INTERNATIONAL JOURNAL OF HIGHER EDUCATION AND RESEARCH
IJHER, Vol., 10 (2), 2020, 210-228.
www.ijher.com

IJHER
(ISSN 2277 260X)

CODEN:IJHER ORIGINAL ARTICLE

Application of Principal Component Analysis (PCA) for Water Quality Evaluation and Management of Lakes in Valsad District

Ayantika Das* and Jigna Desai
Department of Bioscience
Veer Narmad South Gujarat University, Surat, Gujarat, India
*Corresponding Author: ayantika.d2016@gmail.com

ABSTRACT

The present research primarily conceives a model for recognition of the most indicative water quality parameters of five perennial lakes of Valsad district for long term recurrent monitoring and management. Secondarily we will also relate water quality parameters to lake biodiversity. Sixteen water quality parameters considered for our present study were pH, electrical conductivity, color, turbidity, odor, BOD, COD, TDS, total hardness, residual free chlorine, chloride, phenolic compounds, total ammonia, sulfate, sulfite, fluoride, and metals. We take two samples each month from each lake for water collection for a period of six months. Using Principle Component Analysis (PCA), we reduced the number of variables by clustering and reducing variables based on the level of significance. From this study, we will predict the most prototypical water quality parameters that will be considered for future water quality monitoring and to formulate plans on how to improve the pollution scenario. This will, in turn, guide us to formulate a predictive relationship between parameters with floral diversity of plants in lake water and around the lakes, in an effect and cause relationship. This study will act as a reference for future environmentalists, ecologist and government bodies in water quality monitoring and management.
***Keywords**: Water quality of tropical lakes; Principle component analysis; Floristic Diversity*

[Das, A. and Desai, J. Applcation of Principal Component Analysis (PCA) for Water Quality Evaluation and Management of Lakes in Vasad District.. *International Journal of Higher Education and Research* 2020;10(2):210-228]. ISSN 2277 260X (online). http://www.ijher.com. doi:10.7755/ijher170720.01.

INTRODUCTION

Organisms live within their particular niche, which has conditions favorable for the growth and proliferation of the organism. There is a very delicate balance between favorable and unfavorable conditions. Both environmental disasters, as well as anthropogenic pressures, disturb the favorable conditions. Over the last decade due to rapid industrialization and urbanization, environmental imbalances have occurred over many places. Accumulation of pollutants alters the structure and function of the ecosystem (Woodwell, 1970). Plants are ecosystem engineers as they reduce pollution by alteration and sequestering of pollutants to some extent.

When there is any change in the normal composition of water then there is a modification in the state of equilibrium of the aquatic medium. Anthropogenic pressures have adversely affected air, water as well as soil composition. Water both surface, as well as groundwater, polluted these days. Allochthonous inputs alter the composition of lakes and rivers, which in turn pollute groundwater by seepage. The industrial effluent, municipal, and domestic sewage and nutrient load influx in the form of runoff from agricultural land and orchard plantations adversely alter the physiochemical and biological compositions of a water body. Transport of nitrogen and phosphorus from soil to water, though less recognized the substantial process (Anderson and Polis, 1999). Air pollutants precipitate on land and water and change their chemical composition (Mudd, 2012). Acid pollution from mine drainage and acidic rain affects the microbial flora, fauna, and fishes (Heath, 2018).

Self - purification process of lake water occurring naturally includes – (1) Dilution – which occurs during monsoon. (2) Sedimentation – will take time to occur naturally. Sedimentation will reduce the TDS to some extent. (3) Temperature – Temperature plays an important role in the self-purification of surface waters. A lower temperature will increases the DO as low-temperature decreases the decomposition rate. So practically, there will be organic biodegradable nutrient in water, but it will slowly decompose at low temperatures. Such a lake will cause algal bloom on the rise of temperature.

Physiochemical, biological and microbial parameters, condition the biodiversity of a lake and its surrounding ecotone. We have done fieldwork to identify the species richness of the lake ecosystems and its variations. Water quality analysis of various physiochemical parameters that will help us to get a picture of how it affects the biodiversity in and around a water body. Pearson has done such work earlier in 1980. Several classification indexes have used to determine the biodiversity of an ecosystem and the pollution level by Woodivis, 1964; Verneaux, 2003: Kelly, 1998; Seele et al in 2000. Carlson et al in 2006 have used a physiochemical approach.

MATERIAL AND METHODS

Descriptions of the study area – of five perennial lakes of Valsad district, two are in urban (Lake Number 1 and 2) and rest three lakes, (Lake 3, 4, 5) are in rural agriculturally fertile regions.

Urban area lakes surrounding by human habitation and receives untreated domestic sewage animal bathing. Rural areas surrounded by orchards of mango and sapodilla receives runoff from orchard lands also used for fishery purpose also. We studied the location of the perennial lakes of the Valsad district that indicated in Figure 1.

Pollution sources of the lakes can be categorized into five groups – All the lakes are facing anthropogenic pressures (urban and agricultural only, not industrial), which directly affects its biodiversity. These lakes have many non-point sources of pollution such as domestic discharge found in lakes 1and 2, agricultural runoff from fields and orchards surrounding the lakes 3, 4 and 5. Direct solid domestic waste dumping - in the lake, which was a very prominent pollutant source in lakes 1, 2 and 5, bathing of animals in lakes 2, 3, 4 and 5. We can get an idea of the dimensions of the perennial lakes from table 1. We observe that except Lake 1 and 2 all other lakes are relatively large lakes. Table 1 indicates the area.

Procedure for water collection

We collected water samples from the five perennial lakes, every 15 days from five sampling stations, one from each lake. We collected water samples in glass bottles. We have done sampling for six months – two pre-monsoon months, two monsoon months and two post-monsoon months. We collected water samples from fifty-centimeter depth, as surface water is very unstable to parameter change. We collected two liters water for analysis and kept in a cool dark place for further analysis.

Chemical Analysis process

Water samples were analyzed for physicochemical parameters such as pH, biological oxygen demand (BOD), chemical oxygen demand (COD), dissolved oxygen (DO), electrical conductivity (EC), color, odor, turbidity, total dissolved solids (TDS), hardness as $CaCO_3$, chloride, residual chlorine, phenolic compounds, sulfide, sulfate, total ammoniacal nitrogen, metals. For the analysis of lacustrine water samples, we have followed the IS 3025 – 1987 technique for all the parameters. We performed all the analyses in triplicate. Various parameters we considered are as follows:-

pH: - The pH value is dependent upon factors such as - (1) the Presence of dissolved ions (hydroxides, carbonates, and bicarbonates of calcium and magnesium are the sources of alkalinity). Detergents cause hydroxide alkalinity. (2)Fertilizers cause carbonate alkalinity and insecticides cause bicarbonate alkalinity (3) Presence of algal population (pH is highest during the day due to photosynthesis and low at night due to algal respiration and respiration of other aquatic plants and animals). (4) The presence of other substances like humic acid present in soil form salts increasing alkalinity of the soil. At low pH, metals are dissolved in soil solution easily making them biological available. Some of these metals are toxic (such as cyanides and sulfides). Ammonia dissolves at intermediate pH. The pH and DO will below if there is excess macrophytes invasion in the lacustrine ecosystem. The pH range from 6.6 to 7 usually safe for all plants.

Electrical conductivity – Electrical conductivity of water directly related to the concentration of dissolved ions in the water. If the EC of water suddenly increases then we conclude that the source of dissolved ions is in the vicinity. It is a quick way to determine a water quality problem. Electrical conductivity is the degree to which water conducts electricity. We determine it by recording resistance in mm hos/cm or dS/m after passing an electric current through the water sample and. Electric conductivity is used to estimate the TDS of water using the equation TDS(ppm or mg/l) = EC(mmhos/cm or dS/m) $\times$ 640.

Color - APHA, Pt-Co, and Hazen are the three most commonly used color scales. The color of filtered water ranges from 1 to 25 Hazen units, for unfiltered water it ranges between 1 - 85 Hazen units. Color mainly contributed by effluent from paper and pulp industry effluent, discharges from cloth and dye industries. Color is due to the metallic ion such as iron and manganese. Fresh sanitary sewage imparts grey color to the water, which darkens with time.

Odor - Odor in water is mainly due to the decomposition of organic matter. Fresh sanitary sewage is odorless but after anaerobic decomposition of proteins and other organic matter rich in nitrogen, sulfur and potassium water becomes foul-smelling. The presence of foul-smelling water thus may lead to the low BOD due to high organic matter content. Such water will have low aquatic life.

Turbidity- The turbidity of water is due to suspended colloidal particles such as clay and silt, which reduces the light penetrating capacity or vegetable fibers and microbes indicating anaerobic conditions, and by soaps, detergents, dyes, and emulsifying agents.

Dissolved oxygen (DO)- Dissolved oxygen is low in the following cases: – (1) if the bacterial count is high; (2) or if the temperature is high; (3) or if it is a cloudy day when the algae respire but cannot photosynthesize; (4) or when the submerged and rooted plants die and decay in water. Water hyacinth initially increases the DO, but after full colonization of the lake, it decreases the DO in the end; (4) Dissolved solids provide adsorption surface for chemical and biological pollutants, which in turn reduce the DO. Both oxygen and nitrogen are poorly soluble in water and dissolved salts further reduce the solubility. DO is measured in milligram per liter (mg/l) or parts per million (ppm). Dissolved oxygen of freshwater may range from 0 – 18 mg/l. DO of 5 – 10 mg/l is good for fishes, but if DO of water is less than 5 then the fishes are stressed.

Biological oxygen demand (BOD) - Biological Oxygen Demand means oxygen consumed by microbes and organisms in the water**.** Amount of BOD in a water body depends upon factors : - (1) Temperature (less oxygen dissolve in high temperature); (2) Quantity of organic biodegradable sediment; (3) Amount of oxygen taken out from the system by respiration and decaying of organisms; (4) Amount of oxygen put back into the system by photosynthesis; (5) Pollutant load from a point and non-point sources. BOD of unpolluted water is below 1 mg/l; BOD of moderately polluted water ranges between 2 to 8 mg/l; BOD of municipal sewage is 20 mg/l or lower and BOD of untreated sewage is 200 to 6000 mg/l.

Chemical oxygen demand (COD) - Chemical oxygen demand does not differentiate between organic and inorganic substances that microbes can be oxidizing in water. If the BOD/COD ratio is more than 0.6, then the wastewater biodegradable and we can treat such water biologically; if BOD/COD ratio is between 0.3 to 0.6 mg/l, then seeding is required to treat the water biologically if BOD/COD ratio is less than 0.3 then we cannot treat water biologically.

Total dissolved solids (TDS) - TDS of total dissolved solids measures various salts includes several cations and anions of calcium, magnesium, sodium, carbonate, bicarbonate, chloride, sulfate dissolved in water that cannot be removed by traditional methods of filtration but only through the use of a membrane, by reverse osmosis and distillation. TDS measured by EC meter. Total dissolved solids may be also due to organic matters like plant fiber, biological solids such as bacteria and algae and inorganic matters such as clay silt dissolved in water. High TDS can cause a 10 % reduction in optimum crop production. It is different from general hardness is a measure of only calcium and magnesium salts dissolved in water. A water sample can have high TDS but considered as soft water as they have a low concentration of calcium and magnesium. Under the condition of high TDS, the aquatic plants have to use much more energy to absorb nutrients from the soil. Under high TDS, the plants show symptoms of burned leaflets but some plants may visually appear to be healthy but these plants also have to use extra energy to maintain the salt balance inside the plant. High TDS in soil and water affects the seed germination also at a crucial period. High TDS also produces odor, color and imparts taste to water.

Total hardness as $CaCO_3$ - Hardness in water is due to divalent metallic cations like Ca++, Mg++, Fe++, Mn++, Sr+, and anions such as bicarbonates, chlorides, and sulfates (of Ca and Mg). The hardness of water is of two types : - (1) Temporary hardness caused by Ca (HCO) 3, Mg (HCO) 3 or both. This is carbonate hardness. (2) Permanent hardness caused by mostly sulfates of calcium and magnesium. This is non-carbonate hardness cannot be easily removed. According to WHO Guidelines, values of total hardness of the water in mg/L as CaCO3 in - 0-30 is soft water; 30- 60 is moderately soft water; 60- 120 is moderately hard water; 120-180 is hard water; above 180 is very hard water. Expected total hardness of freshwater ranges from 15 – 375 mg/L as CaCo3; of which calcium hardness ranges from 10 – 250 mg/L and magnesium are 5 – 125 mg/L.

Chloride - Chlorides a non-essential micronutrient that occurs naturally in the earth’s crust, often associated with salinity damage and toxicity. Chloride concentrations of 1 – 100 ppm (parts per million) are normal in freshwater. Chloride ions come in solution in water in underground aquifers. Natural spikes in the chloride concentration can occur during the summer low flow period when the evaporation rate is more than the precipitation rate. Presently there has been much increase in chloride concentration due to many anthropogenic activities like sewage contamination, domestic wastewater, use of water softener.

Residual chlorine – Residual chlorine is the low amount of chlorine remaining in the water after the initial application. Chlorination of general water done in water treatment plants to reduce color.

Phenolic compounds – Phenolic compounds are soluble or insoluble compounds of diverse origin such as natural, agricultural, domestic as well as industrial. It naturally produced by the decay of organic matter. Combustion of wood and automobile exhaust also liberates phenols. They considered hazardous in stream and river water due to its toxicity to plants and fishes.

Fluoride - Fluoride is an ingredient of igneous and sedimentary rocks so found in large quantities in groundwater. When plants accumulate fluoride at high concentration, plant starts showing symptoms of injury. The sensitivity of immature parts is much more than mature parts.

Sulfide - Concentration of sulfide in a water body may be a result of both natural as well as anthropogenic causes. Soluble sulfide such as H_2S is toxic to plant roots (Tanaka et al, 1968). Heavy metals if present even in trace amounts will bind with sulfides to form toxic complexes (Wang et al, 1999).

Sulfate - Sulfate is the second major anion after bicarbonate in water and causes permanent hardness, often found in municipal and industrial discharges. Nutrients like nitrates, sulfates, phosphates and iron act as a signal molecule that alters the rate of cell division and differentiation (Bucio et al, 2003). Root hair formation, lateral root formation, primary growth of the roots is dependent on these nutrient concentrations. They found in agricultural runoff as well. Bacteria reduce sulfates and form H_2S gas. Sulfate is not toxic to plants and animals up to a normal concentration of 500 – 750 mg/l; a higher concentration of sulfate causes complex formation and precipitation of metals.

Total ammonia nitrogen - Ammonia (NH_3) absorbed by plants or converted to nitrite and nitrate. Ammonia reacts with water to form a weak base. There are two species of ammonia namely NH_3^+ and NH_4^+, collectively called Total Ammonia Nitrogen. NH_3^+ is toxic at a concentration of 0.53-22.8 mg/l. At higher pH, the toxicity of NH_3^+ increases. Excess ammonia in soil and irrigation water is detrimental for plant growth and development, accumulated in the plant body causing alteration of metabolic process.

Metals - All metals are soluble in water and toxic even in low concentration as an accumulative poison. Heavy metal exposure can lead to growth inhibition and even death in sensitive plants. The extent of toxicity of heavy metals depends upon the time of exposure, pH of the soil so that we can know whether the heavy metal is in bioavailable form or not (Luther et al, 1999). Al, Pb, and Cr found in trace amounts are very toxic, which will in due course of time accumulate in the plant body.

Statistical analysis

Generally, to analyze surface water quality, we use multivariate methods such as factor analysis, cluster analysis, discriminate analysis and principal component analysis. All these techniques used to reduce the number of parameters by reducing and grouping parameters according to their correlation and level of significance. We used Principal Component Analysis to analyze the result of the analysis of physicochemical parameters of water (Arain et al, 2009). PCA alone is not sufficient for prediction and analysis of data so PCA used along with KMO, Bartlett test. Mean, variance, cumulative variance are first determined. As we have a large number of variables, we are using two principal components for better clarity and less complication in understanding. The factor loading of the rotated component graph gives more clarity than the normal version. PCA performed by SPSS.

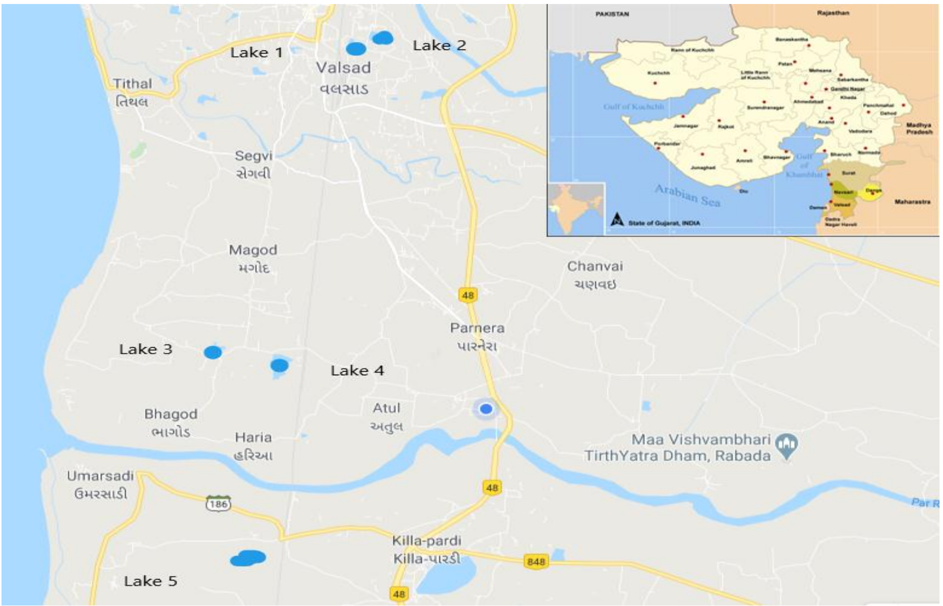


Figure 1. Water sampling locations of Lakes – 1 (Railway Colony), 2 (Mota Talav), 3 (Atar Talav), 4 (Dwivedi Talav), 5 (Highway Talav)

RESULT AND DISCUSSION

Table1. Area of the lake

Lake No	Lake name	Area m^2
1	Railway Colony	8,691
2	Mota Talav	39,887
3	Atar Talav	51,870
4	Dwivedi Talav	171,976
5	Highway Talav	44,033

Vegetation in and around the immediate surroundings of the lakes depicted in table 2. The areas surrounding lake 1 and 2 has few tree vegetation, but abundant in terrestrial weeds and some aquatic weed. Surrounded on three sides by orchards, Lake 3, 4 and 5, also has abundant aquatic and terrestrial weeds. All the lakes weeds became abundant during monsoons. Table 2 depicts the various aquatic plants of the five lakes observed during various seasons. Submerged and rooted vegetation also trap all the nutrients and check algal bloom. All the lakes had water hyacinth during the wet seasons, probably due to agricultural runoff. Runoff water during rainy seasons brings many nutrients as well as minerals to the lake that is present in excess may create a problem with the metabolism of plants and animals in the water.

Table 2. Lake Macrophytes

Name	Lake 1	Lake 2	Lake 3	Lake 4	Lake 5
Eichhornia crassipes	√	√	√	√	√
Nymphea alba			√	√	√
Nymphea pubescence			√	√	√
Nymphea carulea			√	√	√
Nelumbo nucifera			√	√	√
Ceratophyllum sp			√	√	√
Chara sp			√	√	√
Hydrilla sp			√	√	√
Oxalis sp	√	√	√	√	√
Colocasia sp	√	√	√	√	√

Table 3. Descriptive Statistic of Physiochemical parameters

Sl no	Parameter	Unit	Lake 1	Lake 2	Lake 3	Lake 4	Lake 5
1	pH	-	7.23	7.18	7.20	6.95	7.51
2	Electrical conductivity (EC)	us/cm	467	545	620	555	517
3	Color	Hzn/unit	4.7	5.5	5.5	4.3	5.2
4	Dissolved Oxygen (DO)	mg/l	6.9	6.3	6.1	6.3	6.3
5	Biological Oxygen Demand (BOD)	mg/l	14	20	24	19	20
6	Chemical Oxygen Demand (COD)	mg/l	44	59	65	54	57
7	Total Dissolved Solute (TDS)	mg/l	260	303	344	308	287
8	Total Hardness (as $CaCO_3$)	mg/l	164	196	203	212	183
9	Chloride (as Cl)	mg/l	74	68	87	66	77
10	Residual free chlorine	mg/l	0.14	0.35	0.35	0.11	0.45
11	Fluoride	mg/l	0.40	0.50	1.07	0.68	0.64
12	Sulphate (SO_4)	mg/l	17.9	34.5	42.6	24.1	27.4
13	Total Ammoniacal Nitrogen	mg/l	1.9	3.7	6.0	5.5	7.4
14	Zinc	mg/l	1.2	1.0	1.7	1.3	0.7
15	Iron	mg/l	0.21	0.36	0.65	0.44	0.32
16	Boron	mg/l	0.02	0.18	0.29	0.18	0.14

Table 4. KMO and Bartlett's Test data

KMO Measure of Sampling Adequacy	0.72
Approx. Chi-Square	830.80
Bartlett's Test of Sphericity	120
Significance	0.00

The average pH of all the lakes is near to normal pH except Lake 5. The pH all lakes were slightly acidic during dry seasons of pre and post-monsoon, which changed to normal during monsoon due to dilution. As the lakes had slightly acidic to normal pH it can be concluded that there is little fertilizer pollution or eutrophication or detergent pollution or toxic metal pollution or algal bloom or excess macrophytic invasion. The high EC is mostly inorganic is naturally high in high-temperature tropical water bodies. EC is highest in lake numbers 3 and 4 probably due to little agricultural runoff from the surrounding plantation area. The color of the lakes ranged from 4 - 5 Hazen units, which is the color of natural water and fit for drinking, agricultural purposes. Our lakes 3, 4 and 5 had much-submerged hydrophytes, the decay of which contributes to the little odor that we find. Turbidity increases during monsoon due to an increase in sediment load from the surrounding areas. Negligible turbidity year-round found due to detergent and soaps used in bathing. The dissolved oxygen level of the lakes ranged from 6 to 7 mg/l. The BOD level of all the lakes ranged from 14 to 24 mg/l indicating municipal sewage. The COD of all the lakes varied from 20 to 65 mg/l, which shows that there are inorganic wastes in the lakes that consume oxygen. BOD/COD ratio is conducive to the growth of plants. The TDS of various lakes varied from 260 to 344mg/l. The hardness of lake waters varied from 164 to 204 mg/l indicated hard and very hard water. The chloride content of lakes ranged from 66 to 87 ppm, which is quite normal. The residual chlorine content of various lakes was below 1 ppm so it is quite conducive for plant growth. Phenolic compounds were absent in all the lakes which is a good thing for plant growth and development. The fluoride content of water of all the lakes was in around 1mg/L, which is permissible. The fluoride content of water of all the lakes was in around 1mg/L, which is permissible. Sulfide was absent in all the lakes. The maximum sulfate concentration of the lake waters was 67. 5 mg/l, which is conducive to aquatic plant life. The range of total ammonia as nitrogen is less than 1 to 16.68 mg/l. However, as the pH of water is slightly acidic, the total ammonia concentration will not be dangerous for the plants of water and surrounding. Metals such as Fe, Zn, and B considered essential nutrients of plants and found within the tolerable limit of plants so not a deterrent for plant growth. Heavy metals such as Cr, Pb, and Al were, found in lakes in trace amount, which would have been toxic if sulfide was present in the lakes. Heavy metals are accumulative poisons and alter the metabolism of all living organisms in the course of time. The mean value of the physiochemical parameters of our water quality study depicted in Table 3.

Interpretation using PCA

PCA is a non- parametric tool that alone cannot predict the underlying statistical distribution of data (Vega et al, 1998; Lambrakis et al, 2006; Kalin et al, 2012). KMO (Kaiser – Meyer – Olkin) and Bartlett test help us predict whether most of the variables normalized or not and whether the data we are having is fit for factor analysis or not. Only normalized distributions are suitable for principal component analysis. KMO measures sampling adequacies or in other words variance caused by underlying factors. If the variance is high, close to 1 then we can use the variables for factor analysis. If KMO is less than 0.5 then factor analysis is not advisable. On the other hand, the Bartlett test indicates whether the correlation matrix of variables has an identity matrix. Bartlett's test of Sphericity indicates the degree of variance among the variables. Significance value if less than 0.05 indicates that the relationship between variables is quite significant. In Bartlett's Test of Sphericity, if the significance level is less than 0.05 then it rejects the null hypothesis and accepts the alternative hypothesis and the relationship amongst variables is significant. Covariance and correlation matrix coincide in the case of normalized data. The KMO and Bartlett's test value of our parameters shown in table 4. Here KMO sampling adequacy measure is close to 1 and our significance level is 0.00, so relationships amongst our variables are significant.

Table 5. Eigenvalue and Variance

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	10.950	68.436	68.436
2	1.506	9.411	77.848
3	1.178	7.361	85.209
4	0.703	4.391	89.600
5	0.426	2.665	92.265
6	0.327	2.044	94.308
7	0.280	1.752	96.061
8	0.169	1.058	97.119
9	0.146	0.911	98.029

10	0.122	0.760	98.790
11	0.084	0.527	99.316
12	0.065	0.406	99.722
13	0.040	0.249	99.971
14	0.004	0.023	99.994
15	0.001	0.006	100.000
16	3.587E-006	2.242E-005	100.000

Table 6. Rotated component Matrix

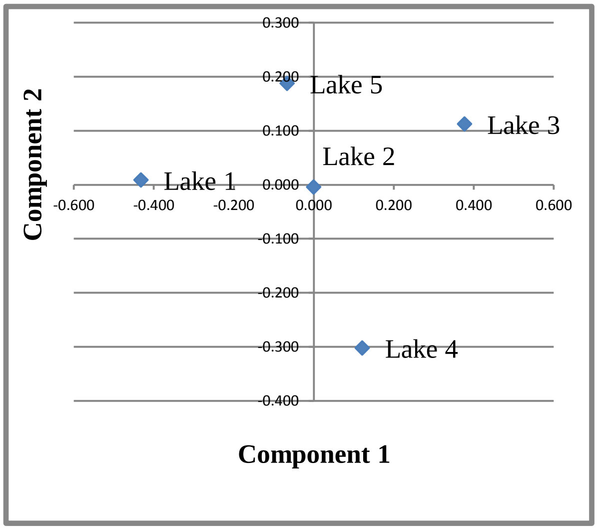
Parameter	Component	
	1	2
BOD	0.929	0.211
COD	0.909	0.228
Sulfate	0.899	0.063
Electrical Conductivity	0.880	0.413
TDS	0.879	0.413
Fluoride	0.877	0.268
Chloride	0.860	0.312
Total Ammonia Nitrogen	0.854	0.231
Boron	0.845	0.086
Dissolved Oxygen	-0.810	-0.278
Total Hardness	0.774	0.538
Iron	0.567	0.548
Color	0.406	0.767
pH	-0.335	0.703
Residual free Chlorine	0.560	0.621
Zinc	0.400	0.570

There is a significant correlation between pH and residual free chlorine. EC has a significant relation with color, DO, BOD, COD, TDS, total hardness, chloride, residual free chlorine, fluoride, sulfate, total ammonia, zinc, iron, and boron. Likewise, color has high significance with all other parameters except pH. Similarly, DO, BOD, COD, TDS, total hardness and chloride have high significance with all other variables except pH.? Iron, boron, and zinc have a high correlation with all other parameters except pH. In the same way, fluoride, sulfate and total ammonia has high significance value with all other parameters except pH and so on.

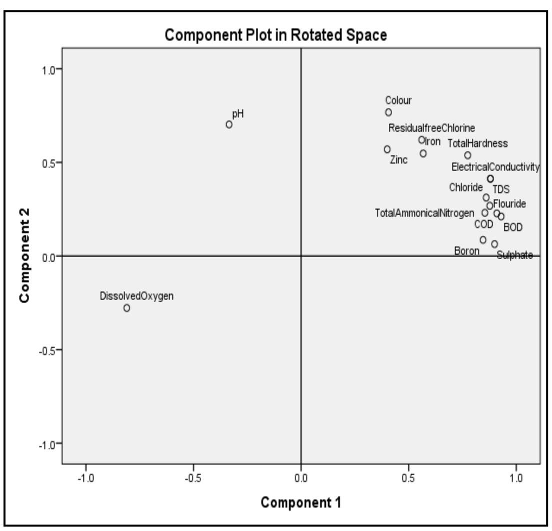
The Eigenvalues, variance percentage, cumulative variance percentage, extracted sums and rotated values of these extracted sums shown in the table, 5. From the above table, we observe that sixteen parameters grouped into few components. The component one, two and three are responsible for maximum variances. Here we have taken into consideration only those variables whose Eigenvalue is more than one as suggested by Kaiser in 1958 and Liu et al in 2007. Component 1 is responsible for 67 % of score loadings. Component 2 is responsible for 9 %. Component one and two together are responsible for 77 % of the total variance and if we include the third component together then they are responsible for 84 % of the total variance. We are going to simplify interpretation by data reduction and will be using the major two components that are the first two major components.

We are using Rotated Component Matrix for a better understanding of the position of various variables in the two major components that we are using. By rotating the factors, we are merely changing the angle of variables for better understanding. Many types of rotation of the component factors present, but popularly used are orthogonal rotation and oblique rotation. For orthogonal rotation, we have used the Varimax technique proposed by Kaiser in 1958. The Rotated Component Matrix table of the variables as shown in Table 6 grouped into two components. This table helps us to better clarify and place the various variables into the two given components. Component 1 has high loadings of parameters such as chemical oxygen demand (COD), sulfate, BOD, boron, fluoride, total ammonia nitrogen, chloride, dissolved oxygen (DO) and hardness - organic constituents. Parameters such as color, electrical conductivity, zinc, residual chlorine, iron, and pH are included in component 2 - inorganic substances. Component one loadings are mostly due to agricultural runoff from orchards. Component 2 is due to domestic sewage.

We observe the component plot of various parameter variables in the rotated space in figure 2 (B) for more clarity. In the component plot, we have placed Component 1 along the X-axis and Component 2 along Y-axis. We have observed that many variables that we can correlate grouped into two components. However, DO and pH not correlated to any of these clusters. We have plotted the lakes in a component plot in figure 2 (A) using a two-factor model to can visualize which component factors are dominant in which lakes and to which extent. Component 1 that is organic component is dominant in lake number 1, 3 and 4, while component 2 that is an inorganic factor is dominant in lakes 2 and 5.

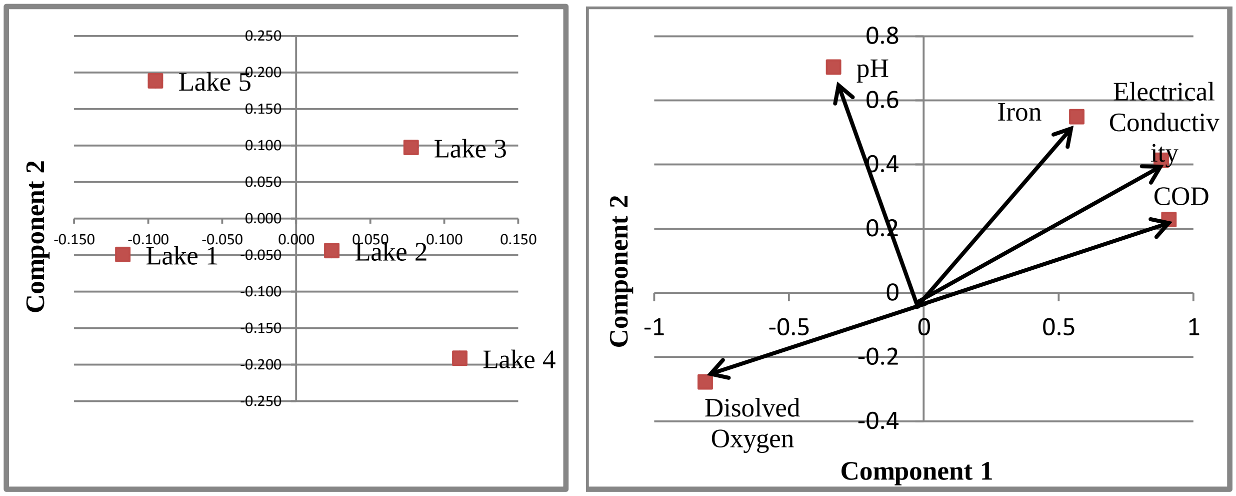


We have repotted the five lakes in a component plot in the figure, 3 (A), using a reduced number of five main parameters. Here we observe that by taking five representative parameters and plotting the various lakes, the position of the various lakes in the component plot has changed a bit. This change in position is due to data reduction. In the figure, 3 (B) we have taken only five variables that are representatives of the entire dataset and plotted parameters- DO, pH, iron, EC and COD of which pH and DO represent themselves only and DO, EC and COD represent one cluster.



A B

Fig. 2. A – Score of the 5 lakes on the axis of principal components 1 and 2, obtained by the sixteen variables, B - Scores of the sixteen variables.



A B

Fig 3. A – Score of the 5 lakes on the axis of principal components 1 and 2, obtained by the five selected variables, B - Scores of the five variables.

CONCLUSION

From our study, of various physical and chemical parameters of water quality of five lakes we conclude that anthropogenic activities happening around the lake ecotone are affecting water quality. Using PCA, we can get a precise description of the water quality of five lakes on a long-term basis by reduction of the number of parameters to just five representative parameters: - pH, electrical conductivity, iron, COD and DO. The use of a reduced set of parameters monitoring becomes cost-effective and less time-consuming. A high concentration of COD and a low concentration of DO in our lakes indicates organic pollutants, which is due to domestic and agricultural pollution. While a moderate concentration of iron and high, the concentration of EC indicates the inorganic component. Thus we see observe codominance of both organic and inorganic components in the lake water. This study using, Principle Component Analysis utilized in environmental engineering and management of water bodies in tropical climates as well as in groundwater analysis.

REFERENCES

Anderson, W. B., & Polis, G. A. (1999). Nutrient fluxes from water to land: seabirds affect plant nutrient status on the Gulf of California islands. *Oecologia*, *118*(3), 324-332.

Bilotta, G. S., Grove, M., & Mudd, S. M. (2012). Assessing the significance of soil erosion. *Transactions of the Institute of British Geographers*, *37*(3), 342-345.

Chelsea Nagy, R., Graeme Lockaby, B., Kalin, L., & Anderson, C. (2012). Effects of urbanization on stream hydrology and water quality: the Florida Gulf Coast. *Hydrological Processes*, *26*(13), 2019-2030.

Heath, A. G. (2018). *Water pollution and fish physiology*. CRC press.

Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., ... & Shah, A. Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicology and environmental safety*, *72*(2), 301-309.

Kelly, M. G., Cazaubon, A., Coring, E., Dell'Uomo, A., Ector, L., Goldsmith, B., ... & Kwandrans, J. (1998). Recommendations for the routine sampling of diatoms for water quality assessments in Europe. *Journal of applied Phycology*, *10*(2), 215.

Papatheodorou, G., Demopoulou, G., & Lambrakis, N. (2006). A long-term study of temporal hydrochemical data in a shallow lake using multivariate statistical techniques. *Ecological Modelling*, *193*(3-4), 759-776.

Schimpff, W. K., & Pearson, H. E. (1980). *Water quality effects related to blending waters in distribution systems* (Vol. 1). Municipal Environmental Research Laboratory, Office of Research and Development, US Environmental Protection Agency.

Shanyengana, E. S., Henschel, J. R., Seely, M. K., & Sanderson, R. D. (2002). Exploring fog as a supplementary water source in Namibia. *Atmospheric Research*, *64*(1-4), 251-259.

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

Verneaux, J., Schmitt, A., Verneaux, V., & Prouteau, C. (2003). Benthic insects and fish of the Doubs River system: typological traits and the development of a species continuum in a theoretically extrapolated watercourse. *Hydrobiologia*, *490*(1-3), 63-74.

Woodwell, G. M. (1970). Effects of pollution on the structure and physiology of ecosystems. *Science*, *168*(3930), 429-433.

Tanaka, A., Mulleriyawa, R. P., & Yasu, T. (1968). Possibility of hydrogen sulfide induced iron toxicity of the rice plant. *Soil Science and Plant Nutrition*, *14*(1), 1-6.

Wang, F., & Chapman, P. M. (1999). Biological implications of sulfide in sediment—a review focusing on sediment toxicity. *Environmental Toxicology and Chemistry: An International Journal*, *18*(11), 2526-2532.

Tian, G., Granato, T. C., Pietz, R. I., Carlson, C. R., & Abedin, Z. (2006). Effect of long-term application of biosolids for land reclamation on surface water chemistry. *Journal of environmental quality*, *35*(1), 101-113.

Zhou, F., Liu, Y., & Guo, H. (2007). Application of multivariate statistical methods to water quality assessment of the watercourses in Northwestern New Territories, Hong Kong. *Environmental Monitoring and Assessment*, *132*(1), 1-13.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187-200.

López-Bucio, J., Cruz-Ramırez, A., & Herrera-Estrella, L. (2003). The role of nutrient availability in regulating root architecture. *Current opinion in plant biology*, *6*(3), 280-287.

Morse, J. W., & Luther Iii, G. W. (1999). Chemical influences on trace metal-sulfide interactions in anoxic sediments. *Geochimica et Cosmochimica Acta*, *63*(19-20), 3373-3378.