International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:10, No:1, 2016

A Multivariate Statistical Approach for Water Quality Assessment of River Hindon, India

Nida Rizvi, Deeksha Katyal, Varun Joshi

Abstract**—**River Hindon is an important river catering the demand of highly populated rural and industrial cluster of western Uttar Pradesh, India. Water quality of river Hindon is deteriorating at an alarming rate due to various industrial, municipal and agricultural activities. The present study aimed at identifying the pollution sources and quantifying the degree to which these sources are responsible for the deteriorating water quality of the river. Various water quality parameters, like pH, temperature, electrical conductivity, total dissolved solids, total hardness, calcium, chloride, nitrate, sulphate, biological oxygen demand, chemical oxygen demand, and total alkalinity were assessed. Water quality data obtained from eight study sites for one year has been subjected to the two multivariate techniques, namely, principal component analysis and cluster analysis. Principal component analysis was applied with the aim to find out spatial variability and to identify the sources responsible for the water quality of the river. Three Varifactors were obtained after varimax rotation of initial principal components using principal component analysis. Cluster analysis was carried out to classify sampling stations of certain similarity, which grouped eight different sites into two clusters. The study reveals that the anthropogenic influence (municipal, industrial, waste water and agricultural runoff) was the major source of river water pollution. Thus, this study illustrates the utility of multivariate statistical techniques for analysis and elucidation of multifaceted data sets, recognition of pollution sources/factors and understanding temporal/spatial variations in water quality for effective river water quality management.

Keywords**—**Cluster analysis, multivariate statistical technique, river Hindon, water Quality.

I. Introduction

RIVER plays an important role in integrating and organizing landscapes and shaping the ecological settings of basins [1]. Rivers also plays a significant role in the assimilation and transportation of domestic and industrial waste waters, which form invariable pollution sources and agricultural runoff, which is temporal and commonly affected by climate [2]. The constant discharges of domestic and industrial wastewater and seasonal surface run-off due to the climate all have a strong effect on the river discharge and water quality. River water quality of a region is mainly determined by the cumulative effect of processes like precipitation, weathering, soil erosion, urban settlement around the bank, agricultural activities, disposal of human-waste, and domestic wastes with increasing exploitation/deterioration of water resources [3]. Rivers are highly prone to pollution, therefore it becomes necessary to keep check on surface water quality and interpret the temporal and spatial variations [4]. Regular monitoring of the quality of water is essential because clean water is obligatory for human health and the integrity of aquatic ecosystems [5].

A regular water quality monitoring program produces reliable data which reflect the state of the water quality of a river. However, producing good data is not sufficient to meet up the objectives of a water quality monitoring program. Data must be processed and presented in a way that provides the understanding of the spatial and temporal patterns in water quality parameters. The aim is to use the collected set of data to explain the current state of the water more widely and make the necessary controls to overcome future water quality issues. One of the problems with many sets of multivariate data generated from a monitoring program is that there are too many variables to analyze and then to draw meaningful conclusion from this [6]. In recent years, principal component analysis (PCA) and cluster analysis (CA) have been widely used in the interpretation of complex data sets to better evaluate the water quality and a variety of environmental issues, including inspecting the spatial and temporal patterns of water quality, chemical species associated with hydrological conditions, assessment pollution sources [7]-[12].

River Hindon has its origin from the lower Himalayas in Saharanpur district, Uttar Pradesh and flows 260 km through six districts, including Muzaffarnagar, Meerut, Baghpat, Ghaziabad and Gautambudh Nagar until its convergence with the Yamuna [13]. It is a major source of water to the highly populated and predominantly rural population of western Uttar Pradesh [13]. Water quality of Hindon is deteriorating at an alarming rate attributable to rapid urbanization, untreated industrial effluents into the river stream. The objective of present study is to find out the possible latent factors responsible for the water quality of river Hindon, and to assess information about the similarity and dissimilarities among the different monitoring stations by using multivariate techniques.

II. Material and Methods

A. Study Area

The River Hindon is one of the important rivers in western Uttar Pradesh (India) having a basin area of about 7000 km^2 and lies between latitude 28^0 30’ to 30^0 15’ N and longitude 77^0 20’ to 77^0 50’E. The river originates from Upper Siwalik (Lower Himalayas) and flows through five major districts, viz., Saharanpur, Muzaffarnagar, Meerut, Ghaziabad and

Deeksha Katyal is with the University School of Environment Management, Guru Gobind Singh Indraprastha University, Dwarka sec 16-C, New Delhi (phone-+91-9971430324; e-mail: deekshakatyal@gmail.com).
Nida Rizvi and Varun Joshi are with the University School of Environment Management, Guru Gobind Singh Indraprastha University, Dwarka sec 16-C, New Delhi. S. (e-mail: nrizvi8@gmail.com, varunj63@yahoo.com).

Greater Noida, a distance of about 200 km before joining the river Yamuna downstream of Delhi [14]. The major land use in the basin is agriculture, with little forest cover. The basin is densely populated because of the rapid industrialization and agricultural growth during the last few decades. The present study includes the stretch of River ranging from its entrance into Ghaziabad to its confluence with the Yamuna River in Tilwada village, Noida. Water samples were collected from eight different sites.

Study sites along the selected river stretch and their latitude and longitude are presented in Table I.

TABLE I
STUDY SITES ALONG WITH THEIR LATITUDE AND LONGITUDE

Site No.	Site Name	Latitude	Longitude
S1	Karehda	$28^0$40'32.7" N	$77^0$24'23.2"E
S2	Road Bridge	$28^0$40'24.3"N	$77^0$24'05.8"E
S3	Railway Bridge	$28^0$40'01.7"N	$77^0$23'58.6"E
S4	Chijarsi Ghaziabad	$28^0$38'11.0"N	$77^0$23'39.4"E
S5	Chijarsi Noida	$28^0$38'08.0"N	$77^0$23'39.2"E
S6	Kulsera	$28^0$31'47.9"N	$77^0$26'00.4"E
S7	Shafipur	$28^0$24'52.52"N	$77^0$29'20.65E
S8	Tilwada	$28^0$24'52.52"N	$77^0$29'50.98"E

B. Sampling and Analysis

Water samples were collected periodically from the selected study sites in the month of December 2013, May 2014 and August, 2014. Total 24 water samples were collected in the plastic bottles by following the procedure described in [15]. All samples were analyzed for 12 water quality parameters which include pH, temperature (Temp), electrical conductivity (EC), total dissolved solids (TDS), total hardness (TH), total alkalinity (TA), calcium (Ca^{2+}), chloride (Cl^-), nitrate (NO_3^-), sulphate (SO_4^{2-}), biological oxygen demand (BOD) and chemical oxygen demand (COD).

The temp of water was recorded on site with the use of thermo probe. pH and EC were also determined on the spot using pH meter and conductivity meter respectively. Calcium and total hardness was estimated by using EDTA titrimetry, alkalinity by titrimetry, chlorides by argentometry, TDS by gravimetry, nitrate and sulphate by spectrophotometry, BOD by incubation method and estimation of COD was done by reflux method.

C. Data Treatment

Physico chemical analysis data collected from eight study sites was normalized by log normal transformation to eliminate the effect of different measurement scales among individual parameters [16]. To test the suitability of data for principal component analysis, Kaiser- Mayer-Olkin (KMO) and Barlett's test of sphericity were performed. River water quality data have been subject to two multivariate techniques namely, PCA and cluster analysis. All statistical calculations were made with the use of SPSS 16.0 software.

D. Principal Component Analysis

PCA is principally used in data reduction and summarization; its objectives are to analyze the interrelationships between a large numbers of variables in terms of their common underlying dimensions, known as factors.

PCA of the normalized variables was performed to extract significant principal components (PCs) and to further reduce the contribution of variables with minor significance; these PCs were subjected to varimax rotation (raw) generating Varifactors (VFs) [17]-[21]. PCs were defined according to the norm that merely factors that account for variance greater than 1 (Eigen value-one criterion) ought to be incorporated. The underlying principle for this is that any component must account for more variance than any single variable in the standardized test score space. Hence, PCA was applied using varimax rotation with Kaiser Normalization. By extracting the Eigen values from the correlation matrix, the number of significant factors and the percent of variance explained by each of them were calculated.

E. Cluster Analysis

Hierarchical agglomerative cluster analysis was performed on the normalized data set by means of the Ward's method, using squared Euclidean distances as a measure of similarity.

The primary objective of cluster analysis is to group objects according to the characteristics they possess or on the basis of similarity. Resultant clusters possess a high level of homogeneity within clusters while heterogeneity between clusters [12].Hierarchical agglomerative clustering technique provides intuitive similarity relationships between any one sample and the entire data set, and is typically illustrated by a dendrogram (tree diagram) [22]. Euclidean distance usually gives the similarity between the two samples and can be represented by the difference between the analytical values from both the samples [23].

III. RESULTS AND DISCUSSION

Statistical summary of physico chemical parameters observed at various study stations of river Hindon is illustrated in Table II. The values of pH vary from 7.03 to 8.1 which show the slightly alkaline nature of river water. Alkaline nature of water could be due to the use of Soaps/detergents during baths of the pilgrims in the river. Temperature is one of the most important parameter for a river system. The variation in the temperature of surface water affects the solubility of salts, content of DO, organic biodegradation rate of materials and other physico-chemical parameters [24]. Temperature values in the present study range from 15.6^{o}C to 34.7^{o}C. Total hardness is caused due to cations of calcium, magnesium, iron and strontium. In present study TH values vary between 110 mg/l to 340 mg/l among various study sites. Values of calcium were recorded in the range of 31.26 mg/l to 123.4 mg/l at study sites. High value of calcium can create health issues as well as scaling in supply pipelines and boilers. Alkalinity in the study area was found to be in range of 200 mg/l to 630 mg/l. The high value of alkalinity indicates the sewerage mixing in the river. BOD gives the measure of amount of oxygen required by the microorganisms to degrade the organic matter. The values of BOD lie in the range of 11.4 mg/l to 82.76 mg/l in the river.

TABLE II
DESCRIPTIVE SUMMARY

Parameters	Minimum	Maximum	Mean	Std. Deviation
pH	7.03	8.10	7.49	.28426
Temp	15.60	34.70	26.69	7.41
TH	110.00	340.00	200.04	59.15
Ca^{2+}	31.26	123.44	53.21	18.54
TA	200.00	630.00	353.33	102.47
BOD	11.40	82.76	50.61	16.30
COD	16.00	192.00	99.67	39.26
NO_3^-	3.10	133.80	35.40	45.82
$SO4^{2-}$	18.50	179.04	65.79	39.84
Cl^-	39.76	309.66	156.08	82.89
TDS	243.00	1031.00	569.21	203.21
EC	295.00	1275.00	707.12	296.70

Higher value of BOD can be attributed to the mixing of waste water from various sources. The COD is the measure of oxygen equivalent to the organic matter content of the water susceptible to oxidation by a strong chemical oxidant and thus is the index of organic pollution in the river [25]. COD values range from 16 to 192 mg/l among various sites. Industrial waste water flow into the stream can be attributed to the higher COD values in the river system. Nitrate is a significant parameter of river water showing pollution status and anthropogenic load on river [13]. Nitrate and sulphate values in the stream range from 3.1 to 133.8 mg/l and 18.5 to 179.04 mg/l respectively. Chloride values were found to be in range of 39.8 to 309.7 mg/l.

TDS quantifies the solids dissolved in the water. This consists of salts, some organic materials ranging from nutrients to toxic materials. High TDS in water adversely affects the dissolved oxygen and increases the biological and chemical oxygen demand. TDS values in the present study range from minimum of 243 mg/l to maximum 1031 mg/l. High values of TDS indicate the mixing of sewerage, cloth washing and garbage dumping. The main sources of TDS in river water are natural sources, sewage, urban runoff, industrial waste water and chemicals used in water treatment process.

Electrical conductivity in water is due to ionization of dissolved inorganic solids and become a measure of TDS. It is the basic index to check the suitability of water for agricultural purposes [25]. EC values in this study vary from 295 to 1133µS/cm. Higher conductivity of water corresponds to mixing of sewerage in river water.

Correlation coefficient of physicochemical parameters of river Hindon is presented in Table III. The correlation matrix shows strong positive correlation between total hardness and chloride (r=0.82), similarly statistically significant correlation has been found between total alkalinity and sulphate (r=0.80), a similar relationship is reported by [26]. Moderate correlation was found between sulphate and TDS (r=0.67), chloride and TDS (r=0.66) and electrical conductivity and TDS (r=0.60).

TABLE III
CORRELATION MATRIX

Parameters	pH	Temp	TH	Ca	BOD	COD	NO_3^-	TA	SO_4^{2-}	Cl^-	TDS	EC
pH	1											
Temp	**-0.57**	1										
TH	0.16	-0.29	1									
Ca	0.20	-0.16	**0.54**	1								
BOD	0.13	-0.26	0.05	0.11	1							
COD	-0.17	-0.18	0.13	0.38	0.31	1						
NO_3^-	**-0.67**	0.19	0.18	-0.21	0.04	-0.00	1					
TA	0.41	0.06	0.40	0.44	-0.07	0.02	**-0.57**	1				
SO_4^{2-}	**0.52**	-0.24	0.33	0.30	-0.12	-0.07	-0.65	**0.80**	1			
Cl^-	0.15	0.02	**0.82**	**0.50**	0.18	0.10	0.16	0.49	0.27	1		
TDS	0.16	0.24	**0.55**	0.29	-0.11	-0.09	-0.30	0.81	**0.67**	**0.66**	1	

Before running PCA, all data set was normalized by using log normal transformation and Kolmogorove Smirnov (K-S) statistics were used to test the goodness of- fit of the data to log-normal distribution. The K-S test shows that all variable follows the log normal distribution.

To test the suitability of data for principal component analysis, Kaiser- Mayer-Olkin (KMO) and Barlett's test of sphericity were performed. KMO measures the sampling adequacy which points out the fraction of variables that are having common variance. KMO Values greater than 0.5 are considered as satisfactory for PCA. In the present study, this value is 0.577 (Table IV) which indicates that the data set is fit for principal component analysis. Barlett's test of sphericity test whether variables are significantly related and correlation matrix is an identity matrix. Significance level which is 0.00 for this study clearly indicates that the correlation matrix is not identity matrix and variables are significantly related.

TABLE IV
KMO AND BARLETT'S TEST

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.577
Bartlett's Test of Sphericity	Approx. Chi-Square	210.652
	df	66
	Sig.	.000

Principal component analysis was applied to the data obtained from eight study sites by using SPSS 16.0 software. PCA results in correlation matrix and factors were extracted by centroid method rotated by varimax rotation [27]. Eigen value indicates the significance of the factor, Eigen value more than one is considered as significant. PCA result shows that the first five Eigen values are higher than one and hence considered as significant.

Through scree plot (Fig. 1) it can be interpreted that greater part of variance in the original data is explained by first five factors.

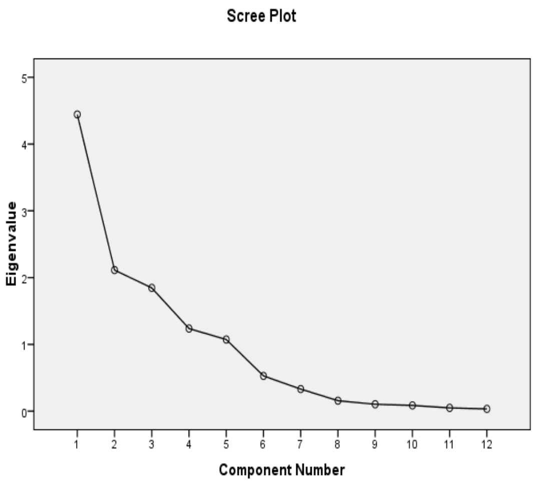

Fig. 1 Scree plot of the Eigen value

PCA yielded the five principal components which accounts for 89.3% of the total variance associated with the 12 parameters. Table V represents the total variance explained by the first five components for both rotated and unrotated factor loadings.

The first component (PC1) explains about 27% of total variance and have strong positive loading of alkalinity and sulphate while negative loading of nitrate. Second component (PC2) is responsible for 23.9% of total variance had strong positive loading of TH, Cl^- and TDS. This factor represents solid and minerals group and parameters under this group are reactive components of fractional anthropogenic origin [28]

TABLE V
EXTRACTED VALUES OF VARIOUS PCA PARAMETERS

PCs	Extraction sum of squared loadings			Rotation sum of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	4.44	37.03	37.03	3.27	27.26	27.26
2	2.11	17.61	54.64	2.87	23.90	51.15
3	1.85	15.38	70.02	1.72	14.32	65.47
4	1.26	10.31	80.33	1.46	12.18	77.65
5	1.07	8.94	89.27	1.40	11.62	89.27

Total hardness reveals influence of soil leaching and erosion. Dissolution of limestone and gypsum soils in the river catchments are major sources of minerals in the water. This factor perpetually indicated that the TDS in the river was mostly contributed by the TH [29]. PC3 had strong loading of pH and temp and accounts for about 14% of variance in the water samples.

Fourth component (PC4) explained 12.18% of total variance within the data set. It has strong positive loading of BOD and EC. BOD is indicative of organic pollution load on the stream which can be attributed to the anthropogenic sources from surrounding areas. PC5 explain about 11.6% of variance and have strong positive loading of Ca^{2+} and COD. COD also quantifies for organic pollution in the river. Table VI represents the factor loadings of various parameters.

TABLE VI
FACTOR LOADING MATRIX

Variables	Principal Components (PCs)				
	1	2	3	4	5
pH	.609		-.650		
temp			.959		
TH		.933			
Calcium					.621
BOD				.874	
COD					.919
nitrate	-.894				
alkalinity	.836				
sulphate	.851				
chloride		.927			
TDS	.610	.639			
EC				.723	

Hierarchical cluster analysis was performed to bunch the sampling sites of certain similarity. The result yielded a dendogram (Fig. 2) which categorized eight sampling sites in to two clusters on the basis of certain similarity between water quality characteristics. Cluster 1 consists of S4, S5, S6, S7 and S8 while cluster 2 consists of all upstream sites that are; S1, S2 and S3. Cluster 1 and cluster 2 corresponded to relatively high and low pollution region respectively. The water quality differences among these sites can be attributed to the difference in morphology and anthropogenic pollution. Site 1 2 and 3 are located near Karehda village which is the entrance point of river Hindon in the Ghaziabad city. Site 4, 5, 6, 7 and 8 are all downstream sites and water quality is deteriorated due to municipal discharge, urban, wastewater and industrial flow from the city which clearly shows that the anthropogenic activities of the Ghaziabad city is imposing immense stress on the river water quality.

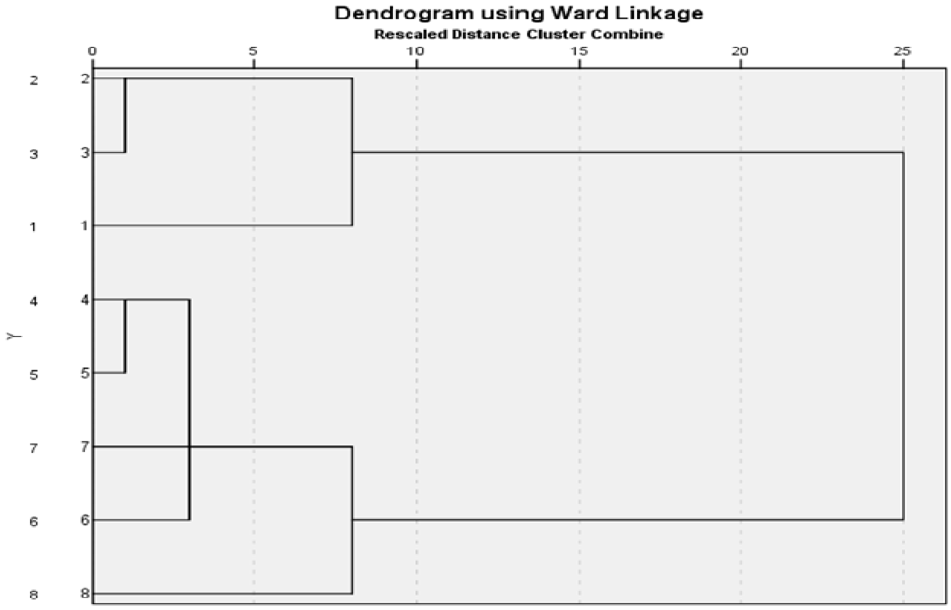

Fig. 2 Dendogram showing the relationships among the sites in the Hindon River

IV. Conclusion

Multivariate statistical technique such as PCA and cluster analysis was successfully applied to the 12 water quality parameters of eight sampling sites over the period of one year

Results obtained can facilitate in scheming optimal sampling frequency and strategy which could reduce the sampling stations. This study illustrated the usefulness of multivariate statistical techniques in water quality assessment and identification of pollution sources. Therefore, priority should be given to minimization of these sources to improve water quality in the river. This information will be helpful to the local authorities for the pollution control and management of Hindon River.

References

[1] V. Kumar, S. Arya, A. Dhaka, Minakshi, Chanchal, “A study on physio-chemical characteristics of Yamuna river around Hamirpur (UP), Bundelkhand region central India,” *International Multidisciplinary Research Journal*, Vol. 1, no.5, pp.14-16, 2011.

[2] F. Zhou, Y. Liu and H. Guo, “Application of multivariate statistical methods to water quality assessment of the watercourses in northwestern New Territories, Hong Kong,” *Environmental Monitoring and Assessment*, vol. 132, pp. 1-13, 2007.

[3] S. R. Carpenter, N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith, “Nonpoint pollution of surface waters with phosphorus and nitrogen,” *Ecol Appl*, vol. 8, no. 3, pp. 559–568, 1998.

[4] S. Yerel and H. Ankara, “Application of Multivariate Statistical Techniques in the Assessment of Water Quality in Sakarya River, Turkey,” *Journal Geological Society of India*, vol. 79, pp. 89-93, 2012.

[5] I. S. Babiker, A. A. Mohamed and T. Himaya, “Assessing groundwater quality using GIS,” Water Resource Management, vol. 21, pp. 699–715, 2007.

[6] U. Kuruppu, A. Rahman, M. Haque and A. Sathasivan, “Water quality investigation in the Hawkesbury-Nepean River in Sydney using Principal Component Analysis,” in *Proc* 20th *International Congress on Modelling and Simulation*, Adelaide, Australia,2013,pp. 2646- 2652.

[7] R. Bouza-Deaño, M. Ternero-Rodríguez and A. J. Fernández-Espinosa, “Trend study and assessment of surface water quality in the Ebro River (Spain),” *J Hydrol*, vol. 361, 227-239, 2008.

[8] O.O. Omo-Irabor , S. B. Olobaniyi, K. Oduyemi and, J. Akunna, “Surface and ground water quality assessment using multivariate analytical methods: a case study of the Western Niger Delta, Nigeria,” *Phys Chem Earth*, vol. 33, pp. 663- 673, 2008.

[9] T. G. Kazi , M. B. Arain, M. K. Jamali, N. Jalbani, H. I. Afridi and R. A. Sarfraz, “Assessment of water quality of polluted lake using multivariate statistical techniques: A case study,” *Ecotox Environ Safe*, vol. 72, pp. 301-309, 2009.

[10] Y. Ouyang, “Evaluation of river water quality monitoring stations by principal component analysis,” *Water Res*, vol. 39, pp. 2621-2635, 2005.

[11] B. Parinet, A. Lhote and B. Legube, “Principal component analysis: an appropriate tool for water quality evaluation and management? application to a tropical lake system,” *Ecol Model*, vol. 178, pp. 295-311, 2004.

[12] S. Shrestha and F. Kazama, “Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan,” *Environ Modell Softw*, vol. 22, pp. 464-475, 2007.

[13] S. Suthar, J. Sharma, M. Chabukdhara and A. K. Nema, “Water quality assessment of river Hindon at Ghaziabad, India: impact of industrial and urban wastewater,” *Env. Mon and Assess*, vol. 165, pp. 103–112, 2010.

[14] C. K. Jain, D. C. Singhal and M. K. Sharma, “Metal Pollution Assessment of Sediment and Water in the River Hindon, India,” *Env. Mon and Assess*, vol. 105, pp. 193-207, 2005.

[15] APHA. Standard methods for the examination of water and waste water, 19th. Ed, American Public Health Association, American Water Works Association & Water Environment Federation, Washington, DC, 1998.

[16] A. Mustapha, A. Z. Aris, M. F. Ramli and H. Juahir, “Spatial-temporal variation of surface water quality in the downstream region of the Jakara River, north-western Nigeria: A statistical approach,” *Journal of Environmental Science and Health*, vol. 47, pp. 1551–1560, 2012.

[17] G. Brumelis, L. Lapina, O. Nikodemus and G. Tabors, “Use of an artificial model of monitoring data to aid interpretation of principal component analysis,” *Env. Mod. and Soft*, vol. 15, no. 8, pp. 755-763, 2000.

[18] K. P. Singh, A. Malik, D. Mohan, and S. Sinha, “Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): a case study,” *Water Res*, vol. 38, pp. 3980-3992, 2004.

[19] K. P. Singh, A. Malik, and S. Sinha, “Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: a case study,” *Analytica Chimica Acta*, vol. 538, pp. 355-374, 2005.

[20] D. Love, D. Hallbauer, A. Amos, and R. Hranova, “Factor analysis as a tool in groundwater quality management: two southern African case

studies," *Physics and Chemistry of the Earth*, vol. 29, pp. 1135-1143, 2004.

[21] S. A. Abdul-Wahab, C. S. Bakheit and S. M. Al-Alawi, "Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations" *Env. Model and Soft*, vol. 20 no. 10, pp. 1263-1271, 2005.

[22] J. E. Jr. McKenna, "An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis," *Env. Model and Soft*, vol. 18, no. 3, pp. 205-220, 2003.

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

[24] G. S. Rao and G. N. Rao, "Study of ground water quality in greater Viskhapatnun city, Andhra Pradesh (India)," *J. of Env. Sc. & Engg*. Vol. 52, pp. 137-146, 2010

[25] S. Gholami and S. Srikantaswamy, "Analysis of agricultural impact on the Cauvery river water around KRS Dam," *World applied Sciences Journal*, vol. 6, no. 8, pp. 1157-1169, 2009.

[26] A. O. Adebola, S. M. Adekolurejo and O. Osibanjo, "Water Quality Assessment of River Ogun Using Multivariate Statistical Techniques," *Journal of Environmental Protection*, vol. 4, pp. 466-479, 2013.

[27] S. Ahmed, M. Hussain and W. Abderrahman, "Using multivariate factor analysis to assess surface/logged water quality and source of contamination at a large irrigation project at Al-Fadhli, Eastern Province, Saudi Arabia," Bulletin of Engineering Geology and the Environment, vol. 64, pp. 232–315, 2005.

[28] S. M. Yidana, "Groundwater classification using multivariate statistical methods: Southern Ghana," *J. Afr. Earth Sci*, vol. 57, pp. 455–469 2010.

[29] A. H. Ismail, S. A. Basim and A. Q. Shahla, "Application of Multivariate Statistical Techniques in the surface water quality Assessment of Tigris River at Baghdad stretch," *Iraq Journal of Babylon University/Engineering Sciences*, vol. 22, no. 2, 2014.