Assessment of water quality on the Yamuna river using principle component analysis: A case study

Bhuri Singh[1], Shahla Khan[2] & Shahjahan[3]
[1]Research Scholar, [2]Associate Professor, [3]Professor
[1,2,3] Department of Chemistry, Al-Falah University, Dhauj, Faridabad, Haryana-121004, India.

Received: September 02, 2018 Accepted: October 25, 2018

ABSTRACT *Present work studies of water quality using PCA (Principal Component Analysis) technique for surface water. Yamuna river is polluted river in North India and highly polluted at Delhi, Gautam Buddh Nagar, Faridabad, Mathura and Agra. Its water is increasingly deteriorated due to anthropogenic pollution from rapid economic development. Domestic and industrial effluents are directly discharge into water. Which has directly affected the water quality of Yamuna river. Fifteen water quality parameters have been analyzed in water sample for three years. A PCA technique has been applied in analysis dataset. Four factors were identified for the data structure explaining 83.73% of the total variance of the dataset, in which nutrient factor (41.79%), domestic sewage contamination (22.41%), physicochemical variability (11.92%) and waste water pollution from industrial (7.61%) in total variance of river water. Assessment of water quality concluded that PCA technique is useful tool for identification of hydrochemicals pollution. River water has really polluted. So, its needs to spread the aware for the pollution status of Yamuna river publicly.*

Keywords: *Industrialization, Principle Component Analysis, Water pollution, Water quality, Yamuna river.*

1.0 Introduction

Water is most valuable resources on earth for all of us and is required for biotic development. Surface water found in rivers, lakes and reservoirs. Rivers play an important role in economic and environment development [1].Yamuna river is polluted river in North India and highly polluted at Delhi region [2].Yamuna river is second largest tributary of Ganga river. Its water flows from Banderpooch (Himalayas), through Yamuna Nagar, Sonipat, Delhi, Gautam Buddh Nagar, Faridabad, Palwal, Aligarh, Mathura, Agra, Etawah and Allahabad [3]. Yamuna river water is used as resources for domestic, agriculture and transport usage. In total, seventy percent water is used for irrigation purposes, twenty two percent is consuming for industrial purpose and remaining eight percent is used for different domestic activities. Some industries are discharging their waste into Yamuna river which includes pulp and paper, sugar, distilleries, textiles, leather, chemical, pharmaceuticals, oil refineries, thermal power plant etc. [4].These areas are mostly polluted and severely affected by industrialization and urbanization. Some point sources has indicated that Yamuna river have polluted at Gautam Buddh Nagar, Faridabad, Palwal, Aligarh, Mathura and Agra [5]. The reason is being urbanization, industrialization and use of fertilizers in agriculture. In recent years, many researchers used PCA technique for identification of hydrochemical quality in different rivers water.(Ganga river, Markanda river, Bharathapuzha river, Gestsi river, Perlis river and Pearl river) [6-12].Some researcher are used PCA techniques for ground and surface water quality [13-15].The large data set has been analyzed by PCA technique to evaluate information about yearly variation in hydrochemical quality of water [16].The object of the present study is to analysis water quality of Yamuna river using PCA technique.

2.0 Material and Method

2.1. Study Area

Yamuna river is polluted river in North India. Its river originated from Himalayas, flows through Gautam Buddh Nagar, Faridabad, Palwal, Aligarh Mathura, and Agra and, its water is highly polluted at these sites. There are many industries in the surrounding area. Which has directly affects the water quality of river. In the present analysis has been carried out for six sites which are chosen as sample station. These sample locations are Hathi Ghat near railway bridge, Agra (S_1) Gokul Barrage near bridge, Mathura (S_2), Yamuna river bank near Pipli village, Aligarh (S_3), Yamuna river bank near Gurwari village, Palwal (S_4) Chhainsa village near bridge, Faridabad (S_5) and Kalindi Kunj near flyover, Gautam Buddh Nagar (S_6).

2.2. Sampling Collection

Fifteen water quality parameters were monitored on yearly basis over the three years (2014, 2015 & 2016). These water quality parameters include, pH, temperature, electrical conductance (EC), dissolved oxygen

(DO), turbidity, total dissolved solid (TDS), salinity, chloride, acidity, alkalinity, hardness, nitrate, biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total coliform. The sampling containers were washed with distilled water and ringed by Yamuna river water. The temperature varies between 27 ^{0}C to 40 ^{0}C during the sample collection. Its waters were immediately transported to the laboratory. Water samples were stored at 4^{0}C in the refrigerator. The water quality parameters were analyzed by standard method for examination of water and wastewater analysis method. [17—19]. PCA technique was used for pollution sources identification. The data toolpak excel and statistical packages social sciences (SPSS.16) have been used to carry out the correlation and factor analysis.

3.0 Result and Discussion

Yamuna river has very polluted from Gautama Buddha Nagar to Agra (S_1, S_2, S_3, S_4, S_5 & S_6). There are many industries in the surrounded area, urban disposal exist and agricultural waste. Which has directly affected the river water quality. So, its needs to water pollution status for publicly .Water samples collected at six sites in three years and analyzed by standard method. PCA technique is used as a tool in data analysis.Correlation is the statistical tool used to study the strength of a relationship between two and more variables. Chloride is strongly positive correlation (0.972.) with salinity which indicates domestic sewage contamination. Acidity is strongly negative correlation (- 0.869) with pH which indicate physicochemical variability. High positive correlation (0.750 to 1.00) can be observed between chloride with salinity, total dissolved solid with EC, salinity with total dissolved solid, hardness with pH, chloride with total dissolved solid, dissolved oxygen with EC which is responsible for domestic sewage contamination and physicochemical variability in river water. High negative correlation (-0.750 to -1.00) can be observed between acidity with pH, physicochemical variability is responsible.

Table.01: Correlation matrix of water quality parameters.

	pH	Temp	EC	DO	Tur	TDS	Sal	Chl	Aci	Alk	Har	Nit	BOD	COD	TC
pH	1														
Temp	-0.337	1													
EC	0.375	-0.132	1												
DO	0.256	-0.240	0.785	1											
Tur	0.238	-0.135	-0.143	-0.161	1										
TDS	0.423	-0.137	0.888	0.699	-0.245	1									
Sal	0.453	-0.269	0.708	0.665	-0.109	0.870	1								
Chl	0.426	-0.262	0.671	0.653	-0.048	0.793	0.972	1							
Aci	-0.869	0.360	-0.011	0.054	-0.542	-0.109	-0.274	-0.265	1						
Alk	-0.172	0.119	0.578	0.546	-0.738	0.572	0.334	0.292	0.609	1					
Har	0.793	-0.278	0.634	0.533	0.113	0.642	0.571	0.540	-0.584	0.147	1				
Nit	0.257	-0.407	0.607	0.638	-0.287	0.656	0.739	0.667	-0.099	0.290	0.423	1			
BOD	-0.431	0.689	-0.326	-0.536	-0.179	-0.357	-0.472	-0.444	0.373	-0.076	-0.562	-0.346	1		
COD	-0.321	0.166	-0.189	-0.440	-0.215	0.067	0.078	0.038	0.174	-0.023	-0.315	-0.021	0.367	1	
TC	-0.347	0.371	-0.136	-0.244	-0.176	0.093	0.073	0.063	0.269	0.087	-0.453	-0.174	0.483	0.626	1

Table.02: Descriptive statistics of water quality parameters.

W.P	Min	Max	Mean	S.D
pH	7.08	7.98	7.551	0.321
Temperature	27	40	34.544	3.683
EC	1022	1776	1559.944	201.852
DO	1.8	4.8	3.527	0.788
Turbidity	12	26	18.556	4.804
TDS	918	1560	1267.833	153.363
Salinity	0.307	0.56	0.402	0.074
Chlorides	170	310	220.388	39.916
Acidity	7.76	42.85	21.893	11.820
Alkalinity	218	540	403.111	89.756
Hardness	226	372	295.5	38.763
Nitrate	18	40	27.222	7.589
BOD	20	56	37.833	12.301
COD	67	130	92.111	17.713
TC	15000	7800000	1974888.889	2112999.259

Factor analysis is related to principal component analysis. Its is designed with the aim to identify certain unobservable factor. Four factors were identified for the data structure explaining 83.73% of the total variance of the dataset, in which nutrient factor (41.79%), domestic sewage contamination (22.41%), physicochemical variability (11.92%) and waste water pollution from industrial (7.61%) in total variance of river water.

Table.03: Total variance explained for components.

Components	Total variance Explained		
	Eigenvalue	Variance (%)	Cumulative frequency (%)
1	6.269	41.793	41.793
2	3.362	22.412	64.205
3	1.788	11.921	76.126
4	1.142	7.612	83.737
5	.777	5.181	88.819
6	.545	3.635	92.553
7	.400	2.670	95.223
8	.319	2.125	97.348
9	.169	1.125	98.473
10	.094	.625	99.098
11	.076	.506	99.604
12	.037	.246	99.805
13	.020	.132	99.982
14	.002	.013	99.995
15	.001	.005	100.00

Table.04: Correlation between variable and factor.

W.P	PCA1	PCA2	PCA3	PCA4
pH	0.631	-0.562	0.211	0.297
Temperature	-0.448	0.404	0.123	0.708
EC	0.835	0.297	-0.074	0.250
DO	0.819	0.228	-0.338	0.039
Turbidity	-0.071	-0.739	0.216	0.096
TDS	0.864	0.366	0.212	0.139
Salinity	0.879	0.203	0.352	-0.093
Chlorides	0.841	0.170	0.341	-0.082
Acidity	-0.356	0.802	-0.413	-0.113
Alkalinity	0.381	0.775	-0.371	0.109
Hardness	0.809	-0.312	-0.001	0.312
Nitrate	0.749	0.228	0.035	0.-320
BOD	-0.661	0.375	0.250	0.376
COD	-0.263	0.418	0.703	-0.280
TC	-0.290	0.538	0.641	0.000

4.0 Conclusion

In present study, fifteen water qualityparameters have been analyzed by standard method at six sites for three years. A PCA technique has been used for pollution sources identification and explains dataset. Four factors explains, in which nutrient factors (41.79%), domestic sewage contamination (22.41%), physicochemical variability (11.92%) and waste water pollution from industrial (7.61%) that represents total variance in river water. Study of PCA technique concluded that River water quality has pollution load. PCA technique is reliably method for large data analysis of surface water.

5.0 Reference

1. Singh. K.P, Malik., A and Sinho., S.(2005). Water quality assessment and apportionment of pollution sources of Gomti river, India using multivariate statistical techniques: A case study.Anal Chem Acta, 538:355-374.
2. Sharma., Keshav.(2015). Pollution study of river Yamuna: The Delhi story. International Journal of Science and Research, 6(10):1718-1622.

3. Rout., Chadetrik.(2017). Assessment of water quality: A case study of river Yamuna. International Journal of Earth Science and Engineering, 10(2):398-403.
4. CPCB.(2006). Assessment and development of river basin series ADSORBS/41/2006-07.Cental Control Pollution Board, New Delhi.
5. Sharma.,Deepshikha and Kansal., Arun.(2011).Water quality analysis of river Yamuna using water quality index in the national capital territory ,India(2000-2009).Applied Water Science, 1:147-157.
6. Fan., Xiaoyum, Cui., Baoshan, Zhao., Hui, Zhang., Zhiming and Zhang., Honggang. (2010). Assessment of river water quality in Pearl river delta using multivariate statistical techniques. Procedia Environmental Science, 2: 1220-1234.
7. Kumar., P, Kaushal., R.K and Nigam., A.K. (2015). Assessment and management of Ganga river water quality using multivariate statistical techniques in India. Asian Journal of Water Environment and Pollution, 12(4): 61-69.
8. Mustapha., Adamu and Nabegu., A.B.(2011).Surface water pollution sources identification using principle component analysis and factor analysis in Getsi river ,KangoNigeria.Australian Journal of Basic and Applied Science, 5(12): 1507-1512.
9. Mishra., A.(2010).Assessment of water quality using principle component analysis: A case study of the river Ganga .Journal of Water Chemistry and Technology, 32(4):227-234.
10. Raj., Nikhal and Azeez., P.A.(2009).Spatial and temporal variation in surface water chemistry of a tropical river the river Bharathapuzha,India.Current Science, 96(2): 245-251.
11. Samsudin., M.S, Juahir., Hafizan, Zain., S.M and Adnan., N.H.(2011).Surface river water quality interpretation using environment techniques : case study at Perlis river basin, Malaysia. International Journal of Environment Protection, 1(5):1-8.
12. Verma., R.K, Verma., K.S and Bhardwaj., S.K.(2015).Pollution potential assessment of Markandar river around Kala-Amb industrial town of Himachal Pradesh, India. International Journal of Science and Nature, 6(4): 606-612.
13. Alhassan., H, Muntasir., A.H, and Reem., J.C. (2015).Groundwater quality assessment in urban area of Bagdad Iran using multivariate statistical techniques. Engineering and Technology Journal, 33(2):463-476.
14. Athimoolam., M and Velayutham., R.B. (2014).Hydrochemistry of groundwater in Chennai city using principal component and factor analysis: A case study. Research Journal of Chemical of Chemical and Environment Science, 2(2):78-88.
15. Kumar., D.S, Kumar., S.P and Gopalkrishan., P. (2011). Ground water quality assessment in paper mill effluent irrigated area using multivariate statistical analysis. World Applied Science Journal, 13(4): 829-836.
16. Dixon., W and Chiswell., B.(1996). Review of aquatic monitoring program design. Water Research, 30:1935-1948.
17. American Public Health Association. (1997). Standard method for the examination of waste water, 19th edition Washington, DC.
18. Maiti., S.K. (2011). Hand book method in environmental studies, Oxford Book Company, New Delhi.
19. Kaul., S.N and Gautam., Ashutosh. (2002). Water and waste water analysis, Daya Publishing House, Delhi.