CrossMark

ORIGINAL ARTICLE

Assessment of groundwater quality of Lakshimpur district of Bangladesh using water quality indices, geostatistical methods, and multivariate analysis

Mohammad Amir Hossain Bhuiyan[1,6] · Md. Bodrud-Doza[1] · A. R. M. Towfiqul Islam[2,3] · M. A. Rakib[3] · M. Safiur Rahman[4,5] · A. L. Ramanathan[6]

Abstract Groundwater evaluation indices, multivariate statistical techniques, and geostatistical models are applied to assess the source apportionment and spatial variability of groundwater pollutants at the Lakshimpur district of Bangladesh. A total of 70 groundwater samples have been collected from wells (shallow to deep wells, i.e., 10–375 m) from the study area. Groundwater quality index reveals that 50 % of the water samples belong to good-quality water. The degrees of contamination, heavy metal pollution index, and heavy metal evaluation index present diversified results in samples even though they show significant correlations among them. The results of principal component analysis (PCA) show that groundwater quality in the study area mainly has geogenic (weathering and geochemical alteration of source rock) sources followed by anthropogenic source (agrogenic, domestic sewage, etc.). Cluster analysis and correlation matrix also supported the results of PCA. The Gaussian semivariogram models have been tested as the best fit models for most of the water quality indices and PCA components. The results of semivariogram models have shown that most of the variables have weak spatial dependence, indicating agricultural and residential/domestic influences. The spatial distribution maps of water quality parameters have provided a useful and robust visual tool for decision makers toward defining adaptive measures. This study is an implication to show the multiple approaches for quality assessment and spatial variability of groundwater as an effort toward a more effective groundwater quality management.

Introduction

Metals' contamination of groundwater is of great concern on lives owing to their toxicity, persistence, and extensive bioaccumulation. Groundwater is an important resource for agriculture, industrial, and other economic sectors in Bangladesh. Rapid urbanization, agricultural, and industrial activities are affecting groundwater day by day. A wide range of public health issues such as cancer, hypertension, hyperkeratosis, peripheral vascular disease, restrictive lung disease, and gangrene occurs due to the consumption of contaminated water (Smith et al. 2000). Approximately 17 % of groundwater in Bangladesh exhibit arsenic (As) concentrations beyond the acceptable limit (10 μg/L) of DoE (1997) for drinking water.

The interpretation of water quality data sets for pollution evaluation is quite difficult by only elemental concentrations (Nimic and Moore 1991). However, WQIs have huge scope to analyze the data sets for better interpretation of pollution. There exists a wide range of WQIs; however, the choice

✉ Mohammad Amir Hossain Bhuiyan
  amirhb75@yahoo.com; amir@juniv.edu;
  amirhossain75@gmail.com

1 Department of Environmental Sciences, Jahangirnagar University, Dhaka 1342, Bangladesh

2 School of Atmospheric Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China

3 Department of Disaster Management, Begum Rokeya University, Rangpur 5400, Bangladesh

4 Department of Environmental Engineering, Dalhousie University, Halifax, NS B3J 1Z1, Canada

5 Chemistry Division, Atomic Energy Center, Dhaka, Bangladesh

6 School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110067, India

depends upon the input variables and the desired results (Handa [1981]; Zou et al. [1988]; Sahu et al. [1991]; Li et al. [2009]). Due to some limitations, WQI values provide better results together with the chemometric techniques. It has been found that chemometric methods are the most reliable approaches for data mining of matrices from environmental quality assessment (Astel et al. [2007], [2008]). Among the available chemometric methods, multivariate statistical analysis has been widely used for source apportionment of metals in soil and water in different parts of the world (e.g., Singh et al. [2005]; Halim et al. [2010]; Bhuiyan et al. [2010]; Li et al. [2013]; Machiwal and Jha [2015]).

On the other hand, geostatistical method has gained importance to evaluate the spatial distribution of pollutant in soils and water. It is also an important tool for spatial dependence/autocorrelation among the sampling points. Subsequently, this type of information is important in estimating the pollutant migration history and spatial distribution of the pollutant at different sites. However, the integrated approaches of multivariate analysis and geostatistics may provide a holistic approach on the complex pollution system.

As the spatial distribution of heavy metal contamination in groundwater is controlled by the geological/geochemical heterogeneity, the spatial interpolation technique has been used to estimate the concentration at unmeasured locations and devise points to show groundwater contamination (Webster and Oliver [2001]). Detailed and extensive explanations of geostatistical method have been reported in different literatures (Isaaks and Srivastava [1989]; Goovaerts [1997]; Webster and Oliver [2001]). The cross-validation results from geostatistics represent that the ordinary kriging technique can predict spatial variability more accurately. The ordinary kriging method deals with spatial correlations between the sample points and has been widely used for mapping spatial variability of elements. An assessment of drinking and irrigation water quality is very essential for understanding the suitability of groundwater for different purposes. In the study area, a limited work has been conducted on groundwater quality. Hence, the integrated approaches of different chemometric methods are considered as important tool for pollution evaluation in this study area. Considering all these aspects, Lakshimpur district of Bangladesh has been selected as the study area for a comprehensive study using the integrated approaches of multivariate analysis and geostatistical methods.

Materials and methods

Study area

Lakshimpur Sadar upazila (a subdistrict, a small administrative unit), located in southeastern Lakshimpur district of Bangladesh, has been selected for this study. Geographically, the study area is positioned between 22°49′–22°03′N and 90°43′–92°00′E (Fig. [1]). It is bounded by Raipur, Ramganj, and Chatkhil upazilas on the north; Daulatkhan, Kamalnagar, and Noakhali Sadar upazilas on the south; Begumganj and Sonaimuri upazilas on the east; and Raipur upazila and Meghna River on the west. Lakshimpur upazila (subdistrict) has an area of 514.78 sq km with a total population of 575,278 (Banglapedia [2006]). The sites are chosen mainly based on their proximity of suspected pollution sources and ecological and environmental importance. Dalal Bazar, Parbatinagar, Dattapara, Hajipara, Jacksinhat sites are densely populated areas. The groundwater quality at Mazuchowdhurihat and Shackchar (western part of the area) is highly dominated by Meghna River. Physiographically, it is a coastal floodplain that experiences tide actions regularly. About 74 % of this area is under water supply coverage provided by several local and national NGOs.

Sample collection and preparation

Groundwater samples are collected from 70 preselected sampling points at the Sadar upazila of Lakshimpur district of Bangladesh (Fig. [1]). The sampling locations are recorded by a GPS device (Explorist model: 200). The information regarding well depths is collected from the record preserved by the well owners and local government offices. Three types of tubewells such as (1) shallow wells (10–60 m depth), (2) deep wells (80–375 m depth), and (3) dug wells (7–318 m) have been selected on the basis of the availability at the study area. Samples are collected in pre-washed high-density polypropylene (HDPP) bottles following the standard method of APHA-AWWA-WEF ([2005]). For metal ions and dissolved organic carbon (DOC) analysis, water samples are preserved following the standard procedures of Rahman and Gagnon ([2014]). All analytical procedures of groundwater samples are conducted following the standard methods (Table [1]). The accuracy and precision of analysis are tested through running duplicate analysis on selected samples, and the average results for all analyses are used to represent the data.

Groundwater pollution evaluation indices

Groundwater quality index (GWQI)

Groundwater quality index (GWQI) method reflects the composite influence of the different water quality parameters on the suitability for drinking purposes (Sahu and Sikdar [2008]). The groundwater quality has been measured by using the following equation for GWQIs Vasanthavigar

Fig. 1 Location map showing the sample sites in the study area

Table 1 List of chemical elemental analysis, methods, and equipments

Variables	Method	Units	Equipment
pH	pH meter		Accumet electrode and Accumet Excel, XL50 (Dual channel pH/ion/conductivity) meter
EC	Conductivity meter		Accumet electrode and Accumet Excel, XL50 (Dual channel pH/ion/conductivity) meter
DOC	Total organic carbon analyzer		TOC-VCHP Total Organic Carbon Analyzer (Shimadzu, Japan)
HCO3	Field titration		HACH digital titrator (HACH 1690, USA)
Cl, F, SO4, NH$_4$–N, Na, K, Ca, Mg	Chromatography		Ion chromatograph (761 Compact IC, Metrohm)
As, Pb, Fe, Mn, Ni, Sb, Ba, Mo, Al, Zn, B, P, Si	Inductively coupled plasma mass spectrometry		Thermo Scientific X-Series2 ICP-MS

et al. (2010) with respect to WHO (2011) and Bangladesh standards (1997).

$$\text{GWQI} = \sum \text{SI}_i$$
$$= \sum (W_i \times q_i) = \sum \left(\left(\frac{w_i}{\sum_{i=1}^{n} w_i} \right) \times \left(\frac{C_i}{S_i} \times 100 \right) \right)$$

where C_i is the concentration of each parameters, S_i is the limit values, w_i is the assigned weight according to its relative importance in the overall quality of water for drinking purposes (Table 2), q_i is the water quality rating, W_i is the relative weight, and SI_i is the subindex of ith parameter. The heavy metal pollution index (HPI) method

has been developed by assigning rating or weightage (W_i) for each chosen parameter (As, Pb, Fe, Mn, Ni, Sb, Ba, Mo, Al, Zn) and selecting the groundwater parameter on which the index has to be based on (Bhuiyan et al. 2010). The rating is nearly zero to one, and its selection reveals the significance of each water quality parameter. It has been developed based on the monitored values, ideal values, and recommended standard values of the studied parameters. It can be defined as inversely proportional to the recommended standard (S_i) for each parameter (Horton 1965; Reddy 1995; Mohan et al. 1996). The concentration limits (i.e., the highest permissible value for drinking water (S_i) and maximum desirable value (I_i) for each parameter) are taken from the Indian drinking water specification standards of 2012 (BIS 2012) for this study. Heavy metal pollution index (HPI) has been used for assigning rating or weightage (W_i) for each selected parameter and can be computed using the following expression (Mohan et al. 1996; Bhuiyan et al. 2015):

$$\text{HPI} = \frac{\sum_{i=1}^{n} W_i Q_i}{\sum_{i=1}^{n} w_i}$$

where Q_i is the subindex of the ith parameter, W_i is the unit weight of the ith parameter, and n is the number of parameters. The subindex Q_i is computed by

$$Q_i = \sum_{i=1}^{n} \frac{\{M_i (-) I_i\}}{(S_i - I_i)} \times 100$$

where M_i, l_i, and S_i stand for the monitored values, ideal values, and standard values of the ith parameter, respectively. The negative sign ($-$) denotes numerical difference in the two values ignoring algebraic sign.

Heavy metal evaluation index (HEI) method provides an insight into the overall quality of the groundwater with respect to heavy metals and metaloids (As, Pb, Fe, Mn, Ni, Sb, Ba, Mo, Al, Zn (Edet and Offiong 2002). It has been calculated by Prasad and Jaiprakas (1999) as follows:

$$\text{HEI} = \sum_{i=1}^{n} \frac{H_c}{H_{\text{mac}}}$$

where H_c is the monitored value and H_{mac} is the maximum admissible concentration (MAC) of ith parameter.

The degree of contamination (C_d/CD) has been adopted from Backman et al. (1997). Prasad and Bose (2001) evaluate the combined effects of several quality parameters which are considered detrimental to household water. The CD/C_d is determined by:

$$C_d = \sum_{i=1}^{n} C_{f_i}$$

where

$$C_{f_i} = \frac{C_{ai}}{C_{ni}} - 1$$

C_{fi} is the contamination factor, C_{ai} is the analytical value, and C_{ni} is the upper permissible concentration for the ith component and n indicates the normative value. Here, C_{ni} is taken as maximum admissible concentration (MAC).

Multivariate statistical analysis

Principal component analysis (PCA) reduces the dimensionality of data by a linear combination of original data to generate new latent variables which are orthogonal and

Table 2 List of parameters, weight factors, and limit values for the water quality index after Vasanthavigar et al. (2010)

Parameters	Units	Weight (w_i)	Relative weight (W_i)	Limit values
pH		4	0.078	6.5–8.5
HCO$_3$	mg/L	1	0.020	600
Cl	mg/L	3	0.059	250
F	mg/L	4	0.078	1.5
SO$_4$	mg/L	4	0.078	400
NH$_4$–N	mg/L	3	0.059	0.5
Na	mg/L	4	0.078	200
K	mg/L	2	0.039	12
Ca	mg/L	2	0.039	75
Mg	mg/L	2	0.039	30
As	µg/L	4	0.078	50
Pb	µg/L	4	0.078	10
Fe	µg/L	4	0.078	1000
Mn	µg/L	4	0.078	300
Al	µg/L	3	0.059	200
Zn	µg/L	3	0.059	5000
		$\sum w_i = 51$	$\sum W_i = 1$	

uncorrelated to each other (Nkansah et al. 2010). It extracts the eigenvalues and eigenvectors from the covariance matrix of original variables (Chabukdhara and Nema 2012). The eigenvalues of the PCs are the measure of their associated variance, the participation of original variables in the PCs is given by the loadings, and the coordinates of the objects are called scores (Helena et al. 2000; Wunderlin et al. 2001; Heberger et al. 2005). PCA provides an objective way of finding indices of this type so that the variation in the data can be accounted for as concisely as possible (Sarbu and Pop 2005). PCA has been performed to extract principal components (PC) from the sampling points and to evaluate spatial variations and possible sources of heavy metals in groundwater.

Factor analysis is similar to PCA except for the preparation of observed correlation matrix for extraction and the underlying theory (Tabachnick and Fidell 2007; Bhuiyan et al. 2011a). The goal of FA can be achieved by rotating the axis defined by PCA, according to the well-established rules, and constructing new variables, also called varifactors (VF) (Shrestha and Kazama 2007).

The correlation coefficient matrix measures how well the variance of each constituent can be explained by relationship with each other (Liu et al. 2003). According to the approach of Liu et al. (2003), the terms "strong," "moderate," and "weak" are applied to factor loadings and refer to absolute loading values as >0.75, 0.75–0.50, and 0.50–0.30, respectively.

The cluster analysis (CA) is applied to identify groups or clusters of similar sites on the basis of similarities within a class and dissimilarities between different classes (Lattin et al. 2003). Hierarchical cluster analysis (HCA) studies distance between parameters of samples. The most similar points are grouped forming one cluster, and the process is repeated until all points belong to one cluster (Danielsson et al. 1999; Birth 2003). The result obtained is shown in a 2D plot called dendrogram. In the study, Ward's method with squared Euclidean distances is used. The experimental groundwater data were subjected to statistical analysis using SPSS software (version 22.0). Pearson's correlation matrix is used to identify the relationship among the pairs of parameters.

Geostatistical modeling

Ordinary kriging (OK) and semivariogram models are applied for spatial distribution of groundwater parameters which are related to groundwater application in hydrological studies. These interpolation techniques are well documented in the recent literature (e.g., Masoud and Atwia 2010; Ahmed et al. 2011; Masoud 2014; Marko et al. 2014). Kriging is one of the most popular and robust interpolation techniques among other techniques. It integrates both spatial correlation and the dependence in the prediction of a known variable. Estimations of nearly all spatial interpolation methods can be represented as weighted averages of sampled data. The equation can be written as follows (Delhomme 1978):

$$\hat{z}(x_o) = \sum_{i=1}^{n} \lambda_i z(x_i)$$

where $\hat{z}$ is the estimated value of an attribute at the point of interest x_o, z is the observed value at the sampled point x_i, λ_i is the weight assigned to the sampled point, and n represents the number of sampling points used for the estimation (Webster and Oliver 2001). The attribute is usually called the primary variable, especially in geostatistics. The semivariance can be estimated from groundwater data by the following equation:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n} \left[z(x_i) - z(x_i + h)^2 \right]$$

where n is the number of pairs of sample points separated by the standard distance calls lag h (Burrough and McDonnell 1998), and $z(x_i)$ is the value of variable z at location x_i. Variogram modeling and estimation are important for structural analysis and spatial interpolation. Among the different kriging techniques, OK has been used in this study because of its easy calculation and prediction accuracy compared to the other kriging methods (Gorai and Kumar 2013). Recently, different variograms or semivariogram models such as linear, exponential, and spherical models are very popular worldwide for spatial analysis of geochemical data sets (Kitanidis 1997; Elogne et al. 2008; Varouchakis and Hristopulos 2013). In this study, circular, spherical, exponential, and Gaussian models have been used to measure spatial autocorrelation or dependence of the groundwater data. The best fit theoretical semivariogram models are prepared based on selecting the trial-and-error basis. Predictive performances of fit models are checked on the basis of cross-validation tests (Gorai and Kumar 2013). The mean error (ME), mean square error (MSE), root mean error (RMSE), average standard error (ASR), and root mean square standardized error (RMSSE) values are assessed to establish the fit models performance. Hu et al. (2004) have discussed several criteria for using error measurements to judge the performance of spatial interpolation methods. Models attain the best goodness-of-fit results in minimum mean error (ME), root mean error (RME), and mean squared error (MSE), attain root mean squared error (RMSE) and average squared error (ASE) close to unity and are considered as the best fit models performance (ESRI 2009). The errors are estimated by the following equations:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)$$

$$RMSE = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2\right]}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(ps_i - os_i)$$

$$RMSSE = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(ps_i - os_i)^2\right]}$$

$$ASE = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(p_i - \left(\sum_{i=1}^{n}p_i\right)\Big/n\right)^2\right]}$$

where n is the number of observation points or samples; o and p are the observed value and predicted or estimated values at location i; os is the standardized observed value; ps is the standardized predicted value. After completing the cross-validation process, kriging offers graphical representation of the distribution of groundwater quality. In this study, Arc GIS (10.2 version) has been used for this interpolation technique.

Results and discussion

Groundwater quality

GWQI values are calculated using different International Standards and BMAC values to determine the suitability of groundwater quality for drinking purposes (Table 3). The GWQI values range from 24.73 to 430.94 with the mean of 113.56. The critical limit (100) for drinking water purposes has been proposed by Vasanthavigar et al. (2010). The results in Table 4 show that 50 % of samples exceed the critical limit (100) of GWQIs. Among the total samples, 17 % of samples belong to excellent water quality and 33 % represent good water quality, 40 % exhibit poor water quality, 7 % of water is of very poor quality, and the rest of 3 % indicate unsuitable water for drinking purposes (Table 4).

The degree of contamination (CD) has been used for estimating the extent of metal pollution (Al-Ami et al. 1987; Bhuiyan et al. 2010). Table 5 shows that range and mean values of CD for groundwater samples are 0.17–66.23 and 11.17, respectively. CD may be categorized (Backman et al. 1997; Edet and Offiong 2002) as follows: low (CD < 1), medium (CD = 1–3), and high (CD > 3). The heavy metal pollution indices (HPI) are computed using the International Standard (BIS 2012)

values of metal content in groundwater samples. The range and mean values of HPI in groundwater samples are 2.19–59.01 and 10.41, respectively. Among the total samples, 77 % of the samples fall above the critical values, i.e., CD 3. According to Edet and Offiong (2002), most of the samples are considered as highly polluted water. Besides the CD values, the water samples are further analyzed by HPI and HEI methods to compare with the results of CD. However, HPI and HEI values for all sample locations fall below the critical values prescribed for drinking purposes (Table 5). The heavy metal evaluation index (HEI) is used to synchronize the criteria for various pollution indices. The HEI criteria for groundwater samples are thus classified as low (HEI < 40), medium (HEI = 40–80), and high (HEI > 80). It is observed that groundwater in the study area exhibits low level of pollution. The results of GWQIs, CD, HPI, and HEI methods show more or less similar trends for most of the samples (Fig. 2). The GWQI values have shown higher spatial variation, whereas HEI values have depicted lower variation. We have also assessed the relationship between metal concentration with the computed indices (GWQIs, CD, HPI, and HEI in Table 5). The GWQIs show positive significant correlations with CD, HPI, and HEI. As, Fe, Mn, and Ni show strong positive correlation (Table 6) with the indices values indicating the metals are the major factors for the pollution in this region.

Source identification of groundwater pollutants

Principal component analysis (PCA) is used for source identification of heavy metals following the standard procedures (Dragović et al. 2008; Franco-Uría et al. 2009). Varimax rotation is used to maximize the sum of variance of the factor coefficients which better explains the possible groups/sources that influence water systems (Gotelli and Ellison 2004). Six factors with eigenvalues >1 are extracted for groundwater data sets which represent 80.61 % of the total variance. The scree plot is used to identify the number of PCs to be retained to understand the underlying parameters' structure (Fig. 3a). The calculated factor loadings together with cumulative percentage and percentages of variance are explained by each factor as listed in Table 7. The positive and negative scores in PCA indicate that most of the water samples are either essentially affected or unaffected by the presence of extracted loads on a specific factor/component, respectively. About 58.38 % of the total variance is represented in the first three loading factors (Fig. 3b). In this study, PC1, PC2, PC3, PC4, PC5, and PC6 explain more than 26, 17, 14, 8, 7, and 6 % of the total variance, respectively.

The first principal component (PC1) in the groundwater data sets explains more than 26.87 % of the total variance. It is loaded with Na, EC, Cl, B, Mg, K, Ca, SO_4, and

Table 3 Descriptive statistics of physiochemical parameters and heavy metals in the study area

Parameters	Minimum	Maximum	Mean	SD	WHO limits (2011)	BMAC (1997)	Acceptable limit of Indian (2012)	Permissible limit of Indian (2012)
pH	6.53	7.59	7.0369	0.24661		6.5–8.5	6.5–8.5	
EC (μs/cm)	384	4640	1135.086	954.7244		700[a]		
DOC (mg/L)	BDL	14	3.5386	3.06104				
HCO_3^- (mg/L)	125	1140	430.18	217.5315		600[a]		
Cl^- (mg/L)	2.7	1550	227.1904	364.6709	250	150–600	250	1000
F^- (mg/L)	0.08	0.45	0.2149	0.07277	1.5		1	1.5
SO_4^{2-} (mg/L)	BDL	265	16.1429	45.05423		400	200	400
NH_4–N (mg/L)	BDL	17.8	1.9434	2.73216		0.5	0.5	
Na_+ (mg/L)	7.9	1090	159.7871	225.6383		200		
K^+ (mg/L)	3.1	41.3	10.8943	7.75802		12		
Ca^{2+} (mg/L)	16.6	186	55.7614	34.72118		75	75	200
Mg^{2+} (mg/L)	15.7	154	46.1486	31.55469		30–35	30	100
As (μg/L)	BDL	404	85.5443	98.45996	10	50	10	50
Pb (μg/L)	0.04	9.23	0.3667	1.09535	10	50	10	
Fe (μg/L)	122	24,800	3235.286	3868.188		300–1000	300	
Mn (μg/L)	51	3830	652	583.4879		100	100	300
Ni (μg/L)	BDL	6.8	1.9086	1.23038	70	100	20	
Sb (μg/L)	BDL	1.18	0.0673	0.15305	20			
Ba (μg/L)	5	657	77.1857	107.0766	700	10	700	
Mo (μg/L)	BDL	20	2.9157	3.67043			70	
Al (μg/L)	BDL	93	6.3143	12.61646		200	30	200
Zn (μg/L)	2	82	21.6429	19.36756		5000	5000	15,000
B (mg/L)	BDL	0.82	0.1716	0.1799	2.4	1	0.5	1
P (mg/L)	BDL	3.4	0.7729	0.78407		0		
Si (mg/L)	8.37	31.6	17.4703	5.55114				

BDL below detection limit, *BMAC* Bangladesh maximum admissible concentration

[a] FAO standard for irrigation purposes

HCO_3. This factor was brought under the purview of many natural hydrogeochemical evolution of groundwater through groundwater–geological medium interaction (Omo-Irabor et al. 2008). PC2, explaining 17 % of total variance, is loaded with P, As, DOC, NH_4–N, Mo, Fe, and HCO_3 (Table 5). PC2 explains leaching of materials from soil horizon to the aquifer which are basically trace elements and are regarded as nonpoint pollutant sources along with partial natural weathering processes. Arsenic comes under this category, i.e., geogenic processes augmented through human-induced activities such as fertilizers and animal feedings (Bhuiyan et al. 2011b). R-mode cluster analyses thus are influenced by both geogenic and anthropogenic activities. Excessive presence of NH_4–N has found to contribute by the use of chemical fertilizer in agricultural fields (Backman et al. 1998). PC3, accounting for 14 % of total variance, is associated with Fe, Ba, and Si, indicating geogenic factors. PC4 is moderately loaded with Mn, accounting for 8.5 % of total variance and is related to geogenic factors. Generally, Mn occurs naturally

as a mineral from sedimentary rocks or from mining and industrial waste products (Bhuiyan et al. 2010). Bacterial activities on Fe and Mn are also responsible for releasing Fe and Mn (Bromfield 1978; Bhattacharya et al. 2002; Naujokas et al. 2013; Islam et al. 2013). PC5 has strong loadings of Al and Pb with 8 % of total variance, indicating anthropogenic pollution from domestic and agricultural sources. Natural causes like geogenic process along with anthropogenic activities (like industrialization) control these enrichments and are shown in high scores of PC5 and S12 and S63 samples. PC6 has moderate to strong loadings of Ni and Sb which is linked to release from stainless steel and alloy product industries.

Spatial similarities and sampling site grouping

In this study, the CA results strongly agree with the PCA results. The R-mode CA retains four main clusters for analyzed parameters (Fig. 4a). Elements belonging to the same cluster are likely to have originated from a common

Table 4 Pollution potential of groundwater samples of the study area based on different quality indices

Index methods	Index values/categories	Degree of pollution	Number of samples	% of samples	Samples
GWQI	50–100	Good water	23	33	S1, S3,S7,S13, S14, S16, S21, S32, S35–S39, S41, S43, S46, S49, S53, S56, S62, S66, S68, S69
	100.1–200	Poor water	28	40	S5, S11, S12, S15, S19, S20, S22–S26, S28, S29, S33, S34, S40, S42, S44, S50–S52, S55, S58, S59, S63, S65,S67, S70
	200.1–300	Very poor water	5	7	S6, S8, S45, S60, S64
	>300	Water unsuitable for drinking purposes	2	3	S10, S18
C_d	<10	Low	42	60	S1–S4, S7, S9, S13, S14, S16, S17, S19, S21, S25, S27, S30–S32, S35–S39, S41, S43, S46–S51, S53–S57, S61, S62, S65–S69
	10–20	Medium	17	24.29	S5, S11, S12, S15, S20, S22, S24, S26, S33, S34, S40, S44, S52, S58, S59, S63, S70
	>20	High	11	15.71	S6, S8, S10, S18, S23, S28, S29, S42, S45, S60, S64
HPI with As	<25	Low	47	67.14	S1, S2, S4, S7, S9, S13, S14, S17, S20, S21, S24–S32, S34–S44, S47–S50, S53, S54, S56–S59, S61, S62, S65, S66, S68–S70
	25–50	Medium	13	18.57	S3, S5, S12, S16, S18, S19, S33, S46, S51, S55, S60, S64, S67
	>50	High	10	14.29	S6, S8, S10, S11, S15, S22, S23, S45, S52, S63
HPI without As	<10	Low	42	60	S1–S5, S7, S9, S13–S17, S21–S22, S26–S27, S30–S37, S39, S41, S46–S50, S53–S57, S59, S61–S62, S66, S69–S70
	10–20	Medium	24	34.29	S6, S8, S11–S12, S18–S20, S23–S25, S28–S29, S38, S40, S42–S45, S52, S58, S60, S65, S67–S68
	>20	High	4	5.71	S10, S51, S63–S64
HEI with As	<10	Low	53	75.71	S1–S5, S7, S9, S12–S17, S19, S21, S22, S25–S32, S34–S39, S41–S43, S46–S50, S53–S57, S59–S62, S65–S70
	10–20	Medium	16	22.86	S6, S8, S11, S18, S20, S23, S24, S33, S40, S44, S45, S51, S52, S58, S63, S64
	>20	High	1	1.43	S10
HEI without As	<5	Excellent water	12	17	S2, S4, S9, S17, S27, S30, S31, S47, S48, S54, S57, S61
	<5	Low	39	55.71	S1–S4, S9, S13–S17, S21–S22, S26–S27, S29–S32, S34, S36–S37, S41–S42, S46–S48, S50, S53–S57, S59, S61–S62, S67–S70
	5–10	Medium	21	30	S5–S7, S11–S12, S19, S23, S25, S28, S33, S35, S38–S39, S43–S45, S49, S51, S60, S65–S66
	>10	High	10	14.29	S8, S10, S18, S20, S24, S40, S52, S58, S63–S64

Table 5 Groundwater pollution indices

Sample ID	GWQI	CD	HPI with As	HPI without As	HEI with As	HEI without As	Sample ID	GWQI	CD	HPI with As	HPI without As	HEI with As	HEI without As	Sample ID	GWQI	CD	HPI with As	HPI without As	HEI with As	HEI without As
S1	84.4	6.16	8.43	7.83	3.87	3.61	S31	35.3	0.17	4.72	3.2	1.65	1.64	S61	49	2.22	**19.4**	5.42	3.8	2.19
S2	27.4	0.3	4.25	2.51	1.32	1.31	S32	51.2	1.43	5.35	4.72	1.62	1.32	S62	64.3	3.19	**20.6**	7.89	**5.47**	3.95
S3	78.4	5.46	**36.4**	8.48	6.59	3.59	S33	116	10.1	**28.1**	9.94	10.2	8.17	S63	139	10.8	67	59	12.2	10.7
S4	42.7	1.66	4.89	3.22	3.01	3	S34	110	10.5	**22**	7.05	4.95	3.23	S64	286	36.6	**43.1**	34.9	17.9	16.6
S5	109	10.2	45.2	9.43	9.14	5.36	S35	90.3	6.93	6.29	5.95	6.85	6.84	S65	100	7.34	16.5	12	6.63	5.9
S6	257	35.5	*81.5*	10.2	15.4	8.13	S36	65.4	2.35	**12.8**	5.95	3.62	2.7	S66	98.6	8	12.8	8.52	7.37	6.68
S7	79.5	5.9	15.2	9.45	8.07	7.23	S37	92.8	7.6	**17**	6.98	3.54	2.3	S67	109	7.64	**26.2**	11	**6.55**	4.78
S8	230	27.1	**93.3**	17.2	19.5	11.7	S38	77.6	4.51	17.5	11.7	7.35	6.49	S68	56.5	3.26	12.5	12.3	3.91	3.72
S9	28	0.4	4.17	2.19	1.46	1.45	S39	87.3	5.88	10.8	9.7	7.23	7.22	S69	53.3	3.55	4.34	2.97	2.93	2.89
S10	394	51.5	*81.2*	30.6	34.4	29	S40	149	14.5	13.5	12.6	13.7	13.7	S70	107	10	**23**	6.16	**4.01**	2.12
S11	134	13.8	**58.7**	11	11.2	6.28	S41	84.7	5.53	**18.2**	9.8	5.25	4.15	Min	24.7	0.17	4.17	2.19	1.11	1.1
S12	137	13.8	**36.7**	11.3	9.29	6.51	S42	171	24.7	14.9	14.1	4.35	4.34	Max	431	66.2	93.3	59	34.4	29
S13	96	7.25	20.2	5.99	4.09	2.45	S43	91.2	6.77	**33.6**	15.1	**9.01**	6.87	Mean	114	11.2	26.1	10.4	7.44	5.73
S14	87.9	7.37	20.7	6.57	4.11	2.47	S44	123	10.3	29.1	19.5	10.7	9.37	CV	100	> 3	100	100		
S15	114	10.4	**57.5**	8.02	8.55	3.43	S45	221	27.9	**64.6**	11.2	**14.6**	9.05	% exceeding CV	50 %	77 %	0 %	0 %		
S16	85.9	6.53	**38.7**	7.69	5.83	2.53	S46	76	5.93	**25.4**	4.55	**5.09**	2.81							
S17	48.3	2.91	6.21	4.65	4.12	4.11	S47	27.7	0.49	5.07	3.41	1.64	1.63							
S18	431	66.2	**47.1**	13.5	16.5	12.9	S48	32.1	0.84	4.91	3.08	2.05	2.04							
S19	111	8.71	**29.8**	15.5	9.47	7.75	S49	78.3	5.47	6.48	4.94	5.51	5.5							
S20	139	13.1	14.6	13.7	12.8	12.7	S50	103	8.98	**18.8**	8.24	4.35	3.05							
S21	56.2	2.44	**9.07**	3.66	1.99	1.23	S51	114	9.22	**49**	24.1	11.6	8.77							
S22	127	13.2	**54.2**	3.06	6.79	1.55	S52	162	17.8	**54.1**	18	**15.5**	11.6							
S23	185	20	**90.3**	10.9	16.2	8.15	S53	65.1	4.14	6.76	5.25	4.37	4.36							
S24	136	14.2	14.7	13.6	11.5	11.3	S54	49.2	1.29	**18**	9.83	4.11	3.03							
S25	116	9.82	**24.6**	10.2	7.08	5.38	S55	102	8.48	**36.2**	7.69	**6.86**	3.8							
S26	105	10.2	**24**	6.54	3.83	1.87	S56	58.3	2.67	**22.8**	8.43	**5.34**	3.66							
S27	43.8	2.17	5.79	4.17	3.54	3.53	S57	43.2	0.93	7.89	4.08	2.22	1.61							
S28	173	23.5	20.2	19.3	7.57	7.13	S58	151	16.9	12.3	11.3	10.1	10.1							
S29	180	26.4	13.5	12.5	4.83	4.82	S59	118	10.6	**16**	8.06	4.77	3.72							
S30	24.7	0.2	4.46	2.74	1.11	1.1	S60	279	39.8	**33.5**	14.5	**9.07**	6.89							

Bold and italic numbers indicate relatively higher values
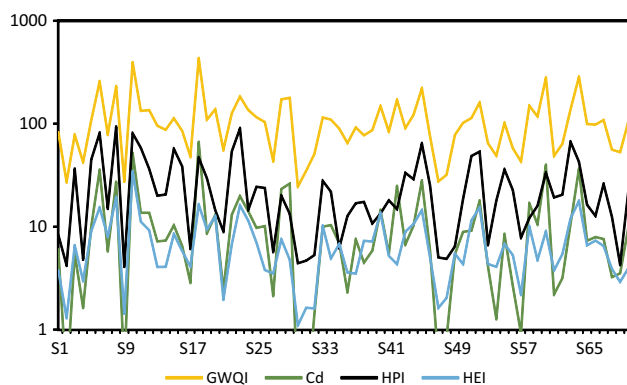
CV critical value

Fig. 2 Spatial variation in groundwater quality index values in the study area

Table 6 Correlation coefficient matrix for indices values and metal concentration

	HPI	HEI	CD	GWQI
HPI	1	0.774**	0.599**	0.661**
HEI	0.774**	1	0.760**	0.830**
Cd	0.599**	0.760**	1	0.988**
GWQI	0.661**	0.830**	0.988**	1
As	0.941**	0.627**	0.498**	0.544**
Pb	0.235	0.116	-0.002	0.042
Fe	0.423**	0.850**	0.615**	0.661**
Mn	0.349**	0.429**	0.366**	0.419**
Ni	0.350**	0.580**	0.506**	0.539**
Sb	0.118	0.1	−0.007	0.016
Ba	−0.108	0.314**	0.261*	0.259*
Mo	0.445**	0.176	0.309**	0.313**
Al	0.281*	0.250*	0.097	0.144
Zn	0.145	0.164	0.331**	0.312**

* Correlation is significant at the 0.05 level (two-tailed)

** Correlation is significant at the 0.01 level (two-tailed)

source. The R-mode CA performed on the groundwater samples produces four clusters. Cluster 1 includes EC, Na, Cl, Mg, SO_4, K, HCO_3, and B which may be explained by combining nonpoint sources and leaching of fertilizers from the soil horizon to the aquifer. Cluster 2 consists of Mn, Zn, Ca, Ni, Fe, and Ba reflecting the influence of domestic and agricultural pollution (Omo-Irabor et al. 2008). Cluster 3 includes As, DOC, NH_4–N, P, Mo, and pH and is explained by the dissolution of minerals under basic condition. Cluster 4 includes Pb, Al, Si, and Sb which represent the presence of anthropogenic and geogenic activities. Although there are some variations between the results of CA and PCA, a good agreement between the two statistical techniques is evident in all the data sets analyzed.

Q-mode CA has been applied to detect the spatial similarities and site grouping among the sampling points (Table 8). Samples are clustered in a particular group which share similar characteristics with respect to the analyzed parameters. The 70 sampling sites fall into four clusters (Fig. 4b). Cluster 1 consists of 37 sampling points. These 41 sampling points are S4–S12, S15–S18, S20, S23, S24, S27, S33, S35, S38–S41, S43–S49, S51–S55, S58, S62, S63, S65, S66, and S69. Among the sites, strong correlation exists among the element pairs: Ca with Cl ($r = 0.679$, $p < 0.01$) and Mg with Na ($r = 0.786$, $p < 0.01$). It is worth noting that moderate concentrations of Ca, Mg, K, SO_4, Cl, HCO_3, EC, B, and Na have been observed at most of the stations. Cluster 2 contains only four samples such as S19, S56, S64, and S67. Cluster 3 includes the following 12 sample sites: S1–S3, S14, S22, S25, S26, S30, S31, S34, S36, and S37. Cluster 4 consists of 13 sites which are S13, S21, S28, S29, S32, S42, S50, S57, S59, S60, S61, S68, and S70. The relationships among the analyzed parameters are also visualized in the factor loading plots of PC1 versus PC2 and PC1 versus PC3 (Figs. 5, 6). Five major clusters are obtained for all parameters from the plotting of PC1 versus PC2 (Fig. 5a). Cluster 1 contains parameters As, P, DOC, Mo, and Fe. Cluster 2 consists of F, Al, pH, Ni, Zn, Mn, Ba, Al, Sb, and Pb. Cluster 3 includes B, Mg, Be, Na, Cl, K, and SO_4. NH_4–N, HCO_3, Ca belong to cluster 4, whereas Si independently remains in cluster 5. Near similar groupings of parameters are observed in the plot of PC1 versus PC3 (Fig. 5b) with some exceptions of Si, pH's groupings. For sampling sites, the PC1 versus PC2 and PC1 versus PC3 plots show same common sources having four distinct clusters (Fig. 6a, b) which differ from the clustering phenomenon that are shown in Fig. 5a, b.

Correlation matrix analysis

Pearson's correlation coefficient matrix is generated in order to identify the rotations among the parameters and sources of groundwater pollution (Table 9). Correlation matrix shows that inter-parameter relationships agree with the results obtained from PCA. It also shows some new associations between the parameters that are not adequately reported in the previous sections. Strong ($p < 0.01$) and significant correlations ($p < 0.05$) are observed in most of the parameters of groundwater samples. EC has a strong positive correlations at $p < 0.01$ with Na^+ ($r = 0.959$), K^+ ($r = 0.682$), Mg^{2+} ($r = 0.847$), Ca^{2+} ($r = 0.645$), NH_4^{2-}–N ($r = 0.538$), SO_4^{2-} ($r = 0.685$), HCO_3^- ($r = 0.608$), Cl^- ($r = 0.939$), and B ($r = 0.842$), and these parameters are also positively correlated with each other (Table 8). These associations indicate mixed sources of geogenic/anthropogenic origin which are described in PC1
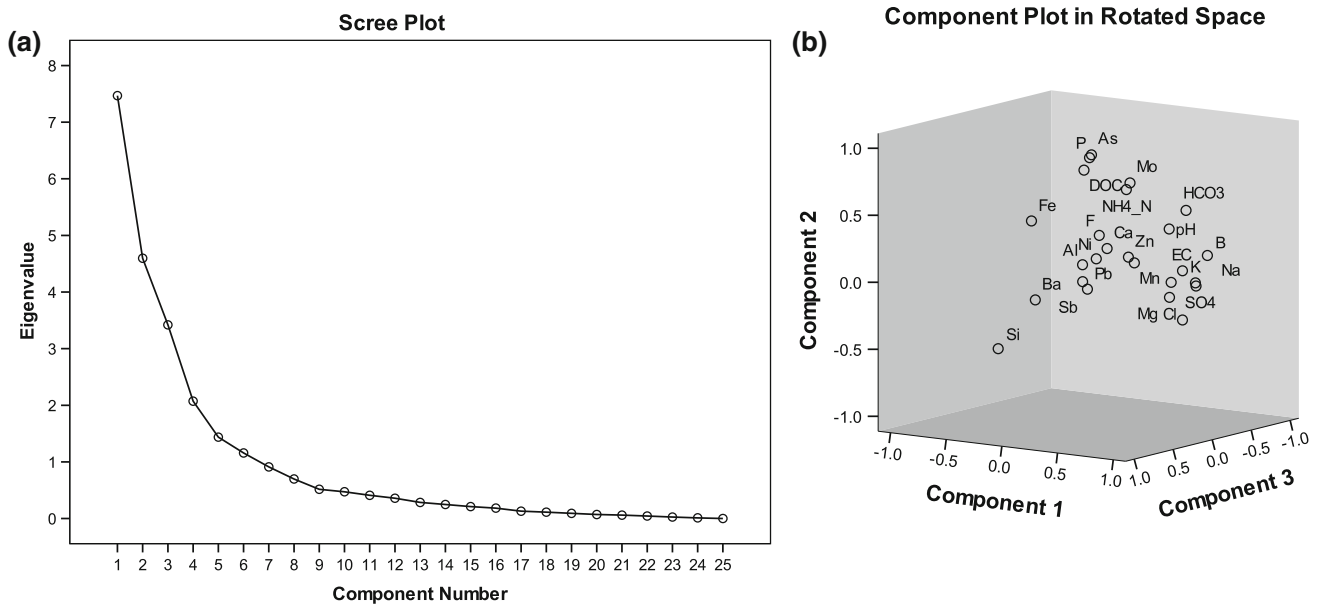
Fig. 3 **a** Scree plot of the characteristic roots (eigenvalues) of principal component analysis, **b** Component plot in rotated space of principal component analysis

section. Na^+, Cl^-, Mg^{2+}, K^+ are the main constituents of groundwater as a result of interaction with minerals or trapped saline fluids in aquifers and chemical weathering of catchment rocks. The partially acidic nature of groundwater is due to leaching of altered rocks and anthropogenic sources. DOC shows significant correlations with As ($r = 0.617$, $p < 0.01$), Mo ($r = 0.479$, $p < 0.05$), Fe ($r = 0.414$, $p < 0.05$), NH_4^{2-}–N ($r = 0.422$, $p < 0.05$), and P ($r = 0.583$, $p < 0.01$). These results are also showing similarity with PC2. These correlation results are attributed to geogenic sources from the basement rocks. Harvey et al. (2002) have reported that DOC in groundwater in Bangladesh is positively correlated with As and has concluded that the mobility of As has been closely related to the recent inflow of carbon or desorption of As by carbonate rock. pH shows negative significant correlations with Fe ($r = -0.488$, $p < 0.05$) and Si ($r = -0.696$, $p < 0.01$). The metal pairs Fe–Ni, Fe–Ba, and Fe–P show significant correlations (at $p < 0.05$) with correlation coefficients (r) 0.493, 0.609, and 0.458, respectively, depicting a similar sources as mentioned in PC3 section. Similar observations are made by Chapagain et al. (2010) in the deep groundwater quality in Kathmando, Nepal, where the occurrence of heavy metals is possibly influenced by redox levels and nature of underlying sediment (i.e., mineral composition and organic matter contents) of groundwater. A poor correlation exists between As with Fe ($r = 0.305$, $p < 0.05$) and Mn ($r = 0.142$, $p < 0.05$), whereas As shows significant correlations with P ($r = 0.773$, $p < 0.01$) and Mo ($r = 0.543$, $p < 0.01$). So,

the correlations between As, Fe, and Mn in groundwater are due to the precipitation of dissolved Fe as siderite solids ($FeCO_3$) or rhodochrosite ($MnCO_3$) under reduced environmental conditions (Reza et al. 2010a). High As, low Fe, low Mn in groundwater of the Meghna flood plain in southeastern parts of Bangladesh are evaluated by Reza et al. (2010b). However, their findings quite differ from the observations in this study where very high concentrations of Fe, Mn, and As are observed. A positive strong correlation is observed between Pb and Al ($r = 0.843$, $p < 0.01$), indicating similar sources as observed in PC5. Ni exhibits significant correlation with Sb ($r = 0.480$, $p < 005$) as observed in P62 section.

Geostatistical modeling

Semivariogram models are employed after normalizing the data using ArcGIS (version 10.2). Among other models, OK (ordinary kriging) is applied in this study. The nugget, sill, and the range values of the best fit semivariogram models for quality parameters are shown in Table 10. Extremely low nugget effects and the absence of variability in groundwater elevation at short distances demonstrate that there is an insignificant small-scale variability measurement error. Thus, the fit semivariogram represents very well the spatial structure of these variables in the groundwater. The error statistics are estimated to ascertain the reproducibility of observed values by the theoretical semivariogram models and developed spatial maps (Table 10). The best fit semivariogram model is chosen

Table 7 Varimax rotated principal component analysis (R-mode) for groundwater samples

| Parameters	PC1	PC2	PC3	PC4	PC5	PC6
R-mode						
pH	0.212	0.265	−0.732	−0.26	−0.058	−0.249
EC	0.957	0.158	0.161	0.047	−0.015	0.023
DOC	−0.016	0.792	0.035	0.021	−0.041	−0.056
HCO₃	0.627	0.5	−0.355	−0.059	0.03	−0.022
Cl	0.917	−0.029	0.271	0.141	−0.035	0.016
F	−0.16	0.233	−0.366	−0.702	−0.032	0.009
SO_4^2	0.762	−0.269	−0.117	0.24	−0.039	0.008
NH₄_N	0.416	0.701	0.109	−0.059	0.012	−0.009
Na	0.977	0.025	0.017	−0.044	−0.046	0.01
K	0.782	−0.01	−0.254	0.208	−0.025	−0.066
Ca	0.529	0.332	0.517	0.486	0.059	0.007
Mg	0.839	0.056	0.139	0.413	0.056	−0.006
As	−0.105	0.865	−0.189	0.11	0.007	0.077
Pb	−0.012	−0.101	−0.003	0.101	0.947	0.034
Fe	0.006	0.517	0.74	0.165	0.076	0.077
Mn	0.267	0.094	−0.207	0.77	0.182	0.208
Ni	0.333	0.215	0.376	0.304	0.23	0.637
Sb	−0.067	−0.054	−0.021	0.019	−0.038	0.944
Ba	0.16	−0.032	0.91	−0.077	−0.071	−0.042
Mo	0.153	0.663	−0.314	−0.49	0.076	0.065
Al	−0.015	0.089	0.054	0.041	0.955	0.013
Zn	0.361	0.177	0.004	0.01	0.118	0.171
B	0.84	0.19	−0.325	−0.202	−0.018	0.007
P	0.029	0.888	0.029	−0.097	0.008	0.019
Si	−0.326	−0.477	0.692	−0.057	0.1	−0.074
Eigenvalues	7.468	4.596	3.42	2.073	1.438	1.158
% of variance	26.869	17.288	14.224	8.546	7.82	5.865
Cumulative %	26.869	44.157	58.382	66.928	74.748	80.613
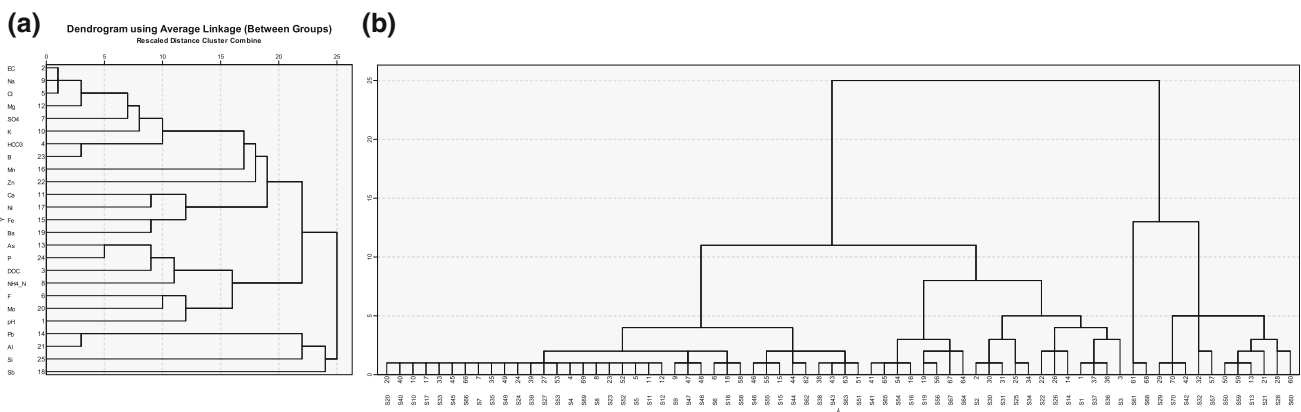
Fig. 4 a Dendrogram showing the hierarchical clusters of analyzed parameters. *Dashed lines* in dendrogram represent *Phenon lines*. **b** Dendrogram showing the hierarchical clusters of analyzed samples site. *Dashed lines* in dendrogram represent *Phenon lines*

based on ME, MSE, RMSE, RMSSE, ASE criterion. A model is considered robust and accurate when ME and MSE are close to zero, RMSE and ASE are small, and RMSSE is close to 1 (Adhikary et al. 2010). Ranges of semivariograms are a measure of extension where auto-correlation existed (Li et al. 2009). The distance where the

Table 8 Varimax rotated principal component (Q-mode) analysis for groundwater samples

Sample	PC1	PC2	PC3	PC4	PC5	PC6	Sample	PC1	PC2	PC3	PC4	PC5	PC6
S1	0.43	−0.21	−0.49	−0.88	0.11	0.23	S36	−0.26	−0.12	−0.71	−1.25	−0.04	−0.21
S2	−0.62	−0.51	0.16	−0.19	−0.40	−0.88	S37	−0.12	−0.04	−0.70	−0.35	−0.03	−0.29
S3	−0.49	0.75	−0.82	1.14	−0.40	−0.56	S38	−0.83	−0.31	−0.40	1.29	−0.51	−0.43
S4	−0.75	−0.62	0.56	−0.21	−0.45	−0.73	S39	−0.36	−0.97	1.33	−0.72	−0.37	1.37
S5	−0.77	0.55	−0.21	0.44	−0.47	1.40	S40	−0.26	−0.48	2.20	−0.14	−0.40	0.05
S6	1.09	3.09	−0.06	−2.04	−0.04	−0.16	S41	−0.45	−0.38	−0.30	0.85	−0.32	0.95
S7	−0.80	−0.30	0.07	0.49	0.16	−0.40	S42	3.57	−1.26	−0.82	0.67	−0.14	0.27
S8	−0.54	2.64	0.10	1.20	−0.12	0.09	S43	−0.53	−0.29	−0.35	0.64	0.46	0.36
S9	−0.62	−0.79	0.33	−0.54	−0.14	−0.32	S44	−0.16	0.03	−0.68	1.98	−0.18	0.02
S10	0.12	3.51	2.62	1.35	1.11	−0.16	S45	−0.46	1.87	0.35	0.52	−0.39	−0.26
S11	−0.62	1.59	−0.48	−0.19	0.01	−0.18	S46	−0.80	0.21	−0.39	−0.04	−0.46	−0.35
S12	−0.09	1.25	−0.45	−0.86	1.35	0.62	S47	−0.74	−0.97	0.35	−0.47	−0.13	−0.68
S13	0.29	0.04	−0.44	−0.68	−0.05	−0.17	S48	−0.71	−0.98	0.31	−0.52	−0.15	−0.43
S14	0.53	0.66	−0.93	−2.58	0.45	0.51	S49	−0.32	−0.81	1.23	−0.30	−0.53	−0.13
S15	−0.77	0.75	−0.54	0.80	−0.23	−0.41	S50	0.55	0.19	−0.75	−1.13	0.27	0.45
S16	−0.67	0.40	−0.92	0.07	0.02	−0.35	S51	−0.79	−0.41	−0.36	0.04	−0.23	7.25
S17	−0.75	−0.91	0.81	−0.60	−0.23	−0.38	S52	−0.85	1.12	0.36	1.29	−0.28	0.40
S18	2.30	2.53	1.63	−1.02	−0.23	−0.23	S53	−0.53	−0.86	0.87	−0.32	−0.31	−0.11
S19	−0.45	−0.15	−0.67	1.36	−0.32	−0.33	S54	−0.41	−0.45	−0.46	−0.66	1.17	0.60
S20	−0.35	−0.26	1.96	0.09	−0.32	−0.39	S55	−0.43	0.58	−0.53	−0.71	−0.24	0.11
S21	−0.08	−0.19	−0.77	−1.13	−0.18	−0.37	S56	−0.78	−0.61	−0.41	0.47	−0.35	0.16
S22	−0.16	0.89	−1.30	−1.34	−0.35	−0.37	S57	−0.03	−0.70	−0.94	−0.82	−0.20	−0.58
S23	−0.83	1.33	−0.73	0.85	−0.53	−0.41	S58	1.23	−0.93	3.68	−0.43	−0.51	−0.59
S24	0.48	−0.31	2.99	−0.24	−0.06	1.29	S59	0.76	−0.42	−0.30	−0.39	−0.19	−0.03
S25	0.08	0.31	−0.69	−0.12	−0.19	−0.20	S60	2.59	0.25	0.07	0.20	−0.18	−0.22
S26	−0.02	0.54	−0.90	−1.89	0.40	0.28	S61	−0.67	−0.37	−0.66	0.40	−0.31	−0.74
S27	−0.73	−0.84	0.89	−0.69	−0.25	−0.32	S62	−0.44	−0.33	−0.66	−0.01	0.19	−0.46
S28	2.36	−0.19	−0.22	1.36	−0.05	0.45	S63	−0.19	−0.84	0.19	0.69	7.68	−0.30
S29	3.93	−0.99	−0.68	0.27	−0.13	0.15	S64	1.00	−0.07	0.00	3.84	−0.20	0.24
S30	−0.67	−0.85	0.01	−0.26	0.04	−0.81	S65	0.46	−0.51	−0.79	0.72	−0.13	−0.26
S31	−0.50	−0.95	0.26	−0.26	−0.37	−0.29	S66	0.20	−0.73	−0.17	0.88	−0.38	−0.08
S32	0.10	−0.29	−1.03	−0.86	−0.14	−0.21	S67	0.01	−0.20	−0.72	1.31	−0.36	−0.47
S33	−0.61	0.29	0.03	0.51	−0.39	−0.39	S68	−0.13	−0.67	−0.54	0.90	0.22	−0.34
S34	0.27	0.36	−0.78	−0.89	0.09	−0.14	S69	−0.67	−0.91	0.70	−0.62	−0.03	−0.63
S35	−0.15	−0.77	1.64	−0.38	0.23	−0.06	S70	1.61	−1.01	−0.98	0.11	−0.37	−0.46

models first flatten out is known as the range which varies in each variable of groundwater quality indices. The range of semivariograms for all the variables is from 0.5 to 22 km showing the lowest range for PC1 and highest range for GWQIs.

The nugget/sill ratio represents spatial dependence of groundwater quality parameters (Nayanaka et al. 2010). There are three classifications used for model explanation: If the ratio is less than 25 %, it shows strong spatial dependence; if the ratio is in between 25 and 75 %, it indicates moderate spatial dependence; and if the ratio is

more than 75 %, it represents weak spatial dependence. Figures 7 and 8 show experimental semivariograms (binned sign) around the omnidirectional semivariogram model (blue line) and average of semivariogram models (plus sign). The best fit semivariogram models for different groundwater quality parameters are shown in Figs. 7 and 8, and Table 10. The Gaussian semivariogram model is identified to be the best fit model for CD, HPI, PC3, PC4, PC5, PC6, while the circular semivariogram model fit best for GWQIs. The exponential semivariogram model is observed to be the best fit model for HEI and PC2, whereas
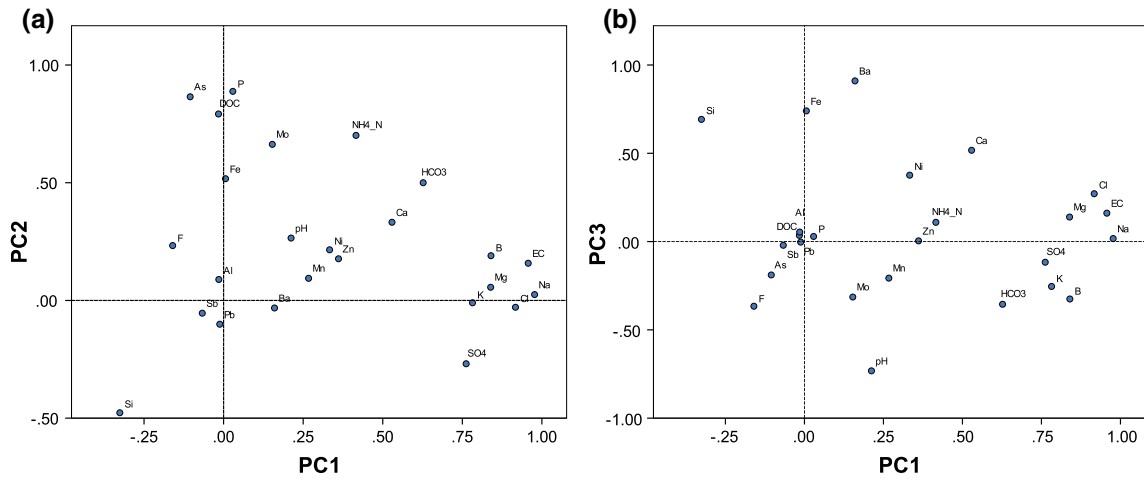
Fig. 5 Plots of first three principal component loadings. **a** PC1 versus PC2 and **b** PC1 versus PC3 for all analyzed parameters
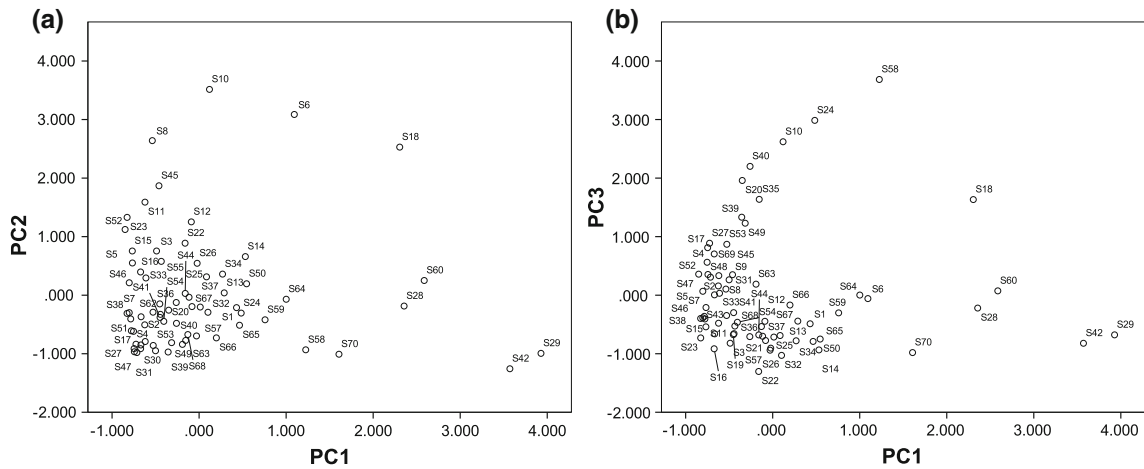

Fig. 6 Plots of first three principal component loadings. **a** PC1 versus PC2, **b** PC1 versus PC3 for all analyzed sampling sites

the spherical model fit well for PC1. In this study, the results show that most of the variables are weakly spatial dependent in semivariogram shapes except PC1 and HEI. PC1 represents strong spatial dependence, whereas HEI shows moderate spatial dependence (Fig. 7). Mostly, the weak spatial dependence has been demonstrated in the large nugget effect in semivariogram shape (Fig. 8) and is due to the high variability of topography of groundwater which varies with the agriculture, residential, and industrial areas.

The OK interpolation techniques are applied to develop spatial distribution maps of groundwater data set ($n = 70$) for each groundwater pollution index and each PC score. After conducting the cross-validation test of best fit models, maps of kriged estimates are generated which provide a visual representation of the distribution of groundwater samples. GWQIs, CD, HPI, and HEI exhibit more or less similar distribution patterns with an increasing trend at the

southeast to northwest direction of the study area (Fig. 9). In general, contamination of groundwater with metal pollution is attributed to anthropogenic sources. Other than this, both geogenic and anthropogenic sources are likely to be contributed by the high scores of GWQIs, CD, HPI, and HEI in the study area. High scores are observed at the northern and northwestern parts of the sampling area, and low scores are observed at the southeastern part, suggesting the existence of similar point sources. The strong significant correlations among GWQIs, CD, HPI, and HEI show similar distribution pattern of the previous findings.

Spatial distribution of PC1 score in Fig. 10a reveals complex pattern where high scores (i.e., values from 0.42 to 3.93) are mostly observed at the southern parts and low scores are found at the northern and central parts. The higher score of PC1 is probably due to the rock–water interaction and ion exchange in groundwater (Guler et al. 2012). Similar findings are obtained from the recent studies

Table 9 Pearson's correlation matrix among physicochemical parameters and metals in the analyzed samples

	pH	EC	DOC	HCO3	Cl	F	SO4	NH4–N	Na	K	Ca	Mg	As
pH	1												
EC	0.098	1											
DOC	0.193	0.14	1										
HCO3	0.458**	0.608**	0.374**	1									
Cl	−0.02	0.939**	−0.023	0.367**	1								
F	0.484**	−0.215	0.126	0.098	−0.319**	1							
SO4	0.124	0.685**	−0.146	0.349**	0.687**	−0.343**	1						
NH4–N	0.198	0.538**	0.422**	0.512**	0.428**	0.114	−0.042	1					
Na	0.2	0.959**	0.032	0.615**	0.922**	−0.158	0.749**	0.432**	1				
K	0.321**	0.682**	−0.013	0.570**	0.645**	−0.169	0.632**	0.279*	0.705**	1			
Ca	−0.246*	0.645**	0.243*	0.293*	0.679**	−0.485**	0.329**	0.458**	0.478**	0.0385**	1		
Mg	−0.004	0.847**	−0.007	0.431**	0.894**	−0.372**	0.635**	0.427**	0.786**	0.697**	0.756**	1	
As	0.318**	0.005	0.617**	0.372**	−0.123	0.21	−0.2	0.510**	−0.064	−0.019	0.166	−0.001	1
Pb	−0.115	−0.025	−0.132	−0.054	−0.008	−0.14	0.004	−0.046	−0.05	−0.01	0.056	0.087	−0.016
Fe	−0.488**	0.213	0.414**	0.021	0.182	−0.252*	−0.14	0.397**	0.014	−0.1	0.620**	0.172	0.305*
Mn	−0.081	0.270*	0.002	0.223	0.306*	−0.349**	0.317**	0.161	0.204	0.320**	0.469**	0.592**	0.143
Ni	−0.345**	0.425**	0.176	0.16	0.433**	−0.308**	0.214	0.240*	0.302*	0.195	0.631**	0.475**	0.145
Sb	−0.222	−0.044	−0.075	−0.071	−0.035	−0.063	−0.011	−0.047	−0.039	−0.068	−0.052	−0.069	0.059
Ba	−0.543**	0.270*	0.018	−0.231	0.371**	−0.301*	0.042	0.062	0.156	−0.121	0.536**	0.234	−0.193
Mo	0.509**	0.152	0.479**	0.567**	−0.047	0.582**	−0.153	0.480**	0.153	0.071	−0.073	−0.043	0.543**
Al	−0.075	−0.005	0.085	0.062	−0.038	−0.079	−0.029	0.034	−0.048	−0.03	0.121	0.044	0.059
Zn	0.098	0.358**	0.111	0.187	0.313**	−0.017	0.22	0.292*	0.299*	0.223	0.269*	0.302*	0.099
B	0.435**	0.766**	0.118	0.853**	0.617**	0.123	0.560**	0.426**	0.842**	0.663**	0.221	0.563**	0.103
P	0.227	0.179	0.583**	0.471**	−0.01	0.194	−0.17	0.633**	0.075	−0.041	0.257*	0.016	0.773**
Si	−0.696**	−0.296*	−0.373**	−0.664**	−0.11	−0.267*	−0.246*	−0.340**	−0.322**	−0.420**	−0.001	−0.19	−0.492**

	Pb	Fe	Mn	Ni	Sb	Ba	Mo	Al	Zn	B	P	Si
pH												
EC												
DOC												
HCO3												
Cl												
F												
SO4												
NH4–N												
Na												

Table 9 continued

	Pb	Fe	Mn	Ni	Sb	Ba	Mo	Al	Zn	B	P	Si
K												
Ca												
Mg												
As												
Pb	1											
Fe	0.022	1										
Mn	0.216	0.025	1									
Ni	0.21	0.493**	0.425**	1								
Sb	0.036	0.037	0.133	0.480**	1							
Ba	−0.061	0.609**	−0.212	0.330**	−00.055	1						
Mo	−0.06	0.01	−0.081	−0.019	−0.019	−0.205	1					
Al	0.843**	0.194	0.17	0.278*	−0.005	−0.03	0.072	1				
Zn	0.054	0.082	0.178	0.300*	−0.012	0.005	0.172	0.067	1			
B	−0.075	−0.156	0.178	0.129	−0.059	−0.163	0.474**	−0.034	0.223	1		
P	−0.079	0.458**	−0.019	0.165	−0.013	−0.022	0.603**	0.086	0.134	0.213	1	
Si	0.12	0.196	−0.257*	0.003	−0.041	0.551**	−0.513**	0.077	−0.219	−0.551**	−0.377**	1

* Correlation is significant at the 0.05 level (two-tailed)

** Correlation is significant at the 0.01 level (two-tailed)

Table 10 Indices and PCs of best fit semivariogram models for groundwater parameters and their variance

Indices and PCs	Fit model types	Nugget	Major range	Sill	Nugget/Sill	Lag size	ME	RMSE	MSE	RMSSE	ASE
GWQI	Circular	5911.38	22,381.50	6055.74	0.976	2797.69	−0.826	80.49	−0.01	1.01	79.52
CD	Gaussian	147.90	22,381.49	147.90	1.00	2797.69	−0.044	12.68	−0.003	1.01	12.55
HPI	Gaussian	411.10	8719.78	411.10	1.00	1089.98	−0.627	21.61	−0.029	1.03	20.95
HEI	Exponential	19.42	8142.02	25.72	0.76	1017.75	−0.139	5.45	−0.025	1.07	5.04
PC1	Spherical	0.20	5304.55	1.07	0.19	663.07	−0.004	0.98	−0.001	1.09	0.89
PC2	Exponential	0.73	8354.05	0.86	0.84	1044.26	−0.023	0.95	−0.023	1.01	0.94
PC3	Gaussian	0.91	6743.03	0.91	1.00	842.88	−0.001	1.01	−0.001	1.01	0.99
PC4	Gaussian	1.04	11,276.01	1.039	1	1409.50	−0.011	1.01	−0.009	0.96	1.05
PC5	Gaussian	0.93	9418.91	0.927	1	1177.37	−0.009	1.03	−0.009	1.03	0.99
PC6	Gaussian	1.04	13,156.66	1.036	1	1644.58	0.005	1.01	0.004	0.97	1.05

ME mean error, *RMSE* root mean square error, *MSE* mean standardized error, *RMSSE* root mean square standardized error, *ASE* average standard error
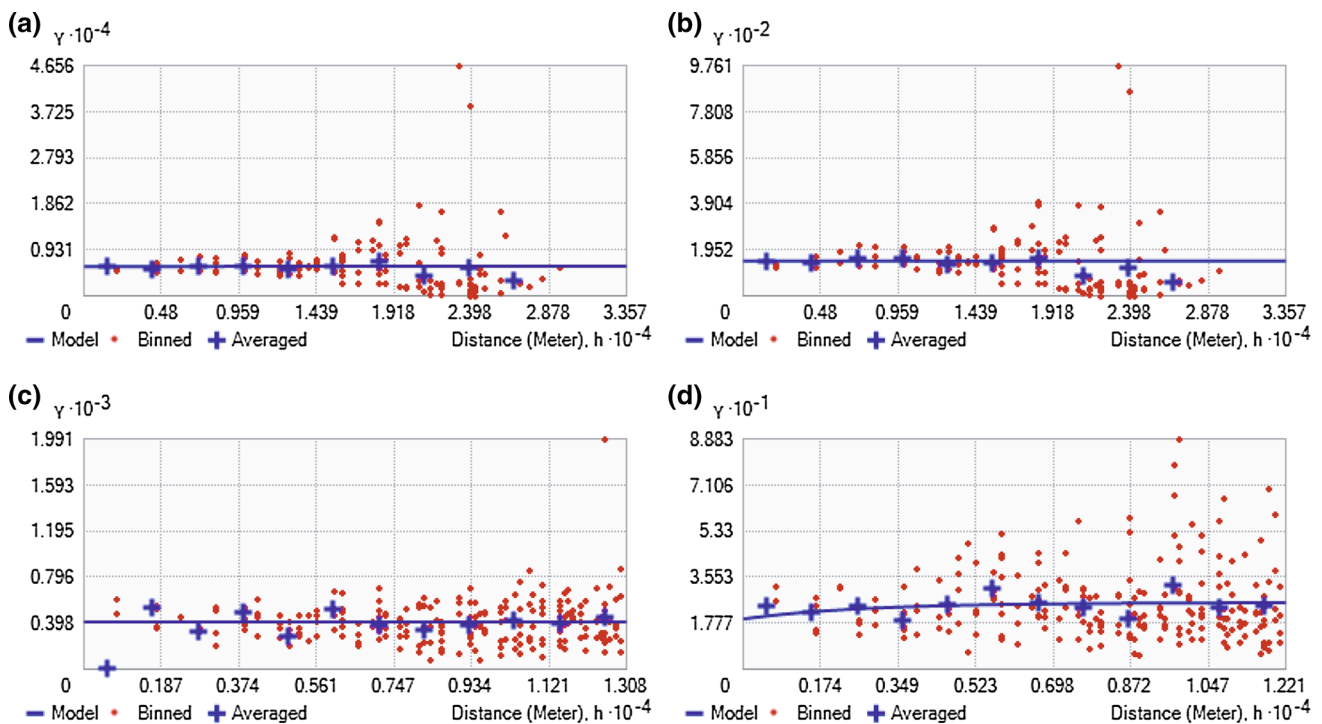

Fig. 7 Best fit semivariogram models of groundwater quality evaluation indices in the study area. **a** GWQI, **b** CD/C_d, **c** HPI, and **d** HEI

conducted in various regions of the world (e.g., Fernandes et al. 2008; Guler et al. 2012). The spatial distribution of PC2 scores shows high scores (values from 0.02 to 3.52) at the northern part and the low scores (i.e., values from −1.26 to −0.38) at the southern part (Fig. 10b). This variation is influenced by the dominance of nonpoint sources such as fertilizer and sewage discharges. For example, high NH_4 and P in PC2 of groundwater samples may be attributed to the use of fertilizers and urban runoff.

Figure 10c shows the distribution of PC3 scores, where moderately higher scores (i.e., values from 0.23 to 1.91) are observed at the northern corner of the study area and the lower scores are found at the southern edge. PC3 contains very high scores of Fe, Ba, and Si which are likely due to geogenic origin. Figure 10d showing the distribution of PC4 scores reverses the distribution pattern of PC3 score, where the higher scores (from 0.12 to 2.18) are observed at the southern part and the lower scores are observed at the
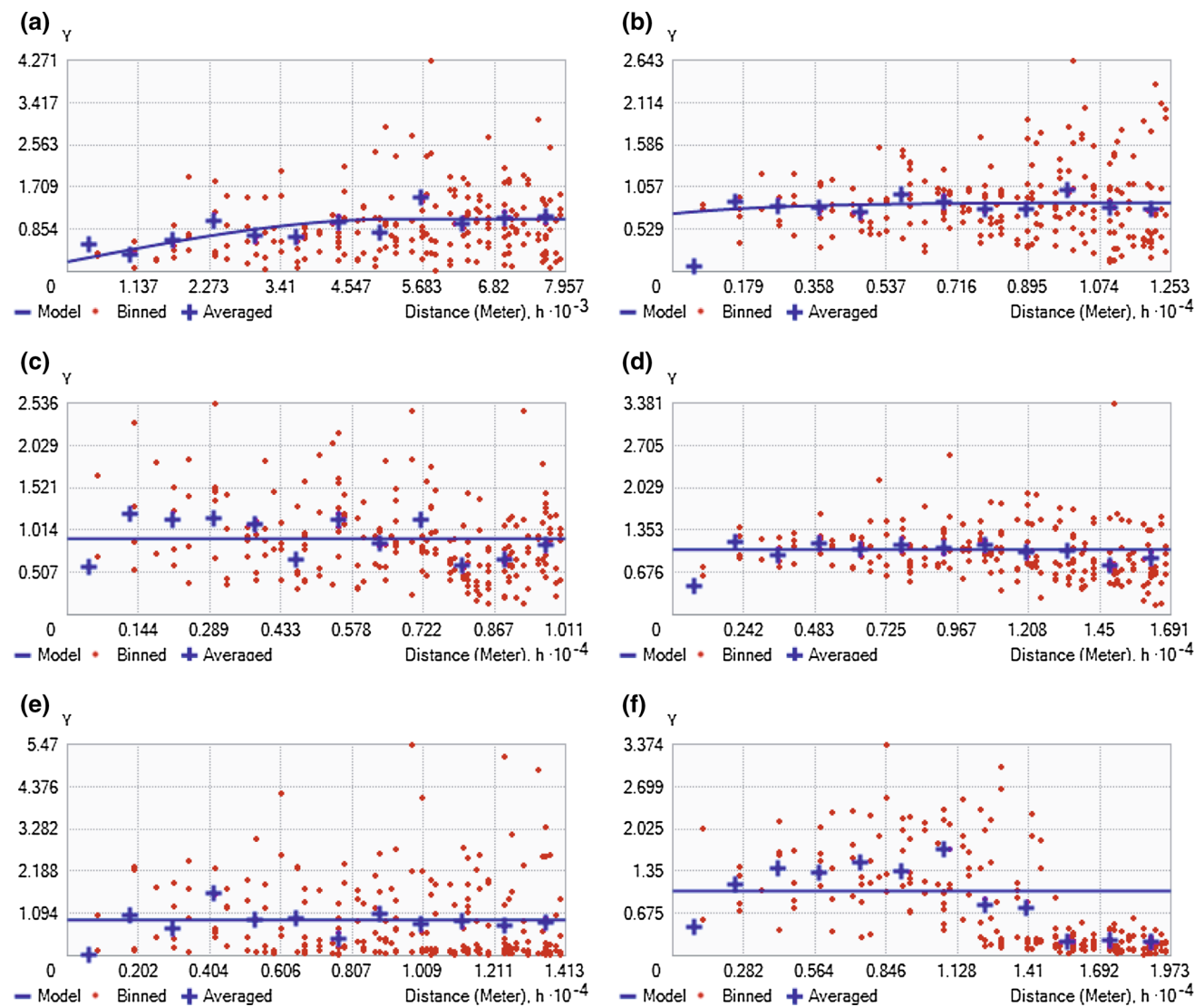
Fig. 8 Best fit semivariogram models of groundwater parameters in the study area. a PC1, b PC2, c PC3, d PC4, e PC5, and f PC6

northern part. PC4 comprises of the high positive score of Mn and negative score of F reflecting the geogenic origin. The distribution of PC5 component score in Fig. 10e shows higher score (values from 0.12 to 2.18) at the eastern part and the lower score at the western part of the study area. Overall, the map shows a decreasing trend toward the east. In PC5, high score of Al is strongly correlated with Pb, indicating the anthropogenic origin. Figure 10f shows the spatial distribution of PC6 component score. The highest score (i.e., values from 1.00 to 7.25) is found at the central part which comprises of high score of Sb and Ni. This component is contributed by the wastewater discharge from urban areas.

Among the geostatistical models, OK is considered as an effective tool for preliminary decision makers of groundwater quality management at the southeastern Bangladesh.

The best fit semivariogram models are selected based on the criterion of ME, MSE, RMSE, RMSSE, and ASE which varied for each PC and index. The Gaussian semivariogram model is identified to be the best fit model for CD, HPI, PC3, PC4, PC5 and PC6 while the circular semivariogram model fit best for GWQIs. The exponential semivariogram model is observed to be the best fit model for HEI and PC2, whereas the spherical model fit well for PC1. In the study, most of the variables have weak spatial dependence except PC1 and HEI. PC1 represents strong spatial dependence, whereas HEI shows moderate spatial dependence. However, the weak spatial dependence indicates high variability of topography in groundwater system (e.g., agriculture, residential, and industrial areas). The cross validation indicates that kriging method is the most accurate interpolation technique for the study. The spatial
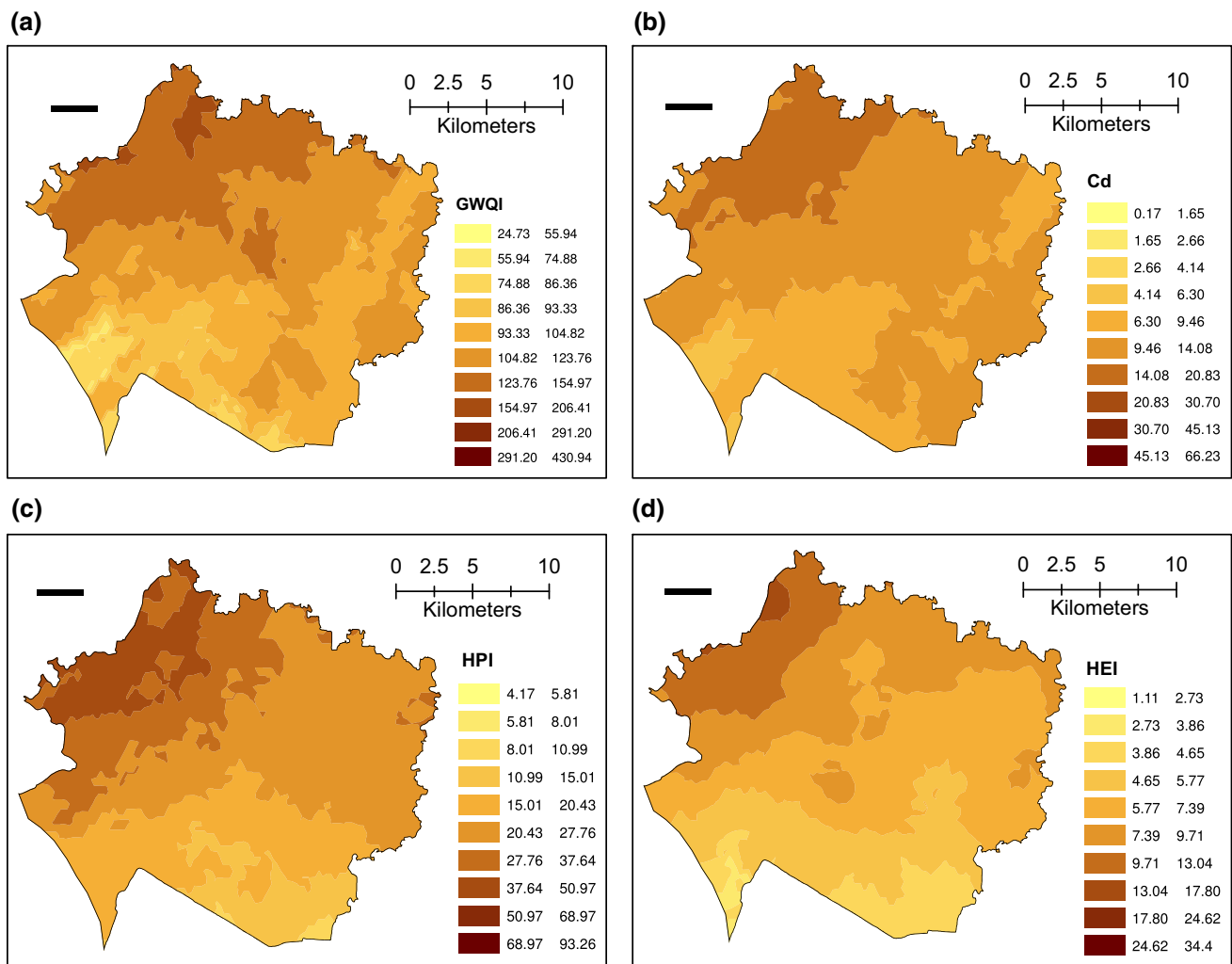
(a)

(b)

(c)

(d)

Fig. 9 Maps showing the spatial distribution of four index scores obtained by the indices of quality evaluation of the groundwater samples. **a** GWQI, **b** CD/C_d, **c** HPI, and **d** HEI. Here, both HPI and HEI are plotted considering with As and other metals

distribution maps are generated using OK method showing the variation in different indices in the study area. Considering the distribution maps, six PC groups (PC1–PC6) and four indices (GWQIs, CD, HPI, and HEI) with different spatial variability patterns can be distinguished (Figs. 9, 10). Indices scores exhibit more or less similar distribution pattern with an increasing trend at the southeast to northwest direction where high scores of GWQIs, CD, HPI, and HEI are observed at the northern and northwestern parts of the study area (Fig. 9a–d). The low scores are found at the southeastern part which suggests the existence of similar point sources. However, spatial analysis of GWQIs depicts that water quality is poor at the western and northwestern parts of the study area. Different distribution patterns of the PC scores have implied the existence of different sources. Distribution maps of PC1 to PC6 have shown some anomalies regarding the mean scores of groundwater variables (Fig. 10a–f). In fact,

salinity observed at the northwestern part of PC1 of the study area is alarming, as households depend upon groundwater for domestic purposes.

Conclusions

Integrated approaches of different multivariate statistical and geostatistical techniques are employed in this study to evaluate the variations in groundwater quality of the Lakshimpur district of Bangladesh. Cluster analysis groups 70 sampling points into four clusters of similar water quality characteristics. Based on the obtained information, it may be easier to design the study area further where optimal sampling strategies can reduce the number of sampling sites and associated costs. Principal component analysis assists for identifying the factors or sources responsible for water quality variations. This study illustrates the
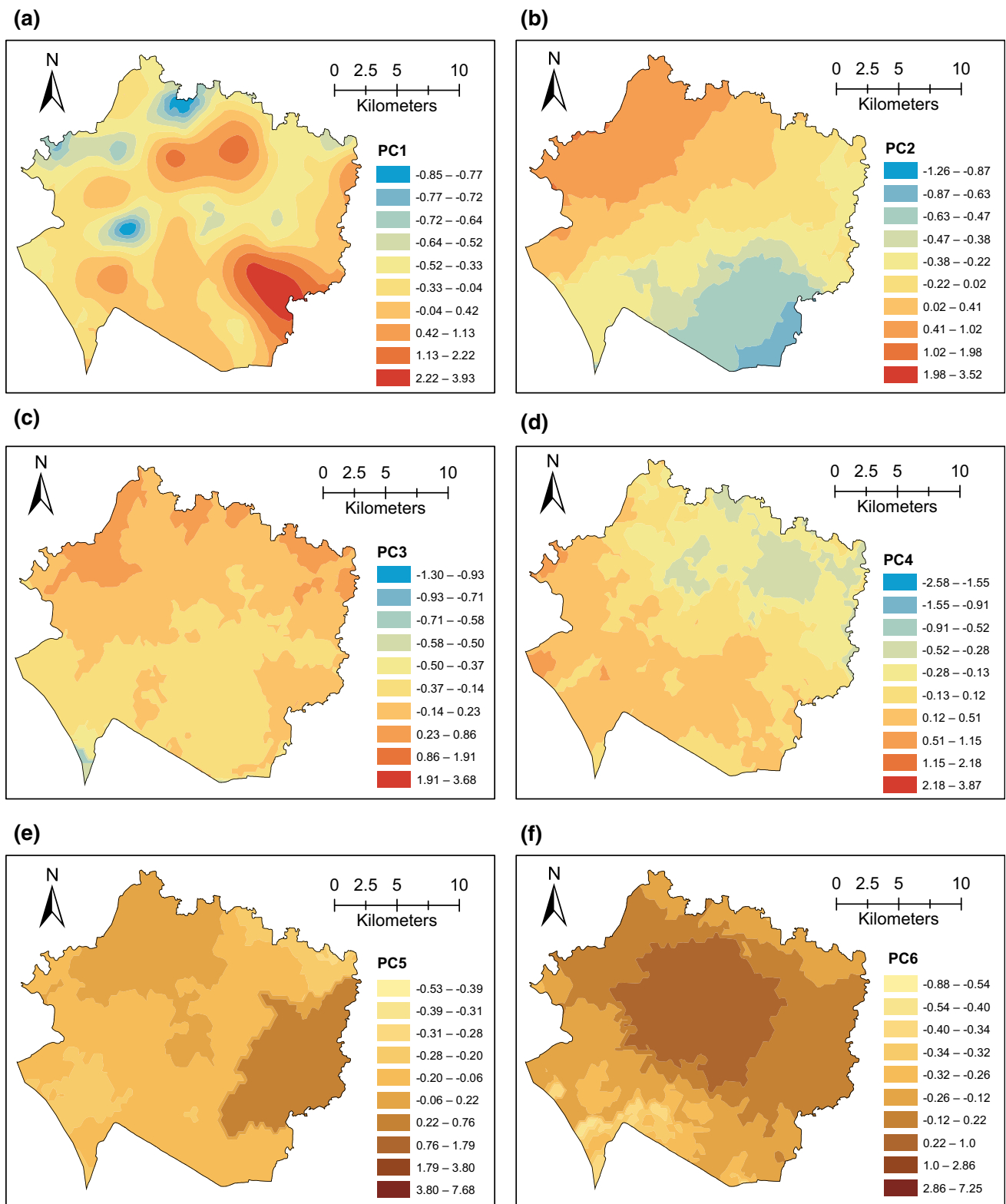
Fig. 10 Maps showing the spatial distribution of six PC scores. **a** PC1: Na, K, Ca, Mg, Cl, SO_4, HCO_3, EC, and B; **b** PC2: P, Mo, As, NH_4N, DOC, and Fe; **c** PC3: Fe, Ba, and Si; **d** PC4: Mn and Fe; **e** Pb and Al; and **f** Ni and Sb

usefulness of multivariate statistical and geostatistical techniques for the analysis and interpretation of complex data sets, identification of pollution sources, and understanding spatial variations in water quality for effective groundwater management. Moreover, the chemometric studies enable us to demonstrate the similarities and differences among the examined variables that are not evidently visible from an examination of the analytical data as shown in the tables. The ordinary kriging is an effective tool for preliminary decisions makers of groundwater quality management in southeastern Bangladesh. Except As, other geogenic pollutants are not much alarming for the groundwater consumption at the study area. However, some anthropogenic processes are quite unpleasant.

The results of principal component/factor analysis indicate that anthropogenic (agrogenic, surface runoff, and domestic sewage) and natural/geogenic sources (weathering of source rock) are responsible for variation in physicochemical parameters and metal contents in groundwater systems at the southeastern coastal region of Bangladesh. The resulting spatial distribution maps provide a helpful and robust visual tool for researchers and policy makers toward defining adaptive measures. The study provides background information on physiochemical parameters, harmful metals, possible sources, and spatial variation in groundwater systems in Lakshimpur district of Bangladesh.

References

Adhikary PP, Chandrasekharan H, Chakraborty D, Kamble K (2010) Assessment of groundwater pollution in West Delhi, India using geostatistical approach. Environ Monit Assess 167(1–4):599–615

Ahmed ZU, Panaullah GM, DeGloria SD, Duxbury JM (2011) Factors affecting paddy soil arsenic concentration in Bangladesh: prediction and uncertainty of geostatistical risk mapping. Sci Total Environ 412–413:324–335

Al-Ami MY, Al-Nakib SM, Ritha NM, Nouri AM, Al-Assina A (1987) Water quality index applied to the classification and zoning of Al-Jaysh canal, Bagdad, Iraq. J Environ Sci Health A22:305–319

APHA-AWWA-WEF (2005) Standard methods for the examination of water and wastewater, 20th edn. APHA, AWWA and WEF, Washington DC, USA

Astel A, Tsakovski S, Barbieri P, Simeonov V (2007) Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Res 41:4566–4578

Astel A, Tsakovski S, Simeonov V, Reisenhofer E, Piselli S, Barbieri P (2008) Multivariate classification and modeling in surface water pollution estimation. Anal Bioanal Chem 390:1283–1292

Backman B, Bodis D, Lahermo P, Rapant S, Tarvainen T (1997) Application of a groundwater contamination index in Finland and Slovakia. Environ Geol 36:55–64

Backman B, Bodiš D, Lahermo P, Rapant S, Tarvainen T (1998) Application of a groundwater contamination index in Finland and Slovakia. Environ Geol 36(1–2):55–64

Banglapedia (2006) National encyclopedia of Bangladesh. Asiatic Society of Bangladesh, Dhaka. https://en.wikipedia.org/wiki/Banglapedia

Bhattacharya P, Frisbie SH, Smith E, Naidu R, Jacks G, Ahmed KM, Khan AA, Routh J (2002) Arsenic in groundwater of the Bengal Delta plain aquifers in Bangladesh. Bull Environ Contam Toxicol 69:538–545

Bhuiyan MAH, Islam MA, Dampare SB, Parvez L, Suzuki S (2010) Evaluation of hazardous metal pollution in irrigation and drinking water systems in the vicinity of a coal mine area of northwestern Bangladesh. J Hazard Mater 179:1065–1077. doi:10.1016/j.jhazmat.2010.03.114

Bhuiyan MAH, Rakib MA, Dampare SB, Ganyaglo S, Suzuki S (2011a) Surface water quality assessment in the central part of Bangladesh using multivariate analysis. KSCE J Civil Engineer 15(6):995–1003. doi:10.1007/s12205-011-1079-y

Bhuiyan MAH, Suruvi NH, Dampare SB, Islam MA, Quraishi SB, Ganyaglo S, Suzuki S (2011b) Investigation of the possible sources of heavy metal contamination in lagoon and canal water in the tannery industrial area in Dhaka, Bangladesh. Environ Monit Assess 175:633–649. doi:10.1007/s10661-010-1557-6

Bhuiyan MAH, Dampare SB, Islam MA, Suzuki S (2015) Source apportionment and pollution evaluation of heavy metals in water and sediments of Buriganga River, Bangladesh, using multivariate analysis and pollution evaluation indices. Environ Monit Assess 187:4075. doi:10.1007/s10661-014-4075-0

Birth G (2003) A scheme for assessing human impacts on coastal aquatic environments using sediments. In: Woodcoffe CD, Furness RA (eds) Coastal GIS. Wollongong University Papers in Center for Maritime Policy, Australia

BIS (Bureau of Indian Standards) (2012) Indian Standard drinking water-specification, 1st rev., pp 1–8

Bromfield SM (1978) The oxidation of manganous ions under acidic conditions byan acidophilousactinomycete from acid soil. Aust J Soil Res 16:91–100

Burrough PA, McDonnell RA (1998) Principles of geographical information systems. Oxford University Press, Oxford

Chabukdhara M, Nema AK (2012) Assessment of heavy metal contamination in Hindon River sediments: a chemometric and geochemical approach. Chemosphere 87:945–953

Chapagain SK, Pandey VP, Shrestha S, Nakamura T, Kazama F (2010) Assessment of deep groundwater quality in Kathmandu valley using multivariate statistical techniques. Water Air Soil Pollut 210:277–288. doi:10.1007/s11270-009-0249-8

Danielsson A, Cato I, Carman R, Rahm L (1999) Spatial clustering of metals in the sediments of Skagerrak/Kattegat. Appl Geochem 14:689–706

Delhomme JP (1978) Kriging in the hydrosciences. Adv Water Res 1:251–266

DoE (Department of Environment) (1997) The environment conservation rules 1997. Government of the People's Republic of Bangladesh, Dhaka

Dragovíc S, Mihailovíc N, Gajíc B (2008) Heavy metals in soils: distribution, relationship with soil characteristics and radionuclides and multivariate assessment of contamination sources. Chemosphere 72:491–549

Edet AE, Offiong OE (2002) Evaluation of water quality pollution indices for heavy metal contamination monitoring. A study case from Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria). GeoJournal 5:295–304

Elogne S, Hristopulos D, Varouchakis E (2008) An application of Spartan spatial random fields in environmental mapping: focus on automatic mapping capabilities. Stochastic Environ Res Risk Assess 22(5):633–646

ESRI (2009) ArcGIS desktop software. ArcGIS Desktop 9.3., Redlands

Fernandes PG, Carreira P, da Silva MO (2008) Anthropogenic sources of contamination recognition—Sines coastal aquifer (SW Portugal). J Geochem Explor 98:1–14

Franco-Uría AC, López-Mateo E, Roca ML, Fernández-Marcos (2009) Source identification of heavy metals in pastureland by multivariate analysis in NW Spain. J Hazard Mater 165:1008–1015

Goovaerts P (1997) Geostatistics for natural resource evaluation. Oxford University Press, New York

Gorai AK, Kumar S (2013) Spatial distribution analysis of groundwater quality index using GIS: a case study of Ranchi Municipal Corporation (RMC) area. Geoinfor Geostat Overview 1:2

Gotelli NJ, Ellison AM (2004) A primer of ecological statistics, 1st edn. Sinauer Associates, Sunderland

Guler C, Kurt MA, Alpaslan M, Akbulut C (2012) Assessment of the impact of anthropogenic activities on the groundwater hydrology and chemistry in Tarsus coastal plain (Mersin, SE Turkey) using fuzzy clustering, multivariate statistics and GIS techniques. J Hydrol 414–415:435–451

Halim MA, Majumder RK, Nessa SA, Oda K, Hiroshiro Y, Jinno K (2010) Arsenic in shallow aquifer in the eastern region of Bangladesh: insights from principal component analysis of groundwater compositions. Environ Monit Assess 61:453–472. doi:10.1007/s10661-009-0760-9

Handa BK (1981) An integrated water-quality index for irrigation use. Indian J Agric Sci 51:422–426

Harvey CF, Swartz CH, Badruzzaman ABM, Keon-Blute (2002) Arsenic mobility and groundwater extraction in Bangladesh. Science 22:1602–1606

Heberger K, Milczewska K, Voelkel A (2005) Principal component analysis of polymer-solvent and filler-solvent interactions by inverse gas chromatography. Colloids Surf A 260:29–37

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

Horton RK (1965) An index systems for rating water quality. J Water Pollut Control Fed 37:300–306

Hu K, Li B, Lu Y, Zhang F (2004) Comparison of various spatial interpolation methods for non-stationary regional soil mercury content. Environ Sci 25(3):132–137

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York

Islam IR, Rahman M, Reza AHMS, Rahman M (2013) Groundwater geochemistry and its implication for arsenic enrichment and mobilization in shallow alluvial aquifers of Pakshi Union, Ishwardi, Pabna, Bangladesh. Int J Chem Mater Sci 1(4):069–078

Kitanidis PK (1997) Introduction to geostatistics: applications in hydrogeology. Cambridge University Press, Cambridge

Lattin J, Carrol D, Green P (2003) Analyzing multivariate data. Duxbury Press, Belmont

Li Z, Fang Y, Zeng G, Li J, Zhang Q, Yuan Q, Wang Y, Ye F (2009) Temporal and spatial characteristics of surface water quality by an improved universal pollution index in red soil hilly region of South China: a case study in Liuyanghe River watershed. Environ Geol 58:101–107

Li F, Huang J, Zeng G, Yuan X, Li X, Liang J, Wang X, Tang X, Bai B (2013) Spatial risk assessment and sources identification of heavy metals in surface sediments from the Dongting Lake, Middle China. J Geochem Explor 132:75–83

Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a black foot disease area in Taiwan. Sci Total Environ 313:77–89

Machiwal D, Jha MK (2015) Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. Region Stud, J Hydrol. doi:10.1016/j.ejrh.2014.11.005

Marko K, Al-Amri NS, Elfeki AMM (2014) Geostatistical analysis using GIS for mapping groundwater quality: case study in the recharge area of Wadi Usfan, western Saudi Arabia. Arab J Geosci 7:5239–5252. doi:10.1007/s12517-013-1156-2

Masoud AA (2014) Groundwater quality assessment of the shallow aquifers west of the Nile Delta (Egypt) using multivariate statistical and geostatical techniques. J Afr Earth Sci 95:123–137. doi:10.1016/j.jafrearsci.2014.03.006

Masoud AA, Atwia MG (2010) Spatio-temporal characterization of the Pliocene aquifer conditions in Wadi El-Natrun area, Egypt. J Environ Earth Sci 62(7):1361–1374

Mohan SV, Nithila P, Reddy SJ (1996) Estimation of heavy metal in drinking water and development of heavy metal pollution index. J Environ Sci Health A31:283–289

Naujokas MF, Anderson B, Ahsan H, Aposhian HV, Graziano JH, Thompson C, Suk WA (2013) The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. Environ Health Persp 121(3):295–302

Nayanaka VGD, Vitharana WAU, Mapa RB (2010) Geostatistical analysis of soil properties to support spatial sampling in a paddy growing alfisol. Trop Agric Res 22(1):34–44. doi:10.4038/tar.v22i1.2668

Nimic DA, Moore JN (1991) Prediction of water-soluble metal concentrations in fluvially deposited tailings sediments, Upper Clark Fork Valley, Montana, U.S.A. Appl Geochem 6:635–646

Nkansah K, Dawson-Andoh B, Slahor J (2010) Rapid characterization of biomass using near infrared spectroscopy coupled with multivariate data analysis: part 1 yellow-poplar (Liriodendron tulipifera L.). Bioresour Technol 101(2):4570–4576

Omo-Irabor OO, Olobaniyi SB, Oduyemi K, Akunna J (2008) Surface and groundwater water quality assessment using multivariate analytical methods: a case study of the Western Niger Delta, Nigeria. Phys Chem Earth 33:666–673

Prasad B, Bose JM (2001) Evaluation of the heavy metal pollution index for surface and spring water near a limestone mining area of the lower Himalayas. Environ Geol 41:183–188

Prasad B, Jaiprakas KC (1999) Evaluation of heavy metals in ground water near mining area and development of heavy metal pollution index. J Environ Sci Health A34:91–102

Rahman MS, Gagnon GA (2014) Bench-scale evaluation of drinking water treatment parameters on iron particles and water quality. Water Res 48:137–147. doi:10.1016/j.watres.2013.09.018

Reddy SJ (1995) Encyclopaedia of environmental pollution and control, vol 1. Karlia, Environmental Media, p 342

Reza AHMS, Jean J-S, Lee M-K, Liu C-C, Bundschuh J, Yang H-J, Lee J-F, Lee Y-C (2010a) Implications of organic matter on arsenic mobilization into groundwater: evidence from north-western (Chapai-Nawabganj), central (Manikganj) and south-western (Chadpur) Bangladesh. Water Res 44:5556–5574

Reza AHMS, Jean J-S, Yang H-J, Lee M-K, Woodall B, Liu C-C, Lee JF, Luo S-D (2010b) a). Occurrence of arsenic in core sediments and groundwater in the Chapai-Nawabganj district, Northwestern Bangladesh. Water Res 44(6):2021–2037

Sahu P, Sikdar PK (2008) Hydrochemical framework of the aquifer in and around East Kolkata Wetlands, West Bengal. India Environ Geol 55:823–835

Sahu BK, Panda RB, Sinha BK, Nayak A (1991) Water quality index of the River Brahmani at Rourkela industrial complex of Orissa. J Ecotoxicol Environ Monit 1:169–175

Sarbu C, Pop HF (2005) Principal component analysis versus fuzzy principal component analysis. A case study: the quality of Danube water (1985e 1996). Talanta 65(5):1215–1220

Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. Environ Model Soft 22:464–475

Singh KP, Malik A, Sinha S (2005) Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques: a case study. Anal Chim Acta 538:355–374

Smith AH, Lingas EO, Rahman M (2000) Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bull World Health Org 78:1093–1103

Tabachnick BG, Fidell LS (2007) Using multivariate statistics. Pearson/Allyn and Bacon, London

Varouchakis EA, Hristopulos DT (2013) Improvement of groundwater level prediction in sparsely gauged basins using physical laws and local geographic features as auxiliary variables. Adv Water Resour 52:34–49

Vasanthavigar M, Srinivasamoorthy K, Vijayaragavan K, Ganthi RR, Chidambaram S, Anandhan P, Manivannan R, Vasudevan S (2010) Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. Environ Monit Ass. doi:10.1007/s10661-009-1302-1

Webster R, Oliver M (2001) Geostatistics for environmental scientists. Wiley, Chichester

WHO (2011) Guidelines for drinking-water quality, 3rd ed, vol. 1, Recommendations. WHO, Geneva

Wunderlin DA, Diaz M, Ame MMV, Pesce SF, Hued AC, Bistoni M (2001) Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia River basin (Cordoba-Artgentina). Water Res 35:2881–2894

Zou J, Zhang J, Wu J, Zhong F, Gu T (1988) On organic pollution and its control in the Haibe estuarine area of the Bohai Bay. Stud Mar Sin 29:1–20