

Spatial variations in water quality of river Ganga with respect to land uses in Varanasi

Shikha Sharma¹ · Arijit Roy² · Madhoolika Agrawal¹

Received: 3 February 2016 / Accepted: 4 August 2016 / Published online: 15 August 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Water quality of a river is a function of surrounding environment and land use due to its connectivity with land, resulting in pollutants finding their way through land. This necessitates a spatially explicit study of river ecology. The paper presents a pioneer study to establish and explore the linkage between land use and water quality of river Ganga in Varanasi district. The land use land cover (LULC) map of 20 km of river stretch for buffer radii of 1000 m in Varanasi revealed that riparian vegetation is negligible in the district. The hierarchical cluster analysis of LULC data suggested that there are two major land use categories, viz., urban and agriculture. The land use wise principal component analysis (PCA) suggested that urbanized areas are major contributor of metals, whereas agricultural land contributes organic matter into the river. The Spearman correlation study revealed that with rising urbanization, the pollutant load into the river increased compared to that from agricultural land use. The statistical analysis of the data clearly concluded that water quality

of river Ganga at Varanasi was a function of adjacent land use. The study provides an insight anticipating the Indian government to embrace the relationship of land use to river water quality while formulating policies for the upcoming River Regulation Zone.

Keywords Agriculture · Hierarchical cluster analysis · Land use and land cover · Urban · Principal component analysis · Spearman correlation · Water quality

Introduction

Water is of prime importance for the sustenance of life and its quality is equally essential. Water possesses a unique property of dissolving and carrying with it in suspension a variety of substances thus easily being contaminated (Tiwari and Ali 1988). Among the forms of available fresh water sources, surface waters are the most easily available and also one of the most exploited, which include rivers. Rivers, besides serving industries and agriculture, act as source of water for drinking, bathing, washing, etc. as well as sink for huge loads of waste from industries, domestic sewage, and agriculture, making it one of the most vulnerable forms of surface water (Ravindra et al. 2003; Singh et al. 2005; Duan et al. 2013a).

River Ganga is lifeline of India with its basin extending internationally to Tibet, Nepal, and Bangladesh with an area of 1,086,000 km² (<http://india-wris.nrsc.gov.in>). The rising pollution of rivers is of global concern presently, with the World Bank disbursing a total of \$66 million for clean Ganga Mission (<http://www.businesstoday.in/>). The Ganga basin exhibits variability in terms of land use. River Ganga is the main trunk with 46 tributaries joining during its 2525-km journey before meeting the Bay of Bengal (<http://india-wris.nrsc.gov.in>). The entire length of river Ganga is divided

Responsible editor: Philippe Garrigues

Electronic supplementary material The online version of this article (doi:10.1007/s11356-016-7411-9) contains supplementary material, which is available to authorized users.

✉ Madhoolika Agrawal
madhoo.agrawal@gmail.com

Shikha Sharma
shikha.sharma190@gmail.com

Arijit Roy
arijitroy@iirs.gov.in

¹ Laboratory of Air Pollution and Global Climate Change, Department of Botany, Banaras Hindu University, Varanasi 221005, India

² Forestry and Ecology Department, Indian Institute of Remote Sensing, ISRO, Kalidas Road, Dehradun 248001, India

into three parts, viz., upper reaches (from origin to Narora), middle reaches (Narora to Ballia), and lower reaches (Ballia to its delta) (<http://india-wris.nrsc.gov.in/>). The Ganga main stem mean annual flow is 84.98 billion cubic meters with December to May as lean flow months (National River Conservation Directorate, Ministry of Environment and Forests, Government of India 2009). At present, river Ganga appears to be a lotic dump yard receiving and transporting waste in its 2525-km journey from Gangotri to Ganga Sagar (Sharma and Agrawal 2014). Such exploitation necessitates the monitoring of water quality not only evaluating the pollution status but also ensuring the efficient management and protection of aquatic life (Strobl and Robillard 2008; Duan et al. 2016).

Water quality of a river is a function of surrounding environment and land use, due to its sharing of boundary with land (Ward 1989; Zheng and Baetz 1999), resulting in pollutants finding their way through land-water interface (Zhang et al. 2010; Duan et al. 2015). Presently, land is experiencing a transit from natural and undisturbed land cover to artificial and disturbed anthropogenically influenced land use. An undisturbed land cover and large areas of natural landscapes are important for stabilizing the hydrology of the river (National River Conservation Directorate, Ministry of Environment and Forests, Government of India 2009). With the advancement in remote sensing and geographical information systems, there has been a paradigm change in assessment of water quality in relation to land use (Griffith 2002; Ierodiaconou et al. 2005; Rothwell et al. 2010). Internationally, there are a number of studies relating land use to water quality (Allan et al. 1997; Sliva and Williams 2001; Giri and Singh 2013; Huang et al. 2013; Hashmi 2013). However, there are only few studies in India mainly confined to small basins in South India (Chattopadhyay et al. 2005; Raj and Azeez 2010; Soman et al. 1997), and river Ganga has yet not been considered sufficiently with this perspective.

On looking at the existing correlation between pollution inputs and land use (Perry and Vanderklein 1996), there is a need for directing the riverine research spatially. The present requirement is to fill the void in research studies on how land use and population growth influence the freshwater at local, regional, and national scales (Ekness 2013), particularly in Indian context. In view of this, the paper presents a pioneer study with an effort to establish and explore the linkage between land use and water quality of river Ganga in Varanasi district. It is a preliminary study initially confined to a small stretch looking at the pollution problem in river Ganga from a landscape perspective, checking the suitability and applicability of such kind of studies first on smaller scale and then conducting it with broader perspective for the entire basin. This study has attempted for the first time to calculate the proportion of different land uses within 1000 m from the river bank along the entire stretch of river in Varanasi.

The focal point of the polluted state of river Ganga at Varanasi is the shift of built-up areas towards the river bank. Since the last century, the city is expanding in an unplanned way along the western side of the river stretch. It was hypothesized that (1) water quality of river has a significant relation with land use of riparian zone, and (2) types and sources of pollutants also change according to adjacent land uses.

Methods

Study area

The study area covers ~20 km of river Ganga stretch extending from 25°12'56.61"N, 83°00'37.98"E (upstream) to 25°20'15.67"N, 83°04'41.32"E (downstream) of Varanasi district. This district is situated on the western bank of the river. Varanasi is at a distance of 1295 km from the source (Gangotri glacier), with an average mean annual flow of river ranging from ~300 to 2500 m³/s (National River Conservation Directorate, Ministry of Environment and Forests, Government of India 2009). Varanasi lies in the middle reach of the Ganga catchment. The river meanders in Varanasi flowing from north to south. The total area of the district is 1535 km², supporting a population of 3.148 million persons (<http://varanasi.nic.in/>). This district is densely populated, with 2063 persons per square kilometer, as against the state average of 689 persons per square kilometer. The urban agglomeration is stretched between 25°16'06.80"N, 83°01'01.19"E (upstream) and 25°19'40.07"N, 83°02'34.67"E (downstream).

Sampling zones

Although there lies the disparity whether the land use of riparian buffer width or the whole of the catchment of an area affects the riverine conditions (Sliva and Williams 2001), still for the purpose of this study, the land use of buffer zone of 1000 m from the river bank frames the basis for selecting the sampling zones. The sampling zones were selected on the western bank of river Ganga in Varanasi with high population density. The entire waste generated in the city finds its way into the river through 23 sewage outfalls located on the western bank. There are 84 ghats (concrete staircases leading into the river) along the western bank, which form the major attractions for the visitors. The river Ganga in Varanasi is deeply revered, visited by innumerable local, domestic, and foreign tourists. The city, known for its festive and religious fervor, witnesses numerous people taking a holy dip in the sacred water of river Ganga. Therefore, monitoring the water quality of the river is of utmost concern near the western bank. At few zones in the rural area both upstream and downstream, agriculture is extensively practiced on the western bank, with

variable crops grown throughout the year. However, the eastern bank of the river has a high proportion of accumulated sand. The agriculture is not much extensively practiced in rural areas both upstream and downstream of this bank (only mustard cultivation during winters). This bank is sparsely populated, and very few people visit this bank. Also, there are no sewage outfalls. Therefore, for the present study, sampling zones were restricted to the western bank having intense activities.

Land use and land cover (LULC)

The land use and land cover map was prepared using software Arc GIS 10.2 from high-resolution Google earth imagery of the area. Onscreen visual interpretation technique was used to prepare the map of the land use and land cover of 1 km buffer on either side of the Ganga along a 20-km river stretch in Varanasi. A total of 23 classes, viz., residential, mixed built-up (comprising both residential and commercial), recreational (parks and garden), public–semi-public (schools and religious places like temples), public utility (ghats), commercial (shops), vegetated area (areas around buildings with major green cover), transportation (roads and narrow passages), open area (playing grounds in urban area without any vegetation), riparian vegetation, rural area, cropped area, fallow land, floriculture, orchards, scrubland land, wetland, ponds, river, other water bodies, river sand, point source, and mining commercial areas, were delineated following the National Resource Census (NRC) of government of India classification system (NRSA 2005). The entire river stretch of the study area was divided into 12 zones from upstream to downstream of Varanasi. The area (in hectares) of each land use feature in each zone was calculated via Arc GIS 10.2. Later, each land use feature was grouped together broadly into agriculture (cropped area, fallow, floriculture, orchards, scrubland), urban (residential, mixed built-up, recreational, semi-public, public utility, commercial, vegetated area, transportation, open area), and rural, and subsequently their percent contribution to the land cover was calculated.

Collection of water samples

Water samples were collected using a modified Swing sampler at 12 sampling zones with an aim to monitor the effects of land use on water quality of the river. Sampling was done in the middle of every month from April 2013 to March 2015 except for the months of August and September in both years and July in 2013 due to flood (Uttarakhand tragedy). Grab sampling method was applied for collecting the samples. Samples were collected within 5 m from the river bank at a depth of ~0.5 m in Teflon bottles of 2 L. The samples were transported and stored at 4 °C until analysis. The samples for dissolved oxygen (DO) analysis were collected in BOD bottles and

fixed onsite. The samples for metal analysis were preserved by acidifying with concentrated HNO₃ to pH <2.

Sample analysis

The water quality parameters, viz., temperature, conductivity, total dissolved solid, and pH, were analyzed on sites using handheld devices such as a digital thermometer, Milwaukee conductivity tester, Hanna TDS meter, and Milwaukee pH meter, respectively, and turbidity was measured in the laboratory using a nephelometer. Other parameters were analyzed in the laboratory using APHA (American Public Health Association) 1995 and other standard methods: dissolved oxygen (DO) and biological oxygen demand (BOD)—Winkler azide method; chemical oxygen demand (COD)—closed reflux method, Thermoreactor 420; hardness—EDTA titrimetric method; alkalinity—sulfuric acid titrimetric method; phosphate—colorimetric, molybdate; and nitrate—colorimetric, brucine. Metals, viz., chromium (Cr), copper (Cu), cadmium (Cd), lead (Pb), and zinc (Zn), were analyzed by atomic absorption spectrometry (Perkin Elmer, Analyst 800). Prior to analysis, samples were digested with concentrated HNO₃ (APHA Standard method 1995). The details of the parameters assessed, methods followed, instruments used, linearity ranges, detection limits, resolution, and accuracy are given in Supplementary Table 1.

The quality assurance of generated water quality data was ensured on a regular basis during sample collection, sample transportation, sample storage, and sample analysis. The chemicals used for analyses were of Merck analytical grade. Further, the reagent blank and standards of variable ranges were analyzed prior to performing the analysis of water samples and also after every five readings to ensure proper calibration of the instruments used. In titrimetric analyses, the titrating solutions were standardized against respective standard solutions. The standards used were from Sisco Research Laboratory Pvt. Ltd., Mumbai, India. The reproducibility was reported to be within ±5 %. Samples were analyzed in triplicates to further enhance the analytical precision of the data.

Statistical analysis

The data was subjected to various statistical tests using SPSS 16. First and foremost, the descriptive analyses, viz., range, mean, standard error, standard deviation, and variance of generated water quality data, were done. Subsequently, the data set was analyzed for the normality check using two methods, visual method (Box Plot) and normality test method [Shapiro-Wilk's test ($p > 0.05$)].

Land use statistics obtained after digitization was subjected to cluster analysis with a view to group zones with similar land use characteristics into a cluster or group. Of the three approaches for clustering, hierarchical agglomerative clustering

(HAC) is most commonly used (Yaari 1997). To measure the dissimilarity between the zones, Euclidean distance measure of dissimilarity was used. Mathematically, Euclidean distance is the square root of the sum of the squared differences in the variable values (Mooi and Sarstedt 2011). The clustering algorithm used was the Ward method. The Ward method performs the clustering by using analysis of variance approach starting with the most similar pair of objects and subsequently forming higher clusters step by step (Ward 1963).

Prior to conducting principal component analysis (PCA), Kaiser–Meyer–Olkin (KMO) and Bartlett’s sphericity tests were conducted to test whether PCA can be performed on data or not. KMO values higher than 0.6 are considered satisfactory (Arumugam et al. 2010), while values above 8 are considered good (Kaiser 1974). Bartlett test of sphericity is based on the null hypothesis that variables are orthogonal, and PCA can be performed only if the null hypothesis is rejected at 5 % level of significance (Jackson 1993). Subsequently, PCA of water quality data sets was performed after z-scale transformation on correlation matrix of data to extract significant principal components (PCs). The z-scale transformation avoids misclassification due to wide differences in data dimensionality (Simeonov et al. 2003; Singh et al. 2004). The rotation used was varimax with Kaiser normalization. During rotation, the axes are being rotated in a form that the clusters of variables fall closely (Osborne 2015). Varimax is the most commonly used rotation due to its development by improving other rotation algorithms (Osborne 2015). The maximum iterations for converging the rotation selected were 25. The scree plot and component with eigenvalue greater than 1 was used to identify the number of PCs to be retained (Jackson 1993; Vega et al. 1998).

Accounting for the non-normal distribution of measured water-quality parameters, Spearman *R* coefficient was used to study the correlation between agriculture land use and water quality and also subsequently for urban land use and water quality. Spearman’s correlation works by calculating Pearson’s correlation on the ranked values of data (McDonald 2009). Ranking is assigned from a higher to lower value, e.g., rank 1 is assigned to the largest value, followed by 2 to the next largest and so on (McDonald 2009). SPSS automatically converts the other forms of data to ranked data and subsequently calculates the correlation between the variables.

Results

Land use and land cover map

The LULC Map of the study region (Fig. 1) had all 23 classes as described in “Land use and land cover (LULC)” section. The percent contribution of each broad land use categories is given in Table 1. The maximum percentage of agriculture was

observed in zone 12 with 92.38 %, followed by zone 2 with 92.04 %, zone 3 with 86.65 %, zone 11 with 75.84 %, zone 1 with 68.11 %, and zone 4 with 58.66 % (Table 1). The maximum percentage of urban land use was observed in zone 8 with 99.80 % followed by zone 9 with 99.47 %, zone 7 with 99.14 %, zone 6 with 95.26 %, zone 10 with 51.26 %, and zone 5 with 46.29 % (Table 1). The rural area was observed in a few zones; maximum rural area was observed in zone 4 with 34.18 %, followed by zone 5 with 27.19 %, zone 11 with 24.15 %, zone 12 with 6.98 %, zone 3 with 5.56 %, zone 10 with 4 %, and zone 2 with 0.19 % (Table 1). However, riparian vegetation formed only 1.76 of the total land area in 12 sampling zones.

Clustering zones on the basis of land use

Hierarchical agglomerative cluster analysis clusters the object considering each object in a separate cluster and joining the clusters together step by step based on similarity until only one cluster remains (Lattin et al. 2003; Mckenna 2003). The clustering procedure generated a dendrogram representing two clusters having similar land use features (Fig. 2). Cluster 1 comprised zones 1, 2, 3, 4, 11, and 12 forming the agriculture land use. Cluster 2, on the other hand, was composed of zones 5, 6, 7, 8, 9, and 10 representing urban land use.

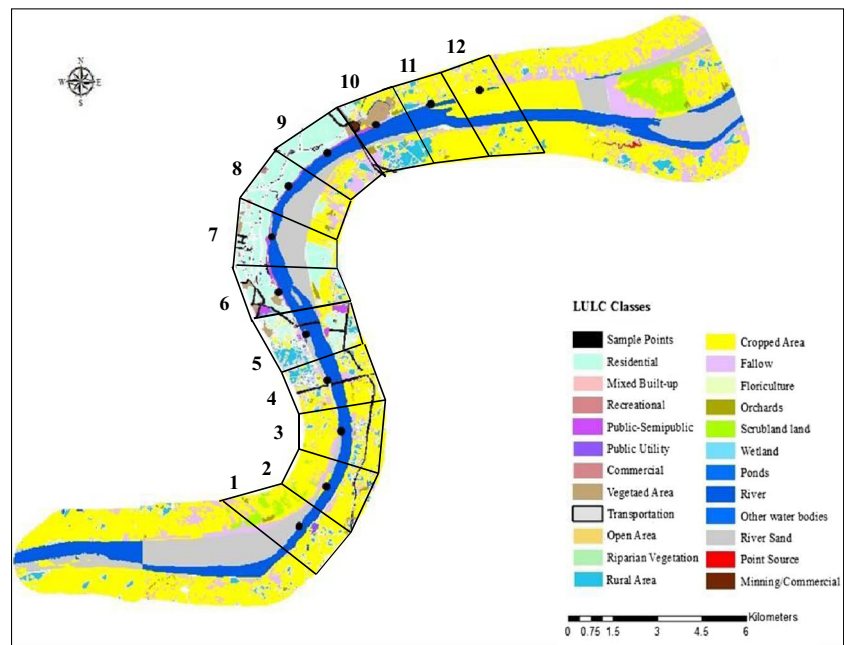
Descriptive statistics and sample characteristics of water quality parameters

The basic statistics of water quality data set is summarized in Supplementary Tables 2, 3, 4, and 5. The normality check of data via box plot suggested that the data in the present study was not normally distributed as box plots formed were not symmetric. The box plot of normally distributed data is symmetric with the median line approximately at the center of the box and also symmetric whiskers slightly longer than the subsections of the center box. The result obtained from the visual test was supported by Kolmogorov–Smirnov and Shapiro–Wilk test as ‘*p*’ value obtained was less than 0.05. Thus, both tests explained for non-normal distribution of data.

Land use wise principal component analysis of water quality parameters

The KMO measure of sampling adequacy in PCA for agricultural and urban land use was 0.83 and 0.81, respectively. Bartlett’s test of sphericity was significant for both agriculture and urban land use with $\chi^2(136) = 7394, p = 0.000$ and $\chi^2(136) = 6309, p = 0.000$, respectively. This suggested that the data was adequate enough to perform PCA. The rotation converged in 18 and 6 iterations in agricultural and urban land use PCA, respectively. PCA is applied to data sets to find low-dimensional combinations from high-dimensional data sets in

Fig. 1 Land use land cover map showing 12 zones of the study region



view to capture maximum variability in original data sets (Vega et al. 1998). The low-dimensional combinations generated are non-correlated (Vega et al. 1998). The components are formed in such a way that the first component captures the largest proportion of variability and the subsequent components capture the next proportion of the variability (Manage and Scariano 2013). The land use wise PCA analysis of water quality data sets resulted in four principal components with eigenvalue >1 (PCs) in both agriculture and urban land uses. The four PCs in agriculture land use and urban land use explained 81.95 and 78.35 % of the total variance, respectively. In the agriculture land use class, first PC represented 29.57 %, second PC 28.01 %, third PC 15.98 %, and fourth PC 8.38 % of the total variances. Similarly, in the urban land use, first PC

represented 26 %, second PC 24.24 %, third PC 16.22 %, and fourth PC 11.88 % of the total variance.

For the agricultural land use class, the parameters such as COD, BOD, pH, DO, Zn, and phosphate formed the first PC (Fig. 3), out of which COD, BOD, and pH showed strong positive loadings and DO with strong negative loadings, while Zn and phosphate were with moderate positive loadings. PC 2 explained strong positive loadings for temperature, Cr, Cu, Cd, and Pb while it explained moderate positive loading for hardness and alkalinity (Fig. 3). The TDS and conductivity formed the third PC with strong positive loading, while temperature and turbidity formed the moderate negative loading (Fig. 4). The fourth PC was formed by nitrate and phosphate with

Table 1 Percent contribution of three major land uses obtained from LULC map

Zones	Agriculture	Urban	Rural
Zone 1	68.11	0.0	0.09
Zone 2	92.04	0.0	0.19
Zone 3	86.65	0.28	5.56
Zone 4	58.66	3.00	34.18
Zone 5	19.62	46.29	27.19
Zone 6	3.44	95.26	0.0
Zone 7	0.0	99.14	0.0
Zone 8	0.0	99.80	0.0
Zone 9	0.0	99.47	0.0
Zone 10	37.31	51.26	4.00
Zone 11	75.84	0.0	24.15
Zone 12	92.38	0.0	6.98

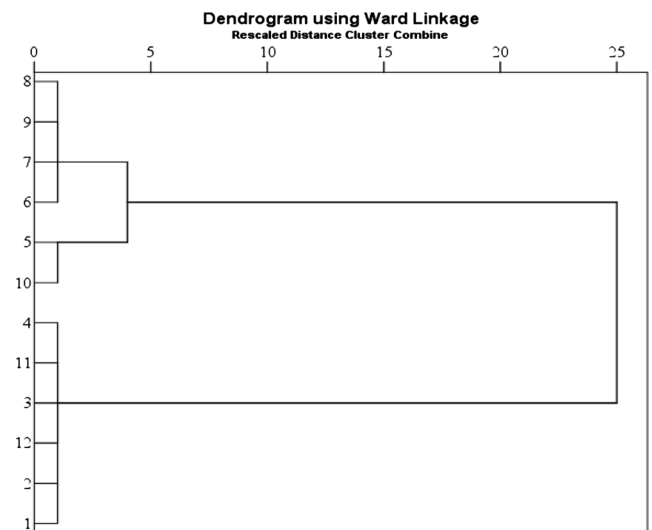


Fig. 2 Cluster formations of 12 sampling zones

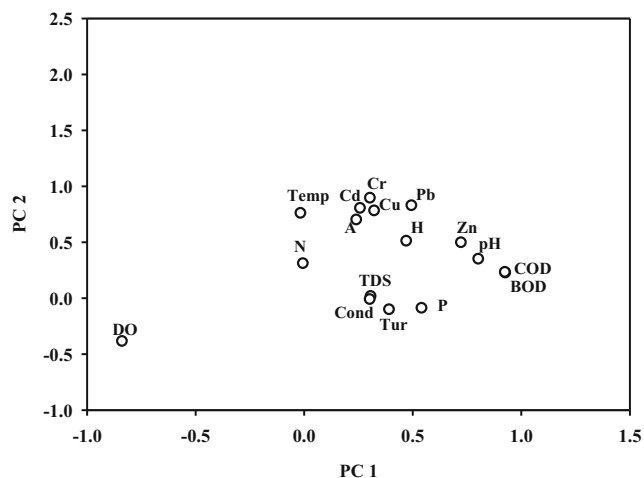


Fig. 3 Loadings of water quality parameters on PC 1 and PC 2 in agriculture land use. *DO* dissolved oxygen, *BOD* biological oxygen demand, *COD* chemical oxygen demand, *Temp* temperature, *Tur* turbidity, *P* phosphate, *N* nitrate, *A* alkalinity, *H* hardness, *TDS* total dissolved solid, *Cond* conductivity, *Cu* copper, *Zn* zinc, *Cr* chromium, *Cd* cadmium, *Pb* lead

moderate positive loading (Fig. 4). However, in the case of urban land use class, Cr, Cu, and Pb with strong positive loadings and temperature, Cd, and Zn with moderate positive loadings formed the first PC (Fig. 5). PC 2 represented positive strong loading for pH, BOD, and COD and moderate loading for Zn (Fig. 5). PC 3 explained strong positive loading for TDS and conductivity, whereas it explained negative moderate loading for turbidity (Fig. 6). The hardness formed the fourth PC with strong positive loading and phosphate and alkalinity with moderate positive loading (Fig. 6).

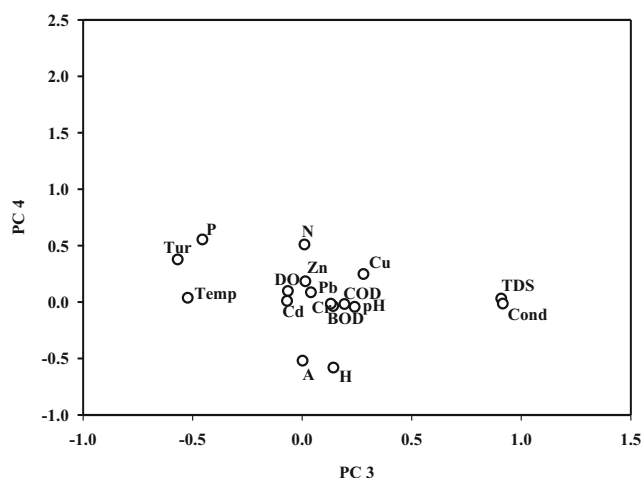


Fig. 4 Loadings of water quality parameters on PC 3 and PC 4 in agriculture land use. *DO* dissolved oxygen, *BOD* biological oxygen demand, *COD* chemical oxygen demand, *Temp* temperature, *Tur* turbidity, *P* phosphate, *N* nitrate, *A* alkalinity, *H* hardness, *TDS* total dissolved solid, *Cond* conductivity, *Cu* copper, *Zn* zinc, *Cr* chromium, *Cd* cadmium, *Pb* lead

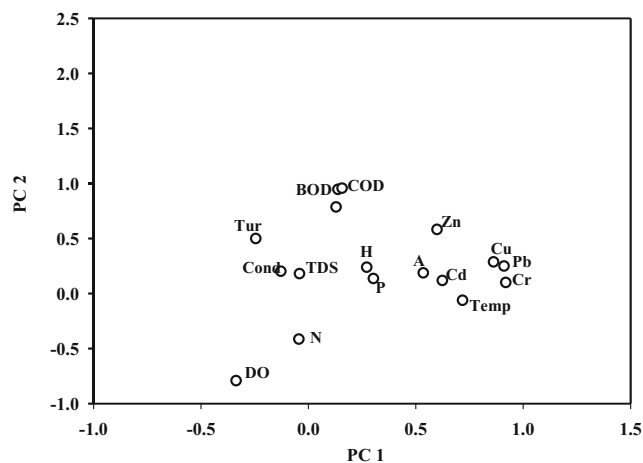


Fig. 5 Loadings of water quality parameters on PC 1 and PC 2 in urban land use. *DO* dissolved oxygen, *BOD* biological oxygen demand, *COD* chemical oxygen demand, *Temp* temperature, *Tur* turbidity, *P* phosphate, *N* nitrate, *A* alkalinity, *H* hardness, *TDS* total dissolved solid, *Cond* conductivity, *Cu* copper, *Zn* zinc, *Cr* chromium, *Cd* cadmium, *Pb* lead

Correlation between land use class and water quality parameters

For the purpose of studying the relationships between land use class and water quality, the Spearman *R* coefficient was used in accordance to non-normal distribution of data. Although in the cluster analysis zones 5 and 10 (Fig. 2) formed the cluster within the urban land use category, these zones were omitted in the correlation analysis due to their heterogeneous nature with percentage contribution of all three land uses (Table 1).

The correlation results showed that agriculture showed a significant positive correlation with turbidity, DO, and nitrate and negative correlation with hardness and COD

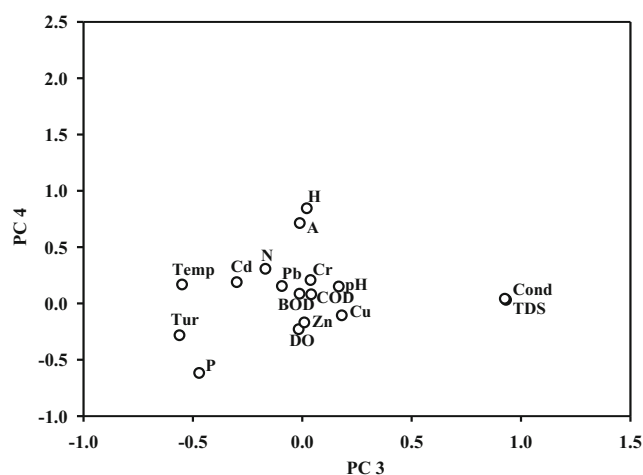


Fig. 6 Loadings of water quality parameters on PC 3 and PC 4 in urban land use. *DO* dissolved oxygen, *BOD* biological oxygen demand, *COD* chemical oxygen demand, *Temp* temperature, *Tur* turbidity, *P* phosphate, *N* nitrate, *A* alkalinity, *H* hardness, *TDS* total dissolved solid, *Cond* conductivity, *Cu* copper, *Zn* zinc, *Cr* chromium, *Cd* cadmium, *Pb* lead

Table 2 Spearman correlation coefficients between land uses and water quality parameters

Parameters	Agriculture	Urban
Temperature	0.003	0.021
TDS	−0.016	0.087
Conductivity	−0.026	0.073
Turbidity	0.156**	−0.388**
pH	−0.052	0.294**
DO	0.136*	−0.320**
BOD	−0.094	0.737**
Hardness	−0.131*	0.141*
Alkalinity	−0.054	0.062
COD	−0.211**	−0.754**
Nitrate	0.620**	0.067
Phosphate	−0.032	0.037
Cr	−0.096	0.216**
Cu	−0.078	0.325**
Cd	−0.133*	0.087
Pb	−0.088	0.277**
Zn	−0.191**	0.526**

Levels of significance: * $p < 0.05$, ** $p < 0.01$

(Table 2). However, the correlation with other parameters, viz., temperature, TDS, conductivity, pH, BOD, alkalinity, and phosphate, was non-significant (Table 2). Among metals, agriculture was negatively correlated with Zn and Cd and non-significantly with Cr, Cu, and Pb (Table 2). The urban land use showed positive correlation with pH, BOD, hardness, and metals (Cr, Cu, Pb, and Zn) (Table 2). However, urban land use was negatively correlated with DO, turbidity, and COD (Table 2). A non-significant relation of urban land use with temperature, TDS, conductivity, alkalinity, nitrate, phosphate, and Cd was observed (Table 2).

Discussion

James Prinsep is accredited with the first map of Varanasi in 1822 (<http://www.varanasi.org.in/>). The map of old Varanasi city from rivers Assi to Varuna revealed that the entire stretch was lined with concrete ghats and built-up areas with few empty spaces in between, which were later occupied forming 84 ghats (<http://www.varanasi.org.in/>; Singh 2004), and also the areas adjoining ghats showed urban built-up. The LULC map of the western bank of Ganga for the buffer radii of 1000 m revealed more or less similar land use features in zones 6, 7, 8, and 9. These zones showed variable urban land use classes such as residential areas, mixed built-up, recreational, public–semi-public, transportation, commercial areas, public utilities, and point sources. However, some vegetated areas were also

observed in zones 6 and 7 representing the marginal green cover in and around the localities. Despite the heterogeneity observed in land use features of zones 5 and 10, these zones formed a cluster with zones 6, 7, 8, and 9 (Fig. 2), which can be explained due to the percent contribution of urban land use, which was maximum at all the zones in this cluster compared to the percent contribution of the other two land uses (Table 1). On the other hand, the heterogeneity observed in LULC statistics for zones 5 and 10 can possibly be attributed to more or less similar contribution from agricultural and rural land use classes (Table 1).

Although there are no studies where LULC of 1000 m from the river bank in Varanasi was prepared, there are two change detection studies in Varanasi region by Ohri and Poonam (2012) and Jaiswal and Verma (2013). Ohri and Poonam (2012) reported 50 % change in area from agricultural land to built-up land from year 1976 to 2010, thus suggesting considerable decline in agricultural land. Further, from 2005–2006 to 2011–2012, the district has experienced a 53 % increase in built-up area and 5.3 % decline in land under agriculture (Jaiswal and Verma 2013). The present study is not a change detection study, but still the maximal percent contribution of agricultural land in the digitized stretch was comparatively lesser than urban land use (Table 1). Zone 12, which represented maximum agricultural land, also contained some portion of built-up areas (in the form of rural land use) (Table 1). Similar observations were made in zones 3, 11, and 4, thus suggesting the initiation of conversion process as reported by Ohri and Poonam (2012) and Jaiswal and Verma (2013). Zone 4, which was under the agricultural practice as observed in map of the year 1976 (Ohri and Poonam 2012), depicted a significant contribution from built-up areas in the form of urban and rural classes apart from agriculture (Table 1), thus suggesting that agricultural land near the river bank is also experiencing a shift.

The descriptive analysis of water quality data depicted the variations in physico-chemical properties of river in different zones of land use categories. These variations in river water quality can be attributed to localized sources and seasonal variations. The river flow also contributes to these variations in physico-chemical properties (Duan et al. 2013b), but was not measured during the present study. Accounting for the large variations in data, the PCA was conducted to reduce a complex data into lower dimension factors and also to extract parameters that are most important in assessing variations in river water quality (Shlens 2014; Ouyang 2005). Nowadays, PCA is also used as an important technique for source apportionment in river water quality monitoring (Juahir et al. 2011; Huang et al. 2010; Shrestha and Kazama 2007). The parameters with variable loadings in the PC was identified using Liu et al. (2003) classification for factor loadings where the factor loadings are designated as “strong,” “moderate,” and “weak” corresponding to absolute loading values of >0.75 , 0.75 – 0.50 ,

and 0.50–0.30, respectively. PCA results suggested that of 18 parameters tested, DO, BOD, COD, pH, temperature, TDS, conductivity, and metals (Cr, Cu, Cd, Pb) with high loadings are of significance for assessing the water quality of river Ganga in agriculture land use and DO, BOD, COD, pH, TDS, conductivity, hardness, and metals (Cr, Cu, Cd) in urban land use.

The loadings of parameters differed in each PCs for different land use categories (Figs. 3, 4, 5, and 6) as represented in the results. PC 1 in agricultural land use explained for the organic matter and fertilizer sources. The residues from crops grown in the area and also use of organic manures can be attributed to sources of organic matter from agricultural land. The sources of phosphate and Zn can be through fertilizers. The agricultural soils of Indo-Gangetic plains are deficient in phosphate and Zn (ICAR 2009; Prasad 2007); therefore, it necessitates providing additional doses of these elements in the form of fertilizers. However, in urban areas, PC 2 represented sources for organic matter. Sources of organic matter in urban areas are mainly from sewage outfalls bringing in domestic waste from the city. Also, in particular to Varanasi, the floral offerings and dead remains of humans and animals which find their way directly into river from Ghats can be possible contributors. Thus, this PC explains the anthropogenic pollution sources.

As this study has attempted for the first time to relate water quality of river Ganga to land uses, therefore it was difficult to correlate the land use wise PCA results with some other studies. However, results can be discussed with studies where PCA was applied on water quality data sets. Similar strong positive loadings of COD and BOD in PCs were observed in Cauvery River basin (Umamaheswari and Saravanan 2009) and river Ganga at Varanasi (Kumari and Tripathi 2014). Mishra (2010) reported negative DO and positive BOD loadings in river Ganga in Varanasi as observed in the present study. The negative loading of DO means the negative correlation with other parameters of PC, which may be due to an increase in organic matter in the river, which requires dissolved oxygen for microbes and chemicals to break down the organic matter resulting in higher BOD and COD (Singh et al. 2005).

The maximum variability in urban land use was explained by PC 1. The PC 1 with strong metal loadings suggested that zones with built-up areas were source of metals. This possibly can be due to sewage outfalls in these zones bringing in the waste from small industries and handlooms. On the other hand, in agricultural land use, PC 2 was heavily loaded with metals suggesting that agriculture is also the source of metals. In a report by Papafilippaki et al. (2008), a significant amount of metals were reported in river Keritis of Greece, with agriculture being the major land use. The use of fertilizers and biocides in agriculture contributed to the metals in the river (Papafilippaki et al. 2008). Temperature also showed strong positive loadings in PC 1 and PC 2 in urban and agricultural

land uses. The solubility and release of metals in surface waters are determined by temperature (Iwashita and Shimamura 2003). Similar results were reported by Singh et al. (2004) in river Gomti.

The third PC of both land use categories explained the sources of solids indicating towards their origin in runoff from the agricultural fields and waste disposal activities from urban areas. The third PC again represented the anthropogenic sources. The presence of ions and ionic compounds in water leads to high loading of these variables. Also, conductivity and TDS are closely related to each other as conductivity is a useful indicator of TDS because the conduction of current in an electrolyte solution is primarily dependent on the concentration of ionic species (Hayashi 2004). On the other hand, conductivity and temperature are inversely related to each other (Hayashi 2004) as also indicated by PCA results of the present study (Figs. 4 and 6). The results of this study are in agreement with the results reported by Singh et al. (2004) in Gomti River (India), Mustapha and Abdu (2012) in Jakara River (Nigeria), and Bhardwaj et al. (2010) in Choti Gandhak River (India).

The fourth PC representing the minimum variability in agricultural land use with moderate loadings explained nutrient group. However, upon looking at the loadings score, it can be deciphered that agricultural land use is not contributing much phosphate and nitrate due to utilization of these essential nutrients by crops. Varol et al. (2012) also reported similar results for Tigris River. In a study conducted in Jakara Basin, these nutrient-related factors showed minimum variability with strong loadings (Mustapha and Abdu 2012). Simeonov et al. (2003) showed that nitrate contributed primarily and phosphate secondarily in PC, thus representing influences from agricultural runoff and atmospheric deposition. In urban areas, fourth PC represented the source of minerals and ions in water. Similar results were reported by Singh et al. (2004). The hardness and alkalinity in surface water is mainly because of ions such as Mg^{+2} , Ca^{+2} , CO_3^{-2} , HCO_3^{-} , OH^{-} , BO_3^{3-} , SiO_4^{-4} , and PO_4^{-3} (APHA 1995). Urban areas may possibly contribute towards alkalinity as concrete structures provide additional weatherable materials (Stets et al. 2014; Barnes and Raymond 2009).

The correlation results suggested that with increasing urbanization, the DO content of river decreased significantly. However, an increase in land under agriculture led to a significant increase in DO content. Bu et al. (2014) also observed a significant negative correlation of DO with built-up areas and positive relation with agricultural areas in Taizi River Basin, China. In Chalakudy River basin of Kerala, DO content of river under plantation and forested land uses was higher compared to the sites of urban areas (Chattopadhyay et al. 2005). Also, increasing agricultural land use resulted in an increase in nitrate and turbidity. Poor and McDonnell (2007) also recorded higher nitrate concentration in agricultural as compared to

residential catchments. Further, the correlation results of the present study suggested significant reductions in hardness, COD, Zn, and Cd with increased agricultural land use.

The significant positive correlations of urban land use with BOD, hardness, pH, and metals suggest that the increasing urbanization near the river bank may lead to an increase in the concentrations of metals and organic matter, thus lowering the water quality. Such a positive correlation of built-up areas with BOD was also observed by Bu et al. (2014), but a negative correlation was found with pH. The correlation results further suggest that the increase in urban areas near the river bank resulted in a decline in turbidity. A significant negative correlation of urban land use with COD may be due to higher BOD, which may have consumed more oxygen compared to oxidation by chemicals. Huang et al. (2013), however, found a positive correlation between COD and built-up region.

In the present study, urban areas seem to contribute more pollutants in the river compared to agricultural land use. Sliva and Williams (2001) observed a very marked increase in chemical fluxes with increasing urban land use; however, more variability was observed in agricultural land use. In a study on comprehensive analysis of land use and water quality of Huangpu river in Shanghai from 1947 to 1996, Ren et al. (2003) recorded a decrease in water quality in highly urbanized city centers, while improvement in water quality was observed in areas where agriculture is practiced. Similarly, water quality of Wang Chhu River, in Thimphu, was better upstream as compared to urban areas (Giri and Singh 2013).

Conclusion

The results of this study clearly demonstrate that there exists a relationship between adjacent land use and water quality of river Ganga in Varanasi. The LULC map of the study stretch revealed that riparian vegetation is nearly absent in the district and there are two major land uses, viz., agriculture and urban. The land use wise PCA on water quality parameters suggests that out of 18 parameters tested, 11 parameters (DO, BOD, COD, pH, temperature, Cr, Cu, Cd, Pb, TDS, and conductivity) in agricultural land use and 10 parameters (DO, BOD, COD, pH, Cr, Cu, Pb, TDS, conductivity, and hardness) in urban land use are of prime importance in monitoring water quality of the study stretch. The maximum variability in urban land use was explained by metals and by organic matter in agricultural land use. The correlations exhibiting the strength of relationship between land use class and water quality suggested that the increase in urban land use resulted in an increase in BOD, hardness, and metals on one hand and a decrease in turbidity, DO, and COD on the other. With the increase in area under agriculture, the rise in turbidity, DO, and nitrate, and the decline in hardness, COD, Cd, and Zn were evident.

The present pioneering study provides a preliminary insight to look at river Ganga in Varanasi from a landscape perspective. Therefore, studies encompassing the entire 2525 km of river Ganga may be undertaken to understand the relationship between land use and water quality, which may provide a holistic approach to “Clean Ganga Mission.”

Acknowledgments The authors are thankful to the Head of the Department of Botany, Banaras Hindu University and Director of the Indian Institute of Remote Sensing for providing necessary facilities to carry out the research work. The DST-INSPIRE fellowship, Government of India is also acknowledged for financial support. The authors also acknowledge the anonymous reviewers for their valuable comments and suggestions which helped in improving the manuscript.

References

- Allan JD, Erickson DL, Fay J (1997) The influence of catchment land use on stream integrity across multiple spatial scales. *Freshw Biol* 37(1): 149–161
- American Public Health Association (APHA), American Water Works Association (AWWA), and Water Environment Federation (WEF) (1995) Standard methods for the examination of water and wastewater 19th edition. United Book Press, Inc., Baltimore, Maryland
- Arumugam N, Fatimah MA, Chiew EFC, Zainalabidin M (2010) Supply chain analysis of fresh fruits and vegetables (FFV): prospects of contract farming. *Agr Econ* 56(9):435–442
- Barnes RT, Raymond PA (2009) The contribution of agricultural and urban activities to inorganic carbon fluxes within temperate watersheds. *Chem Geol* 266(3):318–327
- Bhardwaj V, Singh DS, Singh AK (2010) Water quality of the Chhoti Gandak River using principal component analysis, Ganga plain, India. *J Earth Syst Sc* 119(1):117–127
- Bu H, Meng W, Zhang Y, Wan J (2014) Relationships between land use patterns and water quality in the Taizi River basin, China. *Ecol Indic* 41:187–197
- Chattopadhyay S, Rani LA, Sangeetha PV (2005) Water quality variations as linked to landuse pattern: a case study in Chalakudy river basin, Kerala. *Curr Sci* 89(12):2163–2189
- Duan W, He B, Takara K, Luo P, Nover D, Sahu N, Yamashiki Y (2013a) Spatiotemporal evaluation of water quality incidents in Japan between 1996 and 2007. *Chemosphere* 93(6):946–953
- Duan W, Takara K, He B, Luo PP, Nover D, Yamashiki Y (2013b) Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari River, Japan, 1985 to 2010. *Sci Total Environ* 461:499–508
- Duan W, He B, Takara K, Luo PP, Nover D, Hu MC (2015) Modeling suspended sediment sources and transport in the Ishikari River basin, Japan, using SPARROW. *Hydrol Earth Syst Sc* 19(3):1293–1306
- Duan W, He B, Nover D, Yang G, Chen W, Meng H, Liu C (2016) Water quality assessment and pollution source identification of the eastern Poyang Lake Basin using multivariate statistical methods. *Sustainability* 8(2):1–15
- Eknesh P (2013) Ecohydrological impacts of climate and land use changes on watershed systems: a multi-scale assessment for policy. Dissertation, University of Massachusetts, Amherst Scholar Works @ UMass Amherst, paper 789
- Giri N, Singh OP (2013) Urban growth and water quality in Thimphu, Bhutan. *J Urban Environ Eng* 7(1):82–95

- Griffith JA (2002) Geographic techniques and recent applications of remote sensing to landscape-water quality studies. *Water Air Soil Poll* 138(1–4):181–197
- Hashmi F (2013) Correlation between riparian buffers and water quality in North Carolina watersheds. Doctoral dissertation, Duke University
- Hayashi M (2004) Temperature-electrical conductivity relation of water for environmental monitoring and geophysical data inversion. *Environ Monit Assess* 96(1–3):119–128
- Huang F, Wang X, Lou L, Zhou Z, Wu J (2010) Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. *Water Res* 44(5):1562–1572
- Huang J, Zhan J, Yan H, Wu F, Deng X (2013) Evaluation of the impacts of land use on water quality: a case study in the Chaohu lake basin. *Sci World J* 2013:1–7
- Ierodiaconou D, Laurenson L, Leblanc M, Stagnitti F, Duff G, Salzman S, Versace V (2005) The consequences of land use change on nutrient exports: a regional scale assessment in south-West Victoria, Australia. *J Environ Manag* 74(4):305–316
- Indian Council of Agricultural Research (ICAR) (2009) A Science and Technology Newsletter 15(2). www.icar.org.in
- Iwashita M, Shimamura T (2003) Long-term variations in dissolved trace elements in the Sagami River and its tributaries (upstream area), Japan. *Sci Total Environ* 312(1):167–179
- Jackson DA (1993) Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology* 74(8):2204–2214
- Jaiswal JK, Verma N (2013) Land use change detection in Baragaon block, Varanasi District using remote sensing. *Int J Eng Sci Innov* 2(7):49–53
- Juahir H, Zain SM, Yusoff MK, Hanidza TT, Armi AM, Toriman ME, Mokhtar M (2011) Spatial water quality assessment of Langat River basin (Malaysia) using environmetric techniques. *Environ Monit Assess* 173(1–4):625–641
- Kaiser HF (1974) An index of factorial simplicity. *Psychometrika* 39(1):31–36
- Kumari M, Tripathi BD (2014) Source apportionment of wastewater pollutants using multivariate analyses. *B Environ Contam Tox* 93(1):19–24
- Lattin JM, Carroll JD, Green PE (2003) Analyzing multivariate data. Thomson Brooks/Cole, Pacific Grove, CA
- 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):77–89
- Manage AB, Scariano SM (2013) An introductory application of principal components to cricket data. *J Stat Educ* 21(3):1–22
- McDonald JH (2009) Handbook of biological statistics. Sparky House Publishing, Baltimore, MD
- Mckenna J (2003) An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environ Model Softw* 18(3):205–220
- Mishra A (2010) Assessment of water quality using principal component analysis: a case study of the river Ganges. *J Water Chem Techno* 32(4):227–234
- Mooi E, Sarstedt M (2011) A concise guide to market research—the process, data and methods using IBM SPSS statistics. Springer, Heidelberg Dordrecht, London
- Mustapha A, Abdu A (2012) Application of principal component analysis & multiple regression models in surface water quality assessment. *J Environ Earth Sci* 2(2):16–23
- National Remote Sensing Agency (NRSA) (2005) National Resources census, Department of Space, Government of India
- National River Conservation Directorate, Ministry of Environment and Forests, Government of India (2009) Status paper on River Ganga—state of environment and water quality. Report, Alternate Hydro Energy Centre, Indian Institute of Technology, Roorkee 1–31
- Ohri A, Poonam (2012) Urban sprawl mapping and land use change detection using remote sensing and GIS. *International Journal of Remote Sensing and GIS* 1(1):12–25
- Osborne JW (2015) What is rotating in exploratory factor analysis? *Pract Assess Res Eval* 20(2):1–7
- Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. *Water Res* 39(12):2621–2635
- Papafilippaki AK, Kotti ME, Stavroulakis GG (2008) Seasonal variations in dissolved heavy metals in the Keritis River, Chania, Greece. *Global nest. The Int J* 10(3):320–325
- Perry J, Vanderklein E (1996) Water quality management of a natural resource. Blackwell Science, Cambridge, Massachusetts, pp 639
- Poor CJ, McDonnell JJ (2007) The effects of land use on stream nitrate dynamics. *J Hydrol* 332(1):54–68
- Prasad R (2007) Phosphorus management in the rice-wheat cropping system of the Indo-Gangetic plains. *Better Crops—India*:8–11
- Raj PN, Azeed PA (2010) Land use and land cover changes in a tropical river basin: a case from Bharathapuzha River basin, southern India. *J Geogr Inf Sys* 2(04):185
- Ravindra K, Ameen M, Monika R, Kaushik A (2003) Seasonal variations in physico-chemical characteristics of river Yamuna in Haryana and its ecological best-designated use. *J Environ Monitor* 5(3):419–426
- Ren W, Zhong Y, Meligrana J, Anderson B, Watt WE, Chen J, Leung HL (2003) Urbanization, land use, and water quality in Shanghai: 1947–1996. *Environ Int* 29(5):649–659
- Rothwell JJ, Dise NB, Taylor KG, Allott TEH, Scholefield P, Davies H, Neal C (2010) Predicting river water quality across north West England using catchment characteristics. *J Hydrol* 395(3):153–162
- Sharma S, Agrawal M (2014) Thwarting changing land use—a step towards Ganga rejuvenation. *Curr Sci* 107(2):174
- Shlens J (2014) A tutorial on principal component analysis. arXiv preprint arXiv:1404.1100
- 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 A, Anthemidis SM, Kouimtzis T (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(18):3980–3992
- Singh RP (2004) Cultural landscapes and the lifeworld: literary images of Banaras. Pilgrimage and cosmology. Indica Book, Varanasi
- 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(1):355–374
- Sliva L, Williams DD (2001) Buffer zone versus whole catchment approaches to studying land use impact on river water quality. *Water Res* 35(14):3462–3472
- Soman K, Mahamaya C, Ouseph PP (1997) Status of riverine pollution in south Kerala and its relation to physiography and landuse. In Proc. IX Kerala Science Congress, Thiruvananthapuram, pp 93–95
- Stets EG, Kelly VJ, Crawford CG (2014) Long-term trends in alkalinity in large rivers of the conterminous US in relation to acidification, agriculture, and hydrologic modification. *Sci Total Environ* 488:280–289
- Strobl RO, Robillard PD (2008) Network design for water quality monitoring of surface freshwaters: a review. *J Environ Manag* 87(4):639–648
- Tiwari TN, Ali (1988) Water quality index for Indian rivers. In: Trivedy RK (ed) Ecology and pollution of Indian rivers, 1st edn. Ashish Publishing House, New Delhi, pp. 271–286

- Umamaheswari S, Saravanan NA (2009) Water quality of Cauvery River basin in Trichirappalli, India. *International Journal of Lakes and Rivers* 2(1):1–10
- Varol M, Gökot B, Bekleyen A, Şen B (2012) Water quality assessment and apportionment of pollution sources of Tigris River (Turkey) using multivariate statistical techniques—a case study. *River Res Appl* 28(9):1428–1438
- Vega M, Pardo R, Barrado E, Debán L (1998) Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res* 32(12):3581–3592
- Ward JV (1989) The four-dimensional nature of lotic ecosystems. *J N Am Benthol Soc* 8(1):2–8
- Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. *J Amer Statist Assoc* 58(301):236–244
- Yaari Y (1997) Segmentation of expository texts by hierarchical agglomerative clustering. arXiv preprint [cmp-lg/9709015](https://arxiv.org/abs/cmp-lg/9709015)
- Zhang Q, Xu CY, Tao H, Jiang T, Chen YD (2010) Climate changes and their impacts on water resources in the arid regions: a case study of the Tarim River basin, China. *Stoch Env Res Risk A* 24(3):349–358
- Zheng PQ, Baetz BW (1999) GIS-based analysis of development options from a hydrology perspective. *J Urban Plan D* 125(4):164–180
- <http://varanasi.nic.in/>. Accessed date 31 Dec 2015
- <http://www.varanasi.org.in/>. Accessed date 2 Jan 2016
- <http://www.businesstoday.in/>. Accessed date 1 Jun 2016
- <http://india-wris.nrsc.gov.in/>. Accessed date 1 Jun 2016