Available online at www.elixirpublishers.com (Elixir International Journal)

Pollution

Elixir Pollution 126 (2019) 52424-52429

ISSN: 2229-712X

Multivariate Analysis of water Quality and Identification of Potential Pollution Sources of Lake Hawasa, Ethiopia

Melaku Zigde and Endale Tsegaye

Chemistry Department, College of Natural and Computational Science, Hawassa University, Hawassa, Ethiopia.

ARTICLE INFO

Article history:

Received: 3 December 2018;

Received in revised form: 24 December 2018;

Accepted: 4 January 2019;

Keywords

Multivariate Statistics,
Water Pollution,
Water Quality Parameters.

ABSTRACT

Multivariate statistics was used to categorize the potential sampling sites of Lake Hawassa and identify potential pollution sources by analyzing water quality parameters. Water quality parameters, such as total dissolved solids (TDS), pH, temperature, conductivity, turbidity, dissolved oxygen (DO), five day biological oxygen demand (BOD_5), total hardness as $CaCO_3$, total alkalinity as $CaCO_3$, nitrate, sulfate, orthophosphate, fluoride, K, Mg, Cu, Cd, Cr, Fe, Mn, Pb, and Zn were determined. The results were compared with WHO standards. Principal component analysis (PCA) extracted seven principal components. The first principal component (PC1) accounted for 22.1% of the total variance, and pH, Mn turbidity, specific conductance (SC) and sulfate were strongly loaded on it. Principal component two (PC2) was mainly composed of BOD_5, total hardness, temperature, iron, DO, and TDS. This component accounted for 17.3% of the total variance. The third component (PC3), dominated by potassium, TDS, and zinc, explained 12.4% of the total variance. Copper and fluoride were associated in the fourth principal component, accounting for 10.5% of the total variance. In the fifth component, total alkalinity, phosphate, and zinc were the dominant components, which account for 10.2% of the total variance. Nitrate, chromium, and lead were isolated in the sixth component (PC6), accounting for 9.4% of the total variance. The last component was dominated by magnesium, explaining 7.6% of the total variance. Hierarchical cluster analysis (HCA) divided the sampling sites into four clusters. Cluster A include five sampling sites and it was highly loaded with PC4 and PC6, which showed the presence of a high level of pollution from industrial effluents and agricultural runoff. Cluster B and C consisted of three sampling sites and one sampling site, respectively, and they were highly loaded with PC3 and PC6, which indicated the presence of a high level of pollution from domestic wastewaters, land development and urban runoff. Cluster D comprised two sampling sites and was highly loaded with all component loadings except for PC7. It was considered a highly polluted site from multiple sources of pollution. These results obtained from the multivariate analysis can be very useful for the surrounding rural and urban communities for the proper and safe use of the lake. In addition, it can reduce the cost associated with monitoring the lake by reducing the number of sampling sites.

Introduction

Water is an indispensable and basic element that supports life and the natural environment, a prime component for industry, a consumer item for human beings and animals, and a vector for domestic and industrial pollution. Access to adequate water for domestic purposes, irrigation, and sanitation are the three basic needs that impact significantly on socioeconomic development and the standard of life. In general, urbanization, industrialization, agricultural activities, and tourism, as well as population growth,-and changes in climate and lifestyle, put increasing constraints on water resources and ecosystems [1].

Ethiopia is endowed with a number of lakes and large rivers, which give immense value to the overall economic development. For instance, the country has twelve river basins, eleven fresh lakes, nine saline lakes, four crater lakes and over twelve major swamps/wetlands. The Rift Valley Lakes Basin (RVLB) is a hydrologically closed basin, characterized by terminal lakes; those with no surface water outlet. Four of the seven main lakes of the RVLB are terminal in themselves, and those that are not (Ziway, Langano and Abaya) flow into terminal lakes and are thus part of a terminal lake system. Ethiopian Rift Valley Lakes have significant environmental, economic and cultural importance to the region [2].

Among freshwater resources, Lake Hawassa is one of the major Rift Valley lakes in Ethiopia and is used for various purposes by semi-urban and urban dwellers [3, 4]. It is considered to be the livelihood of all business created along the shore and plays a great role in tourism, investment, and biodiversity conservation. However, the lake faces a high risk of pollution as a result of natural activities, such as erosion and heavy rainfall, and anthropogenic activities, such as urban0ization, intense agriculture, rapid industrialization, growth of population, urban runoff and municipal waste, overfishing, grass cutting along the shores, car washing on the lakeshore, and horticultural farming on the lakeshore [4-10].

Tele: +251912371675
E-mail address: melerevised@gmail.com

Thus, the aim of the study was to assess the current pollution status of Lake Hawassa and to obtain suitable sampling sites and identify major types of pollution using multivariate statistical analysis to carry out appropriate measures.

Materials and Methods

Description of the Study area

This study was carried out in Lake Hawassa, one of the major Rift Valley lakes with a closed basin feature that receives only one perennial river from the eastern escarpment, TikurWuha River. It is located between 06°58′ to 07° 14′ North latitudes and 38° 22′ to 38° 28′ East longitudes with an elevation of 1680 meter above sea level. It has a total surface area of 90 km^2 and a drainage area of 1250 km^2 [11].

Sampling, Sample preparation and Analysis

The sampling sites were selected based on information available about the sampling sites. The selected sampling sites are shown in Figure 1.

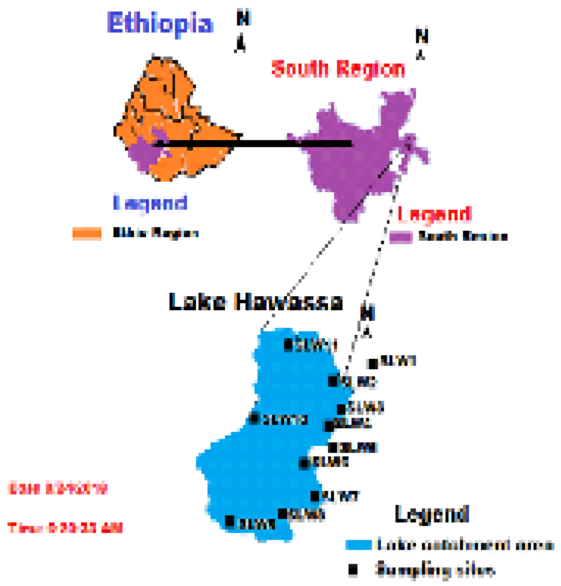

Figure 1. Location map of sampling sites at Lake Hawassa.

A total of twenty two water samples were collected from the selected sampling sites during the wet season (July and August) in 2018. The samples were collected in one liter capacity plastic bottles after they had been thoroughly rinsed with the sample and preserved airtight to avoid evaporation. Physical parameters such as total dissolved solids (TDS), SC, and temperature were determined in situ using a Wagtech Conductivity/TDS Meter. Dissolved oxygen was also determined in situ using a HANA Model HI 9143 Dissolved Oxygen Meter. Turbidity and pH were determined onsite using a Wagtech Turbidimeter, and a pH meter, respectively.

These samples were kept refrigerated prior to the analysis of nutrients, major ions and trace metals. Major ions’- such as K^+ and Mg^{2+} and trace metals such as iron were determined using Photometer 7100 integrated with the Palintest system of water analysis. Nutrients such as NO_3^-, SO_4^{2-}, PO_4^{3-} and F^- were determined using Photometer 7100 integrated with the Palintest system of water analysis. Total alkalinity and total hardness measurements were carried out by acid titration with 0.02 N H_2SO_4 and ethylenediamine tetra acetic acid (EDTA) titration, respectively.

For the analyses of trace metals such as Mn, Cu, Zn, Cr, and Pb, 100.0 mL of unfiltered water sample was taken in a beaker and heated until the volume of the sample solution reached 20.0 mL. Then, the sample solution was cooled and acidified with 2.0 mL of concentrated nitric acid and made up to the mark with deionized water. Then, the analyses were completed using an atomic absorption spectrophotometer (Buck Scientific, Model 210 VGP Atomic absorption spectrophotometer, USA).

Statistical analysis

Descriptive statistics and correlation analysis for the selected physic-chemical water quality parameters were carried out.

Multivariate Statistical Analysis

Multivariate statistical analysis can help to simplify and organize large data sets to provide meaningful insights [12-16]. In the present study, multivariate statistical analysis such as hierarchical cluster analysis (HCA) and principal component analysis (PCA) was used to evaluate the water quality of the Lake and identify potential sources of pollution. HCA was applied with the objective of grouping end-member samples according to their chemical similarities. PCA was then applied to the same subset to reveal details related to end-member physico-chemical characteristics. The statistical software package IBM® SPSS statistic 20 [23] was used for the multivariate statistical analysis.

Hierarchical cluster analysis (HCA)

In [17-19], HCA is described as “an efficient means to recognize groups of samples that have similar chemical and physical characteristics. HCA is one type of cluster analysis that has been used to view water chemistry data for both surface water [17] and ground water [20]. In this study, HCA on the original data set, using Ward’s method as a linkage rule and squared Euclidean distances as a measure of similarities [17-19], was used to classify sampling sites. The results of a hierarchical clustering procedure can be displayed graphically using a dendrogram, which shows all the steps in the hierarchical procedure [21, 22, 23].

Principal Component Analysis (PCA)

PCA was used to explain the variances observed in the data and understand the relationship between the different physic-chemical parameters [22, 24]. It was used to transform original water quality parameters into new uncorrelated variables called principal components (PCs), which are expressed as linear functions of the water quality parameters [25,26].

PCA was applied on standardized data through Z-scale transformation to avoid misclassification due to the wide difference in data dimensionality [14, 27, 31, 32]. The PCs were extracted on the basis of the Kaiser criterion and using a Varimax rotation [28-30]. Before applying PCA to the water quality parameter data set, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy [33] and Bartlett’s tests of sphericity [34,35] were performed on the parameter correlation matrix to validate the PCA.

The materials and methods section should contain sufficient details so that all procedures can be repeated. It may be divided into headed subsections if several methods are described.

Results and Discussion

Descriptive statistics

The results of the analyses of the water samples from the study area are presented in Table 1. The mean concentrations of the different parameters analyzed in the samples were compared with the WHO standards.

Table 1. Summary of the physico-chemical results of the water samples from the study area. All units except for temperature (°C), Turbidity (NTUs) and SC (μS/cm) are in mg/L.

Parameters	Minimum	Maximum	Mean	WHO (2008)
TDS	368	1289	974.5	500
pH	5.79	9.47	8.19	6.5 to 8.5
Temp	20.1	28.6	24.5	> 40.0
SC	730	2154	1564.71	750
Turbidity	8.01	91	25.9	5
DO	2.36	7.33	5.48	5.0 to 7.0
BOD5	3.89	78.6	35.2	2.0 to 5.0
Total hardness	52	74	61.7	300
Total alkalinity	10	82	27.2	120
Nitrate	2.25	7.35	3.86	45
Sulfate	60	192	119.7	250
Phosphate	0.29	1.93	1.06	0.1
Fluoride	10.7	18.5	13.5	1.5
Potassium	51.9	92	74.8	-
Magnesium	16.4	33	25.6	30
Iron	0.09	0.85	0.46	0.3
Lead	0.02	0.17	0.082	0.01
Chromium	0.16	0.75	0.38	0.05
Manganese	0	1.9	0.47	0.1
Copper	1.32	19.2	6.04	2
Zinc	0.2	17.5	4.77	5

All units except temperature (°C), Turbidity (NUTs) and SC (μS/cm) are in mg/L.

Correlation analysis

Pearson's correlation coefficients between the physico-chemical water quality parameters are summarized in Table 2.

Table 2. Pearson correlation coefficient matrix of the physico-chemical water quality parameters.

Pearson Correlations

	TDS	pH	Temp	SC	Turbidity	DO	BOD5	Total Hardness	Total Alkalinity	Nitrate	Sulfate	Phosphate	Fluoride	Potassium	Magnesium	Iron	Lead	Chromium	Manganese	Copper	Zinc
TDS	1																				
pH	-.204	1																			
Temp	.623**	.127	1																		
SC	.134	.913**	.463**	1																	
Turbidity	-.020	-.736**	-.047	-.672**	1																
DO	.578**	.298*	.710**	.562**	.040	1															
BOD5	-.569**	-.016	-.884**	-.329**	-.117	-.734**	1														
Total Hardness	.104	-.398**	.400**	-.269*	.438**	.237	-.632**	1													
Total Alkalinity	-.205	-.188	.020	-.228	.103	-.354**	.112	.321**	1												
Nitrate	-.180	.356**	.005	.225	-.170	.134	.089	-.055	-.135	1											
Sulfate	-.457**	-.681**	-.501**	-.708**	.677**	-.379**	.382**	.379**	.132	-.019	1										
Phosphate	.449**	-.195	.084	-.057	.386**	.644**	-.194	-.083	-.459**	.065	.073	1									
Fluoride	-.029	.373**	.337**	.488**	-.054	.472**	-.422**	.106	-.163	.147	-.082	.303*	1								
Potassium	-.571**	.186	-.218	.019	.175	.029	.097	.325**	.286*	-.007	.335**	-.034	.165	1							
Magnesium	.247*	-.101	.351**	-.007	.066	.219	-.430**	.024	.004	.131	-.336**	.298*	.237	-.095	1						
Iron	.680**	-.362**	.342**	-.133	.144	.313*	-.636**	.490**	.018	-.335**	-.040	.169	.041	-.253*	.038	1					
Lead	.435**	-.294*	.090	-.148	.339**	.300*	-.200	.098	-.183	.230	.197	.693**	.319**	-.372**	.102	.510**	1				
Chromium	.437**	.085	.623**	.266*	-.254*	.164	-.424**	.055	.098	-.468**	-.696**	-.384**	-.314*	-.273*	.019	.251*	-.444**	1			
Manganese	.365**	-.861**	-.111	-.723**	.726**	-.049	-.082	.157	-.113	-.347**	.415**	.636**	-.137	-.319**	.173	.445**	.580**	-.177	1		
Copper	-.257*	-.188	-.439**	-.433**	-.114	-.442**	.343**	.111	.131	.156	.159	-.172	-.606**	.304*	.133	-.098	-.348**	-.125	-.069	1	
Zinc	-.479**	.070	-.286*	-.094	-.103	-.376**	.619**	.010	.686**	.102	.273*	-.217	-.089	.627**	-.119	-.529**	-.419**	-.289*	-.386**	.330**	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

It can be seen that TDS had a significant positive correlation with temperature, and DO. In addition, the three parameters exhibited negative correlations with BOD5, which in turn was negatively correlated with total hardness. However, DO had shown a positive correlation with phosphate. pH had a significant positive correlation with specific conductance, whereas it showed a negative correlation with manganese, sulfate , and turbidity. Specific conductance exhibited positive correlations with DO, but negative correlations with manganese, and sulfate

Multivariate statistical analysis

Principal Component Analysis (PCA)

PCA was conducted on the lake water data to identify the probable sources of water pollution. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy (0.612) and Bartlett's test of sphericity (2306.145, p is 0.000) showed that the water quality parameters were suitable for PCA. The results showed that seven PCs were extracted, which could account for over 89.5% of the total variation in heavy metal concentrations. Table 3 presents the PCs, eigenvalues, and cumulative percent of the water quality parameters.

The first PC (PC1) was strongly and positively loaded with Mn, turbidity, and sulfate and strongly and negatively loaded with pH and SC with strong loadings of 0.892, 0.872, and 0.734, and -0.917 and -0.871, respectively, which accounts for 22.06% of the total variance in the lake water. This factor predominantly consisted of domestic wastewaters, land development, and urban runoff.

The second PC (PC2) accounts for 17.26% of the total variance. This component is moderately and positively loaded with total hardness (0.790), temperature (0.755), iron (0.703), DO (0.637), and TDS (0.531) and strongly and negatively loaded with BOD5 (-0.866). This factor was primarily composed of waste discharge, land development, and urban runoff.

The third component (PC3), explained 12.44% of the total variance, and the factor is strongly and positively loaded with potassium (0.943), moderately loaded with zinc (0.522), and moderately and negatively loaded with TDS (-0.618). This component mainly represented influences from residential activities, household products, and pharmaceuticals.

Copper (-0.898) and fluoride (0.737) were strongly and moderately loaded in the fourth PC (PC4), accounting for 10.52% of the total variance. This factor was mainly composed of anthropogenic impacts, possibly from waste water disposal (human sewage, and industrial waste from ceramics, dyes, plastics and food processing industries) and overland runoff.

The fifth component (PC5) was strongly and moderately loaded with total alkalinity (-0.914), zinc (-0.570), and phosphate (0.637), which accounts for 10.16% of the total variance. This factor might be due to the pollution from agricultural runoff and solid-waste incinerators.

Nitrate (0.842), lead (0.556) and chromium (-0.764) were strongly and moderately loaded in the six component (PC6), accounting for 9.41% of the total variance. This latent factor was mostly composed of anthropogenic impacts, possibly from agricultural runoff, domestic wastewaters, urban stormwater, and land development.

The seventh component (PC7) was strongly and positively loaded with magnesium (0.910), explaining 7.63% of the total variance. This factor consisted of the natural dissolution of magnesium containing minerals-rocks.

Table 3. Principal components, eigenvalues, and cumulative percent for the determined water quality parameters.

Parameters	Principal component loadings						
	PC1	PC2	PC3	PC4	PC5	PC6	PC7
pH	**-0.917**	-0.089	0.18	0.231	0.13	0.114	-0.057
Manganese	**0.892**	0.043	-0.336	0.034	0.226	-0.06	0.12
Turbidity	**0.872**	0.136	0.208	0.154	0.052	-0.016	0.0037
SC	**-0.871**	0.169	0.022	0.413	0.163	-0.005	0.033
Sulfate	**0.734**	-0.016	0.368	-0.024	-0.103	0.268	-0.396
BOD5	0.076	**-0.866**	0.071	-0.251	-0.135	0.136	-0.286
Total Hardness	0.354	**0.79**	0.317	-0.1	-0.267	0.095	-0.137
Temp	-0.257	**0.755**	-0.168	0.318	-0.071	-0.16	0.329
Iron	0.301	**0.703**	-0.411	-0.041	0.08	-0.048	-0.174
DO	-0.215	**0.637**	0.05	0.41	0.465	0.056	0.231
Potassium	0.013	0.045	**0.943**	-0.027	-0.076	0.04	-0.052
TDS	0.054	**0.531**	**-0.618**	0.064	0.254	-0.16	0.275
Copper	0.058	-0.108	0.269	**-0.898**	-0.058	0.163	0.137
Fluoride	-0.162	0.215	0.211	**0.737**	0.112	0.305	0.145
Total Alkalinity	0.13	0.061	0.14	-0.019	**-0.914**	-0.071	0.053
Phosphate	0.13	0.061	-0.022	0.301	**0.637**	0.234	0.4
Zinc	0.385	0.054	**0.522**	-0.067	**-0.57**	0.112	0.104
Nitrate	-0.054	-0.371	0.023	-0.1	0.018	**0.842**	0.117
Chromium	-0.328	-0.031	-0.27	-0.159	-0.103	**-0.764**	0.034
Lead	-0.335	0.374	-0.476	0.361	0.213	**0.556**	-0.014
Magnesium	0.429	0.208	-0.059	-0.036	0.004	0.089	**0.91**
Initial Eigenvalue	4.634	3.625	2.614	2.21	2.134	1.977	1.603
Variance %	22.06	17.26	12.44	10.52	10.16	9.416	7.635
Cumulative %	22.06	39.33	51.77	62.3	72.46	81.87	89.51

PC1 and PC2 were assigned as contributions from anthropogenic sources such as domestic wastewaters, land development, and urban runoff while PC4, PC5 and PC6 were assigned as contributions from industrial effluents, and agricultural runoff. PC3 was assigned as the contribution from domestic wastewaters, and PC7 was contributions from rock-water interactions.

Hierarchical Cluster Analysis

HCA produced four clusters (Cluster A, Cluster B, Cluster C and Cluster D) based on their water quality similarities. One way ANOVA analysis has shown that the four clusters were statistically significant at the 95% confidence interval. The clusters were clearly different from each other in a dendrogram, as illustrated in Figure 2. Cluster A comprised five sampling sites, SLW5, SLW6, SLW8, SLW9, and SLW10, which were near Fikir Haik marsh, Lewi resort, Hawassa Referral Hospital, Loke, and Algirima, respectively. Cluster B included three sampling sites, SLW 2, SLW3, and SLW11, which were near Tikur Wuha Lake, Haile resort, and the airport, respectively. Cluster C consisted of only one sampling site SLW1 (Tikur Wuha River) closer to the inlet of the river. Cluster D contained two sampling sites SLW4 and SLW7 that were near Fikir Haik and Amoragedel, respectively.

These results indicated that HCA offered a reliable categorization of the studied water sampling sites of the Lake, and this categorization is essential to design a sampling strategy that can reduce the number of sampling sites and the cost associated with it.

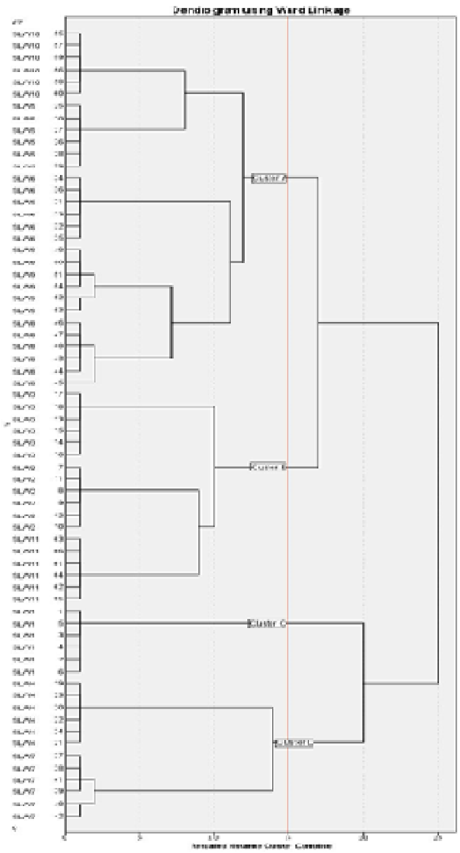

Figure 2. Dendrogram with an added line indicating the optimal stopping point of the clustering procedure.

Descriptive statistics for the determined physico-chemical water quality parameters of the identified clusters are shown in Table 4 and their corresponding stacked graphs are shown in Figure 3a and 3b.

Table 4. Descriptive statistics for each cluster.

Parameters	Cluster A	Cluster B	Cluster C	Cluster D	WHO (2008)
	mean ± SD , n=30	mean ± SD , n=18	mean ± SD , n=6	mean ± SD , n=12	
TDS	1056.5 ± 347.8	896.1 ± 278.1	428.7 ± 8.78	1160.0 ± 7.38	500
pH	8.56 ± 0.407	9.15 ± 0.216	7.27 ± 0.047	6.27 ± 0.469	6.5 to 8.5
Temp	25.4 ± 2.97	25.0 ± 0.48	21.2 ± 0.19	23.1 ± 2.24	> 40.0
SC	1766.6 ± 123.5	1983.8 ± 130.6	843.0 ± 2.68	792.1 ± 61.1	750
Turbidity	15.9 ± 7.6	20.4 ± 7.7	32.3 ± 0.86	55.8 ± 32.96	5
DO	5.73 ± 1.11	6.45 ± 0.56	2.39 ± 0.031	4.92 ± 1.99	5.0 to 7.0
BOD5	27.1 ± 23.6	33.2 ± 5.93	76.2± 1.60	37.7 ± 27.7	2.0 to 5.0
Total Hardness	61.2 ± 5.9	59.9 ± 3.058	67.2 ± 0.75	62.7 ± 10.3	300
Total Alkalinity	24.8 ± 25.3	20.5 ± 11.9	80.2 ± 1.17	16.8 ± 2.6	120
Nitrate	3.34 ± 0.81	5.16 ± 1.34	4.20 ± 0.34	3.03 ± 0.22	45
Sulfate	95.5 ± 20.7	117.1 ± 8.7	178.0 ± 4.47	154.7 ± 30.0	250
Phosphate	0.83 ± 0.51	1.33 ± 0.08	0.35 ± 0.04	1.60 ± 0.31	0.1
Fluoride	12.96± 0.9	15.5 ± 1.8	11.6 ± 0.42	12.9 ± 0.93	1.5
Potassium	74.3 ± 7.9	76.1 ± 8.3	89.0 ± 2.5	66.9 ± 14.5	-
Magnesium	25.6 ± 4.5	25.9 ± 2.7	23.0 ± 1.3	26.2 ± 2.1	30
Iron	0.48 ± 0.19	0.36 ± 0.31	0.24 ± 0.07	0.66 ± 0.12	0.3
Lead	0.05 ± 0.02	0.11 ± 0.04	0.04 ± 0.005	0.14 ± 0.01	0.01
Chromium	0.54 ± 0.10	0.23 ± 0.04	0.24 ± 0.02	0.29 ± 0.03	0.05
Manganese	0.22 ± 0.16	0.18 ± 0.08	0.15 ± 0.008	1.69 ± 0.16	0.1
Copper	6.6 ± 6.2	2.3 ± 0.86	15.2 ± 0.78	5.60 ± 1.98	2
Zinc	3.49 ± 2.97	5.36 ± 1.58	17.4 ± 0.17	0.80 ± 0.39	5
All units except temperature (°C), Turbidity (NUTs) and SC (µS/cm) are in mg/L.					

In Cluster A, the highest values of temperature (25.4 °C) and chromium (0.54 mg/L) as well as the lowest values of turbidity (15.9 NTUs), BOD_5 (27.1 mg/L), and sulfate (95.5 mg/L) were recorded. However, except for the concentration of sulfate, the concentrations of the other parameters were higher than the limit of the WHO standards [36]. These are likely to be linked to waste discharge, failing septic systems, urban stormwater, land development, and deposition of excessive amounts of organic and inorganic matter. Other studies [37, 38, 39] related to these sites also indicated higher BOD5 and turbidity.

In Cluster B, the highest values of pH (9.15), SC (1983.8 µS/cm), nitrate (5.16 mg/L), and fluoride (15.5 mg/L), were sewage, and industrial waste from ceramics, dyes, plastics and food processing industries) or overland runoff. Other studies related to these sites [39] also showed that the value of fluoride was higher than the WHO standard.

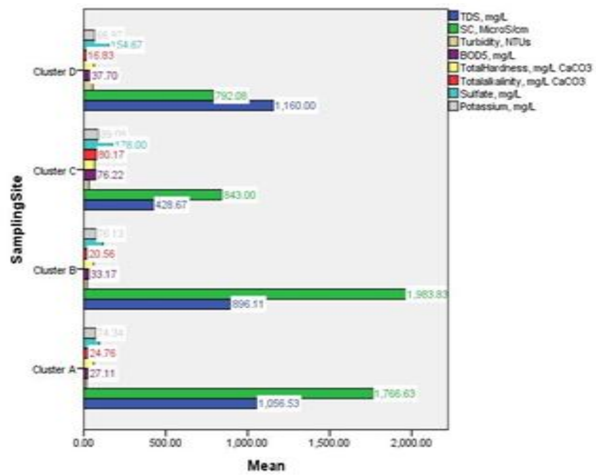

Figure 3a. Mean concentration (mean ± SD in mg/L) of selected water quality parameters for the clusters.

The highest values of BOD5 (76.2 mg/L), total hardness (67.2 mg/L CaCO3), total alkalinity (80.2 mg/L CaCO3), sulfate (178 mg/L), potassium (89.0 mg/L), copper (15.2 mg/L), and zinc (17.4 mg/L) were recorded in Cluster C. The lowest values recorded were for TDS (428.7 mg/L), temperature (21.2 °C), DO (2.39 mg/L), phosphate (0.35 mg/L), fluoride (11.6 mg/L), magnesium (23.0 mg/L), iron (0.24 mg/L), lead (0.04 mg/L), and manganese (0.15 mg/L). However, the values of BOD_5, copper, zinc, phosphate, and fluoride were found to be higher than the limit of the WHO standards, and the DO value was below the limit of the WHO standards.

These might be due to effluents from pulp and paper mills, food-processing plants, solid-waste incinerators, drained wetlands, runoff from agricultural and developed landscapes, household products, and pharmaceuticals. Other studies [37, 39] related to these sites also indicated the highest turbidity, BOD_5, fluoride, manganese and phosphate.

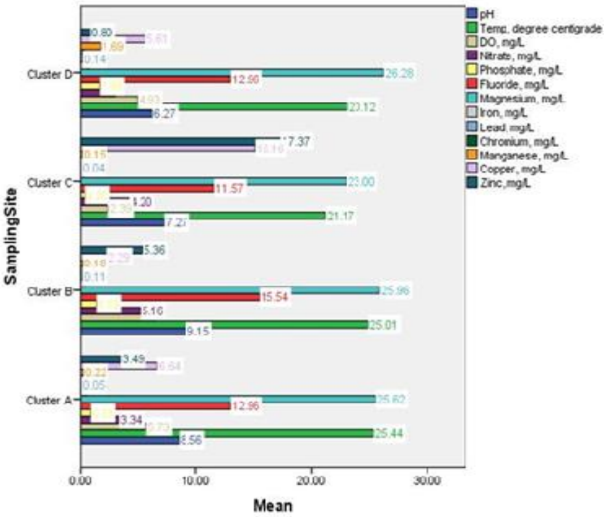

Figure 3b.Mean concentration (mean ± SD in mg/L) of selected water quality parameters for the clusters.

In Cluster D, the highest values of TDS (1160.0 mg/L), turbidity (55.8 NTUs), phosphate (1.60 mg/L), magnesium (26.2 mg/L), iron (0.66 mg/L), lead (0.14 mg/L), and manganese (1.69 mg/L) were recorded. The lowest values recorded were for pH (6.27), SC (792.1 µS/cm), total alkalinity (16.8 mg/L $CaCO3$), nitrate (3.03 mg/L), potassium (66.9 mg/L), and zinc (0.80 mg/L).

However, the values of TDS, turbidity, phosphate, iron, lead, manganese, and SC were found to be higher than the limit of the WHO standards. These might be due to the pollution from deposition of excessive amounts of organic and inorganic matter, industrial waste discharges containing iron, erosion from catchment areas after storm events, land development, and agricultural and urban runoff. In a similar study [37, 38, 39], the highest values of BOD_5, TDS, turbidity, fluoride, nitrate and phosphate were reported, which are similar with the recent findings.

Factor scores for the four clusters were calculated to determine the level of pollution, which are given in Table 5.

Table 5. Factor scores of the four clusters for an eigenvalue greater than the absolute value of 0.3.

Parameters	First Rotated Component Matrix			
	Cluster A	Cluster B	Cluster C	Cluster D
REGR factor score 1	-0.618	**0.938**	-0.595	**0.982**
REGR factor score 2	0.537		**0.9**	**0.984**
REGR factor score 3		**0.932**	**0.827**	**0.994**
REGR factor score 4	**-0.936**	**0.961**	**-0.787**	**0.963**
REGR factor score 5	0.427	-0.561		**0.981**
REGR factor score 6	**0.972**	**-0.941**	**0.915**	**-0.761**
REGR factor score 7				-0.388
Extraction Method: Principal Component Analysis				
Rotation Method: Varimax with Kaiser Normalization				

Cluster A was highly loaded with PC4 and PC6, and these components showed the presence of a high level of pollution from industrial effluents and agricultural runoff. Both Cluster B and Cluster C were highly loaded with PC3 and PC6. Cluster B was also highly loaded with PC1 and PC2 while Cluster C was loaded with PC2. These indicated the presence of a high level of pollution from industrial effluents, agricultural runoff, domestic wastewaters, land development, and urban runoff. Cluster D was highly loaded with all component loadings except for PC7. Therefore, this cluster was affected by multiple sources of pollution.

Conclusions

Multivariate statistical methods such as PCA and HCA are vital to assess water quality parameters and identify potential sources of pollution. The current study examined the water quality parameter data obtained from different sampling sites in Lake Hawassa. The paper suggests that four sampling sites in the Lake are adequate to evaluate and monitor the pollution status of the lake. In general, the lake water is unsuitable for drinking and it has faced many threats from point and nonpoint pollution sources, such as waste water disposal (human sewage, and industrial waste from ceramics, dyes, plastics and food processing industries), failing septic systems, urban stormwater, agricultural runoff, and land development.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1]International Network of Basin Organizations (INBO) and the Global Water Partnership (GWP),"The handbook for integrated water resources management in transboundary basins of rivers, lakes and aquifers," ISBN 978-91-85321-85-8, 2012.

[2]R C Hart, "Ethiopian Rift Valley Lakes," African Journal of Aquatic Science, Vol. 28, no. 1, pp 85-85, 2003.

[3]K. Gebremedhin, and T. Berhanu, "Determination of some selected heavy metals in fish and water samples from lake Hawassa and Ziway Lake," Journal of Analytical Chemistry, Vol. 3, p10-16, 2015.

[4]G.M. Zinabu and D. Zerihun,"The Chemical Composition of the Effluent from Awassa Textile Factory and its Effects on Aquatic Biota,"Ethiopian Journal of Science, Vol. 25, no. 2, 2002.

[5]G.M. Zinabu, K. Elizabeth, and D. Zerihum,"Long-term changes in the Chemical and Biological Features of Seven Ethiopian Rift Valley Lakes," Hydrobiologia, Vol. 477, no. 1-3, pp 81- 91, 2002.

[6]Y. D. Abebe and K. Geheb, "Wetlands of Ethiopia," Proceedings of a seminar on the resources and status of Ethiopia's wetlands, vi + 116pp, 2003.

[7]E. Gugissa, "Urban Environmental impacts in the town of Awassa Ethiopia," Addis Ababa University, M Sc Thesis, Ethiopia, 2004.

[8]G. Katie, "Environmental Policy Review: Lake Water Management in three Ethiopian Rift Valley Watersheds," 2011.

[9]Y. Zenebe, "Accumulation of certain heavy metals in Nile Tilapin (oreochromisniloticus) Fish species Relative to heavy metal concentrations in the water of Lake Hawassa," Addis Ababa University, M Sc Thesis, Ethiopia, 2011.

[10]D. Larissa, M. Mesfin, and D. Elias, "Assessment of heavy metals in water samples and tissues of edible fish species from Awassa and Koka Rift Vally Lakes, Ethiopia," Environ. Monit. Assess, Vol. 185, no. 4, 2012.

[11]T. Girma, and G. Ahlgren,"Seasonal Variations in Phytoplankton biomass and primary production in the Ethiopian Rift Valley lakes: Ziway, Awassa and Chamo-The basis of fish production,"Elsevier Sc. Limnologca, Vol. 40, pp 330-342, 2009. (2010).

[12]A. Papaioannou, A. Mavridou, and C. Hadjichristodoulou," Application of multivariate statistical methods for groundwater physicochemical and biological quality assessment in the context of public health," Environ. Monit. Assess, Vol. 170, pp 87, 2010.
[13]A. Ene, N. Tigau, M. Praisler, and L. Moraru, "Principal component analysis of physico-chemical parameters of river water, "Annals Univ. Dunarea de Jos Galati, Fasc. II., Vol. 33, no. 1, pp 130, 2010.
[14]V. Lovchinov, and S. Tsakovski, "Multivariate statistical approaches as applied to environmental physics studies," Cent. Eur. J. Phys., Vol. 4, no. 2, pp 277, 2006.
[15]H. Boyacioglu, and H. Boyacioglu, "Surface water quality assessment by environmetric methods," Environ. Monit. Assess, Vol. 131, pp 371, 2007.
[16]V. Simeonov, J.A. Stratis, and C. Samara, "Assessment of the surface water quality in Northern Greece," Water Res., Vol. 37, pp 4119, 2003.
[17]C Güler, G.D Thyne, J.E McCray, and K.A Turner,"Evaluation of graphical and multivariate statistical methods for classification of water chemistry data," Hydrogeology Journal, Vol. 10, pp 455-474, 2002.
[18]V Cloutier, R Lefebvre, R. Therrien, and M.M Savard, "Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system,"J. Hydrol, Vol. 353, pp 294-313, 2008.
[19]M Templ, P Filzmoser, and C Reimann, "Cluster analysis applied to regional geochemical data: problems and possibilities," Appl. Geochem, Vol. 23, pp 2198-2213, 2008.
[20]I .M. Farnham, A. K Singh, K..J. Stetzenbach, and K. H. Johannesson, "Treatment of nondetects in multivariate analysis of groundwater geochemistry data," Chemometr Intell. Lab. Syst, Vol. 60, no. 1, pp 265–281, 2002.
[21]R. A. Johnson and D W. Wichem, "Applied multivariate statistical analysis," London,Prentice-Hall, 2002.
[22] C.R Alvin, "Methods of Multivariate Analysis," Wiley Interscience, 2002.
[23]M. J. Norusis, "Inc. SPSS, IBM SPSS Statistics 20. Statistical Procedures," Companion, Prentice Hall, Englewood Cliffs, 2011
[24] F J M Bryan, "Multivariate statistical method: A primer," London, Chapman and Hall, 1991.
[25]U. C Panda, S. K Sundaray, P Rath,., B. B Nayak, and D. Bhatta, "Application of factor and cluster analysis for characterization of river and estuarine water systems - a case study: Mahanadhi River (India)," Journal of Hydrology, Vol. 331, no. 3-4, pp 434-445, 2006.
[26]J. C Davis, "Statistics and Data Analysis in Geology (2nd Ed.)," Wiley, New York, 1986.
[27]T. Spanos, V. Simeonov, P. Simeonova, E. Apostolidou, and J. Stratis, "Environmetrics to evaluate marine environment quality," Environ. Monit. Assess, Vol. 143, no. 215, 2008.
[28]G Brumelis, L Lapina,., O Nikodemus, and G Tabors, "Use of an artificial model of monitoring data to aid interpretation of principle component analysis," Environmental Modeling & Software, Vol. 15, no. 8, pp 755-763, 2000.
[29]Love, D., Hallbauer, D., Amos, A., and R Haranova,"Factor analysis as a tool in ground water quality management: two southern African case studies," Physics and Chemistry of the Earth, Vol. 29, pp 1135-1143, 2004.
[30]S. A Abdul-Wahab, C. S Bakheit, and M Al-Alawi, "Principle component and multiple regression analysis in modeling of ground-level zone and factors affecting its concentrations," Environmental Modeling &Software, Vol. 20, no.10, pp1263-1271, 2005.
[31]J. H Kim, C. M Choi, S. B Kim, and S. K Kwun, "Water quality monitoring and multivariate statistical analysis for rural streams in South Korea," Paddy Water Environ, Vol. 7, pp197-208, 2009.
[32]C. W Liu, K. H Lin, and Y. M Kuo, "Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan," Science of the Total Environment, Vol. 313, pp 77-89, 2003.
[33]H. F. Kaiser,"A Second-Generation Little Jiffy," Psychometrika, Vol. 35, no. 4, pp 401-415, 1970.
[34]M. S. Bartlett,"Tests of significance in factor analysis," British Journal of Psychology, pp77-85, Vol. 3, no. 2, 1950.
[35]A. P. field, "Discovering Statistics using SPSS (2nd Ed edition),"London. New Delhi: Sage publications, 2005.
[36]World Health Organization,"Guidelines for Drinking water Quality, Fourth Edition," ISBN 978 92 4 154815 1, 2008.
[37]B. Abate1, A. Woldesenbet and D. Fitamo, "Water quality assessment of lake Hawassa for multiple designated water uses," Water Utility Journal 9, pp 47-60, 2015.
[38]Praveen and Mukemil, "CCME Water Quality Index and Assessment of Physiochemical Parameters of Lake Hawassa, Ethiopia," International Journal of Recent Scientific Research, Vol.6, Issue 6, pp. 7891-7894, 2015.
[39]W.W. Adimasu, "Evaluation of the water quality status of Lake Hawassa by using water quality index, Southern Ethiopia," International Journal of Water Resources and Environmental Engineering, Vol. 7, no. 4, pp. 58-65, 2015.