Assessment of ground and surface water quality along the river Varuna, Varanasi, India

Pardeep Singh · R. K. Chaturvedi · Ankit Mishra · Lata Kumari · Rishikesh Singh ·
D. B. Pal · Deen Dayal Giri · Nand Lal Singh · Dhanesh Tiwary · Pradeep Kumar Mishra

Abstract Multivariate statistical techniques were employed for monitoring of ground-surface water interactions in rivers. The river Varuna is situated in the Indo-Gangetic plain and is a small tributary of river Ganga. The study area was monitored at seven sampling sites for 3 years (2010–12), and eight physio-chemical parameters were taken into account for this study. The data obtained were analysed by multivariate statistical techniques so as to reveal the underlying implicit information regarding proposed interactions for the relevant area. The principal component analysis (PCA) and cluster analysis (CA), and the results of correlations were also studied for all parameters monitored at every site. Methods used in this study are essentially multivariate statistical in nature and facilitate the interpretation of data so as to extract meaningful information from the datasets. The PCA technique was able to compress the data from eight to three parameters and captured about 78.5 % of the total variance by performing varimax rotation over the principal components. The varifactors, as yielded from PCA, were treated by CA which grouped them convincingly into three groups having similar characteristics and source of contamination. Moreover, the loading of variables on significant PCs showed correlations between various ground water and surface water (GW-SW) parameters. The correlation coefficients calculated for various physiochemical parameters for ground and surface water established the correlations between them. Thus, this study presents the utility of multivariate statistical techniques for evaluation of the proposed interactions and effective future monitoring of potential sites.

Keywords Multivariate analysis · Varuna River · Cluster analysis · Principal component analysis

P. Singh · L. Kumari · D. Tiwary
Department of Chemistry, Indian Institute of Technology
(Banaras Hindu University), Varanasi 221005, India

R. K. Chaturvedi · R. Singh
Institute of Environment and Sustainable Development,
Banaras Hindu University, Varanasi 221005, India

A. Mishra · D. B. Pal · D. D. Giri · P. K. Mishra (✉)
Department of Chemical Engineering and Technology, Indian
Institute of Technology, (Banaras Hindu University),
Varanasi 221005, India
e-mail: pkmishra.che@itbhu.ac.in

N. L. Singh
Department of Chemistry, Banaras Hindu University,
Varanasi 221005, India

Introduction

Surface water hydrochemistry depends on several natural factors like changes in precipitation, erosion and weathering of crustal materials as well as anthropogenic factors such as industrial and agricultural activities and increased water consumption due to rapid urbanization. Furthermore, urbanization has severely affected the water resources by increasing the pressure on urban hydrology (Mishra 2011). These factors have degraded the quality of surface water for drinking purposes and are a

threat to the existing water bodies (Singh et al. 2005a, b, c; Mishra 2010).

The discharge of waste water from municipal, industrial, and agricultural areas is an issue of serious concern as it affects river's ecology (Varol et al. 2013). Surface run off from agriculture and urban areas varies mostly with seasonal changes in precipitation, ground water flow, water interception, and abstraction. These factors strongly affect the concentration of pollutants in river water which in turn deteriorates river water quality (Vega et al. 1998).

The quality of river water at a point is a function of several developmental pressures including the lithology of the basin, atmospheric and anthropogenic inputs, and climatic conditions (Yu et al. 2010). These developmental pressures and increasing human population have made the river of the study area vulnerable to sewage flow, solid waste dumping, etc. and consequently exert pressure on the percolation and infiltration processes responsible for the groundwater recharge (Kumar et al. 2012). Hence, these pressures affect the ground water conditions which make it imperative to analyse the hydrochemical properties of both the stream and the aquifers in order to comprehend the interactions existent between them.

The interaction between ground water and surface water (GW-SW) has been studied focussing on different perspectives. For example: Sophocleous (2002) studied this interaction focusing on geomorphological, hydrogeological, and climatic control; Valett and Sheibley (2009) studied the geomorphic and biological structure and how they influence the direction of GW-SW interaction; Kikuchi et al. (2012) performed spatial telescopic study for improved characterization of GW-SW interaction; Lu et al. (2011) analysed interactions between atmosphere, surface, and ground water using EARTH model in the unsaturated zone; Anibas et al. (2011) adopted a simple thermal mapping method for studying seasonal spatial patterns of GW-SW interactions; Harveya et al. (2006) modeled decadal time scale GW-SW interactions; and Wade et al. (2006) studied nitrate dynamics in these systems.

Multivariate statistical analysis techniques such as principal component analysis (PCA) and cluster analysis (CA) have been applied to characterize and evaluate surface freshwater quality, because they are useful in verifying the temporal and spatial variations caused by natural and anthropogenic factors linked to seasonality (Helena et al. 2000; Shrestha and Kazama 2007; Singh

et al. 2005a, b, c). These techniques effectively compress the data and reveal various correlations between different parameters analysed. Studies have been conducted in the past to study GW-SW interactions by evaluating their hydrochemical properties through such methods (e.g., Woocay and Walton 2008; Mencio and Mas-Pla 2008). The previous studies have focused more on surface water quality (e.g., Sundaray et al. 2006; Sojka et al. 2008; and Selle et al. 2013) and hydrochemistry. For example: Singh et al. (2014) have used ensemble learning approach to evaluate influences on seasonal variations and anthropogenic activities on alluvial groundwater hydrochemistry; Isa et al. (2012) evaluated extent and severity of groundwater contamination based on hydrochemistry mechanism of sandy tropical coastal aquifers. Various other studies have also focused on aquifer dynamics and hydrochemistry, (Stigter et al. 2006; Soulsby et al. 2002). Multivariate statistical analysis techniques facilitate to draw an appropriate profile of the stream hydrology when there is insufficient data available or when the end members are imprecise (Mencio and Mas-Pla 2008).

In the present study river Varuna, a minor tributary of river Ganges was monitored. This river receives huge quantities of untreated sewage, agriculture runoff with pesticides, fertilizers, etc. from catchment areas which lead to degradation of water quality. A study by Singh and Dwivedi (2007) and Kumar et al. (2012) asserted that river Varuna receives sewage from twenty-two municipal drainages located on both sides of its 15-km-long passage through Varanasi city. So there is an urgent necessity to identify the pollution sources and estimate their quantitative contributions, in order to achieve an effective pollution control. Therefore, the objective of the present study is to establish correlations between various sites on the basis of CA and identify the major hydrochemical factors based on the PCA. These results will further help in the investigation of relevant human pressures contributing to degradation of the water quality and also aid in concluding about the GW-SW interactions. Hence, periodic monitoring and assessment of water quality helps to develop management strategies to control surface water pollution inspite of increasing urbanization and anthropogenic pressure on them.

Multivariate statistical analysis was employed to examine GW-SW interactions and the potent hydrochemical factors responsible for altering the quality of the river water in Varanasi. This work is structured as follows: (1) sampling and analysis of physiochemical

parameters from different sites selected throughout the city; (2) PCA analysis on the collected data to identify the extent of ground water or surface water contribution to pollution in the relevant site; and (3) a CA to show the correlations between sites.

Materials and methods

Hydrology and geology of Varuna River

Varuna River basin is situated in Indo-Gangetic plain (Fig. 1) which is underlain by Quaternary alluvial sediments of Pleistocene to recent age. The basin of this river consists of three main geological formations: newer alluvium and older alluvium aging from upper Pleistocene to recent and from middle to upper Pleistocene, respectively (Raju et al. 2009). The rocks associated with the newer alluvium formations are composed mainly of unconsolidated sand, silt, and clay while that associated with older alluvium consist mainly of consolidated clay with kankar, fine to medium grain, and some gravel. In the study area, the unconsolidated sediments form a sequence of clays and sands of various grades. The presence of clay and sand zones controls the availability of ground water in alluvial zones. The sand

Fig. 1 Geographical locations and photographs of water quality monitoring sites
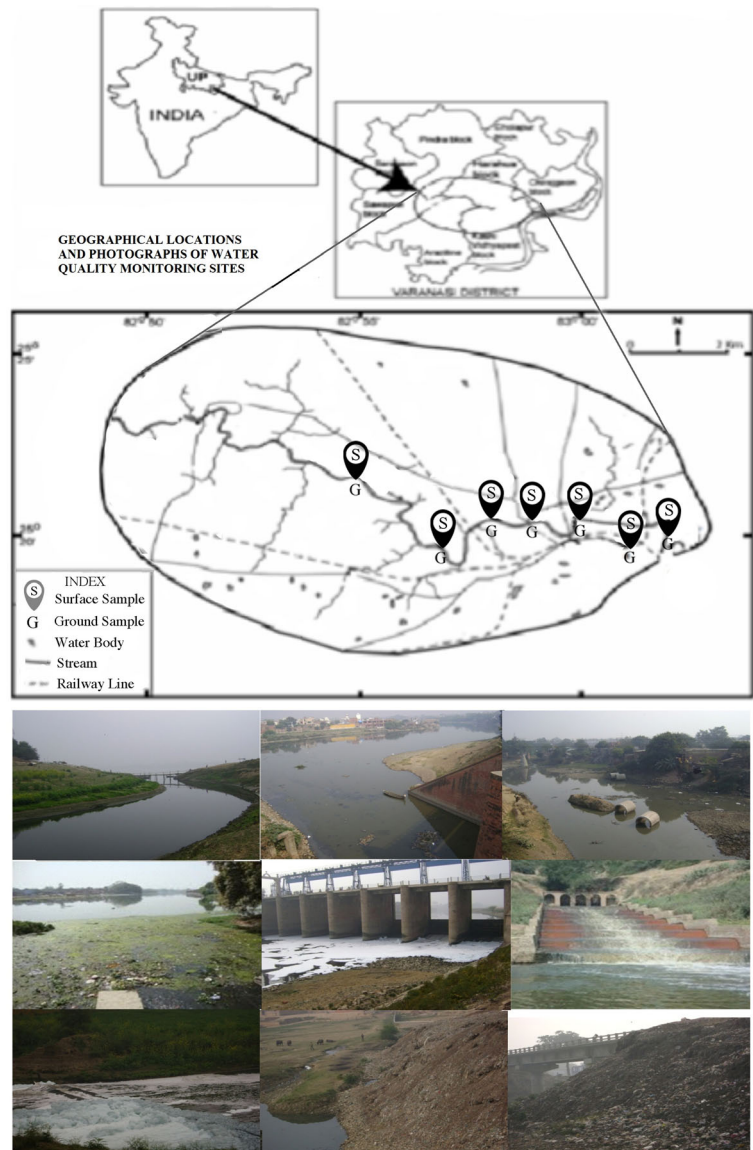
beds with or without kankar in the study area forms the main aquifer zones of the multi-tier aquifer system (Shukla and Raju 2008). Near surface groundwater occurs under water table condition, while deeper aquifers occur in semi-confined to confined conditions. Extraction of ground water is through both dug wells and deep bore wells. The rainfall is the main source of recharge to groundwater. The average hydraulic gradient is 0.35 m/km which indicates porous nature of near surface formations of the area (Pandey 1993).

Sampling and qualitative analysis

Seven sampling sites were randomly selected across the width of the river for the assessment of water quality. Each sampling site is in close proximity to a ground water pumping station to make the analysis more credible. These samples were collected during winter (2010–12) across a period of 3 months to monitor changes caused by anthropogenic as well as natural sources. Sampling, preserving, and transportation of the water samples to the laboratory were done as per the standard methods (APHA 1998). Samples were qualitatively analysed for eight different physiochemical parameters and were tabulated in Table 1. Triplicate samplings were done at each selected site at an interval of 1 month for three consecutive months as mentioned above. The pH, temperature, and electrical conductivity (EC) were determined immediately at the collection site by a portable pH meter, to minimize errors with time due to biological and chemical reactions between the atmosphere and the sample (Hutton 1983). Mean and standard deviation of water quality parameter studied both on ground (G) and surface (S) at different sites on Varuna River during winter season is given in Table 2.

Table 1 The associated water quality parameters, with abbreviation and their units used in study

Parameters	Abbreviations	Units
pH	pH	
Temperature	TEMP	°C
Electric conductivity	EC	$\mu S\ cm^{-1}$
Phosphate	$PO_4^{3-}{}_4^{-3}$	$Mg\ l^{-1}$
Chloride	Cl^-	$Mg\ l^{-1}$
Nitrate	NO_3^-	$Mg\ l^{-1}$
Total hardness	HARD	$Mg\ l^{-1}$
Total dissolved solid	TDS	$Mg\ l^{-1}$

Table 2 Mean and standard deviation of water quality parameter studied both on ground (G) and surface (S) at different sites on Varuna River during winter season

		TEMP	PH	TDS	NO_3^-	$PO_4^{3-}{}_4^{-3}$	Cl^-	HARD	EC
Nakhighat	(G) Mean±SD	19.70±0.03	7.40±0.12	330.60±0.51	21.00±0.50	2.90±0.03	42.40±0.22	305.45±0.69	660.17±0.29
	(S) Mean±SD	19.40±0.00	7.89±0.12	417.45±0.83	10.00±0.19	29.95±0.33	73.15±0.12	323.89±0.51	751.00±0.33
Chaukaghat	(G) Mean±SD	19.78±0.35	7.66±0.13	440.12±0.84	10.21±0.12	3.44±0.13	35.42±0.23	239.56±0.19	750.33±0.58
	(S) Mean±SD	20.08±0.10	7.73±0.03	527.22±0.51	6.78±0.38	41.63±0.32	56.72±0.13	350.78±0.19	845.00±0.32
Old Bridge	(G) Mean±SD	20.17±0.44	7.73±0.06	310.33±0.67	5.41±0.12	3.18±0.08	60.35±0.01	356.61±0.67	530.33±0.58
	(S) Mean±SD	20.14±0.18	7.63±0.06	400.22±0.51	4.10±0.37	31.91±0.34	62.78±0.16	296.67±0.33	764.00±0.67
Rajghat	(G) Mean±SD	19.45±0.25	7.68±0.16	910.45±0.69	18.11±0.42	2.39±0.10	215.93±0.19	324.26±1.48	1122.11±3.95
	(S) Mean±SD	20.47±0.29	8.01±0.05	603.33±0.00	13.82±0.08	42.61±0.48	122.04±0.27	296.22±3.08	1008.77±0.19
Kachahari	(G) Mean±SD	20.22±0.38	7.29±0.05	320.56±0.51	5.33±0.09	3.49±0.15	60.35±0.01	371.78±0.63	319.78±1.71
	(S) Mean±SD	20.91±0.14	7.80±0.06	443.89±0.51	9.85±0.51	19.80±0.33	43.63±0.21	317.33±0.00	849.11±1.07
Dhobighat	(G) Mean±SD	20.00±0.38	7.94±0.05	250.44±0.51	8.30±0.09	1.95±0.15	85.27±0.01	280.52±0.63	392.22±1.17
	(S) Mean±SD	21.41±0.13	7.91±0.07	387.34±0.00	10.78±0.19	43.86±0.02	80.40±0.33	293.67±0.58	716.67±0.00
Hukulganj	(G) Mean±SD	19.50±0.17	7.42±0.08	350.22±0.38	3.21±0.02	2.86±0.03	46.26±0.16	350.26±0.36	620.89±0.84
	(S) Mean±SD	20.86±0.05	7.73±0.07	450.11±0.19	6.04±0.00	39.62±0.00	55.61±0.38	361.45±0.39	780.00±0.00

Statistical analysis

PCA extracts meaningful information from raw data by utilizing the existing correlations among various parameters which are being monitored and hence successfully aid the interpretation of meaningful information (e.g., Chaturvedi and Raghubanshi 2015). Hierarchical CA groups various sites into different clusters based on the similarity in various studied parameters.

The statistical analysis was done using IBM SPSS Statistics software package. PCA technique extracts the eigenvalues and eigen vectors from the covariance matrix of original variables to find association between the variables such that it reduces the dimensionality of the data set. The principal components (PCs) obtained from PCA are the uncorrelated (orthogonal) variables, obtained by multiplying the original correlated variables with the eigenvector (loadings or weightings). The eigenvalues of the PCs are the measure of their associated variance, the participation of the original variables in the PCs is given by the loadings, and the individual transformed observations are called scores (Singh et al. 2005a, b, c) (Arslan 2013). Thus, the PCs are weighted linear combinations of the original variables. These PCs extract meaningful information by capturing maximum variance and hence facilitate the compression of data (Vega et al. 1998; Helena et al. 2000). PCA is designed to transform the original variables into new, uncorrelated variables (axes), called the principal components, which are linear combinations of the original variables. The new axes lie along the directions of maximum variance. So, PCA provides an objective way of finding indices of this type such that the variation in the data can be accounted for as concisely as possible. This purpose can be achieved by rotating the axis defined by PCA, according to well-established rules, and constructing new variables, also called varifactors (VF). PC is a linear combination of observable water quality variables, whereas VF can include unobservable, hypothetical, and latent variables (Vega et al. 1998; Helena et al. 2000). PCA of the normalized variables was performed

Table 3 The PCA axes explaining complete variance of the data

Axis	% of variance	Cum % of variance
1	46.096	46.096
2	18.674	64.770
3	13.698	78.468

to extract significant PCs and to further reduce the contribution of variables with minor significance; these PCs were subjected to varimax rotation (raw) generating VFs. Consequently, a small number of factors will usually account for approximately the same amount of information as does the much larger set of original observations.

In CA, unsupervised pattern of the data is used to uncover intrinsic structure or underlying behavior of a data set without any prior assumption in classification of objects into categories based on similarity (Vega et al. 1998). The most commonly used hierarchical clustering technique is CA in which clusters are formed sequentially, starting with the most similar pair of objects followed by forming higher clusters step by step. The Euclidean distance usually gives the similarity between two samples, and a "distance" can be represented by the "difference" between analytical values from both the samples (Otto 1998). This technique was performed on the normalized data set by means of the Ward's method, using Euclidean distances as a measure of similarity. This method uses the analysis of variance approach to evaluate the distances between clusters, attempting to minimize the sum of squares of any two clusters that can be formed at each step. CA has been applied to assess the quality of river water in relation to its spatial variability at selected sites spread over the river stretch. The linkage distance is reported as Dlink/Dmax=(quotient between the linkage distance for a particular case/ maximal distance)×100 to standardize the linkage distance and represented on y-axis (Wunderlin et al. 2001; Simeonova et al. 2003).

Results and discussion

Principal component analysis

PCA was performed on the combined data set of 3 years (between the sites and parameters) and PCs with eigenvalue >1 were retained. Results from this analysis including the loadings, eigenvalue, and variance were explained by each PC. The extracted three PCs explain 78.5 of the total variance as shown in Table 3, VF1 explains 46 % of total variance, has strong positive loadings of TDS, chloride, and EC both on ground water as well as surface water. VF2 explains 18.7 % of total variance, has strong positive loadings on PO_4^{-3}, on ground and surface water and strong negative loadings

Table 4 Loading of variables on significant principal component of surface and ground water data

variable	VF1	VF2	VF3
TEMP (G)	−0.6615	−0.3687	−0.4092
pH (G)	0.2866	−0.7308	0.3998
TDS (G)	0.9082	0.2919	−0.1262
NO_3^- (G)	0.6797	0.2028	0.0161
PO_4^{3-}(G)	−0.5518	0.7283	−0.1331
Cl^- (G)	0.8962	−0.1536	−0.3264
HARD (G)	−0.2554	0.1119	−0.7608
EC (G)	0.8521	0.4080	0.2224
TEMP (S)	−0.0523	−0.6383	−0.1712
pH (S)	0.8108	−0.2478	−0.2152
TDS (S)	0.7723	0.4783	0.0376
NO_3^- (S)	0.7643	−0.1961	−0.3671
PO_4^{3-} (S)	0.5399	−0.3168	0.7293
Cl^- (S)	0.9445	−0.2203	−0.0352
HARD (S)	−0.3199	0.6286	0.5050
EC (S)	0.7371	0.4224	−0.3095

Significant loading is shown in italic font

on pH, whereas, it has moderate positive loadings on total hardness. VF3 (13.7 % of total variance) has strong negative loadings on total hardness on both ground and surface water as shown in Table 4.

From PCA (Fig. 2), we observed that Chaukaghat, Hakulganj, Nakhighat, and Kachahari are affected by total hardness and $PO_4^{3-}{}_4^{-3}$, whereas, Rajghat site is influenced by TDS, EC, and NO_3^-. The pollutants at these sites are contributed by anthropogenic sources like small dyeing industries, laundries, small scales foundry, battery industries, and haphazard dumping of municipal solid waste along the river (Barai and Kumar 2013).

Clustering of the sites based on spatial similarity

Cluster analysis (CA) was applied to detect spatial similarity among sites under the monitoring network. It rendered a dendrogram as shown in (Fig. 3), grouping

Fig. 2 Result of principal component analysis (PCA) for the site and physiochemical parameters of surface and ground water
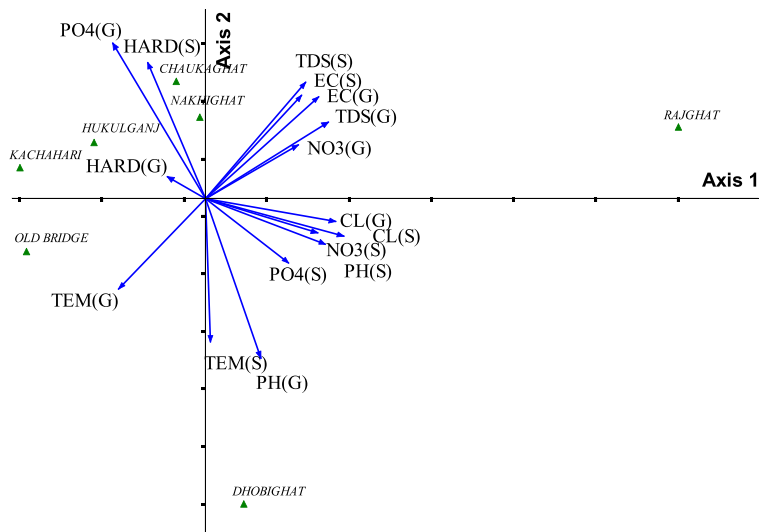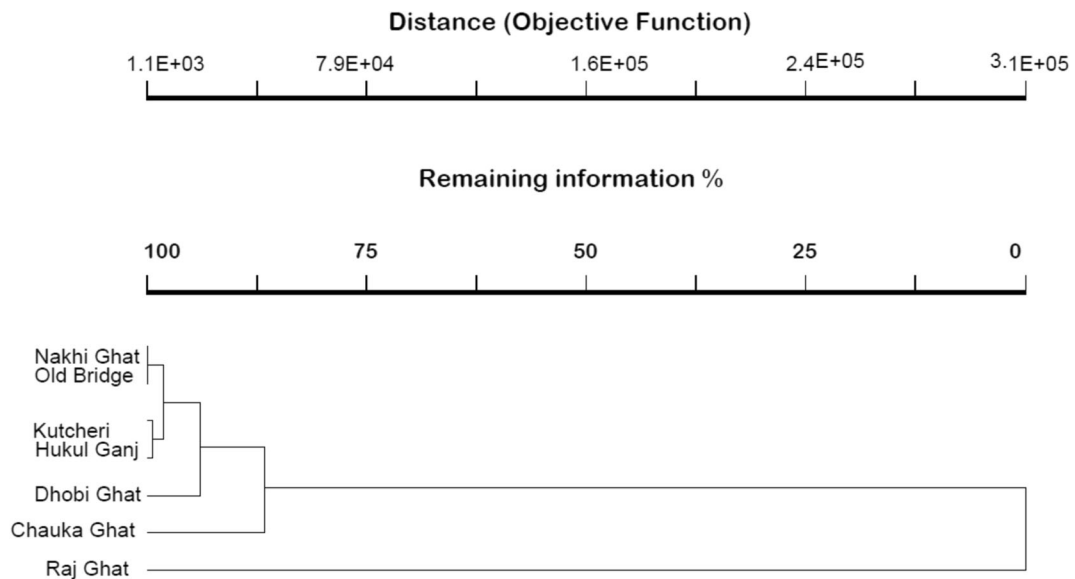
Fig. 3 Dendrogram showing the cluster of sampling sites on Varuna River (linkage method: Ward's method, distance measure: Euclidean (Pythagorean) percent chaining=85.71)

all the seven sampling sites on the river into three statistically significant clusters at $(D\text{link}/D max) \times 100 \leq 75$. The clustering procedure generated three groups of sites in a very convincing way, as the sites in these groups have similar characteristic features and same source of contamination, and their range of correlation coefficient are shown in Table 5. Cluster 1 (Nakhighat, Hukulganj, Old Bridge, Kachahari, Dhobighat), cluster 2 (Chaukaghat), and cluster 3 (Rajghat) are the clusters that correspond to contamination of water. It is evident that this technique is useful in offering reliable classification of surface and ground water in the whole region. So instead of monitoring eight sites, only three sites could be selected in future spatial sampling strategy without affecting the result (Fig. 2). The reduction in sampling sites in the monitoring network will be more economical without losing any significance of the outcome (Singh and Singh 2010; Singh et al. 2005b, c).

Correlation analysis

As the raw data are not normally distributed, the correlation among the eight variables was determined by estimating Spearman's ranked correlation coefficient (r), which is a non-parametric test of the degree of correlation (Shrestha and Kazama 2007). Because of the known association among variables used in the study and the fairly large number of observations (42), the critical value of r, which is 0.259 at $p < 0.05$ was replaced by 0.424 at $p < 0.001$, as the critical value for this study. This is expected to highlight the stronger positive or negative correlations that are likely to be of more utility (Helena et al. 2000). It is also of interest to see how these correlations are reflected in the multivariate statistics, especially PCA, and subsequently in the VFs after varimax rotation of the PCs. Data in Table 4 provide the correlation matrix of the quality parameters

Table 5 The range of correlation coefficient of cluster analysis

	Nakhighat	Chaukaghat	Old Bridge	Rajghat	Kachari	Dhobighat
Chaukaghat	0.9920					
Old Bridge	0.9896	0.9747				
Rajghat	0.9382	0.9562	0.9110			
Kachari	0.9279	0.9143	0.9682	0.8253		
Dhobighat	0.9691	0.9571	0.9897	0.8727	0.9858	
Hukulganj	0.9965	0.9869	0.9957	0.9250	0.9504	0.9808

Table 6 Correlation coefficient for the various physiochemical parameters of surface and ground water

	pH	TDS	NO_3^-	PO_4^{3-}	Cl^-	Hard	EC
TEMP	0.315 ns	−0.145 ns	−0.167 ns	0.542[a]	−0.171 ns	0.020 ns	0.007 ns
pH		0.188 ns	0.056 ns	−0.520[a]	0.271 ns	−0.350[b]	0.354[b]
TDS			0.469[a]	−0.170 ns	0.813[a]	−0.015 ns	0.866[a]
NO_3^-				−0.124 ns	0.472[a]	−0.374[b]	0.465[a]
PO_4^{3-}					0.034 ns	0.437[a]	−0.020 ns
Cl^-						−0.026 ns	0.548[a]
Hard							−0.187 ns

ns $p > 0.05$

[a] $p < 0.01$

[b] $p < 0.05$

obtained from the PCA. Among the various physio-chemical parameters of surface and ground water, only few parameters exhibited significant correlations (Table 6). Strong negative correlations have been observed between pH and PO_4^{3-}, and NO_3^- and hardness (Table 3). We observed significant positive relationship of temperature with PO ($r=0.54$, $p<0.01$). It has been reported that the solubility of calcium phosphate minerals increases with increasing temperature (House 1999). It can implicitly be concluded that the degree of contamination is also dependent on the season of sampling. There exists strong positive correlation of TDS with Cl^- ($r=0.81$, $p<0.01$) and EC ($r=0.87$, $p<0.01$), and a moderate correlation with NO_3^- ($r=0.47$, $p<0.01$) which is clearly comprehensible as TDS accounts for both organic and inorganic impurities. The increase in impurities enhances number of free ions in the water leading to increased EC. Correlation analysis showed the degree of positive and negative correlation between the parameters used for analysis and made necessary conclusions about the interdependence of free ions on EC. NO_3^- was also positively correlated with the EC ($r=0.47$, $p<0.01$). The correlation coefficients of various parameters have been presented in Table 3.

Multivariate statistical techniques (PCA and CA) effectively evaluated the variations in surface water and ground water quality of river Varuna. PCA helped in identifying the factors or sources responsible for water quality variations explaining about 78.5 % of the total variability. CA reduced sampling sites into three groups based on the similarity in the water quality, and this could be helpful in minimizing the sampling location for the future monitoring. The interactions between hydrochemical parameters of ground water and surface water are quite evident by studying loading of variables on significant principal components of ground and surface water.

The previous studies over establishing this kind of interactions were done in different environment but used the same methodology. Mencio and Mas-Pla (2008) performed similar kind of analysis in urbanized Mediterranean streams, and Woocay and Walton (2008) also followed the same methodology and successfully exhibited the interactions which have hypothesized to establish for this study area. Furthermore, different studies have been conducted for analysis of ground water using multivariate statistical techniques but only a few of them proposed for any kind of interactions.

Conclusions

The present study took into account only a particular season spanning across 3 years so as to make the necessary conclusions. The data for different sampling sites collected across the aforementioned time period successfully established an interaction between ground water and surface water. These techniques further aided in finding out relevant parameters for monitoring the task at every site. In conclusion, there is an interaction as hypothesized in our proposal which was exhibited by the statistical technique. Moreover, this analysis will help in future water control management program as it has outlined the parameters contributing to pollution for every site. This will make the future monitoring more economical and also easier to comprehend. It is therefore, needful, to develop a comprehensive river water

quality monitoring program all over the world (Sharma and Kansal 2011).

References

Anibas, C., Buis, K., Verhoeven, R., Meire, P., & Batelaan, O. (2011). A simple thermal mapping method for seasonal spatial patterns of ground water-surface water interaction. *Journal of Hydrology, 397*(1–2), 93–104.

APHA (1998) Standard methods for the examination of water and wastewater, American Public Health Association Washington D.C., Ed.19.

Arslan, H. (2013). Application of multivariate statistical techniques in the assessment of groundwater quality in seawater intrusion area in Bafra Plain, Turkey. *Environmental Monitoring and Assessment, 185*(3), 2439–2452.

Barai, S. R., & Kumar, S. (2013). Evaluation of the physio-chemical characterstics of river Varuna at Varanasi, India. *Journal of Environmental Biology, 34*, 259–265.

Chaturvedi, R. K., & Raghubanshi, A. S. (2015). Assessment of carbon density and accumulation in mono-and multi-specific stands in Teak and Sal forests of a tropical dry region in India. *Forest Ecology and Management 339*, 11–21.

Harveya, W. J., Newlina, T. J., & Krupab, L. S. (2006). Modeling decadal timescale interactions between surface water and ground water in the Central Everglades, Florida, USA. *Journal of Hydrology, 320*, 400–420.

Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J. M., & Fernandez, L. (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research, 34*(3), 807–816.

House, W. A. (1999). The physico-chemical conditions for the precipitation of phosphate with calcium. *Environmental Technology, 20*(7), 727–733.

Hutton, L. G. (1983). Field testing of water in developing countries. Medmenham, England: Water Research Center.

Isa, N. M., Aris, A. Z., & Sulaiman, W. A. N. W. (2012). Extent and severity of groundwater contamination based on hydrochemistry mechanism of sandy tropical coastal aquifer. *Science of the Total Environment, 438*, 414–425.

Kikuchi, C. P., Ferre, T. P. A., & Welker, J. M. (2012). Spatially telescoping measurements for improved characterization of ground water-surface water interactions. *Journal of Hydrology, 446–447*, 1–12.

Kumar, S., Mishra, P.K., Singh, N.L., Singh, K.K. & Srivastava, P. (2012). Water Quality of River Varuna in Varanasi City, Uttarpradesh, India. *Asian Journal of Biochemical and Pharmaceutical Research* (2).

Lu, X., Zhou, Z., & Wang, J. (2011). Analysis on interactions among atmosphere, surface and ground water using EARTH model in the unsaturated zone. *Procedia Environmental Sciences, 8*, 134–139.

Mencio, A., & Mas-Pla, J. (2008). Assessment by multivariate analysis of groundwater-surface water interactions in urbanized Mediterranean streams. *Journal of Hydrology, 352*(3–4), 355–366.

Mishra, A. (2010). Assessment of water quality using principal component analysis: a case study of the river Ganges. *Journal of Water Chemistry and Technology, 32*(4), 227–234.

Mishra, A. K. (2011). Impact of urbanization on the hydrology of Ganga Basin (India). *Water Resources Management, 25*(2), 705–719.

Otto, M. (1998). Multivariate methods. In R. Kellner, J. M. Mermet, M. Otto, & H. M. Widmer (Eds.), *Analytical chemistry*. Weinheim: WileyeVCH.

Pandey, D. S. (1993). Groundwater pollution studies in urban settlements of Varanasi city, UP, Annual work programme report, 1992–1993, (p 35). Allahabad: Central Groundwater Board.

Raju, N. J., Ram, P., & Dey, S. (2009). Groundwater quality in the Lower Varuna River Basin, Varanasi District, Uttar Pradesh. *Journal of the Geological Society of India, 73*(2), 178–192.

Selle, B., Schwientek, M., & Lischeid, G. (2013). Understanding processes governing water quality in catchments using principal component scores. *Journal of Hydrology, 486*, 31–38.

Sharma, D., & Kansal, A. (2011). Water quality analysis of river Yamuna using water quality index in the national capital territory, India (2000–2009). *Applied Water Science, 1*, 147–157.

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin. *Environmental Modelling & Software, 22*(4), 464–475.

Shukla, U. K., & Raju, J. N. (2008). Migration of the Ganga river and its implication on hydro-geological potential of Varanasi area, U.P., India. *Journal of Earth System Science, 117*(4), 489–498.

Simeonova, P., Simeonov, V., & Andreev, G. (2003). Environmetric analysis of the Struma River water quality. *Central European Journal of Chemistry, 2*, 121–126.

Singh, S. J., & Dwivedi, A. K. (2007). Numerical interdependence in pH, acidity and alkalinity of a polluted river water. *Journal of Environmental Biology, 30*(5), 773–775.

Singh, S., & Singh, K. N. (2010). Physio-chemical analysis of sewage discharged into Varunariver at Varanasi. *Current World Environment, 5*(1), 201–203.

Singh, K. P., Malik, A., & Sinha, S. (2005a). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques a case study. *Analytica Chimica Acta, 538*(1–2), 355–374.

Singh, K. P., Malik, A., Singh, V. K., Mohan, D., & Sinha, S. (2005b). Chemometric analysis of groundwater quality data of alluvial aquifer of Gangetic plain, North India. *Analytica Chimica Acta, 550*(1–2), 82–91.

Singh, K. P., Malik, A., Sinha, S., Singh, V. K., & Murthy, R. C. (2005c). Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. *Water, Air, and Soil Pollution, 166*(1–4), 321–341.

Singh, K. P., Gupta, S., & Mohan, D. (2014). Evaluating influences of seasonal variations and anthropogenic activities on alluvial groundwater hydrochemistry using ensemble learning approaches. *Journal of Hydrology, 511*, 254–266.

Sojka, M., Siepak, M., Ziola, A., Frankowski, M., Murat-Blazejwska, S., & Siepak, J. (2008). Application of multivariate statistical techniques to evaluation of water quality in the Mala Welna River (Welna Poland). *Environmental Monitoring and Assessment, 147*, 159–170.

Sophocleous, M. (2002). Interactions between ground water and surface water: the state of the science. *Hydrogeology Journal, 10*, 52–67.

Soulsby, C., Gibbins, C., Wade, A. J., Smart, R., & Helliwell, R. (2002). Water quality in the Scottish uplands: a hydrological perspective on catchment hydrochemistry. *Science of the Total Environment, 294*(1–3), 73–94.

Stigter, T. Y., Carvalho Dill, A. M. M., Ribeiro, L., & Reis, E. (2006). Impact of the shift from groundwater to surface water irrigation on aquifer dynamics and hydrochemistry in a semi-arid region in the south of Portugal. *Agricultural Water Management, 85*, 121–132.

Sundaray, K. S., Panda, C. U., Nayak, B. B., & Bhatta, D. (2006). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of the Mahanadi river-estuarine system (India) a case study. *Environmental Geochemistry and Health, 28*, 317–330.

Valett H.M., & Sheibley, R.W. (2009).Ground water and surface water interactions. Reference Module in Earth Systems and Environmental Sciences, from Encyclopedia of Inland Waters, 691–702.

Varol, M., Gokot, B., Bekleyen, A., & Sen, B. (2013). Geochemistry of the Tigris River basin, Turkey: spatial and seasonal variations of major ion compositions and their controlling factors. *Quaternary International, 304*, 22–32.

Vega, M., Pardo, R., Barrado, E., & Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research, 32*(12), 3581–3592.

Wade, J. A., Butterfield, D., & Whitehead, G. P. (2006). Towards an improved understanding of the nitrate dynamics in lowland, permeable river-systems: applications of INCA-N. *Journal of Hydrology, 330*(1–2), 185–203.

Woocay, A., & Walton, J. (2008). Multivariate analysis of water chemistry: surface and ground water interactions. *Ground Water, 46*(3), 437–449.

Wunderlin, D. A., Diaz, M. P., Ame, M. V., Pesce, S. F., Hued, A. C., & Bistoni, M. A. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). *Water Research, 35*, 2881–2894.

Yu, Y. J., Guan, J., Ma, Y. W., Yu, S. X., Guo, H. C., & Bao, L. Y. (2010). Aquatic environmental quality variation in Lake Dianchi watershed. *Procedia Environmental Sciences 2*, 76–81.