

Surface water quality assessment of Amingaon (Assam, India) using multivariate statistical techniques

Kunwar Raghvendra Singh, Nidhi Bharti, Ajay S. Kalamdhad and Bimlesh Kumar*

Department of Civil Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, India

*Corresponding author. E-mail: bimk@iitg.ernet.in

Abstract

The pollution of surface water has become a global environmental issue. Monitoring of surface water is essential to know the current status of water quality and maintain it at certain desirable level. In this study surface water quality of Amingaon has been analysed. Amingaon is a locality in North Guwahati (Assam, India). In last few decades' the locality has undergone rapid and uncontrolled development activities which have a detrimental impact on its ecology and environment. Samples were collected from 12 lakes and analysed for 24 parameters namely temperature (T), pH, electrical conductivity (EC), turbidity (Tur), total suspended solids (TSS) and total dissolved solids (TDS), total alkalinity (TA), total hardness (TH), chloride ion (Cl^-), fluoride (F^-), sulphate (SO_4^{2-}), sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), dissolved oxygen (DO) biochemical oxygen demand (BOD), chemical oxygen demand (COD), ammonium nitrogen ($\text{NH}_3\text{-N}$), total Kjeldahl nitrogen (TKN), nitrate (NO_3^-) total phosphorus (TP) and available phosphorus (AP). Multivariate statistical techniques were used for the assessment of water quality. Cluster analysis (CA) was used for classification of water quality parameters and principal component analysis (PCA) was used to identify the sources of pollution. CA grouped all the water quality parameters in three cluster. PCA resulted in six useful components which explained 90.54% of the total variance. Based on overall study it was concluded that the sources of pollution of lakes were the use of fertilizers, storm water runoff, land development and domestic waste water discharge. Trophic state of lakes was also evaluated using Carlson's Trophic State Index (TSI).

Key words: cluster analysis, principal component analysis, surface water, trophic state index

INTRODUCTION

Water quality is a state of water system exhibited through physical, chemical and biological parameters (Boyacioglu & Boyacioglu 2008). The physicochemical characteristic of water in a region is a function of combined interactive natural processes and various anthropogenic factors (Yang *et al.* 2010; Bouguerne *et al.* 2017). Surface water such as lakes and rivers are important sources of drinking water, irrigation, fishery and energy production (Iscen *et al.* 2008). In the recent years, quality of surface water has become a very sensitive issue (Simeonov *et al.* 2003; Yang *et al.* 2010). Being highly important to sustain life, water possesses a characteristic property of dissolving and carrying within it a variety of substances thus exposed to get absolutely polluted (Shrestha & Kazama 2007; Bu *et al.* 2010). The rapid urbanization and industrialization puts immense pressure on the aquatic ecosystem and creates a decrement effect on water quality and biodiversity (Iscen *et al.* 2008). Fertilizers, pesticides and other chemicals used in agricultural lands are released during irrigation and the rainfall runoff carries this contaminant load driving it to the approachable surface water body, causing agricultural non-point source pollution (Iscen *et al.* 2008). Rainfall also transports the atmospheric pollution, receiving a wide range of inputs. Waste water from industries is discharged into the

rivers and estuaries, as a result of their impact various nutrients and toxic chemicals aids to increased exploitation of water resources (Ouyang *et al.* 2006; Varol *et al.* 2012). The pollution of surface water has become a global environmental issue (Ouyang *et al.* 2006). Continuous monitoring is very important to maintain the water quality at a certain level. A major drawback of water quality monitoring is the complexity associated with numerous data and large number of variables (Simeonov *et al.* 2003; Iscen *et al.* 2008; Varol *et al.* 2012). The application of various multivariate statistical techniques helps in the interpretation of large and complex data sets for better understanding of water quality and allows the identification of possible factors that influence the water quality. Different multivariate statistical techniques such as cluster analysis (CA), principal component analysis (PCA), factor analysis (FA) and discriminate analysis (DA) have been widely used for assessment of surface water quality (Simeonov *et al.* 2003; Ouyang *et al.* 2006; Shrestha & Kazama 2007; Iscen *et al.* 2008; Zhang *et al.* 2009, 2011; Bu *et al.* 2010; Varol *et al.* 2012; Wang *et al.* 2013; Bouguerne *et al.* 2017). In the present study surface water quality of Amingaon, North Guwahati (Assam, India) has been assessed. CA and PCA was performed to classify the pollutant and identify the possible sources of pollution. Trophic state of lakes was evaluated using Carlson's Trophic State Index (TSI) value.

MATERIALS AND METHODS

Study area

Amingaon, is a locality in North Guwahati bound by $26^{\circ}11'0''$ N latitudes and $91^{\circ}40'0''$ E of equator, located at an elevation of 31 m above main sea level (Figure 1). In a subtropical climate of semidry summer and cold winter, the annual rainfall ranges from 1,500 mm to 2,600 mm, with an average humidity of 76% the maximum temperature ranges from 37°C to 39°C and the minimum temperature remains between 6°C to 7°C . The city is surrounded by wetlands and most of the residing population (approximately 8,500) depends upon it for various domestic as well as agricultural purposes. The locality has undergone rapid and uncontrolled development activities including establishment of numerous industries, various construction activities, utilization of agricultural and forest land for other development purpose. These activities have had a detrimental impact on the ecology and environment of the locality. Over the last few decades, the landscape of the city has been altered and disturbed the water retention capacity as well as the flow dynamics of various surface water sources, which are affecting the infiltration rate. All wetlands are under threat due to the

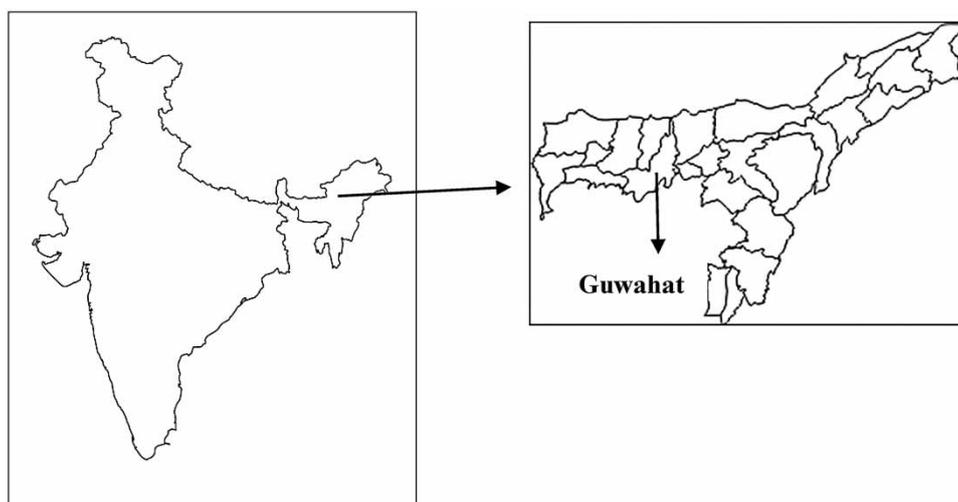


Figure 1 | Map of Assam.

encroachment and unplanned urban development of the Amingaon. Location of sampling sites is shown in Figure 2.



Figure 2 | Location of sampling sites.

Sample collection and analytical procedures

Water samples were collected between June 2016 and May 2017 from 12 different lakes (Figure 2). Four data points were set up at each lake and samples were collected in triplicates from each data point. Sampling, preservation, transportation of the water samples to the laboratory and analysis were performed according to standard methods (APHA 2012). Samples were analyzed for 24 parameters namely temperature (T), pH, electrical conductivity (EC), turbidity (Tur), total suspended solids (TSS) and total dissolved solids (TDS), total alkalinity (TA), total hardness (TH), chloride ion (Cl^-), fluoride (F^-), sulphate (SO_4^{2-}), sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), dissolved oxygen (DO) biochemical oxygen demand (BOD), chemical oxygen demand (COD), ammonium nitrogen ($\text{NH}_3\text{-N}$), total Kjeldahl nitrogen (TKN), nitrate (NO_3^-) total phosphorus (TP) and available phosphorus (AP). pH, T and DO of each water sample were measured *in situ* by field instrument, DO meter. All water samples were analyzed for different physico-chemical parameters within 48 h of collection. Tur and EC were analyzed immediately after the sample was brought to the laboratory. BOD and COD were determined on the same sampling day to avoid time-induced changes in bacterial concentration. BOD was determined by the dilution and seeding method, while COD was determined by the acid titration method. TSS and TDS were determined gravimetrically at a

temperature of 102–105 °C. TH was measured by EDTA titration. TA was determined by acid titration using methyl-orange as endpoint, and chloride by silver nitrate (AgNO_3) titration using potassium chromate (K_2CrO_4) solution as an indicator. The $\text{NO}_3\text{-N}$ and $\text{NO}_2\text{-N}$ were analyzed by phenol disulfonic acid colorimetry and N-(1-naphthyl)-ethylenediamine colorimetry, respectively. Total nitrogen TN and TP were analyzed by absorption spectrophotometry after decomposition with potassium peroxodisulfate ($\text{K}_2\text{S}_2\text{O}_8$). Na^+ , K^+ and Ca^{++} were measured by flame photometer. All of the water quality parameters were expressed in milligram per liter (mg/l) except T (°C), pH and EC (mS/cm).

Trophic state index

The trophic state of a waterbody is defined as the total weight of biomass in a given water body at the time of measurement. It represents the biological response for nutrient additions to the water bodies. Carlson (1977) quantifies the trophic state of a waterbody by its TSI value. Carlson's TSI is one of the more commonly used trophic indices and it can be classified into four basic classes: oligotrophic, mesotrophic, eutrophic, and hypereutrophic. In this study TSI of lakes was calculated by using following equations:

$$\text{TSI (TP)} = 14.42 \ln (\text{TP}) + 4.15 \quad (1)$$

$$\text{TSI (TN)} = 14.42 \ln (\text{TN}) + 54.45 \quad (2)$$

Data treatment and multivariate statistical analysis

All mathematical and statistical computations were made using SPSS 20.0. Multivariate analysis of the lake water quality data set was performed using CA and PCA techniques. The objective of CA was to assign observation to clusters (group) so that observations within each cluster are similar to one another with respect to variables and cluster themselves stand apart from one another. It is used to group the observations into homogeneous and distinct group. For present study hierarchical agglomerative clustering (HAC) method was used. It is the most common approach (Varol *et al.* 2012; Wang *et al.* 2013). In this method, all clusters formed were consist of the merger of previously formed cluster and ended with a single cluster containing all observations. CA was applied to all measured water quality parameters to group the similar parameters. PCA is a pattern recognition technique that attempts to interpret the variance within a large set of intercorrelated variables by converting them into a smaller set of independent variables (Simeonov *et al.* 2003; Iscen *et al.* 2008). It provides information on the most significant parameters to describe the entire data set with a minimum loss of original information (Isцен *et al.* 2008).

RESULTS AND DISCUSSIONS

General chemistry

Temperature is one of the most important parameters in natural surface water systems. Temperature of surface water affects the biological species and their rates of activity. It also influences the various chemical reactions that occur in natural water system and solubilities of gases. Temperature of all lakes varied from 27.5° C to 33° C. Temperature of lakes were primarily affected by ambient temperature. pH is another important parameter. pH of all lakes were in the range of 6.5 to 8.5. To permit all natural processes of aquatic life pH should be between 6.5–9.0 (USEPA 2009). Highest pH was

observed in lake 3 and lake 12. The pH of lakes 2, 7, 8, 10 and 11 were below 7 (slightly acidic), it may be due to decomposition of organic matter from wastewater discharges which can lead to acidification. TSS value ranged from 13.75 mg/l (Lake 5) to 92.50 (Lake 7) mg/l. It affects the aquatic fauna. Major sources of TSS were inorganic solids such as clay and silts, and biological solids such as algal cell and bacteria. Turbidity of all lakes varied from 4.30 NTU to 43.33 NTU. Highest turbidity was observed in Lake 11 because of high concentration of TSS (61.57 mg/l). TDS consists of inorganic salts, small amounts of organic matter, and dissolved materials in water. Concentration of TDS was varied from 127.50 mg/l (Lake 4) to 800 mg/l (Lake 7). This high level of TDS was due to agricultural and urban runoff. EC of water is affected by the presence of inorganic dissolved solids such as Cl^- , NO_3^- , SO_4^{2-} , and PO_4^{3-} anions or Na^+ and Ca^{2+} cations. Highest EC (546 mS/cm) was found in the samples of Lake 11. TA was ranged between 39 mg/l (Lake 4) to 129 mg/l (Lake 11). It is a measure of the buffering capacity of water. The main sources for alkalinity were rocks which contain carbonate, bicarbonate, and hydroxide compounds and limestone applied to fields to increase soil pH. Hardness is a common quality of water which contains dissolved compounds of calcium and magnesium. Its concentration varied from 38.00 mg/l to 163.00 mg/l (Lake 11). In Lake 11 TH was maximum because of high concentration of Ca^{2+} in the lake (44.35 mg/l). Chloride in surface water generally occurs from both natural and anthropogenic sources, such as run-off, use of inorganic fertilizers, animal feeds, and irrigation drainage. It is also a key indicator of sewage discharge. Except in Lake 1, 2, 3 and 5 concentrations of Cl^- was found more than 500 mg/l and maximum concentration was found in lake 7 (1083.19 mg/l). Fluoride is a naturally occurring element found in the earth's crust and most toxic to freshwater aquatic life. Highest concentration was found in Lake 1 (0.61 mg/l). In Lake 6, concentration of F^- was below detection limit (BDL). SO_4^{2-} is one of the major dissolved components of rain and is found in almost all natural water. Concentration of SO_4^{2-} was varied from 2.65 mg/l to 155.00 mg/l. In Lake 8, concentration of SO_4^{2-} was BDL. Na^+ is the most common nontoxic metal found in surface water. Concentration of Na^+ was varied from 3.67 mg/l (lake 6) to 25.78 mg/l (lake 7). Ca^{2+} is one of the major inorganic cations in fresh water. It is an important determinant of water hardness. Maximum concentration of Ca^{2+} was found in lake 11 (44.35 mg/l). K^+ occurs widely in all natural waters. It ranged from 1.10 mg/l to 5.50 mg/l. DO is a most important parameter in natural water systems. It determines the health of ecosystem. DO was varied from 1.42 mg/l to 9.53 mg/l. Lowest DO was observed in Lake 11 which receives untreated domestic waste water. COD a measure of the total quantity of oxygen required to oxidize all organic material into carbon dioxide and water. COD was ranged between 26.70 mg/l and 211.67 mg/l. Highest COD was observed in Lake 9. BOD measures the amount of oxygen consumed by microorganisms in decomposing organic matter in water under aerobic conditions. Highest BOD was found in Lake 1 (52.91 mg/l). In Lake 6 and Lake 8, BOD was greater than 50 mg/l. Leaves and woody debris, dead plants, and animal manure were organic matter present in all these lakes. Discharge of domestic waste water was another source of BOD found in lakes. $\text{NH}_3\text{-N}$ varied between 0.01 mg/l to 18.69 mg/l. In Lake 7, its concentration was BDL. TKN was found between 0.02 mg/l to 52.30 mg/l. Concentration of NO_3^- was ranged between 0.01 mg/l to 6.47 mg/l. In Lake 9, It was BDL. TP was ranged between 0.04 mg/l to 9.15 mg/l. All these parameters ($\text{NH}_3\text{-N}$, TKN, NO_3^- and TP) are mainly associated with use of chemical fertilizers. Variation of water quality parameters in all monitored lakes is shown in Figure 3.

Trophic state index

Trophic states of lake were calculated using Carlson's TSI. Results are summarized in Table 1. Lake 1, Lake 3, Lake 10, Lake 11 and Lake 12 was found in hyper-eutrophic condition. Hyper-eutrophic condition means lakes are extremely productive with noxious surface scums of algae and water is not

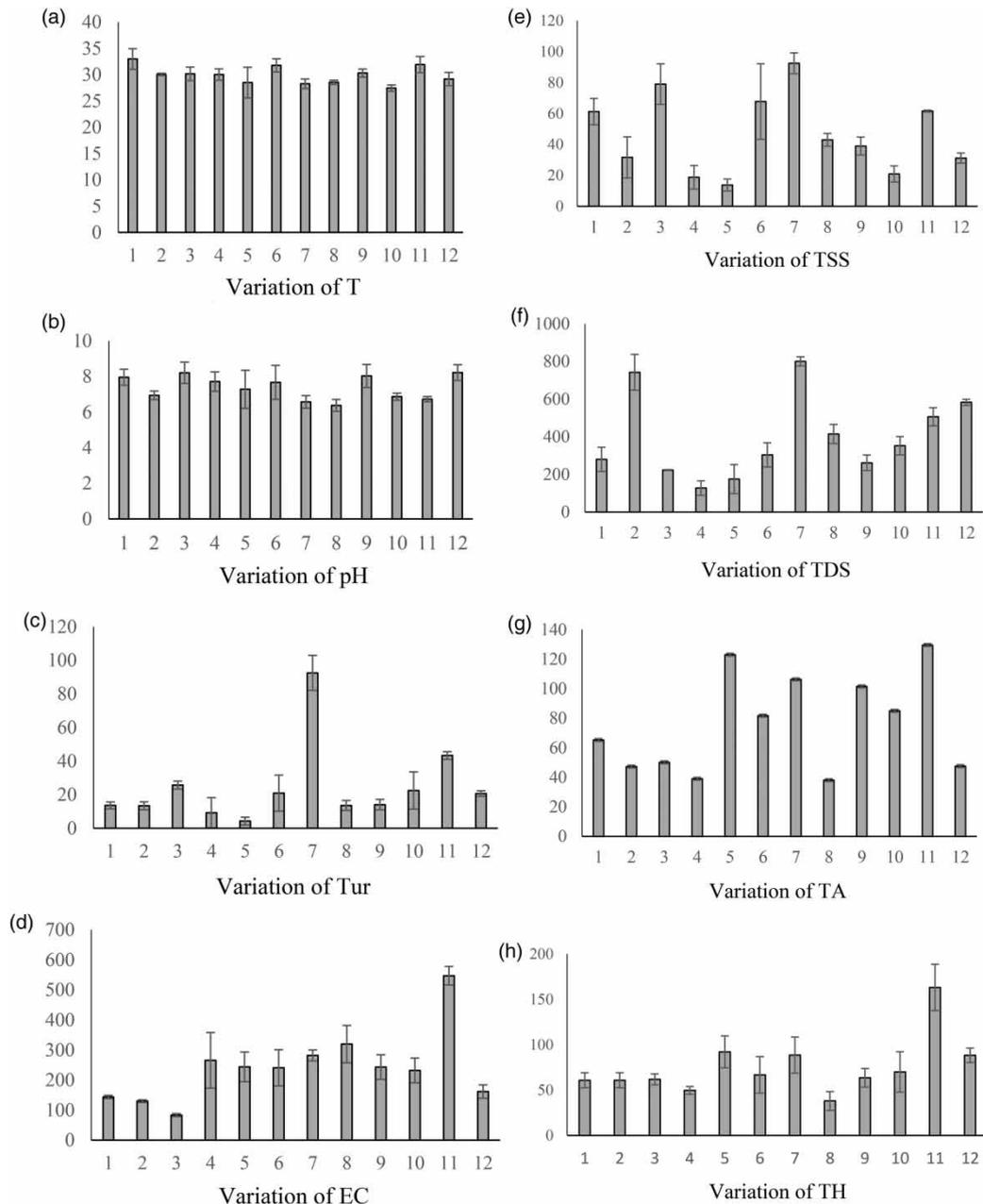


Figure 3 | Spatial variation of water quality parameters. (continued).

suitable for recreation. Lake 7, Lake 8 and Lake 9 are in eutrophic condition. It showed that these lakes are very productive and fertile; low clarity, high nitrogen and phosphorus concentrations. All other remaining lakes are found in oligotrophic condition. Oligotrophic condition means clear water, oxygen throughout the year at the bottom of the lake.

Cluster analysis

In this study, classification of measured water quality parameters was performed by the use of CA. R mode, hierarchical agglomerative cluster analysis (HCA) was performed on the normalized data using Euclidean distance as similarity measure with Ward's method of linkage. CA was applied to water quality data set to group the similar parameters and dendrogram was generated (Figure 4). As

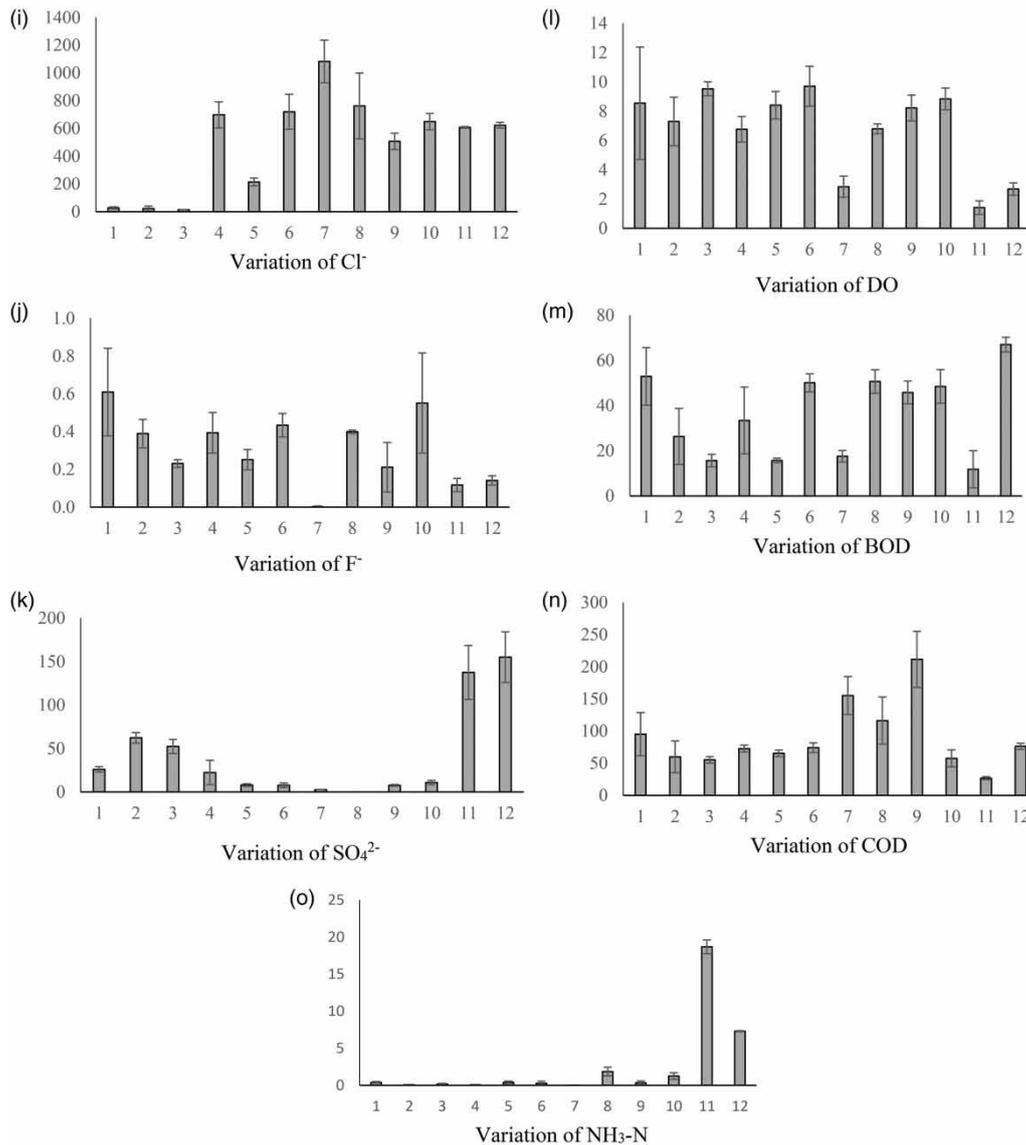


Figure 3 | Continued.

shown in [Figure 4](#), all 22 parameters were classified based on visual examination into three clusters. Cluster 1 was formed by $\text{NH}_3\text{-N}$, TKN, TH, AP, SO_4^{2-} , NO_3^- and TP. All these parameters are associated with fertilizers. Fertilizers used in agricultural lands are released during irrigation and the rainfall runoff carries this load to the approachable surface water body, causing agricultural non-point source pollution. Cluster 2 was associated with turbidity, TSS, TDS, Na^+ , K^+ , COD, EC, Ca^{2+} , Cl^- and TA. Major sources of all these parameters are urban storm water runoff and land development. Third and last cluster was formed with T, pH, F^- , DO and BOD. This group may be associated with domestic waste water discharge. Based on above discussion it was concluded that the sources of pollution of lakes were the use of fertilizers, storm water runoff, land development and domestic waste water discharge.

PCA

PCA was performed on the correlation matrix followed by Varimax rotation, for the data set of 264 samples and 22 parameters. [Table 2](#) shows the communalities of all measured parameters.

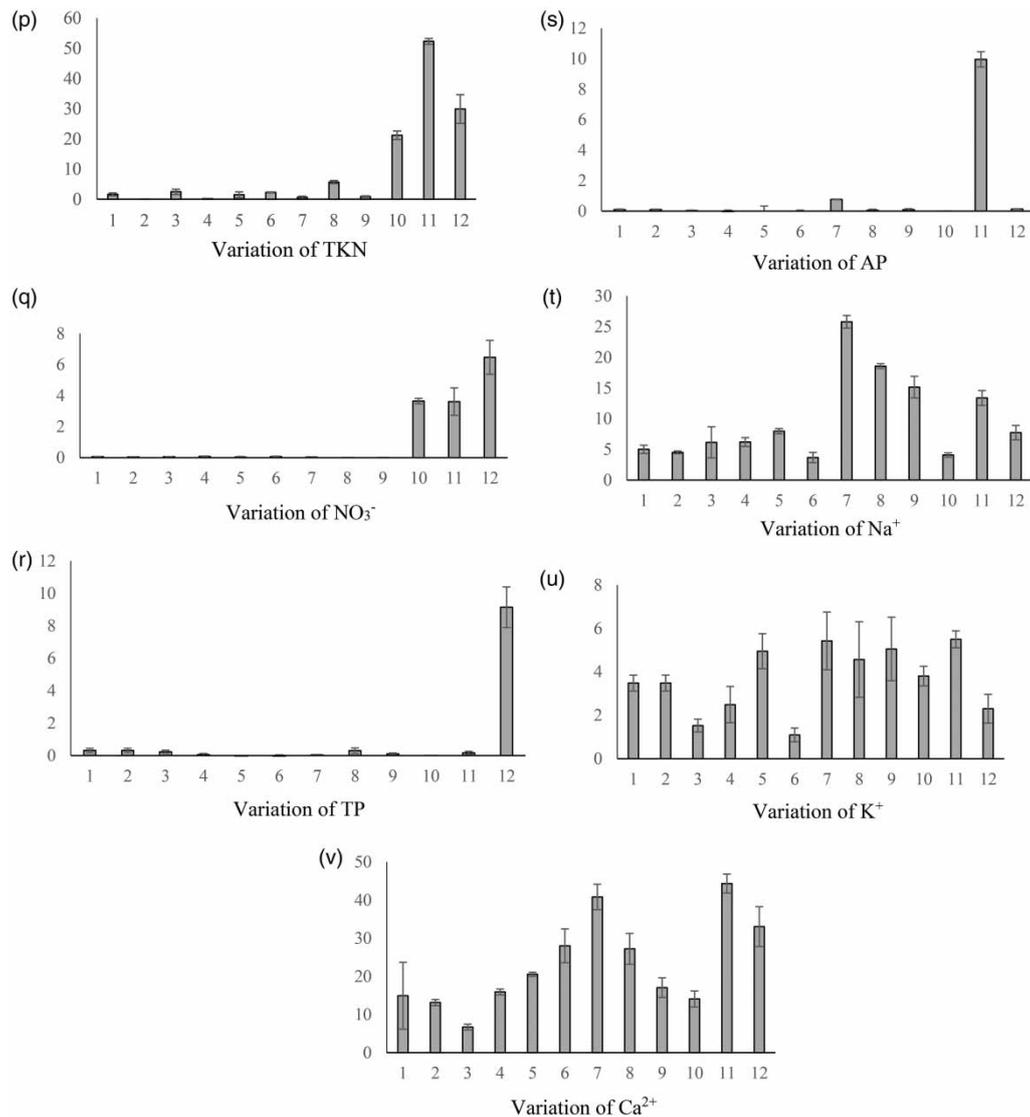


Figure 3 | Continued.

Table 1 | TSI and trophic condition of lakes

Lake	TSI (TP)	TSI (TN)	Condition
Lake 1	186.44	214.81	hyper-eutrophic
Lake 2	186.48	23.85	oligotrophic
Lake 3	116.08	154.41	hyper-eutrophic
Lake 4	47.34	28.05	oligotrophic
Lake 5	–	63.78	oligotrophic
Lake 6	–	68.12	oligotrophic
Lake 7	57.34	48.67	eutrophic
Lake 8	86.39	83.58	eutrophic
Lake 9	70.55	156.51	eutrophic
Lake 10	–	101.51	hyper-eutrophic
Lake 11	109.77	116.67	hyper-eutrophic
Lake 12	135.68	106.93	hyper-eutrophic

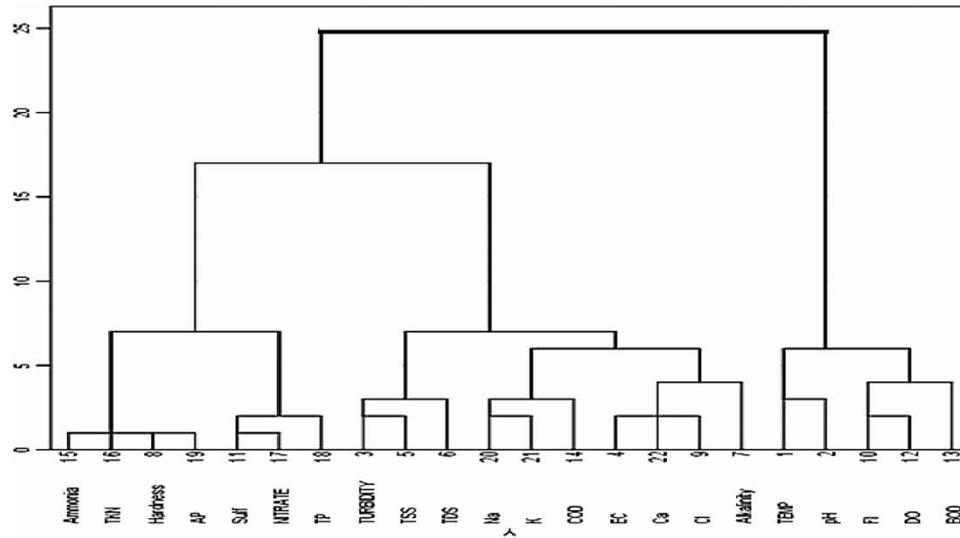


Figure 4 | Dendrogram of water quality parameters.

Communality of a parameter is the sum of the loadings of this parameters on all extracted factors. Communality of all parameters was very high (greater than 0.7) and hence FA was reliable and all the parameters were reflected well via extracted factors.

Table 3 shows the eigenvalues of the extracted factors and the proportion of total sample variance which are explained by the factors. Analysis generated total 22 factors. Eigenvalues greater than 1

Table 2 | Communalities

	Initial	Extraction
T	1.000	0.890
pH	1.000	0.965
Tur	1.000	0.886
EC	1.000	0.958
TSS	1.000	0.913
TDS	1.000	0.770
TA	1.000	0.767
TH	1.000	0.958
Cl ⁻	1.000	0.813
Fl ⁻	1.000	0.897
SO ₄ ²⁻	1.000	0.963
DO	1.000	0.903
BOD	1.000	0.930
COD	1.000	0.909
NH ₃ -N	1.000	0.971
TKN	1.000	0.944
NO ₃ ⁻	1.000	0.912
TP	1.000	0.971
AP	1.000	0.971
Na ⁺	1.000	0.913
K ⁺	1.000	0.814
Ca ⁺⁺	1.000	0.902

Table 3 | Total variance explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.274	37.608	37.608	8.274	37.608	37.608	6.118	27.811	27.811
2	4.345	19.749	57.357	4.345	19.749	57.357	3.891	17.685	45.495
3	2.830	12.863	70.221	2.830	12.863	70.221	3.782	17.191	62.686
4	1.925	8.748	78.968	1.925	8.748	78.968	2.307	10.485	73.171
5	1.338	6.084	85.052	1.338	6.084	85.052	2.103	9.560	82.731
6	1.208	5.491	90.543	1.208	5.491	90.543	1.719	7.812	90.543

were taken as criterion for extraction of the useful principal component. A total of six components had eigenvalues greater 1 as shown in Table 3. These six principal components explained 90.54% of the total variance. The aim of this study was to generate an entirely new set of component much smaller in number when compared to original data sets.

The parameter loading for all six useful components from PCA of the data set are given in Table 4. Loading greater than 0.75 was classified as strong, loading between 0.75 and 0.50 was classified as moderate and loading below 0.50 was classified as weak (Liu *et al.* 2003). The first component explained 27.8% of the total variance and is heavily loaded with EC, TH, NH₃-N, TKN and AP. TA, K⁺, Ca²⁺ had moderate loading. DO had moderately negative loading. This factor is associated with chemical components due to the geologic feature in the water environment (Bu *et al.* 2010).

Table 4 | Rotated component matrix

	Component					
	1	2	3	4	5	6
T					0.884	
pH					0.646	
Tur		0.880				
EC	0.923					
TSS		0.742			0.550	
TDS		0.782				
TA	0.643					
TH	0.838					
Cl ⁻						0.520
F ⁻						
SO ₄ ²⁻			0.801			
DO	-0.552					
BOD						0.797
COD				0.877		
NH ₃ -N	0.870					
TKN	0.793		0.509			
NO ₃ ⁻			0.832			
TP			0.960			
AP	0.923					
Na ⁺		0.627		0.623		
K ⁺	0.519					
Ca ²⁺	0.650	0.582				

The second component explained 17.68% of total variance. This component was strongly loaded with turbidity, TSS and TDS. Na^+ and K^+ had moderate loading. This factor represents the influence of surface runoff. The third component was associated with SO_4^{2-} , TKN, NO_3 and TP and explained 17.19% of total variance. SO_4^{2-} , NO_3 and TP had strong positive loading while TKN had moderate positive loading. This nutrient factor represents the non-point source of pollution due to agricultural runoff (Simeonov *et al.* 2003). The fourth component accounted for 10.48% of total variance and strongly loaded with COD and moderately loaded with Na^+ . The fifth component was associated with T, pH and TSS and explained 9.56% of total variance. This component was strongly loaded with T and moderately loaded with pH and TSS. It may represent the 'physicochemical' source of variability (Simeonov *et al.* 2003; Yang *et al.* 2010). The sixth component explained 7.8% of the total variance and had strong positive loading on BOD and moderate positive loading on Cl^- . It may be due to the point-source sewage discharges to lakes (Phillips 1994).

CONCLUSIONS

In this study, the surface water quality of Amingaon (Assam, India) was assessed. Samples were collected from 12 lakes and analysed for 22 water quality parameters. Carlson's TSI was used to identify the trophic state of all lakes. Lake 1, Lake 3, Lake 10, Lake 11 and Lake 12 were found to have hyper-eutrophic condition. Lake 7, Lake 8 and Lake 9 were eutrophic and other remaining lakes were oligotrophic. CA and PCA was used for classification of water quality parameters and source characterisation. CA grouped all the water quality parameters in three clusters. PCA resulted in six useful components which explained 90.54% of the total variance. Sources of pollution of the lakes were identified as the use of fertilizers, storm water runoff, land development and domestic waste water discharge. This study illustrates that CA and PCA are excellent tools for interpreting the large and complex data sets of water quality monitoring program. It can be useful for policy makers in water quality management.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the helpful suggestions and comments of Dr D. J. Lumley, Editor WPT.

REFERENCES

- APHA 2012 *Standard Methods for Examination of Water and Wastewater*. American Public Health Association, Washington.
- Bouguerne, A., Boudoukha, A., Benkhaled, A. & Mebarkia, A. H. 2017 *Assessment of surface water quality of Ain Zada dam (Algeria) using multivariate statistical techniques*. *International Journal of River Basin Management* **15**(2), 133–143.
- Boyacioglu, H. & Boyacioglu, H. 2008 *Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey*. *Environmental Geology* **54**(2), 275–282.
- Bu, H., Tan, X., Li, S. & Zhang, Q. 2010 *Water quality assessment of the Jinshui River (China) using multivariate statistical techniques*. *Environmental Earth Sciences* **60**(8), 1631–1639.
- Carlson, R. E. 1977 *A trophic state index for lakes*. *Limnology and Oceanography* **22**(2), 361–369.
- Iscen, C. F., Emiroglu, Ö., Ilhan, S., Arslan, N., Yilmaz, V. & Ahiska, S. 2008 *Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey*. *Environmental Monitoring and Assessment* **144**(1–3), 269–276.
- Liu, C. W., Lin, K. H. & Kuo, Y. M. 2003 *Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan*. *Science of the Total Environment* **313**(1), 77–89.
- Ouyang, Y., Nkedi-Kizza, P., Wu, Q. T., Shinde, D. & Huang, C. H. 2006 *Assessment of seasonal variations in surface water quality*. *Water Research* **40**(20), 3800–3810.

- Phillips, P. J. 1994 Chloride concentrations as indicators of point source sewage discharges in the Hudson River basin, New York. American Geophysical Union 1994 fall meeting. *EOS* 75(44 Suppl), 229.
- Shrestha, S. & Kazama, F. 2007 [Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan](#). *Environmental Modelling & Software* 22(4), 464–475.
- Simeonov, V., Stratis, J. A., Samara, C., Zachariadis, G., Voutsas, D., Anthemidis, A., Sofoniou, M. & Kouimtzi, T. 2003 [Assessment of the surface water quality in Northern Greece](#). *Water Research* 37(17), 4119–4124.
- U.S. Environmental Protection Agency (USEPA) 2009 *National Lakes Assessment: A Collaborative Survey of the Nation's Lakes*. EPA 841-R-09-001. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, DC.
- Varol, M., Gökot, B., Bekleyen, A. & Şen, B. 2012 [Water quality assessment and apportionment of pollution sources of Tigris River \(Turkey\) using multivariate statistical techniques – a case study](#). *River Research and Applications* 28(9), 1428–1438.
- Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L. F. & Li, B. L. 2013 [Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China](#). *Journal of Hydro-Environment Research* 7(1), 30–40.
- Yang, Y. H., Zhou, F., Guo, H. C., Sheng, H., Liu, H., Dao, X. & He, C. J. 2010 [Analysis of spatial and temporal water pollution patterns in Lake Dianchi using multivariate statistical methods](#). *Environmental Monitoring and Assessment* 170(1), 407–416.
- Zhang, Q., Li, Z., Zeng, G., Li, J., Fang, Y., Yuan, Q., Wang, Y. & Ye, F. 2009 [Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang watershed, China](#). *Environmental Monitoring and Assessment* 152(1), 123–131.
- Zhang, X., Wang, Q., Liu, Y., Wu, J. & Yu, M. 2011 [Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong](#). *Environmental Monitoring and Assessment* 173(1), 17–27.