



Pollution mapping in the urban segment of a tropical river: is water quality index (WQI) enough for a nutrient-polluted river?

Kashifa Iqbal¹ · Shamshad Ahmad^{1,2} · Venkatesh Dutta¹

Received: 25 April 2019 / Accepted: 11 October 2019 / Published online: 25 October 2019
© The Author(s) 2019

Abstract

Most of the tropical rivers of the world are being affected by multiple sources of pollution. The intensity of pollution is much bigger in the urban stretches due to discharge of untreated or partially treated sewage. A rapid and cost-effective tool is required for identification of water quality problems and their spatial variation for determining the main pollution sources and to detect relationships between various parameters. For this study, Gomti River, a major tributary of River Ganges, India, was considered which has gained substantial attention because of increasing anthropogenic pollution loads that has badly affected its water quality and ecosystem functions. The urban segment is polluted with organic substances, nutrients and heavy metals. The study provides an overview of the quality of water in the Gomti River through water quality index (WQI) and multivariate statistical techniques to check if WQI is enough for a nutrient-polluted river in the urban stretch. The study suggests that periodic monitoring and the water quality index development are not enough as it does not incorporate all the aspect of a rivers water quality. The separate assessment of nitrogenous biochemical oxygen demand, carbonaceous biochemical oxygen demand, sediment oxygen demand and the nitrification inhibition aspects are required to be integrated when developing a WQI. Present study illustrates that water quality of Gomti River has gradually worsened from upstream and downstream to middle stretch. The middle stretch was found to be most polluted as the major drains are concentrated within this stretch. Principal component analysis/factor analysis (PCA/FA) helped in obtaining and recognizing the factors/sources accountable for river water quality differences in the study area. The findings are useful for the decisions regarding water quality management and this can also be applied for speedy and low-cost assessment of water quality of the polluted urban stretch of other tropical rivers for better environmental management and planning perspective.

Keywords Gomti River · Multivariate statistical analysis · Water quality index · Anthropogenic pollution · Contamination · Spatio-temporal

Introduction

Rivers are a vital fresh water source for human consumption, irrigation, industrial purposes and ecosystem services. Temporal and spatial information about the water quality of river and its continuous improvement is essential for the management of these fresh water resources effectively. The decline in river water quality in Indian rivers results occasionally from natural processes (during flood pulses), but more profoundly due to anthropogenic activities like discharge of industrial effluents, domestic wastewater and municipal sewage, besides non-point agricultural drainage. Nonetheless, the industrial, domestic wastewater drainage as well as the agricultural run-off from the fields contributes to the enormous proportion of rivers' pollution (Dutta et al. 2018). From a study by Akbarzadeh et al. (2018), the

✉ Venkatesh Dutta
dvenks@gmail.com

Kashifa Iqbal
kaishsiddiqui786@gmail.com

Shamshad Ahmad
shamshadahmad93@gmail.com

¹ Department of Environmental Science (DES), Babasaheb Bhimrao Ambedkar University, Raebareli Road, Vidya Vihar, Lucknow, Uttar Pradesh 226025, India

² Water Technology and Management Division, National Environmental Engineering Research Institute (NEERI), Nagpur 440020, India

untreated wastewater discharged from wastewater treatment plant (Seine Aval) of Paris municipality into lower Seine River, the dissolved oxygen level dropped significantly in the downstream at 100 km segment (Aissa-Grouz et al. 2015). In China, regulating the sewage release and run-off from urban areas is crucial in recovery of the water quality of the Pearl River estuary (He et al. 2014). Natural processes such as temperature and precipitation along with anthropogenic factors affect the river water quality and steer to distinct attributes between seasons (Kumar et al. 2015). Hence, for an integrated water resources management, continuous monitoring and evaluation of river water quality is a prerequisite (Wang et al. 2013). For monitoring the quality of river water as well as the visualization of the rivers' current condition (locally and temporally), the sampling networks appear as an exceptional provenance of information. Besides giving a general outline of the conditions at temporal scale, these sampling networks also provide enough evidence of the ecosystem progression at seasonal and geographical scale (Barakat et al. 2016).

For an effective management of the surface water bodies, there are certain techniques like modelling approaches and monitoring programs for identification of the water pollution to better check the regulation contraventions and also to empower the authorities for formulating some steadfast decisions in the form of guidelines and regulations for managing the water quality as well as quantity. Within the modelling approaches, water balance, pollutant transport and the water quality are studied comprehensively for an interactive model development (Lorenz et al. 2014). The monitoring approach involves the measurement of the water quality of the surface water body for an extended period of time. Such monitoring programs ends up giving an enormous and vague dataset encompassing huge physico-chemical parameter data, which sometimes becomes problematic for analysing and comprehending because of complexity. A lot of accessible indicators are present to assess the aquatic ecosystem quality and evolution due to various human-induced influences and pressures for extended period of time. However, little attention has been given to the development of some specific tools for rapid assessment of the water quality (Sorche et al. 2019). The surface water monitoring requires certain indicators and descriptors that incorporate the ecosystem specific features and processes ongoing within a system. Also, special consideration should be given to the ecosystem condition when susceptible to some disturbances like pollution (organic, inorganic) from nutrients like nitrates, phosphates, from heavy metals (As, Pb, Cu, etc.) and other toxic constituents like pesticides.

Water resources monitoring for the quality assessment is necessarily important for assessing its worth for healthy ecosystem, hygienic environment and for domestic and other uses like recreation, agriculture, aquaculture, etc. Assessing

all the water quality parameters of a water body is a complicated task, and the bulky data so obtained after the analysis of all the samples and comparing it with the existing guidelines are uneasy to handle. So, the WQI system has been developed to simplify the impression of water quality status of a water body.

The only plus point of the WQI development is that they give us comprehensive idea of the water quality of the studied area. There are existing examples of the various internationally developed water quality indices like the US National Sanitation Foundation Water Quality Index (NSFWQI), British Columbia Water Quality Index (BCWQI), Oregon Water Quality Index (OWQI) and Canadian Council of Ministers of the Environment Water Quality Index (CCMEWQI) and various others existing and being added. These above-mentioned water quality indices meet all the criteria of an all-inclusive good index as they integrate all the parameters that could be studied for an aquatic ecosystem. WQIs are quite user-friendly and could be easily handled in computational tools (Shrestha and Basnet 2018). The principle of WQI has been revised to some extent to advance existing concept for certain other forms of water quality index development (Singh et al. 2018) for reflecting the water quality conditions in various environmental circumstances having only difference with the development of sub-index and aggregation function. There are several other indices like the Transitional Water Quality Index (TWQI) which integrates six variables like the dissolved oxygen, dissolved inorganic nitrogen and phosphorus, phytoplankton chlorophyll-a, benthic phanerogams and macroalgal species. A Groundwater Quality Index (GWQI) was developed by formulating the eight principal parameters of drinking water which are K^+ , Na^+ , Ca^{2+} , Mg^{2+} , SO_4^{2-} , Cl^- , pH and total dissolved solid (TDS) to monitor the groundwater quality in Qazvin plateau area of Iran (Saeedi et al. 2010). Twenty-two water quality parameters were measured for the formulation of WQI in a study carried out in Varanasi, India (Chaurasia et al. 2018).

Physical, chemical and biological parameters can also be helpful for assessing the water quality of any particular area or specific source. If the values of these parameters exceed the limits, it becomes harmful for human health. Similarly, to describe the water quality, WQI is the most efficacious approach used to define the suitability of water sources for human consumption is water quality index (WQI). For the water quality assessment of different water resources, groundwater and surface water, predominantly rivers, WQI has been extensively applied. WQI has done an imperative work in management of water resources (Sutadian et al. 2016; Mohebbi et al. 2013). The multiple environmental parameters are effectively combined and converted into one value which reveals the water quality status. Consequently, in spite of measuring the results of the various parameters repeatedly, WQI syndicates and gives cohesive information

about the overall water quality thus proving to be a potent approach for water quality assessment and management (Wu et al. 2018).

The prevailing water quality of any particular region varies both spatially and temporally depending upon the kind of daily routines and practices adopted by the population of the region, the occupation opted by the people of that region and various other influencing factors like the meteorological and bio-geophysical factors. Hence, water quality is greatly influenced by the combined effects of the above-mentioned factors and these play a major role in its assessment. With the close observation of all these factors that may change with time except to some extent the bio-geophysical factors, it becomes profoundly problematic to develop an index that is unanimously accepted and suitable for world-wide regions. Therefore, it becomes utmost important to develop an index that is specific to a particular region at a point of time. Also, long term (i.e. years) assessing and analysing the various parameters of a water body of a particular region is quite tedious work requiring a lot of labour and giving an index that is developed after years of analysis. Thus, the developed index gives the condition of long-term data though the situation of water body at present may not be same.

A systematic and continuous monitoring of the parameters is required so as to detect the fluctuations in the concentration of the physiochemical parameters. A rapid assessment of some selected parameters is important for WQI development of any region so that one could give a quick overview of the prevailing situation of a water body and convey the alarming situation to the government, policy makers and the general public in a simple and understandable manner as the local population are the real stakeholders. In the present study, a short-term assessment of the water quality is carried out for deciphering the condition of Gomti River in the stretch of Lucknow city by developing a rapid WQI to analyse and interpret an enormous dataset of Gomti River in Lucknow into useful information for its proper management.

Multivariate statistical approaches are the most apposite and extensively used methods for processing and examining the data obtained over time. It has the ability to treat bulk of spatial and temporal data from different monitoring sites into comprehensive information. Different statistical techniques like principal component analysis (PCA), cluster analysis (CA) and factor analysis (FA) are very useful for riverine studies as these have the ability to evaluate spatial as well as temporal river quality variations and have the aptitude to identify the potential water contamination sources (Sabia et al. 2018; Alilou et al. 2018; Li et al. 2017; Iqbal et al. 2017; Villas-Boas et al. 2017; Zhao et al. 2016).

Principal component analysis and cluster analysis were used for carrying out a study on the Thamirabarani river basin situated in South India related to its hydrochemistry (Kumarasamy et al. 2014). Khan et al. (2016) applied

correlation analysis, PCA and CA components for identifying possible pollution causes, studying the seasonal variations and clustering of monitoring places of Yamuna and Ganga Rivers in Uttarakhand State of India. These statistical techniques have previously been used and approved by different workers at different regions/rivers systems (Diamantini et al. 2018; Nguyen et al. 2017; Ojok et al. 2017; Chowdhury et al. 2016). Statistical approaches like PCA and CA methods were used for classifying the sampling sites and identifying the underlying pollution source (Najar et al. 2017; Xue et al. 2017; López-Morales and Mesa-Jurado 2017). The above studies showed that for determining the fundamental relationships amongst various parameters related to water quality, identifying the pollution radices and grouping identical sampling sites into classes with analogous attributes, PCA and CA statistical methods are very important tools (Sharma et al. 2015).

To accomplish this goal, CA, PCA, Pearson's correlation matrix along with water quality index were used to (1) ascertain the similitude as well as dissimilitude amongst the studied sampling sites and associated physiochemical parameters, (2) evaluate the contribution of affected parameters of water quality for temporal and spatial discrepancies in quality at the selected sites, and (3) identify the contaminant affecting water quality and their potential sources, and (4) develop a rapid water quality index for assessment of urban stretch of the river.

Materials and methods

Description of the study area

Gomti River, a groundwater as well as rainfall nourished surface body and also a tributary of the Ganges River system having paramount importance in the plains of Northern India is chosen for the present research. This alluvial river emerges from a lake (Fulhar-Jheel) present naturally within a forest (at an elevation of about 200 m; 28° 13' 40" N and 80° 10' 70" E) adjacent to Pilibhit in Uttar Pradesh province of the Indo-Gangetic Plain, around 50 km southwards of the foothills of the lower Himalayas. This perennial river that flows across the central and eastern part of Uttar Pradesh (U.P) criss-crosses a distance of roughly 960 km before it finally confluences with the Ganges River close to Kaithighat, Varanasi (Dutta et al. 2011) (Fig. 1). The river, with an approx. length of 33 kms in the stretch of Lucknow city, functions as a chief domestic water supply source to the city which is the capital of Uttar Pradesh. Consequently, the Gomti river gets back the domestic wastewater, municipal sewage, agricultural run-off and effluents from the industries (distilleries, sugar



Fig. 1 Study area with location of sampling sites of Gomti River in Lucknow City, Uttar Pradesh, India

mills, agrochemicals, paper and pulp and many others) which is left untreated and discharged directly throughout its course from the Lucknow city along with towns of Jagdishpur, Sultanpur and Jaunpur (Dutta et al. 2011; Singh et al. 2004). Besides washerman and crematoriums, there are many other activities causing pollution in Gomti. Except during the monsoon season, the river is marked by listless flow round the year (Gupta et al. 2014).

Sampling and monitoring of pollutants

The present study was done during the months of January to April to assess the water quality of Gomti River and obtain information of various physico-chemical parameters such as dissolved oxygen (DO), biochemical oxygen demand (BOD), nitrate (NO_3^-), chemical oxygen demand (COD), pH, total dissolved solids (TDS), chloride (Cl^-), electrical conductivity (EC), ammonia (NH_4^+), nitrite (NO_2^-), phosphate

Table 1 Physiochemical parameters of water with their units and methods of measurement

Parameters	Abbreviation	Scale	Methods
pH	pH	–	pH-meter
Electrical conductivity	EC	S/m	EC-meter
Total dissolved solids	TDS	mg/L	TDS-meter
Total hardness	T-Hard	mg/L	EDTA titrimetric method
Alkalinity	Alk	mg/L	H ₂ SO ₄ titrimetric method
Chlorides	Cl ⁻	mg/L	Spectrophotometric
Nitrate	NO ₃ ⁻	mg/L	Spectrophotometric
Ammonia	NH ₄ ⁺	mg/L	Spectrophotometric
Nitrite	NO ₂ ⁻	mg/L	Spectrophotometric
Phosphate	PO ₄ ⁻³	mg/L	Spectrophotometric
Sulphate	SO ₄ ⁻²	mg/L	Spectrophotometric
Dissolved oxygen	DO	mg/L	Winkler–azide method
Biological oxygen demand	BOD	mg/L	Non seeded dilution method
Chemical oxygen demand	COD	mg/L	Potassium dichromate method
Heavy metals (Pb, As, Cd, Cr)	HMs	Ppm	Inductive coupled plasma–mass spectroscopy (ICP-MS)

(PO₄⁻³), sulphate (SO₄⁻²), alkalinity (Alk.), hardness (Hard) and heavy metals (Pb, As, Cd, Cr) (Table 1). The sampling sites were chosen owing to their contamination vulnerability, specifically, such sites in proximity with the place where the stream enters in Lucknow were thought less polluted with civic contamination whilst sites in the mid-stretch of the city were thought to be more impaired by the contamination. Seven sites, namely NH-230 (Ghaila Pul), Pakka Pul, KGMC, Hanuman Setu, Birbal Sahani, Lohia path, NH-27 were chosen for studying the river quality (Fig. 2). The first site (NH-230), upstream of the Lucknow city is situated in reasonably low-pollution zone. Other four sites (Pakka Pul, KGMC, Hanuman Setu, and Birbal Sahani) are in the mid-stretch situated in the high pollution zone because there are several drains of wastewater. The last two sites (Lohia path, NH-27), in the downstream zone have moderate pollution levels because the river substantially recuperates in its progression. The sampling approach was such designed to understand the widespread pollution causes at important sites which characterize the river water quality by considering contributions from various pollution sources that influences the quality of water (Dutta et al. 2018).

Statistical analysis

Reliable assessment of the dataset of numerous variables scrutinized in various month periods at different sampling sites was done using multivariate approach thereby performing Pearson's correlation, principal component analysis/factor analysis (PCA/FA) and cluster analysis (CA) and ultimately a water quality index development was made. Statistical and computational analyses were performed

using Microsoft Excel 2010 with XLSTAT 2015.1 and IBM SPSS 20 software.

Cluster analysis

In order to investigate the similitude of variation in water quality propensities between the sampling sites, a cluster analysis was done. A class of multivariate methods, the foremost resolution of which is to accumulate articles owing to the characteristics retained by them, is cluster analysis. Cluster analysis categorizes entities, so that every entity is alike to the other entity in the cluster with reverence to a preset selection standard. The resultant groups of entity then display strong internal homogeneity and strong external heterogeneity. The most usual method, i.e. hierarchical agglomerative clustering which offers inherent similarity associations amongst the sample and the whole dataset is usually demonstrated with a tree diagram called dendrogram.

Cluster analysis considered as a multivariate analysis technique classifies the entire dataset into various distinct categories. By virtue of the similitudes or extent amongst the various observed values for parent groups, this is further branched out into hierarchical and non-hierarchical techniques. In this study, the squared Euclidean distance technique was used that assembles the squares of the difference of each of the variables, thereby calculating the extent between the target clusters. Ward method has been applied for the cluster combination method (Majkic-Dursun et al. 2018; Nadiri et al. 2018).



Fig. 2 Sampling sites of Gomti River in Lucknow City

Principal components analysis (PCA)

PCA is a method that recognizes the patterns amongst the various related variables. PCA transforms the bulk of inter-correlated variables by transforming them into independent variables of comparatively smaller size, thus interpreting the variance amongst them (Abdi and Williams 2010). It gives knowledge on the important parameters that is used to explain the complete dataset, minimized data and to encapsulate the relationship between water constituents with the slightest harm to the primary information. PCA is applied on a correlation matrix of reorganized data to describe the arrangement of the essential dataset. This is used to ascertain the hidden and obscured pollution sources (Zhao et al. 2012).

In this work, principal components analysis was employed to outline the statistical relationship between the parameters of water quality. The calculations were done on correlation

matrix of the reorganized chemical constituents. The PCA scores were acquired from the uniform methodical data.

Development of water quality index (WQI)

WQI is a beneficial and exceptional assessment to portray the level of water quality through one index (Kansal 2018). The WQI calculation was done by the weighted arithmetic index method using equation:

$$WQI = \sum_{i=1}^n Q_i W_i \quad (1)$$

Q_i ; sub-index for i th water quality parameter, W_i is the concentration of i th water quality parameter's and n is the number of water quality parameters. The ratings of water quality index were given as excellent (100–95), good (94–80); fair

(79–65) marginal (64–45); poor (44–0) by Singh and Saxena (2018).

Results and discussions

Spatio-temporal variations of Gomti River

Box and whisker plots are used to explain the spatio-temporal discrepancies of the water quality parameters (Fig. 3). In the boxplots, large whiskers and outliers specify the extensive deviation in parameter's concentration. Due to great variation in low or high values, the mean concentrations of the distinct parameters are considerably influenced (Dabrowski et al. 2018). Hence, median is used in boxplots to specify the central tendency of the dataset in the diagram. The samples of water of the four different sampling months, i.e. January, February, March and April of seven sites on the river was studied. The EC of the river varied from (240–822 S/m). However, in the month of April, the EC was very low (256 S/m) in NH 230 while a very high value (822 S/m) was found in the Birbal Sahani site. The pH was moderate at all sites except KGMC which was very low (6.22) in January. NH 230 showed a very high pH in the months of February (8.6) and April (8.57) which exceeded the range of pH in those months. The TDS was found maximum in the month of April while at the same time the TDS of the NH 230 was very low in comparison to the rest of the sites in all the months. The NO_3^- , NO_2^- and NH_4^+ concentrations were higher in April.

DO decreased in April at all sites comparatively than the other months while BOD increased in April which clearly indicates that as temperature increases, BOD increases and the dissolved oxygen concentration decreases. BOD is inversely proportional to DO in respect with temperature (Diamantini et al. 2018). The Cl^- concentration increased in the month of April but was found low (11.66 mg/L) at NH 230 shown in Table 2. The heavy metals Pb, As, Cd, Cr showed a comparatively high concentration in April. Because of a momentous increase in temperature, most of the water quality parameters showed a substantial rise in the month of April. This could be the reason of the wide interquartile ranges and outskirts in the boxplot diagrams.

Correlation

Correlation coefficient is a statistical tool that is used to define the extent of dependence of one variable on the other and to quantify the association amongst the two variables (Bajpayee et al. 2012). Correlation loadings measure the extent of familiarity amongst the variables and principal components (Furlan et al. 2019). The major loadings can be either negative or positive which indicates the implication of

the magnitudes. The negative loading specifies that the contribution of the variables decreases with the decreasing loading in magnitude; and positive loading indicates an increase (Tashtoush and Al-Subh 2015). To illustrate the extent of dependence of a variable on the other, correlation coefficient has been applied. Before statistical analysis, the criterion given by Little and Rubin (2019) was used to standardize the data. Both EC and TDS are strongly correlated with each other. It indicates presence of pollution sources by human activities like run-off from agricultural or brick kilns fields into the river. Moderate correlation was observed NO_3^- with BOD, COD and DO and NH_4^+ correlate with PO_4^{3-} and SO_4^{2-} which indicate degradation of organics (Table 3). Such kind of correlation, amid the parameters, is indicative of organic matter decomposition happening in the river along with the industrial discharge.

Cluster analysis

To identify the clusters with spatial similarity amongst the sampling sites, cluster analysis was done into the dataset. CA categorized the sampling sites in two separate clusters, Cluster 1 and Cluster 2 (Fig. 4).

Cluster 1 was formed by sites Lohia Path, NH 27, Pakka Pul, NH 230, while cluster 2 comprises of KGMC, Hanuman Setu and Birbal Sahani. Cluster 1 corresponds to low-pollution comprising of the sites from extreme upstream and downstream of Gomti River in Lucknow and cluster 2 corresponding to high pollution comprises of the sites in the mid-stretch of Gomti River in Lucknow. Moreover, a deep analysis of cluster 1 revealed that Lohia Path, NH 27 and Pakka Pul had similar features compared to NH 230. In cluster 2, KGMC, Hanuman Setu and Birbal Sahani were grouped having similar characteristics. Polluted water sample in Cluster 1 exhibited far superior physical and chemical quality than the polluted water sample in cluster 2. The two clusters were developed including the sampling sites having similar natural characteristics and contaminated by alike pollution sources. The river water quality variances amongst the two clusters revealed variance in types and level of anthropogenic pollution loads. The enhanced quality of water at NH 27 and Lohia Path (downstream) indicated the rivers' self-purification ability and adjustability (Sim and Tai 2018). The worsening of water quality in mid-stretch at KGMC, Hanuman Setu, and Birbal Sahani revealed significant severe pollution due to domestic wastewater, agricultural runoff, municipal solid waste discharge and effluents discharge by industries situated nearby the river bank. The results of the study indicated that two different types of water qualities were existent in the area where river is flowing. It inferred that one sampling site from each cluster is enough and may perhaps assist well in spatial evaluation of the quality of water of the particular area. This could be used for a speedy

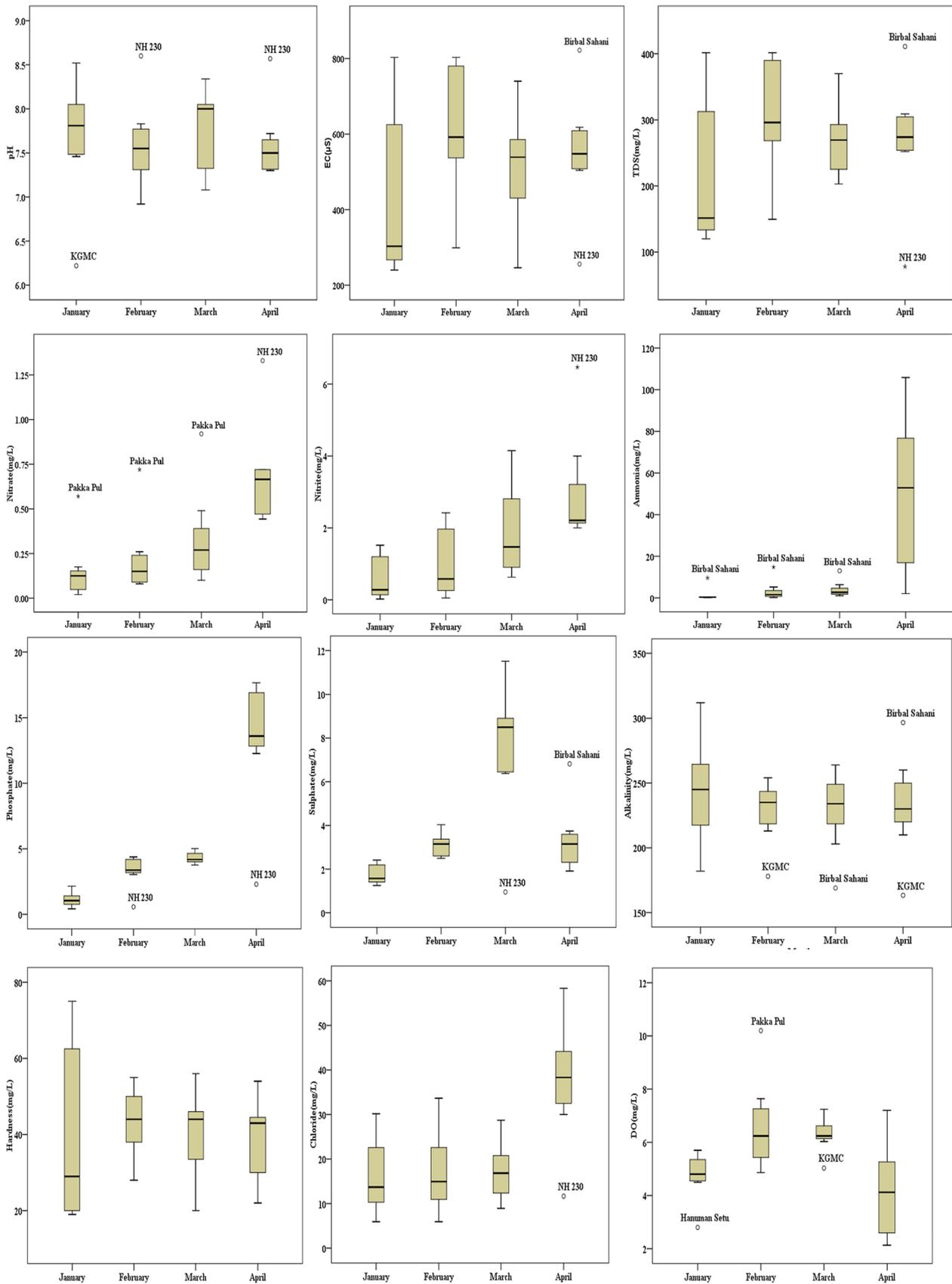


Fig. 3 Box and whisker plots of the physiochemical parameters

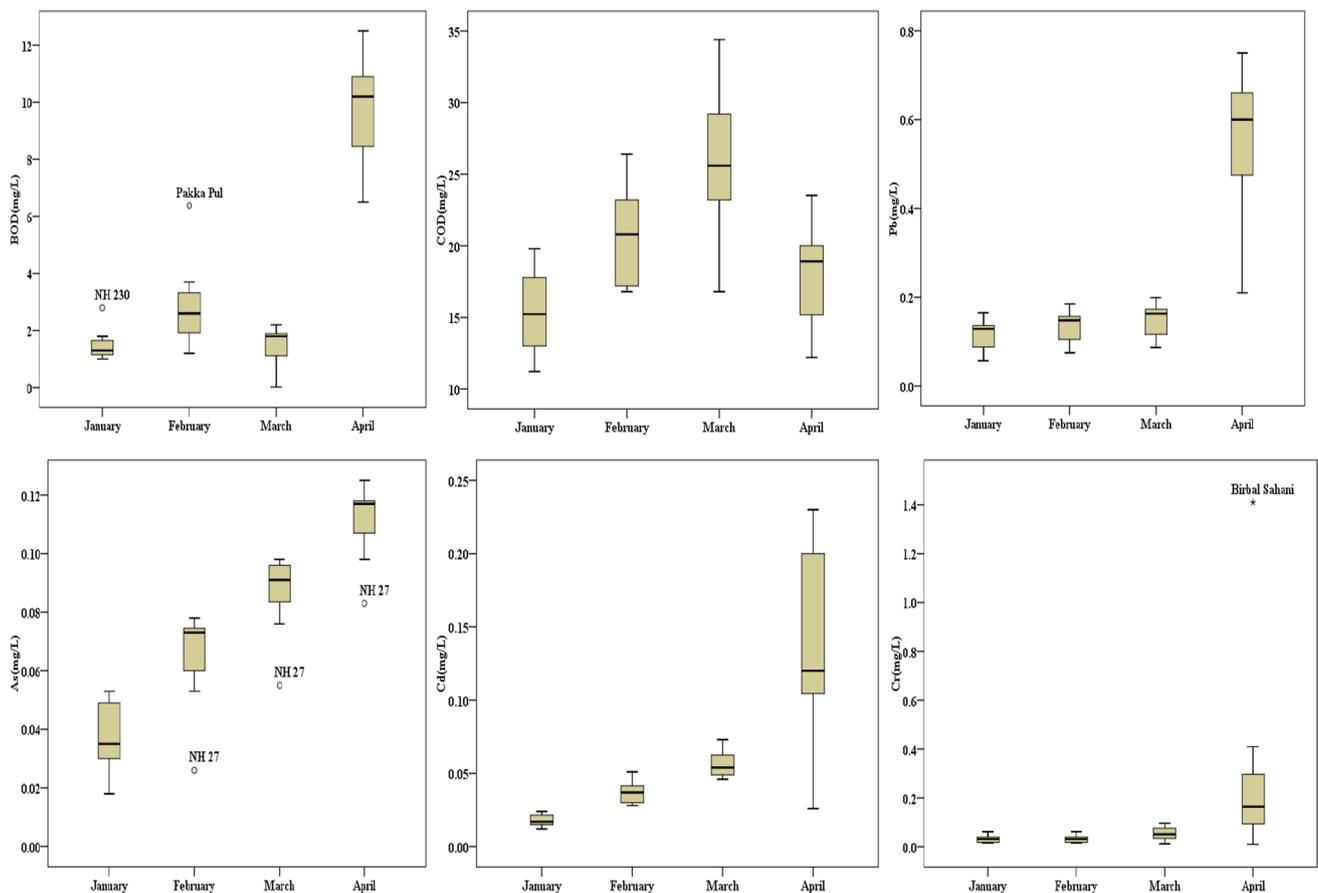


Fig. 3 (continued)

Table 2 Variation in physiochemical parameters of water at various sampling sites (Dutta et al. 2018)

	NH 230	Pakka Pul	KGMC	Hanuman Setu	Birbal Sahani	Lohia Path	NH 27
pH	8.17 ± 0.51	7.76 ± 0.58	7.08 ± 0.61	7.70 ± 0.15	7.84 ± 0.52	7.57 ± 0.41	7.44 ± 0.48
EC	264 ± 23.76	525.75 ± 218.10	662.25 ± 163.15	640 ± 103.68	691 ± 122.86	467.25 ± 140.28	468.5 ± 138.64
TDS	132 ± 11.88	262.87 ± 109.05	331.12 ± 81.57	320 ± 51.84	345.5 ± 61.43	233.62 ± 70.14	234.25 ± 69.32
NO ₃ ⁻	0.50 ± 0.55	0.71 ± 0.147	0.17 ± 0.18	0.24 ± 0.31	0.19 ± 0.21	0.33 ± 0.15	0.31 ± 0.28
NO ₂ ⁻	0.52 ± 0.186	2.08 ± 0.51	0.81 ± 1.01	0.88 ± 0.77	2.85 ± 1.49	2.15 ± 0.72	0.99 ± 1.17
NH ₄ ⁺	1.66 ± 0.98	0.33 ± 0.18	5.44 ± 3.99	13.57 ± 26.20	35.80 ± 46.73	18.05 ± 34.34	22.05 ± 41.19
PO ₄ ⁻³	2.15 ± 2.155	4.64 ± 2.71	4.25 ± 3.21	3.95 ± 3.14	6.32 ± 6.61	4.79 ± 3.22	4.87 ± 3.62
SO ₄ ⁻²	1.89 ± 0.99	3.81 ± 1.75	3.25 ± 2.21	4.32 ± 4.82	5.31 ± 2.92	4.25 ± 3.38	4.49 ± 2.80
Alk	241.25 ± 19.61	247.5 ± 14.10	181.58 ± 16.37	228.25 ± 14.88	253.16 ± 65.23	242 ± 30.27	237.75 ± 16.26
Hard	25.5 ± 9.71	37 ± 10.03	38.25 ± 15.84	48 ± 13.46	58 ± 12.19	48 ± 8.67	29.75 ± 7.41
Cl ⁻	14.05 ± 10.69	22.71 ± 8.96	23.21 ± 10.23	19.60 ± 17.47	26.19 ± 22.07	20.48 ± 12.06	31.11 ± 2.75
DO	6.64 ± 0.65	6.31 ± 2.78	5.34 ± 0.63	4.2 ± 2.03	5.45 ± 2.13	5.32 ± 0.67	4.96 ± 1.87
BOD	3.74 ± 1.96	5.44 ± 4.47	2.89 ± 3.55	3.59 ± 4.56	3.9 ± 4.37	4.10 ± 5.59	3.65 ± 3.40
COD	16.47 ± 5.41	20.13 ± 3.84	20.39 ± 4.96	22.93 ± 8.30	23.40 ± 6.27	17.76 ± 4.30	18.72 ± 6.79
Pb	0.31 ± 0.26	0.23 ± 0.17	0.23 ± 0.14	0.30 ± 0.29	0.24 ± 0.23	0.11 ± 0.06	0.20 ± 0.26
As	0.078 ± 0.039	0.084 ± 0.02	0.078 ± 0.04	0.065 ± 0.027	0.0837 ± 0.03	0.079 ± 0.038	0.045 ± 0.03
Cd	0.031 ± 0.016	0.056 ± 0.033	0.084 ± 0.08	0.073 ± 0.085	0.058 ± 0.04	0.08 ± 0.10	0.054 ± 0.04
Cr	0.035 ± 0.015	0.085 ± 0.05	0.075 ± 0.07	0.034 ± 0.033	0.37 ± 0.68	0.084 ± 0.061	0.031 ± 0.01

Table 3 Correlation of water quality parameters (Pearson coefficient)

Variables	pH	EC	TDS	NO ₃ ⁻	NO ₂ ⁻	NH ₄ ⁺	PO ₄ ⁻³	SO ₄ ²⁻	Alk
pH	1								
EC	-0.52	1							
TDS	-0.52	1.00*	1						
NO ₃ ⁻	0.48	-0.57	-0.57	1					
NO ₂ ⁻	0.13	0.39	0.39	0.04	1				
NH ₄ ⁺	-0.01	0.46	0.46	-0.50	0.68	1			
PO ₄ ⁻³	-0.33	0.70**	0.70**	-0.35	0.81*	0.85*	1		
SO ₄ ²⁻	-0.24	0.68**	0.68	-0.41	0.68	0.88*	0.92*	1	
Alk	0.75**	-0.33	-0.33	0.44	0.55	0.44	0.22	0.31	1
Hard	-0.06	0.74	0.74**	-0.49	0.72	0.69	0.72**	0.75*	0.17
Cl ⁻	-0.55	0.56	0.56	-0.32	0.41	0.66	0.83*	0.76*	-0.01
DO	0.49	-0.60	-0.60	0.70**	0.08	-0.42	-0.35	-0.61	0.26
BOD	0.43	-0.19	-0.19	0.83*	0.54	-0.03	0.17	0.12	0.67
COD	-0.15	0.90*	0.90*	-0.44	0.38	0.51	0.60	0.68	-0.05
Pb	0.46	-0.05	-0.05	0.09	-0.49	-0.33	-0.51	-0.41	-0.07
As	0.28	0.11	0.11	0.24	0.47	-0.16	0.02	-0.20	0.05
Cd	-0.79**	0.68	0.68	-0.53	0.16	0.12	0.41	0.41	-0.57
Cr	0.10	0.48	0.48	-0.35	0.74*	0.86*	0.79*	0.67	0.37
Variables	T-Hard	Cl ⁻	DO	BOD	COD	Pb	As	Cd	Cr
pH									
EC									
TDS									
NO ₃ ⁻									
NO ₂ ⁻									
NH ₄ ⁺									
PO ₄ ⁻³									
SO ₄ ²⁻									
Alk									
Hard	1								
Cl ⁻	0.27	1							
DO	-0.46	-0.39	1						
BOD	0.02	0.00	0.46	1					
COD	0.74**	0.42	-0.55	-0.06	1				
Pb	-0.25	-0.40	0.13	-0.16	0.25	1			
As	0.36	-0.41	0.55	0.32	0.09	0.01	1		
Cd	0.53	0.29	-0.63**	-0.29	0.38	-0.52	0.06	1	
Cr	0.62	0.58	-0.03	0.04	0.53	-0.12	0.20	-0.06	1

Bold characters signify the highest degree of correlation of the variables with each other

assessment of water quality (Martinez-Tavera et al. 2017). To reduce the necessity for several sampling sites, CA technique is a good tool.

Principal component analysis of quality of Gomti River

The PCA/FA depending on the correlation matrix has been implemented to discern the inherent interrelationships

amongst the various water quality variables of all the studied sites, so as to observe their properties (Barakat et al. 2016). The factor loadings were categorized as 'strong', 'moderate', and 'weak' in accordance to the total loading values of 1–0.75, 0.75–0.50 and 0.50–0.30, respectively (Azid et al. 2015). High positive and negative loadings reveal that the variables are significant for the given source of pollution (Burstyn 2004). The factor loadings of all principal components (PC)/factors, eigenvalues (Ev), variability and

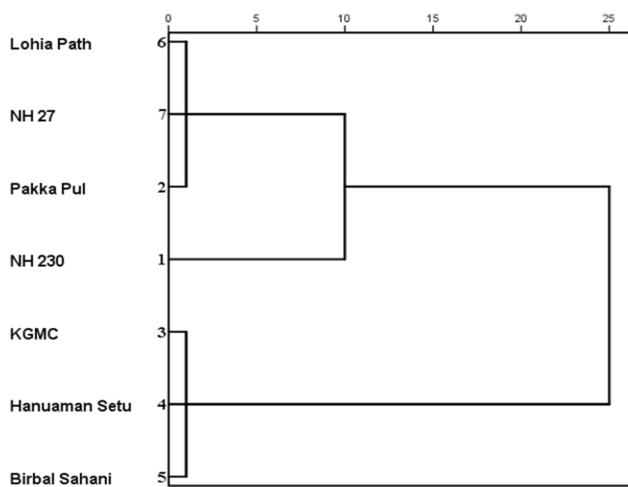


Fig. 4 Dendrogram based on hierarchical clustering (Ward’s method) for seven sampling sites of Gomti River in Lucknow City

entire variance revealed high positive loadings for EC, TDS, NH_4^+ , PO_4^{3-} , SO_4^{2-} , Hard, COD, moderate positive loadings for NO_2^- , Cl^- , Cd, Cr; moderate negative loadings for NO_3^- , and DO; weak negative loadings for pH, and Pb. The presence of such contaminants is indicative of organic as well as inorganic pollution from anthropogenic sources, such as domestic waste water, untreated municipal sewage discharge, release from industrial effluents and water treatment plants. Cd and Cr presence are indicative of industrial effluents resulting from the production of corrosion inhibitors and pigment which in turn could be toxic to aquatic life (Jaishankar et al. 2014). Presence of PO_4^{3-} could be due to the detergent discharges from the municipal sewage and industrial effluents. Presence of NO_2^- and SO_4^{2-} is due to fertilisers and livestock waste run-off from agricultural fields. The second PC responsible for 22.97% of the entire variance where Alk and BOD displayed high positive loadings, pH, NO_3^- , NO_2^- , DO and

Table 4 Loadings of 18 variables on principal components for the whole datasets

PC	pH	EC	TDS	NO_3^-	NO_2^-	NH_4^+	PO_4^{3-}	SO_4^{2-}	Alk	Hard
1st PC	-0.46	0.88*	0.88*	-0.63**	0.61**	0.79*	0.91*	0.91*	-0.04	0.83*
2nd PC	0.69**	-0.22	-0.22	0.54**	0.73**	0.38***	0.32	0.25	0.90*	0.23
3rd PC	0.27	0.35***	0.35***	-0.03	0.05	-0.26	-0.18	-0.23	-0.21	0.35***
4th PC	0.43***	-0.05	-0.05	-0.31	-0.29	0.33	-0.11	0.10	0.25	-0.05
5th PC	-0.09	0.21	0.21	0.47***	-0.11	-0.23	0.03	0.12	0.02	-0.20

PC	Cl^-	DO	BOD	COD	Pb	As	Cd	Cr	Ev	Var (%)	Cum %
1st PC	0.74**	-0.64**	-0.13	0.78*	-0.36	-0.04	0.62*	0.69*	8.08	44.91	44.91
2nd PC	0.02	0.52**	0.78*	0.00	-0.10	0.40***	-0.52**	0.51**	4.13	22.97	67.87
3rd PC	-0.49***	0.16	-0.03	0.47***	0.58**	0.72**	0.07	0.05	1.98	11.00	78.87
4th PC	0.03	-0.26	-0.36	0.26	0.64**	-0.49**	-0.53**	0.26	1.82	10.10	88.97
5th PC	0.25	-0.13	0.48***	0.32	0.29	-0.27	-0.03	-0.26	1.07	5.96	94.93

The factor loadings, i.e., the physicochemical parameters that contribute significantly (either positively or negatively) to the respective principal component have been shown in bold

cumulative variance are given in Table 4. A high value of eigenvalue clearly defines the significance of loadings. Sum of all eigenvalues are equal to variance (Abdi and Williams 2010). The graphs in the scree plot are extensively applied for categorizing the total factors present in the elementary data structure (Liu et al. 2003).

Figure 5 shows sorted Ev from large to small as a function of the factor number. The scree plot graphs are used to ascertain the number of factors or the principal factors of the given data which clearly define that the first four factors are responsible for 88.97% of the total variance.

In the present study, Ev was taken out from the PCs. Four PCs were acquired with $\text{Ev} > 1$ that described approximately 89% of the total variance of the dataset (Table 4). The variables with $\text{Ev} < 1$ were removed due to their low significance. The first PC accounting for 44.91% of the

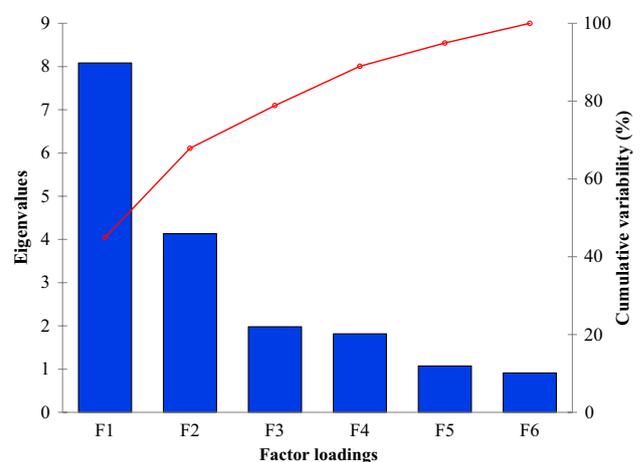


Fig. 5 Scree plot of the eigenvalues (Ev)

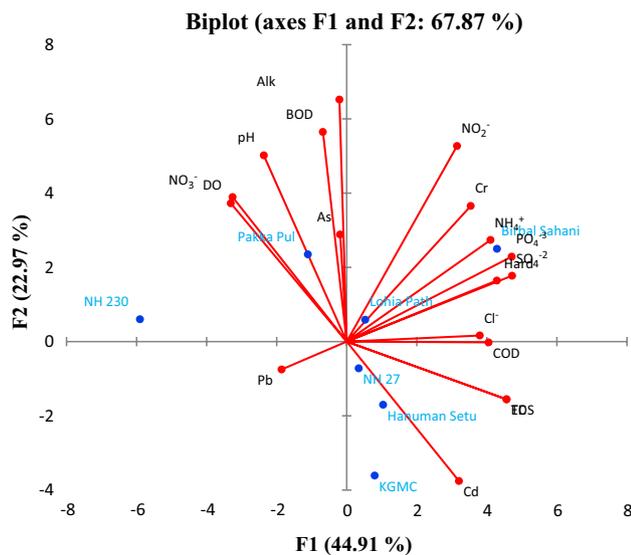


Fig. 6 Biplot between physiochemical parameters and sampling sites

Cr moderate positive loadings, NH_4^+ , PO_4^{3-} and As show weak positive but Cd displayed moderate negative loading. These including organic pollution possibly be related to inputs from point sources (municipal & industrial), non-point sources (agricultural fields run-off), livestock activities coupled with domestic influences (Khatri and Tyagi 2015). Cumulative effects of the anthropogenic pollution from different sources like untreated municipal discharges directly into the river via drain which is responsible for high BOD. Third PC contribute 11% of the total variance in which Pb and As showed moderate positive loadings, EC, TDS, Hard and COD showing weak positive loadings indicate presence of inorganic pollutant and Cl^- showed weak negative loading is indicative of the presence of a huge fraction of non-biodegradable waste by anthropogenic activity.

The fourth PC responsible for 10.10% of entire variance displayed moderate positive loadings for Pb, and moderate negative loading for Cd; weak positive loading for pH and NH_4^+ , weak negative loading for BOD and As indicating of the industrial pollution. The fifth PC indicating for 5.96% of the variance showed weak positive loading for NO_3^- , BOD and COD was indicative of organic as well as inorganic matter decomposition in the river. According to biplot, F1 axis contributes 44.91% and F2 axis 22.97% of total variance and the point signifies physiochemical parameters and observations, i.e. sites. Birbal Sahani (mid-stretch) in the F1 on the right top corner showed loadings for NH_4^+ , SO_4^{2-} , PO_4^{3-} , NO_2^- , Hardness, COD, Cl^- , EC, TDS, Cr and Cd. Pakka Pul in the second axis on the left top corner had loadings for pH, Alk, BOD, DO NO_3^- and As (Fig. 6). Hanuman Setu, NH 27, KGMC comprising a

Table 5 WQI of Gomti river stretch in Lucknow city (Dutta et al. 2018)

Sites	WQI (%)
NH 230	65.26
Pakka Pul	62.81
KGMC	61.87
Hanuman Setu	60.80
Birbal Sahani	58.98
Lohia Path	65.44
NH 27	66.64

group showed same pollution level while NH 230 shows a low pollution than other sites.

Water quality index of Gomti River

The current water quality index designates threatened quality of water in Gomti River attributable to recurrent release of waste effluents by numerous discharges poured down directly into the river by anthropogenic sources (Dutta et al. 2018; Sharma and Kansal 2011). Gomti river has a fair index of water quality when it enters the Lucknow city nearby Ghaila, whilst as it advances within the city premises at sites like Pakka Pul, KGMC, Hanuman Setu, Birbal Sahani the index worsens to marginal (Table 5). At Birbal Sahani, the index of water quality was 58.98 which are clearly evident from Table 5 and gives an impression that this particular site being located in the mid-Lucknow is terribly contaminated ascribed to the numerous drainage influences (domestic as well as municipal) from the main city. Once the river moves away from Birbal Sahani and travels downstream at sites like Lohia Path and NH 27, it is discerned that the index of the Gomti water again ameliorated to fair by virtue of the self-replenishing power of a river as it advances away from the heterogeneous pollution sources.

Discussion

The statistical techniques are used on a vast scale for the temporal as well as spatial assessment of the physico-chemical, bacteriological parameters of the water quality of any water body but the underlying principles governing certain ongoing imperative processes within the riverine or aquatic ecosystem is not investigated thoroughly. Although a number of studies have been published emphasizing on the spatio-temporal variations of water quality parameters (physical, chemical, biological) and these studies have also emphasized on the multivariate statistical approaches like the PCA, FA, CA, DA and many more. In a study done on eight locations of Gomti river for 3 years to analyse 34 parameters, the use of multivariate statistical techniques was found to be quite significant for effectively analyse complex

water quality data and generate results for better monitoring and management of lotic water bodies like Gomti (Singh et al. 2005a, b). A biennium study on 16 polycyclic aromatic hydrocarbons (PAHs) in the water and sediments of Gomti river have also been conducted which suggests that pyrolytic and petrogenic sources are contributing for the PAHs in the river. The two and three ring PAHs in water and three and four ring PAHs in the sediments were dominant in the river (Malik et al. 2011). A monitoring study on the heavy metals (Cd, Cr, Cu, Fe, Mn, Pb, Zn and Ni) contamination for 5 years period revealed that the water and the bed sediments of the Gomti river are severely impacted by the human activities. However, after calculating the geo-accumulation index of the metals at individual sites, it was found that the river bed is severely impacted by the heavy metals which may cause toxicity in the river affecting its aquatic diversity (Singh et al. 2005a, b).

Various water quality indexes have also been developed for quality monitoring purposes utilizing certain parameters including the heavy metals. In the present study also, heavy metals (Pb, As, Cd, Cr) analysis has been done and a water quality index has been developed including these heavy metals. Though the water quality index gives a clear understanding of the condition of any water body but it does not inform us about the background processes in an aquatic ecosystem. The outcome of any particular data greatly depends on certain factors like data quality, handling and data understanding prior to statistical analysis and data interpretation. Most of the research basically deals with the monitoring of the water quality and its management but does not address the major points that drive these problems. The research should be focussed on applying one's own knowledge and wisdom first and then moving further ahead for other statistical complexity for interpreting the outcomes.

Keeping these questions in mind, one such study has been carried out on the Tay Ninh River of Dong Nai River basin in Vietnam where pollutant loaded (nutrients and organic loadings) water body has been studied for nitrification inhibition by means of statistical methods (PCA, CA, DA) (Le et al. 2017). Nearly at all stations, the mean ammonium concentration was found about 1.5 mg N/L, whilst the mean nitrate concentration was around 0.4 mg N/L. As per the study of Wetzel (2001), when the nitrification process operates properly in an unpolluted water body, the $\text{NO}_3\text{-N}$ to $\text{NH}_4\text{-N}$ ratio is 25:1 whilst the ratio fluctuated from 1:2 to 1:12.8 in the Tay Ninh River. The primary productivity of certain ecosystems is regulated by nitrogen because it is a vital nutrient required for the primary production of plants and algae. Nitrogen along with another nutrient phosphorus is needed in a proportion of 30:1 by the aquatic macrophytes. The nitrogen requirement of higher plants is twice greater than that of phytoplankton when compared with phosphorous. Through various studies, it is evidenced that

nitrogen unavailability hampers the primary production in the higher plants (Durand et al. 2011). N in excess or deficit is both harmful for the healthy and balanced well-being of a sustainable aquatic ecosystem. While the excess causes serious problems of water eutrophication, water hypoxia, species diversity decrement, water habitat and ecosystem disruption (Durand et al. 2011); its deficit may result in the low primary productivity of the aquatic flora like the macrophytes and cyanobacteria.

Oxygen deficiency occurs when sufficiently huge quantity of ammonium due to untreated and partially treated wastewater is present in a lotic water body. In European rivers, nearly 0.1 mg N/L (nitrogen per litre) per day of nitrification rate has been documented. In these rivers, the bacterial population varies from 0.005–0.05 to 0.01–0.5 mg C/L (coliform/litre) per day and as high as 2.5 mg C/L per day in the case of unaltered natural forest rivers, large rivers and heavily polluted rivers with municipal sewage, respectively (Brion and Billen 2000). Inhibition occurs when the activity of the ammonium oxidizers and the nitrite oxidizers (chemolithotrophic bacteria) are negatively influenced by some factors like dissolved oxygen, pH, water temperature, substrate concentration and salinity (Chen et al. 2006). These obligate autotrophic bacteria that carry out the biological oxidation converting ammonium to nitrate makes energy via nitrification for fixing CO_2 .

BOD signifies the volume of oxygen used by two major processes: carbonaceous biochemical oxygen demand (CBOD) which is the decrease in the organic carbon matter by the metabolic process to CO_2 with the help of microorganisms. Nitrogenous biochemical oxygen demand (NBOD) is the volume of oxygen consumed in the biological oxidation process of converting ammonia to nitrate during nitrification. The total amount of oxygen required by carbonaceous and nitrogenous matter for extended period of time is the ultimate BOD. Considering this, the CBOD and NBOD should be separately analysed. There is another aspect of free-flowing rivers which is an amalgamation of various processes, which is sediment oxygen demand (SOD). The amount of oxygen consumed for the decay of the settled organic matter in the river is the SOD. In a free-flowing river, the SOD is nearly insignificant as the long-term organic matter aggregation is prohibited due to recurrent abrasion during floods and storm events. As Gomti is not a free-flowing river as it used to be before channelization, so the SOD should also be considered for analysis. The NBOD could significantly increase in the water body during the nitrification process of organic nitrogen and $\text{NH}_3\text{-N}$ conversion to nitrate by consuming oxygen. The CBOD is even smaller than NBOD even though the NBOD does not increase rapidly as NBOD is reliant on various environmental aspects like alkalinity, nitrifying bacterial populace, temperature and dissolved oxygen accessibility. Along with the

CBOD, a comparatively great pressure on oxygen is exerted by NBOD. As for the oxidation of 1 mol of NH_4^+ , 1.89 mol of oxygen are utilized thereby producing 1.98 mol of hydrogen ions. Therefore, NBOD and CBOD should be studied separately and SOD should also be considered.

The researchers have always made an attempt for studying the aquatic ecosystem for extended time span whether the objective is developing the water quality index (Tripathi and Singal 2019; Tian et al. 2019; Ewaid et al. 2018), transitional water quality index development (Giordani et al. 2009) or for long-term analysis of the physico-chemical and biological characterization of a riverine ecosystem (Jameel et al. 2018; Gupta et al. 2017; Le et al. 2017). Pollution is such an anthropogenic nuisance that never occurs for a season or two; it is constantly being added in a river day by day due to poor sanitation activities along the banks, the industrial and agricultural discharge and direct dumping. So, long-term studies are not enough for monitoring purposes. The plume of pollutant load is constantly moving from one place to another as river is a lotic ecosystem. So, the authors suggest short-term studies of a surface water body.

Analysing all the ongoing processes within an aquatic ecosystem, it prompts us to come on a viewpoint that merely the analysis, monitoring and index development does not signify the problematic status of any surface water body. As it is evident from the study of Le et al. (2017), a lot of nitrate formation from ammonium via nitrification is inhibited sometimes due to lack of pH, DO and certain other factors. These behind the scenes problems are not integrated in the development of any kind of water quality index developed so far. Though a rapid WQI development is required for monitoring but aspects such as nitrification inhibition, carbonaceous and nitrogenous biochemical oxygen demand, and sediment oxygen demand should also be considered at the same time. When these vital processes are considered and integrated for the evaluation of water quality status, then only the monitoring studies will give us a comprehensive understanding of the current status of an aquatic ecosystem.

Conclusion

In the current study, multivariate techniques like PCA/FA, CA were found to be efficient and realistic in assessing spatio-temporal differences of river water quality and for identifying contamination sources, site categorization besides identifying substantial parameters of Gomti water quality in Lucknow city. CA grouped the sampling sites into two important clusters (Cluster 1 (Moderately Polluted): Lohia Path, NH 27, Pakka Pul and NH 230; Cluster 2 (Heavily Polluted): KGMC, Hanuman Setu and Birbal Sahani) illustrating that water quality of Gomti river gradually worsened from upstream and downstream to middle

stretch. The middle stretch was found to be most polluted as the major drains are concentrated within this stretch. PCA/FA helped in obtaining and recognizing the factors/sources accountable for river water quality differences in the study area. Boxplots proved to be a potent tool for spatio-temporal variation analysis of the river. The sources responsible for contamination are mixing of semi-treated/untreated urban wastewater through many drains, slum areas adjoining the river bank that release its untreated wastewater and garbage directly to the river, direct dumping of religious and household wastes by the localities, Dhobhighats (sites for washing clothes by the washer men) and crematorium on river banks, along with agricultural runoff. The latent reason behind the deterioration in water quality is the lack of awareness among the locals and their cultural practices such as immersion of idols and religious materials during and after various festivals. Also, the river channelization has deteriorated the water quality to a great extent as the water flow is controlled by barrages and weirs in this stretch, so the river is unable to replenish itself. To monitor rivers like Gomti, short-term monitoring and rapid water quality index development are the need of the hour but at the same time before coming on any conclusion, all the natural ongoing operations like nitrification inhibition, nitrogenous and carbonaceous biochemical oxygen demand, sediment oxygen demand of the river should be considered at the same time for water quality index development. Our study emphasized on short-term continuous monthly measurement of the water quality and index development of the river so that management decisions could focus on gradual improvements in the water quality.

Acknowledgements The authors would like to thank Head, Department of Environmental Science, Babasaheb Bhimrao Ambedkar University, Lucknow for providing the necessary instrumentation facilities required for this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval and informed consent Not applicable. The study does not involve human or animal subjects.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdi H, Williams LJ (2010) Principal component analysis. Wiley Interdiscip Rev Comput Stat 2:433–459
- Aissa-Grouz N, Garnier J, Billen G, Mercier B, Martinez A (2015) The response of river nitrification to changes in wastewater treatment (the case of the lower Seine River downstream from Paris). *Int J Limnol* 51(4):351–364 (EDP Sciences)
- Akbarzadeh Z, Laverman AM, Rezanezhad F, Raimonet M, Viollier E, Shafei B, Van Cappellen P (2018) Benthic nitrite exchanges in the Seine River (France): an early diagenetic modeling analysis. *Sci Total Env* 628:580–593
- Alilou H, Nia AM, Keshtkar H, Han D, Bray M (2018) A cost-effective and efficient framework to determine water quality monitoring network locations. *Sci Total Environ* 624:283–293
- Azid A, Juahir H, Toriman ME, Endut A, Kamarudin MKA, Rahman MNA, Yunus K (2015) Source apportionment of air pollution: a case study in Malaysia. *J Teknologi* 72:83–88
- Bajpayee S, Das R, Ruj B, Adhikari K, Chatterjee PK (2012) Assessment by multivariate statistical analysis of ground water geochemical data of Bankura, India. *Int J Environ Sci* 3:870
- Barakat A, El Baghdadi M, Rais J, Aghezzaf B, Slassi M (2016) Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. *J Soil Water Conserv* 4:284–292
- Brion N, Billen G (2000) Wastewater as a source of nitrifying bacteria in river systems: the case of the River Seine downstream from Paris. *Water Res* 34(12):3213–3221
- Burstyn I (2004) Principal component analysis is a powerful instrument in occupational hygiene inquiries. *Ann Occup Hyg* 48:655–661
- Chaurasia AK, Pandey HK, Tiwari SK, Prakash R, Pandey P, Ram A (2018) Groundwater Quality assessment using Water Quality Index (WQI) in parts of Varanasi District, Uttar Pradesh, India. *J Geol Soc India* 92(1):76–82
- Chen S, Ling J, Blancheton JP (2006) Nitrification kinetics of biofilm as affected by water quality factors. *Aquacult Eng* 34(3):179–197
- Chowdhury S, Mazumder MJ, Al-Attas O, Husain T (2016) Heavy metals in drinking water: occurrences, implications, and future needs in developing countries. *Sci Total Environ* 569:476–488
- Dabrowski JJ, Rahman A, George A (2018) Prediction of dissolved oxygen from pH and water temperature in aquaculture prawn ponds. In: Proceedings of the Australasian joint conference on artificial intelligence-workshops, pp 2–6
- Diamantini E, Lutz SR, Mallucci S, Majone B, Merz R, Bellin A (2018) Driver detection of water quality trends in three large European river basins. *Sci Total Environ* 612:49–62
- Durand P, Breuer L, Johnes PJ, Billen G, Butturini A, Pinay G, Curtis C (2011) Nitrogen processes in aquatic ecosystems. Cambridge University Press, Cambridge, pp 126–146
- Dutta V, Srivastava RK, Yunus M, Pathak VV, Shamsad A, Rai A, Prasad N (2011) Restoration plan of Gomti River with designated best use classification of surface water quality based on river expedition, monitoring and quality assessment. *Earth Sci India* 4(3):80–104
- Dutta V, Sharma U, Iqbal K, Kumar R, Pathak AK (2018) Impact of river channelization and riverfront development on fluvial habitat: evidence from Gomti River, a tributary of Ganges, India. *Environ Sustain* 1(2):167–184
- Ewaid SH, Abed SA, Kadhum SA (2018) Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. *Environ Technol Innov* 11:390–398
- Furlan A, Galeazzo A, Paggiaro A (2019) Organizational and perceived learning in the workplace: a multilevel perspective on employees' problem solving. *Organ Sci* 30(2):280–297
- Giordani G, Zaldívar JM, Viaroli P (2009) Simple tools for assessing water quality and trophic status in transitional water ecosystems. *Ecol Ind* 9(5):982–991
- Gupta SK, Chabukdhara M, Kumar P, Singh J, Bux F (2014) Evaluation of ecological risk of metal contamination in river Gomti, India: a biomonitoring approach. *Ecotoxicol Environ Saf* 110:49–55
- Gupta N, Pandey P, Hussain J (2017) Effect of physicochemical and biological parameters on the quality of river water of Narmada, Madhya Pradesh, India. *Water Sci* 31(1):11–23
- He B, Dai M, Zhai W, Guo X, Wang L (2014) Hypoxia in the upper reaches of the Pearl River Estuary and its maintenance mechanisms: a synthesis based on multiple year observations during 2000–2008. *Mar Chem* 167:13–24
- Iqbal M, Masum Billah M, Nurul Haider M, Shafiqul Islam M, Rajib Payel H, Khurshid Alam Bhuiyan M, Dawood MA (2017) Seasonal distribution of phytoplankton community in a subtropical estuary of the south-eastern coast of Bangladesh. *Zool Ecol* 27:304–310
- Jaishankar M, Tseten T, Anbalagan N, Mathew BB, Beeregowda KN (2014) Toxicity, mechanism and health effects of some heavy metals. *Interdiscip Toxicol* 7:60–72
- Jameel Y, Stein S, Grimm E, Roswell C, Wilson AE, Troy C, Bowen GJ (2018) Physicochemical characteristics of a southern Lake Michigan river plume. *J Great Lakes Res* 44(2):209–218
- Kansal ML (2018) Issues and challenges of River Health Assessment in India. *Water Resour Manag* 105:119
- Khan MYA, Gani KM, Chakrapani GJ (2016) Assessment of surface water quality and its spatial variation. A case study of Ramganga River, Ganga Basin, India. *Arab J Geosci* 9:28
- Khatri N, Tyagi S (2015) Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas. *Front Life Sci* 8:23–39
- Kumar P, Kaushal RK, Nigam AK (2015) Assessment and management of Ganga river water quality using multivariate statistical techniques in India. *Asian J Water Environ Pollut* 12:61–69
- Kumarasamy P, James RA, Dahms HU, Byeon CW, Ramesh R (2014) Multivariate water quality assessment from the Tamiraparani river basin, Southern India. *Environ Earth Sci* 71:2441–2451
- Le TTH, Zeunert S, Lorenz M, Meon G (2017) Multivariate statistical assessment of a polluted river under nitrification inhibition in the tropics. *Environ Sci Pollut Res* 24(15):13845–13862
- Li B, Yang G, Wan R, Hörmann G, Huang J, Fohrer N, Zhang L (2017) Combining multivariate statistical techniques and random forests model to assess and diagnose the trophic status of Poyang Lake in China. *Ecol Indic* 83:74–83
- Little RJ, Rubin DB (2019) Statistical analysis with missing data, vol 793. Wiley, Hoboken
- Liu RX, Kuang J, Gong Q, Hou XL (2003) Principal component regression analysis with SPSS. *Comput Methods Prog Biomed* 71:141–147
- López-Morales CA, Mesa-Jurado MA (2017) Valuation of hidden water ecosystem services: the replacement cost of the aquifer system in Central Mexico. *Water* 9:571
- Lorenz M, Prilop K, Thang MT, Le H, Hieu ND, Meon G, Quan NH (2014) Ecohydrological modelling of the Thi Vai catchment in South Vietnam. In: 2014 EWATEC-COAST: Technologies for environmental and water protection of coastal zones in Vietnam. Contributions to 4th international conference for environment and natural resources, ICENR, pp 2363–7218
- Majkic-Dursun B, Oros I, Boreli-Zdravković Đ (2018) Spatial distribution of groundwater quality parameters in the Velika Morava River Basin, Central Serbia. *Environ Earth Sci* 77:30
- Malik A, Verma P, Singh AK, Singh KP (2011) Distribution of polycyclic aromatic hydrocarbons in water and bed sediments of the Gomti River, India. *Environ Monit Assess* 172(1–4):529–545

- Martinez-Tavera E, Rodriguez-Espinosa PF, Shruti VC, Sujitha SB, Morales-Garcia SS, Muñoz-Sevilla NP (2017) Monitoring the seasonal dynamics of physicochemical parameters from Atoyac River basin (Puebla), Central Mexico: multivariate approach. *Environ Earth Sci* 76:95
- Mohebbi MR, Saeedi R, Montazeri A, Vaghefi KA, Labbafi S, Oktaie S, Mohagheghian A (2013) Assessment of water quality in groundwater of Iran using a modified drinking water quality index (DWQI). *Ecol Indic* 30:28–34
- Nadiri AA, Aghdam FS, Khatibi R, Moghaddam AA (2018) The problem of identifying arsenic anomalies in the basin of Sahand dam through risk-based 'soft modelling'. *Sci Total Environ* 613:693–706
- Najar I, Khan A, Hai A (2017) Evaluation of seasonal variability in surface water quality of Shallow Valley Lake, Kashmir, India, using multivariate statistical techniques. *Pollution* 3:349–362
- Nguyen HQ, Radhakrishnan M, Huynh TTN, Bains-Salingay ML, Ho LP, Steen PVD, Pathirana A (2017) Water quality dynamics of urban water bodies during flooding in Can Tho City, Vietnam. *Water* 9:260
- Ojok W, Wasswa J, Ntambi E (2017) Assessment of seasonal variation in water quality in river Rwizi using multivariate statistical techniques, Mbarara Municipality, Uganda. *J Water Resour Prot* 9:83
- Sabia G, Petta L, Moretti F, Ceccarelli R (2018) Combined statistical techniques for the water quality analysis of a natural wetland and evaluation of the potential implementation of a FWS for the area restoration: the Torre Flavia case study, Italy. *Ecol Indic* 84:244–253
- Saeedi M, Abessi O, Sharifi F, Meraji H (2010) Development of groundwater quality index. *Environ Monit Assess* 163(1–4):327–335
- Sharma D, Kansal A (2011) Water quality analysis of River Yamuna using water quality index in the national capital territory, India (2000–2009). *Appl Water Sci* 1:147–157
- Sharma M, Kansal A, Jain S, Sharma P (2015) Application of multivariate statistical techniques in determining the spatial temporal water quality variation of Ganga and Yamuna Rivers present in Uttarakhand State, India. *Water Qual Expo Health* 7:567–581
- Shrestha AK, Basnet N (2018) An evaluation of physicochemical analysis and water quality index of Ratuwa River of Damak, Jhapa, Nepal. *Int J Recent Res Rev* XI(2):1–9
- Sim SF, Tai SE (2018) Assessment of a physicochemical indexing method for evaluation of tropical river water quality. *J Chem* 1–12
- Singh PK, Saxena S (2018) Towards developing a river health index. *Ecol Indic* 85:999–1011
- 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 (2005a) Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. *Anal Chim Acta* 538(1–2):355–374
- Singh KP, Malik A, Sinha S, Singh VK, Murthy RC (2005b) Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. *Water Air Soil Pollut* 166:321–341
- Singh AN, Shrivastava R, Mohan D, Kumar P (2018) Assessment of spatial and temporal variations in water quality dynamics of river Ganga in Varanasi. *Pollution* 4:239–250
- Sorche NL, McDermott G, O'Boyle S, Wilkes R, Stengel DB (2019) Decoupling abundance and biomass of phytoplankton communities under different environmental controls: a new multi-metric Index. *Front Marine Sci* 6:312
- Sutadian AD, Muttill N, Yilmaz AG, Perera BJC (2016) Development of river water quality indices—a review. *Environ Monit Assess* 188:58
- Tashtoush SM, Al-Subh SA (2015) Interpretation of groundwater quality parameters for springs in Tafleeh area in South of Jordan using principal components analysis. *J Environ Sci* 3:31–44
- Tian Y, Jiang Y, Liu Q, Dong M, Xu D, Liu Y, Xu X (2019) Using a water quality index to assess the water quality of the upper and middle streams of the Luanhe River, northern China. *Sci Total Environ* 667:142–151
- Tripathi M, Singal SK (2019) Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of river Ganga India. *Ecol Ind* 96:430–436
- Villas-Boas MD, Olivera F, de Azevedo JPS (2017) Assessment of the water quality monitoring network of the Piabanha River experimental watersheds in Rio de Janeiro, Brazil, using autoassociative neural networks. *Environ Monit Assess* 189:439
- Wang Y, Wang P, Bai Y, Tian Z, Li J, Shao X, Li BL (2013) Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China. *J Hydro Environ Res* 7:30–40
- Wetzel R (2001) *Limnology: lake and river ecosystems*, vol 3. Elsevier Academic Press, San Diego, pp 227–237
- Wu Z, Wang X, Chen Y, Cai Y, Deng J (2018) Assessing river water quality using water quality index in Lake Taihu Basin, China. *Sci Total Environ* 612:914–922
- Xue W, Huang D, Zeng G, Wan J, Zhang C, Xu R, Deng R (2017) Nanoscale zero-valent iron coated with rhamnolipid as an effective stabilizer for immobilization of Cd and Pb in river sediments. *J Hazard Mater* 341:381–389
- Zhao Y, Xia XH, Yang ZF, Wang F (2012) Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques. *Procedia Environ Sci* 13:1213–1226
- Zhao Y, Guo L, Liang J, Zhang M (2016) Seasonal artificial neural network model for water quality prediction via a clustering analysis method in a wastewater treatment plant of China. *Desalin Water Treat* 57:3452–3465

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.