Statistical Analysis of Water Quality Data of Kukkarahalli Lake, Mysuru

Shruti Kulkarni[1] Pushpa Tuppad[2]
[1]M.Tech. Student [2]Associate Professor
[1,2]Department of Environmental Engineering
[1,2]SJCE Mysuru, Karnataka

Abstract— The water quality data of Kukkarahalli Lake for 13 years obtained from KSPCB was analyzed for trends using Box-and Whisker's plots, Mann–Kendall test, Exceedance Probability plots and Principal Component Analysis. Mann-Kendall test performed for water quality data interprets that there is increasing trend in the BOD concentrations while decreasing trend in DO concentration indicating deterioration of the water quality of lake. The increasing trend for phosphate and nitrate concentration also indicates that the state of the lake is turning eutrophic. The exceedance probability curves indicated that the BOD and phosphate concentrations exceeded the permissible limit for more than 62% and 95% of the total time period. DO values were lower than 4 mg/L for about 36% of the entire time period which indicates that the sustaining aquatic life is adversely affected while nitrate concentrations did not exceed the permissible limits. The variance explained by the principal components varied among the years in the study period.

Key words: Water Quality, Kukkarahalli Lake

I. INTRODUCTION

Due to urbanization, industrialization, infrastructural developments, demographic and climatic factors and increased standard of living, there has been a lot of changes occurred in the land use and land cover in past few decades. These changes have also been affecting the surface waters due their strong influence on them. Surface water, including rivers, lakes, ponds and reservoirs, are essential natural resources for the living beings. Surface water quality causes direct impacts on the ecosystems, human health, and economic activities. The surface water bodies located near the urban areas are subjected to heavy deterioration due to increased anthropogenic pollutants entering into them. As a concern of public health, monitoring of water and study of spatial and temporal changes is of great importance.

Surface waters such as shallow lakes, are dynamic systems and are characterized by a high degree of heterogeneity in space and time (Papatheodorou et al. 2006). Lakes in urban regions are ecologically sensitive zones, important natural resources and true indicators of sustainable urban development. These lakes and their watersheds are subjected to a great pressure due to urbanization and industrialization, ultimately resulting into eutrophication of lakes. The nutrients which are present in the fertilizers as well as domestic and industrial wastewaters have been identified as main cause for changing trophic state of water bodies from Oligotrophic to Mesotrophic to Eutrophic.

South India is basically rich in lakes and tanks since olden days due to the prime importance given to irrigation by the kings. Mysuru city in Karnataka is known for many fresh water lakes. Some of these lakes were artificially constructed by the king to meet the drinking water demands of the kingdom. Presently the lakes are being used to dump domestic, agricultural and industrial wastewater. It is necessary to prevent these lakes from degrading water quality so as to preserve and safeguard the freshwater resources and associated the ecological wealth. In recent years numerous studies have been carried out by academic, non-governmental as well government funded organizations, mainly focusing on determining, analyzing and monitoring the water quality of the existing lakes in Myusru city. But little preventive measures practically taken so far are inadequate bringing back the pristine water quality of the lakes.

The present study focuses on statistical analysis of long-term (2002 to 2014) water quality data of four major parameters: BOD (Biochemical Oxygen Demand), DO (Dissolved Oxygen), Nitrates and Phosphates of Kukkarahalli lake of Mysuru city.

II. MATERIALS AND METHODS

Kukkarahalli lake is a manmade impoundment and is situated at 12°18'27.72"N and 76°38'10.07"E. It is one of the biggest natural water bodies in Mysuru city. It is located within the Manasa Gangothri campus of the Mysuru University bounded by Mysuru-Hunsur road at the south. This lake which was once lifeline of the city, is now a declining lake due to excess inflow of wastewater adding huge amount of nitrate and phosphate. Hence there is a need to protect and safeguard the lake.

For the present study, the secondary water quality data for 13 years (2002 to 2014) collected from Karnataka State Pollution Control Board (KSPCB), Mysuru and from the review of research work conducted in past (Kavya et al., 2014) was considered. The KSPCB had the repository of lake water quality data, one value per month for each parameter (mostly sampled during the first week of the month). The physico-chemical parameters considered in this study include BOD, DO, Phosphate and Nitrate.

Box and Whisker's, Mann-Kendall Test, Probability Exceedance and Principal Component Analysis (PCA) were the methods used for statistical analysis of the four water quality parameters. Box and Whisker's Plot and Mann-Kendall test were performed by using XLSTAT, an extension of MS Excel. PCA was performed by using PAST software. Both these software are available in the public domain, free to use.

III. DATA ANALYSIS

A. Box and Whisker's plot

The Box and Whisker's plot is an exploratory graphic, used to show the distribution of a dataset at a glance. The box plots being non-parametric display variation in samples of a dataset without making any assumptions of the underlying statistical distribution. These plots help to interpret the degree of spread and skewness in the data and identify outliers.

B. Mann-Kendall test

A non-parametric Mann-Kendall test, has been used to analyze the trend in meteorological, hydrological and water quality parameters (Tuppad et al. 2010; Shao 2010; Karpouzos et al. 2010; Chang 2008). It is independent of the magnitude of data and assumptions of distribution, missing data and irregularly spaced monitoring periods. It assesses whether a time-ordered data set exhibits an increasing or decreasing trend, within a predetermined level of significance.

C. Probability Exceedance Plot

Exceedance probability shows the probability of the parameter on the vertical axis that will be equaled or exceeded for the given lead time at a specific location. These plots allow the visual representation of the number of observations exceeding a particular concentration (maximum allowable concentration) for each parameter. This interpretation has a great significance in the assessment of Total Maximum Daily Load (TMDL) for watersheds, which describes the value of the maximum amount of a pollutant that a water body can receive while still meeting the water quality standards for the designated use. Application of TMDL has broadened significantly in the last decade. TMDL is a regulatory term in the U.S. Clean Water Act. The control of point sources is easy while it is difficult to control the non-point sources of pollution. Probability exceedance plots are helpful, especially in assessing the impacts of pollution caused due to non-point sources of pollution. A threshold limit on the percentage of time a pollutant is allowed to exceed could be set up that will in turn help in making decisions on effective allocation of resources for control of pollution.

D. Principal Component Analysis

PCA is a non-parametric, true Eigen vector based multivariate analyses method, the main purpose of which is to identify patterns to reduce the dimensions of the dataset with minimal loss of information. PCA is a mathematical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. The transformation is such that the first principal component has the largest possible variance or it accounts for as much of the variability in the data as possible, and each succeeding component in turn has the next highest variance possible, under the constraint that it be orthogonal, or is uncorrelated with the preceding components. Principal components are guaranteed to be independent only if the data set is jointly normally distributed.

IV. RESULTS AND DISCUSSIONS

A. Box and Whisker's Plot

BOD values ranged from 3.9 mg/L (October, 2011) to 140.3 mg/L (April, 2014) (CHART 4.1.a). The median 41.0 mg/L indicates that 50% of the BOD data is greater than 41.0 mg/L in 2002. For the entire period, majority of BOD values were in the high range of 66.0 to 79.0 mg/L and less number of values were between 20.0 to 22.8 mg/L. For the year 2003, the BOD value lies in the short range of 20.0 mg/L to 44.0 mg/L. About 50% of the BOD values were greater than 29.5 mg/L. For 2005, the box plot shows almost no skewness which represents symmetric distribution of data. The BOD values lie between the minimum of 9.0 mg/L to maximum of 49.0 mg/L. There is no significant difference between the median 30.9 mg/L and mean 30.6 mg/L. In the year 2006, the BOD values were between a minimum of 19.0 mg/L and maximum of 46.0 mg/L. 50% of the BOD values were greater than 24.3 mg/L. Hence it shows the skewness and the median of the boxplot is dragged towards the 1st (lower) quartile. From 2006, the maximum BOD value for each year started to increase till 2011. After the gradual increase a sudden fall was observed in the maximum BOD value for the year 2012. The outlier BOD value is 89.7 mg/L for the month of February which tends to be greater than 1.5 times of the 3rd Quartile (49 mg/L). This outlier found in the month of February in the year 2008 can be a value obtained due to the manual error occurred during the experiments performed or may be due to the sudden BOD load inflow during. For the year 2010, no significant skewness was seen in the plot and the data was symmetrically distributed with no significant difference in the mean (51.9 mg/L) and median (51.6 mg/L). In 2013, significant difference is observed between the median (64.0 mg/L) and mean (75.0 mg/L) indicating skewness. The median is dragged towards the lower quartile due to the majority of the value lying between 26.8 mg/L to 48.5 mg/L.

The box plot of DO (CHART 4.1.b) shows a very clear representation of the variation in DO values for each year. The DO value ranged from 1.3 mg/L (May, 2007) to 11.0 mg/L (January, 2003). After 2003 a gradual decrease is observed till 2008. For the year 2004 and 2006 there is no significant difference between the mean and the median which indicates that there is no skewness in the box plot and the data is normally distributed.

Similarly the box plots for Nitrate (CHART 4.1.c) and Phosphate (CHART 4.1.d) represent the variation of their concentrations throughout the considered time period for 2002 to 2014.

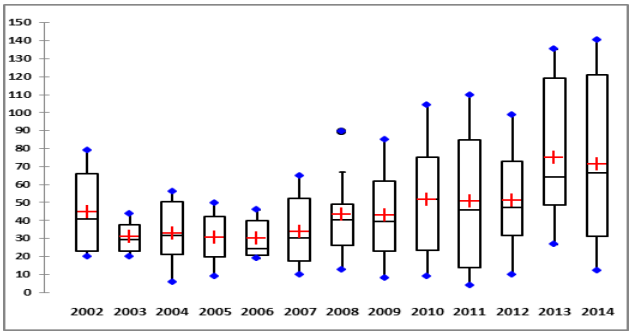

Fig. 1: Chart 4.1.a: Box and Whisker's plot for BOD spanning from 2002 to 2014 of Kukkarahalli Lake

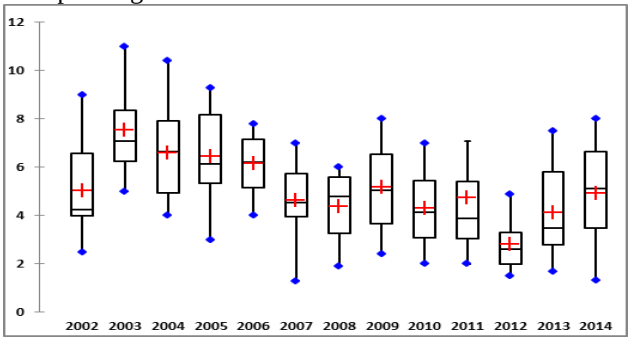

Fig. 2: Chart 4.1.b: Box and Whisker's plot for DO spanning from 2002 to 2014 of Kukkarahalli Lake

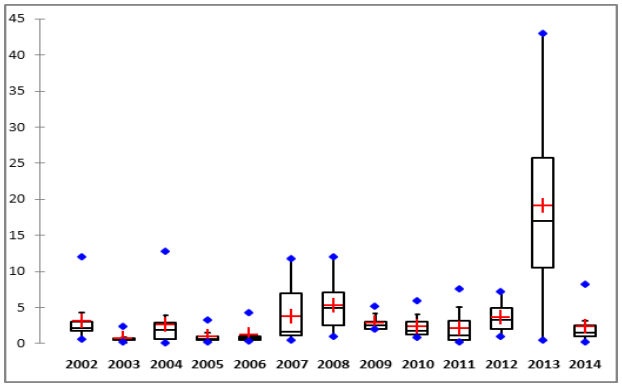

Fig. 3: Chart 4.1.c: Box and Whisker's plot for Phosphate spanning from 2002 to 2014 of Kukkarahalli Lake

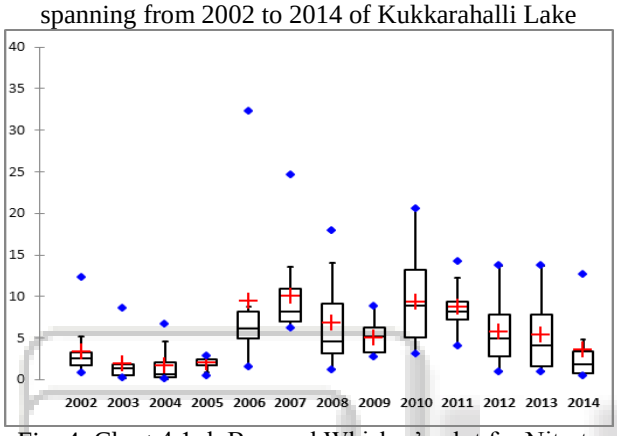

Fig. 4: Chart 4.1.d: Box and Whisker's plot for Nitrate spanning from 2002 to 2014 of Kukkarahalli Lake

The Y axis in the charts denotes the respective parameters in mg/L while X axis denotes the years from 2002 to 2014.

B. Mann-Kendall Test

According to Mann-Kendall test, BOD, DO, Phosphate, and Nitrate have p value less than 0.05, indicating acceptance of alternate hypothesis. This, in turn indicate existence of trend in these parameters.

Increasing trend in BOD (CHART 4.2.a) is a clear indication of heavily degrading water quality. These BOD value are greatly affected due to various factors like external inputs, sewage inflows, and increased demand of oxygen for the degradation of the inflowing organic waste into the lake.

The Sen's Slope is found to be negative for DO (CHART 4.2.b) indicating a decreasing trend. The decreasing trend of DO shows that there is significant degradation in the water quality since 2002 till date. The CHART 4.2.b shows that the DO had increased throughout the year 2002. Every year, as the summer approached, high DO values were observed. The high temperature and duration of bright sunlight has a great influence on the amount or percentage of soluble gases (Oxygen and Carbon-di-oxide). During summer the long days and high intensity of sunlight seem to accelerate the rate of photosynthesis by the aquatic plants and phytoplanktons utilizing CO2 and releasing oxygen. As the aquatic plants and phytoplanktons die at the end of the growing summer season, their decomposition results in heavy oxygen consumption. As a result, DO level gets reduced as it is consumed for the decomposition. The variation in the DO level is also caused by other seasonal factors like seasonal changes in lake water levels, volume of inflows and outflows, volume of sewage inflow which contributes to huge inflow of BOD load.

Parameters	P- value	Sen's slope
BOD	0.004	0.141
DO	< 0.0001	-0.02
Phosphate	0.001	0.011
Nitrate	0.003	0.02

Table 4.2: Results of Mann-Kendall test

The Phosphate (CHART 4.2.c) and Nitrate (CHART 4.2.d) concentration also showed an increasing trend for the considered time period. Phosphate concentrations showed greater values for the winter season of the year 2003. The values again rose to a very high value of around 37.0 mg/L in the year 2013 while the recommended limit for drinking water being 0.1 mg/L. The higher values of nitrate were observed in the year 2005 around 32.7 mg/L. But the nitrate values for the entire period were found to be below the permissible limit for drinking water of 45.0 mg/L.

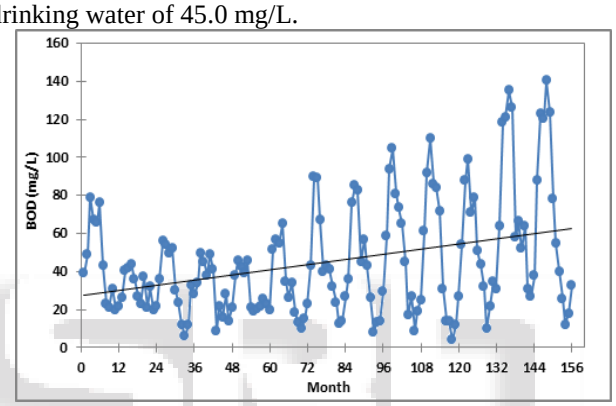

Fig. 5: Chart 4.2.a: Time series plot for BOD from 2002 to 2014 for Kukkarahalli Lake

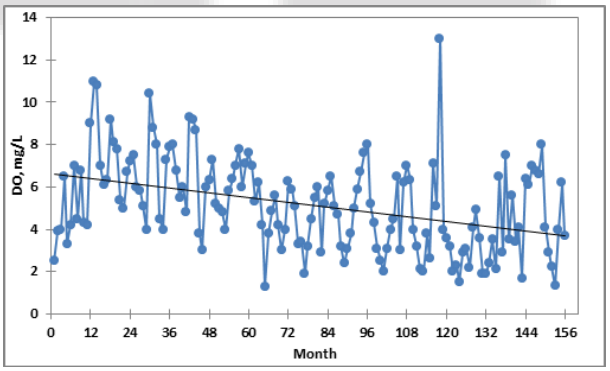

Fig. 6: Chart 4.2.b: Time series plot for DO from 2002 to 2014 for Kukkarahalli Lake

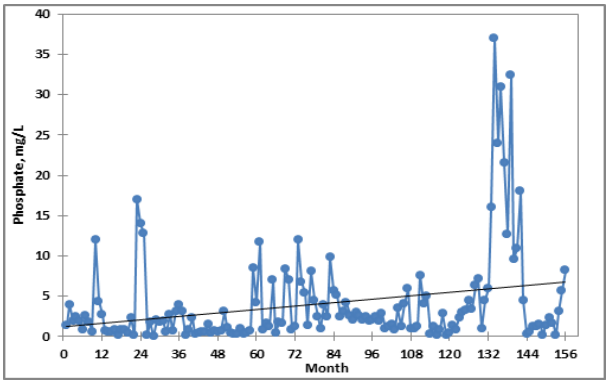

Fig. 7: Chart 4.2.c: Time series plot for Phosphate from 2002 to 2014 for Kukkarahalli Lake

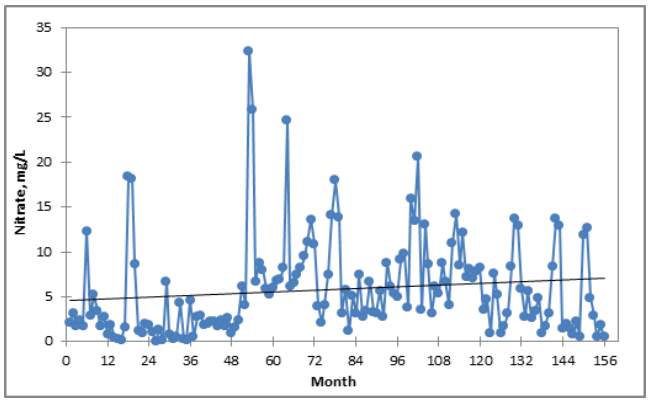

Fig. 8: Chart 4.2.d: Time series plot for Nitrate from 2002 to 2014 for Kukkarahalli Lake

C. Exceedance Probability Plot

The permissible limits of drinking water quality parameters according to IS 10500 are displayed on the exceedance probability plots to identify the length of time for which the parameter concentration equals or exceeds that value.

BOD value exceeded 30.0 mg/L for about 62% of the entire time period considered (CHART 4.3.a). DO values were lower than 4.0 mg/L for about 36% of the entire time period and was lower than 2.0 mg/L for about 6% of the time (CHART 4.3.b). The phosphate concentration exceeded 0.1 mg/L for 95% of the time (CHART 4.3.d). Nitrates (CHART 4.3.c) concentration did not exceed the permissible limit ever in the study period considered.

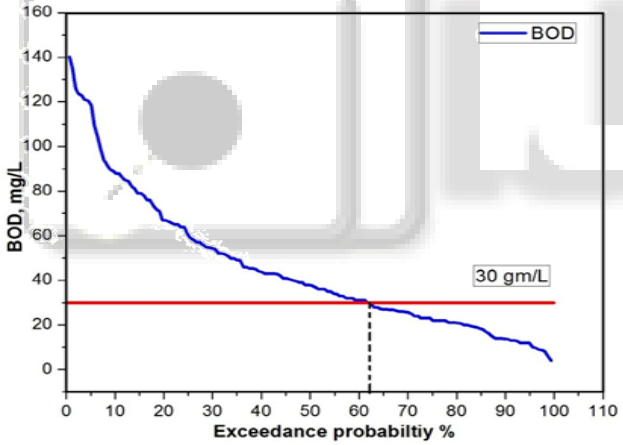

Fig. 9: Chart 4.3.a: Exceedance Probability curve for BOD concentrations

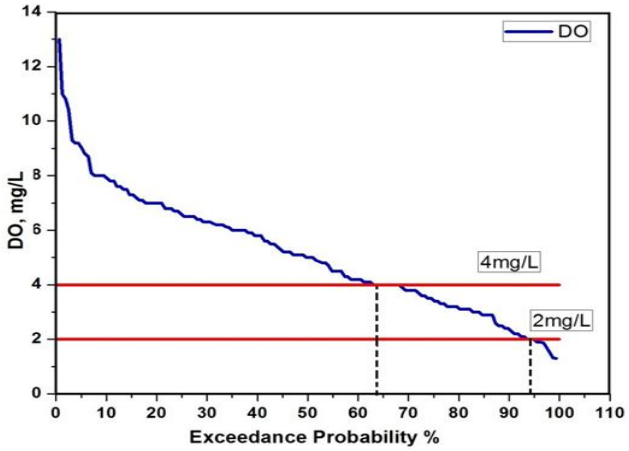

Fig. 10: Chart 4.3.b: Exceedance Probability curve for DO concentrations

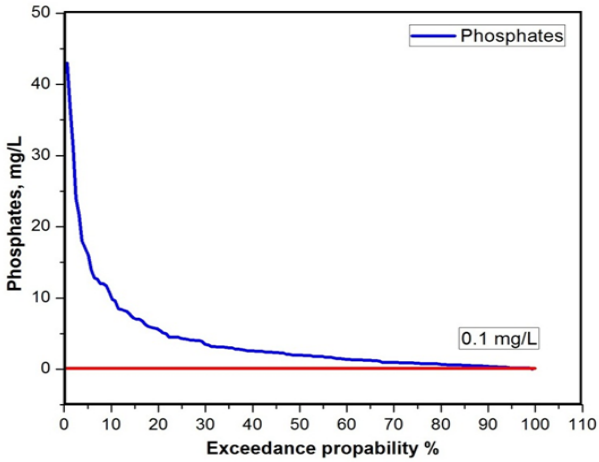

Fig. 11: Chart 4.3.c: Exceedance Probability curve for Phosphate concentrations

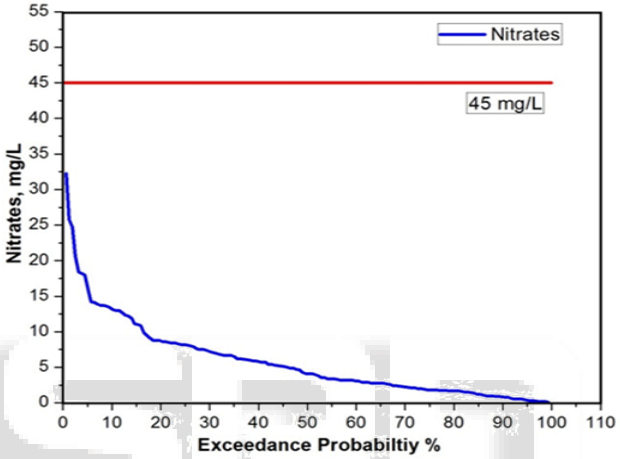

Fig. 12: Chart 4.3.d: Exceedance Probability curve for Nitrate concentrations

D. Principal Component Analysis

According to PCA performed for each year of the considered time period, PC1 and PC2 were the two principal components found for each year except for the year 2010, in which, PC1 was the only principal component found. Among the 4 considered parameters, BOD was the principal component influencing the lake water quality. CHART 4.4.a and CHART 4.4.b display the PCA of Kukkarahalli lake for the year 2002 and 2014 respectively.

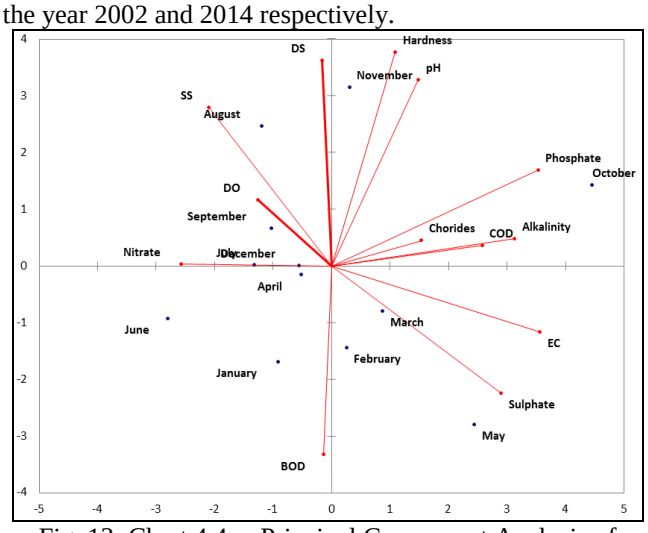

Fig. 13: Chart 4.4.a: Principal Component Analysis of Kukkarahalli Lake for 2002

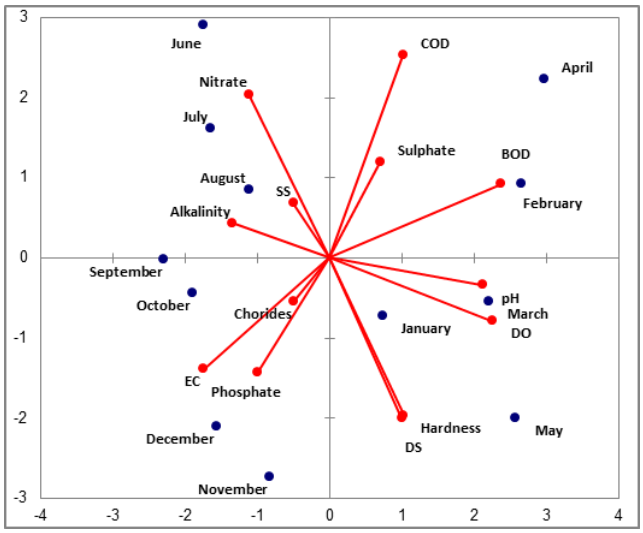

Fig. 14: Chart 4.4.b: Principal Component Analysis of Kukkarahalli Lake for 2014

V. CONCLUSIONS

Various statistical techniques including Box and Whisker's plots, Mann-Kendall test, Exceedance Probability and PCA were applied on BOD, DO, Phosphate, and Nitrate data of Kukkarahalli Lake. Box and Whisker's plots showed that mean concentration values are higher than median indicating that the concentrations values are skewed toward the higher values and hence most of the constituents analyzed in this study showed departures from the normal distribution. The decreasing trend for DO while increasing trend for BOD, phosphate and nitrate interpreted by Mann-Kendall test indicated that the state of the lake is turning eutrophic which is harmful for the lake ecosystem. The exceedance probability curve indicated that the BOD and phosphate concentrations exceeded the standard concentration limit for 62% and 95% of the total time period, respectively. DO values were lower than 4.0 mg/L for about 36% of the entire time period which indicates sustaining aquatic life is adversely affected. While it was lower than 2.0 mg/L for about 5% of the time which indicates death of fishes. According to PCA Analysis, BOD was found to be the principal component among the considered four parameters.

REFERENCES

[1] George Papatheodorou, Gerasimoula Demopouloua, Nicolaos Lambrakis (2006), "A long-term study of temporal hydrochemical data in a shallow lake using multivariate statistical techniques", Ecological Modelling Vol. 193 : 759–776
[2] Guangju Zhao, Georg Hörmann, Nicola Fohrer, Zengxin Zhang, Jianqing Zhai (2010), "Streamflow Trends and Climate Variability Impacts in Poyang Lake Basin, China Water Resour Manage Vol. 24 : 689–706
[3] Heejun Chang (2008),"Spatial analysis of water quality trends in the Han River basin, South Korea", Water Research Vol. 42 : 3285 – 3304
[4] Karpouzos D.K, Kavalieratou S. and Babajimopoulos C. (2010), "Trend Analysis of precipitation Data in Pieria Region (Greece)", European Water pg. no. 31-40
[5] Kavya S. and Savitha Ulavi (2014), " Basillariophyceae as Ecological indicators of Water Quality in two lakes of Mysore", Universal Jouranal of Environmental Research and Technology, Vol 4, Issue 1 : 1-11
[6] Pushpa Tuppad, Chinnasamy Santhi, Raghavan Srinivasan (2010), "Assessing BMP effectiveness: multi-procedure analysis of observed water quality data", Environ Monit Assess Vol 170 : 315–329
[7] Sudevi Basu and Lokesh K., (2012), "Evaluation of Cauvery River Water Quality at Srirangapatna in Karnataka using Principal Component Analysis", International Journal of Engineering and Science ISSN: 2278-4721, Vol. 1, Issue 4, Pg,no. 6-12

