Malaysian Journal of Chemistry, 2026, Vol. 28(3), 100-111
(Special Issue: *International Conference on Chemistry and Sustainability, IC-CS 2025*)

Water Quality Evaluation of the Himalayan Lake through Chemometric, Indexical, and Multivariate Statistical Approach: A Case Study

Sunaina Alok[1], Kishan Singh Rawat[2]* and Sudhir Kumar Singh[3]
[1]Department of Environmental Science, Graphic Era (Deemed to be University), Dehradun, Uttarakhand, India
[2]Department of Civil Engineering, Graphic Era (Deemed to be University) Dehradun, Uttarakhand, India
[3]Atmospheric & Ocean Science (KBCAOS), University of Allahabad, Uttar Pradesh, India
*Corresponding author (e-mail: ksr.kishan@gmail.com, kishan.ce@geu.ac.in)

The primary aim of this study was to evaluate the water quality of a Naini Lake and classify the lake water using the Water Quality Index (WQI). The study also aimed to identify the factors influencing water quality and examine relationships among water quality parameters using multivariate statistical techniques. Water quality data for the period 2017–2023 were obtained from the Uttarakhand Pollution Control Board, India. The selected water quality parameters for analysis include pH, Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Electrical Conductivity (EC), Temperature (Temp), Dissolved Oxygen (DO), Alkalinity, Chloride (Cl), Calcium (Ca), Magnesium (Mg), Total hardness (TH) and Total Dissolved Solids (TDS). The rating of WQI is grouped into three categories like excellent, good, poor water. According to the WQI the water can be termed as poor when WQI value is higher. Principal Component Analysis (PCA) identified three principal components (PCs) PC1 (EC, Alka, Cl, Ca, Mg, TH, TDS & Temp), PC2 (EC, Alka, Cl, Ca, Mg, TH, TDS & Temp) and PC3 (EC, Alka, Cl, Ca, Mg, TH, TDS) which explicate the sources of pollutions.

Keywords: Water quality, electrical conductivity, Chemometric, multivariate statistical approach

Received: March 2026; Accepted: March 2026

Lakes are aquatic water bodies and play a significant role in economic development and ecological balance. In many regions, lakes support fisheries, tourism, and drinking water supply [1]. Human activities, particularly uncontrolled tourism, climate change, and agricultural practices in surrounding areas, adversely affect lake health [2]. Surface water is currently highly vulnerable to contamination due to its readily accessible nature for the deposition of pollutants and effluents [3]. Surface water is characterized by a complex interplay of anthropogenic and natural processes occurring globally [4, 5]. These processes include weathering, erosion, hydrological dynamics, climate variability, precipitation patterns, industrial activities, agricultural practices, sewage discharge, and the water resources exploitation [4, 5, 6]. Increased inputs of sediment due to erosion, nutrients, and impurities from logging, farming, urbanization, and industry are causing freshwaters to degrade [7, 8]. When natural elements such as rainfall amount, vegetation type, landscape, terrain, soil kinds, and river discharge have an impact on any location, it becomes more difficult to preserve water quality [9, 10, 11]. Monitoring water quality serves a variety of functions, including ensuring that specific chemical concentrations are not exceeded, preventing harmful biological effects, highlighting species richness and/or diversity, ensuring that human society can sustainably use the aquatic ecosystem for a long time, and assessing compatibility with other elements of the ecological landscape [12].

The water quality index (WQI) is a widely used tool for classifying water quality. The Naini Lake, which simplifies the evaluation of the network data and offers a comprehensive judgement of water quality [13]. The WQI is a unified value used for comparing the water quality of two or more sources or over a period [5]. WQI is able to summarize huge information into quantitative value [14]. Jamion et al. [15], evaluated water quality of ex-mining lake amenities being converted to manmade wetlands in the Paya Indah wetlands of Selangor Malaysia with the WQI. Singh et al. [16], developed a fuzzy analytical process-based model for water quality for assessment of suitability of surface water for various uses. Zhou et al. [17], investigated the effects of diversion on improvement of water quality of Lake Dianchi, China through numerical modeling technique. Hu et al. [18], restored water quality parameters based on the Intelligent Optimization Algorithms-Machine Learning (IOA-ML) models and response to the short-term hydrometeorological factor for two large inland lakes and found a rise the water quality parameters. Singh et al. [19], employed Principal Component Analysis (PCA) and Cluster Analysis (CA) to evaluate water quality.

†Paper presented at the International Conference on Chemistry and Sustainability, IC-CS 2025

Many researchers recommended the use of combination of chemometric, and indexical approach for assessing the water quality [20, 21, 22]. It is pivotal to regularly assess the lake water quality and identify the driving factors which controls the quality of water in different seasons. This work's main objective was to investigate the Nainital Lake water quality utilizing several physio-chemical variables (temperature, pH, TDS, turbidity, alkalinity, BOD, DO), indexical and statistical method. Additionally, the different causes of the declining quality of lake water are being determined in order to calculate the appropriate countermeasures and preserve the geo-environmental stability of the area.

MATERIAL

The Nainital district is situated at 29.2794° N latitude and 79.4704° E longitude (Figure 1). Nainital town is also situated at an altitude of 1937 m and is sustained by Nainital Lake, which provides drinking water for a population of over 50,000 and an ever-expanding number of tourists. This tectogenically formed and thermally divided monomictic water body has a maximum and mean depth being 27.3 and 16.2 m respectively. The lake object s maximum width and depth is 1.4 km and 20.0 meters respectively, 0.45 km and 27.32 m, respectively, with average depth of 18.5 m, and it has a catchment area of 4.9 km^2, of which 10.4% constitutes water body, 18.3% barren, 19.3% inhabited and 48.4% vegetation. It has an annual rainfall of 2,030 mm. A transverse underwater ridge (110 m wide; 7 m below the surface) separates Tallital and Mallital. Water is provided to the lake through springs, precipitation, and inflows of several drainage nullahs, a few serving as principal conduits of impure silt and sewage. As per 2011 Census of Government of India, the district had a population of 954,605.

METHOD

Water quality data were obtained from the Uttarakhand Pollution Control Board (UKPCB), Uttarakhand, India (https://ueppcb.uk.gov.in/). The study period covered January 2017 to December 2023. Twelve physicochemical water quality parameters were analyzed, including pH, BOD, COD, EC, DO, alkalinity, chloride, calcium, magnesium, total hardness, and TDS. A water quality index (WQI), using this data set was then calculated. The WQI were determined using the quality standards provided by BIS (2012) and weighted arithmetic index method.

Statistical Analysis

To evaluate and justify the water quality immediately, the WQI is used. It is a numerical representation or rating system that considers physical, chemical, and biological parameters [19, 23]. While the specific parameters used in calculating the WQI may vary depending on the region or organization, common parameters include pH level, DO, BOD, TDS, temperature, conductivity, turbidity, and the presence of pollutants or contaminants like heavy metals, nitrates, phosphates, and faecal coliform bacteria.

By giving each metric, a weight based on how important they are in determining the quality of the water. These weights are determined through expert judgment or statistical analysis. Each parameter is then given a sub-index value based on its measured value and a standard or reference value. The total WQI score is computed by summing the sub-index values using the relevant mathematical procedures [19].

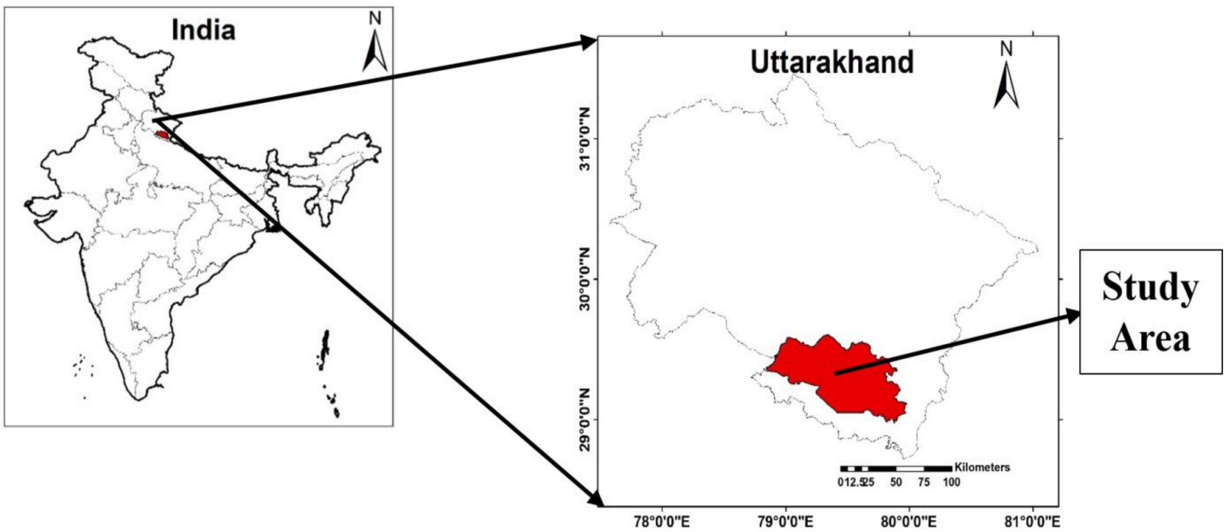

Figure 1. Location map of study area showed with red color.

The WQI for the sample was calculated using twelve physicochemical water quality criteria established by WHO guidelines for human consumption. The WAI method of Brown et al. [24] was employed to determine the WQI.

WQI Calculation, Involved Three Stages

Step-1. Assigning of unit weight (W_n) for the n^{th} parameter.

$$W_n = \frac{K}{S_n} \quad (1)$$

where, K is a constant and it is derived as eq. 2.

$$K = \frac{1}{\frac{1}{S1}+\frac{1}{S2}+\frac{1}{S3\ldots}+\frac{1}{Sn}} = \frac{1}{\Sigma 1/Sn} \quad (2)$$

Where Sn = Expected standard value of the nth parameter.

Step 2. Assigning of sub-index (Q_n) value to n^{th} parameter using the following eq. 3.

$$\mathrm{Qn} = 100 * [\frac{Vn-Vi}{Vs-Vi}] \quad (3)$$

Where Qn is the sub-index or rating of water quality value, Vn, the actual concentration of the water sample, Vs the standard value of the nth water quality parameter, Term n is the number of quality parameters, Vi is the ideal value for a pure water sample, and V0 is 0 for all parameters except for pH (which is 7) and DO (which is 14).

Step-3. Finally, WQI was derived using eq. 4.

$$\mathrm{WQI} = \Sigma Q_n . W_n / \Sigma W_n \quad (4)$$

since ΣW_n= 1

$$\text{Hence, WQI} = \Sigma Q_n . W_n \quad (5)$$

Multivariate Analysis

Tradition illustrates multivariate analysis as the statistical analysis of experiments in which each experimental unit is subjected to several measurements, and the structure and relationship between the multivariate measurements are critical for understanding the experiment [25]. Certain features of multivariate analysis resemble classic univariate analysis, such as hypothesis tests that compare distinct populations. The water quality characteristics were analysed multivariate methods on SPSS (version 16, Chicago, IL, USA) [11]. Basic statistics, such as correlation and multivariate analyses, were used to investigate the selected characteristics of Naini Lake water. Further, Pollutant sources identification was conducted by PCA.

Correlation Matrix

The following formula was used to calculate Karl Pearson's r in the present work, x and y are the water quality measures and n denotes the number of observations. The latter equation gives the correlation coefficient \(r\) between the variables \(x\) and \(y\) as:

$$r = n\sum(xiyi) - (\sum xi).(\sum yi)/\sqrt{[n\sum xi^2 - (\sum xi)^2][n\sum yi^2 - (\sum y)i]^2} \quad (6)$$

where two distinct water quality characteristics are represented by the variables x (x = values of the x-variable) and y (y = values of the y-variable).

There is n = number of data points, and all summations must range from 1 to n. The following straight-line equation can be used to define a straight linear regression

$$y = ax + b \quad (7)$$

where x and y denote independent and dependent variables (or unknowns), respectively, and a and b are the slope and the y-intercept of the line.

The following formula can be used to define the slope, a, and y-intercept, b:

$$a = n\sum(xy) - \sum x\sum y/n\sum(x^2) - (\sum x)^2 \quad (8)$$

$$b = \sum y - (a\sum x)/n \quad (9)$$

In order to establish a correlation between x and y, the experimental data pertaining to the variables x and y must be utilized to construct parameters a and b. Constants a and b have values determined by the relations in accordance with the accepted least squares approach. Correlation analysis is a statistical technique for displaying connections between variable pairs. Closer to + 1 or 1, correlation coefficient values indicate a good possibility that the variables X will occur, and Y is linearly related.

RESULTS AND DISCUSSION

Hydro-geochemistry of Lake

The results of a statistical analysis of surface water data obtained from UK PCB, Uttarakhand, are presented (Table 1). The chloride content ranged from 11 to 28 mg/l, averaging 19.37 mg/l. EC values ranged from 310 to 510 µS/cm, with a mean value of 427.86 µS/cm. The nature, quality, and type of surface water are mostly determined by its physiochemical characteristics [16]. Surface water quality greatly influences its suitability for domestic, agricultural, and industrial uses and human consumption [26]. Rusydi et al. [27] according to which the EC measures the water ability to transmit the electric current. It instantly indicates the number of ionised particles in the test sample. The mean TDS concentration was

291.61 mg/l, which is within the permissible limits for drinking water. Freshwater generally contains TDS concentrations below 500 mg/l. Some international factors like pH, value of water is influenced by temperatures, equivalency, solubility of oxygen and concentration of carbon dioxide. Magnesium ion concentration fluctuated between 86 to 158 mg/l, calcium ion concentration fluctuated between 94 to 192 mg/l and the highest concentrations of calcium and magnesium are much higher not higher than the maximum levels recommended by BIS [28] of 75 and 30 mg/l, respectively. Calcium and magnesium are abundant in carbonate phases like calcite and dolomite [29]. Magnesium is largely derived from water by exchange with minerals in rocks and soils.

Lakes are more vulnerable to water contamination than any other kind of water source due to the existence of motionless water that lacks the ability to purify itself [4]. Surface water, readily accessible for the discharge of contaminants and effluents, is presently the most susceptible to pollution [16]. Surface water quality is influenced by a variety of factors. weathering, erosion, hydrological features, precipitation, climate change, exploitation by people, of water resources, natural processes (like the weathering process, erosion, and agricultural practices), and human activities (like manufacturing activities, farming procedures, and sewage discharge) [17, 21]. The degradation of freshwater sources is a consequence of increased inputs of silt, nutrients, and pollutants from various sources such as farming, logging, urbanization, and industry [10, 15]. These effects lead to a decrease in the biological diversity of aquatic ecosystems and a reduction of water quality [6]. The climate throughout the basin has a major role in surface runoff being a seasonal phenomenon, whereas human discharges provide a continuous source of pollution [30]. In addition to altering the physical and chemical characteristics of soils, erosion of land within the lake catchment has added nutrients to the lake's water Upadhyay et al. [31]. The growth of settlement surrounding a lake without an appropriate sewage infrastructure exacerbates organic and nutrient loading in the lake, a problem that affects all urban water bodies [32]. The Naini Fault, which splits the synclinorium formed by the Naintal Hill into two pieces diagonally. Lithologie of the Krol Formation (Precambrian) which mainly compose of such as limestone, dolomite, gypsum, argil-limestone argil-slate, ferrigenous shale, greywackes and so forth. envelopes the catchment area [33]. The southwest is composed by dolomite together limestone and black carbonaceous slates and the northwest by argillaceous limestone and marlites [30]. Now-blocked upper reaches of the Balia River, which formed Lake Naini, is interpreted to be result of recent rotations. The peripheral drainage basin, forming the hilly slopes and combs, contributes its flows to the lake. Hydrological studies water balance: net gain and loss modules and to estimate and measure the elements of the input and output of the lake were worked out using radio isotopes [34, 35]. The study of Sediment and hydro biogeochemistry of Lake Nainital, Kumaun Himalaya, India was investigated by Purushothaman et al. [34]. Except when an exceptionally large decrease in the number of fish was observed, the inflow and outflow were important, accounting for 63-44% and 41-44%, respectively, of the influx undersea and outflow multiple times. The outflows were categorized as surface outflow, underwater outflow through springs on the downstream side, draft through wells for Nainital City water supply and evaporation loss from lake surface [35].

Table 1. Descriptive statistical summary of water quality parameters.

Parameters	BIS (Limits)	Min	Mean	Median	StD	Max
pH	6.5-8.5	7	7.59	7.57	0.26	8.19
BOD	3 (CPCB)	1	1.61	1.6	0.36	2.6
COD	250	4	5.76	6	0.87	8
DO	6	6	7.99	7.93	0.94	9.8
EC	400	310	427.86	430	40.45	510
Temp	30	6	13.83	14	4.14	24
Alka	200-600	144	214.64	213	30.23	284
Cl	250-1000	11	19.37	19	4.41	28
Ca	75-200	94	141.90	141	20.59	192
Mg	30-100	86	116.74	116	16.17	158
TH	200-600	208	258.64	257	32.40	344
TDS	500-2000	209	291.61	291	27.56	360

Note: Numerical values presented in pH (pH unit), EC (µs/cm), all other in (mg/l). Min, Minium, StD; Standard Deviation, Max170; Maximum CEC; Electrical conductivity, Temperature, Alka; Alkalinity; Cl, Chloride, Ca; Calcium, Mg; Magnesium, TH; Total Hardness, TDS; Total dissolved solid

Water hardness is a result of calcium and magnesium cations and such anions as sulphates, chloride, and bicarbonates [4]. Natural water often includes high levels of calcium and magnesium ions. Calcium and magnesium are typically found in carbonate minerals like dolomite and calcite [19]. The primary source of magnesium in Naini Lake water is the ion exchange of minerals between water and rocks and soil.

WQI Interpretation

WQI ranges from 70 to 110; as a consequence, they are categorized into three categories as excellent, good, and poor water (Table 2; Figure 2). Based on WQI classification of entire year 2022 to 2023 (24 months), Naini Lake fell under the poor water quality category class which is 50% of study time period. While 33 (39.29% of 84 months) time Naini Lake fall in class good, 5 (5.95% of 84 months) time Naini Lake comes under class E (excellent), while 54.76 % time period Naini Lake fall under the poor class. Water quality has deteriorated considerably in recent years, very much due to the high influx of tourists, rising demand of water and the high inflow of water into the lake during the monsoon season [30]. The conversion of forest land into haphazard built-up areas and habitational centers, the shifting of the lake's riparian zone into cemented areas, and the construction of concrete roads [30]. As a result, runoff has significantly increased, affecting the region's ability to retain water and percolate, raising groundwater levels, negatively impacting groundwater recharge zones, and creating ecological imbalances that affect the existence and natural growth of flora and fauna in the area.

Higher CO_2 levels lower pH in natural water pH is governed by the system carbon dioxide-bicarbonate-carbonate. A 25 °C rise in temperature causes the pH of pure water to drop by roughly 0.45. Minimize this temperature effect by buffering ions such as CO_3^-, HCO_3^-, and OH^-. The pH minimum and maximum for the lake is 7 and 8.19, respectively. This shows that the lake is slightly alkaline in nature. Oxygen utilized by microbes to decompose organic materials in water is measured by the BOD. It aids in the evaluation of how organic matter affects a reservoir's dissolved oxygen levels [16]. Maintaining water quality effectively requires this. BOD for the lake is towards the higher range, which is not suitable consumption and also not conducive for aquatic life. The minimum and maximum BOD was 1 mg/l and 2.6 mg/l with a mean value of 1.61 mg /l. Quantity of oxygen required to oxidize the organic compounds in water is indicated by COD [16].

It is employed to assess a water sample's organic matter content and oxidation potential. The estimation of inorganic chemical content is also aided by COD testing. COD for the lake was under desirable limit. Water quality is reflected in the DO level, which show the amount of oxygen accessible for aquatic life. Overabundance of organic matter can damage aquatic environments by creating eutrophication, which depletes oxygen. There is an inverse link between the level of DO and water temperature, with DO level being lowest in the summer. DO of the lake is 6mg/l to 9.8 mg/l. Two aerators have been installed in the lake to maintain dissolved oxygen concentrations. There 30 discs all over the lake which supply 01 billion of oxygen bubbles per day. The measurement of EC in water is determined by the presence of dissolved inorganic ions such as calcium, sodium, and chloride. Organic constituents do not significantly affect EC because of their low conductivity. Because EC enhances with temperature, it is frequently measured at 25 °C for consistency. EC of the lake was always above the permissible limit. The mean EC was 428 µs/cm. Alkalinity is the measurement of ions in water, primarily from hydroxide, carbonate, and bicarbonate, that balance out hydrogen ions. It comes from organic waste decomposition, mineral breakdown, and CO_2 solubility. When combined with certain cations, high alkalinity can cause sediment development, which can clog pipes and give water an unpleasant taste. The alkalinity of the lake varied from 215mg/l to 284mg/l. The temperature of the water affects biological activity, growth, and the kinds of animals that live in aquatic habitats. Every species has a range of temperatures that works well for it; straying from those limits can cause population declines or even extinction. Measurements of amount and water chemistry are also impacted by temperature. The mean temperature of the lake was 13.83 °C. Animals as well as plants need calcium to survive, as it is essential for the framework of skeletal systems, shells, and plants. Water is contaminated by deposits of limestone and dolomite. While calcium carbonate in modest amounts can prevent corrosion in pipes, too much calcium can cause harmful calcite buildup, which is typically treated with water softening. The minimum and maximum concentration of calcium in the lake water was 94 mg/l to 192 mg/l. The eighth most common element in the crust of the Earth, magnesium is found naturally in water and is necessary for all living things. Magnesium is found in rocks like dolomite and magnetite. In the human body, muscles and tissues contain 40% while bones contain 60% of magnesium. The concentration of magnesium is high in the lake water which is between 86 mg/l to 158 mg/l which should be between 30 mg/l to 100 mg/l. Calcium and magnesium dissolved concentrations are used to measure water hardness. High concentrations of these minerals in hard water cause "soap scum" to form when soap combines with calcium Hardness varied from 208 mg/l [36] to 344 mg/l. Sewage, the ocean, and industrial waste are the main sources of chloride, which is primarily introduced into water by the dissolution of salts such as NaCl. Surface waters usually have less chloride than do ground waters.

Although chloride is necessary for human metabolism, excessive amounts of it can harm plants and infrastructure. Chloride in the lake water was11 mg/l to 28 mg/l. Minerals like K, Na, Cl, Ca, and Mg that are dissolved in water and are classified as TDS. High TDS levels are a sign of extremely mineralized water, that can lead to discolouration and bad taste. This makes it crucial to consider while using water. The TDS of the lake varied from 209 mg/l to 360 mg/l as minimum and maximum respectively. In our study it was observed that the BOD, calcium, magnesium, total hardness and TDS are in a range.

Statistical Results

The quality of Naini Lake water and its appropriateness for consumption may be evaluated. According to hydro-chemical analysis of the Naini Lake, the water ranges from slightly acidic to neutral. pH, EC, Mg and physicochemical parameters were observed to be above the permissible limits as per BIS [28] for drinking purpose. Descriptive statistics were created using seven years of monthly Naini Lake water quality data and the range of values obtained. Correlation matrices are produced to determine the significance level of each parameter across a research period. Correlation analysis revealed significant positive relationships among COD, Ca, and Mg. The correlation study also shows that COD, Ca, and Mg are produced from the same source. Mineral dissolution from rock water interfaces, the influence of anthropogenic activities, and ion exchange mechanisms within Naini Lake water are the primary explanations for the observed variability in Naini Lake water quality, as revealed by correlation analysis. Table 3 showed the Pearson's correlation study of Naini Lake water's physiochemical characteristics. COD, Ca, and Mg exhibited strong correlations and significant associations with a few physicochemical variables; however, COD showed moderate to high correlations with BOD. Magnesium and calcium revealed strong positive connections. According to Namiotko et al. [37], the acidic environment is very corrosive in comparison to the host soil and surrounding rock, increasing the concentration of most ions in Naini Lake water. According to the findings of Pearson's correlation between physicochemical variables, TDS revealed a substantial association with EC, whereas TH and Ca had a somewhat positive correlation with alkalinity.

PCA results suggest that weathering processes, leaching, and anthropogenic activities are the major factors influencing Naini Lake water quality. PCA is the most effective approach for identifying the high-dimensional variable responsible for the phenomenon [2, 23]. Table 4 reveal the first three frequent factors (as principal components are 2827.259, 1800.914 and 474.788, respectively) which having maximum variance (and corresponding percentages of variance are 52.455%, 33.413% and 8.809 %.). And the combined contribution of these three primary components is up to 94.68%, which corresponds to the first two extraction principles discussed before. As a result, it is possible to conclude that these three frequent elements encompassed all of the information from the 12 indicators used for this study. These are the only three components that have been gathered and flipped. Rotating the factors closes the gap between their variances by redistributing the variation that describes each factor's starting variables. As a result, these three components make up the essential details of the original data set.

The data regarding the eigenvalues of each factor is shown in the scree plot (Figure 3). It is helpful to determine the right amount of principal component factors when selecting factors by applying the principles component analysis concepts. As seen in Figure 3, only eigenvalues larger than one are taken into account. It is revealed that the first three components offer more pertinent information by using them as the primary structure component to be considered rather than all variables. The scree plot of the PCA eigenvalue curve (Figure 3) shows that the eigenvalues of the fourth component are less than one, and the curve gradually gets softer. This implies that the explanatory capability decreases and mutation occurs at eigenvalue λ 4. Above importantly, the number of major components is set at three.

Tables 5 (before rotation) and 6 (after rotation) display the factor loading matrix. These tables indicate which components two distinct variables load on. As mentioned above, the load factor has polarised right after the rotation. EC, Temp, Alka, Cl, Ca, Mg, TH, TDS as the loading coefficient of these parameters in the first component is larger (in a mathematical sense). EC, Temp, Alka, Cl, Ca, Mg, TH, and TDS have a higher load coefficient for the second component (second programme), showing a higher association with the first. Both EC, Alka, Ca, Mg, TH, and TDS have stronger loading coefficients in the 3rd component, in absolute terms. To further clarify the possible group or sources that may change the Naini Lake water chemistry, the varimax rotation method was applied to maximize the sum of the square of the coefficient components (Figure 4).

Based on Table 6, the first factor is organic pollution factor which presented the quantity of water polluted by EC, Alka, Cl, Ca, Mg, TH, and TDS. But Temp is the pollution factor (Table 7) due to nature. The second factor, EC, Alka, Cl, Ca, Mg, TH, TDS and Temp, are indicating a general pollution factor that mostly reflects the extent of water contaminated by reducing chemicals, while the third factor EC, Alka, Cl, Ca, Mg, TH and TDS, is representative of a pollution factor.

The first VF1 (Varimax Factor), where EC, Temp, Alka, Cl, Ca, Mg, TH and TDS contributed optimally to represent it, explained 52.46% of the total variance. Temp, Alka, Ca, Mg, and TH are heavily positively loaded, and the EC and TDS are heavily negatively loaded. This component causes pollution coming from the human (industrial and domestic, Table 7) and from natural sources (natural pollution, Table 7). In addition to explaining 33.41% of the total variation, VF2 had a positive high loading on EC, Temp, Alka, Cl, Ca, Mg, TH, and TDS. VF3 contains a positive loading of EC, Cl, Ca, Mg, and TH but a negative loading of Alka and TDS.

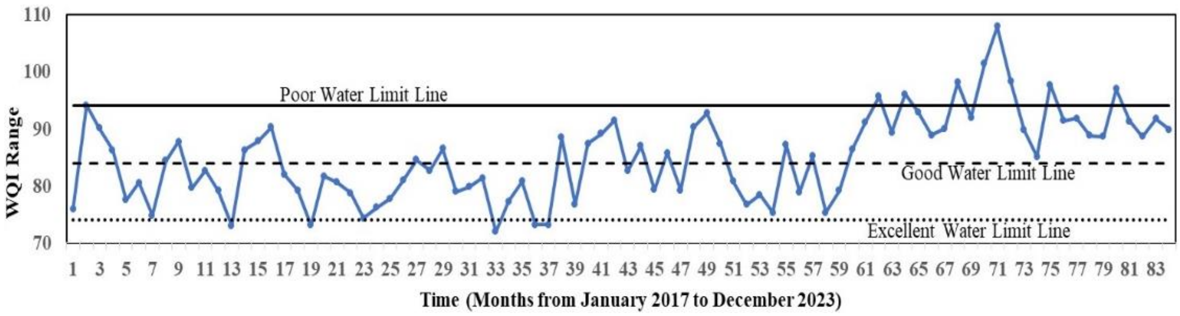

Figure 2. Temporal variation of water quality index categories in the study area during the study period (January 2017 to December 2023).

Table 2. Classification of water based on WQI value.

Month	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Jan	G	E	G	E	P	P	P
Feb	P	P	G	P	P	P	P
March	P	P	P	G	G	P	P
April	P	P	G	P	G	P	P
May	G	G	P	P	G	P	P
June	G	G	G	P	G	P	P
July	G	E	G	G	P	P	P
Aug	P	G	G	P	G	P	P
Sept	P	G	E	G	P	P	P
Oct	G	G	G	P	G	P	P
Nov	G	G	E	G	G	P	P
Dec	G	G	G	P	P	P	P

Note: - Poor (P); Good (G); Excellent (E)

Table 3. Pearson's correlation analysis.

Parameters	**pH**	**BOD**	**COD**	**EC**	**Temp**	**Do**	**Alka**	**Cl**	**Ca**	**Mg**	**Ha**	**TDS**
pH	1	-0.17	-0.13	0.33	-0.23	-0.14	-0.03	0.23	0.08	0.00	0.05	0.35
BOD	-0.17	1	**0.62**	0.07	0.45	-0.08	0.49	-0.04	0.35	0.37	0.41	0.13
COD	-0.13	**0.62**	1	0.12	0.28	-0.07	0.46	0.06	0.39	0.32	0.41	0.10
EC	0.33	0.07	0.12	1	-0.10	-0.38	-0.01	0.39	-0.27	-0.25	-0.29	**0.85**
Temp	-0.23	0.45	0.28	-0.10	1	-0.03	0.21	-0.05	0.22	0.33	0.30	-0.18
Do	-0.14	-0.08	-0.07	-0.38	-0.03	1	-0.28	-0.51	-0.31	-0.02	-0.21	-0.39
Alka	-0.03	0.49	0.46	-0.01	0.21	-0.28	1	0.07	0.57	0.40	0.56	0.21
Cl	0.23	-0.04	0.06	0.39	-0.05	-0.51	0.07	1	0.16	-0.03	0.09	0.34
Ca	0.08	0.35	0.39	-0.27	0.22	-0.31	0.57	0.16	1	0.56	0.91	-0.20
Mg	0.00	0.37	0.32	-0.25	0.33	-0.02	0.40	-0.03	0.56	1	**0.85**	-0.33
TH	0.05	0.41	0.41	-0.29	0.30	-0.21	0.56	0.09	**0.91**	**0.85**	**1**	-0.29
TDS	0.35	0.13	0.10	0.85	-0.18	-0.39	0.21	0.34	-0.20	-0.33	-0.29	1

Table 4. Total variance explained.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative percentage
1	2827.259	52.455	52.455	2827.259	52.455	52.455
2	1800.914	33.413	85.867	1800.914	33.413	85.868
3	474.788	8.809	94.677	474.788	8.809	94.677
4	164.192	3.046	97.722			
5	92.694	1.72	99.442			
6	14.986	0.278	99.72			
7	13.936	0.259	99.979			
8	0.635	0.012	99.99			
9	0.417	0.008	99.998			
10	0.063	0.001	99.999			
11	0.041	0.001	100			
12	-3×10^{-13}	-6×10^{-15}	100			

Note: Extraction Method: Principal Components

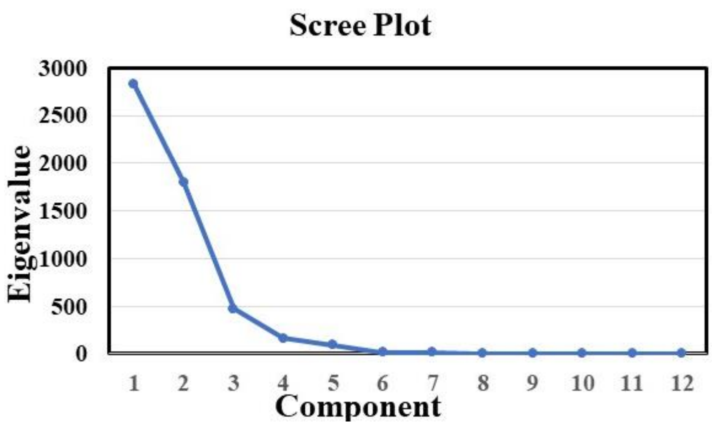

Figure 3. Scree plot diagram showed based on eigen value three PCs.

Table 5. Factorial load/Component matrix.

Parameters	Component (PCs)		
	1	2	3
pH	-0.06	0.07	0.05
BOD	0.06	0.18	-0.04
COD	0.13	0.42	-0.02
EC	-35.13	19.73	8.37
Temp	1.05	0.70	0.26
DO	0.15	-0.49	-0.05
Alka	9.31	21.88	-16.12
Cl	-1.01	1.53	0.83
Ca	14.85	13.47	2.73
Mg	10.83	7.86	5.88
TH	25.69	21.33	8.61
TDS	-22.50	15.22	-5.30

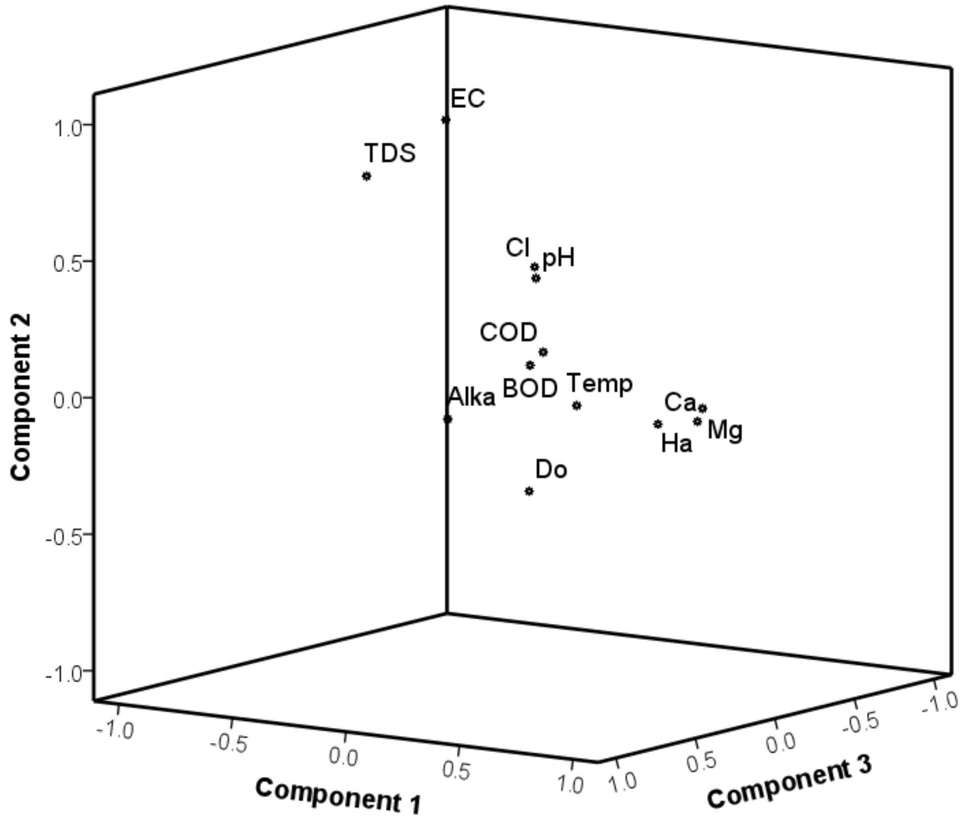

Figure 4. Component plot in rotated space.

Table 6. Rotated Component Matrix.

Parameters	Component		
	VF1	VF2	VF3
pH	-0.06	0.07	0.05
BOD	0.06	0.18	-0.04
COD	0.13	0.42	-0.02
EC	-35.13	19.73	8.37
Temp	1.05	0.70	0.26
DO	0.15	-0.49	-0.05
Alka	9.31	21.88	-16.12
Cl	-1.01	1.53	0.83
Ca	14.85	13.47	2.73
Mg	10.83	7.86	5.88
TH (Ha)	25.69	21.33	8.61
TDS	-22.50	15.22	-5.30

Note: Rotation Method: Varimax with Kaiser Normalization

Table 7. Predicted sources of pollution during study time period.

Principal Component	Typical Loadings	Predicted Sources
PC1	EC, Alka, Cl, Ca, Mg, TH, TDS	Industrial and domestic wastewaters, Soil leaching/ Soil erosion/ Agricultural runoff/rock and weathering.
	Temp	Natural pollution
PC2	EC, Alka, Cl, Ca, Mg, TH, TDS	Industrial and domestic wastewaters, Soil leaching/ Soil erosion / Agricultural runoff/rock and weathering.
	Temp	Natural pollution
PC3	EC, Alka, Cl, Ca, Mg, TH, TDS	Industrial and domestic wastewaters, Soil leaching/Soil erosion/Agricultural runoff/rock and weathering.

CONCLUSION

Based on WQI the majority of water samples fall in the category of poor quality in recent years. Pearson's correlation analysis helped evaluate the relationships among the measured physicochemical parameters. Correlation analysis showed a strong link between calcium, magnesium, and hardness, suggesting that they originate from the same source. Using factor analysis, it was determined that dissolution of minerals from rock-water interactions was the leading cause of the correlation between the variations in surface water quality. The spread of human-driven activities and the release of ion exchange phenomena within the Naini Lake water have resulted in both. PCA identified the major physicochemical variables contributing to water quality variations. Factor analysis revealed that the primary cause of the variance in Naini Lake was rock-water, which displayed dissolution of minerals from rock-water interactions. Anthropogenic activities and ion-exchange processes were identified as important contributors to water quality variation in the lake. The Nainital lake survey region's decrease is linked to numerous causes. The primary source of pollution in the water of Naini Lake is contamination from the town's framework of stormwater drains and sewer lines Solid waste is frequently deposited into waterbodies before it is able to reach its goal. The lakes, and sewage overflows are exacerbated by rainwater runoff.

ACKNOWLEDGEMENT

This project work was financially supported by the Uttarakhand Biotechnology Council (UBC), Govt. of Uttarakhand, (Sanction order no. UBC/R&D Project/ 2021/245-A, dated: 19-02-2021) and Indian Space Research Organization (ISRO, sanction order no. ISRO/RES/3/865/22-23), Govt. of, India. We are grateful to the Uttarakhand Pollution Control Board, India for providing the data at no cost.

REFERENCES

1. Zhang, X., Li, B., Xu, H., Wells, M., Tefsen, B. and Qin, B. (2019) Effect of micronutrients on algae in different regions of Taihu, a large, spatially diverse, hypereutrophic lake. *Water Research*, **151**, 500–514.

2. Shen, D., Huang, S., Zhang, Y. and Zhou, Y. (2021) The source apportionment of N and P pollution in the surface waters of lowland urban area based on EEM-PARAFAC and PCA-APCS-MLR. *Environmental Research*, **197**, 111022.

3. Dendukuri, S. C., Chandra, D. S., Raju, M. V. S. and Asadi, S. S. (2017) Estimation of Water Quality Index By Weighted Arithmetic Water Quality Index Method: A Model Study. *International Journal of Civil Engineering and Technology*. http://iaeme.com/Home/journal/IJCIET.

4. Singh, S. K., Singh, P. and Gautam, S. K. (2016) Appraisal of urban lake water quality through numerical index, multivariate statistics and earth observation data sets. *International Journal of Environmental Science and Technology*, **13**, 445–456.

5. Liu, L. and You, X. Y. (2023) Water quality assessment and contribution rates of main pollution sources in Baiyangdian Lake, Northern China. *Environmental Impact Assessment Review*, **98**, 106965.

6. Andrabi, S., Bakhtiyar, Y., Yousuf, T., Akhtar, M. and Nissar, S. (2024) Water quality assessment in relation to fish assemblage using multivariate analysis in Manasbal Lake, Kashmir. *Water Science*, **38(1)**, 92–108.

7. Dzwairo, R., Singh, S. K. and Patel, A. (2024) Soil erosion susceptibility assessment through morphometric analysis and morphotectonic

implications in Rietspruit sub-basin, South Africa. *Environment, Development and Sustainability*, 1–22.

8. Tiwari, M., Shukla, S., Rawat, K.S., Mishra, V.N., Singh, S.K., Shukla, K. (2026) Temporal Analysis of Urban Waterlogging Hotspots Using Landsat-Derived Spectral Indices: Kolkata, India. *Acta Scientiarum Polonorum Formatio Circumiectus*, **24(2)**.

9. Amin, A., Fazal, S., Mujtaba, A. and Singh, S. K. (2014) Effects of land transformation on water quality of Dal Lake, Srinagar, India. *Journal of the Indian Society of Remote Sensing*, **42**, 119–128.

10. Liu, S., Xu, J., Wang, R., Fu, X., Liu, X., Zhao, Y. and Zhang, X. (2025) Investigating the causal effects of anthropogenic factors on urban streams and lakes water quality by integrating causal inference with interpretable machine learning. *Journal of Cleaner Production*, **488**, 144626.

11. An, L., Liu, C., Fan, Z., Liao, K., Wang, W. and Wang, N. (2025) Effects of water level variations on the water quality of Huayang Lakes, China. *Journal of Geographical Sciences*, **35(1)**, 173–188.

12. Alok, S., Rawat, K.S. and Nathani P. (2025) Bridging the Gap: Application of Earth Observational Data Set and Remote Sensing Techniques for Monitoring of Water Quality. *Ecology, Environment and Conservation*, http://doi.org/10.53550/EEC.2025.v31i02.019.

13. Kumar, R., Singh, S. and Sharma, R. C. (2019) Application of WQI for assessment of water quality of high altitude lake Dodi Tal, Garhwal Himalaya, India. *Sustainable Water Resources Management*, **5(3)**, 1033–1042. https://doi.org/10.1007/s40899-018-0281-1.

14. Rawat, K.S., Mishra, A.K. and Sehgal, V.K. (2013) Identification of Geospatial Variability of Fluoride contamination in Ground Water of Mathura District, Uttar Pradesh, India. Journal of Applied and Natural Science **4(1)**, 117-122.

15. Jamion, N. A., Lee, K. E., Mokhtar, M. and Goh, T. L. (2025) Water Quality Assessment of Ex-Mining Lakes Converted Man-made Wetlands in Paya Indah Wetlands, Selangor, Malaysia. *Water, Air, & Soil Pollution*, **236(2)**, 1–23.

16. Singh, V. K., Kumar, D., Singh, S. K., Pham, Q. B., Linh, N. T. T., Mohammed, S. and Anh, D. T. (2021) Development of fuzzy analytic hierarchy process based water quality model of Upper Ganga river basin, India. *Journal of Environmental Management*, **284**, 111985.

17. Zhou, X. Q., Shen, Y. M. and Tang, J. (2025) Numerical modeling of water diversion impacts on water quality improvement in Lake Dianchi. *Environmental Modelling & Software*, 106375.

19. Singh, S. K., Srivastava, P. K., Singh, D., Han, D., Gautam, S. K. and Pandey, A. C. (2015) 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. *Environmental Geochemistry and Health*, **37**, 157–180.

20. Akiner, M. E., Chauhan, P. and Singh, S. K. (2024). Evaluation of surface water quality in the Betwa River Basin through the water quality index model and multivariate statistical techniques. *Environmental Science and Pollution Research*, **31(12)**, 18871–18886.

20. Hsu, C. C. and Sandford, B. A. (2007) The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research, and Evaluation*, **12**, 10. https://doi.org/10.7275/pdz9-th90.

21. Hassan, A., Samy, G., Hegazy, M., Balah, A. and Fathy, S. (2024) Statistical analysis for water quality data using ANOVA (Case study–Lake Burullus influent drains). *Ain Shams Engineering Journal*, **15(4)**, 102652.

22. Adejuwon, E. O., Ogwueleka, T. C., Ogungbemi, E. O., Prabhu, R., Rendon-Nava, A. and Yates, K. (2025) Assessment of Surface Water Quality Using Chemometric Tools: A Case Study of Jabi Lake, Abuja, Nigeria. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 1–24.

23. Iyer, S., Kaushik, S. and Nandal, P. (2023) Water Quality Prediction Using Machine Learning. *MR International Journal of Engineering and Technology*, **10(1)**, 59–62. https://doi.org/10.58864/mrijet.2023.10.1.8.

24. Yang, W., Zhao, Y., Wang, D., Wu, H., Lin, A. and He, L. (2020) Using principal components analysis and IDW interpolation to determine spatial and temporal changes of Surfacewater quality of Xin'Anjiang river in Huangshan, China. *International Journal of Environmental Research and Public Health*, **17(8)**. https://doi.org/10.3390/ijerph17082942.

25. Brown, R. M. and McClelland, N. I. (1970) Deininger, R.A.; Tozer, R. G. A Water Quality Index: Do We Dare? *Water Sew. Works*, **117**, 339–343.

26. Alam, A. and Singh, A., (2023) Statistical approach to evaluate groundwater contamination for drinking and irrigation suitability.

Groundwater for Sustainable Development, 9. https://doi.org/10.1016/j.gsd.2019.100251.

27. Akhtar, N., Ishak, M. I. S., Ahmad, M. I., Umar, K., Md Yusuff, M. S., Anees, M. T., Qadir, A. and Almanasir, Y. K. A. (2021) Modification of the water quality index (Wqi) process for simple calculation using the multi-criteria decision-making (mcdm) method: A review. *Water (Switzerland)*, **13(7)**. https://doi.org/10.3390/w13070905.

28. Rawat, K.S., Singh, S.K. and Gautam, S.K. (2018) Assessment of groundwater quality for irrigation use: a peninsular case study. *Applied water science*, 8(233), doi.org/ 10.1007/s13201-018-0866-8.

29. BIS (2012) https://cpcb.nic.in/wqm/BIS_ Drinking_ Water_Specification.pdf.

30. Rawat, K. S., Singh, S. K., Jacintha, T. G. A., Nemčić-Jurec, J. and Tripathi, V. K. (2017) Appraisal of long -term groundwater quality of peninsular India using water quality index and fractal dimension. *In Journal of Earth System Science, Indian Academy of Sciences*, **126(8)**. https://doi.org/10.1007/s12040-017-0895-y.

31. Shah, S., Sen, S. and Sahoo, D. (2024) State of Indian Northwestern Himalayan lakes under human and climate impacts: A review. *Ecological Indicators*, **160**, 111858.

32. Upadhyay, R., Bakshi, S., Bisht, H. and Kotlia, B. (2021) Assessment of Water Quality of Nainital Lake and surrounding Springs, using Water Quality Index (WQI) and Heavy Metal Pollution Index (HPI). https://doi.org/10.31870/ ESI.14.1.2021.3

33. Rawat, K. S. and Singh, S. K. (2024) Monitoring water spread and aquatic vegetation using earth observational data for Nani-High Altitude Lake (N-HAL) of Uttarakhand State, India. *Journal of Engineering Research*, **12(1)**, 64–74.

34. Dutt Maindoli, S., Maindoli, S. D., Pant, M., Singh, P. and Gupta, P. K. (2019) Physico-Chemical and Biological Responses of Aeration for Eutrophication Control in Nainital Lake, Uttarakhand. *Pollution Research*, **38(1)**, 145–150. https://www.research gate.net/publication/ 350810378.

35. Rawat, K.S. and Singh S.K. (2018) Water Quality Indices and GIS-based evaluation of a decadal groundwater quality. Geology, Ecology, and Landscapes, **2(4)**, DOI: 10.1080/24749508.2018.1452462.

36. Shrivastava, N. G. (2021) Assessment of lake water quality by using trophic state index indicators parameters for ecological lake restoration-a case study of Naukuchiatal, Kumoun Region, Uttrakhand, India. *International Journal of Multidisciplinary and Current*, **3(2)**, 101–122.

37. Brown, D. M., Cameron, B. A., Meyer, S. S. and Umber, J. J. (1991) The effects of water hardness level on washing quality using commercial laundry detergents. *Journal of Consumer Studies & Home Economics*, **15(3)**, 215–222.

38. Jacintha, T.G.A., Rawat, K.S., Mishra, A., and Singh, S.K. (2016) Hydrogeochemical Characterization of Groundwater of Penninsular Indian Region using Multivariate Statistical Techniques. Applied Water Science, **7,** 3001-3013 doi/10.1007/ s13201-016-0400-9