An Innovative Approach for Groundwater Quality Assessment with the Integration of Various Water Quality Indexes with GIS and Multivariate Statistical Analysis—a Case of Ujjain City, India

Usman Mohseni[1] · Nilesh Patidar[1] · Azazkhan Ibrahimkhan Pathan[1] · P. G. Agnihotri[1] · Dhruvesh Patel[2]

Abstract

In India, a majority of the populace relies on groundwater for drinking. For this, the determination of groundwater quality (GWQ) is of great importance. The water quality index (WQI) is an effective technique that determines the suitability of water for drinking. In the present study, 54 groundwater samples consisting of eight physicochemical parameters were evaluated to assess water quality using four indexing methods: Numerow's pollution index (NPI), Weighted Arithmetic Water Quality Index (WA WQI), Groundwater Quality Index (GWQI), and the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI). A Geographic Information System (GIS) was employed to outline the spatial distribution maps of eight physicochemical parameters and WQI maps using the Inverse Distance Weighted (IDW) technique. Multivariate statistical analysis such as correlation analysis, principal component analysis (PCA), and cluster analysis (CA) were used for the evaluation of large and complicated groundwater quality data sets in the study. The results of the WQI indicate that 43% (NPI), 96% (WAWQI), 74% (GWQI), and 94% (CCME WQI) of groundwater samples had poor to unsuitable drinking water quality. Using Karl Pearson's correlation matrix, correlation analysis reveals a strong positive correlation of 0.9996 between EC and TDS. The application of PCA resulted in three major factors with a total variance of 72.5%, explaining the causes of water quality degradation. With the help of dendrogram plots, CA classifies eight groundwater parameters and 54 sampling locations into three major clusters with similar groundwater characteristics. According to the integrated approach of different water quality indexes with GIS, it is concluded that samples from wards 20, 44, and 47 are the most common and in the excellent-to-good category, and samples from wards 17, 34, and 43 are the most common and in the poor-to-very poor category. In view of the above, it is recommended to monitor the physicochemical parameters on a regular basis in order to safeguard groundwater resources and to prioritize management strategies in order to maintain the drinking quality of water.

Keywords Groundwater quality · GIS · Physicochemical parameters · Water quality index · Multivariate statistical analysis

✉ Usman Mohseni
  usmanmohsenialiakbar@gmail.com

  Nilesh Patidar
  nileshpatidar2671996@gmail.com

  Azazkhan Ibrahimkhan Pathan
  pathanazaz02@gmail.com

  P. G. Agnihotri
  pgasvnit12@gmail.com

  Dhruvesh Patel
  dhruvesh1301@gmail.com

[1] Department of Civil Engineering, Sardar Vallabhbhai National Institute of Technology, Icchanath, Surat, Gujarat 395007, India

[2] Present Address: Civil Engineering Department, Pandit Deendayal Energy University, Gandhinagar, India

Introduction

Water is the elixir of life and is among the most precious resources. Water makes up over 70% of the body weight of practically all living things. Without water, life on this planet is impossible. About 97.2% of the water on the planet is saline, with only 2.8% being fresh water, of which about 20% is groundwater. Because it has qualities that surface water does not, groundwater is held in high esteem [1]. Groundwater is the most significant supply of drinking water in the country, and it is essential for the country's development. It also helps people meet their numerous needs. The unrestricted use of groundwater, urbanization, industrialization, and agricultural activities all result in a massive amount of contaminated water [2]. Also inappropriate use of chemical

fertilizers, particularly nitrogen fertilizers, to boost crop yields, as well as the transport of urban and industrial effluent are believed to be factors that raise nitrate concentrations in groundwater [3]. The purity of replenished water, atmospheric showers, inland surface water, and underground geochemical activities all have an effect on groundwater quality [4]. However, the nature and quantity of contaminants is determined by the geology of the river's course and the quality of the water it supplies [5].

Groundwater quality (GWQ) assessment is of utmost importance due to the fact that consumption of polluted water is detrimental to people's health, corporate growth, and societal welfare [4]. The quality of groundwater is determined using the water quality index (WQI). Horton [6] was the first to develop the WQI on the basis of a weighted arithmetic approach. The WQI is a simple and effective approach for determining water quality [7]. It is a one-dimensional number that ranges from 0 to 100. It is a numeric rating system that shows the quality status of water (excellent, good, bad, etc.) at a specific location based on a variety of water quality factors. As a result, the WQI is being used as a crucial tool for comparing groundwater quality [8]. Many researchers have conducted studies over the last two decades using WQI to assess groundwater quality [9–11]. In the past, various indexing methods, such as Prati's Index of Pollution, Bhargava's Index, Oregon WQI, Dinius' Second Index, Weighted Arithmetic Water Quality Index (WA WQI), Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI), and National Sanitation Foundation (NSF), were adopted for the evaluation of groundwater quality [12]. However, in this study, based on the physico-chemical parameters available, four different indexing methods were adopted to determine the groundwater quality of the study area, viz., Numerow's pollution index (NPI), WA WQI, Groundwater Quality Index (GWQI), and CCME WQI.

In recent years, GIS technique has been used to monitor and evaluate groundwater quality frequently [13]. Groundwater assessment has traditionally relied on laboratory testing, but GIS have made it much easier to connect multiple databases [14]. The IDW interpolation method, along with the (GIS) technology, has been shown to be an effective approach for interpreting and analyzing spatial information of groundwater. It is an economical and time-saving method for converting large data sets into different spatial distribution maps and projections that indicate patterns, correlations, and sources of pollutants [15]. In this study, the spatial evaluation of all eight groundwater quality parameters was done with the help of the GIS technique. Also, spatial distribution maps based on different WQIs were prepared. Several studies have been conducted to assess groundwater quality using WQI within the context of a GIS framework [13, 15–17].

Multivariate statistical analysis (MSA) is an efficient tool for analyzing the properties of physico-chemical parameters in groundwater and determining the relationship between them [18]. These techniques could be used to find correlations between parameters and sample locations, highlight relevant variables and sources that influence groundwater quality, and offer effective tools for both water resource management and groundwater quality monitoring [19]. In this study, MSA such as correlation analysis, principal component analysis (PCA), and cluster analysis (CA) were adopted to elucidate the relationships between water quality variables and possible factors, as well as their effects on water quality. Sadat-Noori et al. [20] used correlation analysis to determine the correlation coefficient (r), which depicts the correlation between variables. PCA is a useful tool for elucidating massive data sets in complicated forms and reducing distortion in processes. It also encourages us to be aware of potential pollution sources or variables that affect water quality [21]. CA is used to examine the spatial groupings of the sampling locations. It is a widely used approach for grouping variables into clusters [22]. Various researchers have used this concept.

In the past, GWQ assessment was done with the help of different water quality index. Hamlat et al. [23] have adopted 10 such WQIs to evaluate the water quality of the Tafna basin. The results from the study reveal that CCME WQI and BC WQI were the best indices to describe the water quality of the basin. Also, numerous studies have been carried out using MSA to determine the GWQ. Acikel et al. [24] employed various MSA techniques such as FA, CA, and correlation analysis to determine the quality of water in the Azmak spring zone, Turkey. The study reveals that MSA is an important technique for describing the groundwater flow mechanism. However, with the recent advancement in technology such as GIS, much work has been done under the framework of GIS to evaluate the water quality. Ram et al. [25] used GIS and WQI for the assessment of groundwater quality in Mahoba district, Uttar Pradesh, India. The study concluded, with the help of WQI map, that the overall quality of water in the area is suitable for drinking. Recently, many combined approaches have been adopted for evaluating the GWQ. Roy et al. [26] suggested combined application of WQI and MSA for evaluation of GWQ West Tripura, India.

In the present study area, groundwater is the primary source of drinking water. However, minimal work has been performed on GWQ assessment of Ujjain City [e.g., 27, 28]; not even a single research has been done with an integration of WQI, GIS, and MSA to assess the GWQ of this region. Thus, there is a research gap here, and more discussion is needed to have a better understanding of the extent and causes of GWQ degradation. In light of these considerations, the city of Ujjain in Madhya Pradesh, India, has been chosen

for the purpose of an exhaustive study using the integrated approaches of different water quality indexes with GIS and multivariate statistical analysis. The purpose of this study was to accomplish the aforementioned objectives: (1) to use different indexing methods and GIS techniques to analyze the groundwater quality for its suitability for drinking as per BIS 10500:2012; (2) to categorize all 54 samples as excellent, good, or poor based on ratings from various indexing methods such as NPI, WA WQI, GWQI, and CCME WQI; (3) to develop thematic maps for individual physicochemical parameters and also WQI maps based on indexing methods using GIS; (4) to evaluate the disparities and clarification of a huge and complicated GWQ data set using MSA techniques such as correlation analysis, PCA, and CA; and (5) to find the wards that are most common in all indexing methods that are unfit for the consumption of drinking water.

Study Area

Ujjain is a historic city in Madhya Pradesh, India, that sits on the banks of the Shipra River. During ancient times, the city was known as Ujjayini. It is one of the most populous cities in Madhya Pradesh. It serves as the administrative hub of the Ujjain district. The Ujjain Municipal Corporation comprises a total of 54 wards. The latitude and longitude of the city are 23°10′58″N and 75°46′38″E, respectively. The city of Ujjain covers a total area of 93 km². According to Census India's provisional reports, the population of Ujjain City was 515,215 (5.15 lakhs) in 2011, and the forecasted population of Ujjain in 2022 is 5.70 lakhs, with a population growth rate of 10% over the decade. The city of Ujjain is segmented into pedeplains (shallow, deep, and moderate), residual hills, valley fills, flood plains, and other geomorphological features, a few of which have good groundwater potential, such as pedeplains (deep) and valley fills [28]. It has a pleasant monsoon climate.

However, winter begins in mid-November and is comforting and cool, with a daytime temperature of 20 °C, while the night-time temperature can drop dramatically. The annual rainfall of Ujjain city is 892.9 mm. On average, the elevation is 491 m. The temperature ranges between 8 and 40 °C.

Materials and Methodology

Sampling and Analysis

In this study, eight physicochemical properties such as pH, turbidity, electrical conductivity (EC), total dissolved solids (TDS), alkalinity, hardness, chloride (Cl^-), and fluoride (F^-) were selected from 54 groundwater samples that were collected from dug wells, bore wells, and hand pumps, which were assessed and compared with BIS 10500:2012 for drinking purposes. The samples were collected at distances with reference to other locations to provide a broad investigation of the study area's water quality. The samples were gathered in clean and dry plastic bottles from different sources after draining the water for a few minutes. All 54 samples were examined for the eight parameters using the procedures outlined by the American Public Health Association (APHA 2017). Table 1 summarizes the methodology, which comprises the analytical techniques, software, and instruments used to complete the work. The overall evaluation of GWQ was done using four different indexing methods, such as NPI, WA WQI, GWQI, and CCME WQI. Moreover, ArcGIS 10.8 was used to prepare the digitized base map of the study area as shown in Fig. 1. Using the spatial analyst tool from the tool box, Inverse Distance Weighted (IDW) technique was selected for preparing interpolated maps. The MSA techniques such as correlation analysis, PCA, and CA were executed using Minitab statistical software.

Table 1 Analytical techniques adopted for analysis

| Parameters	Methods
pH	Electrometric using digital pH meter
EC (μmho/cm)	Electrometric using digital conductivity meter
Hardness (mg/l)	Titration by EDTA
TDS (mg/l)	Electrometric using TDS Meter Hanna instruments
Alkalinity (mg/l)	Titration by H_2SO_4
Chloride (mg/l)	Titration by $AgNO_3$
Fluoride (mg/l)	Spectrophotometric using UV–Vis spectrophotometer
Turbidity (NTU)	Turbidimetric using digital nephaloturbidometer
Statistical analysis	Minitab statistical software
Karl Pearson's correlation analysis	MS Excel 2016
WQI method calculation	Drinking standards of BIS 10500:2012
Spatial distribution maps	Inverse Distance Weighted (IDW) interpolation technique
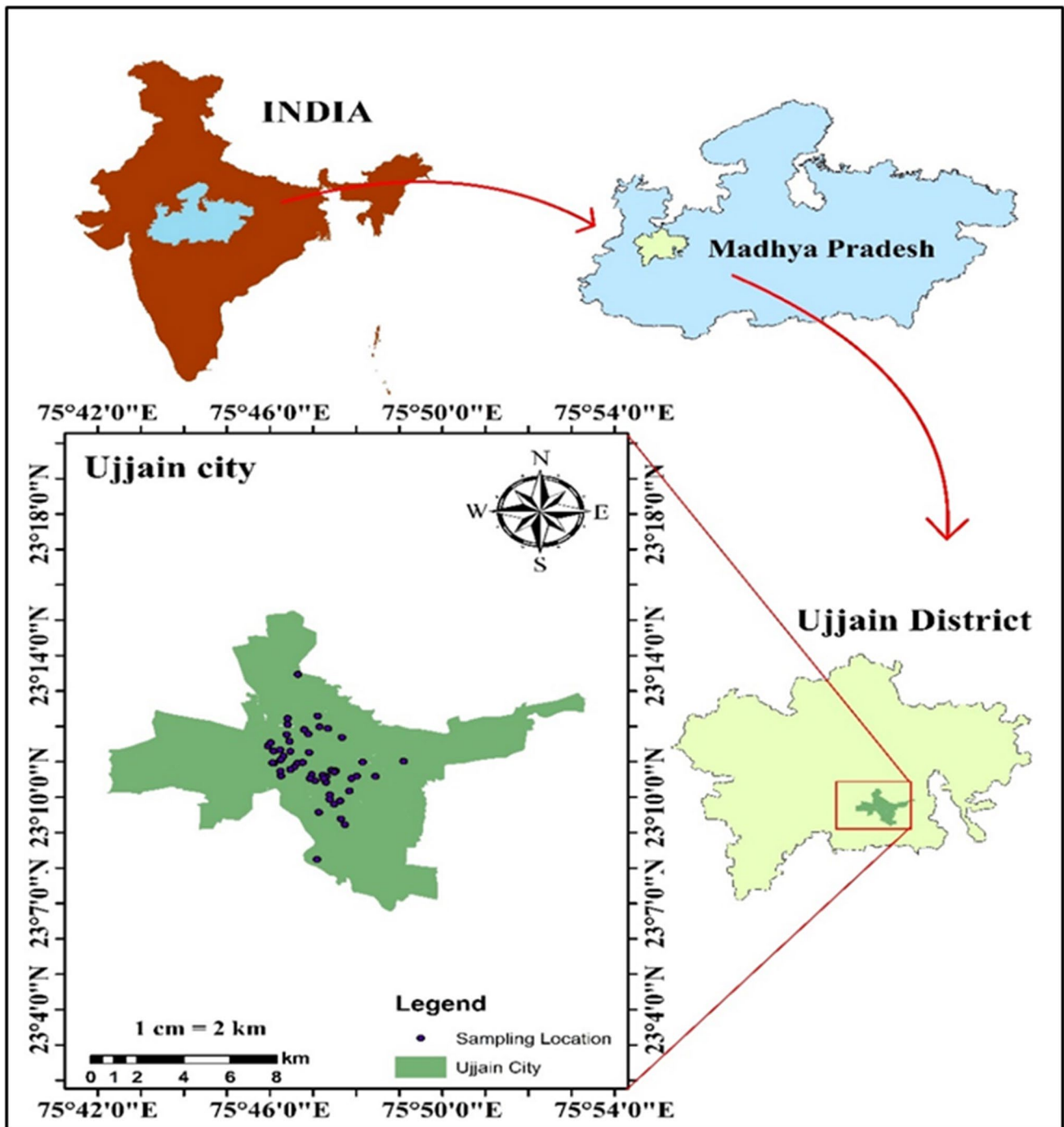
Fig. 1 Index map of study area

Methodology

The methodological flowchart illustrating the details of various steps involved in evaluating the GWQ is shown in Fig. 2. The methodological details for GIS analysis and mapping of groundwater parameters, GWQ modeling, and GWQ analysis have been presented in the flowchart.

GIS Analysis and Mapping of Groundwater Parameters

GIS aids in the interpolation of various experimental data in order to create thematic and spatial maps. It allows for the statistical development of a relationship in order to summarize the GWQ of the area in a simplified visual form.

Fig. 2 Methodological flow-chart
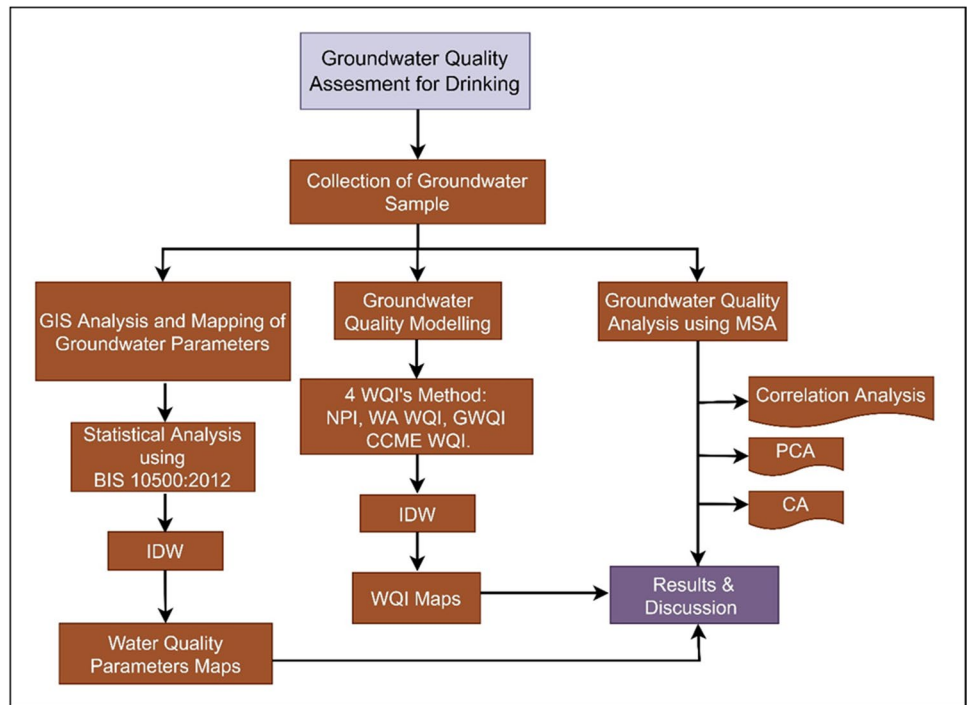
The most commonly used and acknowledged methods for generating spatial distribution maps are Inverse Distance Weighted (IDW), kriging, and cokriging. However, in this study, the (IDW) interpolation method in ArcGIS-10.8 software has been utilized to create spatial distribution maps of all parameters. The IDW interpolation method calculates undetermined values in relation to a distance, with the closest point receiving more weightage and decreasing as the distance increases. Furthermore, numerous researchers [25, 29] had employed this technique to create spatial distribution maps for different parameters. Table 2 shows statistical analysis of analyzed physicochemical properties. Table 3 shows NPI calculation for ward 1. Table 4 shows classification of 54 groundwater samples for drinking based on WQI values obtained using different indexing methods.

Groundwater Quality Modeling

In this study, the GWQ of all 54 samples were modeled using four indexing methods: NPI, WA WQI, GWQI, and CCME WQI [30–33]. All these methods are discussed in detail in the following sections.

NPI

The NPI was created to measure the impact of each individual particle and is used to calculate the total harm caused by pollution. It is an overall pollutant indicator that takes into account the combined influence of several pollutants for a given application. For the establishment of an index for any given function, such as drinking and

Table 2 Statistical analysis of analyzed physicochemical properties

Parameter	Minimum	Maximum	Mean	Standard deviation
pH	6.8	8.2	7.335	0.349
EC (µmho/cm)	195	8673	1288.407	1332.390
Hardness (mg/l)	148	990	395.22	203.352
TDS (mg/l)	125	5604	831.907	854.581
Alkalinity (mg/l)	200	532	307.019	73.229
Chloride (mg/l)	38	1320	270.519	291.621
Fluoride (mg/l)	0.4	1.7	0.77	0.256
Turbidity (NTU)	0.1	5.9	0.549	0.775

Table 3 Calculation of NPI for Ward No. 1

Parameters	Standard value (S_n)	Observed value (C_n)	Pollution index ($PI_n = C_n/S_n$)	NPI = $\sum PI_n$	Status
pH	8.5	7.12	0.84		
Turbidity	5	0.2	0.04		
EC	300	1081	3.6		
TDS	500	690	1.38	12.82	Good
Alkalinity	200	532	2.66		
Chloride	250	310	1.24		
Hardness	300	660	2.2		
Fluoride	1	0.86	0.86		

Table 4 Classification of 54 groundwater samples for drinking based on WQI values obtained using different indexing methods [30–33]

Method	Classes (rating)	Ward no	No. of samples	% of samples
NPI	Excellent (< 10)	3, 4, 5, 8, 9, 11, 12, 13, 14, 15, 19, 20, 21, 22, 23, 28, 29, 30, 31, 32, 33, 35, 36, 40, 44, 45, 47, 49, 50, 51, 54	31	57
	Good (10–20)	1, 2, 6, 7, 10, 16, 18, 24, 25, 26, 27, 38, 39, 41, 42, 46, 48	19	35
	Poor (20–30)	37, 43	2	4
	Very Poor (30–40)	34	1	2
	Unsuitable (> 40)	17	1	2
WA WQI	Excellent (0–25)	44, 51	2	4
	Good (26–50)	1, 2, 3, 4, 5, 6, 8, 10, 12, 13, 15, 16, 18, 20, 23, 24, 26, 28, 29, 32, 35, 37, 38, 46, 47, 48, 49, 50, 53	29	53
	Poor (51–75)	7, 9, 11, 14, 17, 19, 22, 21, 25, 27, 30, 31, 33, 34, 36, 40, 41, 42, 43, 52, 54	21	39
	Very Poor (76–100)	39, 45	2	4
	Unsuitable (> 100)	-	-	-
GWQI	Excellent (< 50)	-	-	-
	Good (50–100)	5, 13, 20, 21, 22, 23, 29, 31, 33, 44, 47, 50, 51, 54	14	26
	Poor (100–200)	1, 2, 3, 4, 8, 9, 10, 11, 12, 14, 15, 19, 24, 26, 28, 30, 32, 35, 36, 38, 40, 45, 48, 49, 53	25	46
	Very Poor (200–300)	6, 7, 25, 27, 39, 41, 42, 46, 52	9	17
	Unsuitable (> 300)	16, 17, 18, 34, 37, 43	6	11
CCME WQI	Excellent (95–100)	-	-	-
	Good (80–94)	20, 44, 47	3	6
	Fair (65–79)	5, 8, 11, 13, 14, 21, 22, 23, 28, 29, 30, 31, 32, 33, 36, 45, 50, 51, 54	19	35
	Marginal (45–64)	2, 3, 4, 5, 9, 10, 12, 15, 19, 24, 25, 26, 35, 38, 40, 41, 42, 46, 48, 49, 53	21	39
	Poor (0–44)	1, 7, 16, 17, 18, 27, 34, 37, 39, 43, 52	11	20

irrigation, the index value includes relevant particles [30]. Based on NPI the water quality is classified into five classes [Excellent (< 10), Good (10–20), Poor (20–30), Very Poor (30–40), Unsuitable (> 40)]. Gauns et al. [34] have adopted this approach to determine the quality of drinking water**.** NPI is calculated with the help of the following equation:

$$PI_n = C_n/S_n \tag{1}$$

$$NPI = \sum PI_n \tag{2}$$

where PI_n is the nth parameter's pollution index, C_n is the nth parameter's observed value, S_n is the nth parameter's permissible value, and NPI is the Nemerow's pollution index.

WA WQI

WA WQI is used to determine the quality of water for drinking by using the selected physicochemical parameters. In this study, eight parameters were taken into consideration to compute the WA WQI. The quantitative assessment of GWQ using WA WQI was carried out using Brown's method [31]. Based on WA WQI the water quality is classified into five classes [Excellent (0–25), Good (26–50), Poor (51–75), Very Poor (76–100), Unsuitable (> 100)]. This approach has been widely adopted in the past by many researchers [35, 36]. The following steps were used to calculate the WA WQI:

1. Unit weight (W_n): To calculate W_n, a quantity which is inversely proportional to S_n of the suitable parameter was utilized. The W_n of each parameter is given in Table 5.

$$W_n = K/S_n \tag{3}$$

where W_n is the nth parameter's unit weight, S_n is the nth parameter's standard value, and K is the proportionality constant.

2. Subindex (q_n): Subindex is calculated by the following equation:

$$q_n = \frac{(V_n - V_o)}{(S_n - V_o)} \times 100 \tag{4}$$

where V_n is the mean value of nth parameter, S_n is the standard value of the nth parameter, and V_o is the actual value of parameter.

3. By linearly combining the q_n and W_n, the overall water quality index is calculated.

Table 5 Unit weights assigned to each parameter

Parameter	Unit weights (W_i)			
	NPI	WA WQI	GWQI	CCME WQI
pH	No weights are assigned	0.0882	0.0455	Each variable is given the same amount of weight
Turbidity		0.15	0.2273	
EC		0.0025	0.1818	
TDS		0.0015	0.1364	
Alkalinity		0.00375	0.0909	
Chloride		0.003	0.1364	
Hardness		0.0025	0.0909	
Fluoride		0.75	0.0909	

$$WQI = \frac{\Sigma q_n}{\Sigma W_n} \tag{5}$$

GWQI

This method of determining the WQI is easy and trustworthy. GWQI is among the most extensively employed index to assess GWQ for drinking purposes throughout the world [33]. GWQI classify the water quality into five classes [Excellent (< 50), Good (50–100), Poor (100–200), Very Poor (200–300), Unsuitable (> 300)]. It is one of the most widely adopted approach; Agarwal et al. [37] have employed GWQI for evaluating quality of water. The following five steps were carried out to find GWQI:

1. Assigning weightage (w_i): To calculate the GWQI, eight parameters were selected: pH, EC, TH, TDS, alkalinity, Cl⁻, F⁻, and turbidity. As stated in Table 5, parameters were assigned a weightage (w_i) on a scale of 1 to 5 depending on their relative significance to GWQ.
2. Calculation of relative weights (W_i): The following expression was used to calculate W_i for each parameter. The W_i of each parameter is given in Tables 6, 7 and

8 as Tables illustrating/showing calculation of WA WQI, GWQI, and CCME WQI for ward 1 respectively.

$$W_i = \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{6}$$

3. Calculating quality rating scale (Q_i): It represents the percentage of the parameter's actual value to its standard value.

$$Q_i = \frac{C_i}{S_i} \times 100 \tag{7}$$

where C_i is the ith parameter's actual value and S_i is the ith parameter's standard value.

4. Subindex (SI_i): It is calculated for an individual parameter and is given by the following equation:

$$SI_i = W_i \times Q_i \tag{8}$$

5. Calculation of GWQI: By combining all of the subindices for each parameter, an overall groundwater quality index was calculated.

$$GWQI = \sum SI_i \tag{9}$$

Table 6 Calculation of WA WQI for Ward No. 1

Parameters	Standard value (S_n)	Observed value (V_o)	Quality rating (q_n)	Unit weight (w_n)	$\sum q_n . w_n$	WA WQI	Status
pH	6.5–8.5	7.12	8	0.0882	0.080		
Turbidity	5	0.2	4	0.15	0.090		
EC	300	1081	360	0.00225	0.00639		
TDS	500	690	138	0.0015	0.00321	**44.85**	**Good**
Alkalinity	200	532	266	0.00375	0.0091		
Chloride	250	310	220	0.0025	0.0059		
Hardness	300	660	124	0.003	0.0063		
Fluoride	1	0.86	86	0.75	1.451		

Table 7 Calculation of GWQI for Ward No. 1

Parameters	Standard value (S_i)	Observed value (C_i)	Quality rating ($Q_i = C_i/S_i \times 100$)	Relative weight (W_i)	Subindex $SI_i = W_i \times Q_i$	GWQI	Status
pH	6.5–8.5	7.12	83.7647	0.0455	3.81		
Turbidity	5	0.2	4	0.0909	0.36		
EC	300	1081	360.3333	0.2273	81.89		
TDS	500	690	138	0.1364	18.82	193.81	Poor
Alkalinity	200	532	266	0.0909	24.18		
Chloride	250	310	124	0.1364	16.91		
Hardness	300	660	220	0.1818	40.00		
Fluoride	1	0.86	86	0.0909	7.82		

Table 8 Calculation of CCME WQI for Ward No. 1

Parameters	Observed value (V_o)	Standard value (S_n)	Scope value (F_1)	Frequency value (F_2)	Amplitude value (F_3)	CCME WQI	Status
pH	6.5–8.5	7.12					
Turbidity	5	0.2					
EC	300	1081					
TDS	500	690	62.5	62.5	43.19529	**43.19**	**Marginal**
Alkalinity	200	532					
Chloride	250	310					
Hardness	300	660					
Fluoride	1	0.86					

The CCME WQI

The CCME WQI is centered upon several water uses, such as drinking, leisure, agriculture, animals, and sea species. The CCME WQI is a standardized approach for assessing water quality that was developed by Canadian authorities. A committee constituted inside the CCME designed the WQI. The index value is graded on a scale of 0 to 100, with 0 being the lowest and 100 representing the finest water quality. CCME WQI divides the quality of water into five classes [Excellent (95–100), Good (80–94), Fair (65–79), Marginal (45–64), Poor (0–44)]. Wagh et al. [38] have adopted CCME WQI to determine the quality status of water. The CCME WQI mathematical formula is shown below [32].

$$CCMEWQI = 100 - \left[\frac{\sqrt{F1^2 + F2^2 + F3^2}}{1.732} \right] \quad (10)$$

where F_1 is the number of variables with unmet aims (failed variables), F_2 is the fraction of individual tests with failed tests, and F_3 is the percentage of failed test values with unmet objectives.

1. Calculating scope value (F_1):where

$$F_1 = \frac{\text{No. of failed variables}}{\text{No. of variables}} \times 100 \quad (11)$$

2. Calculating frequency value (F_2):where

$$F_2 = \frac{\text{No. of failed test}}{\text{No. of test}} \times 100 \quad (12)$$

3. Calculating amplitude value (F_3):where

$$F_3 = \frac{\text{nse}}{0.01 \, \text{nse} + 0.01} \quad (13)$$

$$\text{nse} = \frac{\sum_{i=1}^{n} \text{excursion}}{\text{Total no. of test}} \quad (14)$$

$$\text{Excursion} = \frac{\text{Failed test value}}{\text{Objective}} \quad (15)$$

GWQ Analysis Using MSA

MSA

Simeonov et al. [39] suggest that MSA is the most effective method for avoiding misinterpretation of large amounts of

complex pollution monitoring data. Correlation analysis, PCA, and CA were performed to evaluate spatial variability and discover pollution sources. In this study, these methods were used on 54 samples for eight variables. To eliminate misclassification due to dimensionality differences, all 8 variables were standardized by computing their standard scores (z-scores) as follows:

$$Z_i = \frac{X_i - X}{SD} \qquad (16)$$

where Z_i is the standard score of ith variables, X_i is the actual value of ith variable, X is the mean value of variable, and SD is the standard deviation of variable [40].

Correlation Analysis

Correlation analysis is a popular and helpful statistical technique for evaluating the strength of a relationship between two variables. In this study, the correlation coefficients (r) of the variables were employed to determine the correlation between them. The correlation coefficients of all the 8 variables were determined using Karl Pearson's correlation matrix and are presented in Table 9. Its value (r) can be positive or negative, ranging from -1 to 1 [41]. A few correlation coefficients are positive, expressing similarity in the same direction, and a few are negative, expressing dissimilarity, as seen in Table 10, where $+1$ denotes a perfect

positive relationship and 0 shows no relationship between the correlated variables.

PCA

PCA describes the variation of a vast data set of variables by compressing them into a reduced data set of independent variables [42]. PCA decreases the dimensionality of data by creating new hidden variables that are perpendicular and uncorrelated to each other by combining original data in a linear manner [43]. From the covariance matrix of given variables, it derives the eigenvalues and eigenvectors [44]. The PCs' eigenvalues represent their associated variance, whereas the loadings represent the given variable contributions to the PCs [45]. The correlation matrix determines how well each constituent's variance may be described by their relation with one another [46].

CA

CA is a renowned classification tool that aims to determine either the distance or similarity between the variables to be grouped. It is commonly shown with the help of a dendrogram, which is a two-dimensional graph that displays a clear pictorial description of the process [47]. The distance between the parameters of samples is studied using hierarchical cluster analysis (HCA). The points that are the most similar are

**Table 9** Karl Pearson's correlation matrix between the parameters

Parameter	pH	Turbidity	EC	TDS	Alkalinity	Chloride	Hardness	Fluoride
pH	1							
Turbidity	−0.1754	1						
EC	−0.2236	0.4181	1					
TDS	−0.2244	0.4059	**0.9996**	1				
Alkalinity	−0.1538	−0.0290	0.1093	0.1081	1			
Chloride	−0.2784	0.4605	**0.7067**	**0.7028**	0.2083	1		
Hardness	−0.0853	0.1421	0.4856	0.4826	**0.6015**	0.4470	1	
Fluoride	0.0599	−0.0457	0.0457	0.0455	0.0562	−0.1095	−0.1372	1

The significance of "bold" emphasis that values are showing strong correlation between the corressponding parameters. For example 0.9996 is showing strong correlation between EC and TDS

Table 10 Summary of parameters based on the BIS 10500:2012 permissible limit [34, 53]

Parameter	Permissible limit for drinking BIS 10500:2012	No. of samples under permissible limit (%)	No. of samples over permissible limit (%)	Undesired effect
pH	6.5–8.5	100	0	Taste
EC	300 μmho/cm	1.85	98.15	Gastrointestinal irritation
Hardness	300 mg/l	16.67	83.33	Corrosion tendency and scaling
TDS	500 mg/l	37.04	62.96	Taste, discoloration of water
Alkalinity	200 mg/l	1.85	98.15	Taste
Chloride	250 mg/l	74.08	25.92	Taste, corrosion in pipes
Fluoride	1 mg/l	90.75	9.25	Skeletal fluorosis
Turbidity	5 NTU	98.15	1.85	Cloudiness

Fig. 3 Spatial distribution map of water quality parameters. (**a**) pH. (**b**) EC. (**c**) Hardness. (**d**) TDS. (**e**) Alkalinity. (**f**) Chloride. (**g**) Fluoride. (**h**) Turbidity
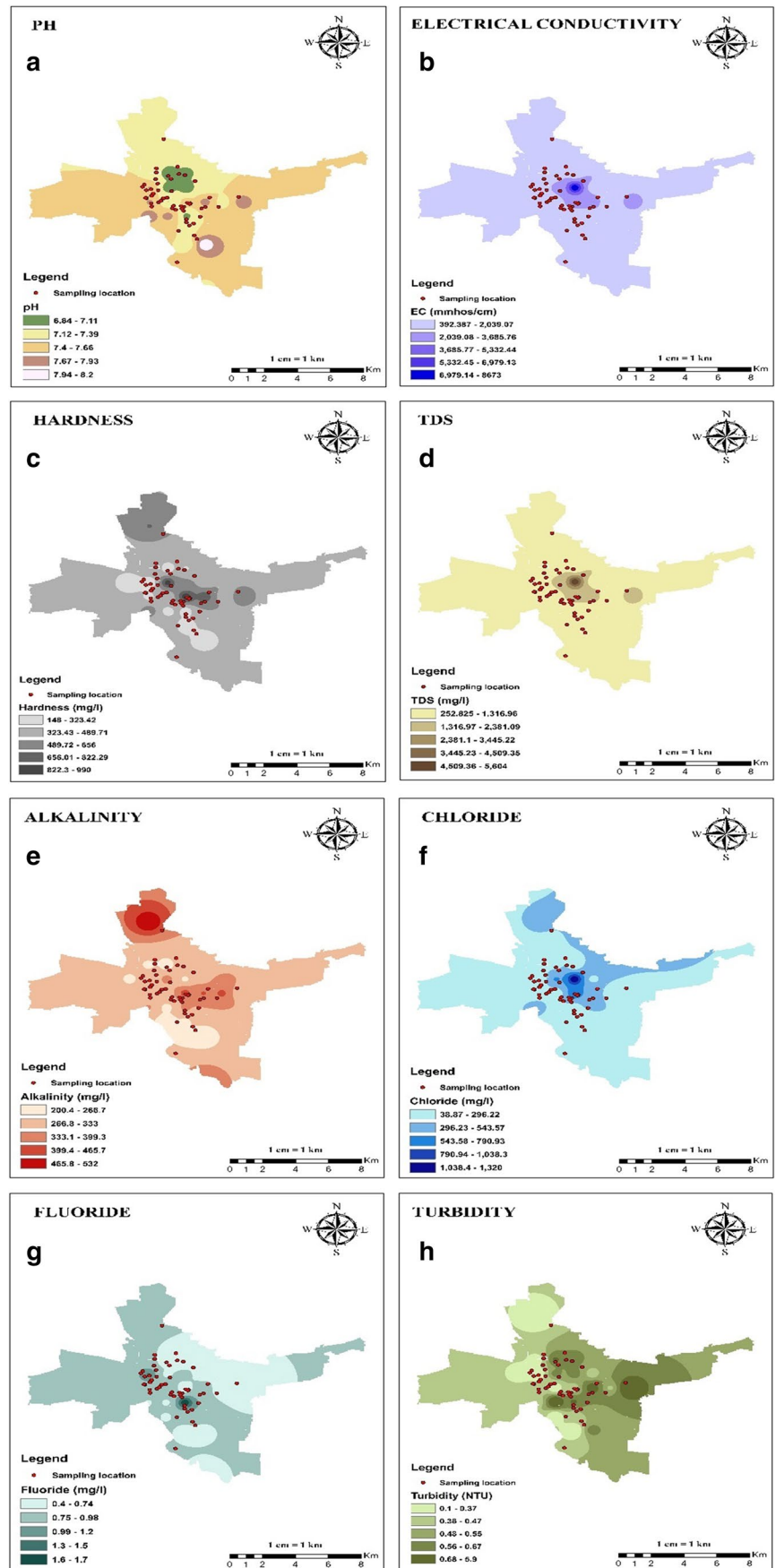
joined together to form a cluster, and this is continued till all the points fit into the same cluster [48]. In this study, Ward's method, using squared Euclidean distances, is employed.

Results

Groundwater Quality Parameters

As shown in Fig. 3a–h, groundwater quality mapping was implemented using IDW in ArcGIS 10.8 for the eight physicochemical parameters. The several parameters taken into account in the study are explained in the following lines:

pH The amount of hydrogen ions in water is indicated by the pH, which is neutral. It represents the acidic or basic nature of water. For drinking purposes, the pH must be between 6.5 and 8.5 (BIS, 2012). The pH of pure water is neutral, indicating that hydrogen ions are present. However, the pH in the collected groundwater sample varies from 6.8 (minimum) to 8.2 (maximum), indicating that it is well within the permissible range (6.5 to 8.5) as shown in Fig. 3a.

EC The amount of dissolved material in a medium determines the EC; the more dissolved material in a medium, the higher the EC. High EC values can be caused by a variety of geochemical processes, such as reverse and direct ion exchange, significant evaporation, silicate degradation, and rock–water interaction [49]. However, for drinking purposes,

EC should not exceed 300 μmho/cm. The EC in this study ranges from 8673 to 195 μmho/cm as shown in Fig. 3b.

Hardness It refers to the quantity of calcium and magnesium ion dissolved in the water. The total hardness of water is a vital parameter in the household sector. It occurs as a result of the existence of calcium and magnesium in the body. The maximum hardness that can be tolerated is 300 mg/l. Range of hardness is from 148 to 990 mg/l as shown in Fig. 3c.

TDS The majority of TDS is made up of inorganic salts and a little quantity of organic molecules dissolved in water. High TDS levels in water can alter the taste and hardness of the water. On the other hand, water with exceptionally low TDS has a bland taste [8]. The acceptable TDS for drinking water, according to the BIS, is less than 500 mg/l. It lies in the range of 125 to 5604 mg/l as shown in Fig. 3d.

Alkalinity It is caused by the presence of carbonate, bicarbonate, and hydroxide ions in the water. Water has a better capacity to neutralize acids when it has a higher alkalinity, and vice versa. Alkalinity and bicarbonate are associated in neutral water [50]. It should not exceed 200 mg/l. The taste of water becomes harsher further than this point. For the study area, alkalinity ranges from 200 to 532 mg/l which can be seen in Fig. 3e.

Chloride The higher the chloride concentration in water, the more dangerous it is to human health. The associated

Fig. 4 Plot showing NPI for each ward
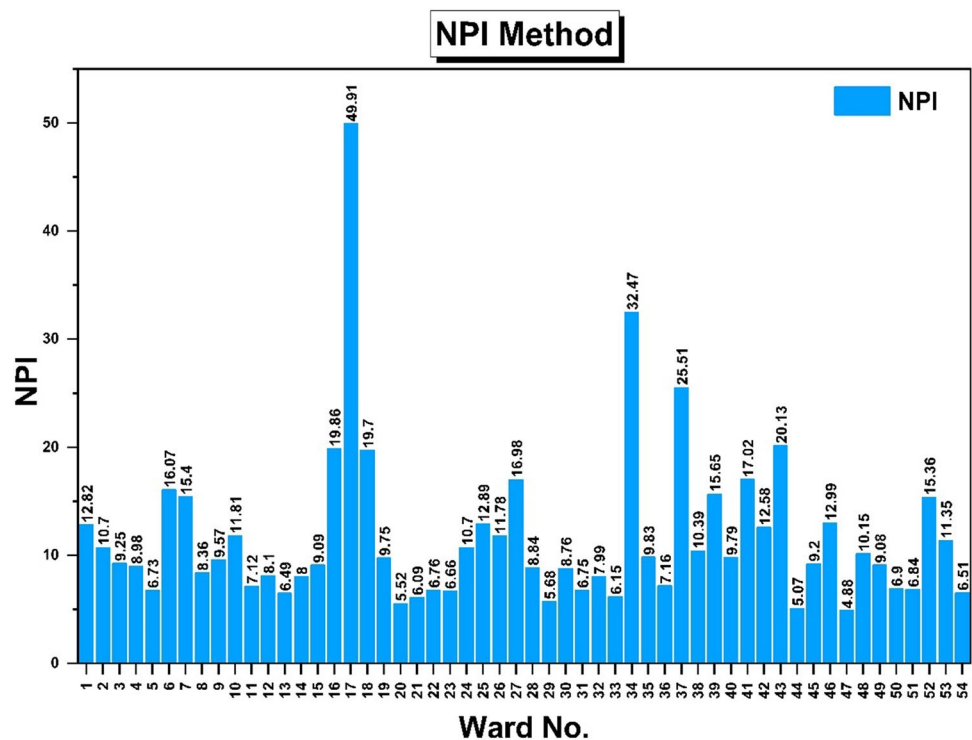
cation influences the taste threshold of the chloride ion in water. Geogenic or anthropogenic processes seems to be to responsible for the increasing Cl concentrations in groundwater [51]. Chloride levels should not exceed 250 mg/l. The amount of chloride in this study varies from 38 to 1320 mg/l as shown in Fig. 3f.

Fluoride It is primarily found in water as a result of geological processes. Fluoride in high concentrations (> 1.0 mg/l) can cause skeletal fluorosis. The concentration of fluoride in the study ranges from 0.4 to 1.7 mg/l, as shown in Fig. 3g.

Turbidity It describes the foggy appearance of water caused by particles, often known as "suspended matter." Drinkable water that is turbid loses its aesthetic appeal. Turbidity can have a variety of appearances and colors [52]. For drinking water, the maximum turbidity allowed is 5 NTU. The turbidity in this study area ranges from 0.1 to 5.9 NTU as shown in Fig. 3h.

Integrated Approach of Different Water Quality Indices with GIS

NPI

NPI was determined for the 54 samples for the provided water quality parameters. It also takes into account the effects of a number of variables that influence water quality. A pollution index for individual parameters was determined for each ward. The summation of the pollution indexes of all parameters in each ward gives the NPI value (Fig. 4). Table 3 shows the NPI calculation for Ward 1. The obtained NPI range is 4.88 to 49.91, i.e., from excellent to unsuitable. Figure 4 depicts the variation of NPI throughout the wards. The categorization of water quality was divided into five categories depending on the water quality rating given in Table 4. As per the NPI ratings, out of the 54 samples, 31 fall into the excellent category, 19 fall into the good category, 2 fall into the poor category, 1 in the very poor category, and 1

Fig. 5 NPI map for Ujjain City
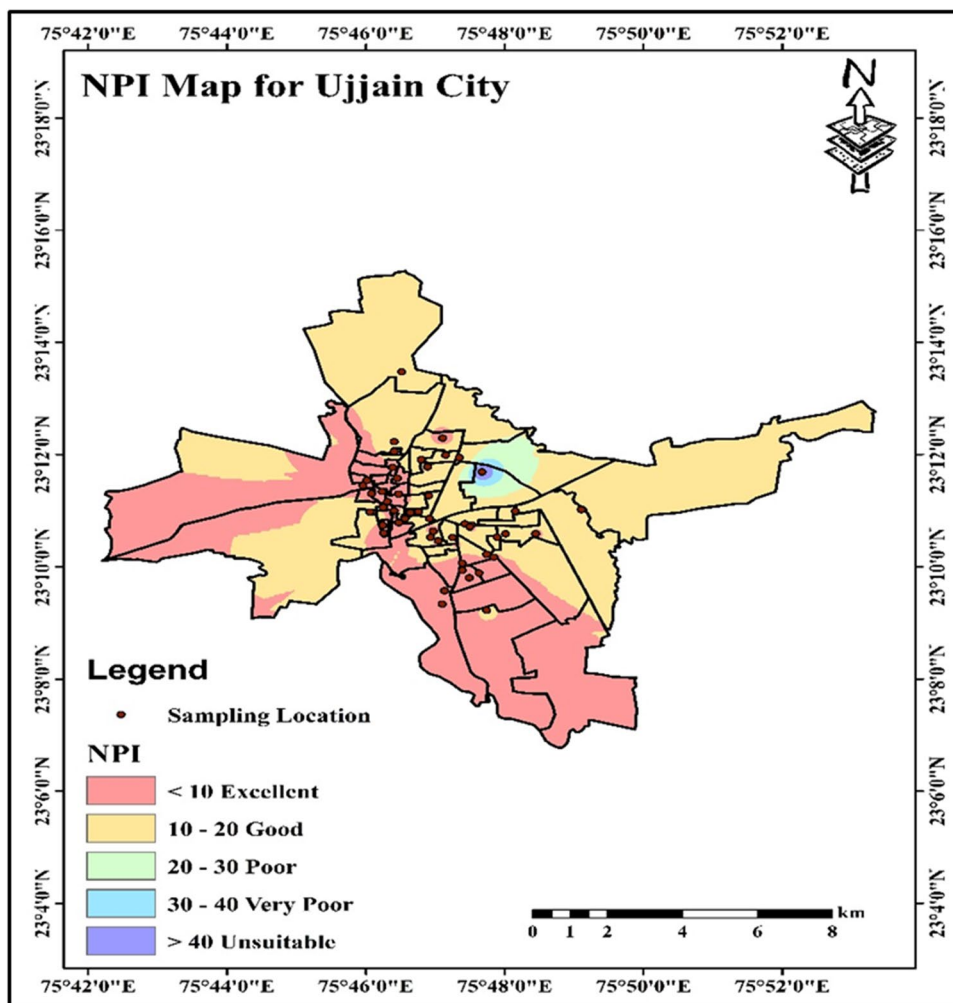
in the unsuitable category. The percentage wise distribution of samples shows that "Excellent Water" is found in 57% of samples, "Good Water" is 35%, "Poor Water" is 4%, "Very Poor Water" is 2%, and "Unsuitable Water" is 2%, which can be seen in Fig. 6a. To create a water quality index map in ArcGIS, NPI values were interpolated over the entire study region. The developed map can be seen in Fig. 5. The map clearly indicates that wards 34 and 17 lie in very poor to unsuitable categories.

The WA WQI

According to WA WQI, the readings show that numerous parameters have a greater impact on water quality than a single one (Fig. 7). The unit weights (W_i) of all parameters are given in Table 5. The WQI calculation for Ward 1 is shown in Table 6. The obtained WQI range is 23.7 to 90, i.e., from excellent to very poor categories. Figure 7 depicts the variation of WA WQI throughout the wards. The water quality ratings in Table 4 for WA WQI were used to classify the entire area. According to the results of the weighted arithmetic method, 2 samples fall into the excellent category, 29 fall into the good category, 21 fall into the poor category, and 2 into the very poor category. The percentage wise distribution of samples reveals that 4% of samples are Excellent Water, 53% are Good Water, 39% are Poor Water, and the remaining 4% are Very Poor Water, which can be seen in Fig. 6b. The IDW tool in ArcGIS was used to interpolate spatial data based on location coordinates and quality parameters. Figure 8 depicts the WA WQI map interpolation across the study area. The map clearly indicates that wards 39 and 45 lie in the very poor category.
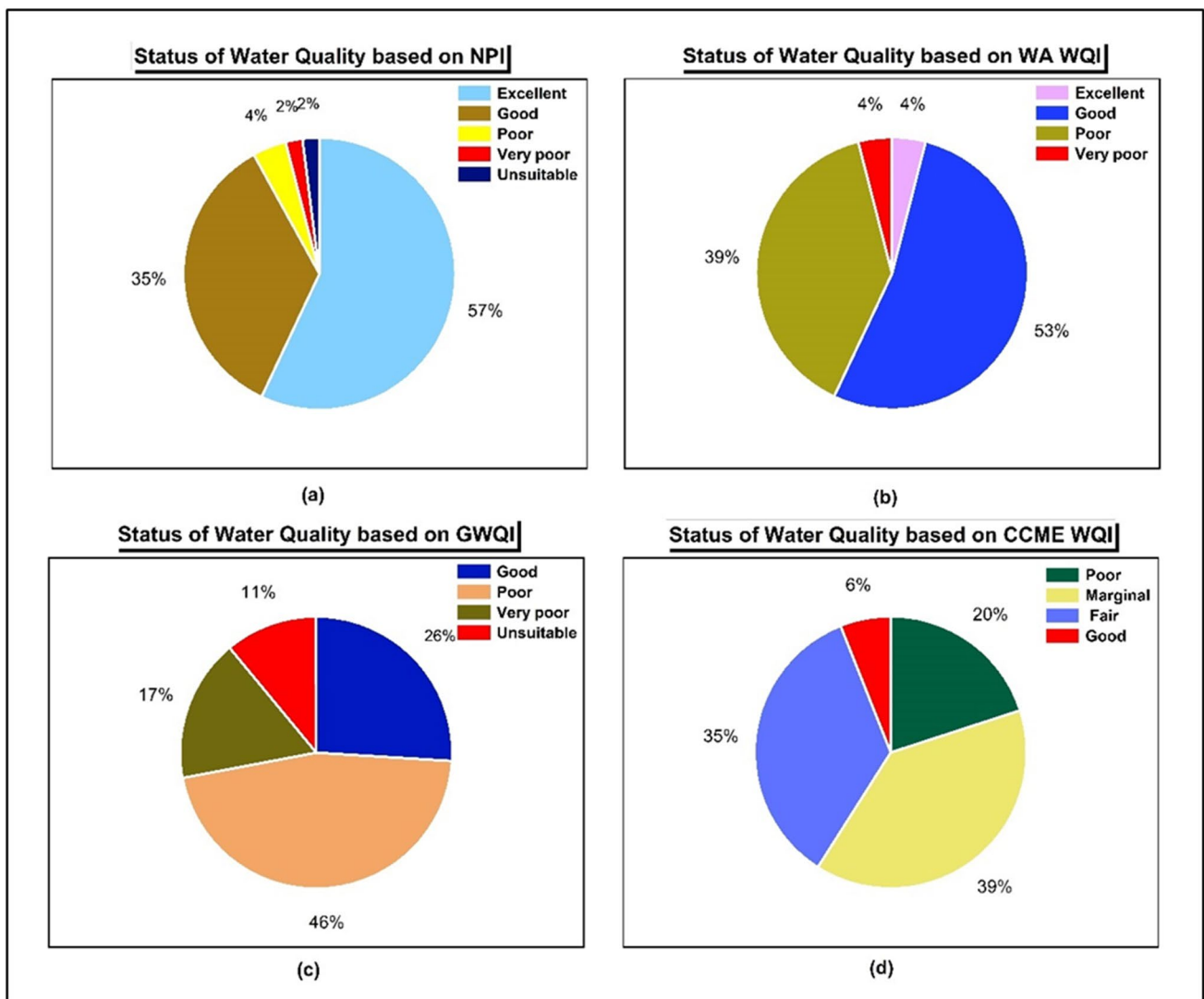

Fig. 6 Percentage wise status of water quality based on (**a**) NPI, (**b**) WA WQI, (**c**) GWQI, and (**d**) CCME WQI

GWQI

GWQI was determined for the 54 samples for the provided water quality parameters. An overall GWQI was calculated by combining all of the subindices for each parameter. The W_i of all individual parameters is specified in Table 5. Table 7 shows the GWQI calculation for Ward 1. The obtained GWQI range is 60.04 to 934.08, i.e., from good to unsuitable categories. Figure 9 depicts the variation of GWQI throughout the wards. The water quality ratings in Table 4 for GWQI were used to classify the entire area. According to the results, 14 samples are in the good category, 25 are in the poor category, 9 fall into the very poor category, and 6 fall into the unsuitable category. The percentage wise distribution of samples shows that 26% of samples are Good water, 46% are Poor water, 17% are Very poor water, and the remaining 11% are Unsuitable water, which can be seen in Fig. 6c. The IDW tool in ArcGIS was used to interpolate spatial data based on location coordinates and quality parameters. Figure 10 depicts the GWQI map interpolation across the study area. The map clearly indicates that wards 16, 17, 18, 34, 37, and 43 lie in the unsuitable category.

The CCME WQI

The CCME WQI was calculated for all samples gathered from 54 wards of Ujjain City. Table 8 shows the CCME WQI calculation for Ward 1. The obtained CCME WQI range is 8.52 to 89.43, i.e., from good to poor categories. Figure 11 depicts the variation of CCME WQI throughout the wards. The categorization of CCME WQI was done into five categories as per the rating given in Table 4. CCME WQI revealed that out of 54 samples, 3 samples fall into the good category, 19 fall into the fair category, 21 fall into the marginal category, and 11 into the poor category. The percentage wise distribution reveals that just 6% of samples were in the good category. Thirty-five percent of the samples were of fair quality. Similarly, 39% of samples were marginal, while 20% of samples were found to be poor, which can be seen in Fig. 6d. To create a water quality index map in ArcGIS, CCME WQI values were interpolated over the entire study region. The developed map can be seen in Fig. 12. The map clearly indicates that wards 1, 7, 16, 17, 18, 27, 34, 37, 39, 43, and 52 lie in the poor category.

MSA

MSA such as correlation analysis, PCA, and CA were used to evaluate variations and interpret a large complicated GWQ data set. MSA allows for the extraction of hidden information from a data set concerning the environment's potential effects on water quality. In this study, these methods were used on 54 samples for 8 variables. MSA was performed on a physicochemical data matrix. Minitab statistical software was used to accomplish the statistical analysis.

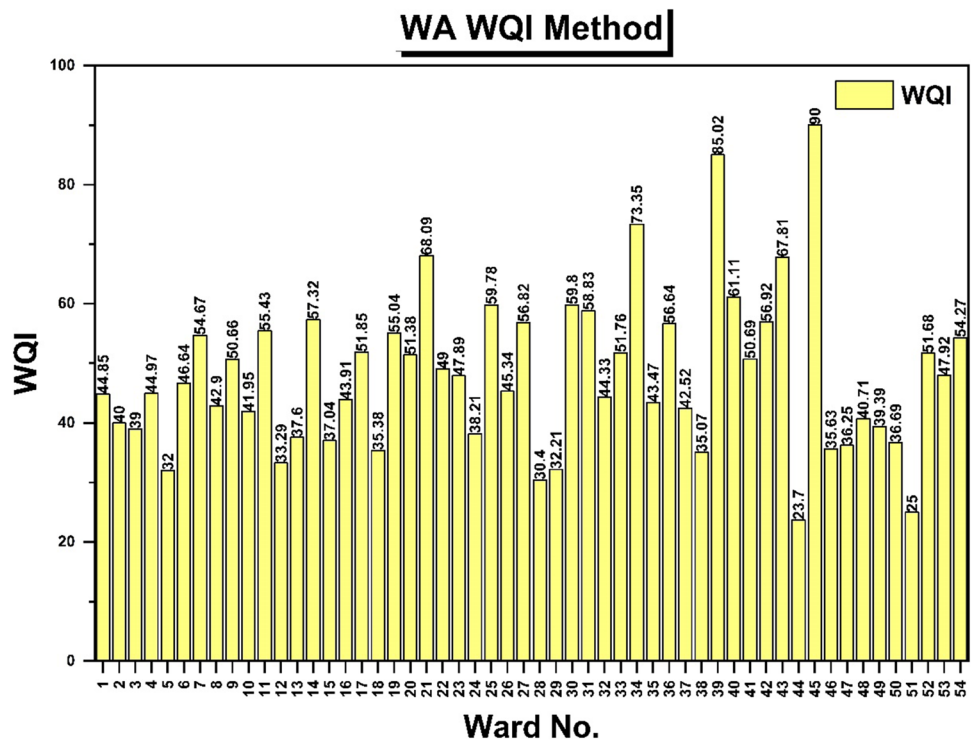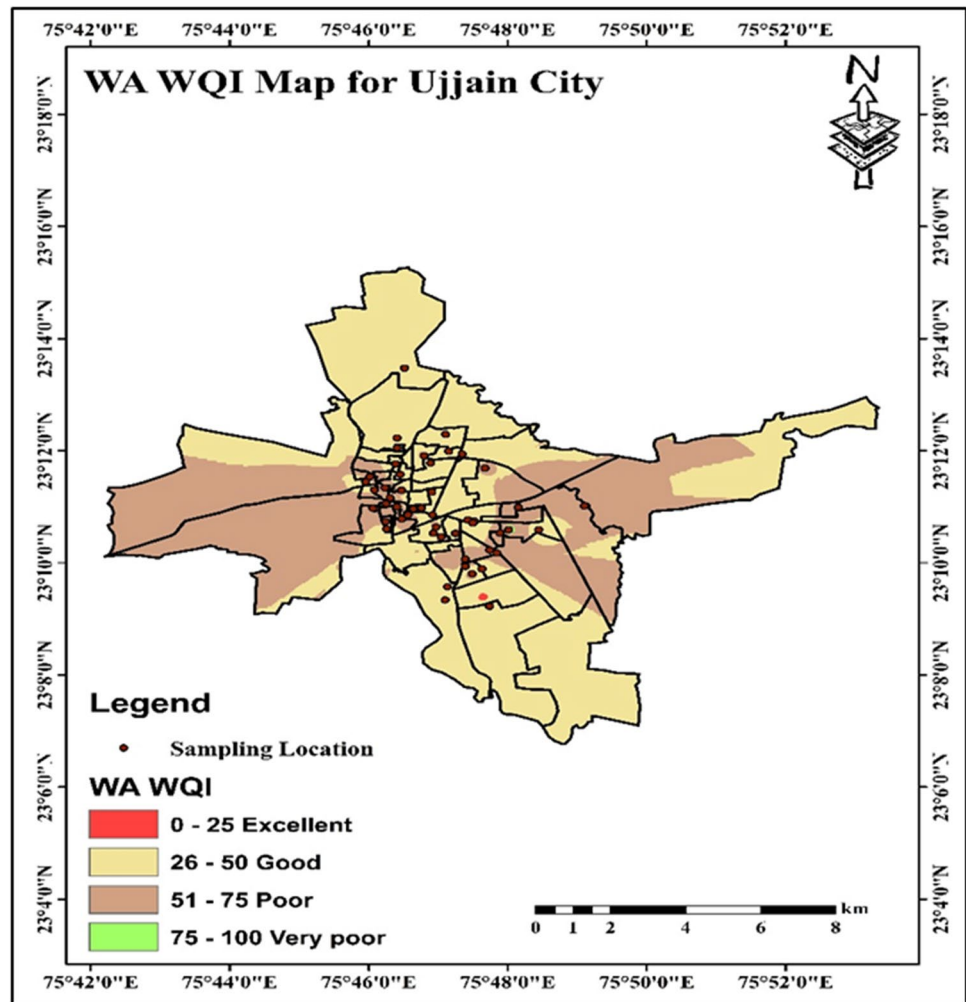**Fig. 7** Plot showing WQI for each ward

Fig. 8 WA WQI map for Ujjain City

Correlation Analysis

Table 9 shows the developed Karl Pearson's correlation matrix using MS Excel 2016 for the eight analyzed physico-chemical parameters. Among these, the most positive correlation (0.9996) was found between EC and TDS of water samples, which supports the fact that conductivity measurement is commonly used to estimate TDS, while the most negative correlation (−0.2784) was found between pH and chloride. According to the results, strong positive correlations are observed between EC, TDS, and chloride; TDS and chloride; and alkalinity and hardness. Except for fluoride, pH is negatively correlated to all the parameters. Apart from alkalinity and fluoride, turbidity has a positive relationship with all other parameters. Chloride is positively correlated to hardness and negatively correlated to fluoride. Hardness is negatively correlated with fluoride. EC, TDS, and alkalinity are the parameters that are positively correlated with other parameters. The most significant correlation is between EC,

TDS, and alkalinity, which has a greater impact on the total assessment of groundwater quality than any other parameter.

In Table 10, the parameter's standard permissible limit as per BIS 10500:2012 is stated. The pH value of drinking water is the concentration of H^+ ions, which defines the water as acidic or basic. It is observed that all of the samples were found to be in the BIS range of 6.5 to 8.5, and the majority of them were somewhat alkaline in nature. The turbidity ranged from 0.1 to 5.9 NTU. The turbidity of all samples is below the BIS permissible level of 10 NTU. From an aesthetic standpoint, turbid water is unappealing. The TDS is calculated using the EC of water. When the EC of water is high, the TDS value is similarly high. It is seen that the permissible limit is exceeded in 98.15% of EC samples and 62.96% of TDS samples. At ward number 17, the highest concentration of TDS was reported to be 5604 mg/l. Total alkalinity lies between 200 mg/l and 532 mg/l. Alkalinity is limited to 200 mg/l by the BIS. The permissible limit is exceeded by 98.15%

Fig. 9 Plot showing GWQI for each ward
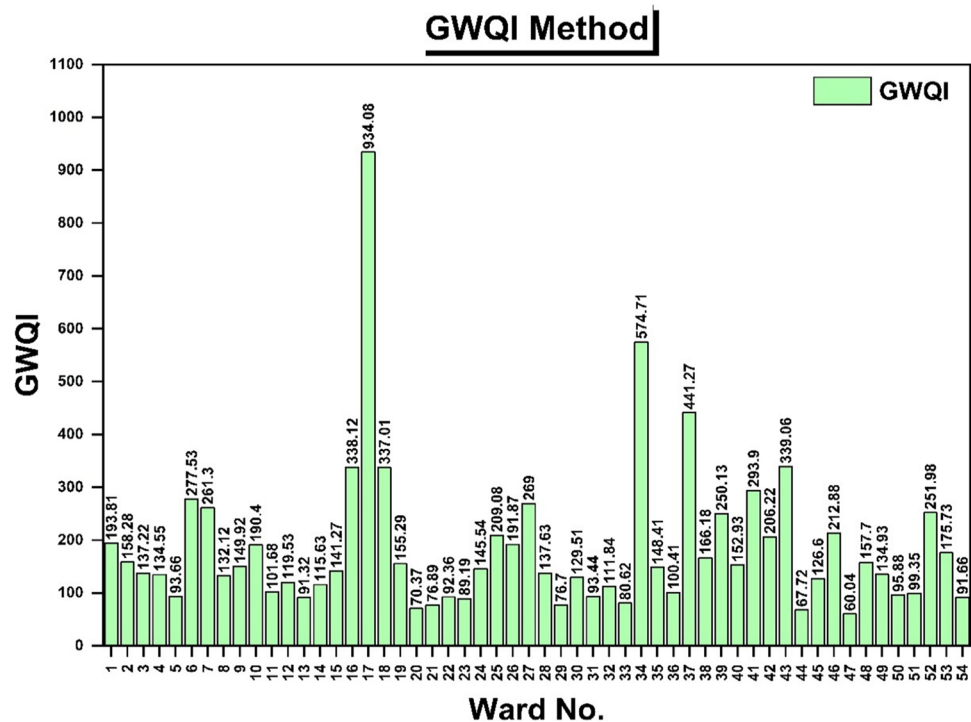
of alkalinity samples. The main source of alkalinity is dissolved carbon dioxide, which is found in large concentrations in most water sources. The permissible limit is exceeded by 83.33% of the hardness sample from the study area. At ward number 37, the maximum concentration of total hardness was reported to be 990 mg/l. The permissible limit for chloride and fluoride was exceeded in 25.92% and 9.25% of the samples, respectively.

PCA

In the present study, PCA was conducted on the correlation matrix of a water sample consisting of 8 physico-chemical parameters. It is used to identify individual loadings of eight variables in water quality. Eigenvalues are commonly used to obtain principal components (PCs). The eigenvalue of a significant variable defines its maximum value. The most significant eigenvalues are those greater than 1. PCs with eigenvalues of less than 1 were excluded due to their low essentiality [54]. As seen in a scree plot diagram (Fig. 13), the first three factors have eigenvalues greater than 1. After the third eigenvalue, the slope begins to decline slightly. As a result, only the first three components have been decided, with a total variance of 72.5%. Table 11 shows the percentage variance, eigenvalues, and loadings of three PCs. In addition, these loadings were classified as strong (> 0.75), moderate (0.75–0.50), and weak (0.50–0.30) [55]. PC1 featured positive loadings for turbidity, EC, TDS, alkalinity,

chloride, and hardness and negative loadings for pH and fluoride, accounting for 42.2% of the overall variance. This also suggests that these factors were tightly associated with one another, as evidenced by the correlation matrix, which signifies complete dominance in terms of water quality influence. Because of the dominance of solids in the groundwater, the first principal component, PC1, is referred to as the "solid component." Water quality was influenced the most by EC, TDS, and chloride in PC1. PC2 explicates about 17% of the total variance. Except pH and hardness, all other variables are positively loaded, with turbidity playing a significant role. PC3 had positive coefficients for turbidity, chloride, and hardness but was negatively loaded with pH, EC, TDS, alkalinity, and fluoride, accounting for only 13.3% of the total variation. It could be attributed to the dissolving of solids, which results in water turbidity and contamination due to soluble salts of chloride.

Cluster Analysis

Water samples were grouped using cluster analysis at each sampling location based on chemical and physical characteristics. All variables were standardized for hierarchical cluster analysis (HCA), and dendrograms were generated using Ward's technique with Euclidean distances. Using eight variables (pH, EC, TH, TDS, alkalinity, Cl⁻, F⁻, and turbidity), HCA was applied to the 54 sampling locations. The results are depicted in Fig. 14 as dendrograms.
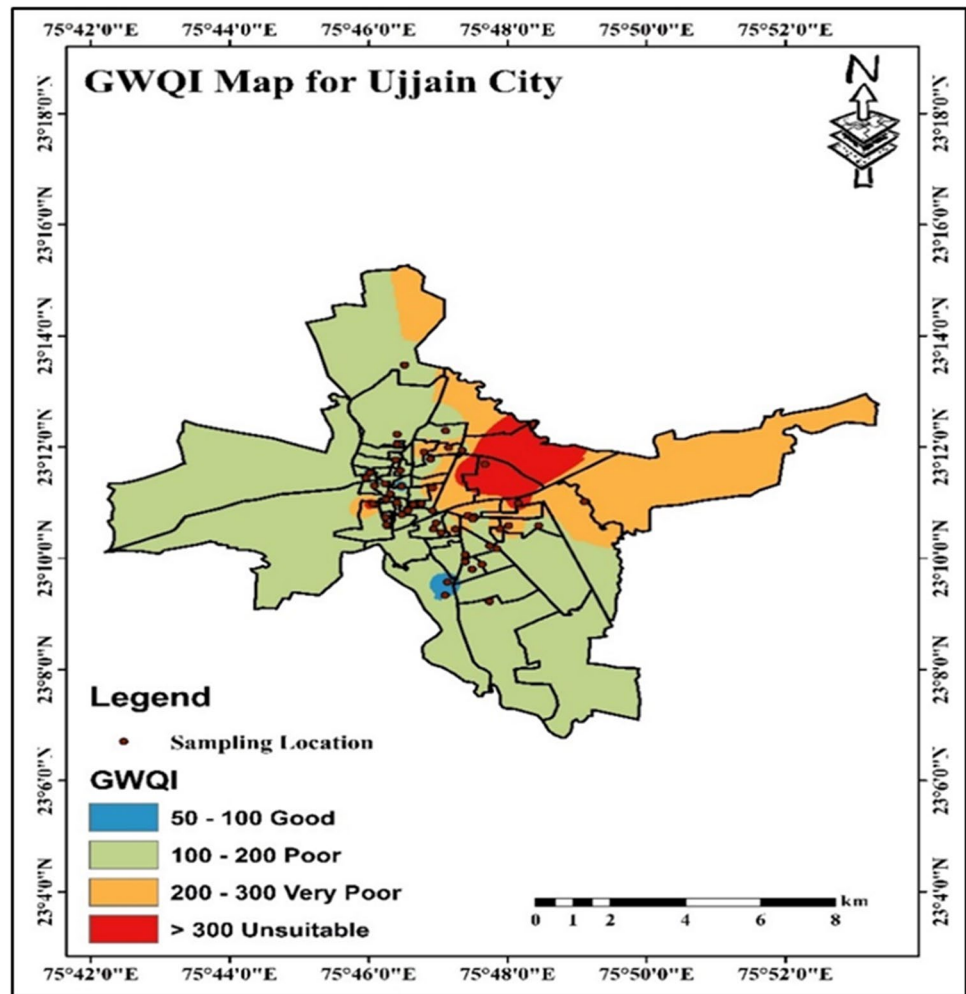
Fig. 10 GWQI map for Ujjain City

Figure 14a shows a dendrogram depiction of parameter cluster analysis. Here, cluster 1 includes two parameters: pH and fluoride. Turbidity, EC, TDS, and chloride are the four parameters covered by Cluster 2. Cluster 3 includes alkalinity and hardness.

Figure 14b shows a dendrogram view of the cluster analysis of sampling locations. CA creates three clusters from 54 sampling locations. Cluster 1 consists of 23 sampling locations, which are grouped into two subgroups (1, 4, 6, 8, 10, 16, 2, 3, 5, 12, 13, 15, 18) and (7, 9, 11, 14, 17, 20, 19, 22, 23, 26). Because 7 samples are of poor quality and 16 are of good quality, this cluster is classified as a moderately polluted zone. Cluster 2 is made up of 14 sampling locations that are divided into two subgroups (21, 25, 27, 30, 31, 33, 36, 41, 40, 42) and (34, 43, 39, 45). Because 12 samples are of poor quality and two samples are of very poor quality, this cluster is referred to as a highly polluted zone. Cluster 3 is formed of 17 sampling stations divided into two subgroups (24, 28, 29, 32, 35, 37, and 38) and (44, 51, 46, 47, 48, 49, 50, 52, 54, 53). This cluster is known

as a "low-pollution zone." There are 13 stations with good quality (24, 28, 29, 32, 35, 37, 38, 46, 47, 48, 49, 50), two stations with excellent quality (44, 51), and two stations with poor quality (52, 54).

Discussion

The results revealed that the quality of groundwater varies considerably depending on the location. After evaluating NPI, WA WQI, GWQI, and CCME WQI, it was determined that majority of the water samples were found to be safe for drinking. Because numerous physicochemical properties of the samples are below acceptable limits as per BIS 10500:2012, the WQI value is likewise lower, indicating that the water is fit for human consumption. Also, a spatial distribution map of the WQI were developed with the help of ArcGIS 10.8 software which can be seen in Figs. 5, 8, 10, and 12, demonstrating which parts of the groundwater are fit for drinking. However, all four

Fig. 11 Plot showing CCME WQI for each ward
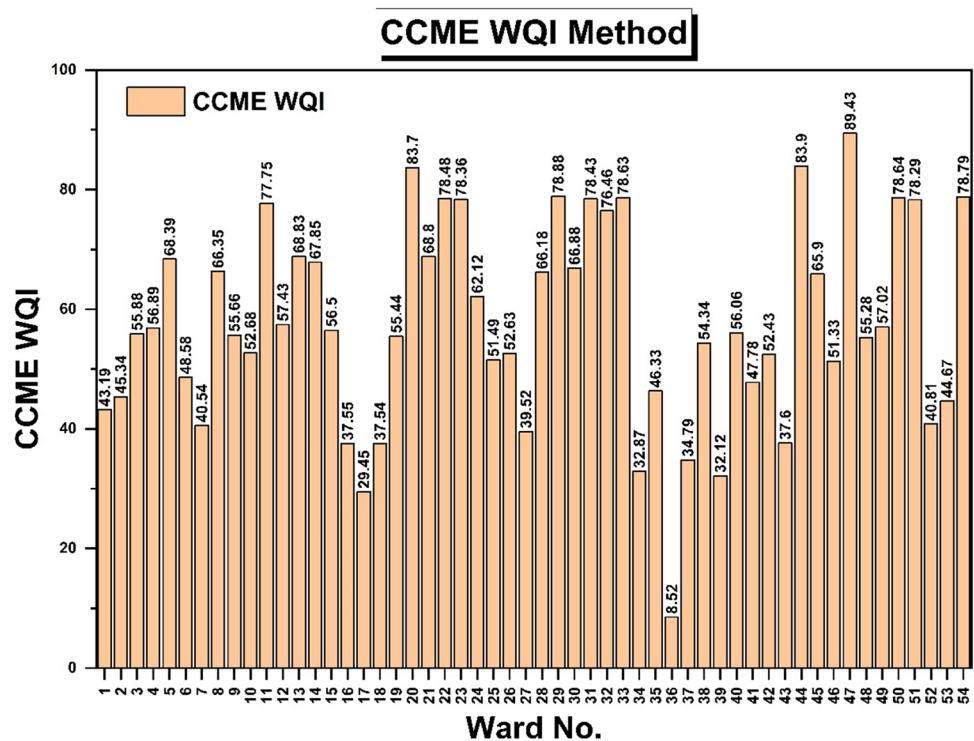
indexing methods and WQI maps reveal that samples from wards 20, 44, and 47 are in the excellent-to-good category, and samples from wards 17, 34, and 43 are in the poor-to-very poor category. As a result, the groundwater samples collected from these three wards are unfit for direct human consumption and must be treated before being consumed. This is despite the fact that numerous researchers have attempted to apply and compare various WQI methods. The findings of this study when it comes to the relative responses of these four WQI methods are in good accordance with some other publications [47, 56, 57].

As shown in Fig. 3a–h, thematic maps of all groundwater quality parameters have been created. The pH distribution pattern shows the presence of alkaline groundwater except for the central part (Fig. 3a). The EC is at 8673 mmhos/cm, with tiny spots in the central east (Fig. 3b). Due to inadequate fluxing and severely worn rock formations, a tiny patch in the city's central east exhibits TDS of 5604 mg/l in groundwater (Fig. 3d). This is in accordance with the higher EC (strong positive relation between EC and TDS). The alkalinity map clearly shows that it is largely in the middle portion of the city, with a few higher values in the northern and southern parts (Fig. 3e). The alkalinity of groundwater and hardness are found to have a strong positive correlation. This is reflected in the hardness map (Fig. 3c), which shows that the study area has hard groundwater. Chloride levels are highest in the central east, at 1320 mg/l, and it is clearly visible as a small patch (Fig. 3f). The correlation matrix

clearly shows that EC, TDS, and chloride have a high positive correlation, which is evident in their spatial distribution maps. F^- is an essential parameter that is found in isolated patches. The maximum concentration of 1.7 mg/l is found in the central part of the area (Fig. 3g). Keshavarzi et al. [58] state that fluoride in high concentrations (> 3.0 mg/l) can cause skeletal fluorosis. The turbidity map illustrates that the highest values are found in the eastern portion of the city, while the lowest values are seen in the western and central parts (Fig. 3h).

Positive values of the correlation coefficient indicate a strong and direct relationship between variables, whereas negative values indicate inverse relationship. As per Tirkey et al. [59] correlation analysis, $0.9 \leq R^2 \leq 1$ is considered to be strong, $0.9 \leq R^2 \geq 0.5$ is moderate, and $R^2 < 0.5$ is poor. The obtained results show a strong correlation of 0.9996 between EC and TDS, a moderate correlation between EC and chloride 0.7067, TDS and chloride 0.7028, and alkalinity and hardness (0.6015). The key contributors to water quality degradation are EC, alkalinity, hardness, and TDS as per Table 10, which reveals the summary of parameters based on the BIS 10500:2012 permissible limits. The high concentration of TDS is due to an increase in salts containing carbonates, bicarbonates, sulfates, calcium, chloride, sodium, potassium, and other ions [60]. Because these parameters can reduce water clarity, decrease photosynthesis, and induce gastrointestinal irritation in people, it is necessary to treat the water before
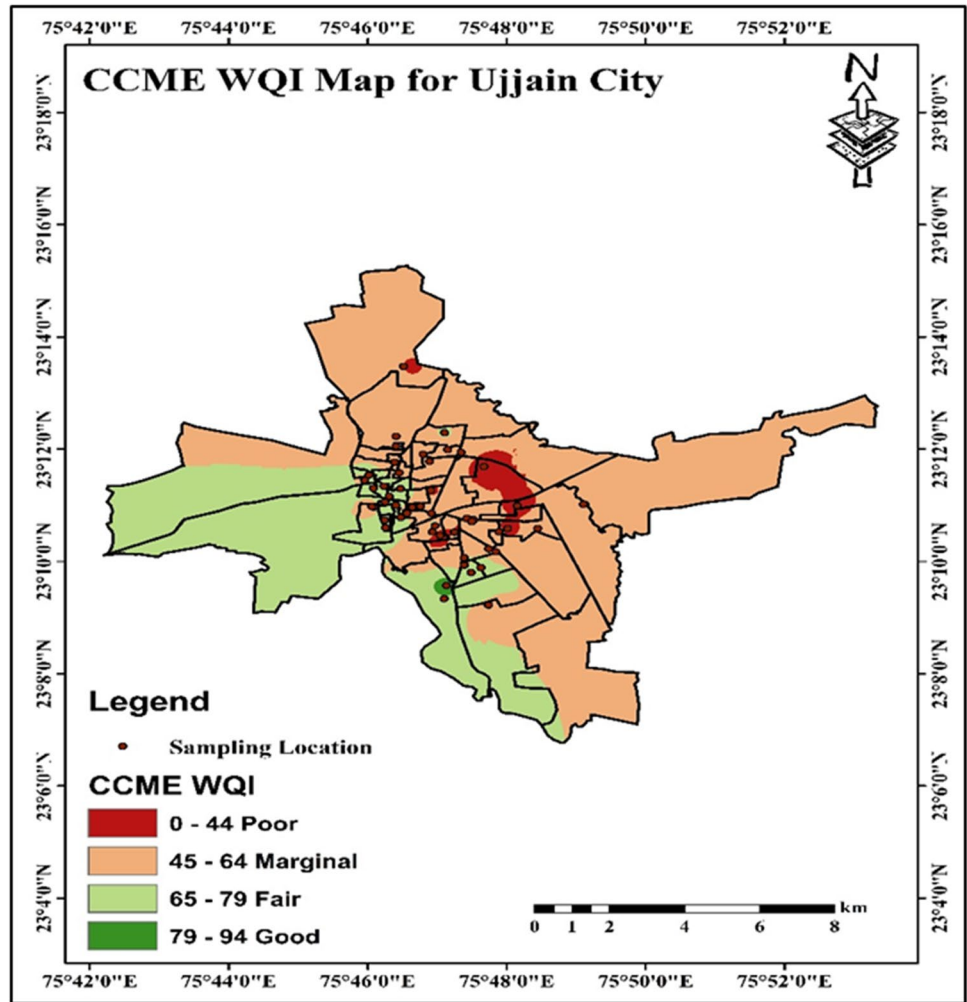
Fig. 12 CCME WQI map for Ujjain City

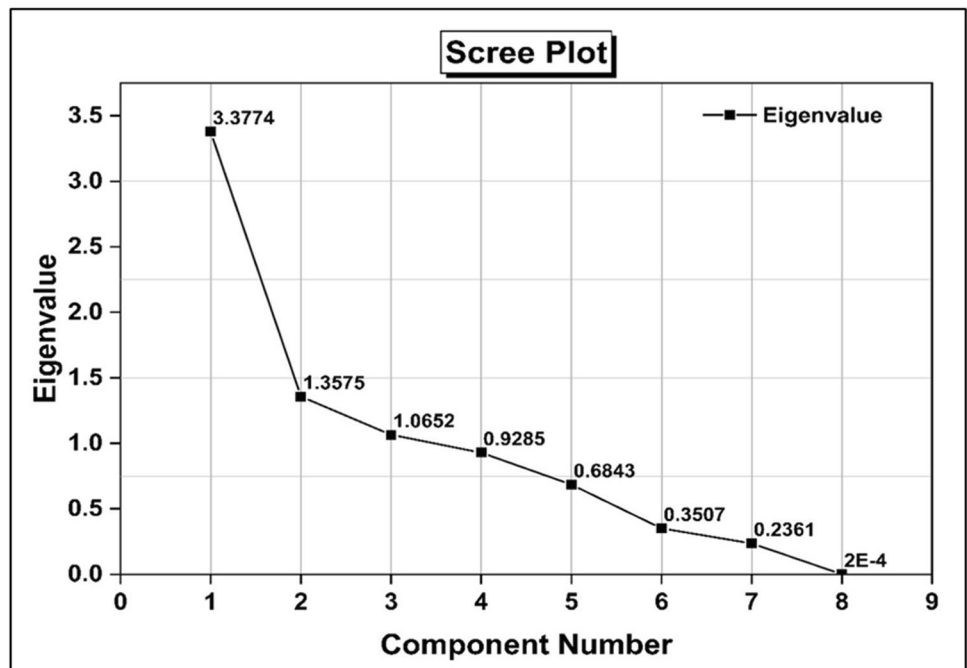**Fig. 13** Scree plot of principal component analysis

Table 11 Component loading and eigenvalues from PCA

Parameters	PC1	PC2	PC3
pH	−0.195	−0.019	−0.304
Turbidity	0.296	**0.366**	**0.132**
EC	**0.498**	0.174	−0.166
TDS	**0.496**	0.173	−0.168
Alkalinity	0.186	−0.727	−0.081
Chloride	**0.461**	0.085	**0.097**
Hardness	0.363	−0.516	0.031
Fluoride	−0.029	0.052	−0.904
Eigenvalue	3.3774	1.3575	1.0652
% of variance	42.2	17	13.3
Cumulative %	42.2	59.2	72.5

The bold values signifies the highest loading value for the respective principal component

it is consumed. However, pH that has a direct impact on water taste was found to be within acceptable limits. Drinking water with a high chloride concentration has a salty taste to it, making it unfit for consumption. Infants and children may be harmed if they drink chloride-rich water [20]. Fluorosis in the teeth and bones can be caused by an overabundance of fluoride in the water. However, the concentrations of Cl^-, F^-, and turbidity are all found to be within acceptable limits, indicating that they have little impact on human consumption.

According to the PCA, the variables were correlated to three PCs that were reported to account for 72.5% of the overall variation in groundwater samples. Positive loadings for EC, TDS, hardness, and Cl^- were found in the first component (PC1), which contributed 42.2% of the total variation. This can be attributed to the natural water source and is referred to as "water hardness salinity" [39]. Water quality is positively influenced by dissolved chloride salts in PC1. PC2 explains about 17% of the total variance, with turbidity, EC, TDS, chloride, and fluoride all having a positive influence. Turbidity shows the most significant loading on PC2. The physical characteristics of water, such as cloudiness, could be represented by PC2 [52]. PC3 displayed substantial positive coefficients (turbidity, chloride, and hardness) but was negatively loaded with pH, EC, TDS, alkalinity, and fluoride, accounting for just 13.3% of the total variance. This factor PC3 may be attributed to chloride solids dissolving, which causes water to become more turbid. Figure 13a shows the dendrogram view of the cluster analysis of 8 physicochemical parameters that are clustered into three groups of the same water quality characteristics. The CA allowed three clusters to be established between the sampling locations, indicating variances in water quality at various locations. According to the dendrograms, the study area was classified into three main groups: low, moderate, and high polluted zones based

on correlation between physicochemical parameters and sampling locations. CA supports the results of the correlation matrix.

In the present study, only four indexing methods were used to access the GWQI. This is owing to the availability of only eight physicochemical parameters. If more parameters are available, more indexing methods can be used. Furthermore, the current work is done for the summer of 2020. Future work can be done for the monsoon and post-monsoon periods to give a comparative result. The assessment of groundwater quality is of utmost importance as it is the primary source of drinking water in the study area. Therefore, a competent management plan must be enacted before its quality deteriorates.

Conclusion

The study used 54 groundwater samples taken during the summer of 2020 from 54 wards of Ujjain City, Madhya Pradesh, India, for water quality assessment using an integrated approach of different water quality indexes with GIS and multivariate statistical analysis. WQI categorizes water based on various parameters, culminating in a composite unit that may be used to determine the quality of water with a single numeric value. The results of different indexing methods, such as NPI, reveal 92% of samples to be in the excellent to good category, WA WQI shows that 57% of samples are in the excellent to good category, GWQI illustrates 72% of samples to be in the good to poor category, and CCME WQI reveals 80% of samples to be in the good to marginal category. According to the outcome of all the indexing methods, it is clear that the majority of the water is in good condition and thus suitable for drinking, with the exception of a few locations that require treatment. Table 4 reveals the classification of WQI values for each ward, calculated using different indexing methods. GIS aids in the conversion of point data into special data, allowing for the classification of areas with excellent and poor water quality. Groundwater quality mapping was implemented using IDW in ARCGIS 10.8 for the eight physicochemical parameters shown in Fig. 3a–h. The WQI spatial distribution map (Figs. 5, 8, 10, and 12) clearly depicts the area's finest drinking locations. According to BIS 10500:2012 standards, physicochemical parameters such as EC, TDS, and TH are above permissible limits. Within the study area, correlation analysis reveals strong and positive correlations between EC and TDS. PCA was used to determine three components in the water quality data set generated in this study. The first component appears to be associated with the presence of solid components, the second component appears to be related to the impact of water cloudiness due to turbidity, and the third component appears to be associated

Fig. 14 (**a**) Cluster analysis of 8 physicochemical parameters. (**b**) Cluster analysis of 54 sampling locations
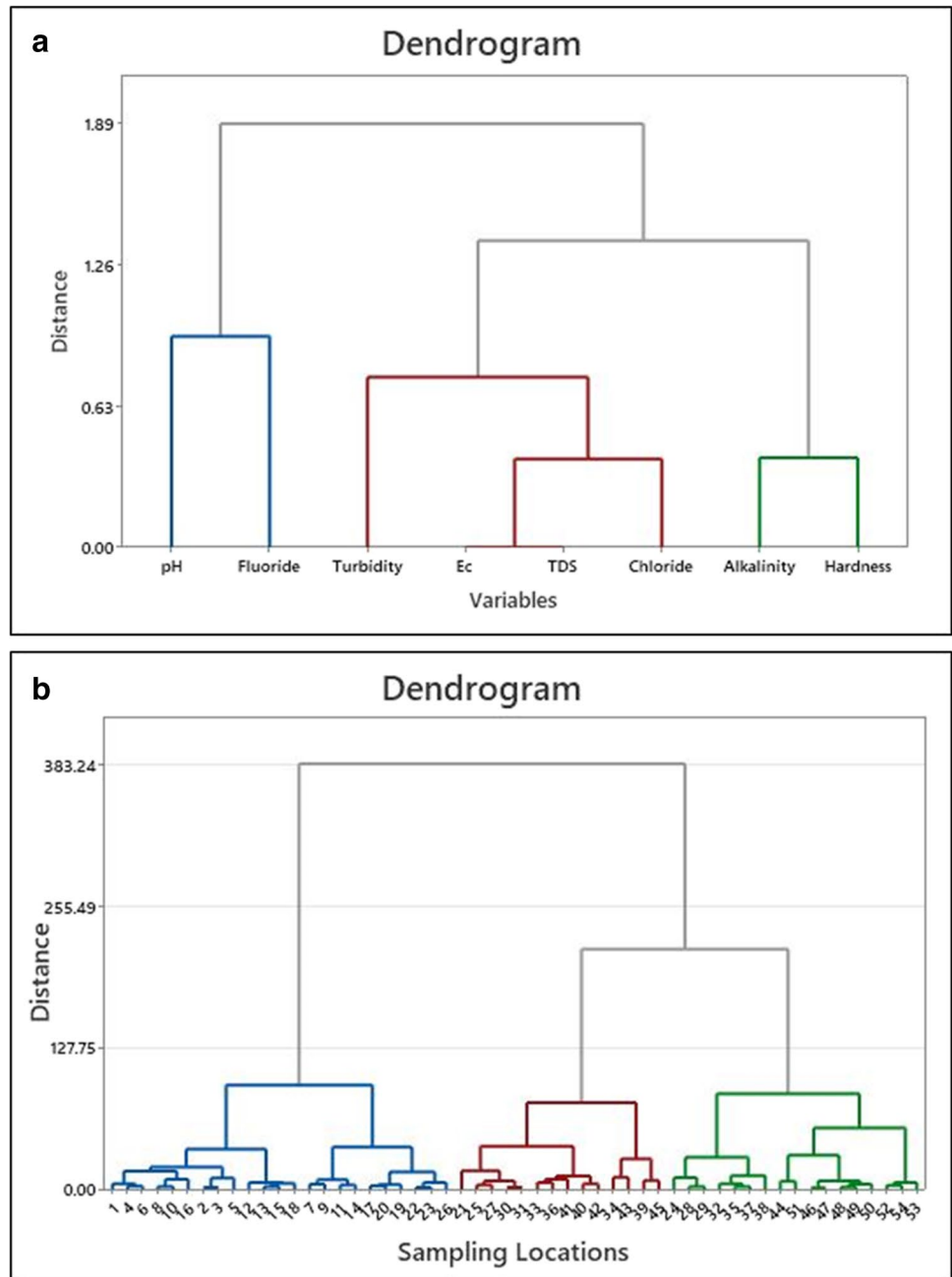
with the changes in water quality due to the presence of soluble chloride salts. CA groups 54 sampling locations into 3 clusters of similar water quality characteristics as seen in Fig. 14a, and 3 clusters of low, medium, and high polluted zones as per sampling locations in Fig. 14b. The PCA and CA supported the results of the correlation matrix. The study confirms that multivariate statistical analysis techniques, including correlation analysis, PCA, and CA, are effective in evaluating spatial variability and identifying contamination sources in the studied area. Moreover, an integrated approach of different water quality indexes using GIS reveals the most common wards that are fit or unfit for human consumption according to all indexing methods. According to the findings, samples from wards 20, 44, and 47 are the most common and in the excellent to good category, whereas samples from wards 17, 34, and 43 are the most common and in the poor to very poor category. It is therefore recommended that monitoring and management

should be prioritized in order to safeguard the groundwater resource from pollution and to provide technologies to make groundwater suitable for drinking purposes.

Declarations

Ethics Approval The research effort is original and does not include any findings from other sources, with the exception of cited work. As a result, the Ethical Standards are followed throughout the work.

Conflict of Interest The authors declare no competing interests.

References

1. Rajankar PN, Gulhane SR, Tambekar DH et al (2009) Water quality assessment of groundwater resources in Nagpur region (India) based on WQI. E-J Chem 6:905–908. https://doi.org/10.1155/2009/971242

2. Mahvi AH, Nouri J, BabaeiA AA, Nabizadeh R (2005) Agricultural activities impact on groundwater nitrate pollution A. Int J Environ Sci Technol 2:41–47

3. Ostad-Ali-Askari K, Shayannejad M, Ghorbanizadeh-Kharazi H (2017) Artificial neural network for modeling nitrate pollution of groundwater in marginal area of Zayandeh-rood River, Isfahan, Iran. KSCE J Civ Eng 21:134–140. https://doi.org/10.1007/s12205-016-0572-8

4. Milovanovic M (2007) Water quality assessment and determination of pollution sources along the Axios/Vardar River, Southeastern Europe. Desalination 213:159–173. https://doi.org/10.1016/j.desal.2006.06.022

5. Ostad-Ali-Askari K, Shayannejad M (2021) Quantity and quality modelling of groundwater to manage water resources in Isfahan-Borkhar Aquifer. Environ Dev Sustain 23:15943–15959. https://doi.org/10.1007/s10668-021-01323-1

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

7. Asadi SS, Vuppala P, Reddy MA (2007) Remote sensing and GIS techniques for evaluation of groundwater quality in municipal corporation of hyderabad (zone-V), India. Int J Environ Res Public Health 4:45–52. https://doi.org/10.3390/ijerph2007010008

8. Khan R, Jhariya DC (2017) Groundwater quality assessment for drinking purpose in Raipur city, Chhattisgarh using water quality index and geographic information system. J Geol Soc India 90:69–76. https://doi.org/10.1007/s12594-017-0665-0

9. Sirajudeen J, Abdul Vahith R (2014) Applications of water quality index for groundwater quality assessment on Tamil Nadu and Pondicherry, India. J Environ Res Dev 8:443–450

10. Kamboj N, Matta G, Bharti M et al (2017) Water quality categorization using WQI in rural areas of Haridwar, India. Int J Environ Rehabil Conserv 8:108–116

11. Chaurasia AK, Pandey HK, Tiwari SK et al (2018) Groundwater quality assessment using Water Quality Index (WQI) in parts of Varanasi District, Uttar Pradesh, India. J Geol Soc India 92:76–82. https://doi.org/10.1007/s12594-018-0955-1

12. Zotou I, Tsihrintzis VA, Gikas GD (2020) Water quality evaluation of a lacustrine water body in the Mediterranean based on different water quality index (WQI) methodologies. J Environ Sci Health A Toxic Hazard Subst Environ Eng 55:537–548. https://doi.org/10.1080/10934529.2019.1710956

13. Balamurugan P, Kumar PS, Shankar K et al (2020) Non-carcinogenic risk assessment of groundwater in southern part of Salem District in Tamilnadu, India. J Chil Chem Soc 65:4697–4707

14. Ketata M, Gueddari M, Bouhlila R (2012) Use of geographical information system and water quality index to assess groundwater quality in El Khairat deep aquifer (Enfidha, Central East Tunisia). Arab J Geosci 5:1379–1390. https://doi.org/10.1007/s12517-011-0292-9

15. Soujanya Kamble B, Saxena PR, Kurakalva RM, Shankar K (2020) Evaluation of seasonal and temporal variations of groundwater quality around Jawaharnagar municipal solid waste dumpsite of Hyderabad City, India. SN Appl Sci 2:1–22. https://doi.org/10.1007/s42452-020-2199-0

16. Magesh NS, Krishnakumar S, Chandrasekar N, Soundranayagam JP (2013) Groundwater quality assessment using WQI and GIS techniques, Dindigul district, Tamil Nadu, India. Arab J Geosci 6:4179–4189. https://doi.org/10.1007/s12517-012-0673-8

17. Kawo NS, Karuppannan S (2018) Groundwater quality assessment using water quality index and GIS technique in Modjo River Basin, central Ethiopia. J Afr Earth Sci 147:300–311. https://doi.org/10.1016/j.jafrearsci.2018.06.034

18. Sharma A, Ganguly R, Kumar Gupta A (2020) Impact assessment of leachate pollution potential on groundwater: an indexing method. J Environ Eng 146:05019007. https://doi.org/10.1061/(asce)ee.1943-7870.0001647

19. 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. https://doi.org/10.1016/j.pce.2008.06.019

20. Sadat-Noori SM, Ebrahimi K, Liaghat AM (2014) Groundwater quality assessment using the Water Quality Index and GIS in Saveh-Nobaran aquifer, Iran. Environ Earth Sci 71:3827–3843. https://doi.org/10.1007/s12665-013-2770-8

21. Kazi TG, Arain MB, Jamali MK et al (2009) Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotoxicol Environ Saf 72:301–309. https://doi.org/10.1016/j.ecoenv.2008.02.024

22. Massart D, Kaufman L (1984) The interpretation of analytical chemical data by the use of cluster analysis. pp 393–401

23. Hamlat A, Guidoum A, Koulala I (2017) Status and trends of water quality in the Tafna catchment: a comparative study using water quality indices. J Water Reuse Desalin 7:228–245. https://doi.org/10.2166/wrd.2016.155

24. Acikel S, Ekmekci M (2018) Assessment of groundwater quality using multivariate statistical techniques in the Azmak Spring Zone, Mugla, Turkey. Environ Earth Sci 77:1–14. https://doi.org/10.1007/s12665-018-7937-x

25. Ram A, Tiwari SK, Pandey HK et al (2021) Groundwater quality assessment using water quality index (WQI) under GIS framework. Appl Water Sci 11:1–20. https://doi.org/10.1007/s13201-021-01376-7

26. Roy B, Roy S, Mitra S, Manna AK (2021) Evaluation of groundwater quality in West Tripura, Northeast India, through combined application of water quality index and multivariate statistical techniques. Arab J Geosci 14.https://doi.org/10.1007/s12517-021-08384-6

27. Vishwakarma V, Thakur LS (2011) Assessment of ground water quality status in Ujjain City, Madhya Pradesh, India. Pollut Res 30:523–527

28. Vishwakarma V, Thakur LS (2012) Multivariate statistical approach for the assessment of groundwater quality in Ujjain City, India. J Environ Sci Eng 54:533–543

29. Zolekar RB, Todmal RS, Bhagat VS et al (2021) Hydro-chemical characterization and geospatial analysis of groundwater for

drinking and agricultural usage in Nashik district in Maharashtra, India. Environ Dev Sustain 23:4433–4452. https://doi.org/10.1007/s10668-020-00782-2

30. Nemerow NL, Sumitomo H (1970) Benefits of water quality enhancement report no. 16110 DAJ, prepared for the US Environmental Protection Agency. Retrieved from Link

31. Brown RM, McClelland NI, Deininger RA, O'Connor MF (1972) A water quality index — crashing the psychological barrier. Indic Environ Qual:173–182.https://doi.org/10.1007/978-1-4684-1698-5_15

32. Saffran K, Cash K, Hallard K (2001) CCME Water Quality Index 1.0 user's manual. Can Water Qual Guidel Prot Aquat Life:1–5

33. Ramakrishnaiah CR, Sadashivaiah C, Ranganna G (2009) Assessment of water quality index for the groundwater in Tumkur taluk, Karnataka state, India. E-J Chem 6:523–530. https://doi.org/10.1155/2009/757424

34 Gauns A, Nagarajan M, Lalitha R, Baskar M (2020) GIS-based assessment of groundwater quality for drinking and irrigation by water quality index. Int J Curr Microbiol Appl Sci 9:2361–2370. https://doi.org/10.20546/ijcmas.2020.903.269

35. In B, Madan Kumar P, Shivakumar M (2012) An assessment of groundwater quality using water quality index in Chennai, Tamil Nadu, India. Chronicles Young Sci 3:146. https://doi.org/10.4103/2229-5186.98688

36. Harshan S, Thushyanthy M, Gunatilake J et al (2016) Assessment of water quality index of groundwater quality in Chunnakam and Jaffna Town, Sri Lanka. 13:84–91

37. Agarwal M, Singh M, Hussain J (2020) Evaluation of groundwater quality for drinking purpose using different water quality indices in parts of Gautam Budh Nagar District, India. ASIAN J Chem 9:1128–1138

38. Wagh VM, Panaskar DB, Muley AA, Mukate SV (2017) Groundwater suitability evaluation by CCME WQI model for Kadava River Basin, Nashik, Maharashtra, India. Model Earth Syst Environ 3:557–565. https://doi.org/10.1007/s40808-017-0316-x

39. Simeonov V, Stratis JA, Samara C et al (2003) Assessment of the surface water quality in Northern Greece. Water Res 37:4119–4124. https://doi.org/10.1016/S0043-1354(03)00398-1

40. Güler C, Thyne GD, McCray JE, Turner AK (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeol J 10:455–474. https://doi.org/10.1007/s10040-002-0196-6

41. Nath A V., Selvam S, Reghunath R, Jesuraja K (2021) Groundwater quality assessment based on groundwater pollution index using Geographic Information System at Thettiyar watershed, Thiruvananthapuram district, Kerala, India. Arab J Geosci 14.https://doi.org/10.1007/s12517-021-06820-1

42. Singh KP, Malik A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) — a case study. Water Res 38:3980–3992. https://doi.org/10.1016/j.watres.2004.06.011

43. 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 101:4570–4576. https://doi.org/10.1016/j.biortech.2009.12.046

44. Chabukdhara M, Nema AK (2012) Assessment of heavy metal contamination in Hindon River sediments: a chemometric and geochemical approach. Chemosphere 87:945–953. https://doi.org/10.1016/j.chemosphere.2012.01.055

45. Héberger K, Milczewska K, Voelkel A (2005) Principal component analysis of polymer-solvent and filler-solvent interactions by inverse gas chromatography. Colloids Surf A Physicochem Eng Asp 260:29–37. https://doi.org/10.1016/j.colsurfa.2005.02.029

46. Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci Total Environ 313:77–89. https://doi.org/10.1016/S0048-9697(02)00683-6

47. Jankowska J, Radzka E, Rymuza K (2017) Principal component analysis and cluster analysis in multivariate assessment of water quality. J Ecol Eng 18:92–96. https://doi.org/10.12911/22998993/68141

48. George R, Martin GD, Nair SM et al (2016) Geochemical assessment of trace metal pollution in sediments of the Cochin backwaters. Environ Forensic 17:156–171. https://doi.org/10.1080/15275922.2016.1163623

49. Alfaifi HJ, Kahal AY, Abdelrahman K et al (2020) Assessment of groundwater quality in Southern Saudi Arabia: case study of Najran area. Arab J Geosci 13.https://doi.org/10.1007/s12517-020-5109-2

50. Nag SK, Lahiri A (2012) Hydrochemical characteristics of groundwater for domestic and irrigation purposes in Dwarakeswar Watershed Area, India. Am J Clim Chang 01:217–230. https://doi.org/10.4236/ajcc.2012.14019

51. Pius A, Jerome C, Sharma N (2012) Evaluation of groundwater quality in and around Peenya industrial area of Bangalore, South India using GIS techniques. Environ Monit Assess 184:4067–4077. https://doi.org/10.1007/s10661-011-2244-y

52. Kumar A, Krishna AP (2021) Groundwater quality assessment using geospatial technique based water quality index (WQI) approach in a coal mining region of India. Arab J Geosci 14.https://doi.org/10.1007/s12517-021-07474-9

53. Standards I (2012) Indian Standard Drinking Water — specification (second revision) IS 10500 (2012): Drinking water. Water Supply 25:1–3

54. Muangthong S, Shrestha S (2015) Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. Environ Monit Assess 187.https://doi.org/10.1007/s10661-015-4774-1

55. Reghunath R, Murthy TRS, Raghavan BR (2002) The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. Water Res 36:2437–2442. https://doi.org/10.1016/S0043-1354(01)00490-0

56. Akkoyunlu A, Akiner ME (2012) Pollution evaluation in streams using water quality indices: a case study from Turkey's Sapanca Lake Basin. Ecol Indic 18:501–511. https://doi.org/10.1016/j.ecolind.2011.12.018

57. Finotti AR, Finkler R, Susin N, Schneider VE (2015) Use of water quality index as a tool for urban water resources management. Int J Sustain Dev Plan 10:781–794. https://doi.org/10.2495/SDP-V10-N6-781-794

58. Keshavarzi B, Moore F, Esmaeili A, Rastmanesh F (2010) The source of fluoride toxicity in Muteh area, Isfahan, Iran. Environ Earth Sci 61:777–786. https://doi.org/10.1007/s12665-009-0390-0

59. Tirkey P, Bhattacharya T, Chakraborty S, Baraik S (2017) Assessment of groundwater quality and associated health risks: a case study of Ranchi City, Jharkhand, India. Groundw Sustain Dev 5:85–100. https://doi.org/10.1016/j.gsd.2017.05.002

60. Backer LC (2000) Assessing the acute gastrointestinal effects of ingesting naturally occurring, high levels of sulfate in drinking water. Crit Rev Clin Lab Sci 37:389–400. https://doi.org/10.1080/10408360091174259

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