

Application of Multivariate Statistical Techniques in Determining the Spatial Temporal Water Quality Variation of Ganga and Yamuna Rivers Present in Uttarakhand State, India

Madhuben Sharma¹ · Ankur Kansal² · Suresh Jain¹ · Prateek Sharma¹

Received: 24 April 2015/Revised: 14 July 2015/Accepted: 15 July 2015/Published online: 30 August 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Primary monitoring of 18 water quality parameters for rivers Ganga and Yamuna of Uttarakhand State was carried out to study the seasonal variation of these parameters, identify potential sources of pollution, and clustering of monitoring stations with similar characteristics. Wilcoxon signed-rank test, Paired *t* test and multivariate statistical techniques—principal component analysis (PCA) and cluster analysis (CA) were used to analyse the collected data. Separate analyses were conducted for summer and winter periods. The Wilcoxon signed-rank test and paired *t* test revealed seasonal variability in the data set with high pollution levels during summer period as compared to winter period. The CA grouped 15 monitoring stations of river Ganga and 5 monitoring stations of river Yamuna into 2 clusters of similar characteristics. The PCA resulted in the identification of four major sources of pollution for river Ganga, and three for river Yamuna. The findings of the study provide useful information in interpretation of complex datasets and for water quality assessment, identification of pollution sources/factors and understanding of temporal and spatial variations of water quality for effective river water quality management.

Keywords Water quality · *t* test · Wilcoxon signed-rank test · Principal component analysis and cluster analysis

Introduction

The surface water quality in a region is largely determined by two factors—the natural processes such as precipitation rate, soil erosion, vegetation etc. and the anthropogenic processes such as domestic and industrial effluents, agricultural runoff etc. (Singh et al. 2004). While the municipal and industrial wastewater discharge constitutes the constant polluting source, the surface runoff is affected by climatic factors. Seasonal variation in precipitation, surface runoff, groundwater flow, water interception and abstraction has a strong effect on river water quantity and quality (Zhao et al. 2011; Razmkhah et al. 2010; Kannel et al. 2008). Escalated anthropogenic activities in the basin and reduced river discharges registered during the last decades have caused multifold increase in the organic pollution load in the surface water bodies in India (Singh et al. 2004). Hence, for the long-term effective management of a river, an understanding of hydro-morphological, chemical and biological characteristic is required (Shrestha and Kazama 2007). Continuous water quality monitoring programmes are carried along different river stretches in order to study the seasonal variations in the physico-chemical and biological parameters and impact of different pollutant sources. Such programmes generate large amount of multiple parameter data distributed in space and time. Multivariate statistical techniques such as the principal component analysis (PCA) and cluster analysis (CA) are useful tools for extracting meaningful information from such large data set. These techniques have been applied successfully to analyse and interpret water quality data, study seasonal variations and identification of potential sources of river pollution (Bonansea et al. 2014; Zhao et al. 2011; Razmkhah et al. 2010; Li et al. 2009).

✉ Madhuben Sharma
madhubensharma06@gmail.com

¹ Department of Natural Resources, TERI University, 10, Institutional Area, Vasant Kunj, New Delhi 110070, India

² Uttarakhand Environment Protection and Pollution Control Board, Roorkee, India

Uttarakhand is a newly created state in India. New infrastructure facilities are being created which include housing, roads, industries etc. The state is also in the process of preparing zoning atlas for siting industry. A study of water quality for the two major rivers that flow through the state—Ganga and Yamuna—would help in providing a database for the zoning atlas in locating appropriate sites for industries, establishments of river rehabilitation plan and designing comprehensive water quality monitoring network to assess the variations of different water quality parameters. For this purpose, primary monitoring of 18 water quality parameters at 15 monitoring stations for river Ganga and 5 monitoring stations for river Yamuna was carried in the months of January and April, representing winter and summer months, respectively. This large multiple parameter data have been analysed statistically to study the (a) seasonal variation of these parameters; (b) clustering of monitoring stations with similar characteristics; and (c) identify potential sources of pollution. Specifically, Wilcoxon signed-rank test, paired *t* test, PCA and CA have been used to achieve the above three objectives, respectively.

Materials and Methods

Site Description, Data Collection and Analytical Methods

The state of Uttarakhand has diverse resources in terms of flora and fauna in the foothills of Himalayas. It lies between geo-coordinates from 30°19'48"N to 78°03'36"E. The geographical area is 53566 sq. km, out of which, about 85 % is hilly/forest/glacier area and the remaining 15 % area buzzes with commercial activities. Uttarakhand is divided into two geographical areas, namely Garhwal and Kumaon area. Garhwal area is nourished by rivers Ganga, Yamuna and Tons; while the rivers Kosi, Ramganga, Gola etc. flow through the Kumaon area. The water quality sampling was carried out for rivers Ganga and Yamuna. The sampling was done for 15 and 5 monitoring stations for river Ganga and Yamuna, respectively (Fig. 1). The Ganga originates from the western Himalayas in the state of Uttarakhand, and flows in the south-east direction through the Gangetic Plain of Northern India (about

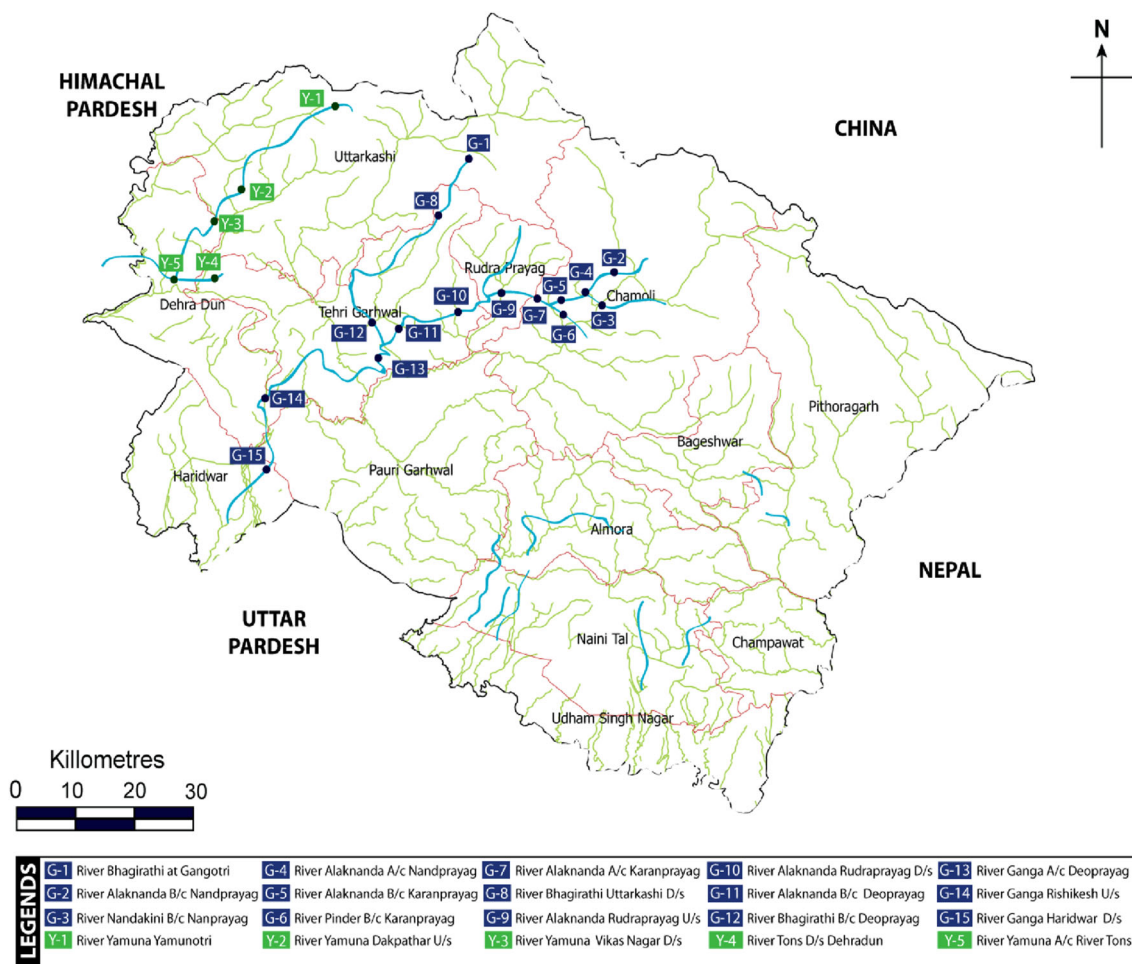


Fig. 1 Map showing study area and sampling stations of the rivers Ganga and Yamuna

2525 km), and finally enters the Bay of Bengal. The river Ganga actually arises due to the confluence of Bhagirathi and Alaknanda rivers at Devprayag. The Ganga receives domestic sewage and heavy effluent discharges from industries such as paper manufacturing, distillery, sugar manufacturing, electroplating and automobiles (Table 1). The river Yamuna on the other hand originates from the Yamunotri glacier at Banderpooch peak in the Uttarkashi district of Uttarakhand. The catchment of the river extends to states of Uttarakhand, Uttar Pradesh, Himachal Pradesh, Haryana, Rajasthan, Madhya Pradesh and the entire union territory of Delhi. The river flows a stretch of 1367 km till its confluence with Ganga at Allahabad. During its course, it receives heavy discharge and pollutants from small-scale industries which include paper, sugar, spices, chemical, leather, distillery, television manufacturing, pharmaceutical and power; some waste from the pesticide industry also enters just before the confluence of river Tons with Yamuna (Table 1).

In order to assess the seasonal variation in the water quality, parameters sampling was done in January 2012 and April 2012 and sampling station network was designed in such a way so as to cover wide range of factors such as domestic wastewater drains, industrial effluent, erosion of mountains and mucks from hydroelectricity projects that influence the water quality both rivers. Samples were collected in triplicates from three points (1/4, 1/2 and 3/2) across the river width at same sites in 2000 ml polyethylene sterilized bottles (cleaned with metal free soap, washed with distilled water, soaked in 10 % nitric acid for 24 h. and finally, rinsed with ultrapure water). Water samples were stored in an insulated, ice cooled container and delivered to the laboratory on the same day. All the samples were kept at 4 °C to avoid physico-chemical changes and interference. The sampling, preservation, transportation and analysis of water samples was carried out according to the methods described in APHA-AWWA-WEF (1998). The polyethylene bottles used for storing samples were sealed with tape to make the sample free from air contact.

A total of eighteen water quality parameters were measured at different monitoring locations. The pH, turbidity, dissolved oxygen (DO), electrical conductivity (EC), total dissolved solids (TDS) and temperature (Temp.) were measured onsite by portable pH meter, turbidity meter, DO meter, conductivity meter, TDS meter and thermometer respectively; whereas the other parameters viz. total hardness (TH), 5-day biochemical oxygen demand (BOD₅), total alkalinity (TA), total suspended solid (TSS), Chloride (Cl⁻), Sodium (Na⁺), Potassium (K), Phosphate (PO₄⁻³), lead (Pb), Copper (Cu), Iron (Fe) and Zinc (Zn) were analysed in laboratory using standard

protocol APHA-AWWA-WEF (1998). All the parameters along with the analytical methods for their measurement are summarized in Table 2.

Data Treatment and Multivariate Statistical Techniques

The raw data for all 18 water quality parameters monitored at various stations for rivers Ganga and Yamuna were subjected to statistical analysis. Exploratory data analysis was first carried out which included computation of summary statistics, assessment of strength of association between different parameters using correlation analysis. The data were further checked for normality by conducting goodness-of-fit tests; the tests revealed that most of the parameters in river Ganga do not follow the normal distribution. Thus, in order to assess seasonal variation in the water quality parameters, Wilcoxon signed-rank test (non-parametric) was carried out for river Ganga and paired t test was performed on normal distributed data of river Yamuna. The correlation analysis revealed that statistically significant association exists between the parameters, and therefore more elaborate PCA can be carried out to identify underlying interrelationship amongst the parameters. The CA was carried out for grouping similar sampling stations into common clusters. All the statistical analysis was carried out using statistical package SPSS, version 10 (SPSS 1999).

The PCA is a variable reduction technique, which is used for understanding the underlying relationships between the variables. The data reduction helps in extracting information about the most meaningful parameters that describe the entire data set without loss of original information. It allows interpretation that is not revealed by simple summary statistics and reveals principal tendencies present in the data matrix. The technique involves transformations of p correlated variable into p uncorrelated orthogonal principal components (PCs), which are weighed linear combinations of the original variables. The total number of PCs extracted from PCA indicates the total number of possible sources of variation in the data set. The successive linear combinations are extracted in a manner that they are orthogonal to each other, and account for successive small amount of the total variance that has not been accounted for by its predecessors. Thus, the first PC has the highest eigen value and represents the most important source of variation in the data. The last PC accounts for the least variation in the data set. The eigenvalues of the PCs are given by the loadings, and the individual transformed observations are called scores. The projections of the original variables on the subspace of PC result in loadings; these loadings are

Table 1 Sampling Locations and site characteristics of River Ganga and Yamuna

Sampling point	Location	Altitude (mt)	Main source of water pollution
River Ganga			
G-1	River Bhagirathi at Gangotri	3048	Origin of river Bhagirathi (Clean water—no source of pollution, no pesticide use, bio-fertilizers are used here)
G-2	River Alaknanda B/c Nandprayag	1358	Sewage and muck of hydro-electric projects from Chamoli (disposal of waste from hotels/restaurants, no STP's, stone crushers, raw sewage is thrown directly in the water bodies)
G-3	River Nandakani B/c Nandprayag	1358	Sewage and muck of hydro-electric projects from Nandprayag (disposal of waste from hotels/restaurants, no STP's, raw sewage is thrown directly in the water bodies)
G-4	River Alaknanda A/c Nandprayag	914	Sewage (disposal of waste from hotels/restaurants, no STP's, raw sewage is thrown directly in the water bodies)
G-5	River Alaknanda B/c Karanprayag	1450	Sewage from Karanprayag (disposal of waste from hotels/restaurants, no STP's, raw sewage is thrown directly in the water bodies)
G-6	River Pinder B/c Karanprayag	1450	Sewage from Karanprayag (disposal of waste from hotels/restaurants, no STP's, cottage industries, raw sewage is thrown directly in the water bodies)
G-7	River Alaknanda A/c Karanprayag	1450	Sewage (disposal of waste from hotels/restaurants, no STP's, raw sewage is thrown directly in the water bodies)
G-8	River Bhagirathi Uttarkashi D/s	1165	Sewage and muck of hydro-electric projects from Uttarkashi town (disposal of waste from hotels/restaurants, no STP's, raw sewage is thrown directly in the water bodies, presence of Lohri nagpal mega hydroelectricity project which was shut down by the Govt. after 75 % construction and continuous erosion of Varnavat mountain in this region)
G-9	River Alaknanda Rudraprayag U/s	895	Sewage and muck of hydro-electric projects (Lohri nagpal mega hydroelectricity project present but shut down by the Govt. after 75 % construction, stone crushers.)
G-10	River Alaknanda Rudraprayag D/s	895	Sewage and muck of hydro-electric projects and agricultural runoff
G-11	River Alaknanda B/c Deoprayag	472	Sewage and muck of hydro-electric projects, pilgrims visit here, between points G11 to G13, there are 3 NTP'S hydro-electric projects and effect is visible only on G13, cottage industries and stone crushers
G-12	River Bhagirathi B/c Deoprayag	472	Sewage and muck of hydro-electric projects, pilgrims visit here, between points G11 to G13, there are 3 NTP'S hydro-electric projects and effect is visible only on G13, cottage industries and stone crushers
G-13	River Ganga A/c Deoprayag	472	Sewage and muck of hydro-electric projects, pilgrims visit here, between points G11 to G13, there are 3 NTP'S hydro-electric projects, tourist influx is very high, G1–G13 highly degradable sewage with high BOD, Impact of Tehri dam and devotee offerings)
G-14	River Ganga Rishikesh U/s	532	Sewage and industrial sources of water pollution (ACE glass industry, closed IDPL pharmaceutical industry, sewage lines from different towns)
G-15	River Ganga Haridwar D/s	250	Sewage and industrial sources of water pollution (electroplating, automobile, pharmaceutical, phosphating industries, sewage from G14 and G15, use of surfactants, industrial plants are present, no ETP, huge industrial discharge from CETP along with sewage)
River Yamuna System			
Y-1	River Yamuna Yamunotri	3293	Origin of river Yamuna (no source of water pollution)
Y-2	River Yamuna Dakpathar U/s	490	Sewage and muck of hydro-electric projects from Dhakphatar area (agriculture is practiced, but without fertilizers, Govt. projects are present in this region, Lakhwadvyasi hydro-electric project is present)
Y-3	River Yamuna Vikas Nagar D/s	430	Sewage and industrial sources of water pollution (TV manufacturing, Tube light manufacturing, flour mill, pulse mill, spice mill, biggest human settlement, pharmaceutical industry (Elder Pharma & Intas Pharma), automobile suspension unit (Talbro), metal industries, sewage lines from towns and use of fertilizers)
Y-4	River Tons D/s Dehradun	399	Sources of water pollution are sewage and industries (pharmaceutical and suspension unit, pesticide formulation, dairy and poultry farm, shoe making, soap and detergent, home cleaner unit (phenyl), miscellaneous units and flow rate is low during winter)
Y-5	River Yamuna after confluence with river Tons	427	Sources of water pollution are sewage and industries—(Dhobighat (open air laundromat), animal washing, receiving tributaries from Selquai hill, Shakumbari hill, Sheetal river and same source of water pollution as mentioned at Y4)

Table 2 Water quality parameters, units and analytical method as measured during 2010 for the Ganga and Yamuna river water

Water quality	Unit	Analytical method
pH	pH unit	pH meter
Turbidity	NTU	Turbidity meter
Dissolved oxygen	mg/l	DO meter
Electrical conductivity	µs/cm	Conductivity meter
TDS	mg/l	TDS meter
Total suspended solid	mg/l	Gravimetric
Temperature	Degree Celsius	Thermometer
Biochemical oxygen Demand	mg/l	Winkler Azide method
Total Hardness	CaCO ₃ mg/l	Titrimetric
Total Alkalinity	CaCO ₃ mg/l	Titrimetric
Chloride	mg/l	Titrimetric
Sodium	mg/l	Flame photometer
Potassium	mg/l	Flame photometer
Phosphate	mg/l	UV spectrophotometer
Lead	mg/l	Atomic absorption spectrophotometer
Copper	mg/l	Atomic absorption spectrophotometer
Iron	mg/l	Atomic absorption spectrophotometer
Zinc	mg/l	Atomic absorption spectrophotometer

correlation coefficients between PCs and variables. The PCs are expressed by the following Eq. (1):

$$PC_i = a_{1i}x_1 + a_{2i}x_2 + \dots + a_{ni}x_n, \quad (1)$$

where PC_i is principal component i and a_{ji} is the component loading (correlation coefficient) of the original variable x . In order to maximize the variance of loadings, different matrix rotation methods are employed. In the present study, the varimax rotation methods were used for maximizing the variance of loadings. Sharma (1996), and Johnson and Wichern (2002) provide more details on the mathematical background of PCA. The PCA has been used in water quality assessment by various researchers to identify the most important water quality parameters and their potential pollution sources at various monitoring stations (Zhao et al. 2011; Razmkhah et al. 2010; Reid and Spencer 2009; Ouyang 2005).

Cluster analysis is another data reduction method that can be used to classify the entities with similar properties, so that objects within a group are similar to each other, but different from those in other groups (Shrestha and Kazama 2007). The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between analytical values from the samples. In this investigation, hierarchical agglomerative CA was performed on the normalized data set by means of Ward's method, using squared Euclidean distance as a measure of similarity (Zhao et al. 2011). The spatial variability of water quality in the whole river basin was determined by CA, using the linkage distance, D_{link}/D_{max} , which represents the quotient between the linkage distances for a

particular case divided by the maximal linkage distance. The quotient is then multiplied by 100 as a way to standardize the linkage distance (Zhao et al. 2011; Simeonov et al. 2003; Wunderlin et al. 2001).

Result and Discussion

Seasonal Variation of Water Quality

The summary statistics in terms of mean \pm standard deviation for the eighteen parameters for the summer and winter seasons is provided in Table 3. In order to assess the seasonal variability of the water quality parameters, Wilcoxon signed-rank test for paired observations was carried out for river Ganga and paired t test performed on river Yamuna. Table 3 provides the results of Wilcoxon signed-rank test and t test (Table 4).

The test shows that for river Ganga, all the water quality parameters, except BOD, pH is statistically significant ($p < 0.05$). All the parameters, except BOD, pH was found to be higher during summer as compared to winter period. This is an interesting observation considering that high flow rates are observed during summer period as compared to winter period due to melting of ice. Field investigations revealed that the major industries discharge untreated effluents in the river during summer period when the flow rate is high, while during winter period the industries treat their effluent and then discharge them in the river body, resulting in lower values. For Yamuna, in addition to pH, turbidity, TA, Cl^- , Na, K, Pb, phosphate, BOD, Fe and Zn

Table 3 Descriptive statistics of water quality parameters for rivers Ganga and Yamuna in summer and winter seasons

Parameter	River Ganga			River Yamuna		
	Summer Mean \pm SD	Winter Mean \pm SD	Number of water Samples (<i>N</i>)	Summer Mean \pm SD	Winter Mean \pm SD	Number of water Samples (<i>N</i>)
PO ₄ ⁻³	0.143 \pm 0.2236	0.1076 \pm 0.1918	15	0.124 \pm 0.1516	0.064 \pm 0.827	5
Temp.	19.3 \pm 2.9	17.4 \pm 2.9	15	18.6 \pm 4.8	17.4 \pm 4.03	5
DO	5.5 \pm 2.5	7.4 \pm 2.1	15	7.1 \pm 1.5	7.6 \pm 1.3	5
pH	8.1 \pm 0.2	8.0 \pm 0.33	15	7.9 \pm 0.12	8.0 \pm 0.22	5
BOD	4.5 \pm 1.7	3.7 \pm 1.0	15	2.4 \pm 0.69	3.9 \pm 0.89	5
Turbidity	21.6 \pm 15.7	18 \pm 13.3	15	16.8 \pm 12.4	10.8 \pm 6.7	5
TH	201.3 \pm 45.9	185.8 \pm 42.3	15	187.6 \pm 33.1	176.0 \pm 33.7	5
TA	102.1 \pm 18.7	97 \pm 20.03	15	125.2 \pm 36.0	110.8 \pm 21.3	5
TDS	268.2 \pm 115.7	169.8 \pm 68.9	15	197.4 \pm 35.6	155.4 \pm 35.9	5
TSS	153.2 \pm 137.0	118.1 \pm 121.9	15	98.8 \pm 112.1	79.8 \pm 104.2	5
Cl	10.5 \pm 3.6	11.9 \pm 4.8	15	6.8 \pm 6.5	10.8 \pm 6.9	5
EC	264.9 \pm 85.4	221.7 \pm 72.1	15	243.6 \pm 66.1	210.0 \pm 54.6	5
Na	11.9 \pm 2.4	11.1 \pm 2.5	15	8.6 \pm 3.9	8.6 \pm 3.82	5
K	60.9 \pm 40.7	54.6 \pm 35.2	15	42.1 \pm 25.3	41.2 \pm 23.5	5
Pb	0.3 \pm 0.3	0.15 \pm 0.18	15	0.16 \pm 0.21	0.07 \pm 0.11	5
Cu	0.9 \pm 0.7	0.64 \pm 0.54	15	0.59 \pm 0.49	0.39 \pm 0.43	5
Fe	0.27 \pm 0.08	0.24 \pm 0.09	15	0.17 \pm 0.06	0.15 \pm 0.05	5
Zn	0.09 \pm 0.1	0.06 \pm 0.11	15	0.05 \pm 0.05	0.01 \pm 0.03	5

Table 4 Wilcoxon signed-rank test for river Ganga and *t* test for river Yamuna

S. no.	Parameters	Ganga (Wilcoxon signed-rank test)		Yamuna (<i>t</i> test)	
		Z	<i>p</i> value	<i>t</i> _{cal}	<i>p</i> value
1	PO ₄	-3.296	0.00	-1.91	0.12
2	Temp.	-3.015	0.00	-3.21	0.03
3	DO	-3.411	0.001	2.903	0.04
4	pH	-1.623	0.10	2.77	0.05
5	BOD	-1.762	0.78	-1.69	0.16
6	Turbidity	-3.108	0.00	-2.07	0.10
7	TH	-3.413	0.00	-3.79	0.01
8	TA	-1.908	0.00	-1.75	0.15
9	TDS	-3.411	0.00	-30.47	0.00
10	TSS	-3.408	0.00	-3.82	0.01
11	Cl	-2.753	0.00	1.93	0.12
12	EC	-3.408	0.00	-3.09	0.03
13	Na	-2.783	0.00	0.10	0.92
14	K	-3.067	0.00	-1.07	0.34
15	Pb	-3.206	0.00	-1.74	0.15
16	Cu	-3.297	0.00	-3.05	0.03
17	Fe	-2.134	0.03	-1.20	0.29
18	Zn	-2.726	0.00	-2.39	0.07

were also found to be statistically insignificant ($p > 0.05$) and all other remaining parameters were found to be statistically significant. All statistically significant parameters were found to have higher values during summer as compared to winter period. It was due to the same reason as mentioned for Ganga. No seasonal variation was observed in the statistically insignificant parameters for the two rivers.

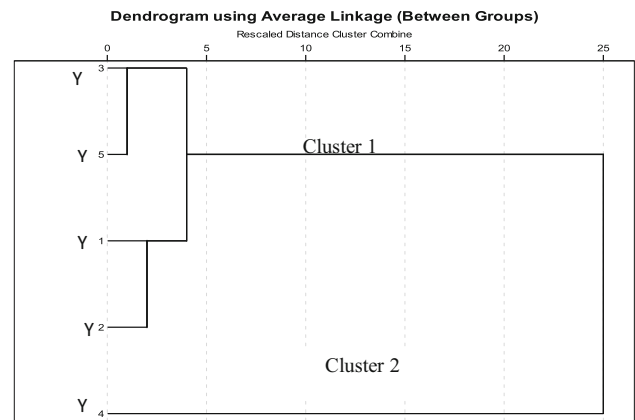
Clustering of Similar Monitoring Stations

Cluster analysis was applied to find out the monitoring stations having similar characteristics in terms water quality parameters. The dendrogram in Figs. 2 and 3 shows grouping of all 15 monitoring stations of river Ganga and 5 monitoring stations of river Yamuna into two statistically significant clusters at $(D_{Link}/D_{max})_{100} < 5$. For the summer data, eight monitoring stations of the river Ganga (1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13 and 14) formed the first cluster; these 13 stations primarily receive raw sewage from domestic sources and sewage muck from hydro-electric projects. The stations 8 and 15, together form the second cluster. These two stations receive effluents from domestic sources as well as from automobile, electroplating and pharmaceutical industries. Similar clusters were formed for winter data.

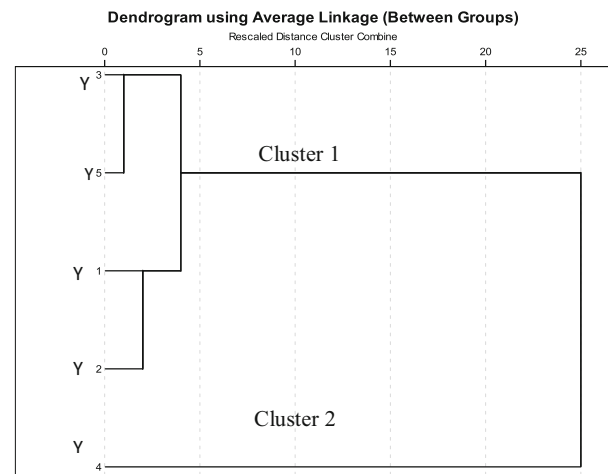
Summer data for river Yamuna resulted in two clusters (Fig. 3). Cluster 1 comprises stations 1, 2, 3 and 5, while cluster 2 comprises station 4. Stations 1, 2, 3 and 5 receive sewage from domestic sources, whereas station 4 receives pollution load primarily from dairy waste in addition to the domestic sewage. The winter data also resulted in two clusters, with station 1, 2, 3 and 5 forming the first cluster, and station 4 forming the second cluster. These two clusters represent low and high pollution sites.

Sources Attribution of Pollution Sources

The PCA was carried out for the summer and winter data set for the two rivers (Tables 5, 6). The Kaiser’s criterion (Kaiser 1960) was followed to determine the number of

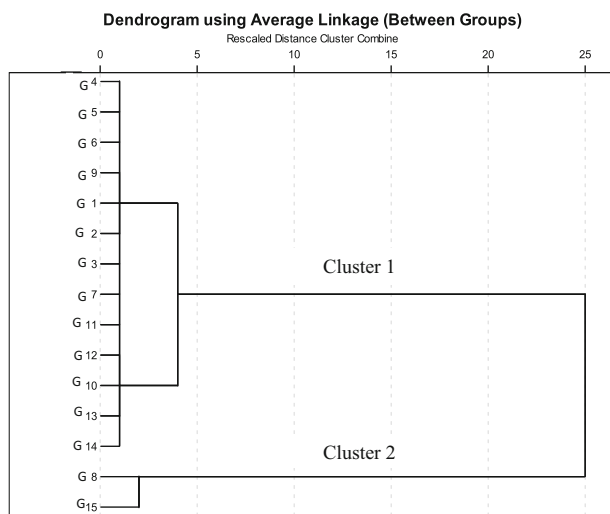


(a) summer

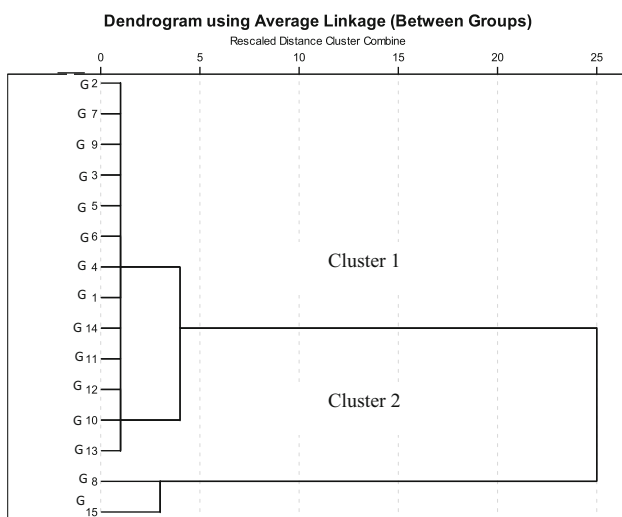


(b) winter

Fig. 3 Dendrogram of cluster analysis for sampling stations according to surface water quality of Yamuna



(a) summer



(b) Winter

Fig. 2 Dendrogram of cluster analysis for sampling stations according to surface water quality of Ganga

PCs to be extracted. According to these criteria, the number of PCs having eigen value equal to or greater than 1 is to be accepted as the possible sources of variance in the data set. The analysed summer data for Ganga resulted in four PCs with eigenvalues greater than 1, explaining 85 % of the variance. A reasonable interpretation of the four PCs can be done using the values of the respective loadings augmented with the knowledge of major contributors of pollution at various sites. The loadings with absolute values greater than 0.5 of the maximum value are selected for the PC interpretation (Jolliffe 1986). Further, the loadings were classified as ‘strong’, ‘moderate’ and ‘weak’ corresponding to absolute loading values of > 0.75, 0.75–0.50 and 0.5–0.30, respectively (Liu et al. 2003). The first component is a measure of PO_4^{-3} , TDS, TSS, K, Pb, Cu, Fe and Zn, parameters, which is indicative of mixed source of pollution comprising of natural and anthropogenic source (domestic sewage as well as industrial discharge) of pollution. The G8 receives large amount of debris due to continuous

Table 5 Results of PCA for rivers Ganga and Yamuna (for summer and winter seasons)

Parameters	River Ganga								River Yamuna						
	Summer				Winter				Summer			Winter			
	1	2	3	4	1	2	3	4	1	2	3	1	2	3	
PO ₄ ³⁻	0.86				0.91					0.98			0.98		
Temp.		0.67		0.50			0.68		0.60		0.63			0.84	
DO		-0.79		-0.63		-0.78			-0.53	-0.83		-0.80			
pH				0.78					0.57		0.77			0.96	
BOD		0.81						0.89	-0.82					-0.57	
Turbidity		0.64	0.73			0.97			0.61	0.74		0.83			
TH		0.83				0.57		0.54	0.78			0.50	0.84		
TA				0.72		0.53			0.98				0.99		
TDS	0.89				0.81				0.77	0.50		0.50	0.85		
TSS	0.52		0.76		0.62	0.68				0.88		0.80			
Cl ⁻				0.85			0.80		0.69	0.61		0.86			
EC	0.53		0.56		0.63	0.51			0.84			0.63	0.73		
Na ⁺				0.57	0.58		0.60		0.59	0.50		0.73			
K	0.92				0.94					0.92		0.93			
Pb	0.94				0.88						0.96			0.94	
Cu	0.69		0.57		0.91						0.94			0.99	
Fe	0.74				0.75						0.89			0.98	
Zn	0.91				0.93						0.99			0.77	
Eigen value	6.2	3.5	3.1	2.5	7.0	3.6	2.9	1.6	11.2	4.3	1.7	10.5	4.7	2	
% of variance	34.3	19.5	17.2	14	39.2	19.9	16.0	8.9	62.3	24.3	9.7	58.6	26.5	11.4	
Cumulative %	34.3	53.9	71	85.1	39.2	59.2	75.3	84.3	62.3	86.6	96.4	58.6	85.2	96.6	

Component loadings with value $\geq|0.5|$ have been shown in bold

Table 6 Extracted values of various factor analysis for water quality parameters

Component	River Ganga								River Yamuna					
	Summer				Winter				Summer			Winter		
	1	2	3	4	1	2	3	4	1	2	3	1	2	3
Eigen value	6.2	3.5	3.1	2.5	7.0	3.6	2.9	1.6	11.2	4.3	1.7	10.5	4.7	2
% of variance	34.3	19.5	17.2	14	39.2	19.9	16.0	8.9	62.3	24.3	9.7	58.6	26.5	11.4
Cumulative %	34.3	53.9	71	85.1	39.2	59.2	75.3	84.3	62.3	86.6	96.4	58.6	85.2	96.6

Component loadings with value $\geq|0.5|$ have been shown in bold

weathering of the mountains; the site also has a dysfunctional mega-hydroelectricity plant in river water, which continuously contributes Cu and Fe to river water due to corrosion and dismantling of machinery. Jain (2002) also reported the presence of phosphate in the same river stretch due to breakdown of the apatitic rocks. The TSS due to their composition—CO₃²⁻, Fe, MnO, clay and organic detritus—show positive correlation with Zn, Cu and Fe. The constituents of TSS play a significant role by providing an active surface for trace metals to be adsorbed (Boughriet et al. 1992); Zn, for instance, is positively

correlated with Fe due to Zn adsorption by hydrous Fe (Adbo 2004). High erosion rate of the unstable rocks, drainage basin with a number of steep elevation and several tributaries are merging in this stretch, leading to high sediment in this belt (Das 1991). The major industries upstream of G15 site that contribute to pollution load are electroplating, pharmaceuticals and automobiles. In the correlation matrix (Table 7) also, it was found that Pb was significantly correlated with Fe, Cu and Zn. Oketola et al. (2013) also reported that Pb shared strong relationship with Cu. Temperature, DO, BOD, turbidity and

TH are the main contributors of PC2, which contrast DO to other parameters and accounts for 19 % of the variance. This component points to an organic source of pollution such as agro-based industries. In correlation matrix, it was found that temperature was negatively correlated to DO. Kumari et al. (2012) reported that as temperature increases, it reduces the dissolution of ambient DO in river water. The third PC explains 17 % of variance and is a measure of turbidity, TA, TSS, EC and Cu. Turbidity, TA and TSS show very high loading factors, which is indicative of natural source of pollution—continuous erosion of soil from the river banks. Table 7 shows that TSS is significantly correlated to Na, K, EC, Pb, Cu and Fe. Temperature, pH, DO, Cl^- and Na^+ are the main contributors in PC4, which also contrast DO to other parameters; it may be attributed to food processing industry and slaughterhouses. PC4 explains 14 % of the total variance. In correlation matrix (Table 7), chloride was reported to be moderately correlated to Na.

The results of the summer data for river Yamuna are also shown in Tables 5 and 6. The first three PCs explain 96.4 % of the variance. The PC1 is a measure of temperature, DO, pH, BOD, turbidity, TH, TA, TDS, Cl^- , EC and Na^+ in which, DO and BOD contrast with other parameters, it accounts for 62.3 % of the variance. In correlation matrix (Table 8), TA shows strong positive correlation with TDS, Cl, EC and Na^+ . The main industries contributing to the pollution load are detergent making, poultry farms and pharmaceutical industries. The second PC comprises PO_4^{3-} , DO, turbidity, TSS, Cl^- , Na^+ , K and TDS, with DO contrasting with other parameters. A very strong loading of PO_4^{3-} and K (above 0.9) is due to abundance of K naturally in that region and PO_4^{3-} is mainly contributed by agricultural, sewage and natural processes. This component accounts 24.3 % of the variance. It was found in the correlation matrix (Table 8) that phosphate share strong positive correlation with K, TSS, turbidity, Cl and negative correlation with DO. Temperature, pH, Pb, Cu, Fe and Zn contributes to the third PC, which accounts for 9.7 % of the variance. This can be attributed as industrial source, with effluents contributed by tube light manufacturing, electronics and fertilizer industry. Peng et al. (2010) and Odoemelam et al. (2014) reported that high pH promotes the precipitation of the heavy metals from the insoluble hydroxides, oxides and bicarbonates present in sediments, whereas increases in temperature accelerate the dissolution of bicarbonates, hydroxides in turn releasing metals from water-soluble, carbonate, and ex-changeable fraction present in the sediments (Li et al. 2013). Correlation matrix (Table 8) shows that Pb shared good positive correlation with other heavy metals (Cu, Fe and Zn) also.

The results of PCA for the winter data are also presented in Tables 5 and 6. For Ganga, the first four components account for 84.3 % of the variance. The first component is a measure of PO_4^{3-} , TDS, TSS, EC, Na^+ , K, Cu, Pb, Fe and Zn pointing again to the same source of variations as one of the summer data—natural and anthropogenic source. The PC2 explains 19 % of the variance and comprises DO, TH, turbidity, TA, TSS, Fe and EC; with DO contrasting with other parameters and it was found that turbidity share significant correlation with TH, TA, TSS and EC (Table 9). The PC2 may be attributed to organic source of pollution, similar to summer data. The third PC is a measure of temperature, Cl^- and Na^+ contribute to 16 % of the variance. High loading of Na^+ and Cl^- may be attributed to food processing industry, slaughterhouses and water purification technologies. Chloride was found to be strongly correlated to Na (Table 9). The fourth PC explains 8.9 % of the variance and comprises BOD and TH again pointing to anthropogenic pollution (domestic wastewater).

The results of PCA for the Yamuna river data for the respective winter period gave three components, which accounts for 96.6 % of the total variance (Tables 5, 6). The first PC contrasts DO with other parameters— PO_4^{3-} , turbidity, TH, TSS, TDS, Cl^- , EC, K and Na^+ , indicating anthropogenic source from domestic sewage and effluents from industries similar to the one observed for summer data. Singh et al. (2005) reported that high TDS and TSS concentration is the indication of domestic sewage. The second PC is a measure of temperature, pH TH, TA, TDS, EC. It accounts for 26.5 % of variance. In correlation matrix (Table 10), it was found that TA was highly correlated to TDS and EC. The third PC explains 11.4 % of the variance and comprises Pb, Cu, Fe and Zn contrast to BOD as the main constituents. This is similar to the sources attributed in summer data.

Overall, the first four PCs in both summer and winter data for the river Ganga account for 86 % of the total variance. This proportion of accounted variance is substantial to provide adequate representation of data. The nature of the first two components is approximately the same. The first PC in both the datasets can be attributed to a mixed source of pollution comprising natural (weathering of rocks, soil erosion) and anthropogenic sources (domestic, agricultural runoff and industrial) of pollution. The second PC points to an organic source of pollution, such as agro-based industries. The nature of the fourth PC for the summer data is similar to the third PC of the winter data, both representing an industrial source of pollution such as food processing industry.

The difference exists between the two, PC3 for summer data and PC4 for winter data. The PC3 for summer data represents a natural source of pollution—soil erosion from

Table 7 Correlation analysis of Ganga River in summer

	Phosphate	Temp	DO	pH	BOD	Turbidity	TH	TA	TDS	TSS	Cl	EC	Na	K	Pb	Cu	Fe	Zn
Phosphate	1.000																	
Temp	0.389	1.000																
DO	-0.485	-0.758	1.000															
pH	0.341	0.433	-0.499	1.000														
BOD	0.252	-0.557	-0.757	0.463	1.000													
Turbidity	0.094	0.599	-0.744	0.405	0.518	1.000												
TH	0.298	0.560	-0.586	0.089	0.532	0.583	1.000											
TA	0.469	0.321	-0.432	0.503	0.143	0.527	0.248	1.000										
TDS	0.935	0.335	-0.061	0.353	0.188	0.154	0.364	0.354	1.000									
TSS	0.668	0.521	-0.395	0.535	0.301	0.709	0.418	0.694	0.711	1.000								
Cl	0.426	0.602	-0.505	0.646	0.395	0.260	0.345	0.611	0.360	0.382	1.000							
EC	0.731	0.706	-0.589	0.504	0.388	0.620	0.601	0.692	0.693	0.847	0.592	1.000						
Na	0.405	0.831	-0.516	0.537	0.481	0.590	0.410	0.558	0.331	0.598	0.663	0.705	1.000					
K	0.914	0.504	-0.058	0.431	0.233	0.051	0.336	0.381	0.881	0.620	0.558	0.684	0.508	1.000				
Pb	0.859	0.400	-0.015	0.299	0.165	0.163	0.456	0.416	0.884	0.692	0.413	0.677	0.519	0.913	1.000			
Cu	0.732	0.526	-0.117	0.397	0.191	0.386	0.292	0.577	0.711	0.827	0.424	0.774	0.718	0.756	0.844	1.000		
Fe	0.650	0.550	-0.128	0.223	0.258	0.456	0.563	0.413	0.692	0.741	0.281	0.655	0.534	0.729	0.811	0.735	1.000	
Zn	0.710	0.198	0.108	0.234	0.091	-0.075	0.419	0.243	0.743	0.442	0.378	0.473	0.282	0.852	0.878	0.567	0.660	1.000

Table 8 Correlation analysis of Yamuna River in summer

	Phosphate	Temp	DO	pH	BOD	Turbidity	TH	TA	TDS	TSS	Cl	EC	Na	K	Pb	Cu	Fe	Zn
Phosphate	1.000																	
Temp	0.587	1.000																
DO	-0.878	-0.668	1.000															
pH	-0.071	0.761	-0.093	1.000														
BOD	0.458	0.473	-0.824	0.152	1.000													
Turbidity	0.793	0.555	-0.968	0.000	-0.765	1.000												
TH	0.537	0.962	-0.711	0.727	-0.978	0.663	1.000											
TA	0.280	0.679	-0.655	0.544	-0.887	0.721	0.843	1.000										
TDS	0.627	0.943	-0.824	0.636	-0.996	0.762	0.973	0.843	1.000									
TSS	0.940	0.676	-0.861	0.062	-0.688	0.827	0.685	0.459	0.710	1.000								
Cl	0.738	0.932	-0.858	0.531	-0.966	0.807	0.963	0.789	0.972	0.845	1.000							
EC	0.593	0.887	-0.842	0.580	-955	0.814	0.956	0.908	0.988	0.694	0.958	1.000						
Na	0.586	0.822	-0.786	0.549	-0.865	0.661	0.785	0.658	0.893	0.519	0.801	0.864	1.000					
K	0.968	0.737	-0.911	0.143	-0.720	0.795	0.672	0.398	0.768	0.903	0.830	0.723	0.763	1.000				
Pb	0.106	0.742	-0.042	0.877	-0.474	-0.138	0.589	0.182	0.507	0.148	0.453	0.389	0.486	0.285	1.000			
Cu	0.156	0.541	-0.056	0.621	-0.279	-0.183	0.317	-0.082	0.343	0.012	0.246	0.212	0.559	0.337	0.851	1.000		
Fe	0.394	0.878	-0.415	0.808	-0.703	0.220	0.738	0.377	0.746	0.363	0.671	0.646	0.810	0.592	0.900	0.874	1.000	
Zn	0.111	0.572	0.024	0.692	-0.272	-0.253	0.352	-0.087	0.329	0.025	0.257	0.192	0.467	0.282	0.928	0.977	0.872	1.000

Table 9 Correlation analysis of Ganga River in winter

	Phosphate	Temp	DO	pH	BOD	Turbidity	TH	TA	TDS	TSS	Cl	EC	Na	K	Pb	Cu	Fe	Zn
Phosphate	1.000																	
Temp	0.391	1.000																
DO	-0.279	-0.688	1.000															
pH	-0.063	0.227	-0.217	1.000														
BOD	0.059	0.476	-0.820	0.180	1.000													
Turbidity	0.088	0.448	-0.841	0.267	0.183	1.000												
TH	0.376	0.432	-0.670	0.043	0.390	0.570	1.000											
TA	0.493	0.495	-0.505	0.317	0.060	0.612	0.339	1.000										
TDS	0.945	0.522	-0.421	0.136	0.089	0.296	0.398	0.615	1.000									
TSS	0.704	0.468	-0.609	0.161	-0.096	0.677	0.444	0.756	0.824	1.000								
Cl	0.427	0.725	-0.530	0.366	0.285	0.298	0.479	0.628	0.501	0.343	1.000							
EC	0.756	0.692	-0.707	0.197	0.115	0.572	0.640	0.724	0.827	0.841	0.641	1.000						
Na	0.603	0.769	-0.628	0.222	-0.132	0.443	0.552	0.659	0.640	0.625	0.773	0.805	1.000					
K	0.905	0.443	-0.281	0.049	-0.152	0.053	0.405	0.501	0.839	0.645	0.532	0.706	0.708	1.000				
Pb	0.732	0.276	-0.286	-0.099	-0.050	0.191	0.708	0.344	0.665	0.552	0.337	0.596	0.629	0.779	1.000			
Cu	0.803	0.385	-0.317	0.095	-0.234	0.188	0.457	0.546	0.776	0.735	0.445	0.724	0.718	0.880	0.794	1.000		
Fe	0.669	0.520	-0.586	0.147	-0.256	0.519	0.583	0.565	0.720	0.845	0.406	0.747	0.748	0.771	0.755	0.879	1.000	
Zn	0.772	0.152	-0.129	0.135	0.170	0.007	0.468	0.342	0.718	0.578	0.314	0.580	0.482	0.867	0.832	0.884	0.744	1.000

Table 10 Correlation analysis of Yamuna River in winter

	Phosphate	Temp	DO	pH	BOD	Turbidity	TH	TA	TDS	TSS	Cl	EC	Na	K	Pb	Cu	Fe	Zn
Phosphate	1.000																	
Temp	0.494	1.000																
DO	-0.848	-0.600	1.000															
pH	0.290	0.954	-0.532	1.000														
BOD	0.805	-0.093	0.838	0.015	1.000													
Turbidity	0.892	0.593	-0.994	0.506	-0.845	1.000												
TH	0.537	0.943	-0.763	0.926	-0.310	0.732	1.000											
TA	0.105	0.858	-0.457	0.963	0.072	0.405	0.886	1.000										
TDS	0.505	0.967	-0.708	0.944	-0.233	0.679	0.995	0.897	1.000									
TSS	0.914	0.670	-0.903	0.559	0.694	0.940	0.701	0.388	0.670	1.000								
Cl	0.797	0.739	-0.699	0.530	-0.486	0.704	0.753	0.404	0.756	0.688	1.000							
EC	0.682	0.957	-0.806	0.892	-0.376	0.801	0.971	0.791	0.969	0.827	0.808	1.000						
Na	0.638	0.716	-0.733	0.588	-0.499	0.696	0.830	0.560	0.817	0.567	0.923	0.798	1.000					
K	0.939	0.739	-0.867	0.554	-0.662	0.889	0.769	0.394	0.753	0.903	0.931	0.868	0.824	1.000				
Pb	-0.324	0.345	0.499	0.288	0.775	-0.493	0.096	0.248	0.188	-0.325	0.163	0.078	0.060	-0.095	1.000			
Cu	0.001	0.235	0.365	0.043	-0.461	-0.333	0.003	-0.064	0.083	-0.182	0.399	0.054	0.201	0.135	0.869	1.000		
Fe	0.025	0.461	0.266	0.306	0.515	-0.237	0.207	0.188	0.293	-0.050	0.446	0.254	0.257	0.217	0.931	0.958	1.000	
Zn	-0.364	0.498	0.394	0.522	0.802	-0.399	0.265	0.502	0.352	-0.224	0.097	0.225	0.047	-0.095	0.939	0.665	0.814	1.000

Table 11 Significant water quality parameters contributing to water quality in each season to rivers Ganga and Yamuna

River	Season	Parameters with strong positive factor loading	Parameters with strong negative factor loading
River Ganga	Summer	PO ₄ ⁻³ , TDS, K, Pb, Zn, TH, TSS, pH, Cl	Nil
	Winter	PO ₄ ⁻³ , TDS, K, Pb, Cu, Fe, Zn, BOD, Turbidity, Cl.	DO
River Yamuna	Summer	PO ₄ ⁻³ , TH, BOD, TA, TDS, EC, DO, TSS, K, pH, Pb, Cu, Fe, Zn	DO
	Winter	PO ₄ ⁻³ , Turbidity, TSS, Cl, K, Temperature, pH, TH, TA, TDS, Pb, Cu, Fe, Zn	DO

river banks, while PC4 for winter data is attributed due to BOD and TH. For the river Yamuna, three PCs were identified for both datasets, which explain 96 % of the total variance. The nature of the three components is about the same; PC1 representing anthropogenic source of pollution from domestic and industrial sources, PC2 representing agricultural runoff and PC3 representing industrial source. The most significant parameters for summer and winter data for the two rivers based on the PCA are presented in Table 11.

Conclusions

Uttarakhand is a newly carved out state in India. The state is in the process of preparing zoning atlas for identifying appropriate locations for industry. An essential requirement for preparing the zonal atlas is primary data of water quality along the stretches of the major rivers that flow through the state. This data once collected and analysed would help in taking informed decisions related to designing of comprehensive water quality monitoring network, identification of appropriate sites for locating industries and rehabilitation plans for the rivers. The primary data collected and subsequently analysed in the present study provide useful insights. For instance, the water quality parameters were found to be higher during summer season compared to the winter season despite of higher dilution capacity achieved due to increased flow rate resulting from ice melting; on investigation, it was found that the industries in the catchments of the two rivers discharge their untreated effluents during summer period relying on natural scavenging capacity of the rivers. This is not an approved practice. Findings of the study can help the implementing authorities to take corrective measures by forcing the polluting industries to discharge the effluents after recommended treatment, as is done during the winter period. Further, the CA findings reveal formation of two similar clusters for river Ganga and two clusters for river Yamuna. The two clusters identified for Ganga receive pollution loads from domestic and industrial sources and

sewage muck from hydro-electric projects, respectively; whereas the two clusters formed for river Yamuna can be categorized as low and high pollution sites, with the former receiving major contribution from domestic sources and the latter from domestic and dairy sources, respectively. Thus, CA findings would help in grouping similar monitoring stations into clusters, which would further help in facilitating the design of optimal number of monitoring stations; this would help in reducing the number of monitoring stations and hence the cost of sampling for both the rivers. Finally, the PCA revealed underlying relationships amongst the water quality parameters and helped identifying sources of pollution. For the river Ganga, four components or sources of pollution were identified that can be attributed to mixed sources (natural and anthropogenic), organic source of pollution—agro based industries, natural source of pollution—riverbank erosion and food processing industry; for river Yamuna, three major sources of pollution were identified—anthropogenic sources of pollution from domestic and industrial sources, agricultural runoff and industrial source. Thus, the PCA helped in identifying the point and nonpoint sources of pollution for the entire stretch of rivers Ganga and Yamuna in Uttarakhand. This essentially will help in prioritizing in control measures with respect to different sources of pollution.

Acknowledgments The authors gratefully acknowledge the cooperation of Mr. Neeraj Sharma for editorial assistance. The editors and anonymous referees are gratefully acknowledged.

References

- Adbo MH (2004) Distribution of some chemical elements in the recent sediments of Damietta Branch, River Nile, Egypt. *Egypt Acad Soc Environ Develop (D-Environmental Studies)* 5(2):125–146
- APHA, AWWA, WEF (1998) *Standard Methods for the Examination of Water and Wastewater*, 20th edn. American Public Health Association/American Water Works Association/Water Environment Federation, Washington
- Bonanssea M, Ledesma C, Rodriguez C, Pinotti L (2014) Water quality assessment using multivariate statistical techniques in

- Río Tercero Reservoir, Argentina. *Hydrol Res.* doi:[10.2166/nh.2014.1](https://doi.org/10.2166/nh.2014.1)
- Boughriet A, Ouddane B, Fischer JC, Wartel M, Leman G (1992) Variability of dissolved Mn and Zn in the Seine estuary and chemical speciation of these metals in suspended matter. *Water Res* 26(10):1359–1378
- Das TM (1991) Some aspects of the biogeography of the Ganga. In: Krishna Murti CR, Bilgrami KS, Das TM, Mathur RP (eds) *The Ganga a scientific study*. The Ganga Project Directorate, New Delhi, pp 187–193
- Jain CK (2002) A hydro-chemical study of a mountainous watershed: the Ganga, India. *Water Res* 36:1262–1274
- Johnson RA, Wichern DA (2002) *Applied multivariate statistical analysis*. Pearson Education Inc., Upper Saddle River
- Jolliffe IT (1986) *Principal component analysis*. Springer, University of Geneva, New York
- Kaiser HF (1960) The application of electronic computers to factor analysis. *Educ Psychol Meas* 20:141–151
- Kannel PR, Leea S, Leeb YS (2008) Assessment of spatial–temporal patterns of surface and ground water qualities and factors influencing management strategy of groundwater system in an urban river corridor of Nepal. *J Environ Manag* 86:595–604
- Kumari M, Tripathi S, Pathak V, Tripathi BD (2012) Chemometric characterization of river water quality. *Environ Monit Assess* 185:3081–3092. doi:[10.1007/s10661-012-2774-y](https://doi.org/10.1007/s10661-012-2774-y)
- Li Y, Xu L, Li S (2009) Water quality analysis of the Songhua River basin using multivariate techniques. *J Water Resour Prot* 2:110–121
- Li H, Shi A, Li M, Zhang X (2013) Effect of pH, temperature, dissolved oxygen, and flow rate of overlying water on heavy metals release from storm sewer sediments. *J Chem.* doi:[10.1155/2013/434012](https://doi.org/10.1155/2013/434012)
- Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. *Sci Total Environ* 313(1–3):77–89
- Odoemelam SA, Okorie DO, Oko IO (2014) Physicochemical parameters and heavy metal content of water, fish and sediments from cross river at afikpo north local Government area of Ebonyi state, Nigeria. *J Chem* 3(1):1–11
- Oketola A, Adekolurejo SM, Osibanjo O (2013) Water quality assessment of River Ogun using multivariate statistical techniques. *J Environ Protect* 4(5):466–479. doi:[10.4236/jep.2013.45055](https://doi.org/10.4236/jep.2013.45055)
- Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. *Water Res* 39:2621–2635
- Peng J, Chen S, Dong P (2010) Temporal variation of sediment load in the Yellow River basin, China, and its impacts on the lower reaches and the river delta. *Catena* 83(2–3):135–147
- Razmkhah H, Abrishamchi A, Torkian A (2010) Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: a case study on Jajrood River (Tehran, Iran). *J Environ Manag* 91:852–860
- Reid MK, Spencer KL (2009) Use of principal components analysis (PCA) on estuarine sediment datasets: the effect of data pre-treatment. *Environ Pollut* 157:2275–2281
- Sharma S (1996) *Applied multivariate techniques*. John Wiley & Sons Inc, Hoboken
- Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ Model Softw* 22(4):464–475
- Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsas D, Anthemidis A (2003) Assessment of the surface water quality in Northern Greece. *Water Res* 37(17):4119–4124
- Singh KP, Malik A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): a case study. *Water Res* 38:3980–3992
- Singh KP, 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
- SPSS-10 (1999) *Statistical package for the social sciences*. Spss Inc, Chicago
- Wunderlin DA, Diaz MP, Ame MV, Pesce SF, Hued AC, Bistoni MA (2001) Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). *Water Res* 35(12):2881–2894
- Zhao J, Fu G, Lei K, Li Y (2011) Multivariate analysis of surface water quality in the three Gorges area of China and implications for water management. *J Environ Sci* 23(9):1460–1471