RESEARCH ARTICLE

European Journal of Environment and Earth Sciences
www.ej-geo.org

Spatio-Temporal Water Quality Assessment and Pollution Source Apportionment of Lake Chamo using Water Quality Index and Multivariate Statistical Techniques

Solomon Ghebrehiwet Ghebremedhin and Bhaskar Sen Gupta

ABSTRACT

Despite the current management effort to reduce pollution of water bodies, lakes are far more vulnerable to pollution than the ocean and rivers. Point and non-point pollution are ending on the lake bodies and concerned bodies should trigger interventions. This study aims to assess the effect of pollution on the water quality of Chamo lake using the weighted arithmetic water quality index (WQI) and multivariate statistical methods. Fifteen monitoring points are used to collect water samples from the lake for dry and rainy periods. Twenty-three physical-chemical parameters were analyzed in the water quality laboratory of Arba Minch University, out of which five parameters were analyzed using a multi-meter at the field. The physicochemical water quality analysis result of Lake Chamo revealed that the constituent's concentration did not show significant variation spatially. The factor analysis result in the Lake Chamo revealed that it had extracted seven principal components that can be explained with a total variance of 86.3%. More than 65% of the monitoring stations in Lake Chamo have an acceptable water quality for irrigation. Moreover, some of the monitoring stations, closer to high human activities, are relatively more polluted than others. Besides, some sensitive parameters such as turbidity have clearly shown that the WQI value can be changed with the parameters changing. Based on clustering analysis, most (47%) of the monitoring are characterized by the recipients of uncontrolled public and state farm and agricultural wastes. The study recommends integrated watershed management and source-based pollution control, which could significantly decrease the pollution level in Lake Chamo after identifying the pollution hotspot areas. Continuous Lake water quality monitoring is necessary for deciding lake pollution.

Keywords: Chamo Lake, Cluster Analysis, Monitoring Stations, Multivariate Statistics, WQI.

ISSN: 2684-446X

DOI: 10.24018/ejgeo.2023.4.1.368

S. G. Ghebremedhin *
Freelance WASH Consultant, Eritrea
(e-mail: Sgwot@yahoo.com)
B. S. Gupta
Professor in Water Technology, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh Campus, UK

**Corresponding Author*

I. INTRODUCTION

When it comes to contamination, surface water sources are more susceptible than underground ones. Lakes are far more vulnerable to pollution than the ocean and rivers since they are close to towns and human activities, such as fishing. Silt and organic waste accumulate over time, affecting the lake's depth, biological productivity, oxygen levels, and clarity. Human activity and natural occurrences may accelerate environmental changes in even the most stable lakes. Discharges of industrial, municipal, and agricultural trash into lakes and other waterways considerably impact lake water quality owing to human activity (intentional or accidental). Surface water resources' quality and long-term viability were significantly impacted by geological, hydrological, and climatic factors in the earth's environment. The ecology was affected by any substantial changes in water quality.

Arba Minch's commercial fishing relies on the waters of Lake Chamo. Recreational and tourist attractions might also be found here. Because of runoff and silt from rain, lake Chamo has become severely contaminated. Resorts and other companies dump waste into the lake, contaminating it. Fish carcasses are left in the lake to rot after the people have caught them. As the aquatic plant population proliferated, it finally filled the lake. Despite these concerns, this lake is nevertheless vulnerable to contamination from urbanization, agricultural intensification, increasing industry, urban runoff, and natural processes like erosion. Chamo Lake's pollutants and impacts were investigated in this project to see how far they can be measured. From observation and experience, it is learned that most municipal sewerage plants and industries in the area discharge partially treated or untreated wastewater containing high levels of organic, metals, and other toxic substance directly into surface watercourses.

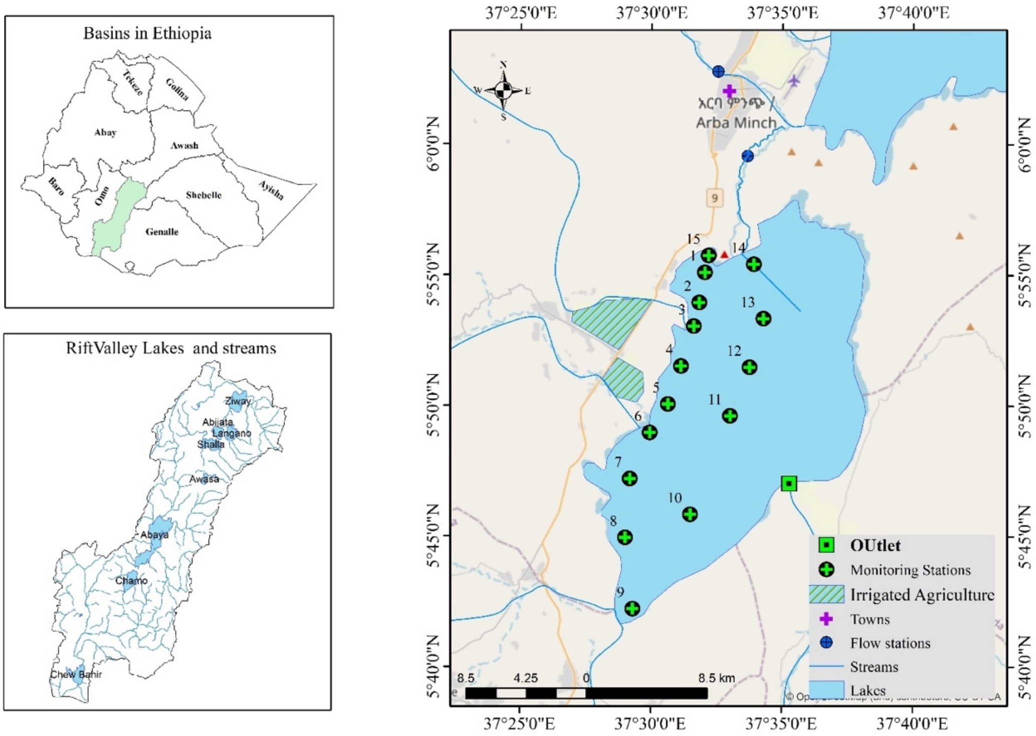

Fig. 1. Location map of the study area and proposed water sample points.

In addition, open sewages prevailing in the sub-basin are currently contaminating groundwater resources. It is, therefore, critically important to incorporate the planning and management of pollution control systems by identifying, controlling, and regulating point and non-point sources of pollution to safeguard the quality of the nation's water resources.

This study aims to assess the effect of pollution on the water quality of Chamo lake. The specific objectives of the study are:

a) To characterize the spatio-temporal water quality characteristics of Chamo Lake based on the assessment of physico-chemical and nutrient constituents.
b) To investigate the potential pollution sources contributing to the pollution of Lake Chamo using multivariate statistical analysis.
c) To discover the pollution status and evaluate the suitability of Lake Chamo for drinking and aquatic life based on the application of weighted arithmetic WQI.

A. Research Questions

i. What are the limiting physico-chemical and nutrient constituents that play great role in characterizing Chamo Lake water quality status?
ii. What is the main source of pollution in Chamo Lake?
iii. Can we use Chamo Lake for our daily activities based on the current status?

II. METHODS AND DATA

A. Description of the Study Area

1) Location of the study area

Lake Chamo is a well-known lake for fish production in Ethiopia. It is located in the main Ethiopian Rift valley basin. Astronomically the study region is found in the Southern Nations Nationalities and People's Regional (SNNPR) State of Ethiopia with an elevation of 1,110m. The lake is located south of Lake Abaya and Arba Minch city, east of the Guge Mountains and west of the Amaro Mountains. Geographically, the study region is bounded by a latitude of 5˚40'27"-5˚58'13" N and a longitude of 37˚25'16"-37˚30'38" E by covering a total area of 329 km^2 (Fig. 1).

2) Perennial and intermittent tributaries of the lake

The catchment surface area of Lake Chamo is 1109 km^2. Likewise, the total surface area of the Lake is 329 km^2.

However, based on the satellite images captured in 2022, the total surface area of the lake increased to 341 km^2, which indicates the area of the lake region is increasing. Rivers such as Kulfo, Sile, Elgo, Argoba, Wezeka, Gato and Sego are the major rivers that feed the Lake.

3) Agriculture practices and lithology of the lake catchment

The flood plains and the deltas of Lake Chamo are fertile and hence have been under extensive agricultural cultivation [1]. Studies also discussed that most of the lake catchment areas are covered with agricultural lands used for growing crops such as wheat, corn, teff, coffee, sorghum and banana. The majority of the lithological unit found in the catchment region includes Pleistocene basalts, mostly vesicular, Rhyolitic Lava flows, lacustrine predominantly, Pumice & unwelded tuffs, Lacustrine sediments, silts, clay diatomites and ignimbrites [2].

4) Climate, and soil types of the area

The catchment of this lake is characterized by a humid to hot semi-arid tropical climate with a bimodal rainfall pattern including two wet seasons; the first is from end-March to mid-June.

TABLE I: LOCATION OF WATER QUALITY MONITORING STATION

Monitoring Station	Latitude (Deg)	Longitude (Deg)	Monitoring Station	Latitude (Deg)	Longitude (Deg)
1	5.92135343	37.5468043	9	5.82446523	37.50414358
2	5.90848568	37.54051829	10	5.81838655	37.52056188
3	5.89998545	37.53628304	11	5.84343644	37.53596167
4	5.89018705	37.52963262	12	5.87224896	37.54306385
5	5.88934835	37.52923277	13	5.88778831	37.54818043
6	5.86294764	37.52115627	14	5.909171	37.55297434
7	5.84822591	37.51563781	15	5.92913081	37.53634593
8	5.83798228	37.51235155	-	-	-

TABLE II: LABORATORY ANALYTICAL METHOD FOR WATER QUALITY PARAMETER CONSTITUENTS ON CHAMO LAKE

No	Parameter (unit)	Measurement location	Analytical method	Instrument
1	pH	Onsite	Multiparameter probe	HQ40D
2	DO (mg/L)	Onsite	Multiparameter probe	HQ40D
3	TDS (mg/L)	Onsite	Multiparameter probe	HQ40D
4	EC (µs/cm)	Onsite	Multiparameter probe	HQ40D
5	Water Temperature, °C	Onsite	Multiparameter probe	HQ40D
6	BOD (mg/L)	Laboratory	Modified Winkler	-
7	COD (mg/L)	Laboratory	Open reflex	-
8	Phosphate (mg/L)	Laboratory	Spectrophotometer	UV-VIS spectrophotometer
9	Nitrate (mg/L)	Laboratory	Spectrophotometer	UV-VIS spectrophotometer
10	Nitrite (mg/L)	Laboratory	Spectrophotometer	DR-2800
11	Total Phosphorus (mg/L)	Laboratory	Spectrophotometer	UV-VIS spectrophotometer
12	Total Suspended Solids (mg/L)	Laboratory	-	-
13	Other	Laboratory	-	-

While the second one is from mid-September to late November; and two dry seasons the first from December to mid-March, the second is from end-June to mid-September [8], [14]. The catchment was dominantly covered with soil textural class of loam and sandy loam.

B. Water Quality Monitoring and Analytical Techniques

In order to characterize the lake water quality and determine the pollution status, a monthly water quality sample is collected from 15 monitoring stations during both the dry and wet seasons (Table I).

The sampling was conducted based on the purposive approach where both composite and grab sampling techniques were used based on factors such as accessibility, the impact of anthropogenic factors, and the location of pollution sources from the monitoring stations. A total of 23 water quality parameters were analysed for physical, chemical, and heavy metal constituents. Parameters such as dissolved oxygen (DO), Total Dissolved Solids (TDS), Electrical Conductivity (EC), Water Temperature (T), and pH were measured using a multiparameter probe as indicated on Table II.

The water quality samples were collected using polyethylene and glass bottles, placed in a cooler box, kept under 4 °C, and immediately transported to the laboratory for analysis. Pretreatment of all sample containers was done with a mixture of distilled water and dilute hydrochloric acid. The sampling bottles were washed with distilled water 24 hours before sample collection and rinsed three times with the sample water before collection. All the analytical techniques for all of the parameters were according to the standard methods for the examination of water and wastewater [17]. The analytical techniques and instruments used for the analysis are shown in Table II.

1) Weighted arithmetic water quality index

The calculation of the WQI was done using a weighted arithmetic water quality index, which was originally proposed by Horton [6] and developed by Brown *et al.* [5]. Equation (1) gives the weighted arithmetic water quality index (WQI_A).

$$WQI_A = \frac{\sum_{i=1}^{n} w_i q_i}{\sum_{i=1}^{n} w_i} \quad (1)$$

where n is the number of variables or parameters, w_i is the relative weight of the i^{th} parameter and q^i is the water quality rating of the i^{th} parameter. The unit weight(w_i) of the various water quality parameters is inversely proportional to the recommended standards for the corresponding parameters.

The value of q_i is calculated using (2) [5].

$$qi = 100\,[(Vi - Vid) / (Si - Vid)]$$

$$q = 100[\frac{V_i - V_{id}}{S_i - V_{id}}] \quad (2)$$

where V_i is the observed value of the i^{th} parameter, S_i is the standard permissible value of the i^{th} parameter and V_{id} is the ideal value of the i^{th} parameter in pure water.

All the ideal values (V_{id}) are taken as zero for drinking water except pH and dissolved oxygen [13]. For pH, the ideal value is 7.0 (for natural/pure water) and the permissible value is 8.5 (for polluted water). Therefore, the quality rating for pH is calculated from (3).

$$q_{pH} = 100[\frac{V_{pH} - 7.0}{8.5 - 7.0}] \quad (3)$$

where V_{pH} = the observed value of pH.

For dissolved oxygen, the ideal value is 14.6 mg/L and the standard permissible value for drinking water is 5 mg/L. Therefore, its quality rating is calculated from (4).

$$qDO = 100\,[(VDO - 14.6) / (5.0 - 14.6)]$$

$$q_{DO} = 100\left[\frac{V_{DO}-14.6}{5.0-14.6}\right] \qquad (4)$$

where V_{DO} = observed value of dissolved oxygen.

2) *WQI Index value classification*

The category value ranges from 0 to 100. The lower the index value, the better the water quality. Table III shows a classification of water quality, based on its quality index [5], [18].

C. *Multivariate Statistical Methods*

Miura *et al.* [19] recommend two/three multivariate analysis techniques (MSTs) to statistically analyze lake water quality discussed as follows. Different MSTs have been used for analyzing water quality data because they are capable of treating numerous data from a variety of monitoring sites. In this study, cluster analysis and principal component analysis (PCA were applied.

1) *Cluster analysis*

This method classifies objects with clusters with respect to already set selection criteria. In this study, hierarchical agglomerative cluster analysis was carried out using squared Euclidean distances as a means of measure of similarity.

2) *Principal component analysis*

PCA uses the idea of dimensionality reduction and several comprehensive variables to obtain major information based on the original variables without any overlap [20]-[22]. In this study, PCA was used to identify the main components and sources in different seasons, which comprehensively reflect the lake's level of pollution [23].

The statistical correlation among water quality parameters was summarised using a correlation matrix of chemical components; and standardized analytical data [10], [3].

TABLE III: CLASSIFICATION OF WATER QUALITY BASED ON WEIGHTED ARITHMETIC WQI METHOD

Rank	WQI Value	Interpretation
Excellent	0 to 25	A water quality condition is unharmed, and it is very close to natural or desired levels. These index values can only be obtained if all measurements comply with the guidelines almost all the time.
Good	26 to 50	Water quality is intact and only one minor threat or deterioration is observed; conditions rarely differ from the natural or desirable levels.
Poor	51 to 75	Water quality is usually intact, but occasionally endangered or deteriorated; conditions sometimes deviate from the natural or desirable levels.
Very Poor	75 to 100	Water quality is frequently endangered or deteriorated; conditions often deviate from the natural or desirable levels.
Unsuitable for Drinking	>100	Water quality is almost always endangered or deteriorated; conditions usually deviate from natural or desirable levels.

TABLE IV: PHYSICO-CHEMICAL WATER QUALITY CONSTITUENTS IN THE LC

MSs	pH	T	DO	Turb.	EC	TDS	TH	COD	BOD	NO_3	NO_2	NH_3	PO4	TA	TSS
MS1	8.95	26.4	6.505	48.5	1397.5	705	92	128	26.183	9.95	4.6	4.7	0.46	538	63.35
MS2	8.99	25.4	6.725	52	1374	684.5	94	144	35.465	24.18	6.7	6.7	0.46	530	74.55
MS3	9.00	24.85	6.73	59.5	1374	683	110	128	25.346	18.39	5.65	6.1	0.47	536	79.45
MS4	8.99	25.95	6.52	50	1363.5	678.5	114	160	28.53	19.72	7.75	11.6	3.45	528	63.85
MS5	8.99	26.05	6.78	48.5	1367.5	680	92	128	28.251	14.68	4.85	5.9	3.45	534	54.45
MS6	9.00	25.85	6.565	43	1364	679.5	104	160	30.758	12.68	2.35	7.4	0.54	540	60.75
MS7	8.99	25.75	6.48	47	1383.5	680.5	90	144	29.287	22.36	7.3	11.9	0.49	550	55.05
MS8	8.95	25.85	6.615	49.5	1365	676.5	100	112	31.161	23.84	7.9	12.4	0.48	540	64.5
MS9	8.96	26.4	6.495	49	1360	676.5	92	128	35.199	14.79	6.55	11.7	0.68	550	52.3
MS10	8.99	28.3	6.56	50	1365	679	90	144	31.146	28.02	8.55	13.4	0.66	530	57.5
MS11	8.98	26.75	6.78	43	1364	678.5	92	144	27.002	23.74	10.2	13.0	0.96	540	47.35
MS12	8.99	27.45	6.985	43	1363	678	94	112	30.905	19.26	5.5	8.9	0.68	526	51.55
MS13	9.00	27.05	7.25	40	1367.5	677	100	144	50.376	17.68	6.2	8.3	0.59	534	51.05
MS14	8.98	27.25	7.055	45	1366.5	680	144	112	39.974	19.88	4.55	8.2	0.79	530	53.7
MS15	8.93	27.75	6.72	49	1429	712	102	144	31.612	20.69	5.15	8.9	5.44	534	56.15

MSs: Monitoring station; all units in mg/L except pH (unitless), T (°C), turbidity (NTU) and EC (µS/cm).

III. RESULTS AND DISCUSSION

A. *Hydrochemistry of Lake Chamo*

The Physico-chemical water quality analysis of Lake Chamo (LC) revealed that the constituent's concentration did not show significant variation spatially (Table IV). Accordingly, the pH showed slight variation and followed a similar trend across all the monitoring stations. However, it has shown a slight increment at some of the monitoring stations, potentially due to the photosynthetic activities of algae that consume carbon dioxide dissolved in the water. Khettaf *et al.* [24] also use a similar interpretation. A mean (standard deviation) pH concentration of 8.98±0.02 is slightly higher than the standard national guideline where the LC water is characterized by its alkaline nature.

DO is the essential water quality constituent that maintains the equilibrium of aquatic life and ecosystems [25]. In LC, all the monitoring stations have met the minimum DO standard in the lake. Similar to pH, the DO showed slight variation across the monitoring station. A minimum concentration of 6.49 mg/L of concentration was recorded at monitoring station 7, which is close to the upstream state farm outlet. In addition, the maximum DO concentration was also registered at the monitoring station 13, where minimum anthropogenic influence was observed. A mean DO concentration of 6.72±0.22 mg/L in LC could indicate that the water resource is relatively suitable for aquatic life despite the recent threat in the watershed.

The EC, a proxy measurement of salt concentration in water resources, is one of the indicator parameters in the evaluation of surface and groundwater hydrochemistry [26]. Although the monitoring stations have shown closer concentrations in LC, higher values are recorded. Accordingly, a relatively higher EC concentration of 1429 ± 106.1 µS/cm was recorded at the monitoring station 15, located at the public inlet to the LC. In addition, the TDS of the LC also showed a similar trend to the EC, having a mean concentration of 683.2±55.15 mg/L. Higher values of EC and TDS could be attributed to anthropogenic activities such as intensified agriculture in the LC vicinity and natural process. A very high positive correlation (r=0.95) was observed between TDS and EC (Table IV).

The high presence of organic pollutants such as BOD and COD in a surface water resource are good indicators of the pollution level in the water. Mishra *et al.* [27] state that water resources with a BOD concentration of more than 15 mg/L are regarded as heavily polluted. However, in the LC water quality monitoring stations, the BOD ranged from 25.34±3.51 mg/L-50.3±0.2 mg/L, indicating a higher organic pollution level in the Lake (Table IV). Similarly, the COD in the LC has recorded a higher concentration with a mean value of 135.5±22.62 mg/L, slightly higher than the national guideline standard [28].

Optimum nutrient concentrations are one of the requirements for aquatic plants to grow. The excessive presence of such contaminations could lead to uncontrolled production of the plants and eutrophication in the water resources [26]. In addition, recent agricultural intensification and increased anthropogenic influences in the LC are becoming a threat. Accordingly, although many of the monitoring stations have a PO_4 concentration within the guideline limit, station 15, which is located at the entrance to the lake, has recorded a relatively higher concentration (5.44 mg/L) than other monitoring stations were some signs of eutrophication in the area being observed in recent days (Table IV). Similarly, monitoring stations 4 and 5 have also registered higher phosphate concentrations where the areas are located near the outlet to the state farm drainage. The mean phosphate concentration of 1.31 mg/L and an increased and intensified agriculture in the area could trigger eutrophication in the LC. The Nitrate concentration in the LC was found to be within the national guideline standard across all monitoring stations [29]. Although irregularities in nitrate concentration were observed between the monitoring stations, they are all below the standard limit of 50 mg/L. For example, a higher concentration of 10.2±12.4 mg/L was observed at the monitoring station 11, and a lower concentration was recorded at monitoring station 6 with a mean concentration of 2.35±2.33 mg/L. On the other hand, ammonia concentration was found to have a mean value of 9.3±4.98 mg/L, which is within the national guideline standard.

B. Spatial Distribution of Cation and Anion Concentrations of LC

Hardness (TH) in water resources is a significant property that determines the applicability of water for domestic consumption [30]. It is mainly determined due to calcium and magnesium in the water and is often expressed as calcium carbonate. In surface and groundwater resources, the harness is not such a health problem, but long-term exposure could lead to complexities. Accordingly, in LC, the TH concentration in some of the monitoring stations have recorded to be slightly higher than the guideline limit (100 mg/L). The mean TH concentration of 100.7±14.9 mg/L could indicate that the water resources are relatively excellent and classified as moderately hard water. Accordingly, the Ca^{2+} and Mg^{2+} have remained the dominant cations controlling the LC's TH. A very high correlation between TH and Mg^{2+} (r=0.94) could reveal that the hardness in LC was mainly attributed to the high presence of Mg^{2+} in the Lake.

Similarly, sodium concentration in the LC for all the monitoring stations was within the guideline limit, whereas potassium concentration surpassed this. The mean concentration of K on all monitoring stations in LC was found to be 19.7±1.06 mg/L, slightly higher than drinking water's standard limit. On the other hand, the chloride concentration in the LC was below the national guideline standard, having a mean concentration of 86.4±4.6 mg/L. in addition, the sulfate concentration in LC was under the desirable limit with a mean concentration of 15.5±9.2 mg/L. In LC, the mean concentrations of CO_3^-, HCO_3, TSS, and turbidity were 141 mg/L, 394.93 mg/L, 59.4 mg/L, and 47.8 NTU, respectively (Table V).

TABLE V: SPATIAL DISTRIBUTION OF CATION AND ANION CONCENTRATIONS OF LC

MSs	Ca	Mg	Na	K	Cl	SO_4	CO_3	HCO_3
MS1	16.8	12	350.5	16.45	86.97303	15.6855	116	422
MS2	14.4	13.92	343.75	17.1	92.97117	15.55405	176	354
MS3	13.6	18.24	358	16.9	85.97334	14.24693	144	392
MS4	17.6	16.8	351.125	18.3	85.97334	14.1855	148	380
MS5	16	12.48	379.875	18.45	84.97365	14.6855	156	378
MS6	12	17.76	361.5	18.1	75.97644	15.74693	140	400
MS7	14.4	12.96	365.5	19.8	97.96962	13.74693	144	406
MS8	16	14.4	390.75	19.55	87.97272	13.1941	140	400
MS9	13.6	13.92	352.875	19.65	83.97396	15.58968	140	410
MS10	15.2	12.48	363.5	20.4	87.97272	14.48403	136	394
MS11	15.2	12.96	410.25	20.35	85.97334	14.54545	124	416
MS12	12.8	14.88	371.75	19.75	83.97396	13.60688	140	386
MS13	12	16.8	383.25	20.7	88.97241	27.62899	136	398
MS14	14.4	25.92	404	24.75	83.97396	13.86978	152	378
MS15	9.6	18.72	386.75	25.35	82.97427	15.6683	124	410

MSs: Monitoring station; all units in mg/L except pH (unitless), T (°C), turbidity (NTU) and EC (µS/cm).

TABLE VI: Bivariate Correlation in the LC Water Quality Constituents

	pH	T	DO	Turb.	EC	TDS	TH	Ca	Mg	Na	K	Cl	SO$_4$	COD	BOD	NO$_3$	NO$_2$	NH$_3$	PO$_4$	TA	CO$_3$	HCO$_3$
pH	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T	-0.23	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DO	0.33	0.32	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Turb.	-0.15	-0.48	-0.46	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
EC	-0.61	0.18	-0.14	0.18	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TDS	-0.65	0.19	-0.17	0.20	**0.95**	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TH	0.13	0.01	0.37	0.02	-0.08	-0.06	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ca^{2+}	0.03	-0.29	-0.37	0.24	-0.41	-0.28	-0.03	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mg^{2+}	0.11	0.11	0.47	-0.07	0.07	0.04	**0.94**	-0.37	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
Na^{+}	-0.14	0.41	0.54	-0.48	0.00	-0.09	0.31	-0.17	0.35	1.00	-	-	-	-	-	-	-	-	-	-	-	-
K^{+}	-0.31	0.69	0.40	-0.36	0.34	0.21	0.44	-0.50	0.58	0.68	1.00	-	-	-	-	-	-	-	-	-	-	-
Cl-	0.05	-0.20	-0.08	0.19	0.07	-0.08	-0.31	0.30	-0.39	-0.16	-0.15	1.00	-	-	-	-	-	-	-	-	-	-
SO$_4$	0.25	0.16	0.57	-0.44	0.01	-0.03	-0.07	-0.37	0.06	0.06	0.07	0.06	1.00	-	-	-	-	-	-	-	-	-
COD	0.30	-0.10	-0.32	-0.08	0.11	0.07	-0.20	-0.10	-0.15	-0.32	-0.16	0.00	0.26	1.00	-	-	-	-	-	-	-	-
BOD	0.21	0.31	0.68	-0.49	-0.16	-0.23	0.27	-0.39	0.39	0.22	0.41	0.08	**0.78**	-0.04	1.00	-	-	-	-	-	-	-
NO$_3$	0.05	0.26	0.05	0.14	-0.09	-0.25	-0.04	0.05	-0.05	0.27	0.31	0.48	-0.21	0.05	0.05	1.00	-	-	-	-	-	-
NO$_2$	-0.08	0.11	-0.14	0.08	-0.24	-0.31	-0.31	0.40	-0.43	0.20	0.01	0.53	-0.10	0.10	-0.11	**0.73**	1.00	-	-	-	-	-
NH$_3$	-0.12	0.32	-0.26	-0.16	-0.28	-0.41	-0.16	0.15	-0.20	0.26	0.31	0.22	-0.20	0.14	-0.02	0.66	0.77	1.00	-	-	-	-
PO$_4$	-0.36	0.22	-0.05	0.10	0.55	0.49	0.08	-0.18	0.14	0.15	0.41	-0.24	-0.08	0.26	-0.16	-0.03	-0.09	-0.04	1.00	-	-	-
TA	-0.24	-0.34	-0.48	-0.01	0.04	-0.06	-0.35	-0.05	-0.31	-0.04	-0.11	0.18	-0.04	0.06	-0.14	-0.22	0.09	0.25	-0.26	1.00	-	-
CO$_3$	0.48	-0.43	0.12	0.27	-0.40	-0.45	0.22	0.15	0.15	-0.27	-0.20	0.26	-0.13	-0.02	0.20	0.23	-0.07	-0.17	-0.09	-0.26	1.00	-
HCO$_3$	-0.49	0.21	-0.29	-0.22	0.34	0.34	-0.32	-0.14	-0.25	0.20	0.12	-0.14	0.09	0.04	-0.21	-0.28	0.09	0.24	-0.03	0.62	**-0.92**	1.00
TSS	0.08	-0.66	-0.31	**0.81**	0.10	0.15	0.11	0.18	0.04	-0.58	-0.56	0.14	-0.22	0.06	-0.31	0.02	-0.14	-0.40	-0.13	-0.16	0.39	-0.38

TABLE VII: Total Variance Explained during FA in LC

PC	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %	Total	% of Variance	Cum. %
1	4.96	21.56	21.56	4.96	21.56	21.56	3.76	16.36	16.36
2	4.15	18.05	39.61	4.15	18.05	39.61	3.54	15.38	31.74
3	3.67	15.97	55.57	3.67	15.97	55.57	3.00	13.02	44.76
4	2.58	11.22	66.79	2.58	11.22	66.79	2.89	12.54	57.31
5	1.88	8.18	74.98	1.88	8.18	74.98	2.83	12.29	69.60
6	1.39	6.06	81.03	1.39	6.06	81.03	2.35	10.23	79.82
7	1.22	5.32	86.36	1.22	5.32	86.36	1.50	6.53	86.36

PC: principal component.

A. Pollution Sources Apportionment in Lake Chamo

The pollution source identification in the LC was based on the factor analysis (FA) where the suitability of the collected water samples was checked by two factors: Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity after normalization of the raw data followed by Z-scale standardization. Accordingly, the KMO in LC was found to be 0.77 > 0.5 revealing the FA is suitable for the analysis and interpretation of water quality variables in the LC.

The FA result in the LC revealed that it has extracted seven principal components that can be explained with a total variance of 86.3%. According to Barakat *et al.* [25] component loadings > 0.75, 0.75–0.50 and 0.50–0.30 is classified as 'strong', 'moderate', and 'weak' loadings, respectively. Each of the extracted components (factors) had an eigen value greater than 1 as shown in Table VII and Table VIII. Accordingly, the first principal component (factor 1) that explained 16.36% of the total variance had strong positive loading for temperature (0.78) and TSS (0.89) and strong negative loading for turbidity (-0.75). These could partly be due to the high contribution of and impact of runoff on the lakes water quality. Besides, the sediment washed out from agricultural lands is negatively impacting the LC turbidity.

TABLE VIII: VARIMAX ROTATED COMPONENT MATRIX FOR EACH WATER QUALITY CONSTITUENT

	Component						
	1	2	3	4	5	6	7
pH	-	-0.72	-	-	-	-	-
T	0.78	-	-	-	-	-	-
DO	-	-	-	-	0.66	-	-
Turbidity	-0.75	-	-	-	-	-	-
EC	-	0.97	-	-	-	-	-
TDS	-	0.94	-	-	-	-	-
TH	-	-	-	0.88	-	-	-
Ca	-	-	-	-	-	-	-
Mg	-	-	-	0.94	-	-	-
Na	-	-	-	-	-	-	-
K	-	-	-	0.63	-	-	-
Cl	-	-	-	-	-	-	-
SO_4	-	-	-	-	0.88	-	-
COD	-	-	-	-	-	-	0.94
BOD	-	-	-	-	0.86	-	-
NO_3	-	-	0.92	-	-	-	-
NO_2	-	-	0.85	-	-	-	-
NH_3	-	-	0.76	-	-	-	-
PO_4	-	0.60	-	-	-	-	-
TA	-	-	-	-	-	-0.94	-
CO_3	-	-	-	-	-	-	-
HCO_3	-	-	-	-	-	-0.76	-
TSS	-0.89	-	-	-	-	-	-

The second principal component (factor 2) had strong positive loading for EC (0.97) and TDS (0.94) and moderate negative loading for pH (0.72) and PO_4 (0.6) which explains the 15.38% of total variance after varimax rotation. This could be a good indication of a mixed source of pollution from nature such as mineral dissolution and anthropogenic influence. Similarly, the third component (PC3) which explained a total variance of 13.02% showed a strong positive loading for NO_3 (0.92), NO_2 (0.85) and NH_3 (076). The high component loading for the constituents could be an indication of the high contribution of agrochemicals to the LC. The LC catchment is characterised by intense agriculture where inorganic nitrogen fertilizers are commonly used, and this could have triggered this component's strong loading for the constituents (Table V).

Components corresponding to moderate and strong loadings are shown in Table VIII.

The fourth principal component had strong positive loading for TH (0.88) and Mg^{2+} (0.94) and moderate loading for K (0.63) and explaining a total variance of 12.54%. The component loadings for TH and Mg^{2+} could be attributed to the rock sedimentation and other magnesium-bearing formations and soil runoff from the nearby catchments. Moreover, the fifth principal component has strong positive loadings for SO_4 (0.88) and BOD (0.86) explaining a total variance of 12.29%. the component loading is more interpreted by the increased impact of organic pollution mainly from domestic sources. The six principal components extracted by FA with a total explained variance of 10.23% had strong negative loading for TH and HCO_3 with a component loading of -0.94 and -0.76. On the other hand, the seventh component (PC 7) had strong positive loading for COD with a loading of 0.94. This could be attributed by biogenic and anthropogenic (urban wastewater) (Table V).

B. WQI and Pollution Status in Lake Chamo

For the determination of pollution status in the Lake Chamo, the arithmetic water quality index approach was used. The weighted arithmetic water quality index method classified the water quality in the LC based on the degree of purity by using the most commonly measured water quality variables. The overall WQI using the weighted arithmetic method is shown on Table IX.

The water quality parameters selected for the analysis are pH, T, DO, EC, TDS, TH, Ca^{2+}, Mg^{2+}, Na^+, Cl-, K^+, SO_4-, BOD, COD, NO_3, NO_2, PO_4, and TA. Accordingly, most of the monitoring stations in the LC fall under very poor water quality index rating evaluated based on drinking purpose as shown on Fig. 3. However, the result revealed that most of the monitoring station WQI values are for irrigation only.

TABLE IX: WATER QUALITY INDEX OF 15 MONITORING STATIONS IN LAKE CHAMO

Monitoring stations	Water quality index
MS1	78.98135
MS2	79.35623
MS3	82.63097
MS4	252.9591
MS5	244.0741
MS6	82.35533
MS7	90.49319
MS8	89.07536
MS9	98.42284
MS10	102.975
MS11	122.2492
MS12	94.55878
MS13	96.22838
MS14	102.897
MS15	359.4861

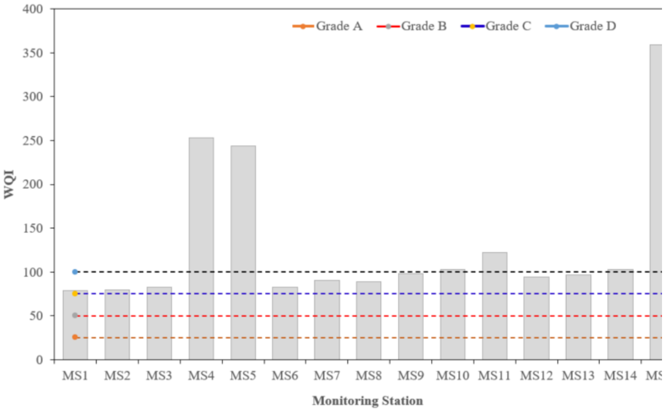

Fig. 3. WQI at LC monitoring stations and their corresponding water quality grades.

The overall WQI value calculated revealed five of the monitoring stations M4, M5, M11, M14 and M15 with a WQI of 252.96, 244.07, 122.25, 102.9 and 359.48, respectively fall under water quality grade-E (WQI > 100) and recommended for restricted irrigation. All other stations can only be used for irrigation but not suitable for drinking and industrial purposes. The five stations with identified WQI not suitable for any purposes except restricted irrigation could be due to the excess presence of phosphate in the area predominantly from agriculture intensive land uses nearby. More specifically, the monitoring station 15 was recorded the highest WQI where the station is characterized by the entry point to the public water navigation and above all inlet to Kulfo river that wash off urban waste to the Lake. Besides, stations 4 and 5 are mainly characterized by the higher agricultural waste influence that could trigger the rise in WQI value in the LC.

From Fig. 3, it can easily be seen that more than 65% of the monitoring stations in LC have an acceptable water quality for irrigation. Both natural and anthropogenic factors could have influenced for some of the monitoring station with a water quality grade of unsuitable. Moreover, some of the monitoring station which are found at a closer range to the high human activities are clearly relatively polluted than others. Besides, some of the sensitive parameters such as turbidity have clearly showed that the WQI value can be changed with the changing of the parameters (Table X).

TABLE X: WQI CLASSIFICATION AND POSSIBLE USAGE

WQI Value	Water Quality Rating	Water Quality Grade	Possible Usage
0-25	Excellent	A	Drinking, Irrigation, Industrial
26-50	Good	B	Domestic, Irrigation, Industrial
51-75	Poor	C	Irrigation and Industrial
76-100	Very poor	D	Irrigation
Above 100	Unsuitable	E	Restricted use for irrigation

C. Cluster Analysis of LC Water Quality

The water quality monitoring stations in LC are grouped based on the cluster analysis according to their chemical characteristics and similarities. The clustering is represented by the dendrogram as shown on Fig. 4. Accordingly, the cluster analysis grouped the LC monitoring stations in to three clusters.

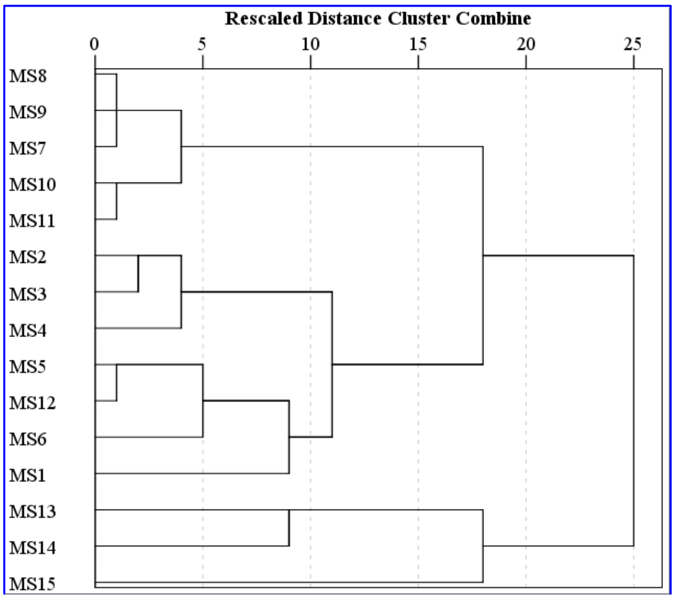

Fig. 2. Dendrogram using Wards method for the cluster analysis of LC water quality monitoring stations.

Cluster 1 is composed of station MS13, MS14 and MS15. All the three monitoring stations are physically located near the inlet to Kulfo River, and this could have attributed to the grouping of the monitoring stations in the same cluster. On the other hand, cluster 2 constitutes majority (47%) of the monitoring stations such as MS1, MS2, MS3, MS4, M5, MS6 and MS12. All these monitoring stations are characterized by the recipients of uncontrolled public and state farm and agricultural wastes. It is found on the western and southwestern flank of the LC close to the shore. In addition, the third cluster is mainly formed by monitoring stations MS7-11 where all of the monitoring stations are located close to the center of the lake. The cluster analysis result is validated by physical locations of the monitoring stations observed during the field visit.

IV. CONCLUSION

The study's objective was to calculate the Water Quality Index of Lake Chamo to assess its suitability for drinking purposes. The physicochemical water quality analysis result of the Lake Chamo revealed that the constituent's concentration did not show significant variation spatially. The factor analysis result in the Lake Chamo revealed that it had extracted seven principal components that can be explained with a total variance of 86.3%. Therefore, more than 65% of the monitoring stations in Lake Chamo have an acceptable water quality for irrigation. Moreover, some monitoring stations closer to high human activities are relatively polluted than others. Besides, some sensitive parameters, such as turbidity, have clearly shown that the WQI value can be changed with the parameters changing. Based on clustering analysis, most (47%) of the monitoring is characterized by the recipients of uncontrolled public and state farm and agricultural wastes.

The study recommends integrated watershed management and source-based pollution control, which could significantly decrease the pollution level in Lake Chamo after identifying the pollution hotspot areas. In addition, continuous Lake water quality monitoring is necessary for deciding on lake pollution.

REFERENCES

[1] Ababu T. Water quality monitoring in Lake Abaya and Lake Chamo region. Arbaminch, Ethiopia. University of Siegen; 2005.

[2] Aman Y, Bezayit M, Berhanu B, Tesfaye GT, Tarekegn D, Tefri Z, Yonas T, *et al.* Geological map of Dila map sheet (NB37-6); Geological map of Hosaena area and Geological map of Nazret, Geological survey of Ethiopia; Addis Ababa, Ethiopia. To accompany memoir No.37; 2012.

[3] Andrea M, Domagoj R. Evaluation of river water quality variations using multivariate statistical techniques. *Environmental Monitoring and Assessment*, 2015; 1: 187-215.

[4] Awulachew SB. Investigation of physical and bathymetric characteristics of Lakes Abaya and Chamo, Ethiopia, and their management implications. *Lakes & Reservoirs: Research & Management*, 2006; 11(3): 133-140.

[5] Brown RM, McClelland NI, Deininger RA, O'Connor MF. (1972). A water quality index—crashing the psychological barrier. In: *Indicators of environmental quality*. Boston, MA: Springer, 1972, pp. 173-182.

[6] Horton IRK. An index number for rating water quality. *Journal of Water Pollution and Control Federation*, 1965; 37(3): 300-306.

[7] Jeevanandam M, Talelign W, Biru A, Sakthi JS, Silva JD, Saravanan, P, Jonathan MP. Evidences of microplastics in Hawassa Lake, Ethiopia: A first-hand report. *Chemosphere*, 2022: 133979.

[8] Makin M, Kingham T, Waddams A, Birchall C, Teferra T. Development prospects in the Southern Rift Valley, Ethiopia. Land Resource Study, Land Resources Division, Ministry of Overseas Development, UK (21); 1975.

[9] Mereta ST, Ambelu A, Ermias A, Abdie Y, Moges M, Haddis A, Mulat WL, *et al.* Effects of untreated industrial effluents on water quality and benthic macroinvertebrate assemblages of Lake Hawassa and its tributaries, Southern Ethiopia. *African Journal of Aquatic Science*, 2020; 45(3): 285-295.

[10] Sayadi M, Rezae A, Rezaei M, Nourozi K. Multivariate statistical analysis of surface water chemistry: A case study of Gharasoo River. *Process of International Academic of Ecological Environmental Science*, 2014; 4: 114-122.

[11] Teffera FE, Lemmens P, Deriemaecker A, Deckers J, Bauer H, Gamo FW, Brendonck L, De Meester L. Why are Lake Abaya and Lake Chamo so different? A limnological comparison of two neighboring major Ethiopian Rift Valley lakes. *Hydrobiologia*, 2018; 829(1): 113–124. https://doi.org/10.1007/s10750-018-3707-8.

[12] Tiruneh TA. Water Quality Monitoring in Lake Abaya and Lake Chamo Region: A Research Based on Water Resources of the Abaya-Chamo Basin-South Ethiopia. PhD thesis, University of Siegen; 2006.

[13] Tripathy JK, Sahu KC. Seasonal hydrochemistry of groundwater in the Barrier Spit system of the Chilika Lagoon, India. *Journal of Environmental Hydrology*, 2005; 13.

[14] Wagesho N. Catchment dynamics and its impact on runoff generation: coupling watershed modelling and statistical analysis to detect catchment responses. International *Journal of Water Resources and Environmental Engineering*, 2014; 6(2): 73–87.

[15] Zhao J, Fu G, Lei K, Li Y. Multivariate analysis of surface water quality in the Three Gorges area of China and implications for water management. *Journal of Environmental Science*, 2011; 23:1460-1471.

[16] Zigde HM, Tsegaye ME. Evaluation of the current water quality of Lake Hawassa, Ethiopia. *International Journal of Water Resources and Environmental Engineering*, 2019; 11(7): 120-128.

[17] APHA. *Standard Methods for the Examination of Water and Wastewater* (23rd ed.). Washington DC: American Public Health Association; 2017.

[18] Chatterjee C, Raziuddin M. Determination of water quality index of a degraded river in Asanol Industrial area, Raniganj, Burdwan, West Bengal. *Nature, Environment and Pollution Technology*, 2002;1(2):181-189.

[19] Miura K, Ban M, Ohba T, Fujinawa A. Sequence of the 1895 eruption of the Zao volcano, Tohoku Japan. *Journal of Volcanology and Geothermal Research*, 2012; 247–248: 139–157. https://doi.org/10.1016/j.jvolgeores.2012.08.005.

[20] Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. *Bioinformatics*, 2005; 21(20): 3940–3941. https://doi.org/10.1093/bioinformatics/bti623.

[21] Rakotondrabe F, Ndam Ngoupayou JR, Mfonka Z, Rasolomanana EH, Nyangono Abolo AJ, Ako Ako A. Water quality assessment in the Bétaré-Oya gold mining area (East-Cameroon): Multivariate Statistical Analysis approach. *Science of the Total Environment*, 2018; 610–611: 831–844. https://doi.org/10.1016/j.scitotenv.2017.08.080.

[22] Pärn E, Edwards D, Sing M. The building information modelling trajectory in facilities management: A review. *Automation in Construction*, 2017; 75, 45–55. https://doi.org/10.1016/j.autcon.2016.12.003.

[23] Huang H, Yu Z, Zhang S, Liang X, Chen J, Li C, Ma J, Jiao R. Drosophila CAF-1 regulates HP1-mediated epigenetic silencing and pericentric heterochromatin stability. *J. Cell Sci.*, 2010; 123(16): 2853-2861.

[24] Khettaf S, Khouni I, Louhichi G, Ghrabi A, Bousselmi L, Bouhidel KE, Bouhelassa M. Optimization of coagulation–flocculation process in the treatment of surface water for a maximum dissolved organic matter removal using RSM approach. *Water Supply*, 2021; 21(6): 3042–3056. https://doi.org/10.2166/ws.2021.070.

[25] Barakat A, El Baghdadi M, Rais J, Aghezzaf B, Slassi M. Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. *International Soil and Water Conservation Research*, 2016; 4(4): 284–292. https://doi.org/10.1016/j.iswcr.2016.11.002.

[26] Angello Z, Behailu B, Tränckner J. Selection of Optimum Pollution Load Reduction and Water Quality Improvement Approaches Using Scenario Based Water Quality Modeling in Little Akaki River, Ethiopia. *Water*, 2021; 13(5): 584. https://doi.org/10.3390/w13050584.

[27] Mishra B, Rafiei Emam A, Masago Y, Kumar P, Regmi R, Fukushi K. Assessment of future flood inundations under climate and land use change scenarios in the Ciliwung River Basin, Jakarta. *Journal of Flood Risk Management*, 2017; 11, S1105–S1115. https://doi.org/10.1111/jfr3.12311.

[28] World Health Organization. *World Health Statistics 2017: Monitoring Health for the Sustainable Development Goals (SDGs)* (1st ed.). World Health Statistics Annual; 2017.

[29] Yakob G, Gebremicheal A, Aklilu A, Melaku E. Participatory Evaluation of Different Multipurpose Grass Species for Graded Soil Bund Stabilization in Gimbo District, South West Ethiopia. *OALib*, 2015; 02(06): 1–10. https://doi.org/10.4236/oalib.1101627.

[30] Abendong AA, Emmanuel E, Fongoh EJ, Akoachere RA. Assessment of groundwater quality in Bamenda–Cameroon for suitable applications. *SN Applied Sciences*, 2019; 1(11). https://doi.org/10.1007/s42452-019-1351-1.