Environ Geochem Health (2015) 37:157–180
DOI 10.1007/s10653-014-9638-z

ORIGINAL PAPER

Modeling groundwater quality over a humid subtropical region using numerical indices, earth observation datasets, and X-ray diffraction technique: a case study of Allahabad district, India

Sudhir Kumar Singh · Prashant K. Srivastava · Dharmveer Singh · Dawei Han · Sandeep Kumar Gautam · A. C. Pandey

Received: 25 March 2014 / Accepted: 17 July 2014 / Published online: 3 August 2014

Abstract Water is undoubtedly the vital commodity for all living creatures and required for well-being of the human society. The present work is based on the surveys and chemical analyses performed on the collected groundwater samples in a part of the Ganga basin in order to understand the sources and evolution of the water quality in the region. The two standard indices such as water quality index and synthetic pollution index for the classification of water in the region are computed. The soil and sediment analysis are carried out with the help of X-ray diffractometer (XRD) for the identification of possible source of ions in water from rock and soil weathering. The dominant minerals which include quartz, muscovite, plagioclase, and orthoclase are reported in the area. The study further utilizes the multivariate statistical techniques for handling large and complex datasets in order to get better information about the groundwater quality. The following statistical methods such as cluster analysis (CA), factor analysis (FA), and principal component analysis (PCA) are applied to handle the large datasets and to understand the latent structure of the data. Through FA/PCAs, we have identified a total of 3 factors in pre-monsoon and 4 factors in post-monsoon season, which are responsible for the whole data structure. These factors explain 77.62 and 82.39 % of the total variance of the pre- and post-monsoon datasets. On the other hand, CA depicted the regions that have similar pollutants origin. The average value of synthetic pollution index of groundwater during pre-monsoon is 9.27, while during post-monsoon, it has been recorded as 8.74. On the other hand, the average values of water quality index of groundwater during pre-monsoon and post-monsoon seasons are found as 217.59 and 233.02,

S. K. Singh · A. C. Pandey
K. Banerjee Centre of Atmospheric and Ocean Studies, IIDS, Nehru Science Centre, University of Allahabad, Allahabad 211002, India

P. K. Srivastava (✉)
Hydrological Sciences (Code 617), NASA Goddard Space Flight Center, Room No. G208, Building 33, Greenbelt, MD 20771, USA
e-mail: prashant.k.srivastava@nasa.gov; prashant.just@gmail.com

P. K. Srivastava
Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD, USA

P. K. Srivastava · D. Han
Department of Civil Engineering, University of Bristol, Bristol, UK

D. Singh
Department of Chemistry, University of Allahabad, Allahabad 211002, India

S. K. Gautam
School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India

respectively. The study indicates that there occurs an extensive urbanization with gradual vast development of various small- and large-scale industries, which is responsible for degradation in water quality. The overall analysis reveals that the agricultural runoff, waste disposal, leaching, and irrigation with wastewater are the main causes of groundwater pollution followed by some degree of pollution from geogenic sources such as rock and soil weathering, confirmed through XRD analysis.

Keywords Water quality index · XRD · Synthetic pollution index · Ganga basin · Remote sensing and GIS · Multivariate analysis

Introduction

In the present scenarios, many countries are facing the problem of water scarcity, even the good quality of drinking water is not available for the human society (Gleick 2000). This situation is wide spreading day by day specially in most of the developing countries such as a country like India where majority of population depends on the availability of ground and surface water (Srivastava et al. 2012b). Water pollution in developing countries increases after industrialization, unprecedented population growth, and urbanization after the globalization era, i.e., 1990 onwards (Singh et al. 2013a); consequently, groundwater is also contaminated and affected by many factors such as the expansion of irrigation activities, industrialization, and urbanization (Krishna et al. 2009; Foster et al. 2002). The groundwater quality of most of the cities in India are now polluted and unsafe for drinking at many locations (Singh et al. 2004, 2013b). Hence, monitoring and conserving surface and groundwater resources are very much required for sustainable environment and to fulfill the freshwater demand (Sun et al. 1992).

Remote sensing and GIS are very useful tools that could be used for synoptic representation of different attributes of any area (Gupta and Srivastava 2010; Patel and Srivastava 2013; Patel et al. 2013). Land-use/land-cover (LULC) changes quantification through satellite remote sensing is one of the major application, and it is important for assessing global environmental change processes, making policies and optimizing the use of natural resources (Srivastava et al. 2012c). The LULC types, such as agricultural land and urban area associated with human activities, often affect both the surface and groundwater quality; hence, monitoring spatiotemporal changes is essential to understand the driving factors influencing the water quality of an area (Merchant 1994; Wu and Segerson 1995; Srivastava et al. 2013). On the other hand, geographic information system (GIS) is an important tool for spatial analysis and integration of spatial and nonspatial data to derive useful outputs (Singh et al. 2013a; Patel et al. 2011). It can be used for formulating a simple and robust water quality pollution assessment tool for rapid information generation and broadcasting to water resources managers and the public (Singh et al. 2012; Vasanthavigar et al. 2010). The concept of water quality to categorize water according to its degree of purity or pollution dated back to 1848 (Vidyalakshmi et al. 2013). Around the same time, the importance of water quality to public health was recognized in the United Kingdom (Snow 1856). Nowadays, a number of water quality index (WQI) methodologies have been developed to provide a simple method for expressing the quality of water based on the user needs (Vasanthavigar et al. 2010; Avvannavar and Shrihari 2008) and can be used to provide the overall summaries of water quality on a scientific basis (Kaurish and Younos 2007).

In conjunction with remote sensing and GIS, the application of multivariate analysis offers a detailed understanding of water quality parameters and possible factors that influence the water quality behavior (Srivastava et al. 2012b). The multivariate statistical techniques such as principal component analysis (PCA), factor analysis (FA), and cluster analysis (CA) offer a valuable tool for reliable and effective management of water resources (Srivastava et al. 2012a; Singh et al. 2004). Several authors use multivariate statistical techniques to characterize and evaluate surface and groundwater quality and have found them very useful for studying the variations caused by natural and anthropogenic factors (Shrestha and Kazama 2007; Singh et al. 2005; Vega et al. 1998). However, the multivariate techniques can give only intimation about the geogenic contribution; the results can be confirmed only if some more thorough analysis such as of X-ray diffraction (XRD) is incorporated (Singh et al. 2013a). XRD is a versatile technique that

can be used to identify any crystalline substances, such as most minerals and indeed many other substances when they are present in a mixture (Moore and Reynolds Jr 1989).

For understanding the groundwater quality, multivariate statistical techniques integrated with remote sensing, GIS, water quality indices (Srivastava et al. 2012b), and XRD could be used for the identification of the possible factor/sources that influence the water system. The present study assessed most of the above mentioned methodologies, by applying different valid standard techniques such as primary data (water quality) generation for the geologically heterogeneous area following the standard laboratory procedure, XRD analysis, and categorization of samples based on the WQI, SPI, and multivariate analysis in correspondence with remote sensing and GIS techniques. Therefore, the foremost objective of this research focuses on the utilization of different water quality indices formulations and using them for estimating the current pollution status of groundwater of the Gangetic basin and surroundings. Further, this study also presents the current land-use/land-cover (LULC) change pattern in the district to identify the potential zones where most of the pollution is occurring. The highly polluted areas are depicted using the GIS-based maps prepared from the water quality indices. The GIS-interpolated maps are further employed to see the seasonality in the water quality of the area. A detailed analysis of PCA/FA/CA is presented to uncover the latent structure of datasets. The possible driving factors are also studied using the XRD and satellite datasets to appraise the groundwater quality in the area with respect to geogenic and anthropogenic factors. At the last, recommendation measures are suggested for protection and restoration of the groundwater quality.

Materials and methodology

Study area

The study area (district Allahabad) is located between 24°47′N and 25°47′N latitudes and between 81°19′E and 82°21′E longitudes covering an area of 5,246 km^2 and lies in the southern part of the State in the Gangetic plain (older and younger alluvial plain) and adjoins the Vindhyan Plateau of India (Singh et al. 2013b). About 88 % of the annual rainfall is received during the monsoon season. July and August are the months of maximum rainfall. The normal annual rainfall in the district is 975.4 mm but the variation from year to year is appreciable. On average, there are about 48 rainy days in a year. From about the middle of November, the temperature begin to fall rapidly, and in January (the coldest month), the mean daily maximum is 23.7 °C. The heat in the summer season particularly in May and the early part of June is intense. May is usually the hottest month of the year with the mean daily maximum temperature at 41.8 °C and the mean daily minimum at 26.8 °C. The hot dry and often dusty westerly winds make the heat more intense during the daytime especially in the Trans-Yamuna tract due to the radiation from the stony outcrops. The relative humidity is 70–80 % during monsoon and progressively decreases in humidity (during the summers humidity is very low, i.e., 15–20 % only). The geographical location of the study area along with sampling points is indicated in Fig. 1.

Soil and hydrogeological characteristics of the study area

The soil, geomorphologic, and hydrogeological setup impact significantly the water quality of any area. During weathering, the dissolution of minerals adds their own contribution to the ground or surface water (often referred as geogenic contributions). An assessment of these factors can significantly help to understand the overall structure of the area. The stratigraphy of the area is represented through Table 1. All the thematic maps used in this study are acquired from the Geological Survey of India (GSI) http://www.portal.gsi.gov.in/.

Soil properties

Agro-climatic zone is a land unit in terms of major climate, superimposed on length of growing period (moisture availability period; FAO 1983), whereas an agro-ecological zone is the land unit carved out of agro-climatic zone superimposed on landform which acts as modifier to climate and length of growing period. The study area comes under the agro-ecological region with alluvium derived soils in most part of the study area except there is a small part of soil derived from Vindhyan region (Singh et al. 2013a). The soils of the study region are predominantly

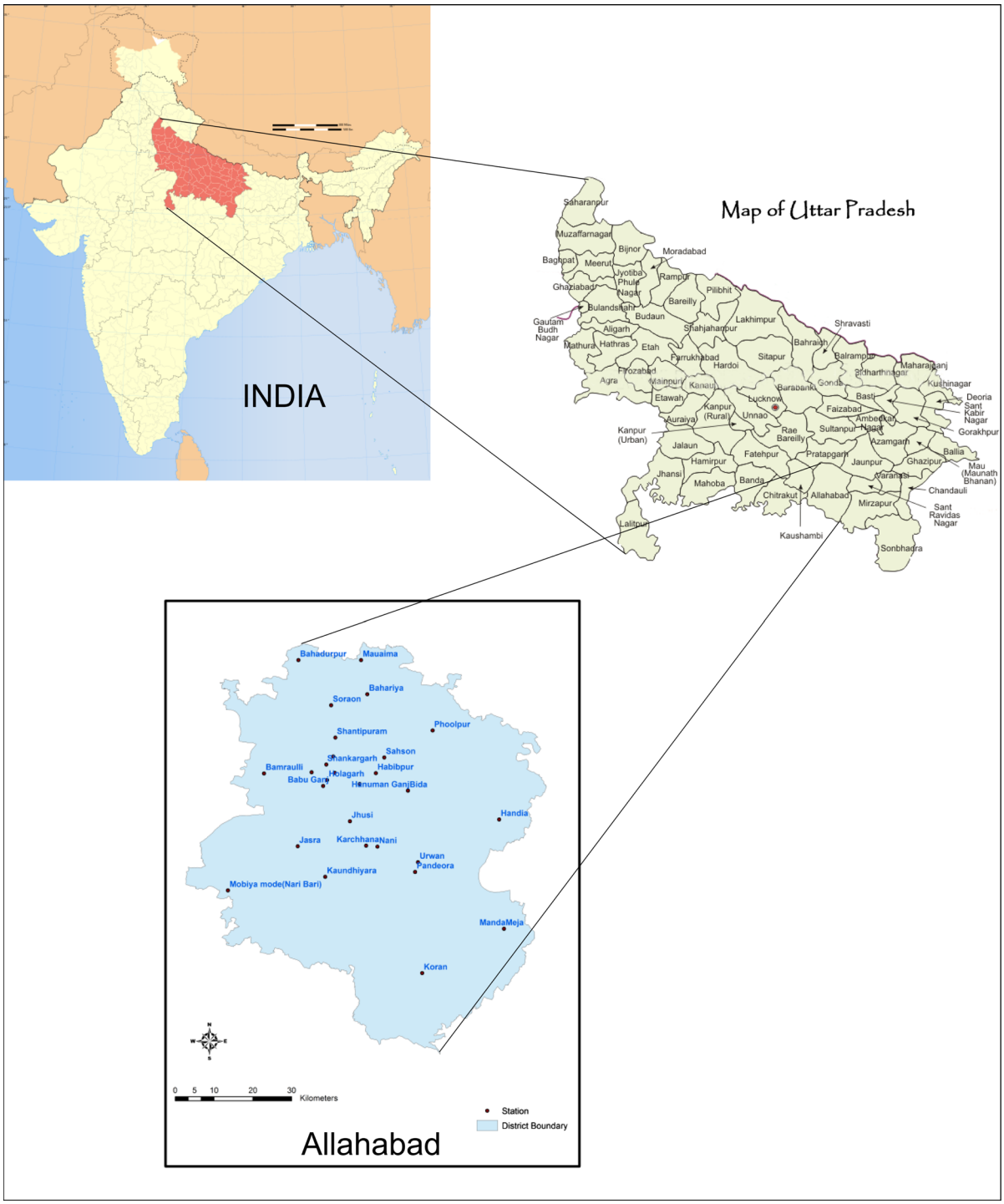

Fig. 1 Geographical location of the study area with sampling stations

medium-textured (loamy sand) followed by fine-textured (loamy clay). Soil drainage in the area can be categorized as well to moderately drained. The average pH value of the soil is 7.18, which means slightly alkaline; the average electrical conductivity is 0.10 dS/m; the average organic carbon percent is 0.74;

Table 1 Stratigraphy of Allahabad district

Group	Formation	Lithology	Age	Thickness (m)
Quaternary	Newer alluvium	Clay, silt, sand	Holocene	130.50
	Older alluvium	Polycyclic sequence of silt, clay and sand with kankar	Early to Late Pleistocene	–
Unconformity				
Vindhyan supergroup	Rewa group	Shale and sandstone	Meso to Neoproterozoic	15 76 18 1
	Kaimur group	Quartzite	Meso to Neoproterozoic	

the average total nitrogen is 235.87 kg/ha; the average total potash is 218.64 kg/ha; and the average available potassium 24.81 kg/ha. The overall condition of the soil is quite productive in nature. The dominant soil group is Typic Ustochrepts in the region.

Geomorphology

The region is mainly divided into two parts, i.e., the Ganga and the Yamuna alluvial plain, and the Vindhyan plateau. According to the Geological Survey of India 2001, commonly the following geomorphic features are present in the area. (a) Active Floodplain: It is quite localized and confined only to the river system (low relief characterized by thick deposition of clay, kankar, sand, and gravel). (b) Older Alluvial Plain: It is characterized by depositional and erosional terraces found in patches along the active plain. (c) Rocky surface (Denudational hills): These are prominent in Trans-Yamuna area formed mainly of quartzitic nature lies in Shankargarh, Koraon, Meja, and Manda. The palaeo-channel and abandoned channel are also found scattered in the region.

Denudational hills

These are the hill ranges, formed by differential erosion and weathering, and occupy the southwestern alignment of the area. The groundwater pollution of this zone is also considered to be in the poor category. The area under artificial recharge is quite considerable particularly in Trans-Yamuna area. Entire Shankargarh, Koraon, Manda, Meja blocks are suitable for artificial recharge. Construction of check dams, contour bunds, and ponds is under planning in the above-mentioned blocks. There is a need of Bori dam (type of earth fill dam) across the small drains (nalas), which is low cost and feasible from socioeconomic point of view. It is also suggested to implement subsurface dyke in Trans-Ganga area. The construction of Bori dam, check dams, and subsurface dyke could be useful specifically in regulating water quality in Trans-Yamuna area.

Alluvial plains

Most of the northern part of the study area is known as Doab region (a region lying between and reaching to the confluence of two rivers) and covered by alluvial plains, which are usually formed by the deposition of soil or sediments by the River Ganga and Yamuna, or their tributary sources. Due to high porosity and permeability, these zones are considered good for groundwater recharge and hence, are also very vulnerable to groundwater pollution in comparison with the confined aquifer zone. Water-level data of National Hydrological Services (NHS) for the last 10 years show that 60 % of the Trans-Ganga areas are prone to water logging while it is declining in Bahadurpur and Chaka block.

Unconsolidated sand, silt, and clay

Northern part of the district is covered by unconsolidated sand, silt, and clay. Due to physical properties

(high hydraulic conductivity) of the sand, silt, and clay, the unconsolidated aquifer systems are more susceptible to pollution. The unconsolidated aquifer needs special attention during pollutant removal.

River channel and floodplain deposits

River Ganga and River Yamuna are the two main rivers that along with a number of tributaries transport a large amount of silt, clay, and sand in the basin. The rivers are considered as good zones for groundwater recharge. The Ganga and the Yamuna rivers change their courses over a period of time and formed palaeo-channel that has good prospects of groundwater. The river channels are responsible for bringing a high amount of diffuse pollutants from upstream areas. Generally, near to the river system, their low-lying area bordering a stream is covered by coarse and fine sand, clay, and silt, which are directly deposited at the time of floods. The floodplain deposits are developed at the constructive side of river channels.

Hydrogeology

Geologically, the age of the Quaternary alluvium and the Vindhyan Plateau ranges from Proterozoic to recent. Quartzite of Kaimur group forms the basement in the area which is unconformably overlain by Quaternary alluvium. In the unconsolidated or alluvial formation, groundwater occurs under unconfined to confined conditions in the shallow and deeper aquifers, respectively, and depth to water ranges between 2 and 20 m during pre-monsoon period, while in the post-monsoon period, it stands between 1 and 18 m. In the consolidated formation (Vindhyans), water table ranges between 3 and 10 m below ground level (mbgl) during pre-monsoon period and 2–8 mbgl during post-monsoon period with seasonal fluctuation ranges between 1 and 4 m. The borehole data give information about the subsurface geological formations. The surface lithology in Trans-Ganga and Trans-Yamuna region has distinct characteristic. The alluvium plain is classified as Younger and Older alluvial plain. Older alluvium is again classified into two subdivisions (1) Banda Older alluvium and (2) Varanasi alluvium. Subsurface geological characters of hard rock area of Vindhyan region are quite distinct from Trans-Ganga area.

Hydrogeological setup

Groundwater exploration studies reveal three-tier aquifer systems in the alluvial area that have distinct granular zones, they are as follows: (1) shallow aquifers (I Aquifer Group) ranging from 0.0 to 110 mbgl; (2) middle aquifer (II Aquifer Group) ranging from 120 to 250 mbgl (meter below ground level); and (3) deeper aquifer (III Aquifer Group) lying below 260 down to depth 400 mbgl. The aquifer material is medium- to coarse-grained sand admixed with gravel at places. The tube well in alluvium thickness yield is of 2,000–3,000 lpm (liter per minute) while the tube well in Vindhyans region, fractures zones, exists down to 125 meters only and yield is of 500–1,000 lpm; within this sandstone domain, a silica sand horizon exists having a thickness of 5–40 m which also contains groundwater.

Field and laboratory analysis

Field samplings were taken place during the period from 2011 to 2013 (total 120 water samples). The selected sites were carefully chosen and fixed by global positioning system (GPS) (Garmin Model 780). From the periods 2011–2013, during pre-monsoon and post-monsoon seasons, a total of 60 groundwater samples from each year were collected in polypropylene bottles. All the groundwater samples are collected from the Ganga basin (Allahabad district), i.e., the alluvial and nonalluvial parts of the study area. Stratified random samplings were performed to ensure that all areas are covered. A total of 20 blocks were selected in this study on the basis of population density maintaining the same depths during each of the seasons. With the help of GPS, the sampling points were recorded in Universal Transverse Mercator projection along with the WGS 1984 coordinate system so that similar sites were sampled during next visits. Higher number of groundwater sample was collected from the denser areas, and therefore, maximum number of samples was collected from the Chaka block that consists of wards such as Shantipuram, Bamraulli, Jhusi, Sashon, and Teliyar Ganj. The sampling sites provide the data reported in this paper and are shown in Fig. 1. Collected samples are divided into three aliquots for subsequent analysis. After the appropriate preservation, all samples are stored at 4 °C until analyses. Collection and analyses are performed as specified in

the standard international methods (APHA 1998). The instrumentations used for the analysis remain same throughout the study periods to avoid any uncertainties and error in water quality analysis. Total alkalinity (as HCO_3^-) and physical parameters such as electrical conductivity (EC), total dissolved solids (TDS), and pH are measured in the field. Alkalinity is measured using a Hach field titration kit (through titration with 0.1 M HCl). The samples were acidified using conc. nitric for cation analysis. The major cations are (Mg^{2+}, Ca^{2+}, Na^+, K^+) analyzed using an atomic absorption spectrometer, and the F^- anions are analyzed using an ion chromatograph. Bicarbonate (HCO_3^-) is determined by titration method as described in the standard methods for the examination of water and wastewater (APHA 1998). Major anions (Cl^-, SO_4^{2-}, NO_3^-) for groundwater samples are undertaken by ion chromatography. The precision and accuracy of the analyses are within 5 % (evaluated through repeated analyses of standards and samples).

Remote sensing and GIS implementation

The land-use/land-cover classification of the area has been estimated on ENVI version 4.8 platform using the Landsat TM and ETM satellite data. The specifications about the satellite datasets used in this study are represented through Table 2, which are used here to accomplish the land-use/land-cover change. A supervised classification technique (maximum likelihood classification) as defined by Eq. 1 is used to determine the LULC of the study area. Thus, seven classes are delineated from the study area namely agricultural land, built-up area, cultivable land, forest, water body, wasteland, and other fallow land. Arc GIS (9.3) software is used for the preparation of primary thematic layers. The image projection parameters are used as Universal Transverse Mercator, World Geodetic System 84 (WGS84), Zone 44 N coordinate system. In this study, a post-classification change detection method is applied in order to find out the land-cover change over time. It requires the comparison of independently produced classified images (Banerjee and Srivastava 2013). The post-classification comparison method proved to be the most effective technique for comparing the changes occurred in two dates (Banerjee and Srivastava 2014). Post-classification method is the most suitable for detecting land-cover changes because it enables the estimation of the amount, location, and nature of change (Singh et al. 2014). This method separately classifies multitemporal images into thematic maps and implements the comparison of the classified images on a pixel-by-pixel basis (Lu et al. 2004).

$$D = \ln(a_c) - [0.5\ln(|\mathrm{cov}_c|)] - [0.5(\mathbf{X} - \mathbf{M}_c)T \quad (\mathrm{cov}_c - 1)(\mathbf{X} - \mathbf{M}_c)] \tag{1}$$

where D is the weighted distance; c is a particular class; X is the measurement vector of the particular pixel; M_c is the mean vector of the sample of class; a_c is percent probability that any particular pixel is a member of class c, (Defaults to 1.0); Cov_c is the covariance matrix of the pixels in the sample of class c; |Covc| is the determinant of Cov_c; $Cov_c - 1$ is the inverse of Cov_c; ln is natural logarithm function; and T = transposition function. For the proper utility of classified images, accuracy assessment is a much needed process. In order to evaluate the performance of the MLC classifiers, the accuracy assessment of the classified images was provided using the ground control points, assuring distribution in a rational pattern so that a specific number of observations was assigned to each category on the classified image. The Kappa accuracy was computed, as given by Eq. 2 (Bishop et al. 1975).

$$\kappa = \frac{N\sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r}(x_{i+})(x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+})(x_{+i})} \tag{2}$$

where r is the number of rows in the matrix, X_{ii} is the number of observations in row i and column i (the diagonal elements), x_{+i} and x_{i+} are the marginal totals of row r and column i, respectively, and N is the number of observations.

For GIS implementation, the Arc GIS version 9.3 was used for all the cartographic analysis and spatial interpolations. Inverse distance-weighted (IDW) procedure is used for preparing spatial interpolation maps based on water quality indices, which is a versatile, easy to use program, and fairly accurate under a wide range of conditions (Johnston et al. 2001) and used when the set of point is dense enough to capture the extent of local spatial variations (Srivastava et al. 2011). It is an exact method that enforces that the estimated value of a point is influenced more by

Table 2 Satellite datasets used in this study

Year	Satellite/sensor	Spatial resolution (m)	Path/row	Available band combination	Date of acquisition
1990	Landsat, TM	30	p143 r042 and p143 r043	1, 2, 3, 4, 5, 6, 7	November 17, 1990 and October 16, 1990
2011	Landsat, ETM$^+$	30	p143 r042 and p143 r043	1, 2, 3, 4, 5, 6, 7	October 18, 2011 and October 18, 2011

nearby points than those farther away (Chang 2002). By applying this method, the property at each unknown location for which a solution sought can be expressed mathematically (Johnston et al. 2001; Mitas and Mitasova 1999). The WQI obtained at the various groundwater collection sites is interpolated to obtain the index for the study area. The expression for IDW can be depicted using the Eqs. 3, 4, and 5:

$$\hat{Z}(S_o) = \sum_{i=1}^{n} \lambda_i(S_o) Z(S_i) \tag{3}$$

$$\hat{Z}(S_o) = \lambda_o^T Z \tag{4}$$

$$\lambda_i(S_o) = \frac{\frac{1}{d^{\beta}(S_o,S_i)}}{\sum_{i=0}^{n} \frac{1}{d^{\beta}(S_o,(S_i)}}; \quad \beta > 1 \tag{5}$$

where λ_i is the weight for neighbor i (the sum of weights must be unity to ensure an unbiased interpolator); $d(S_o, S_i)$ is the distance from the new point to a known sampled point; β is a coefficient that is used to adjust the weights; and n is the total number of points in the neighborhood analysis.

Water quality indices estimation

Water quality index provides a clear picture about the usability of water for different purposes such as drinking, irrigation, and industrial usage. However, it is difficult to simplify surface and groundwater quality to a specific index because of its sensitive nature to inputs received from sources such as geogenic contribution, water–rock reactions, agricultural runoff, domestic, and industrial wastes (Singh et al. 2012). However, the WQI by (Tiwari and Mishra 1985) and synthetic pollution index (SPI) (Ma et al. 2009) are very useful and efficient methods for assessing the quality of water and presently used by many scientists and water managers. To determine the suitability of the water for drinking purposes, WQI can be estimated by using the following methodology:

$$\mathrm{WQI} = \mathrm{Anti}\log\left[\sum_{i=1}^{n} W_i \log_{10} q_i\right] \tag{6}$$

where W_i is the weighting factor computed using the Eq. 7

$$W_i = K/S_i \tag{7}$$

K is proportionality constant derived from Eq. 8

$$K = \left[\frac{1}{\left(\sum_{i=1}^{n} 1/S_i\right)}\right] \tag{8}$$

where S_i is the World Health Organization (WHO 1984)/Indian Council Medical Research (ICMR 1975) standard values of the water quality parameter.

Quality rating (q) is calculated using the formula (Eq. 9),

$$q_{ni} = [(V_{\text{actual}} - V_{\text{ideal}})/(V_{\text{standard}} - V_{\text{ideal}})] \times 100 \tag{9}$$

where q_{ni} is the quality rating of ith parameter for a total of n water quality parameters, V_{actual} is the value of the water quality parameter obtained from laboratory analysis, V_{ideal} is the value obtained from the standard tables, and V_{standard} is Indian Standard Institution (ISI) (ISI 1991) standard of the water quality parameters (Table 3). The rating and category chart for WQI is represented through Table 4.

Another index which can be used to integrate the impact of various pollutants on the water quality is synthetic pollution index (SPI) (Ma et al. 2009). This approach is widely used because it provides a simpler overview (Ouyang et al. 2006). The index is calculated using the following Eqs. 10 and 11:

$$P_r = \sum_{i=1}^{n} P_i \tag{10}$$

$$P_i = \frac{C_i}{C_{io}} \tag{11}$$

Table 3 Different standards given by Indian Standard Institute (ISI), WHO, and ICMR with their permissible and desirable limits

Parameter	ISI highest desirable b (V_{ideal}, mg/L)	ISI maximum permissible ($V_{standard}$, mg/L)	ICMR/WHO standards (S_n)
pH	7	8.5	8.5
BOD	4	6	5
COD	10	15	14.5
Electric conductivity (μS/cm)	1,000	2,250	600
Total hardness	500	1,500	1,000
Fluoride	0.6	1.2	1
Chloride	250	1,000	250
Alkalinity	200	600	120
Sulfate	200	400	250
Nitrate	10	45	50
Calcium	75	200	75
Magnesium	30	100	75
Sodium	30	200	200
Potassium	20	100	100
Phosphate	0.001	0.05	1.5

Table 4 Rating and category chart of WQI

Sl. no.	WQI	Water quality
1	0–25	Suitable
2	26–50	Polluted
3	51–75	Moderately polluted
4	76–100	Highly polluted
5	>100	Unfit

Table 5 Rating and category chart of SPI

Sl. no.	Synthetic pollution index (SPI)	Category pollution
1	<0.5	Suitable
2	0.5–3	Polluted
3	3–5	Moderately polluted
4	5–10	Highly polluted
5	>10	Unfit

where P_r is the synthetic pollution index, P_i is the pollution index of pollutant i, C_i is the measured concentration of pollutant i, C_{io} is the evaluation criteria of pollutant i. The criteria used in monitoring sections are from the corresponding standards in the Water Quality Standard (BIS). The rating and category chart for SPI is represented through Table 5.

X-ray diffraction technique (XRD)

The XRD method is ideal for the qualitative assessment of mineralogy. Samples of <200 mesh size have been used for the mineralogical studies of soil and sediments to decipher the mineral assemblages, using XRD. The XRD method explains the geometry or shape of crystalline materials using X-rays, and is based on the elastic scattering of X-rays from structures that have a long-range order. The interaction of the incident rays with the sample produces constructive interference (and a diffracted ray) when the conditions satisfy Bragg's Law (Kacher et al. 2009; $n\lambda = 2d \sin \theta$). Here, d is the spacing between diffracting planes, θ is the incident angle, n is any integer, and λ is the wavelength of the beam. This law relates the wavelength of electromagnetic radiation to the diffraction angle and the lattice spacing in a crystalline sample (Cole 1970). These diffracted X-rays are then detected, processed, and counted. Conversion of the diffraction peaks to d-spacing allows identification of the mineral because each mineral has a set of unique d-spacing. Typically, this is achieved by comparison of d-spacing with the standard reference patterns (Moore and Reynolds Jr 1989). To estimate the geogenic contribution of ions to groundwater, soil and sediments samples are analyzed on a XRD machine at Sophisticated Analytical Instrumentation Facility (SAIF), Chandigarh, Punjab, India laboratory, to uncover the mineralogical properties of the area.

Multivariate statistical method

The application of multivariate statistical method is very useful for classification, modeling, and interpretation of large datasets which allow the reduction in dimensionality of the large datasets (Singh et al. 2005). CA and FA/PCA techniques are applied for multivariate analysis of datasets of groundwater quality. In CA analysis, it groups the objects (cases) into classes (clusters) on the basis of similarities/dissimilarities within or between classes (Singh et al. 2009), respectively. CA in this study is used to see the pattern in the datasets implemented using hierarchical agglomerative clustering technique by means of the

squared Euclidean distances following Ward's method (Srivastava et al. 2012b). PCA is applied after standardizing the datasets through the z-scale transformation to avoid any misclassification (Krishna et al. 2009). The principal component (PC) is expressed as (Eq. 12):

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{3j} + \cdots + a_{im}x_{mj} \quad (12)$$

where a is the component loading, z the component score, x the measured value of a variable, i the component number, j the sample number, and m the total number of variables. The FA analysis attempts to reduce the contribution of less significant variables obtained from PCA and the new group of variables known as varifactors (VFs). VFs are extracted through rotating the axis defined by PCA. In FA, the basic concept is expressed in Eq. 13,

$$z_{ji} = a_{f1}f_{1i} + a_{f2}f_{2i} + a_{f3}f_{3i} + \cdots + a_{fm}f_{mi} + e_{fi} \quad (13)$$

where z is the measured value of a variable, a the factor loading, f the factor score, e the residual term accounting for errors or other sources of variation, i the sample number, j the variable number, and m the total number of factors.

Results and discussion

Hydrochemistry of groundwater

The descriptive statistics of 12 physicochemical parameters at the 30 locations of groundwater of the Ganga River Basin during the pre- and post-monsoon are given in Table 6. They present the univariate overview of the chemistry of groundwater in the area. The mean pH (potential hydrogen) values of the 30 locations showed very little variation in both the pre- and post-monsoon seasons, ranging from 7.30 to 7.40, respectively, implying slight alkalinity of groundwater. The highest pH of 7.9 and 8.01 is reported in Urwa both during pre- and post-monsoon. This may be the result of the industrial and agricultural activities in the region. The mean electrical conductivity is 829 µS/cm and 891.07 µS/cm during both pre- and post-monsoons. Highest EC 1289 µS/cm is reported at Babu Ganj in pre-monsoon samples and 1,805 µS/cm at Shankargarh samples in post-monsoon season. The mean TDS values are 718 and 671 mg/l in pre- and post-monsoon, respectively. The highest values of TDS are 1,242 and 1,364 mg/l at Babuganj in both pre- and post-monsoon seasons respectively. The mean fluoride value are 0.55 and 0.71 mg/l in pre- and post-monsoon seasons. The highest value of fluoride is 1.3 mg/l at Karchhana during pre-monsoon and 1.65 mg/l at Holagarh during post-monsoon. The mean chloride values are 102 and 91.23 mg/l in pre- and post-monsoon seasons, and highest value of chloride in pre-monsoon is the value 216 mg/l at Shantipuram and in post-monsoon, the value is 280 mg/l recorded at Phoolpur. The mean values of bicarbonate are 366 and 340.36 mg/l in pre- and post-monsoon seasons. The highest value of bicarbonate value is 596 mg/l in pre-monsoon and value is 534 mg/l in post-monsoon at Babu Ganj, respectively. The mean values of sulfate are 64.7 and 71.7 mg/l in pre- and post-monsoon, and highest value is 189 mg/l in Phoolpur and 196 mg/l in Babuganj during pre- and post-monsoon, respectively. The mean values of nitrate are 0.77 and 0.62 mg/l in pre- and post-monsoon seasons with highest values 1.12 mg/l at Pandeora and 0.98 mg/l at Mobiya mod. The mean values of calcium are 39.5 and 19.6 mg/l in pre- and post-monsoon seasons, and the highest values of 69 and 514 mg/l are recorded at Bamraulli and Teliyar Ganj. The mean values of magnesium are 54.9 and 37.96 mg/l in pre- and post-monsoon seasons, respectively, and the highest values of 127 and 108 mg/l are recorded at Babu Ganj in both the seasons. The mean values of sodium are 75.6 and 96.53 mg/l in pre- and post-monsoon seasons with the highest value of 137 mg/l at Kaundhiyara and 221 mg/l at Mobiya Mod. The mean values of potassium are 13.7 and 12.16 mg/l in pre- and post-monsoon seasons with highest values of 56 mg/l at Bamraulli and 55 mg/l at Teliyar Ganj. There is no block in the district identified as clean area as nearly on all site above permissible limit of water quality parameters are recorded.

Water Quality Index (WQI) and Synthetic Pollution Index (SPI)

Water quality index values and synthetic pollution index are computed for groundwater. The WQI table of different stations for groundwater is indicated in Table 7 estimated using the water quality index equations given in "Water quality indices estimation" section. All the values calculated are explicitly higher than the limits given in "Water quality indices estimation" section, indicating very high pollution status of the groundwater during the monsoon period. Through the

Table 6 Physicochemical properties of the groundwater samples during the two seasons (except EC all other parameters units are in mg/l and EC units is μS/cm)

Parameter	Pre-monsoon season			Post-monsoon season		
	Min	Max	Average	Min	Max	Average
pH	6.80	7.90	7.30	8.01	7.04	7.40
EC	545.00	1,289.00	829.23	1,805.00	476.00	891.07
TDS	425.83	1,242.23	717.65	1,364.31	352.96	670.91
F^-	0.12	1.30	0.55	1.65	0.21	0.71
Cl^-	16.00	216.00	102.30	280.00	17.00	91.23
HCO_3^-	203.00	596.00	365.60	534.00	157.00	340.37
SO_4^{2-}	16.00	189.00	64.73	196.00	16.00	71.70
NO_3^-	0.25	1.12	0.77	0.98	0.11	0.63
Ca^{2+}	26.00	69.00	39.50	51.00	8.00	19.60
Mg^{2+}	19.00	127.00	54.87	108.00	12.00	37.97
Na^+	16.00	137.00	75.63	221.00	13.00	96.53
K^+	3.00	56.00	13.70	55.00	2.00	12.17

application of GIS and spatial interpolation method (IDW), pollution zone categorization of groundwater can be provided for better management of water resources. These interpolation maps are created for pre- and post-monsoon seasons separately to see the seasonal effects. The analysis indicates that the maximum (max) (14.14) and minimum (min) (6.54) value of synthetic pollution index is reported at site number 9 and at site number 19, respectively, with standard deviation of 1.67 in pre-monsoon season. The maximum (14.75) and minimum (5.11) value of synthetic pollution index during post-monsoon is reported with a standard deviation of 2.51. The average value of synthetic pollution index of groundwater during pre-monsoon is 9.27, and in post-monsoon, it is recorded as 8.74. This suggests that the average number of groundwater samples has more pollution during the pre-monsoon condition as compared to post-monsoon. The maximum (1,182.6) and minimum (8.92) value of water quality index is reported at site number 20 and at site number 25, respectively. The average value of water quality index of groundwater during pre-monsoon and post-monsoon seasons is 217.59 and 233.02. All the analysis suggests that groundwater lies in the category of highly polluted water.

LULC- and GIS-based assessment of groundwater quality

Land-use and land-cover change analysis (Fig. 2) of the study area has been achieved through Landsat satellite image after classification. On the basis of the overall accuracy results (OA), it can be seen that the highest accuracy was achieved for the year 2011 classified image (89.05 %) followed by 1990 (85.39 %) satellite images. The accuracy assessment results indicate that the land-use changes have been accurately identified and extracted after classification during two periods, confirmed by the reasonable overall accuracies values. The data presented in Table 8 represent the area of each land-use and land-cover category of the different years. The change detection results show that the seven land-use categories (cultivable land, forest, built-up area, other fallow land, agriculture, wasteland, and water body) have changed significantly in the study area during the last 20-year period. The land-use patterns in the district are changing slowly but steadily. Specifically, the built-up area increases from 555.89 km^2 in 1990 to 744.16 km^2 in 2011 with a percentage increase of 33.87 %. This increase has probably taken place due to the migration of population toward Allahabad due to better educational activities, business opportunity, and availability of better urban infrastructure facility. The total area of cultivable land decreases from 246.19 km^2 in 1990 to 116.97 km^2 in 2011 with a percentage decrease of 52.49 %. The decrease may be mainly due to expansion in urban area. The area of other fallow land decreased from 186.04 km^2 in 1990 to 125.88 km^2 in 2011 with a percentage decrease of 32.34 %, which can be attributed to urbanization, some reforestation, and industrial activities. Some change in water body class is found to

Table 7 WQI and SPI values estimated during the two seasons

Site	Pre-monsoon (WQI)	Post-monsoon (WQI)	Pre-monsoon (SPI)	Post-monsoon (SPI)
Site1	352.30	39.36	8.29	10.33
Site2	351.89	27.92	9.59	11.29
Site3	238.44	83.91	8.45	7.59
Site4	152.08	68.56	8.87	8.09
Site5	141.98	92.03	6.94	5.84
Site6	374.67	100.04	7.04	5.79
Site7	180.32	98.91	9.10	9.05
Site8	163.30	109.19	9.02	8.53
Site9	381.56	88.92	14.14	14.75
Site10	248.98	40.41	12.84	14.64
Site11	60.00	21.01	10.71	9.62
Site12	8.24	131.29	10.48	10.25
Site13	414.61	125.40	8.23	5.78
Site14	370.72	242.44	7.86	5.76
Site15	19.50	221.56	10.37	12.41
Site16	44.77	103.57	10.59	12.40
Site17	616.80	139.39	9.93	10.91
Site18	192.54	318.84	8.75	9.92
Site19	36.89	331.24	6.54	5.11
Site20	16.32	1,182.60	6.70	7.40
Site21	123.57	945.33	9.29	8.52
Site22	106.82	130.71	9.57	8.11
Site23	141.81	100.09	10.02	7.20
Site24	141.45	32.16	10.63	6.48
Site25	509.47	8.92	10.58	7.64
Site26	619.19	105.33	9.85	7.52
Site27	121.75	1,118.76	7.80	6.83
Site28	106.27	928.11	8.41	7.90
Site29	143.88	21.65	9.00	8.79
Site30	147.46	33.01	8.65	7.73
Max	619.19	1,182.6	14.14	14.75
Min	8.24	8.92	6.54	5.11
Avg	217.59	233.02	9.27	8.74
SD	170.97	335.47	1.67	2.51

be 328.08 km^2 in 1990, decreased to 298.00 km^2 in 2011 with a percentage decrease of 9.17 %. This decrease in water body is due to unplanned use of water and may be due to climate change. The forest area is 335.87 km^2 in 1990, which decreased to 311.92 km^2 in 2011 indicating a percentage decrease of 7.13 %. This decrease can be attributed to some deforestation activities. There is a change seen in the agriculture area, which increased from 3,638.40 km^2 in 1990 to 3,793.25 km^2 in 2011 with a percentage increase of 4.26 % that can be attributed to decrease in fallow and wasteland. The area of wasteland has shown a declining trend from 1990 to 2011, and in year 1990, it is 279.62 km^2, which decreases to 179.91 km^2 in two decades with percentage change of 35.66 %.

The impact of LULC changes on the water quality of the area has been evaluated using the "best guess" method given by (Reynard et al. 2001; Srivastava et al. 2013) for future land-use change scenarios. A total of 6 cases are produced for detecting the impact of LULC on groundwater quality. These cases are generated on the change detection possibilities from the classified images. All the possible substantial change cases in LULC are considered in this study such as cultivable land to built up (Case I), wasteland to agricultural land (Case II), water body to built up (Case III), other fallow land to built up (Case IV), fallow land to agriculture (Case V), and fallow land to forest (Case VI). To analyze the result more on a remote sensing platform, DEM is also taken into account. DEM of the Allahabad district (Fig. 3) is utilized in this study for generating the slope. Slopes affect a lot the movement of pollutants and show the direct influence of the rainfall on the contaminant transport (Srivastava et al. 2012b). The wells at the highest location will generally have the safer water supply. The wells furthest downslope would receive the combined effluent from the other houses, industries as well as from agricultural fields, because the liquid effluent follows the same path as the surface runoff or snowmelt and hence contaminates water more toward downside slopes (Waller 2001). The high rainfall and associated runoff in monsoon facilitate the movement of contaminants and affect the post-monsoon season water quality. The analysis of DEM and surface elevation map suggests that, in general, the surface runoff movement is toward northwest direction. Hence, the stations on the down slope receive higher pollutants after rainfall than upstream stations due to runoff. Most of the regions in the northwest downstream are well occupied by urban population and are already categorized as polluted because of high WQI and SPI values. The runoff from high altitude could be an additional possible source of pollutants in these regions.

The GIS-interpolated figures (Fig. 4) are used to visualize the cartographic influence of the water quality in area. The analysis indicates that none of the area is in the category under excellent class. WQI indicates that large areas are found above the limit of 100 indicating that the large percentage of water is

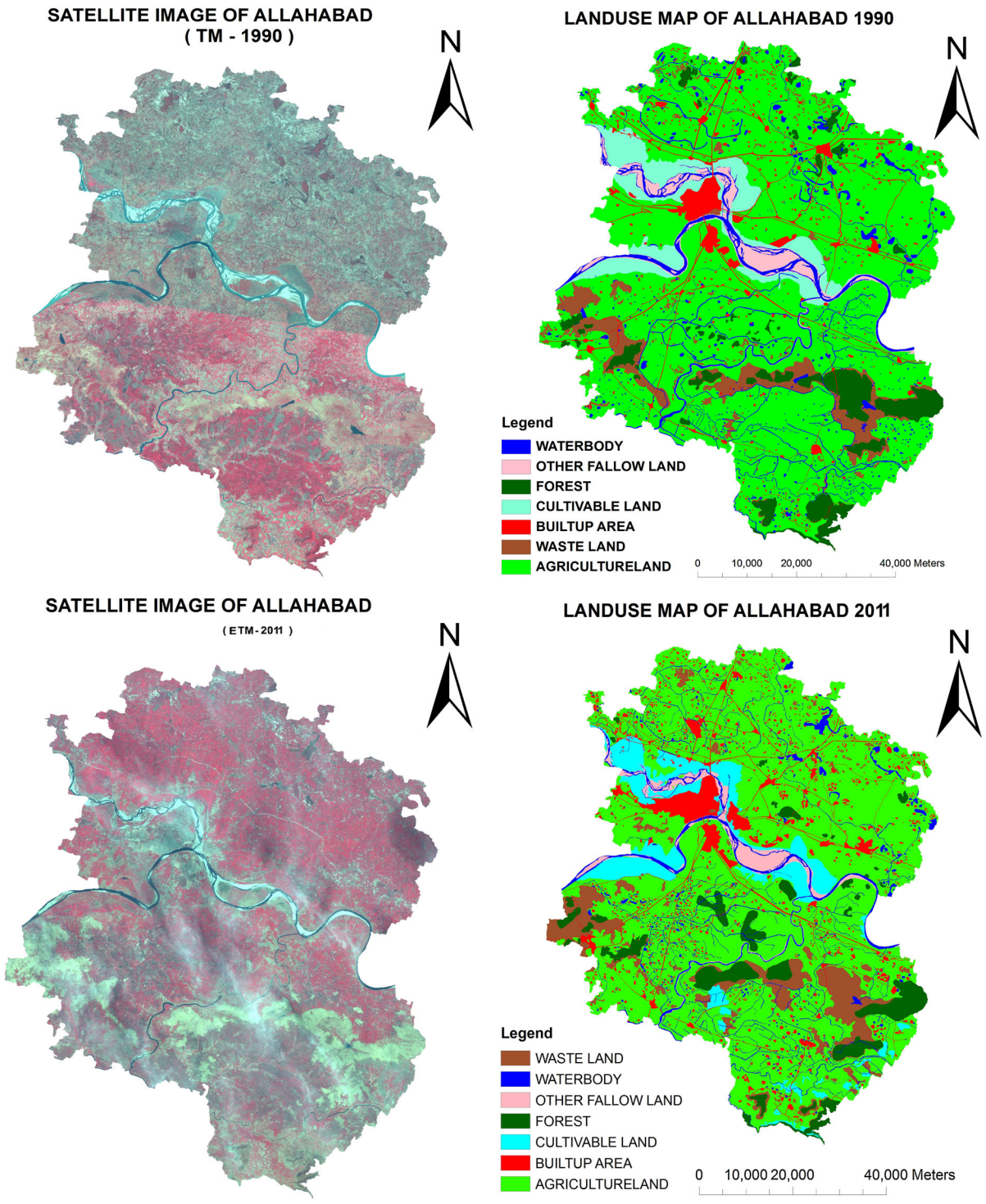

Fig. 2 Land use/land cover of the study area estimated during the year 1990–2011

Table 8 Land-use and land-cover change distribution from 1990 to 2011

LULC categories	AREA in 1990 (km^2)	AREA in 2011 (km^2)	1990–2011 (km^2)	%Difference
Forest	335.87	311.92	−23.95	−7.13
Wasteland	279.62	179.91	−99.71	−35.66
Built-up area	555.89	744.16	188.27	33.87
Cultivable land	246.19	116.97	−129.22	−52.49
Agriculture	3,638.40	3,793.25	154.85	4.26
Water body	328.08	298.00	−30.08	−9.17
Other fallow land	186.04	125.88	−60.16	−32.34

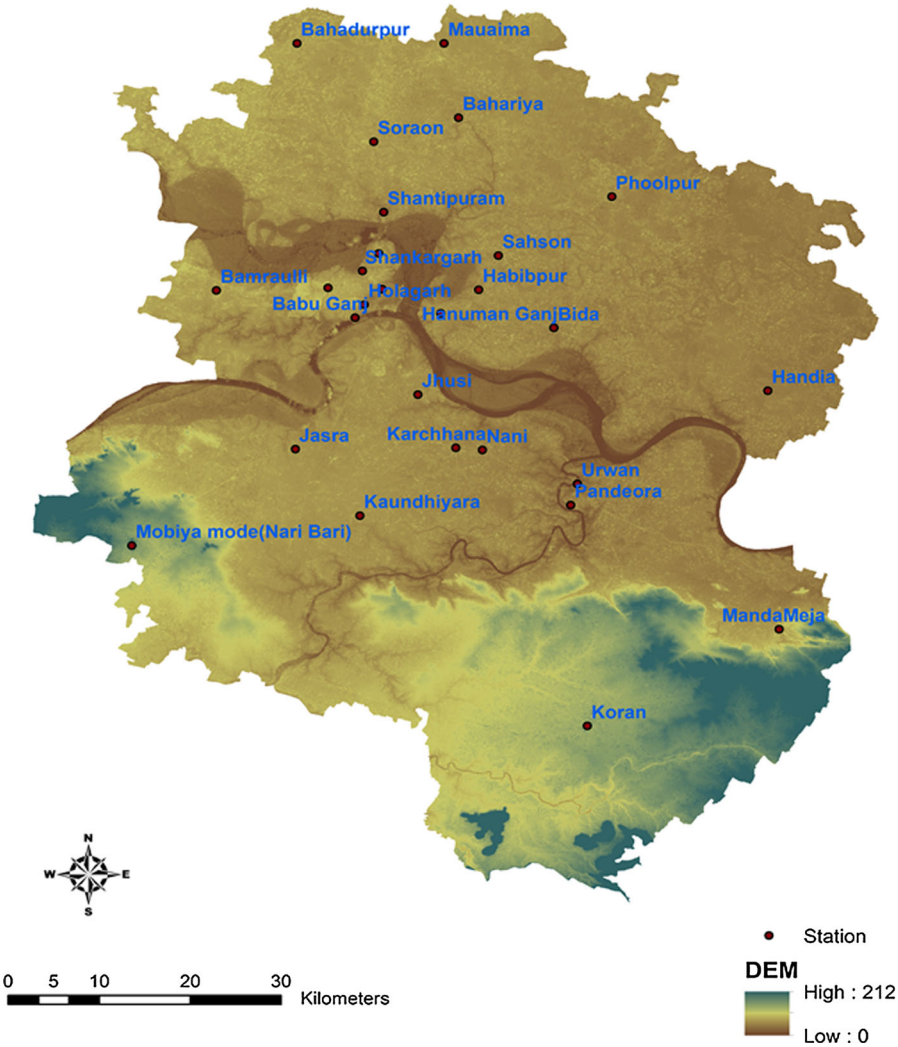

Fig. 3 Digital elevation model (DEM) of the area

unfit for drinking. Only a very small percentage of the area is found suitable for drinking water. The WQI values in the pre-monsoon season indicate that the Koraon, Naini, and Bahariya stations are found as the most contaminated and groundwater over here is nearly unfit for drinking. These areas are mostly

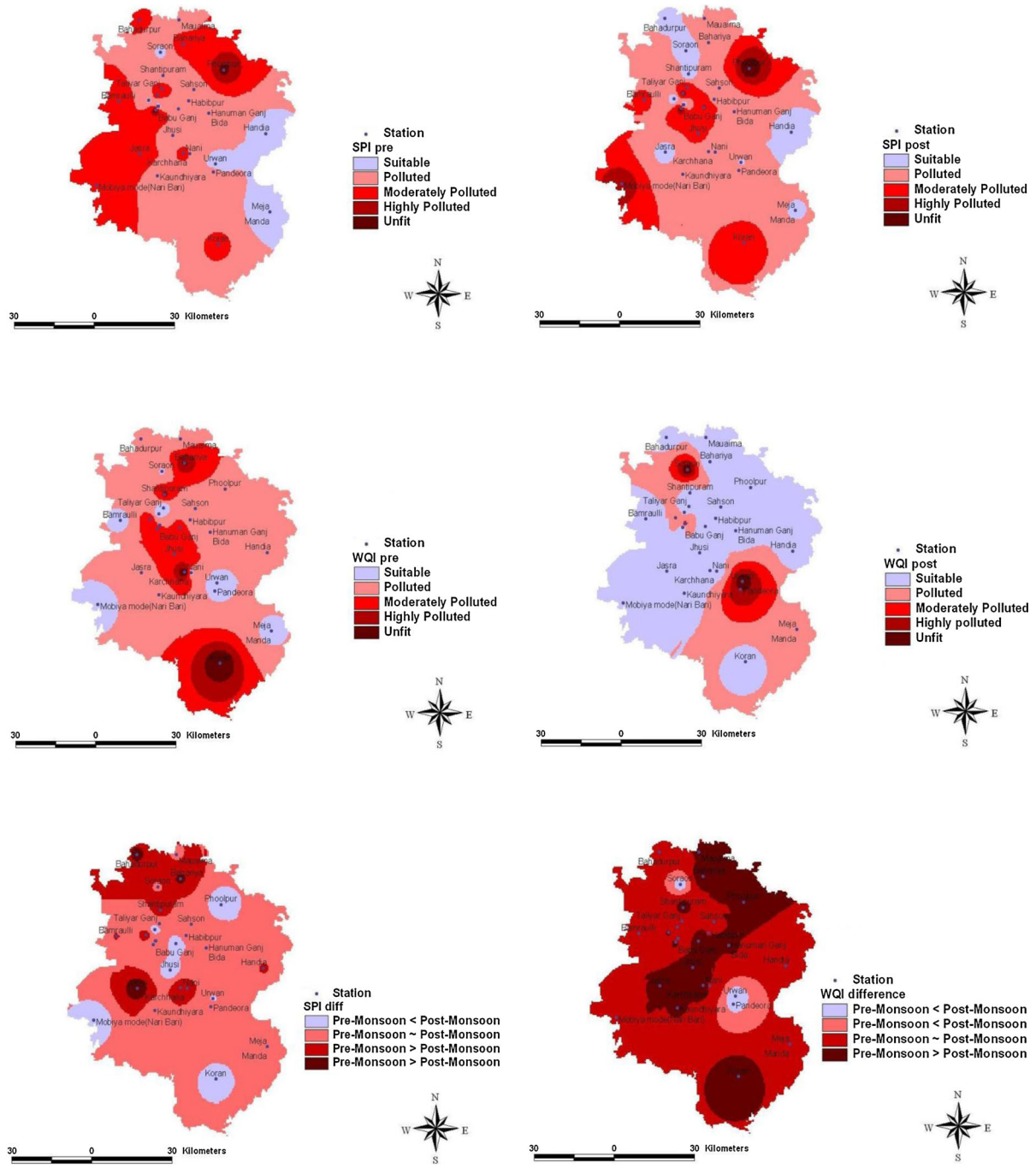

Fig. 4 Spatial interpolation of the WQI and SPI over the region

affected by agricultural and industrial practices (possible LULC Case I and Case IV). High agricultural influence is found in the areas considered under highly polluted zone (possible LULC Case II and Case V). The surrounding areas such as Shantipuram, Jhusi, Babuganj, and Karchhana are also found polluted with

moderate level of pollutant load (possible LULC Case IV). Rest of the stations has a limited influence with the discharge from agricultural fields, industrial, and domestic area (possible LULC Cases II, III, and IV). The suitable areas which are found good for drinking water extraction with very low level of pollution are Taliyar Ganj, Bamrauli, Urwa, Pandeora, Meja, and Mobiya mod (possible LULC Case VI). However, in post-monsoon season, a different pattern is obtained with some stations such as Urwa and Pandeora that are good area for ground extraction but after monsoon, they moved to highly contaminated area. It may be ascribe to high leaching nitrate- and phosphate-based ions in the water and indicate that these sites come under shallow to slightly deep aquifer category. On the other hand, with Koraon, its groundwater quality gets improved in post-monsoon season, which may be because of high infiltration of freshwater in the groundwater well. Some areas such a Shantipuram, Jhusi, Babuganj, and Karchhana, which are in moderately polluted zone, are now in suitable category in post-monsoon season. To understand the results more clearly, a different map between pre- and post-monsoon seasons is prepared to see which areas undergo most notable changes. The result shows that the maximum changes are found with the sites Urwa, Pandeora, and Soraon in the post-monsoon season. Nearly 50 % of areas are remained unchanged from pre- to post-monsoon seasons. Some of the areas such as Mau Aima, Bahariya, Jhusi, Jasra, and Karchhna do not changed much during the post-monsoon season.

To assess the performance of groundwater quality for irrigation purposes, SPI is calculated for the area and depicted through spatially interpolated plots. The analysis indicates that in pre-monsoon season, the Phoolpur area is not fit for irrigation purpose followed by Bahariya, Jasra, Bamrauli, Taliyar Ganj, Bahadurpur, Karchhna, and Koraon. The areas such as Handia, Urwa, Meja, and Soraon are found suitable for agricultural purposes. Rest of the area is found polluted for irrigation purpose. In post-monsoon season, the Phoolpur remains unfit for the irrigation purpose along with one more addition, i.e., Mobiya mod. The stations such as Babuganj, Jhusi, Koraon, and Bamrauli come under moderately polluted category in post-monsoon season. Some stations such as Bahadurpur, Soraon, Shantipuram, Jasra, Meja, Urwa, and Handia are found suitable in post-monsoon season for irrigation purposes. Rest of the stations is found in the category polluted for irrigation purposes. To see the maximum changes in SPI during the seasonal transition, a different map is generated, which indicates that Phoolpur has undergone maximum changes from pre- to post-monsoon season along with the stations such as Babuganj, Jhusi, Koraon, and Mobiya mod. Three stations Bahadurpur, Bahariya, and Jasra indicate a substantially higher value in pre-monsoon season than post-monsoon season. Owing to excess extraction of groundwater, harmful substances are getting concentrated and affecting the quality of water. Rest of the areas shows nearly no significant changes from pre- to post-monsoon season transition.

Fig. 5 XRD estimates of soil mineralogy **a** quartz, muscovite, plagioclase, orthoclase, vermiculite at location Bahadurpur; **b** quartz, muscovite, plagioclase, orthoclase at locations Soraon, Mobiya mode (Nari Bari), Jasra, Koraon, Jhusi, Naini, Karchhana, Urwa, Pandeora, Habibpur, Bahariya, Teliar Ganj, Hanuman Ganj, Babu Ganj, and Holagarh; **c** quartz, muscovite, plagioclase, orthoclase, calcite at locations Bamraulli, Saraye Inayat, Shankargarh, and Bara Thesil; **d** quartz, muscovite, plagioclase, orthoclase, halite at locations Kaundhiyara; **e** quartz, muscovite, orthoclase at locations Meja, Manda, Sahson; **f** quartz, muscovite at locations Handia and Phulpur; **g** quartz, muscovite, plagioclase, orthoclase, calcite, dolomite at locations Shantipuram and Andava mode ▶

Geogenic contribution and XRD estimates

The weathering and dissolution of minerals can be used to see the contribution from the geogenic sources into the groundwater. To estimate the possible sources of ions from rock weathering, XRD analysis is taken into account. The soil and sediment samples are analyzed on XRD (PANalytical), and the information about the minerals present is obtained through X'pert High Score Plus software. The analysis by XRD indicates the presence of muscovite, plagioclase, orthoclase, vermiculite, calcite, halite, and dolomite along with major mineral quartz (Fig. 5). Among the sampling locations, the minerals quartz, muscovite, plagioclase, orthoclase, and vermiculite are found dominant at sampling site 1. At site 7, quartz, muscovite, plagioclase, orthoclase, and calcite and at site 8 quartz, muscovite, plagioclase, orthoclase, and halite are detected in the XRD spectra. The sites 23 and 30 indicate the presence of minerals such as quartz, muscovite, plagioclase, orthoclase, calcite, and dolomite. At the sites 2-12, 18, 20, 21, 22, 25, 26, and

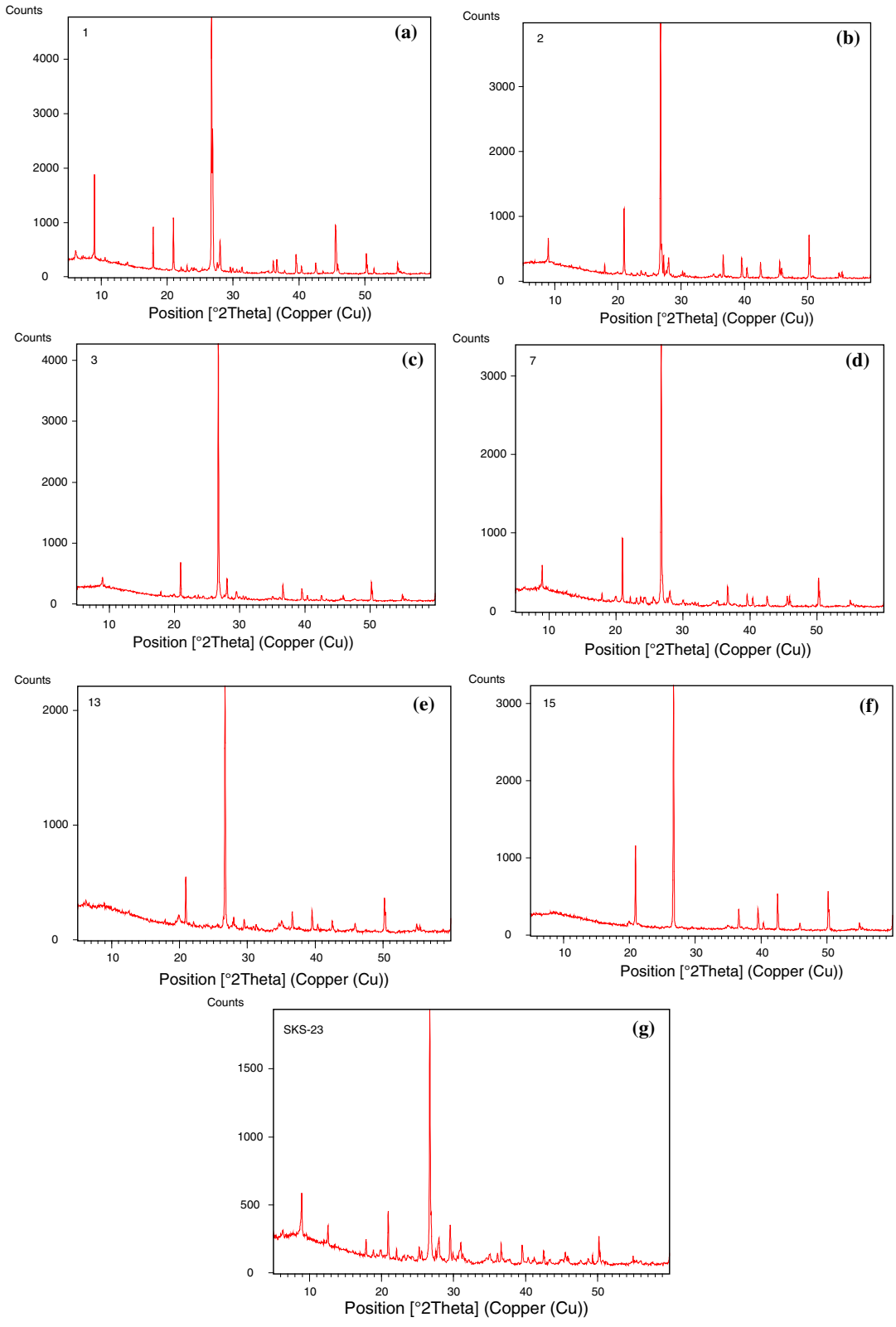

Table 9 Rotated component matrix of (a) pre-monsoon and (b) post-monsoon (varimax with Kaiser normalization)

Variables	Components		
	1	2	3
(a)			
pH	−.850	−.331	.019
EC	**.788**	**.511**	.074
TDS	**.815**	**.563**	−.044
F^-	.005	−.106	**.930**
Cl^-	.246	**.693**	.085
HCO_3^-	**.840**	.040	−.022
SO_4^{2-}	**.564**	**.601**	−.123
NO_3^-	**.770**	.043	.241
Ca^{2+}	−.043	**.684**	**.541**
Mg^{2+}	**.621**	**.619**	−.026
Na^+	**.812**	−.082	−.156
K^+	−.053	**.915**	−.262
Eigenvalue	4.736	3.239	1.339
Variance %	39.466	26.988	11.159
Cumulative variance	39.466	66.453	77.613

Variables	Components			
	1	2	3	4
(b)				
pH	−.216	−.699	.081	.318
EC	**.918**	.288	.003	−.092
TDS	**.865**	.462	.169	−.050
F^-	−.196	−.137	−.093	**.905**
Cl^-	**.546**	**.731**	.271	.066
HCO_3^-	**.918**	.140	.002	−.166
SO_4^{2-}	**.560**	**.618**	.337	.226
NO_3^-	.108	**.747**	−.034	−.238
Ca^{2+}	−.134	.127	**.813**	−.020
Mg^{2+}	.397	**.711**	.175	.224
Na^+	**.892**	.230	−.185	−.150
K^+	.124	−.018	**.885**	−.056
Eigenvalue	4.129	2.872	1.741	1.146
Variance %	34.409	23.931	14.504	9.554
Cumulative variance	34.409	58.341	72.845	82.399

28, the minerals such as quartz, muscovite, plagioclase, and orthoclase are found dominant. At two sites 13 and 14, a total of three minerals such as quartz, muscovite, and orthoclase are found dominant with some visible peak in the spectral analysis. At the sites 15 and 16, the minerals such as quartz and muscovite in the samples are detected. At the sites 24, 27, 29, minerals such as quartz, muscovite, plagioclase, orthoclase, and calcite are detected. The overall analysis of the XRD results suggested the presence of quartz and muscovite minerals nearly on all sites.

The main agent responsible for chemical weathering reactions of above-mentioned minerals is water and weak acids formed in water. The most stable mineral is quartz, which does not chemically weathers but only abrades into smaller grains of quartz (Bennett and Siegel 1987). The main reaction for chemical weathering includes hydrolysis, oxidation, dissolution, and acidolysis. (Chamley 1989). The decomposition of primary minerals leads to the formation of secondary mineral reactions, which are ultimately responsible for the release of ions into the water. The most common weak acid is carbonic acid, which is available almost everywhere, produced by the reaction of the rainwater with carbon dioxide (CO_2) gas in the atmosphere (Meybeck 1987). During the breakdown process, carbonic acid H^+ and bicarbonates are produced. The free H^+ is a small ion and can easily enter crystal structures, responsible for weathering of other minerals in the soil–rock–water interface (Nesbitt and Young 1984). For example, the weathering of orthoclase produces K^+, kaolinite, and quartz in the soil, and similar reaction are also found with muscovite, which may be the reason for some geogenic contribution of K^+ into the groundwater. Over here, hydrolysis is a dominant phenomenon for weathering of orthoclase mineral (Clayton 1988). Due to the replacement of potassium, ions in the presence of water are either by H^+ or OH^- ions. The K^+ ions are removed by dissolution into water and moved down to the aquifers; the process can be referred as leaching. The decomposition of carbonic acid, orthoclase, and muscovite generally follows the reactions:

$$H_2O + CO_2 \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$

$$4KAlSi_3O_8 + 4H^+ + 2H_2O \rightarrow 4K^+ + Al_4Si_4O_{10}(OH)_8 + 8SiO_2$$

Orthoclase Hydrogen ion Water Potassiumion Kaolinite (clay mineral) Quartz

$$2KAl_3Si_3O_{10}(OH)_2 + 2H^+ + 3H_2O \rightarrow 3Al_2Si_2O_5(OH)_4 + K^+$$

Muscovite Hydrogen ion Water Kaolinite (a clay) Potassium ions

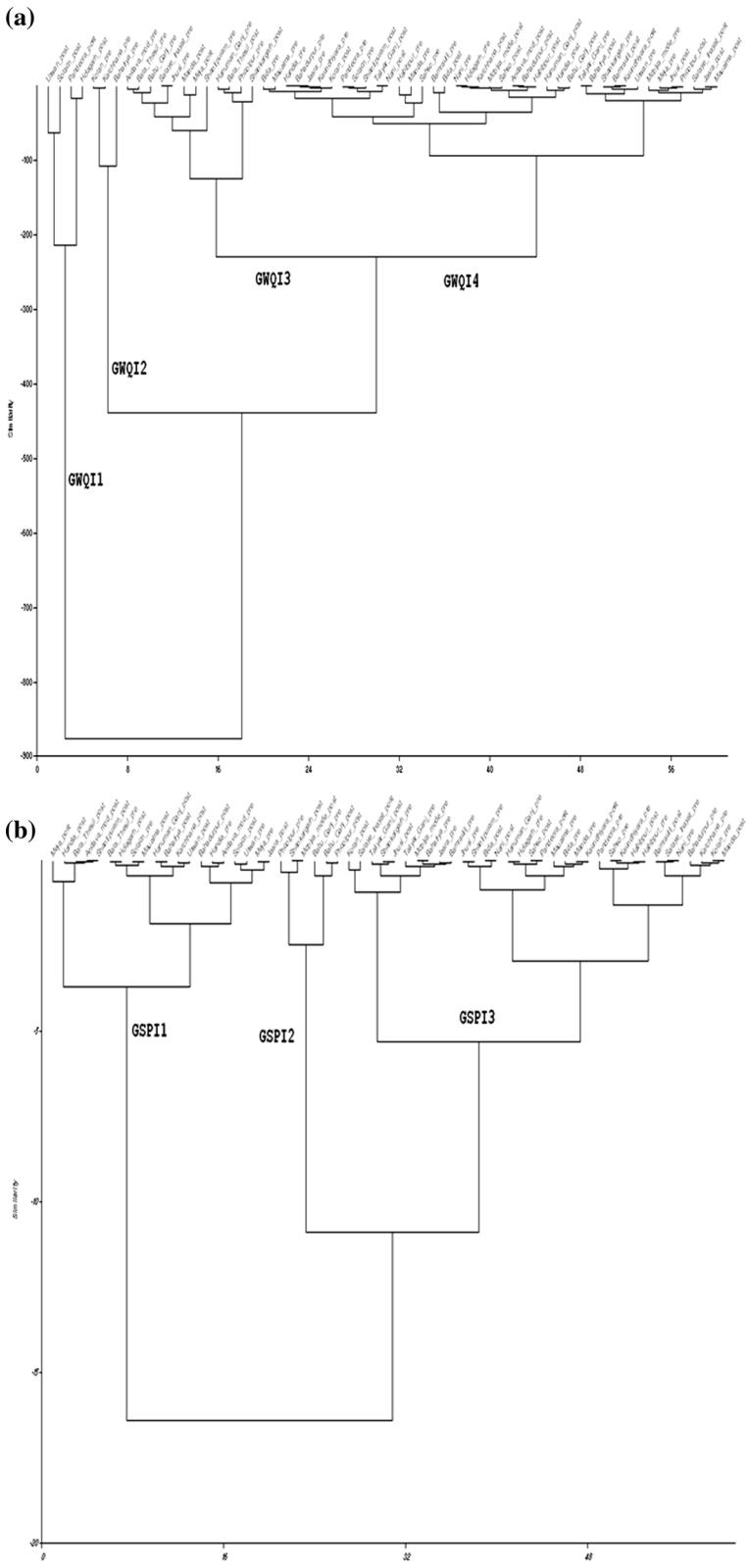

Fig. 6 Dendrogram representing **a** WQI of pre- and post-monsoon seasons, **b** SPI of pre- and post-monsoon seasons

The other minerals which are good source of Ca^{2+} ions are calcite, plagioclase, and dolomite. These minerals are also weathered due to carbonic acid reaction. The weathering of calcite by carbonic acid produces Ca^{2+} ion as well as bicarbonate. However, the weathering of calcium plagioclase produces a mixture of kaolinite, Ca^{2+}, and bicarbonate. On the other hand, the dolomite weathering produces both Ca^{2+} and Mg^{2+} in the water, as this mineral is composed of both Mg^{2+} and Ca^{2+}.

$$CaCO_3 + H_2CO_3 \rightarrow Ca^{2+} + 2(HCO_3)^-$$
Calcite Cabonic acid Calcium ion Bicarbonate ions

$$CaAl_2Si_2O_8 + 2H^+ + H_2O \rightarrow Al_2Si_2O_5(OH)_4^{2-} + Ca^{2+}$$
Plagioclase Hydrogen ion Water Kaolinite (a clay) Calcium ions

$$CaMgCO_3 + H_2O \rightarrow Ca^{2+} + Mg^{2+} + 2HCO_3^- + 2OH^-$$
Dolomite Water Calcium ions Magnessium ions Bicarbonate ions Hydroxyl ions

$$SiO_2 + 2H_2O \rightarrow H_4SiO_4$$
Quartz

Dissolution is a process whereby a mineral passes completely into solution, like salt dissolving in water (Stallard and Edmond 1987). A complete dissolution of halite minerals generally occurs, which releases the Na^+ and Cl^- in the water. Similarly, sodium plagioclase which is an important source of Kaolinite, silica, and Na^+ is also weathered due to hydrolysis and H^+.

$$NaCl \rightarrow Na^+ + Cl^-$$
Halite Sodium Chloride

$$2NaAlSi_3O_8 + 2H^+ + H_2O \rightarrow Al_2Si_2O_5(OH)_4 + 4SiO_2 + Na^+$$
Plagioclase Hydrogen ion Water Kaolinite (a clay) Silica Sodium ion

Multivariate statistical techniques

The multivariate statistical analysis involves PCA/FA and CA. The results of PCA are indicated in Table 9a, b. In the pre-monsoon season, three principal components are extracted. The first PC, accounting for ~39.46 % of total variance, is correlated with representing influences from point sources such as municipal and industrial effluents and soil leaching. This factor is characterized by very high loadings of EC, TDS, sodium, magnesium, bicarbonate, nitrate, and sulfate and thus, accounting for the temporary hardness of the water. The second factor (which accounts for 26.98 % of the total variance) is mainly associated with very high loading of EC, TDS, chloride, sulfate, and cations. The analysis of second component represents the influences from point sources such as from industries. The third PC shows high loading of fluoride and calcium. This factor (~11.15 % variance) probably represents geogenic contribution (XRD detected calcite, dolomite, and fluorite) (Table 9a, b). In the post-monsoon season, four components are extracted in which the first PC, accounting for ~34.40 % of the total variance, is correlated with representing influences from point sources such as municipal (possibly laundry industries) and industrial effluents. This factor is characterized by very high loadings of EC, TDS, chloride, bicarbonate and sulfate, and sodium and, thus, accounts for the salinity of the water. The second factor (which accounts for 23.931 % of the total variance) is mainly associated with the very high loading of chloride, sulfate, nitrate, and magnesium. The analysis of the second component represents influences from nonpoint sources such as agriculture runoff. The third PC (~14.50 % variance) is influenced by calcium, and potassium represents the laundry influence on groundwater. However, over here, the contribution from geogenic sources also cannot be denied (in XRD calcite and dolomite are detected). The fourth PC 9.55 % shows the loading of fluoride only, which represents pollution from pesticide industries and from geogenic sources, and in XRD, some fluorite sources are also detected (see section).

Cluster analysis has been used in this study to identify the similar groups between the sampling sites. The datasets are treated with hierarchical agglomerative clustering following the Ward's methods with Euclidean distance for measuring of similarity. A dendrogram rendered by CA is represented through Fig. 6a, b. It can be seen that cluster analysis successfully generates the distinct groups, and a clear-cut class structure for characterizing the datasets is formed. There are four clusters detected such as GWQI1, GWQI2, GWQI3, and GWQI4 from ground

WQI. This grouping gives evidence that some sites have similar sources of pollution from point or nonpoint sources. The dendrogram reveals that all the stations in GWQI1 are similar in their behavior during post-monsoon season; similarly, 3 stations in GWQI2 are purely identified as pre-monsoon samples with no matching properties with post-monsoon samples. In cluster GWQI3, many samples in pre- and post-monsoon seasons share the same properties, indicating that they are influenced by similar types of sources. From LULC, it appears that these areas are mostly suffered from urbanized waste and agricultural runoff. In cluster GWQI4, a mixed response obtained with WQI shows that these stations are more or less have similar sources of pollution. In this cluster, many samples are from industrial area, urban as well as agricultural land so a mixing of contamination cannot be denied. These stations are mainly polluted due to extensive agricultural and industrial activities followed by domestic and municipal discharges.

In the case of SPI, a total of three clusters are detected named as GSPI1, GSPI2, and GSPI3. The cluster GSPI1 has mixed performance as some of the stations are from pre-monsoon and other are from post-monsoon season. These stations are mainly polluted due to extensive agricultural and industrial activities. The stations in GSPI2 cluster are mostly from residential areas, and hence, that is the reason they clustered in the same group with respect to groundwater quality as they belong to the same source of pollution from domestic and municipal discharges. The cluster 3 (i.e., GSPI3 members) has the same irrigation quality because of similar sources of pollution from runoff. In this group, a high variability is obtained as the samples belong to industrial sites, cultivable land, built-up land, and agricultural land. Because of high alluvial soil dominance in the area, a mixing of contaminant cannot be denied. Over here, the samples suffered from all sort of pollution such as industrial waste, agricultural runoff, leaching from waste dumping sites, and urban waste.

Suggestions and recommendations

The highly populated country like India is facing scarcity of water. Good water quality and sufficient amount is critical to the health of public as well as flora and fauna dwellings in surroundings. In many ways, groundwater is also getting contaminated and polluted in many parts of India. Changes in water's natural quality are due to increase in levels of specific nutrients, pesticides, fertilizers by-products and in many other forms that can have serious negative effects on human lives those who directly or indirectly dependent on those contaminated and polluted water resources. There are great challenges in front of scientists and policy makers that how to minimize the negative impacts, effects, and influences and how to protect and conserve these vital resources judiciously. To understand how changes in the groundwater quality impact humans health, short- and long-term measures of groundwater quality parameters are used together to ensure that water quality will be properly evaluated for human consumption. There is a need of development of new rapid cost-effective in situ and ex situ testing methods, which can be used to determine the quality of water whether it is safe or unsafe. If the samples fall in unsafe category, then proper restoration measurements should be taken to prevent any losses to the human or environment. The appropriate treatment measures must be taken before releasing the effluent and sewage waste in shallow aquifers. More studies and tools are recommended for upstream basins to monitor the sources of pollutant in groundwater. The developed tools must allow water managers to predict water quality problems and identify sources of pollution, so that they can take appropriate measures to solve those problems.

There are many existing and new emerging issues related with quality and quantity of groundwater but need careful understanding. Therefore, it is recommended that issue-based practices should be taken into account for combating problems of groundwater. To counter the rising trend, some of the pollutants in the Trans-Gangetic basin, and to control and understand the contamination of groundwater, the hard rock areas should be surveyed by advanced and modern techniques such as aeromagnetic and geophysical survey to identify fractures present below 100 m depth. The conventional methods like creation of check dams could be used for the sites such as Shankargarh, Koraon, Manda, Meja blocks of the study area to control the urban runoff-induced pollutants in the groundwater. Apart from alternative arrangement and in quality problem are potable water pipeline project, and the information, education and communication

(IEC) methods are required to combat water quality issues. The deep boring should be required in the polluted areas. The cost-effective water testing kit should be provided to the users for knowing the status about the water used for drinking.

Conclusions

The utility of the WQI and SPI can be appreciated as the water resources play a vital role in the development of better society. Based on the analytical results obtained from the laboratory, XRD, water quality indices are applied to assess the groundwater quality of the area and the case study proved that the proposed WQI and SPI are very informative for long-term monitoring of groundwater. The proposed WQI and SPI clearly identify the type of water quality impairment through the group quality system that helps in initiating the immediate water pollution control actions. The satellite imagery can be used to estimate the land-use class and their change over a period of time for relatively large area, and these changes can be linked with the groundwater quality. From the results of the interpretation of Landsat TM and ETM images, the built-up area increased drastically from 1990 to 2011. Further, the land-use/land-cover analysis and field survey in the study area illustrate a high influence of domestic and agricultural waste during post-monsoon condition. The XRD results showed the presence of many dominant minerals such as quartz, muscovite, plagioclase, orthoclase, calcite, dolomite, halite, and vermiculite in the soil samples. The physical and chemical weathering of these minerals shows that the ions present in groundwater have some components from soil weathering. The study results reveal that the leaching and runoff, municipal and industrial wastewater, and waste disposal sites leaching are the main factors responsible for groundwater quality deterioration with some geogenic contribution from soil and rock weathering confirmed through XRD analysis.

Acknowledgments Authors are grateful to School of Environmental Sciences, Jawaharlal Nehru University and University Grant Commission, New Delhi, Grant No. (F. No. 42-74/2013(SR) for their technical and financial support, respectively. The views expressed here are those of the authors solely and do not constitute a statement of policy, decision, or position on behalf of NASA or the authors' affiliated institutions.

References

APHA (1998). *Standard methods for the examination of water and wastewater* (20th ed.). Washington, DC: American Public Health Association Inc.

Avvannavar, S. M., & Shrihari, S. (2008). Evaluation of water quality index for drinking purposes for river Netravathi, Mangalore, South India. *Environmental Monitoring and Assessment, 143*(1–3), 279–290.

Banerjee, R., & Srivastava, P. K. (2013). Reconstruction of contested landscape: Detecting land cover transformation hosting cultural heritage sites from Central India using remote sensing. *Land Use Policy, 34*, 193–203.

Banerjee, R., & Srivastava, P. K. (2014). Remote sensing based identification of painted rock shelter sites: Appraisal using advanced wide field sensor, neural network and field observations. In P. K. Srivastava, S. Mukherjee, M. Gupta, & T. Islam (Eds.), *Remote sensing applications in environmental research* (pp. 195–212). Society of Earth Scientists Series. Berlin: Springer. doi:10.1007/978-3-319-05906-8_11.

Bennett, P., & Siegel, D. (1987). Increased solubility of quartz in water due to complexing by organic compounds.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis theory and practice* (557 p). Cambridge, Massachusetts: MIT Press.

Chamley, H. (1989). Clay formation through weathering. In *Clay sedimentology* (pp. 21–50). Berlin: Springer.

Chang, K. (2002). *Introduction to geographic information systems*. New York.

Clayton, J. L. (1988). Some observations on the stoichiometry of feldspar hydrolysis in granitic soil. *Journal of Environmental Quality, 17*(1), 153–157.

Cole, H. (1970). Bragg's law and energy sensitive detectors. *Journal of Applied Crystallography, 3*(5), 405–406.

Foster, S., Hirata, R., Gomes, D., D'Elia, M., & Paris, M. (2002). *Groundwater quality protection: A guide for water utilities, municipal authorities, and environment agencies*. Washington, DC: World Bank.

Gleick, P. H. (2000). A look at twenty-first century water resources development. *Water International, 25*(1), 127–138.

Gupta, M., & Srivastava, P. K. (2010). Integrating GIS and remote sensing for identification of groundwater potential zones in the hilly terrain of Pavagarh, Gujarat, India. *Water International, 35*(2), 233–245.

Johnston, K., Ver Hoef, J. M., Krivoruchko, K. & Lucas, N. (2001). Using ArcGIS geostatistical analyst, vol. 380. Esri Redlands.

Kacher, J., Landon, C., Adams, B. L., & Fullwood, D. (2009). Bragg's Law diffraction simulations for electron backscatter diffraction analysis. *Ultramicroscopy, 109*(9), 1148–1156.

Kaurish, F. W., & Younos, T. (2007). Developing a standardized water quality index for evaluating surface water quality. *JAWRA, 43*(2), 533–545.

Krishna, A. K., Satyanarayanan, M., & Govil, P. K. (2009). Assessment of heavy metal pollution in water using multivariate statistical techniques in an industrial area: a case study from Patancheru, Medak District, Andhra Pradesh, India. *Journal of Hazardous Materials, 167*(1), 366–373.

Lu, D., Mausel, P., Brond, E., Iacute, Z., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*, 2365–2401.

Ma, J., Ding, Z., Wei, G., Zhao, H., & Huang, T. (2009). Sources of water pollution and evolution of water quality in the Wuwei basin of Shiyang river, Northwest China. *Journal of Environmental Management, 90*(2), 1168–1177.

Merchant, J. W. (1994). GIS-based groundwater pollution hazard assessment: A critical review of the DRASTIC model. *Photogrammetric Engineering and Remote Sensing, 60*(9), 1117–1128.

Meybeck, M. (1987). Global chemical weathering of surficial rocks estimated from river dissolved loads. *American Journal of Science, 287*(5), 401–428.

Mitas, L., & Mitasova, H. (1999). Spatial interpolation. *Geographical Information Systems: Principles, Techniques, Management and Applications, 1*, 481–492.

Moore, D. M., & Reynolds, R. C, Jr. (1989). *X-ray diffraction and the identification and analysis of clay minerals*. Oxford: Oxford University Press.

Nesbitt, H., & Young, G. (1984). Prediction of some weathering trends of plutonic and volcanic rocks based on thermodynamic and kinetic considerations. *Geochimica et Cosmochimica Acta, 48*(7), 1523–1534.

Ouyang, T., Zhu, Z., & Kuang, Y. (2006). Assessing impact of urbanization on river water quality in the Pearl River Delta Economic Zone, China. *Environmental Monitoring and Assessment, 120*(1–3), 313–325.

Patel, D. P., Dholakia, M. B., Naresh, N., & Srivastava, P. K. (2011). Water harvesting structure positioning by using geo-visualization concept and prioritization of mini-watersheds through morphometric analysis in the Lower Tapi Basin. *Journal of the Indian Society of Remote Sensin,*1–14.

Patel, D. P., Gajjar, C. A., & Srivastava, P. K. (2013). Prioritization of Malesari mini-watersheds through morphometric analysis: A remote sensing and GIS perspective. *Environmental Earth Sciences, 69*(8), 2643–2656.

Patel, D. P., & Srivastava, P. K. (2013). Flood hazards mitigation analysis using remote sensing and GIS: Correspondence with town planning scheme. *Water Resources Management, 27*(7), 2353–2368. doi:10.1007/s11269-013-0291-6.

Reynard, N. S., Prudhomme, C., & Crooks, S. M. (2001). The flood characteristics of large UK rivers: Potential effects of changing climate and land use. *Climatic Change, 48*(2), 343–359.

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

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Research, 38*(18), 3980–3992.

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

Singh, S. K., Singh, C. K., Kumar, K. S., Gupta, R., & Mukherjee, S. (2009). Spatialtemporal monitoring of groundwater using multivariate statistical techniques in Bareilly district of Uttar Pradesh, India. *Journal of Hydrology and Hydromechanics, 57*(1), 45–54.

Singh, S. K., Srivastava, P. K., Gupta, M., & Mukherjee, S. (2012). Modeling mineral phase change chemistry of groundwater in a rural-urban fringe. *Water Science and Technology, 66*(7), 1502.

Singh, S. K., Srivastava, P. K., Gupta, M., Thakur, J., & Mukherjee, S. (2014). Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine. *Environmental Earth Sciences, 71*(5), 2245–2255. doi:10.1007/s12665-013-2628-0.

Singh, S. K., Srivastava, P. K., & Pandey, A. C. (2013a). Fluoride contamination mapping of groundwater in northern India integrated with geochemical indicators and GIS. *Water Science and Technology: Water Supply*. doi:10.2166/ws.2013.160.

Singh, S. K., Srivastava, P. K., Pandey, A. C., & Gautam, S. K. (2013b). Integrated assessment of groundwater influenced by a confluence river system: Concurrence with remote sensing and geochemical modelling. *Water Resources Management, 27*(12), 4291–4313.

Snow, J. (1856). Cholera and the water supply in the south districts of London in 1854. T. Richards.

Srivastava, P. K., Gupta, M., & Mukherjee, S. (2012a). Mapping spatial distribution of pollutants in groundwater of a tropical area of India using remote sensing and GIS. *Applied Geomatics, 4*(1), 21–32. doi:10.1007/s12518-011-0072-y.

Srivastava, P. K., Han, D., Gupta, M., & Mukherjee, S. (2012b). Integrated framework for monitoring groundwater pollution using a geographical information system and multivariate analysis. *Hydrological Sciences Journal, 57*(7), 1453–1472. doi:10.1080/02626667.2012.716156.

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012c). Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research, 50*(9), 1250–1265. doi:10.1016/j.asr.2012.06.032.

Srivastava, P. K., Mukherjee, S., Gupta, M., & Singh, S. (2011). Characterizing monsoonal variation on water quality index of River Mahi in India using geographical information system. *Water Quality, Exposure and Health, 2*(3–4), 193–203.

Srivastava, P. K., Singh, S. K., Gupta, M., Thakur, J. K., & Mukherjee, S. (2013). Modeling impact of land use change trajectories on groundwater quality using remote sensing and GIS. *Environmental Engineering and Management Journal, 12*(12), 2343–2355.

Stallard, R., & Edmond, J. (1987). Geochemistry of the Amazon: 3. Weathering chemistry and limits to dissolved inputs. *Journal of Geophysical Research: Oceans (1978–2012), 92(C8)*, 8293–8302.

Sun, H., Bergstrom, J. C., & Dorfman, J. H. (1992). Estimating the benefits of groundwater contamination control. *Southern Journal of Agricultural Economics, 24*, 63.

Tiwari, T., & Mishra, M. (1985). A preliminary assignment of water quality index of major Indian rivers. *Indian Journal of Environmental Protection, 5*(4), 276–279.

Vasanthavigar, M., Srinivasamoorthy, K., Vijayaragavan, K., Ganthi, R. R., Chidambaram, S., Anandhan, P., et al. (2010). Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. *Environmental Monitoring and Assessment, 171*(1–4), 595–609.

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

Vidyalakshmi, R., Brindha, B., Roosvelt, P. S. B., Rajakumar, S., & Devi, M. P. (2013). Determination of land use stress on drinking water quality in Tiruchirappalli, India Using Derived Indices. *Water Quality, Exposure and Health, 5*(1), 11–29.

Waller, R. M. (2001). Ground water and the rural homeowner. US Geological Survey (USGS).

Wu, J., & Segerson, K. (1995). The impact of policies and land characteristics on potential groundwater pollution in Wisconsin. *American Journal of Agricultural Economics, 77*(4), 1033–1047.