources, which may be attributed to the domestic waste water discharge and industrial activities.

Station wise, the parameters like T.Alk, Ba,TN and Cu & Ni are showing high average values in station one and two. The parameters like Temp, TO&G, Zn, and COD (BOD), TKN, Pb, TP, T.Cr, TSS, Turb and Ar are showing high average concentrations in the station three and four respectively, whereas in the station five pH, DO and Cd are showing high average concentrations. The presence of only heavy metal parameters were observed in station C2, whereas, stations C1, C3, C4 & C5 having both the physico-chemical and heavy metal parameters.

Table. 2.
Mean and S.D. of water quality parameters at different locations of the Hussainsagar basin during 2012-2013.

Parameter	Station-C1	Station-C2	Station-C3	Station-C4	Station-L5
pH	7.49±0.15	7.35±0.17	7.37±0.16	7.46±0.20	7.91±0.17
WT	28.18±2.32	28.25±3.12	32.92±6.44	29.08±2.12	26.56±3.03
T.Alk	559±42.17	400±58.96	405±52.93	486±80.22	314±60.67
DO	1.01±0.30	1.32±0.42	1.59±0.87	0.24±0.39	2.90±0.55
BOD	172±26.46	87±20.52	101±27.10	252±39.18	39±4.39
COD	483±58.54	247±58.97	282±76.62	536±58.06	116±10.56
Ba	1.88±1.60	0.95±1.34	1.33±1.50	1.65±1.34	0.31±0.29
TN	77±33.96	43±24.58	55±28.95	66±38.73	26±10.87
TKN	59±25.04	54±21.97	54±28.49	64±33.70	22±9.88
Ca	0.01±0.00	0.02±0.02	0.01±0.00	0.05±0.04	0.05±0.11
Cu	0.06±0.06	0.14±0.33	0.13±0.33	0.07±0.06	0.05±0.06
Pb	0.06±0.05	0.11±0.03	0.11±0.03	0.13±0.03	0.11±0.04
Ni	0.13±0.14	0.20±0.14	0.19±0.14	0.17±0.18	0.16±0.12
TP	7.33±1.65	7.04±3.99	7.24±3.27	7.66±5.02	2.89±1.06
T.Cr	0.01±0.01	0.00±0.01	0.00±0.01	0.02±0.02	0.01±0.01
TO&G	192±169.69	316±370.49	294±351.39	8±1.79	4±2.68
TSS	174±76.49	172±71.84	194±81.72	185±62.69	80±17.12
Turb	56.58±22.04	46.46±16.78	32.50±6.61	72.88±44.13	25.50±3.54
Ar	0.01±0.00	0.02±0.02	0.01±0.01	0.03±0.05	0.02±0.02
Zn	0.15±0.14	0.62±0.62	0.71±0.71	0.65±0.91	0.54±0.43

C1-Picket channel, C2-Balkpur channel, C3-Banjara channel, C4-Kukatpally channel, L5-Lake water.

Overall, the discharge of municipal sewage, industrial effluents, the storm water discharge containing diluted sewage and other impurities on the land surface from over 240 square kilometers area of watershed have resulted in dumping of high amounts of organic matter, nitrogen and phosphorous in to the water and indicating the increased eutrophication. This situation suggests a strong variability due to presence of anthropogenic sources from the catchment affecting the water quality.

4.1. Spatial Similarity with CA

Cluster analysis was used to detect the similarity groups between the sampling sites. It yielded a dendrogram (Fig. 2), grouping all 5 sampling sites of the basin into three statistically significant clusters. Cluster 1 formed by Balkapur (C2) and Banjara stations (C3) because of similar or low distances based on the water quality parameters average concentrations and corresponds to relatively moderately polluted sites within the group and these stations are had a secondary level sewerage treatment plant at upstream side of the lake and carrying less dry whether flows in to the lake. Cluster 2 formed by Picket (C1) and Kukatpally stations (C4) because of similar or dissimilar and showing moderate distances based on the water quality parameters average concentrations and correspond to highly polluted sites within the group.

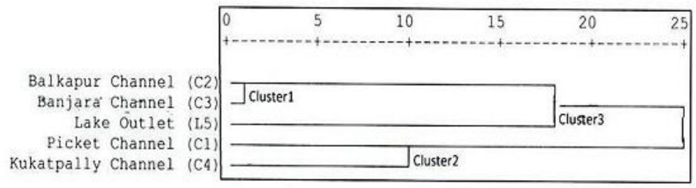

Fig. 2. Dendrogram showing clustering of sampling sites accordingly to water quality characteristics of the Hussainsagar basin

These stations receive pollution either sides of the channels from domestic and industrial areas. Cluster 3 Lake water (L5) combined with cluster I or cluster II and showing higher distances and showing less polluted site, this is due to self purification and assimilative capacity of the lake water. The results indicate that the clusters are showing similarities and dissimilarities between the stations. There are other reports [25], [26] with similar approach has successfully been applied to water quality programs.

4.2. Principal Component Analysis

Pattern recognition of correlations among 20 parameters was best summarized by PCA/FA. In this study, the covariance matrix coincided with the correlation matrix which was presented in table 3, because FA/PCA was applied to normalized data. Overall, the correlations between variables were relatively weak. Overall, the correlations between variables were relatively weak. There are some positive correlations between some variables such as DO, COD, BOD, TA, T.Cr, Ni, and so on. The negative correlations were revealed between some variables such as TKN, TN, TP, As, Cu and so on. Correlation coefficients of two elements were very useful, because they numerically represented the similarity between two elements of the two water quality variables. This also indicated that PCA could successfully reduce the dimensionality of the original data set. Therefore factor analysis of the

ISSN 2229-5518

present data set further reduced the contribution of less significant variables obtained from PCA.

Table. 3.
Correlation matrix of water quality parameters of Hussainsagar basin.

	pH	Temp	Turb	DO	COD	BOD	TSS	TA	TKN	TN	TP	TO&G	Pb	TC	As	Ba	Cd	Cu	Ni	Zn
pH	1																			
WT	-0.24**	1																		
Turb	-0.15*	0.04	1																	
DO	0.48**	-0.29**	-0.42**	1																
COD	-0.37**	0.14	0.48**	-0.92**	1															
BOD	-0.35**	0.13	0.45**	-0.92**	0.96**	1														
TSS	-0.29**	0.43**	0.24**	-0.35**	0.38**	0.33**	1													
TA	-0.28**	0.21*	0.41**	-0.63**	0.70**	0.65**	0.44**	1												
TKN	-0.21*	0	0.01	-0.35**	0.46**	0.40**	0.25**	0.42**	1											
TN	-0.06	0	0.06	-0.37**	0.49**	0.42**	0.29**	0.53**	0.81**	1										
TP	-0.25**	0.29**	0.15	-0.43**	0.40**	0.38**	0.28**	0.45**	0.46**	0.37**	1									
TO&G	-0.17*	0.36**	0.12	-0.14	0.09	-0.04	0.48**	0.20*	0.09	0.12	0.33**	1								
Pb	0.02	-0.02	-0.06	0.06	-0.15	-0.03	-0.07	-0.28**	-0.13	-0.24**	0.02	-0.25**	1							
T.Cr	0.04	-0.21**	-0.08	-0.15	0.20*	0.27**	-0.20*	0.01	-0.06	-0.03	-0.13	-0.32**	0.17*	1						
As	-0.09	0.01	0.03	-0.18*	0.19*	0.26**	-0.11	0.1	-0.05	-0.05	0.32**	-0.11	0.17*	0.23*	1					
Ba	-0.12	0.21*	0.14	-0.48**	0.45**	0.39**	0.27**	0.52**	0.41**	0.45**	0.37**	0.32**	-0.13	-0.03	0.03	1				
Cd	0.13	-0.02	0.02	-0.01	-0.01	0.02	-0.11	0.05	-0.03	-0.01	-0.02	-0.18*	0.13	0.14	0.08	0.03	1			
Cu	-0.11	-0.06	0.08	0.1	-0.03	-0.03	0.09	-0.08	0.17*	0.14	0.16*	0.11	0.07	0.02	0.07	-0.1	0.02	1		
Ni	-0.06	-0.06	-0.07	0.09	-0.16*	-0.12	0.01	-0.13*	0.01	-0.07	-0.07	-0.22**	0.12	-0.21*	-0.08	-0.31**	-0.01	-0.19*	1	
Zn	-0.1	0.03	-0.14	0.09	-0.09	-0.02	0.13	-0.16*	-0.13	-0.17*	0.20*	-0.03	0.21*	-0.01	0.68**	-0.16*	0.01	0.08	0.01	1

**Values are statistically significant at $p<0.01$
*Values are statistically significant at $p<0.05$

This also indicated that PCA could successfully reduce the dimensionality of the original data set. Therefore factor analysis of the present data set further reductions the contribution of highly influenced variables obtained from PCA. Elements belonging to a given factor were defined by factor matrix after varimax rotation, with those having strong correlations grouped in to factors. The identification of factors is based on dominant influence.

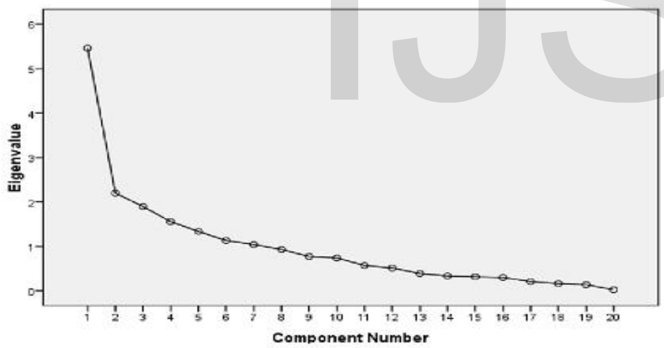

Fig-3. Scree plot of the characteristic roots (eigenvalues) of principal component analysis

The Scree plot (fig. 3) was applied to identify the number of PCs to be retained to understand the underlying data structure. Based on the Scree plot and the eigenvalues >1 criterion, seven factors were chosen as principal factors, explaining 73.05% of the total variance in the factor model. The corresponding VFs, variables loadings, eigenvalues, and explained variance are presented in table 4.

As per the Liu *et al*. [27] classification, the factor loadings classified as "strong", "moderate", and "weak", corresponding to absolute loading values of >0.75, 0.75–0.50, and 0.50–0.30, respectively. The first factor (VF1), explaining 27.28% of total variance, had strong loadings on DO, COD, BOD and TA. VF1, represented organic factor, influenced from point sources, such as domestic waste water. VF2, which explained 11% of total variance, had a moderate loading on To & G, T.Cr, and As represented industrial pollution source.

Table. 4. Loading of 20 parameters on significant VFs for water quality data set.

Parameters	Seven significant PCs						
	VF1	VF2	VF3	VF4	VF5	VF6	VF7
pH	-.453	.012	-.369	.156	-.251	.282	.151
Temp	.326	-.378	.468	-.217	-.066	.282	.171
Turb	.445	.082	.020	-.434	-.095	-.418	.264
DO	-.827	-.221	-.073	.281	-.047	-.024	.015
COD	.891	.282	-.093	-.152	.026	-.075	-.079
BOD	.827	.416	-.041	-.172	.059	-.076	-.048
TSS	.555	-.389	.314	-.068	.088	-.092	.154
TA	.837	.011	-.137	-.116	-.035	.085	.023
TKN	.613	-.093	-.318	.536	.318	-.002	.006
TN	.633	-.129	-.419	.488	.177	.041	.008
TP	.609	-.042	.342	.388	.057	.156	.036
Total Oil & Grease	.324	-.653	.274	.028	-.350	-.021	-.016
Pb	-.207	.402	.351	.091	.170	.020	.404
T.Cr	.018	.691	-.189	.017	-.266	-.049	-.124
Ar	.150	.570	.482	.236	-.120	.162	-.212
Ba	.642	-.142	-.157	.088	-.271	.374	.028
Cd	-.031	.317	-.094	.053	-.109	.284	.749
Cu	.047	-.029	.148	.527	-.184	-.694	.252
Ni	-.191	-.037	.056	-.147	.865	.053	.084
Zn	-.091	.241	.699	.305	.050	.098	-.166
Eigen value	5.456	2.201	1.894	1.554	1.334	1.132	1.040
Percentage of total variance	27.281	11.004	9.472	7.770	6.670	5.659	5.201
Cumulative percentage of variance	27.281	38.286	47.758	55.528	62.198	67.857	73.058

VF3, explaining 9.4% of total variance, had almost strong loading on Zn, and represented industrial pollution source. VF4, explaining 7.7% of total variance, had a moderate loading on TKN and represented the nitrate factor. VF5, explaining 6.6% of total variance and had strong loading on Ni and represented industrial pollution source, this may be due to intensive precipitation of heavy metals and acid oxides with in lake catchment area. VF6, explaining 5.6% of total variance, had a moderate loading on Cu and represented industrial pollution source. VF7, explaining 5.2% of total variance, had almost strong loading on Cd and represented industrial pollution source.

5. CONCLUSION

The water quality data have been examined by different multivariate statistical techniques were used to evaluate spatial and temporal variations of the lake basin. Euclidian distance average method analysis grouped 5 sampling stations into three clusters of mostly similar water quality characteristics and confirmed the existence of three types of water quality (moderately, highly and less polluted stations). Based on obtained information, it is possible to design a future, more optimal sampling strategy, which could reduce the number of sampling stations and associated costs. The PCA and FA assisted to extract and recognize the factors or origins responsible for water quality variations.

ISSN 2229-5518

PCA/FA identified seven latent factors that explained 73.06% of total variance, broadly organic matter, organic nitrogen and industrial pollution factors controlling their variability in waters of Hussainsagar basin.

The multivariate statistical approaches show that, the domestic sewage and uncontrolled industrial effluent discharges are highly influencing the water quality. Migration patterns of heavy metals released in to the environment in the form of untreated effluents by industries in Hussainsagar lake catchment indicate the point sources of pollution.

Thus, this study illustrates the usefulness of multivariate statistical techniques for analysis and interpretation of complex data sets, and in water quality assessment, identification of pollution factors and understanding spatial variations in water quality for effective lake water quality management.

6 RECOMMENDATIONS

The uncontrolled and untreated waste water from the catchment must be diverted through Interception and Diversion facilities. The effluents must be monitored for maintaining the standards prescribed by the pollution control board for various industries in the catchment area. The present study provides the baseline data for assessment of polluting sources and indicating the increased eutrophication.

ACKNOWLEDGEMENTS:

The authors are thankful to Major Shiv Kiran, President, M/s. Sukuki Exnora, Hyderabad, for his continuous support throughout this study.

REFERENCES:

[1]. Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications* 8: 559-568.

[2]. Fergusson, J. E., The heavy element chemistry, in: Fergusson, J. E., (Ed.), (1990). Environmental Impact and Health Effects, Pergamon, New York, p. 720.

[3]. Ahmet, D., Fevzi, Y., Tuna, A. L., and Nedium, O., (2006). Heavy metals in water, sediment and tissues of *Leuciscus cephalus* from a stream in southwestern Turkey, Chemosphere 63, p. 1451-1458.

[4]. Abbas Alkarkhi, F.M., Ismail, N., and Mat Easa, A., (2008). Assessment of arsenic and heavy metal contents in cockles (*Anadara granosa*) using multivariate statistical techniques. *Journal of Hazardous Materials* 150: 783-789.

[5]. Zeng X, Rasmussen T.C., (2005). Multivariate statistical characterization of water quality in Lake Lanier, Georgia, USA. J. Environ. Qual. 34:1980–1991.

[6]. Reghunath, R., Murthy, T. R. S., & Raghavan, B. R. (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India. Water Research, 36, 2437–2442. doi:10.1016/S0043-1354(01)00490-0.

[7]. Palma, P., Alvarenga, P., Palma, V.L., Fernandes, R.M., Soares, A.M.V.M., Barbosa IR., (2010). Assessment of anthropogenic sources of water pollution using multivariate statistical techniques: A case study of Alqueva's reservoir, Portugal. Environ. Monit. Assess. 165:539–552.

[8]. Boyacioglu, H., & Boyacioglu, H., (2007). Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey. Environmental Geology, 54, 275–282.

[9]. Li R, Dong M, Zhao Y, Zhang L, Cui Q, He W., (2007). Assessment of water quality and identification of pollution sources of plateau lakes in Yunnan (China) J. Environ. Qual. 36:291–297.

[10]. Kazi TG, Arain MB, Jamali MK, Jalbani N, Afridi HI, Sarfraz RA, Baig JA, Shah AQ., (2007). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. Ecotox. Environ. Safe. 72:301-309.

[11]. Singh, K.P., Malik, A., Sinha, S., (2005). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: A case study. Anal. Chim. Acta. 538:355–374.

[12]. Kim, J.H., Choi, CM., Kim, S.B., Kwun, S.K., (2009). Water quality monitoring and multivariate statistical analysis for rural streams in South Korea. Paddy Water Environ. 7:197–208.

[13]. Wunderlin, D.A., Diaz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C., Bistini, M.A., (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba-Argentiana), Water Research 35: 2881-2894.

[14]. Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Vousta, D., Anthemidis, A., Sofoniou, M., Kouimtzis, T., (2003). Assessment of the surface water quality in Northern Greece. Water Research 37: 4119-4124.

[15]. Singh, K. P., Malik, A., Mohan, D., Sinha, S., (2004). Multivariate statistical techniques for the evalution of spatial and temporal variations in water quality of Gomti River (India): a case study. Water Research 38: 3980-3992.

[16]. Liu, CW., Lin, KH., Kuo, Y.M., 2003. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Sci. Total Environ. 313:77–89.

ISSN 2229-5518

[17]. Simeonov V, Stratis JA, Samara C, Zachariadis G, Vousta D, Anthemidis A, Sofoniou M, Kouimtzis Th., (2003). Assessment of the surface water quality in Northern Greece. Water Res.2003; 37:4119–4124.

[18]. Otto, M., (1998). Multivariate methods. In: Kellner, R., Mermet, J.M., Otto, M., Widmer, H.M. (Eds), Analytical Chemistry. Wiley-VCH, Weinheim.

[19]. Simeonov V, Stratis JA, Samara C, Zachariadis G, Vousta D, Anthemidis A, Sofoniou M, Kouimtzis Th., (2003). Assessment of the surface water quality in Northern Greece. Water Res.2003; 37:4119–4124.

[20]. Zaharescu DG, Hooda PS, Soler AP, Fernandez J, Burghelea CI., (2009). Trace metals and their source in the catchment of the high altitude Lake Respomuso, Central Pyrenees. Sci. Total Environ. 407:3546–3553.

[21]. Vega, M., Pardo, R., Vega, M., Barrado, E., Deban, L., (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis, Water Research 32: 3581-3592.

[22]. Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M., Fernandez, I., (2000). Temporal evaluation of ground water composition in an alluvial aquifer (Pisuerga river, Spain) by principle component analysis, Water Research 34: 807-816.

[23]. Panda, U.C., Sundaray, S.K., Rath, P., Nayak, B.B., Bhatta, D., (2006). Application of factor and cluster analsis for characterization of river and esturine water systems-a case study: Mahanadhi River (India), Journal of Hydrology 331 (3-4): 434-445.

[24]. Davis, J.C., (1986). Statistics and Data Analysis in Geology, 2nd edn., Wiley, New York.

[25]. Kim, J.H., Choi, CM., Kim, S.B., Kwun, S.K., (2009). Water quality monitoring and multivariate statistical analysis for rural streams in South Korea. Paddy Water Environ. 7:197–208.

[26]. Shrestha, S., Kazama, F., (2007). Assesment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan, Environmental Modelling and Software 22 (4): 464-475.

[27]. Liu, CW., Lin, KH., Kuo, Y.M., 2003. Application of factor analysis in the assessment of groundwater quality in a lack foot disease area in Taiwan. Sci. Total Environ. 313:77–89.