Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrating Multivariate Analysis and Explainable Machine Learning to Assess Eutrophication: Chlorophyll-a and Dissolved Oxygen Dynamics in Ziway and Langano Lakes, Ethiopia

A.A Awasho

afework.alemayehu@aau.edu.et

Arba Minch Water Technology Institute, Arba Minch University

B.E Cholo

Arba Minch Water Technology Institute, Arba Minch University

A.H Gelebo

Arba Minch Water Technology Institute, Arba Minch University

Research Article

Keywords: Spatio-temporal variability, Chlorophyll-a, Dissolved oxygen, Eutrophication, Explainable machine learning, Lake water quality

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11234746/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Integrating Multivariate Analysis and Explainable Machine Learning to Assess Eutrophication: Chlorophyll-a and Dissolved Oxygen Dynamics in Ziway and Langano Lakes, Ethiopia

A. A. Awasho[ac]*, B. Elias[a], and Ayano Hirbo Gelebo[b]

[a] Faculty of Meteorology and Hydrology, Arba Minch Water Technology Institute, Arba Minch University, Arba Minch, Ethiopia
[b] Faculty of Water Resources and Irrigation Engineering, Arba Minch Water Technology Institute, Arba Minch University, Arba Minch, Ethiopia
[c] Urban Transformation and Development Program, Training Institute, Ethiopian Public Service University, Addis Ababa, Ethiopia

*Corresponding author.

E-mail: afework.alemayehu@aau.edu.et

Abstract

Eutrophication, driven by rising chlorophyll-a concentrations and declining dissolved oxygen (DO), is an escalating threat to lake ecosystems worldwide, and multivariate statistical analysis combined with explainable machine learning offers a powerful framework for characterizing its dynamics and drivers to support evidence-based watershed management. The Ethiopian Rift Valley has experienced substantial anthropogenic pressure and climate variability, increasingly affecting Lake Ziway (450 km^2, mean depth 5 m) and Lake Langano (230–250 km^2, depth 46–55 m). This study assessed spatio-temporal dynamics of chlorophyll-a and dissolved oxygen (DO) at both lakes using Principal Component Analysis (PCA), Cluster Analysis, Mann-Kendall trend analysis, Random Forest (RF) machine learning, and SHAP-based Explainable Artificial Intelligence (XAI), with turbidity and total suspended solids (TSS) included as supporting covariates. Mean chlorophyll-a concentrations were 15.09–29.09 µg/L (Langano) and 19.63–37.83 µg/L (Ziway), indicating high phytoplankton productivity in both lakes, with pronounced increases during the wet season. Dissolved oxygen showed statistically significant declining trends (Tau = -0.526, Langano; -0.595, Ziway), while chlorophyll-a increased significantly (Sen's slope: +0.963 and +1.249 µg/L, respectively). DO explained 30–68% of Random Forest model variability and was consistently among the top two predictors of chlorophyll-a at both lakes. RF models achieved high predictive accuracy (R^2 = 0.968–0.986; RMSE = 1.230–1.435). Both lakes were classified as mesotrophic during the dry season, shifting to eutrophic during the wet season, with Lake Ziway showing greater ecological pressure and variability than Lake Langano. Targeted watershed management interventions to curb nutrient loading and protect oxygen balance are recommended.

Keywords: Spatio-temporal variability; Chlorophyll-a; Dissolved oxygen; Eutrophication; Explainable machine learning; Lake water quality.

Graphical Abstract

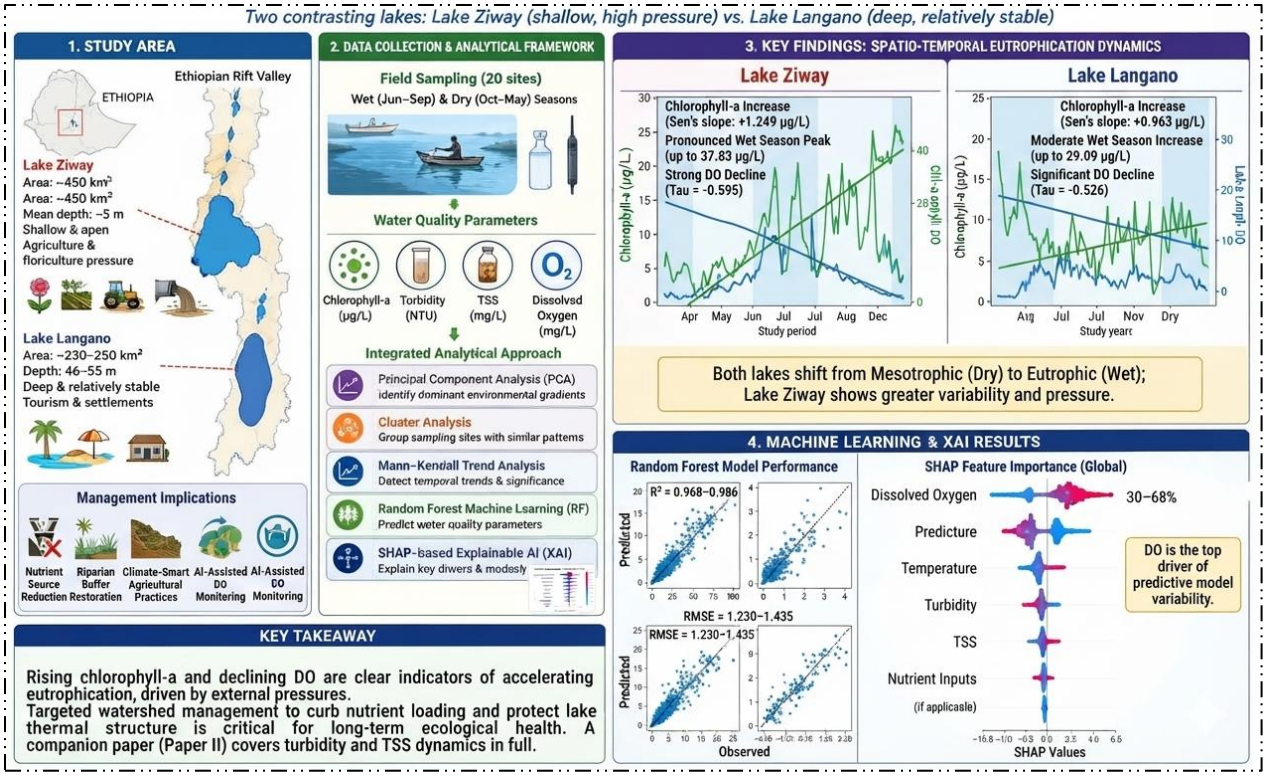


Figure 1. Graphical Abstract for Integrating Multivariate Statistics and Explainable Machine Learning to Assess the Spatio-Temporal Eutrophication Dynamics in Ethiopian Rift Valley Lakes

1. Introduction

Freshwater lake ecosystems provide important services such as drinking water, irrigation, fish populations, recreation supporting biodiversity, and help maintain a healthy environment [1], [2]. Yet, they confront multiple threats including nutrient enrichment, sedimentation, pollution, and climate variability [3], [4]. The Ethiopian Rift Valley, one of the fastest growing regions in the world with annual population growth exceeding 2.5% [5], is under substantial pressure from land use change within the surrounding rural lake catchments.

Chlorophyll-a, a marker of algal biomass, and dissolved oxygen (DO) are two of the most fundamental indicators of a lake's trophic condition and ecosystem health, together capturing the balance between primary production and the oxygen available to support aquatic life [6], [7]. Because chlorophyll-a can be measured remotely using satellite and in situ optical technologies in real time, it is particularly valuable for water quality assessment in resource-poor regions, while DO provides a direct, ecologically critical measure of the consequences of that productivity.

Eutrophication, caused primarily by nitrogen and phosphorus additions to freshwater ecosystems through agriculture, urbanization and industrial discharges, is the greatest threat to freshwater ecosystems worldwide [4], [8], [9]. In particular, the floriculture industry surrounding Lake Ziway adds large quantities of fertilizer to the lake through

irrigation return flows, demonstrating how intensified land use increases local eutrophication. Wet-season runoff results in elevated chlorophyll-a while decreasing DO; this pattern has been documented in lakes throughout eastern Africa, including Lake Victoria [10] and Lake Tana [11].

In the Ethiopian Rift Valley, Lakes Ziway and Langano are vital both regionally and for local communities, particularly for agricultural, fishing and tourism activities [12]. Lake Ziway is shallow and hydrologically open, and is experiencing tremendous government-driven agricultural and floriculture expansion [7], [13]. Historically, Lake Langano has been a stable lake with only limited stressors; however, preliminary results from this research indicate that both lake productivity and stressors (managed tourism and land-use change) have increased [14]. There is a lack of comparative studies that encompass both Lakes Ziway and Langano [15].

Current analytical methods for assessing these lakes have significant limitations. Principal Components Analysis (PCA) identifies the most significant environmental gradients, but does not quantify the predictive strength of individual parameters nor provide a measure of uncertainty associated with those predictions [16]. Machine learning models offer considerable improvement in predictive accuracy [17], but many of these models are "black-box" models that do not describe the reasons for a given prediction [18], [19]. The Random Forest type of predictive model offers a measure of native (global) feature importance; however, it does not explain individual predictions a gap that can be addressed using SHAP-based Explainable Artificial Intelligence (XAI) techniques.

In this study, PCA, cluster analysis, Mann-Kendall trend analysis, Random Forest and SHAP are integrated to assess the spatio-temporal dynamics of chlorophyll-a and dissolved oxygen, and to identify the variables including turbidity and TSS as supporting covariates that most influence chlorophyll-a prediction, in Lakes Ziway and Langano.

2. Materials and Methods

2.1. Description of Study Area

The Ziway-Shala sub-basin, situated within the Ethiopian Central Rift Valley (ECRV) and covering roughly 15,000 km^2, forms a closed hydrological chain of four lakes Ziway, Langano, Abijata, and Shala arranged along a north-to-south hydraulic gradient [20]. This study focuses on the two uppermost lakes in that chain, Ziway and Langano, whose contrasting morphometry and hydrologic regime provide a natural basis for comparing how productivity and oxygen balance respond to differing levels of catchment pressure.

Lake Ziway, the larger and shallower of the two systems (approximately 08°06′44″N to 07°51′19″N and 38°43′30″E to 38°56′44″E; Fig. 2B), receives most of its inflow from the perennial Meki and Katar Rivers before draining south through the Bulbula River toward Lake Abijata [21]. Its shallow depth and comparatively short water-residence time leave the lake hydrodynamically responsive to inflow pulses a property directly relevant to the productivity and oxygen shifts examined here while also underpinning irrigation, fisheries, domestic water supply, and wetland habitat for the surrounding population.

Lake Langano (approximately 07°42′39″N to 07°30′48″N and 38°40′56″E to 38°49′56″E; Fig. 2B) is fed mainly by seasonal streams and direct rainfall and discharges through the Horakela River into Lake Abijata [22], [23]. Its considerably greater depth and more stable water levels, sustained by groundwater recharge and limited abstraction, favor a more buffered oxygen regime than that of Ziway. The lake is free of schistosomiasis (bilharzia) and supports a well-established tourism sector [24].

Both catchments are dominated by a mosaic of agriculture, grazing land, settlement, and remnant natural vegetation, with expanding irrigation, urbanization, and tourism progressively altering catchment hydrology. Because these pressures ultimately register as measurable shifts in productivity and oxygen availability the two indicators examined in this study physical and hydrological details of both lakes are summarized in Supplementary Table 1.

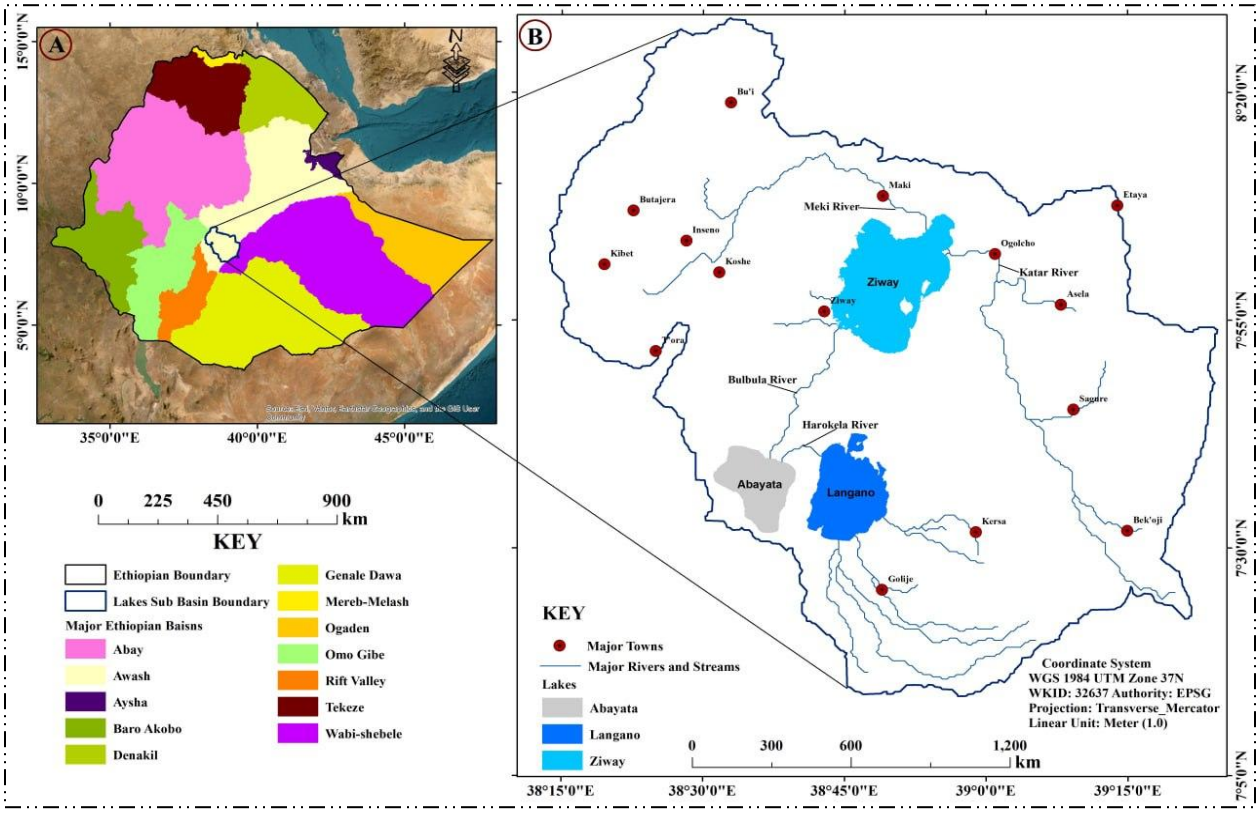


Figure 2. A) Major Ethiopian Basins and B) Ziway and Langano Lakes

2.2. Calibration of In-Situ Water Quality Sensors

Before field deployment and throughout the monitoring campaign, in situ sensors measuring chlorophyll-a and dissolved oxygen (DO) were calibrated with certified standards and analytical grade reagents according to internationally accepted guidelines [25], [26]. The chlorophyll-a sensor (fluorescence detection) was calibrated using a two-point method, with deionized water as the zero reference and rhodamine dye solution as the standard; this sensor was frequently validated against measurements of chlorophyll-a in acetone-extracted samples (see Supplementary Table 2).

DO was calibrated using a two-point technique: one point at zero oxygen using a sodium sulfite solution, and a second point at 100% saturation using air-equilibrated water. All sensors were rinsed before calibration and were routinely checked for accuracy through regular calibration checks to mitigate drift and ensure reliable measurements (see Supplementary Table 2) [25]. Turbidity and TSS sensors were calibrated concurrently as part of the same field campaign.

2.3. Water Sample Collection, Labeling, Storage, and Transportation

A stratified composite sampling design was used to capture both vertical and horizontal water quality variability at each lake. Over the study period, 240 composite samples were collected in total (120 per lake, split evenly between the dry and wet seasons), drawn from ten locations per lake (Z1–Z10 on Ziway; L1–L10 on Langano). Sites were chosen to reflect the spatial heterogeneity of each system accounting for lake size, accessibility, hydrological variability, inflow influence, and anthropogenic pressure consistent with prior limnological work in the Ethiopian Rift Valley [25], [27], [28] (Fig. 3A and 3B).

At each site, a single 600 mL composite sample was built from three 200 mL sub-samples drawn from the upper, middle, and lower water column using a Van Dorn sampler [25], [26], collected into pre-cleaned polyethylene containers that had been rinsed three times on-site beforehand to prevent contamination. Because both chlorophyll-a and dissolved oxygen are subject to pronounced diurnal cycling, sampling was repeated twice per site per day at every location [18], [25]. Containers were labelled by site code, date, time, and season, and water depth was logged with a GPS-equipped echo sounder to support traceability.

Samples were held on ice (4 ºC) immediately after collection to limit biological and photochemical change in chlorophyll-a and dissolved oxygen prior to analysis, and were processed the same day on arrival at the laboratory to preserve measurement accuracy and repeatability [25], [26].

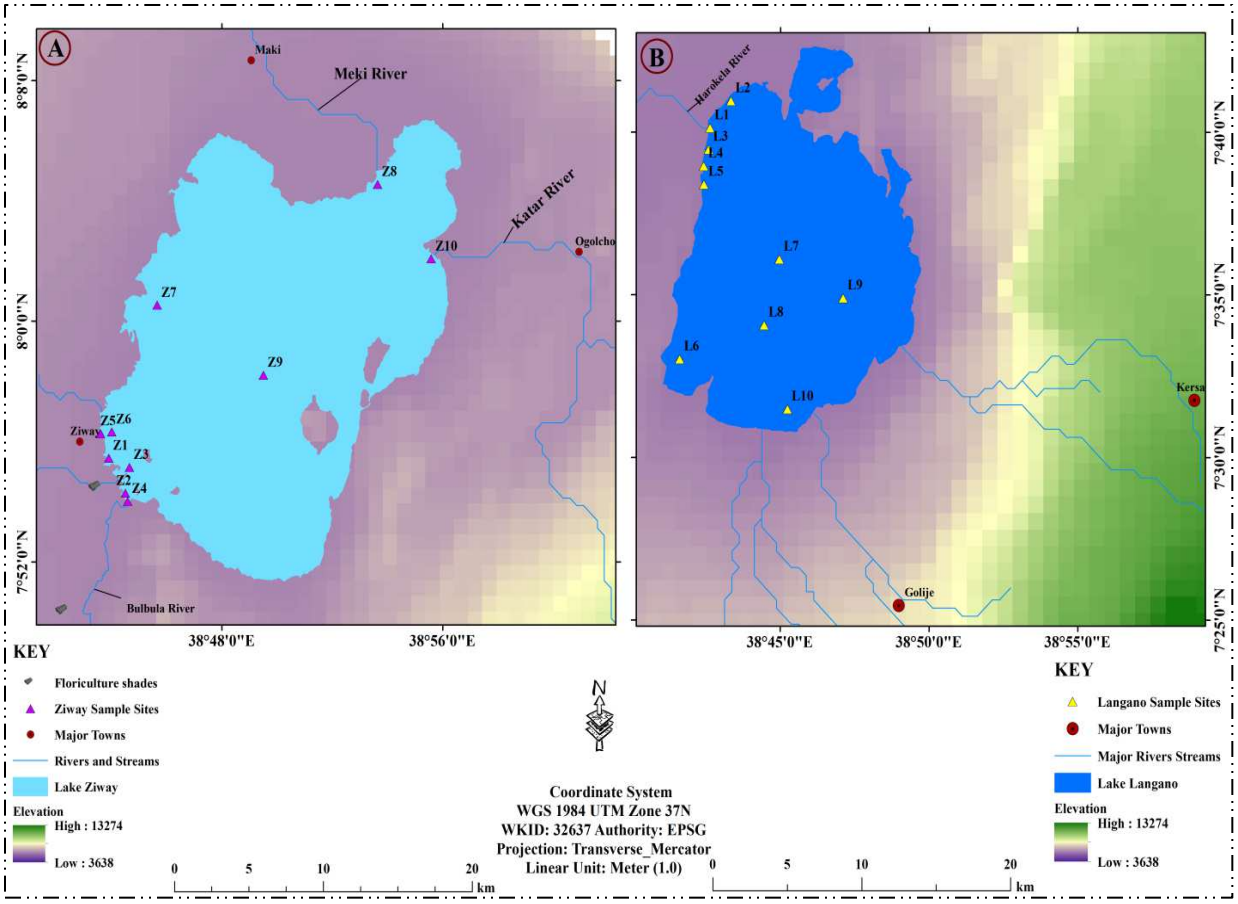


Figure 3. A) Sample sites on Lake Ziway and B) Sample sites on Lake Langano

2.4. Data Collected

The primary sources of data were direct field sampling and laboratory analysis. This paper focuses on chlorophyll-a and dissolved oxygen (DO); turbidity and TSS were collected concurrently and are used here only as supporting covariates in the machine learning framework.

Chlorophyll-a was expressed in µg/L using the [29] corrected spectrophotometric method, accounting for turbidity interference and accessory pigments in phytoplankton biomass estimation (see Supplementary Equation S1). DO was measured in situ via optical probes and confirmed with the Winkler titration method per the [25] standard method (Supplementary Equation S5). Temperature was documented based on its effect on DO saturation and aquatic system processes [30] (see Supplementary Equation S6).

The dataset was subject to statistical analysis and machine learning algorithms in order to evaluate relationships between water quality parameters and produce predictive models.

2.5. Model Description

This research used multiple statistical methods and machine learning approaches to analyze how chlorophyll-a and dissolved oxygen in Lakes Ziway and Langano vary over

time and across locations. This work encompassed PCA, Hierarchical Cluster Analysis (HCA), Mann-Kendall trend analysis, the Sen's slope estimator, and Random Forest (RF), combined with an Explainable Artificial Intelligence (XAI) framework. Turbidity and TSS were retained in the underlying multivariate matrix as supporting covariates because they load onto the same productivity/sediment gradient as chlorophyll-a and DO.

To identify the main environmental gradient(s) driving water quality variation between Lakes Ziway and Langano and to simplify the data for analysis, PCA was used to reduce dimensionality by decomposing the data covariance matrix into eigenvalues (Supplementary Equation S7). PCA was completed using the standardized correlation matrix (i.e., z-score normalized variables) to remove the effect of different measurement units and scales between parameters (chlorophyll-a in μg/L, DO in mg/L, turbidity in NTU, TSS in mg/L).

HCA grouped sampling sites and seasons based on water quality similarity using the Euclidean distance method, with Ward's linkage used to create clusters that minimized within-cluster variance and produced ecological classifications of sampling site-cluster pairs [31], [32].

Temporal trends in chlorophyll-a and DO were evaluated using the Mann-Kendall non-parametric test and the Sen's slope estimator (Supplementary Equation S8 – S10). The Mann-Kendall statistic identifies an overall trend (increase or decrease) across a time series without assuming normality in the data [33], [34].

2.6. **Training of ML and Verification Approach**

Before performing any analysis, the datasets were checked for missing or unusual values. Standardization and normalization were then applied to each dataset to ensure reliable model predictions [35].

In the machine learning component of the study, the cleaned dataset was split into training (70%) and test (30%) subsets [36] to develop a Random Forest (RF) algorithm. The RF model was trained to predict chlorophyll-a using dissolved oxygen, turbidity, and TSS as predictor variables, and its predictive accuracy and generalization were assessed [37]. To understand how each predictor contributed to chlorophyll-a variability, Explainable Artificial Intelligence (XAI) techniques based on SHAP values were applied, with particular attention to the role of dissolved oxygen.

The overall process of calibrating, validating, and modeling is depicted in the Conceptual Framework (Supplementary Fig. 1) and Methodological Workflow (Supplementary Fig. 2).

3. Results

3.1. Lake Ziway

3.1.1. Spatio-Temporal Variation of Chlorophyll-a and Dissolved Oxygen

Lake Ziway demonstrates a high degree of spatio-temporal variation in chlorophyll-a and dissolved oxygen (see Supplementary Fig. 3). This variability is driven by seasonal forcing and related differences in sampling locations (see Supplementary Table 5). Mean

chlorophyll-a (Chl-a) levels rose from 19.63 µg/L in the dry season to 37.83 µg/L in the wet season, demonstrating an increase in phytoplankton productivity during the wet season (Fig. 4A and Supplementary Table 3 and 5). Dissolved oxygen (DO) levels were lowest in the wet season (mean of 4.22 mg/L) and showed the highest variability during the wet season compared to the dry season (mean 6.72 mg/L) (Fig. 4D and Supplementary Table 3 and 5).

There was considerable spatial variation in both parameters among sampling sites. Sites near agricultural runoff and sediment inflow zones consistently had higher chlorophyll-a and lower DO, whereas offshore and central lake sites had lower chlorophyll-a and better-maintained oxygen levels. These differences indicate that localized runoff, sediment transport, and land use strongly influence productivity and oxygen dynamics in Lake Ziway (see Supplementary Table 3 and 5).

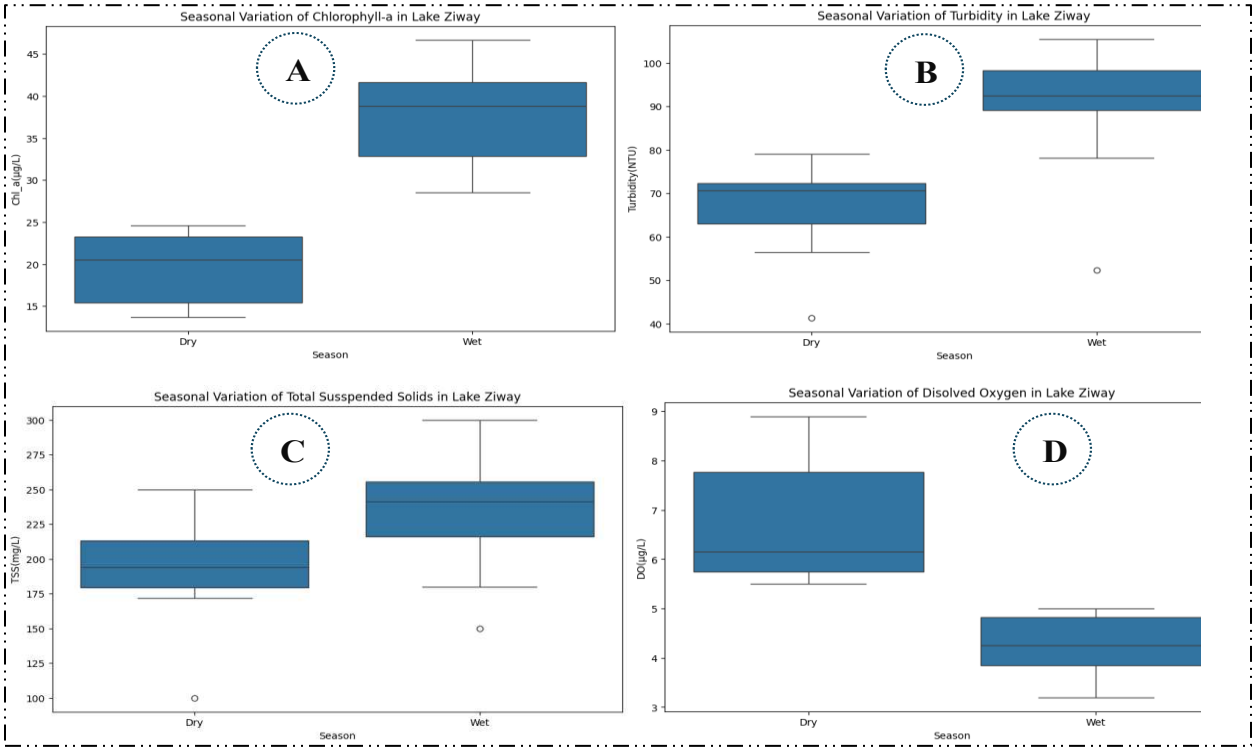


Figure 4. Temporal variations of: A) Chlorophyll-a, B) Turbidity, C) Total suspended Solids, and D) Dissolved Oxygen for Lake Ziway

3.1.2. Multivariate Structure: Principal Component Analysis (PCA)

Chlorophyll-a (Chl-a) attributed one of the strongest positive loadings on principal component one (PC1), at 0.52, alongside turbidity (0.50) and TSS (0.47) as supporting covariates. Dissolved oxygen (DO) showed the strongest negative loading on PC1, at -0.51 (see Supplementary Table 7).

PC1 can therefore be interpreted as a productivity-oxygen axis: sites with high phytoplankton biomass had correspondingly low dissolved oxygen concentrations. PC1

accounted for 80.93% of total variability (eigenvalue = 3.41), indicating strong seasonal co-variation between chlorophyll-a and DO (see Fig. 4A and Supplementary Table 8). PC2 accounted for an additional 9.47% of total variability (see Supplementary Table 8 and Fig. 4D). Together, PC1 and PC2 represented 90.40% of total variation, confirming that seasonally driven hydrological conditions govern the joint variability of chlorophyll-a and DO in Lake Ziway (see Fig. 5 and Supplementary Table 8).

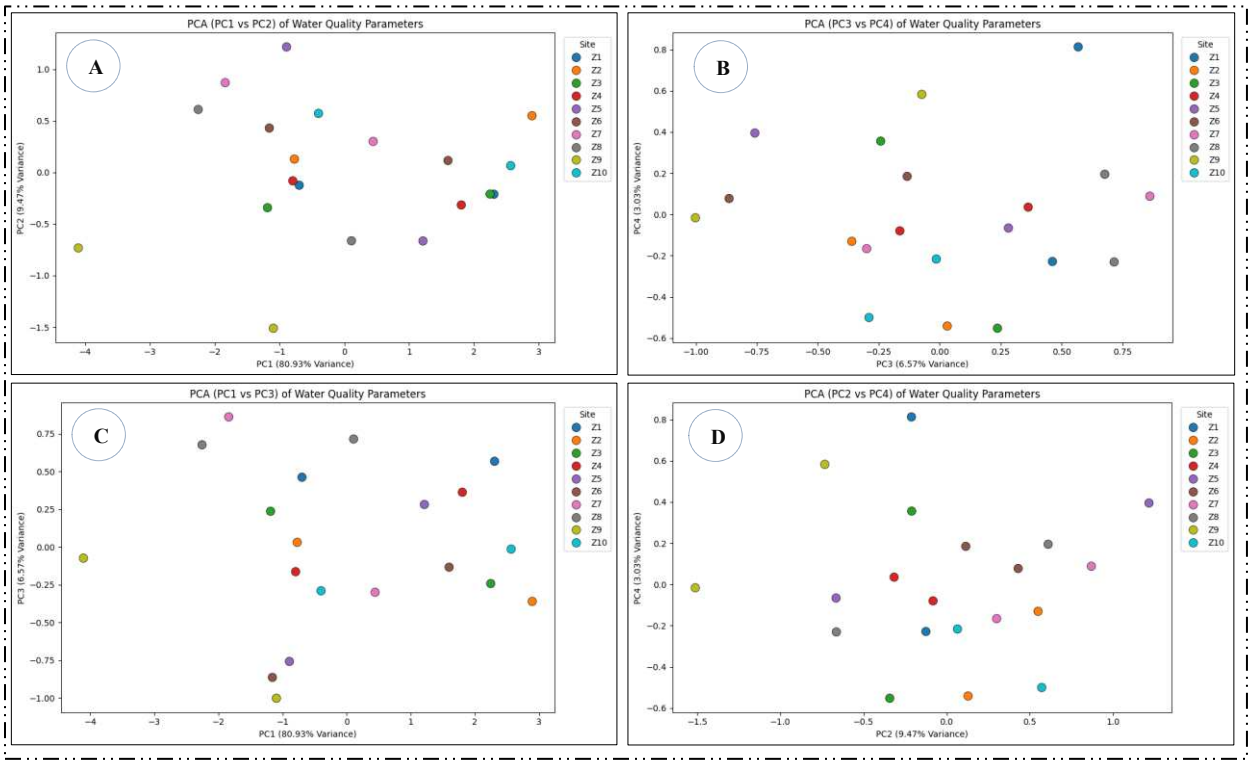


Figure 5. Multivariate Principal Component Analysis of selected water quality parameters for Lake Ziway

3.1.3. Cluster Analysis

Hierarchical cluster analysis using Ward's method with Euclidean distance produced two main clusters. The first cluster consisted of wet season samples from sites strongly influenced by human activity and hydrology (Z8 and Z10 near the Meki and Katari River outlets; Z1 through Z6 beach sites under urban, agricultural, and floricultural influence) (Fig. 6 /Yellow/). This cluster was characterized by elevated chlorophyll-a and depressed dissolved oxygen. The second cluster contained dry season samples, primarily from mid-lake locations (Z7 and Z9) and forested shoreline sites, which exhibited lower chlorophyll-a and higher, more stable dissolved oxygen than the first cluster (Fig. 6 /Green/).

Sub-clustering within each group confirmed significant spatial variability among the ten sampling locations (Z1 through Z10), consistent with the different pressures affecting each site: Z1 and Z2 experience floriculture and urban effluent pressure in the southern basin, while Z3 through Z6 reflect river inflow pressure in the northern basin (Fig. 6).

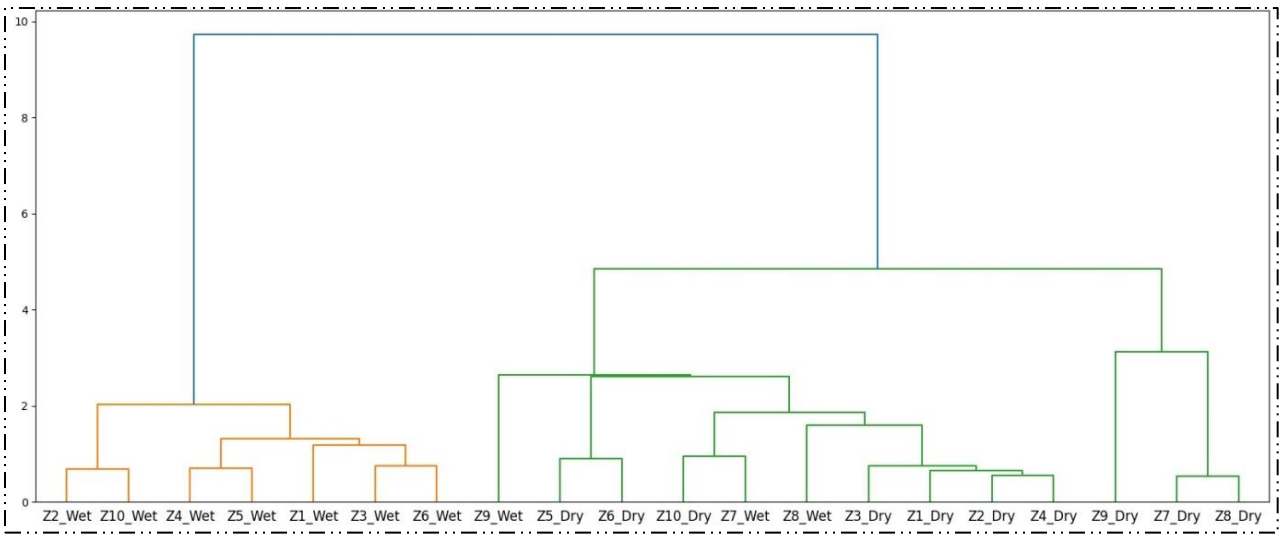


Figure 6. Cluster Analysis of selected water quality parameters for Lake Ziway

3.1.4. Trend Analysis

Mann-Kendall trend analysis for Lake Ziway indicates statistically significant increases in chlorophyll-a and a statistically significant decrease in dissolved oxygen. Chlorophyll-a increased significantly ($Z = 3.017$, $p = 0.003$, Sen's slope = +1.249 µg/L). Dissolved oxygen showed a very strong downward trend ($Z = -3.640$, $p < 0.001$, slope = -0.211 mg/L), indicating progressive oxygen depletion. Relatively large Tau (τ) values (maximum 0.495 for chlorophyll-a) provide evidence of a consistent, monotonically increasing eutrophication signal alongside declining oxygen availability (Table 1).

Table 1. Mann–Kendall Trend Test Results for Water Quality Parameters in Lake Ziway

Lake	Parameters	Trend	h	p	Z	Tau	S	Var(S)	Sen's Slope	Intercept
Lake Ziway	Chll-a(µg/L)	Increasing	True	0.003	3.017	0.495	94	950	1.249	14.68
	Tur (NTU)	Increasing	True	0.025	2.239	0.368	70	950	1.666	61.47
	TSS (mg/L)	Increasing	True	0.012	2.498	0.411	78	950	4.096	174.79
	DO (mg/L)	Decreasing	True	0.000	-3.640	-0.595	-113	947	-0.211	7.25

3.1.5. Trophic Status

Lake Ziway shows a pronounced seasonal shift in trophic state. Average dry-season (mesotrophic) chlorophyll-a concentrations of 19.63 µg/L increase significantly during the wet season to 37.83 µg/L, demonstrating a clear seasonal trend toward greater productivity. This pattern holds across space: all 20 of 20 sites in the lake are classified as mesotrophic during the dry season, while wet-season sites are predominantly classified as eutrophic, indicating elevated nutrient loading and algal growth during the wet season (Table 2 and Supplementary Fig. 4).

Table 2. Seasonal Trophic Status in Lake Ziway and Lake Langano

Lakes	Season	Mean Chll-a (µg/L)	Trophic Status
Ziway	Dry	19.630	Mesotrophic (Moderate)
	Wet	37.830	Eutrophic (High productivity)
Langano	Dry	15.090	Mesotrophic (Moderate)
	Wet	29.090	Eutrophic (High productivity)

3.1.6. Model Performance and Prediction Uncertainty

The Random Forest model predicting chlorophyll-a in Lake Ziway performed well (R^2 = 0.986; RMSE = 1.230; MAE = 0.886; Bias = -0.013) (Fig. 7A), with some prediction uncertainty (approximately 1.2–5.1 µg/L), particularly at higher chlorophyll-a levels (Fig. 7B). Slight overestimation of residuals at higher levels suggests that the relationship between chlorophyll-a and dissolved oxygen becomes ecologically more complex at elevated productivity (see Supplementary Table 10).

Feature importance identified dissolved oxygen (SHAP importance = 0.681) as the strongest predictor of chlorophyll-a, followed by turbidity (0.256) and TSS (0.063) as supporting covariates. This indicates that oxygen depletion dynamics are the primary driver of the productivity signal captured by the model, with sediment-related variables playing a secondary role. Residual error patterns further suggest that the water quality system is undergoing dynamic processes such as heterogeneous mixing (Fig. 6A and B).

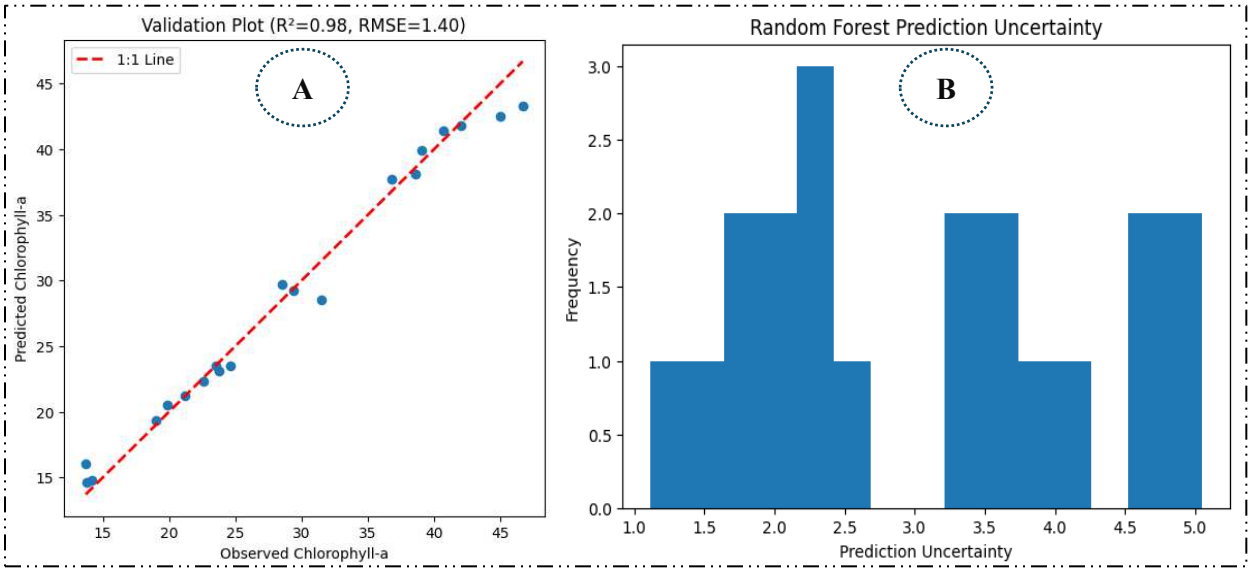


Figure 7. A) Validation of model performance and B) Prediction uncertainty of model for Lake Ziway

3.2. Lake Langano

3.2.1. Spatio-Temporal Variation of Chlorophyll-a and Dissolved Oxygen

Lake Langano shows a moderately varying spatial and temporal pattern in chlorophyll-a and dissolved oxygen, primarily due to seasonal influences, with less spatial variation between sampling sites than Lake Ziway (see Supplementary Fig. 7). Mean chlorophyll-a

concentration shows a clear seasonal relationship, 15.09 µg/L in the dry season and 29.09 µg/L in the wet season, with the productivity peak occurring near the Harokela River inflow and at agriculturally influenced sites L1 through L3 (see Fig. 8A and Supplementary Table 6). Dissolved oxygen concentrations were lower in the wet season, ranging 2.1–7.4 mg/L, compared to 6.1–9.9 mg/L in the dry season (see Fig. 8D and Supplementary Table 6). The lowest DO concentrations occurred at near-shore, settlement-influenced sites (L4–L6), while mid-lake open-water sites (L7–L9) had the highest concentrations (see Fig. 3B).

Although overall differences across Lake Langano are less pronounced than in Lake Ziway, spatial differences remain significant (see Supplementary Fig. 12). Near-shore sampling locations (L1–L6, L10) affected by agricultural runoff and riverine inflows showed elevated chlorophyll-a and depressed DO relative to mid-lake sites (L7–L9). Overall, localized land use, sediment transport, and inflows continue to shape productivity and oxygen dynamics in Lake Langano even though the ecosystem remains relatively stable (see Supplementary Table 4 & 6 and Fig. 3B).

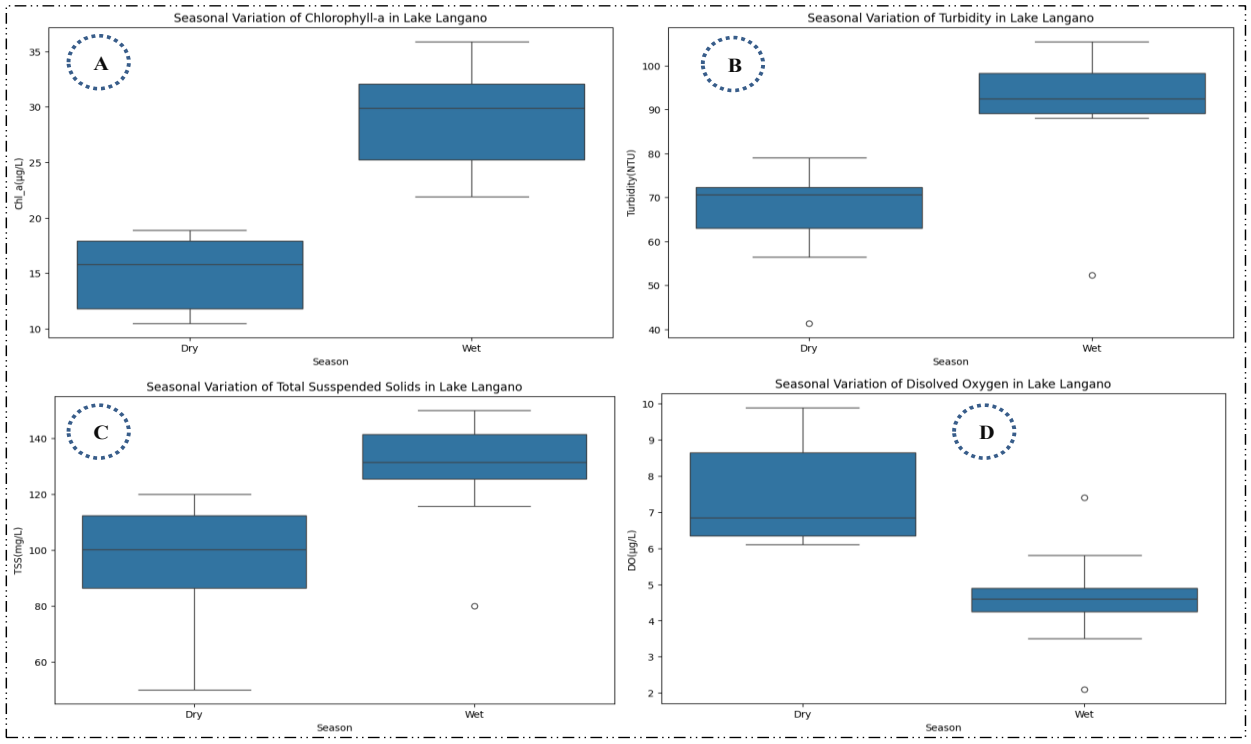


Figure 8. Temporal variations of: A) Chlorophyll-a, B) Turbidity, C) Total suspended Solids, and D) Dissolved Oxygen for Lake Langano

3.2.2. Multivariate Structure of Principal Component Analysis (PCA)

As with Lake Ziway, all variables for Langano were z-score standardized. PC1 explained 88.74% of total variance (eigenvalue = 3.74). Chlorophyll-a (Chl-a) showed a strong negative loading on this axis (-0.48), alongside turbidity (-0.50) and TSS (-0.51) as covariates, while dissolved oxygen exhibited the strongest positive loading (+0.52) (see Supplementary Tables 8).

These results indicate that PC1 defines a productivity-oxygen axis analogous to that in Lake Ziway. PC2 explained an additional 7.28% of variance (eigenvalue = 0.31) (see Supplementary Fig. 9D). Together, PC1 and PC2 accounted for over 96% of total variance, with clear seasonal separation along PC1: dry-season samples generally scored above zero (associated with lower chlorophyll-a and higher DO), while wet-season samples clustered at or below zero (higher productivity, lower DO) (Fig. 9A–D and Supplementary Table 8).

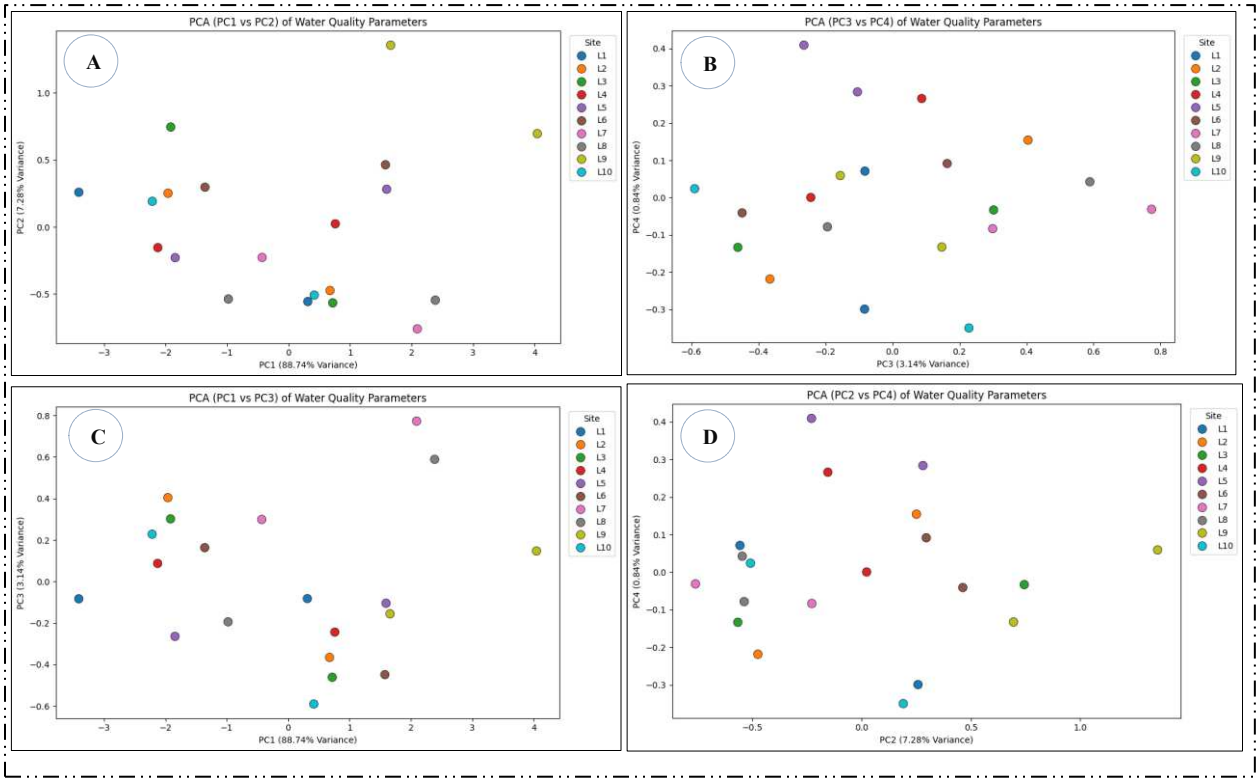


Figure 9. Multivariate Principal Component Analysis of selected water quality parameters for Lake Langano

3.2.3. Cluster Analysis

Hierarchical clustering using Ward's linkage with Euclidean distance identified two distinct clusters reflecting seasonal and spatial differences. Cluster 1 (yellow) comprised wet season samples from inflow-influenced and land-use-affected sites (L1, L10, and shoreline sites L2–L6), characterized by high chlorophyll-a and low dissolved oxygen. Cluster 2 (green) comprised primarily dry season samples from mid-lake (L7–L9) and undisturbed shoreline sites, along with some wet-season mid-lake samples, all exhibiting relatively better productivity-oxygen balance than Cluster 1.

Sub-clusters within each group demonstrated the influence of local pressures on productivity and oxygen dynamics: L1 and L10 are influenced by river inflow, L2–L3 by adjacent agricultural land use, and L4–L6 by near-shore settlement (Fig. 10).

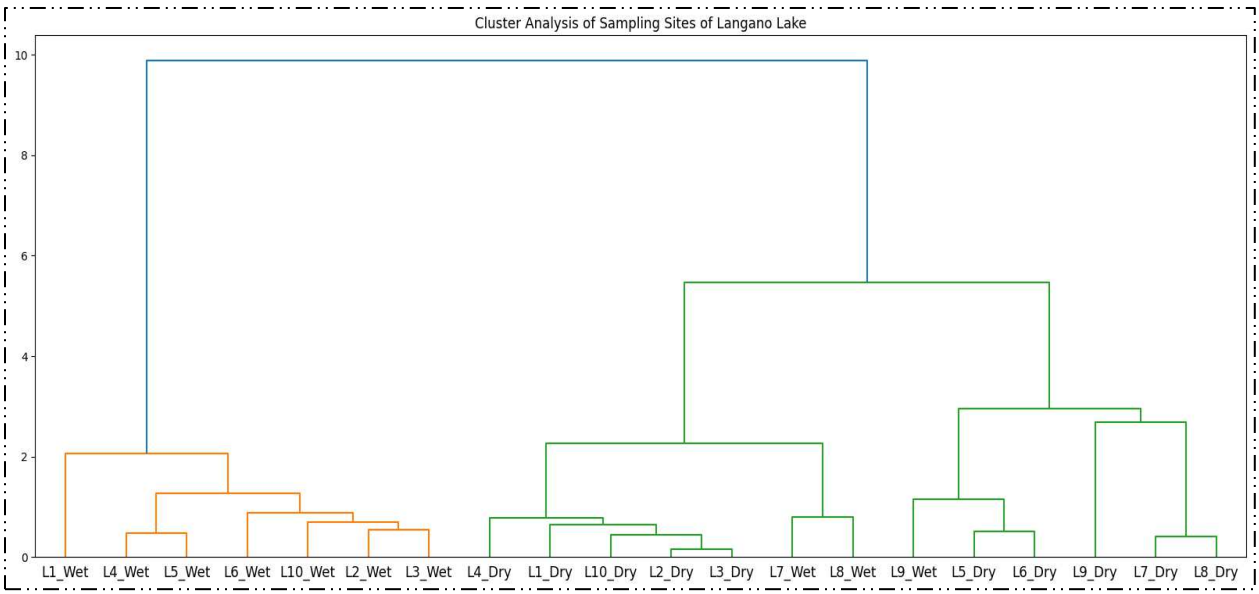


Figure 10. Cluster Analysis of selected water quality parameters for Lake Langano

3.2.4. Trend Analysis

Lake Langano exhibited statistically significant trends paralleling those at Lake Ziway: a significant increase in chlorophyll-a and a significant decline in dissolved oxygen through the Mann-Kendall test. Chlorophyll-a increased significantly (Z = 3.017, p = 0.003, Tau = 0.495, Sen's slope = +0.963 µg/L/year). Dissolved oxygen declined significantly (Z = -3.215, p = 0.001, Tau = -0.526, Sen's slope = -0.235 mg/L/year), indicating an ongoing decline in oxygen availability (Table 3). While both lakes show comparable directional trends, the magnitude of change differs, with Lake Langano's rate of degradation significantly slower than that of Lake Ziway (Table 1).

Table 3. Mann–Kendall Trend Test Results for Water Quality Parameters in Lake Langano

Lake	Parameters	Trend	h	p	Z	Tau	S	Var(S)	Sen's Slope	Intercept
Lake Langano	Chll-a (µg/L)	Increasing	True	0.003	3.017	0.495	94	950	0.963	11.248
	Tur (NTU)	Increasing	True	0.002	3.147	0.516	98	950	0.655	7.178
	TSS (mg/L)	Increasing	True	0.001	3.407	0.558	106	950	1.000	17.600
	DO (mg/L)	Decreasing	True	0.000	-3.666	-0.600	-114	950	-0.235	8.078

3.2.5. Trophic Status

Using the [38] and [39] OECD definitions of mesotrophy and eutrophy (mean chlorophyll-a of 8–25 µg/L indicating mesotrophic conditions, ≥25 µg/L indicating eutrophic conditions), the seasonal variation in trophic status in Lake Langano adheres closely to these thresholds. Mean chlorophyll-a in Lake Langano was 15.09 µg/L during the dry season, with all 10 sampling locations within the mesotrophic range, while in the wet season chlorophyll-a increased to 29.09 µg/L, with 7 of 10 sampling locations classified as eutrophic (Fig. 11 and Table 2). This seasonal shift reflects increased nutrient loading and

algal growth from agricultural, urban, and floriculture-related runoff during the wet season.

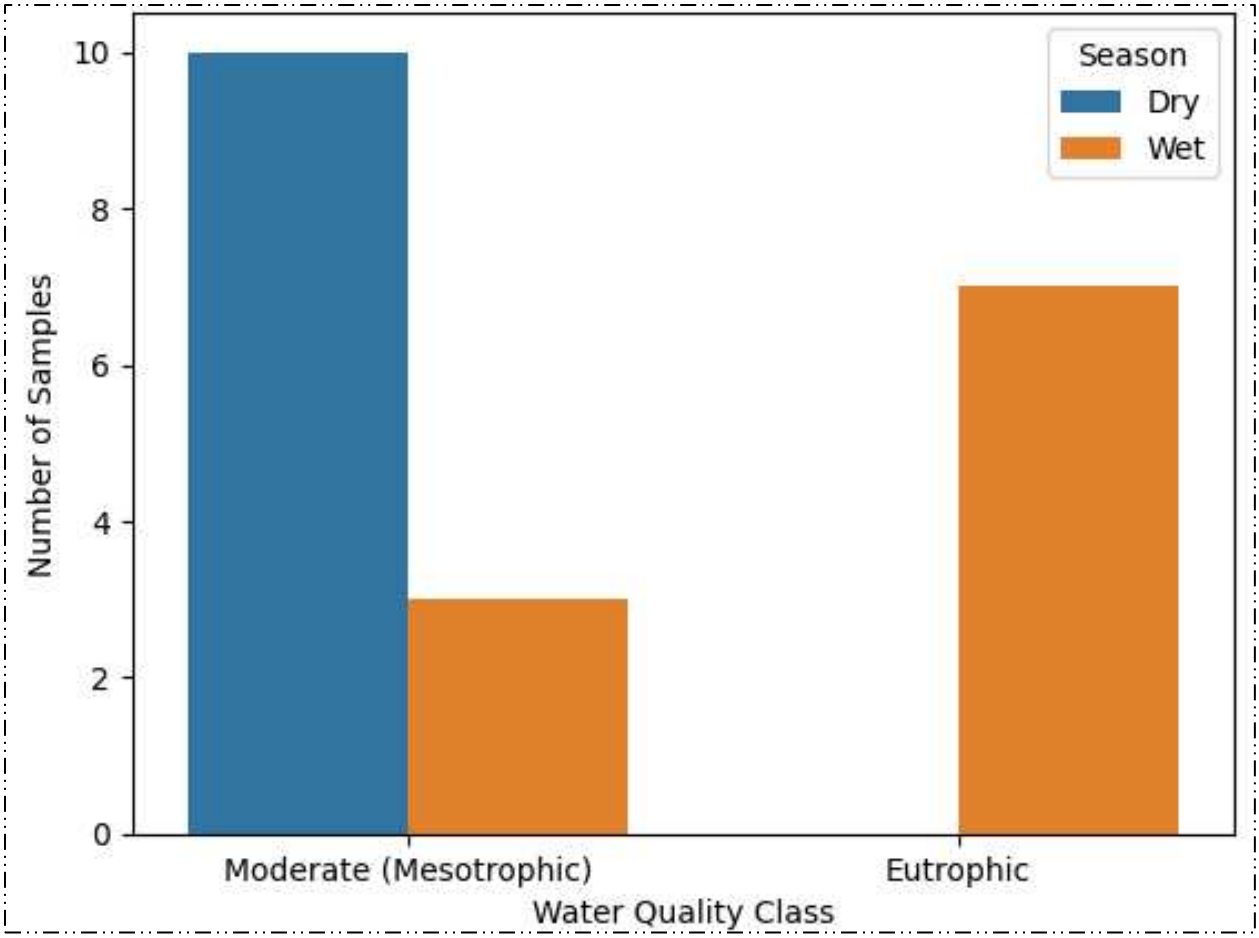


Figure 11. Tropic Status Classification of Lake Langano

3.2.6. Model Performance and Prediction Uncertainty

The Random Forest model predicting chlorophyll-a in Lake Langano also performed well (R^2 = 0.968; RMSE = 1.435; MAE = 1.209; Bias = -0.092). Prediction uncertainty for Langano was substantially lower than for Ziway (1.0–3.5 µg/L, predominantly < 3 µg/L), with more uniformly distributed residuals indicating balanced and consistent predictions (see Supplementary Table 10).

Feature importance for Langano was led by turbidity (SHAP importance = 0.552), followed by dissolved oxygen (0.298) and TSS (0.150), indicating that sediment-related covariates play a somewhat larger relative role here than in Ziway, though the variation among predictors was smaller overall (see Supplementary Table 9). The model indicates a more stable, homogeneous environment than Lake Ziway, supporting greater reliability of the model's projections for Langano (Fig. 12A and B).

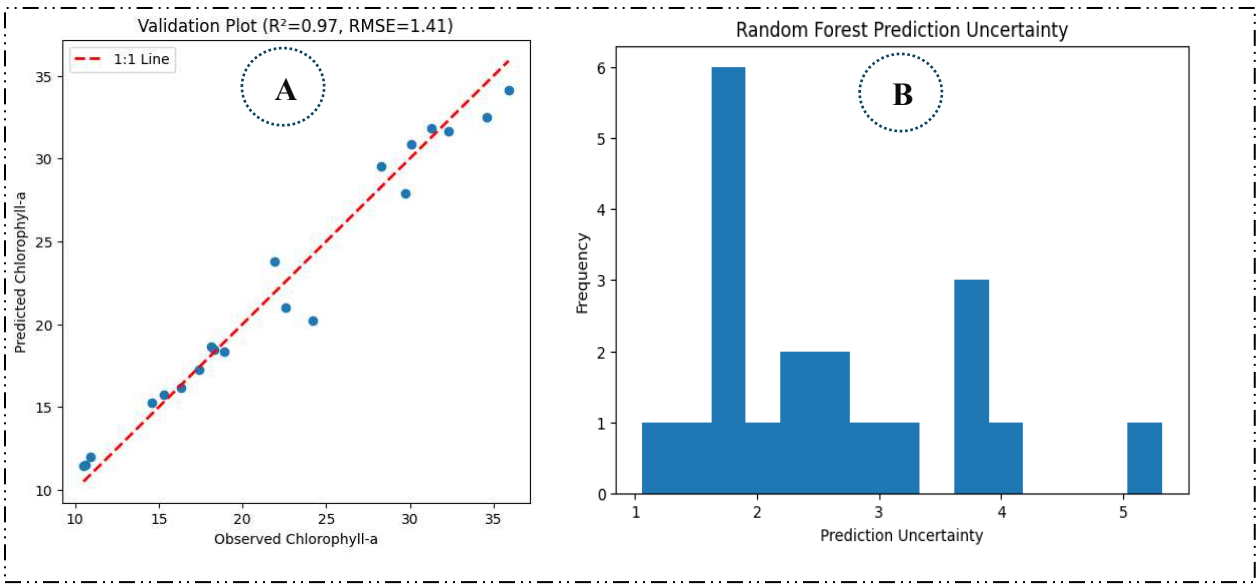


Figure 12. A) Validation of model performance and B) Prediction uncertainty of model for Lake Langano

4. Discussion

4.1. Seasonal Forcing as the Dominant Driver of Chlorophyll-a and Dissolved Oxygen

Hydrological processes occurring throughout the seasons were the leading cause of fluctuation in chlorophyll-a and dissolved oxygen at both Lakes Ziway and Langano, with both lakes experiencing increased chlorophyll-a as rainfall increased runoff and nutrient loading during wet periods. This finding agrees with other studies in the Ethiopian Rift Valley: [40] and [27] both found that seasonality in rainfall is the largest predictor of lake productivity and water quality fluctuation.

The increase in chlorophyll-a concentration was almost identical between Lake Ziway (+18.2 µg/L) and Lake Langano (+14.0 µg/L), similar to findings for Lake Hawassa [14], supporting the conclusion that lakes across the Ethiopian Rift Valley exhibit a shared pattern of nutrient-induced phytoplankton stimulation from wet-season runoff, consistent with findings in other lake ecosystems experiencing seasonal nutrient loading variation [41].

Concomitantly, decreased dissolved oxygen concentrations were recorded during the wet season in both lakes, consistent with the well-documented inverse relationship between eutrophication and dissolved oxygen in highly productive lakes, [18], [30], [42] driven by decomposing organic matter and algal respiration under increased nutrient loading.

4.2. Trophic Status and Eutrophication Dynamics

Lake Ziway transitions from mesotrophic to eutrophic conditions between dry and wet periods, driven by increased nutrient loading from agricultural runoff, urbanization, and water abstraction (Ayenew and Legesse, 2007; Fetahi, 2019; Musie et al., 2020). The high wet-season chlorophyll-a values observed in Ziway (up to 37.83 µg/L) are consistent with

excess phytoplankton growth associated with algal blooms [4], [45]. Given the lake's shallow average depth (about 5 m), Lake Ziway is highly susceptible to water column mixing, raising the risk of anoxia under continued nutrient loading a situation that warrants urgent management attention.

Lake Langano, by contrast, shows a mixed mesotrophic-eutrophic status reflecting comparatively lower pressure from human activities and better internal nutrient and oxygen regulation [46], [47], [48]. The lake's greater depth (~55 m) provides additional stability relative to rapid shifts in trophic status.

Both lakes demonstrate the broader global trend of accelerating eutrophication driven by land use change, underscoring the urgent need for integrated watershed management across the Rift Valley lake systems, with a particular focus on protecting dissolved oxygen balance.

4.3. Model Performance, Uncertainty, and Ecological Complexity

The Random Forest model's predictive performance for chlorophyll-a was highly effective (R^2 = 0.968–0.986; RMSE = 1.230–1.435), exceeding that achieved by alternative modeling efforts for Ethiopian lakes, such as the R^2 of 0.91 reported by [49] for Lake Tana, indicating that the integrated spatial and temporal sampling strategy used in this study improved predictive accuracy.

Lake Ziway showed a wider range of prediction uncertainty (1.2–5.1 μg/L), particularly at high chlorophyll-a values, consistent with nonlinear dynamics and threshold effects common to heterogeneous eutrophic systems [3], [50]. Prediction uncertainty was consistently lower (< 3 μg/L) for Lake Langano, reflecting its greater spatial homogeneity and more stable internal processes [4].

Dissolved oxygen was the dominant predictor of chlorophyll-a in Lake Ziway (SHAP importance = 0.681), while turbidity ranked first in Lake Langano (0.552), with DO second (0.298). This lake-specific difference suggests that, while oxygen depletion dynamics dominate in the shallower, more heavily loaded Lake Ziway, sediment-related processes play a comparatively larger explanatory role for productivity in the deeper, more stable Lake Langano [51], [52].

4.4. Limitations

There are four key limitations when interpreting these results. First, the lack of direct nutrient measurements (total nitrogen and total phosphorus) limits the extent to which causal relationships between nutrient inputs and eutrophication can be established. Future research should measure both nutrients and optical parameters together to strengthen causal interpretation. Second, several physicochemical parameters including pH, water temperature, and electrical conductivity have previously been linked to productivity and oxygen dynamics in lake systems but were not measured in this study [52]. Third, only one seasonal cycle was sampled, so inter-annual variability related to climate variability, drought cycles, and long-term land-use change could not be evaluated, despite their

importance for understanding climate change impacts in the Ethiopian Rift Valley [53], [54].

In addition, SHAP-based explainability identifies important variables only at a correlational level; relationships among variables cannot be interpreted mechanistically without experimental support [50]. Finally, the Random Forest model developed here should not be applied to other lake systems with different morphometric, hydrologic, or land-use characteristics without retraining and validation on those systems.

5. Conclusion and Recommendations

5.1. Conclusion

Chlorophyll-a and dissolved oxygen dynamics in Lakes Ziway and Langano are strongly explained by seasonal variability. During the wet season, chlorophyll-a increased while dissolved oxygen decreased in both lakes, indicating intensified eutrophication pressure.

Lake Ziway exhibited greater variability and more severe degradation than Lake Langano, with lower dissolved oxygen and higher chlorophyll-a. These patterns were confirmed through PCA, which showed that phytoplankton productivity and oxygen depletion define the dominant axis of variability in both lakes.

Random Forest modeling provided high predictive accuracy for chlorophyll-a in both lakes (R^2 = 0.986 for Lake Ziway; R^2 = 0.968 for Lake Langano). SHAP analysis identified dissolved oxygen as the leading or co-leading driver of chlorophyll-a variability across both lakes, with turbidity as a secondary but ecologically meaningful covariate.

The combination of multivariate statistics, Random Forest modeling, and SHAP interpretation provides a robust, transparent, and transferable framework for monitoring productivity and oxygen balance in tropical freshwater ecosystems.

5.2. Recommendations

- Nutrient management should be implemented now, in the form of riparian buffer zones, soil conservation practices, terracing, and climate-smart agriculture, to reduce wet-season nutrient runoff and eutrophication pressure in Lake Ziway [55], [56].
- Preventive watershed and land-use management strategies should be strengthened to maintain Lake Langano's relatively stable oxygen and productivity balance, given the risk of degradation similar to Lake Ziway if preventive measures are not taken. These strategies include controlling nutrient inflow, stabilizing shorelines, and regulating agricultural expansion around inflow areas [52].
- Integrated real-time monitoring of chlorophyll-a and dissolved oxygen, combined with machine learning-based early warning systems using remote sensing and in-situ observations, will assist in early detection of algal blooms and oxygen depletion events in both lakes.

6. Funding

This research was supported by Arba Minch University, Ethiopia, through institutional research funding provided to the corresponding author that no external grant number applies. The funder had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to publish the results.

7. Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The corresponding author has retained the raw field measurements, laboratory analysis records, and processed datasets used to run the statistical and machine learning analyses in order to provide them to other researchers for replication or verification. The corresponding author also holds the Python code used to perform all statistical and machine learning analyses (including PCA, Hierarchical Cluster Analysis, Mann-Kendall Trend Analysis, Random Forest Modeling, and SHAP model explainability), available upon reasonable request.

8. Acknowledgment

Arba Minch University is thanked for providing financial assistance to the author in support of this study. The Ministry of Water and Energy, through its data provision and the issuance of an official letter of support to its respective regional and local Bureau of Water and Energy, provided significant assistance to the author. The author also appreciates the technical assistance provided by senior experts and professionals from the Ministry, who contributed to the production of a reliable dataset for in-situ sample collection and laboratory analyses.

Local administrative offices and the Regional Bureau of Water assisted the author with logistical support and cooperation during the field portion of the research. The author also extends genuine appreciation to Dr. Bisrat Elias (Associate Professor, Principal Advisor) and Dr. Ayano Hirbo (Assistant Professor, Co-Advisor), whose invaluable advice, constructive comments, and continued academic assistance contributed significantly to the completion of this research project.

This study is part of a broader investigation of Lakes Ziway and Langano; a related analysis of turbidity and total suspended solids dynamics in these same lakes is presented in a companion manuscript (Awasho et al., in preparation), submitted separately to a regional journal.

9. Authors Contribution

Afework Alemayehu Awasho (Primary Writer) conceived and conducted the entire research, performed field data gathering with guidance from specialists at the Ministry of Water and Energy, analyzed laboratory samples, led all data processing, analysis, and machine-learning modeling (PCA and Random Forest), and wrote and submitted the manuscript. External technical help was used for some laboratory analysis.

Dr. Bisrat Elias (Associate Professor) provided oversight, refinement of ideas, and feedback on the manuscript. Dr. Ayano Hirbo (Assistant Professor) provided methodological advice,

technical input, and continued oversight of the research and writing. All authors have reviewed and approved the manuscript.

10. Competing Interests

The authors declare that they have no financial or non-financial competing interests in relation to the work described in this manuscript.

11. Ethics Declarations

This study did not involve human participants, human data, or animal subjects, and therefore did not require institutional ethics approval. Field sampling was carried out with an official letter of support and permission from the Ethiopian Ministry of Water and Energy and the respective regional and local Bureaus of Water and Energy.

12. Dual Publication

This manuscript shares its field sampling campaign, study sites, and general survey methodology with a companion manuscript (Paper II, in preparation) that addresses turbidity and total suspended solids (TSS) dynamics in the same two lakes. The two manuscripts analyze distinct water-quality parameters (chlorophyll-a and dissolved oxygen here; turbidity and TSS in the companion paper) with independent statistical and machine-learning treatments, and neither reproduces the results, figures, tables, or conclusions of the other. This overlap in study design and sampling is disclosed here in accordance with the journal's dual-publication policy.

13. Declarations

Consent to publish and participate are not applicable in this study. This is because the only focus of the research was upon sampling the environmental quality of water and conducting laboratory analysis of Lakes Ziway and Langano. There was no inclusion of human subjects or data and no animals were involved in the study.

13. References

[1] B. Aylward *et al.*, “Freshwater Ecosystem Services,” *Ecosyst. Hum. Well-being Curr. State Trends*, pp. 213–255, 2005, [Online]. Available: http://books.google.com/books?hl=en&lr=&id=QYJSziDfTjEC&oi=fnd&pg=PA195&dq=Freshwater+Ecosystem+Services&ots=YewIPMzTzi&sig=UuEYLp3QAzdVhnDIvak0UrdLyG8

[2] M. Getnet, H. Hengsdijk, and M. van Ittersum, “Disentangling the impacts of climate change, land use change and irrigation on the Central Rift Valley water system of Ethiopia,” *Agric. Water Manag.*, vol. 137, pp. 104–115, 2014, doi: 10.1016/j.agwat.2014.02.014.

[3] S. R. Carpenter, N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith, “Nonpoint pollution of surface waters with phosphorus and nitrogen,” *Ecol. Appl.*, vol. 8, no. 3, pp. 559–568, 1998, doi: 10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2.

[4] V. H. Smith and D. W. Schindler, “Eutrophication science: where do we go from

here?," *Trends Ecol. Evol.*, vol. 24, no. 4, pp. 201–207, 2009, doi: 10.1016/j.tree.2008.11.009.

[5] H. Jansen, H. Hengsdijk, D. Legesse, T. Ayenew, P. Spliethoff, and P. Hellegers, "Ecosystems for water , food and economic development in the Central Rift Valley , Ethiopia : First conclusions," no. January, pp. 3–6, 2007.

[6] A. O. Hughes, R. J. Davies-Colley, and A. H. Elliott, "Measurement of light attenuation extends the application of suspended sediment monitoring in rivers," *IAHS-AISH Proc. Reports*, vol. 367, no. December 2014, pp. 170–176, 2014, doi: 10.5194/piahs-367-170-2015.

[7] M. Mamun, U. Atique, and K. G. An, "Assessment of water quality based on trophic status and nutrients-chlorophyll empirical models of different elevation reservoirs," *Water (Switzerland)*, vol. 13, no. 24, 2021, doi: 10.3390/w13243640.

[8] M. J. Brush *et al.*, "Eutrophication, harmful algae, oxygen depletion, and acidification," *Coast. Ecosyst. Transit. A Comp. Anal. North. Adriat. Chesap. Bay*, pp. 75–104, 2020, doi: 10.1002/9781119543626.ch5.

[9] P. Larsson, R. Rosenberg, P. Jonsson, M. Olsson, and D. Broman, "Eutrophication and Contaminants in Aquatic Ecosystems," vol. 29, no. 4, pp. 184–194, 2000.

[10] M. R. Msomphora, "Eutrophication of the East African Great Lakes," pp. 279–289, 1993.

[11] E. S. Leggesse, F. A. Zimale, D. Sultan, T. Enku, and S. A. Tilahun, "Corrigendum: Advancing non-optical water quality monitoring in Lake Tana, Ethiopia: insights from machine learning and remote sensing techniques (Frontiers in Water, (2024), 6, (1432280), 10.3389/frwa.2024.1432280)," *Front. Water*, vol. 6, 2024, doi: 10.3389/frwa.2024.1489691.

[12] T. Ayenew, "Environmental implications of changes in the levels of lakes in the Ethiopian Rift since 1970," *Reg. Environ. Chang.*, vol. 4, no. 4, pp. 192–204, 2004, doi: 10.1007/s10113-004-0083-x.

[13] D. Tibebe, Y. Tesfaye, and Y. Kassa, "The impact of sewage effluents on water quality of Lake Hawassa, Ethiopia," *BMC Chem.*, vol. 17, no. 1, pp. 1–12, 2023, doi: 10.1186/s13065-023-00954-x.

[14] S. M. Lencha, J. Tränckner, and M. Dananto, "Assessing the water quality of lake hawassa Ethiopia—Trophic state and suitability for anthropogenic uses—applying common water quality indices," *Int. J. Environ. Res. Public Health*, vol. 18, no. 17, 2021, doi: 10.3390/ijerph18178904.

[15] D. W. Goshime, A. T. Haile, R. Absi, and B. Ledésert, "Impact of water resource development plan on water abstraction and water balance of Lake Ziway, Ethiopia," *Sustain. Water Resour. Manag.*, vol. 7, no. 3, 2021, doi: 10.1007/s40899-021-00516-w.

[16] T. Miller *et al.*, "Predictive Modeling of Urban Lake Water Quality Using Machine Learning: A 20-Year Study," *Appl. Sci.*, vol. 13, no. 20, 2023, doi: 10.3390/app132011217.

[17] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, "Random Forests," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12343 LNCS, pp. 503–515, 2020, doi: 10.1007/978-3-030-62008-0_35.

[18] S. Shrestha and F. Kazama, “Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan,” *Environ. Model. Softw.*, vol. 22, no. 4, pp. 464–475, 2007, doi: 10.1016/j.envsoft.2006.02.001.

[19] K. P. Singh, A. Malik, D. Mohan, and S. Sinha, “Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) - A case study,” *Water Res.*, vol. 38, no. 18, pp. 3980–3992, 2004, doi: 10.1016/j.watres.2004.06.011.

[20] D. Legesse and T. Ayenew, “Effect of improper water and land resource utilization on the central Main Ethiopian Rift lakes,” *Quat. Int.*, vol. 148, no. 1, pp. 8–18, May 2006, doi: 10.1016/J.QUAINT.2005.11.003.

[21] Z. C. Hao, L. Li, J. H. Wang, Z. H. Wang, and X. L. Shi, “Impact of climate change on surface water resources,” *Diqiu Kexue - Zhongguo Dizhi Daxue Xuebao/Earth Sci. - J. China Univ. Geosci.*, vol. 32, no. 3, pp. 425–432, 2007.

[22] D. Abera, B. Berehanu, and T. Ayenew, “Heliyon Groundwater flow system of the Abijata-Langano-Ziway lakes basin , Ethiopia,” *Heliyon*, vol. 10, no. 23, p. e40831, 2024, doi: 10.1016/j.heliyon.2024.e40831.

[23] T. Ayenew, “Water management problems in the Ethiopian rift: Challenges for development,” *J. African Earth Sci.*, vol. 48, no. 2–3, pp. 222–236, Jun. 2007, doi: 10.1016/J.JAFREARSCI.2006.05.010.

[24] A. Sisay, “Remote Sensing Based Water Surface Extraction and Change Detection in the Central Rift Valley Region of Ethiopia,” *Am. J. Geogr. Inf. Syst.*, vol. 5, no. 2, pp. 33–39, 2016, doi: 10.5923/j.ajgis.20160502.01.

[25] APHA, “Standard methods for the examination of water and wastewater,” *Water Pollut. Control*, 2017.

[26] U. EPA, “National recommended water quality criteria,” *United States Environ. Prot. Agency’, Off. Water, Off. Sci. Technol.*, 2009.

[27] D. Tibebe, F. Zewge, B. Lemma, and Y. Kassa, “Assessment of spatio-temporal variations of selected water quality parameters of Lake Ziway, Ethiopia using multivariate techniques,” *BMC Chem.*, vol. 16, no. 1, pp. 1–18, 2022, doi: 10.1186/s13065-022-00806-0.

[28] O. E. Okoth, M. Mucai, W. A. Shivoga, S. N. Miller, J. Rasowo, and C. C. Ngugi, “Spatial and seasonal variations in phytoplankton community structure in alkaline-saline Lake Nakuru, Kenya,” *Lakes Reserv. Sci. Policy Manag. Sustain. Use*, vol. 14, no. 1, pp. 57–69, 2009, doi: 10.1111/j.1440-1770.2009.00392.x.

[29] S. W. Jeffrey and G. F. Humphrey, “New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton,” *Biochem. und Physiol. der Pflanz.*, vol. 167, no. 2, pp. 191–194, 1975, doi: 10.1016/s0015-3796(17)30778-3.

[30] B. B. Benson and D. Krause, “The concentration and isotopic fractionation of oxygen dissolved in freshwater and seawater in equilibrium with the atmosphere,” *Limnol. Oceanogr.*, vol. 29, no. 3, pp. 620–632, 1984, doi: 10.4319/lo.1984.29.3.0620.

[31] P. K. Yadava, H. Kumar, A. Singh, V. Kumar, and S. Verma, “Impact of climate change on water quality and its assessment,” *Vis. Tech. Clim. Chang. with Mach. Learn. Artif. Intell.*, pp. 39–54, Jan. 2023, doi: 10.1016/B978-0-323-99714-0.00002-9.

[32] D. Pooja, P. Singh, S. Patil, and P. Kumar, *Role of Materials in Sensors for Water Pollutants Monitoring*. 2020. doi: 10.1007/978-981-15-0671-0_10.

[33] D. R. Helsel, R. M. Hirsch, K. R. Ryberg, S. A. Archfield, and E. J. Gilroy, "Statistical methods in water resources," *U.S. Geol. Surv. Tech. Methods*, vol. 2020, no. 4-A3, pp. 1–484, 2020, doi: 10.3133/tm4a3.

[34] C. Libiseller and A. Grimvall, "Performance of partial Mann-Kendall tests for trend detection in the presence of covariates," *Environmetrics*, vol. 13, no. 1, pp. 71–84, 2002, doi: 10.1002/env.507.

[35] B. Ratner, S. Day, and C. Davies, *Statistical and Machine-Learning Data Mining*. 2011. doi: 10.1201/b11508.

[36] L. Breiman, "RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis," *Kluwer Acad. Publ.*, vol. 12343 LNCS, pp. 503–515, 2020, doi: 10.1007/978-3-030-62008-0_35.

[37] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, no. 1, pp. 79–82, 2005, doi: 10.3354/cr030079.

[38] M. Karydis, "Eutrophication assessment of coastal waters based on indicators: A literature review," *Glob. Nest J.*, vol. 11, no. 4, pp. 373–390, 2009, doi: 10.30955/gnj.000626.

[39] E. S. P. B. V and D. C. Grobler, "A S S E S S M E N T OF THE IMPACT OF EUTROPHICATION CONTROL," vol. 31, pp. 237–247, 1986.

[40] T. Ayenew and D. Legesse, "The changing face of the Ethiopian rift lakes and their environs: Call of the time," *Lakes Reserv. Sci. Policy Manag. Sustain. Use*, vol. 12, no. 3, pp. 149–165, 2007, doi: 10.1111/j.1440-1770.2007.00332.x.

[41] N. Jargal, U. Atique, M. Mamun, and K. G. An, "Seasonal and long-term connections between trophic status, sestonic chlorophyll, nutrients, organic matter, and monsoon rainfall in a multipurpose reservoir," *Water (Switzerland)*, vol. 13, no. 13, 2021, doi: 10.3390/w13131720.

[42] J. Rakocevic-Nedovic and H. Hollert, "Phytoplankton community and chlorophyll a as trophic state indices of Lake Skadar (Montenegro, Balkan)," *Environ. Sci. Pollut. Res.*, vol. 12, no. 3, pp. 146–152, 2005, doi: 10.1065/espr2005.04.241.

[43] T. Fetahi, "Eutrophication of Ethiopian water bodies: a serious threat to water quality, biodiversity and public health," *African J. Aquat. Sci.*, vol. 44, no. 4, pp. 303–312, 2019, doi: 10.2989/16085914.2019.1663722.

[44] M. Musie, S. Sen, and I. Chaubey, "Hydrologic responses to climate variability and human activities in Lake Ziway Basin, Ethiopia," *Water (Switzerland)*, vol. 12, no. 1, pp. 1–27, 2020, doi: 10.3390/w12010164.

[45] V. H. Smith, "Eutrophication of freshwater and coastal marine ecosystems: A global problem," *Environ. Sci. Pollut. Res.*, vol. 10, no. 2, pp. 126–139, 2003, doi: 10.1065/espr2002.12.142.

[46] J. Lan, P. Liu, X. Hu, and S. Zhu, "Harmful Algal Blooms in Eutrophic Marine Environments :," pp. 1–64, 2024.

[47] G. Beneberu and S. Mengistou, "Oligotrophication trend of lake Ziway, Ethiopia,"

Ethiop. J. Sci., vol. 32, no. 2, pp. 141–148, 2009.

[48] E. Zinabu, P. Kelderman, J. van der Kwast, and K. Irvine, "Monitoring river water and sediments within a changing Ethiopian catchment to support sustainable development," *Environ. Monit. Assess.*, vol. 191, no. 7, 2019, doi: 10.1007/s10661-019-7545-6.

[49] E. S. Leggesse, F. A. Zimale, D. Sultan, T. Enku, and S. A. Tilahun, "Advancing non-optical water quality monitoring in Lake Tana, Ethiopia: insights from machine learning and remote sensing techniques (Frontiers in Water, (2024), 6, (1432280), 10.3389/frwa.2024.1432280)," *Front. Water*, vol. 6, 2024, doi: 10.3389/frwa.2024.1489691.

[50] J. D. Olden, M. K. Joy, and R. G. Death, "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data," *Ecol. Modell.*, vol. 178, no. 3–4, pp. 389–397, 2004, doi: 10.1016/j.ecolmodel.2004.03.013.

[51] T. P. Simon, R. Morehouse, and S. L. Worden, "Physical and Chemical Limnology of Three Lakes Within Hoosier National Forest," 2012.

[52] D. W. Schindler, "Recent advances in the understanding and management of eutrophication," *Limnol. Oceanogr.*, vol. 51, no. 1 II, pp. 356–363, 2006, doi: 10.4319/lo.2006.51.1_part_2.0356.

[53] M. Geng *et al.*, "Inter-annual and intra-annual variations in water quality and its response to water-level fluctuations in a river-connected lake, Dongting Lake, China," *Environ. Sci. Pollut. Res.*, vol. 29, no. 10, pp. 14083–14097, 2022, doi: 10.1007/s11356-021-16739-5.

[54] Y. E. Moon and H. S. Kim, "Inter-annual and seasonal variations of water quality and trophic status of a reservoir with fluctuating monsoon precipitation," *Int. J. Environ. Res. Public Health*, vol. 18, no. 16, 2021, doi: 10.3390/ijerph18168499.

[55] M. Musie, A. Momblanch, and S. Sen, "Exploring future global change-induced water imbalances in the Central Rift Valley Basin, Ethiopia," *Clim. Change*, vol. 164, no. 3–4, 2021, doi: 10.1007/s10584-021-03035-x.

[56] S. T. Tantu, S. M. Yilma, and K. Q. Mary, "Water quality characterization of Lake Abaya, a water body under increasing anthropogenic pressure in Southern Ethiopia," *Sustain. Water Resour. Manag.*, vol. 10, no. 4, 2024, doi: 10.1007/s40899-024-01119-x.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFile.docx