


Article

Evaluation of Seasonal and Spatial Variations in Water Quality and Identification of Potential Sources of Pollution Using Multivariate Statistical Techniques for Lake Hawassa Watershed, Ethiopia

Semaria Moga Lencha ^{1,2,*} , Mihret Dananto Ulsido ^{2,3} and Alemayehu Muluneh ²¹ Faculty of Agriculture and Environmental Sciences, University of Rostock, 18051 Rostock, Germany² Faculty of Biosystems and Water Resource Engineering, Institute of Technology, Hawassa University, Hawassa P.O. Box 05, Ethiopia; mihret@gmail.com (M.D.U.); muluneh96@yahoo.com (A.M.)³ Center for Ethiopian Rift Valley studies, Hawassa University, Hawassa P.O. Box 05, Ethiopia

* Correspondence: semaria.lencha@uni-rostock.de or semaria@hu.edu.et; Tel.: +49-1521-121-2094

Abstract: The magnitude of pollution in Lake Hawassa has been exacerbated by population growth and economic development in the city of Hawassa, which is hydrologically closed and retains pollutants entering it. This study was therefore aimed at examining seasonal and spatial variations in the water quality of Lake Hawassa Watershed (LHW) and identifying possible sources of pollution using multivariate statistical techniques. Water and effluent samples from LHW were collected monthly for analysis of 19 physicochemical parameters during dry and wet seasons at 19 monitoring stations. Multivariate statistical techniques (MVST) were used to investigate the influences of an anthropogenic intervention on the physicochemical characteristics of water quality at monitoring stations. Through cluster analysis (CA), all 19 monitoring stations were spatially grouped into two statistically significant clusters for the dry and wet seasons based on pollution index, which were designated as moderately polluted (MP) and highly polluted (HP). According to the study results, rivers and Lake Hawassa were moderately polluted (MP), while point sources (industry, hospitals and hotels) were found to be highly polluted (HP). Discriminant analysis (DA) was used to identify the most critical parameters to study the spatial variations, and seven significant parameters were extracted (electrical conductivity (EC), dissolved oxygen (DO), chemical oxygen demand (COD), total nitrogen (TN), total phosphorous (TP), sodium ion (Na⁺), and potassium ion (K⁺) with the spatial variance to distinguish the pollution condition of the groups obtained using CA. Principal component analysis (PCA) was used to qualitatively determine the potential sources contributing to LHW pollution. In addition, three factors determining pollution levels during the dry and wet season were identified to explain 70.5% and 72.5% of the total variance, respectively. Various sources of pollution are prevalent in the LHW, including urban runoff, industrial discharges, diffused sources from agricultural land use, and livestock. A correlation matrix with seasonal variations was prepared for both seasons using physicochemical parameters. In conclusion, effective management of point and non-point source pollution is imperative to improve domestic, industrial, livestock, and agricultural runoff to reduce pollutants entering the Lake. In this regard, proper municipal and industrial wastewater treatment should be complemented, especially, by stringent management that requires a comprehensive application of technologies such as fertilizer management, ecological ditches, constructed wetlands, and buffer strips. Furthermore, application of indigenous aeration practices such as the use of drop structures at critical locations would help improve water quality in the lake watershed.

Keywords: monitoring; mitigations; spatial and temporal variabilities; principal component analysis; cluster analysis; discriminant analysis; water quality; pollution; correlation



Citation: Lencha, S.M.; Ulsido, M.D.; Muluneh, A. Evaluation of Seasonal and Spatial Variations in Water Quality and Identification of Potential Sources of Pollution Using Multivariate Statistical Techniques for Lake Hawassa Watershed, Ethiopia. *Appl. Sci.* **2021**, *11*, 8991. <https://doi.org/10.3390/app11198991>

Academic Editor:
Jorge Rodríguez-Chueca

Received: 2 August 2021
Accepted: 23 September 2021
Published: 27 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Studies have shown that urban, agricultural, and industrial discharges have a direct effect on surface water quality. Similarly, urban wastewaters cause fecal contamination of surface waters, and urban stormwater runoff, which contains large amounts of fecal microbes, also affects surface water quality [1]. Surface water bodies are vital natural resources that are vulnerable to pollution. The contaminants are chemical, physical, and biological constituents resulting from anthropogenic activities and are of greater environmental consideration [2]. Surface water bodies are extensively used as the major sources for domestic, non-domestic, industrial, and irrigation purposes. Therefore, monitoring and assessment of water bodies is imperative to obtain reliable information on water quality for effective management [3]. Anthropogenic uses of the waterbodies in the study basin can degrade the quality of surface water and impair its usability as potable water supply or for industry, agriculture, recreation, or other purposes. Hence, regular monitoring of water quality of rivers and lake is indispensable [4,5]. The most affected river stretches are those that flow through urbanized and exceedingly populated urban areas where there is no adequate sanitation. Upstream rural areas are mainly affected by pollutants from non-point sources such as agricultural runoff, whereas urban areas are polluted by point sources, sewage discharges, urban runoff, and pollutants from upstream areas [6,7].

Studies have shown that some lakes and wetlands around the world have disappeared or are showing changes in their ecosystem. Furthermore, factors such as intensive land use for urbanization and agriculture have had significant impact on the hydrology, ecology, and ecosystem services of lakes, which has eventually led to a decline in lake levels [8]. In addition, pollutants have long been a concern, as their accumulation can have serious effects on fauna, flora, and human health when the huge amount of urban and industrial wastewater reaches the shores [9].

Lake Hawassa is located near the city of Hawassa and is surrounded by agricultural land, industries and residential areas. Therefore, it is susceptible to a variety of pollutants that enter the lake directly or indirectly. On the other hand, the Lake Hawassa Watershed is experiencing rapid land cover change, and natural resources have overwhelmingly diminished. The lake is hydrologically closed and has no apparent outlet, so all pollutants entering the lake are retained. As a result, the lake faces numerous problems, and the water quality deteriorates over time, threatening biodiversity [10].

Significant industrialization, augmented with rapid urbanization and increasing economic development, has increased the extent of pollution [11]. The pollution is mainly from non-point sources caused by urban and agricultural runoff, overgrazing, deforestation, soil erosion, land development, and industrial effluents. This leads to numerous environmental concerns that have resulted in substantial hydrological disturbances. The main factories in the study area are a ceramics factory, a flourmill, a cement products factory, a Moha soft drink factory, a BGI (St. George Brewery factory), an Etabs soap factory, an industrial park in Hawassa, and other small-scale industries. They are virtually all concentrated along the main road, which is close to the shallow swamp, and discharge their effluents into the lake through streams. On the other hand, deforestation and irrigation of the land have caused the drying up of Lake Cheleleka by reducing the streamflow [12].

Various studies have been conducted to examine water quality in the LHW catchment and identify sources of pollution. Teshome [11] investigated the eastern catchment of Lake Hawassa Watershed to assess the seasonal water quality and its suitability for the designated uses. The findings revealed that the rivers in the eastern part of Lake Hawassa Watershed are suitable for agriculture and livestock but unpleasant for aquatic life, and the lake is hypereutrophic.

Amare [13] investigated the primary sources of non-point source pollution and their relative contribution in Lake Hawassa Watershed using the Annualized Agricultural Non-Point Source (AnnAGNPS) model. The pollutant-loading model revealed non-point source pollutants originating from agricultural lands and associated with deleterious anthropogenic activities responsible for the water quality impairment of Lake Hawassa.

These point sources have been determined to be the source of numerous pollutants in the lake ecosystem if the effluent control system put in place is unsuitable [14].

Kebede [15] studied the impact of land cover changes on water quality and streamflow in Lake Hawassa Watershed and concluded that water quality in the upper watershed of the three rivers was better than the lower sections of the catchment with respect to the parameters studied, which might be correlated to the observed land use.

A study conducted by Lencha et al. [16] at Lake Hawassa revealed that most of the population, including the inner part of the city, are using latrines. Larger buildings have conventional flushing systems but without any wastewater treatment. Furthermore, industrial and commercial point sources are known to discharge their effluents into streams or rivers that end up in the Lake. In addition, Hawassa Industrial Park and Referral Hospital discharge their effluents directly into the lake. This is a threat to the people who rely on rivers, streams, and the lake for domestic and other purposes and to the survival of aquatic life as well.

To sum up, some studies regarding the water quality have been conducted in either the eastern or the western catchment of Lake Hawassa, while others have been carried out only at Lake Hawassa. Nonetheless, there is no sufficient water quality study to connect agricultural and urban land use with the watershed pollution level to identify the sources of pollution. The previous studies mainly relied on random monitoring and data from literature and focused only on a few water quality parameters, which cannot reflect the whole picture of water quality in the watershed. Additionally, some previous studies also obtained contradictory findings. On the other hand, urbanization, industrialization, commercial activities, and population growth are increasing rapidly, which could increase sewage and effluents production. Through monitoring data, consistent data analysis, and homogenization of parameters, this study aimed to (1) statistically analyze multiple-parameter data by using principal component analysis (PCA), cluster analysis (CA), and discriminant analysis (DA); (2) investigate the broad-spectrum variation in the parameters of LHW; and (3) cluster monitoring stations with similar characteristics and identify potential sources of pollution in LHW.

2. Materials and Methods

2.1. Study Area

Lake Hawassa Watershed (LHW) is located 275 km from the capital Addis Ababa, in the capital of Sidama regional state, on the main road leading to Nairobi, Kenya via Moyale. LHW has a total area of 1431 km² and lies between 6°45' to 7°15' N latitude and 38°15' to 38°45' E longitude (Figure 1). LHW comprises five sub-watersheds [17].

The watershed is known for its flat plains and dismembered undulating landscape with elevation ranging from 1571 to 2962 m above sea level [18]. The area comprises mountains and low-lying areas, with a wide flat wetland called Cheleleka. Perennial rivers and streams on the north and northeast sides of the catchment and runoff on the east wall feed Cheleleka. The sub-basin of Tikur-Wuha consists of only a tributary called Tikur-Wuha that flows into Lake Hawassa. In this lake system, no surface water flows out from the lake except by evaporation and abstraction, so the catchment can be considered hydrologically closed [15]. The climate of the Hawassa sub-basin is sub-humid and distinctly seasonal. The months from April to October are wet and humid, and the main rainy season is between July and September, with a mean annual precipitation of about 955 mm. The mean minimum precipitation is 17.8 mm in December (dry season) and the mean maximum precipitation is 119.8 mm in August (rainy season) [19].

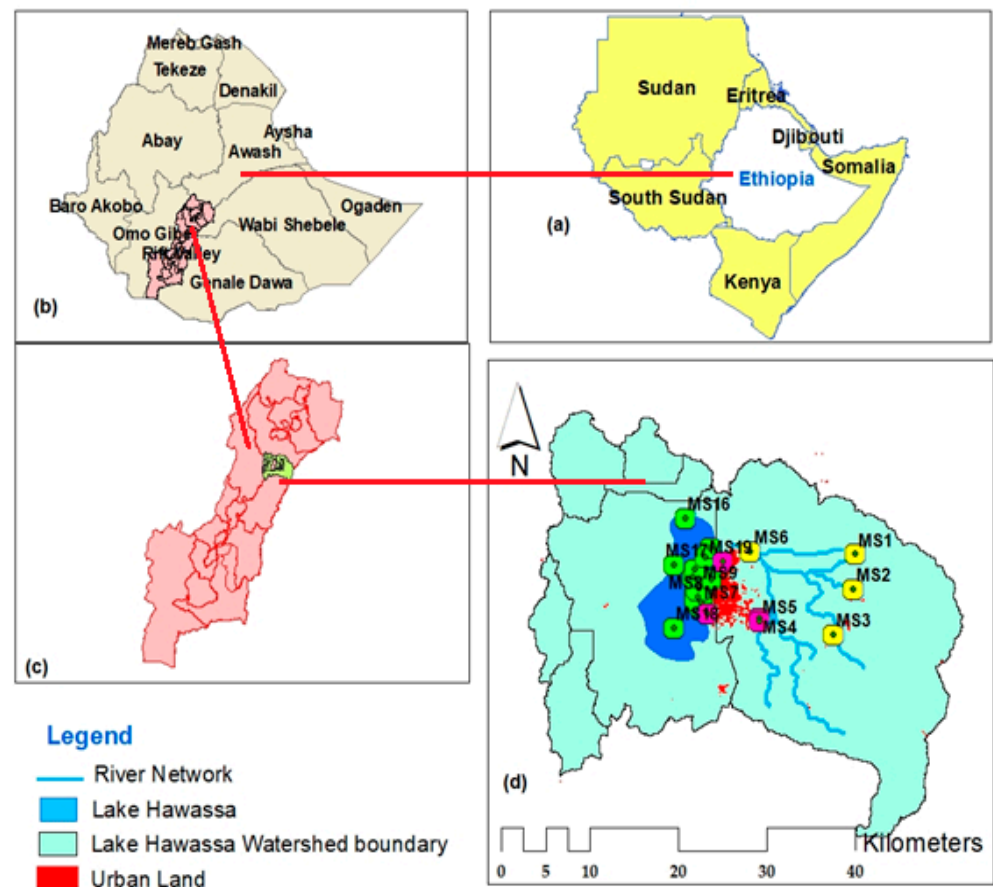


Figure 1. Study area map and monitoring station locations (a) = countries sharing boundaries with Ethiopia, (b) = major river basins in Ethiopia, (c) = Rift Valley lake basin and (d) = Lake Hawassa sub-basin and monitoring stations).

2.2. Sampling and Monitoring Parameters

The monitoring sites and sampling strategy were planned to cover a wide range of factors contributing to the water quality of the river, taking into account tributaries and point sources whose effluents end up in the lake and have a substantial impact on the water quality of the lake. The criteria for selecting monitoring points were hydrological, with confluence of sub-basins having distinct characteristics and land use types, with the intention of transferring parameters to unmonitored sub-basins. Furthermore, factors such as availability of point and non-point sources, land use type, and urban and wastewater drains were considered in the selection of monitoring sites.

Hence, a total of nineteen (19) monitoring stations were selected (Table 1 and Figure 1). Four (4) monitoring sites were selected purposively at the Weshu, Hallo, Wedessa, and Tikur-Wuha river mouths of the respective sub-watersheds.

Eleven (11) monitoring sites were distributed evenly along the entire course of Lake Hawassa for water quality monitoring. Three (3) monitoring sites were selected near the industrial disposal site, and one (1) was at the health care center.

The monitoring sites in the Tikur-Wuha catchment were Weshu River (MS1), Hallo River (MS2), and Wedessa River (MS3), which are located in the upstream part of Lake Hawassa, where agricultural runoff from the catchment flows directly or through its tributaries into the Cheleleka wetland. The three rivers were purposively selected based on their size and spatial location to represent their respective sub-basins. Monitoring station 6 (MS6) is a critical area with mostly fresh water where factories discharge their effluent into the Tikur-Wuha River, and the river eventually flows into Lake Hawassa. This is an area where river inputs to the lake are high.

Table 1. Monitoring stations in Lake Hawassa Watershed.

No	Monitoring Stations	Site Code	Location
1	Wesha River	MS1	LHW upstream
2	Hallo River	MS2	LHW upstream
3	Wedessa River	MS3	LHW upstream
4	BGI effluent discharge site	MS4	LHW middle
5	Pepsi factory oxidation pond	MS5	LHW middle
6	Tikur-Wuha River	MS6	LHW middle
7	Amora-Gedel (fish market)	MS7	Eastern side of LH
8	Amora-Gedel (Gudumale)	MS8	Eastern side of LH
9	Nearby Lewi resort	MS9	Eastern side of LH
10	Fikerhayk center (FH)	MS10	Center of LH
11	Fikerhayk (meznegna)	MS11	Eastern side of LH
12	Center of LH (Towards HR)	MS12	Center of LH
13	Nearby Haile resort	MS13	Eastern side of LH
14	Tikur-Wuha site	MS14	Eastern side of LH
15	Referral Hospital	MS15	Eastern side of LH
16	Ali-Girma site (opposite to HR)	MS16	Western side of LH
17	Sima Site (opposite to Mount Tabor)	MS17	Western side of LH
18	Dore-Bafana Betemengist	MS18	Southern part of LH
19	Hawassa Industrial Park	MS19	LHW middle

The site codes are indicated in Figure 1. FH designates Fikerhayk, HR labels Haile Resort, LHW designates Lake Hawassa Watershed, LH designates Lake Hawassa.

Monitoring sites for point sources were selected from available industries in the catchment that directly or indirectly feed Lake Hawassa. The selected sites were the St. George Brewery factory, BGI (MS4), and the Moha soft drink factory (MS5), whose effluents discharge into the Cheleleka wetland and eventually enter Lake Hawassa via Tikur-Wuha River, as well as the Referral Hospital (MS15) and Hawassa Industrial park (MS19), which discharge their effluents directly in to Lake Hawassa.

The monitoring stations for Lake Hawassa were selected based on the presence of major pollution sources in the lake, existence of point sources, health facilities, industrial effluent emission sites, availability of boating and recreational activities, presence of service rendering facilities such as Haile and Lewi resorts, fish market (Amora-Gedel and Gudumale), and also the central part of the lake where the disturbance is minimum.

For this purpose, eight (8) monitoring sites were selected in the eastern part (northeast to southeast) of the lake and designated as MS7, MS8, MS9, MS10, MS11, MS12, MS13, and MS14.

The other three (3) monitoring sites were located on the western (northwest to southwest) sides of the lake and were designated as MS16 for the local village Ali-Girma site (opposite Haile Resort), MS17 for Sima site that is opposite side of Mount Tabor, and MS18 for Dore-Bafana Betemengist site. In this part of the lake, although there is no point source pollution, there is enormous anthropogenic activity in the form of non-point source pollution from recreational activities, agricultural runoff, and animal waste.

The analyses of physicochemical water quality parameters at selected sites and periods were conducted from May 2020 to January 2021 to see seasonal variation. Sample collection for the wet season was event-based, i.e., samples were collected after rainfall events. The coordinates of each sampling stations was determined using GNSS.

Composite samples were collected in pre-cleaned 2L polyethylene plastic bottles (sterilized glass bottles were used for biochemical oxygen demand (BOD) and chemical oxygen demand (COD) analyses) for different parameters. The bottles were washed with concentrated nitric acid and distilled water before sample collection and thoroughly rinsed with sample water during collection to avoid possible contamination. The water samples were aseptically handled, labelled, preserved in sterile glass bottles, stored in the cooler (Mobicool v30 AC/DC, Germany) and ice box, and transported to the laboratory of Hawassa University Environmental Engineering, Addis Ababa City Government Environmental Protection, and Green Development Commission and Engineering Corporation of Oromiya for analysis.

The collection, handling, preservation, and treatment of the water samples followed the standard methods outlined for the examination of water and wastewater by the American Public Health Association guidelines [20] and all the parameters were presented with their respective analytical methods and instruments used for analysis in Table 2 below.

Table 2. Analytical methods and instruments used for analysis.

Parameter	Analytical Method and Instrument
pH, EC, TDS, and Temperature	Portable multi-parameter analyzer (Zoto, Germany)
Turbidity	Nephelometric (Hach, model 2100A)
DO	Modified Winkler
BOD	Manometric, BOD sensor
COD	Closed Reflux, colorimetric
SRP and TP	Spectrophotometrically by molybdovanadate (Hach, model DR 3900)
TN	Spectrophotometrically by TNT Persulfate digestion (Hach, model DR 3900)
NO ₂ ⁻ and TAN (NH ₃ -N + NH ₄ -N)	Spectrophotometrically by salicylate (Hach, model DR 3900)
NO ₃ ⁻	Photometric measurements, Wagtech Photometer 7100 at 520 nm wavelength
SS	Filtration by standard glass fiber filter
Mg ²⁺ , Na ⁺ , Ca ²⁺ , and K ⁺	Atomic Absorption Spectrophotometer, AAS, model NOVAA400

Total ammonium nitrogen (TAN), electrical conductivity (EC), total dissolved solids (TDS), dissolved oxygen (DO), biochemical oxygen demand (BOD₅), chemical oxygen demand (COD), soluble reactive phosphorous (SRP), total phosphorous (TP), nitrate (NO₃⁻), nitrite (NO₂⁻), magnesium ion (Mg⁺²), sodium ion (Na⁺), potassium ion (K⁺), calcium ion (Ca⁺²), and suspended solids (SS).

Un-Ionized Ammonia Determination from Total Ammonium Nitrogen (TAN)

The un-ionized free ammonia was calculated by the mass action law in its logarithmic form (1). The pK_a as function of temperature was taken from Emerson et al. [21]:

$$\% \text{ Un-ionized NH}_3 - \text{N} = \frac{1}{(1 + 10^{(\text{pK}_a - \text{pH})})} \quad (1)$$

$$\text{pK}_a = \frac{0.09108 + 2729.92}{(T_k)} \quad (2)$$

where T_k is temperature in Kelvins (273 + °C).

3. Multivariate Statistical Techniques and Data Treatment

3.1. Multivariate Statistical Techniques

Multivariate statistical techniques (MVST) are a valuable tool to estimate efficiently the spatio-temporal variability in a watershed and the influences of human intervention on the characteristics of physicochemical parameters at monitoring stations [22]. In addition, MVST like cluster analysis (CA), discriminant analysis (DA), and PCA/factor analysis can be implemented to interpret complex databases to offer better visualization of water quality in the studied watershed [23]. The statistical techniques PCA, CA, and DA are vital to determine the primary relationships among the physicochemical parameters measured in experimental data standardized to the Z-scale to avoid inaccurate grouping because of the huge variability in the data dimensionality [5,24–26].

Principal component analysis (PCA), cluster analysis (CA) and discriminant Analysis (DA) were carried out to examine the seasonal variations, identify possible pollution sources, and analyze and interpret surface water quality data to draw meaningful information in China [2,7,27–30], South Asia—Bangladesh [31], the Middle East—Iran [3], India [23,32,33], South African [34], Ethiopia [22,35], South Asia Malaysia [36], the Middle East—Lebanon [6,37], Spain [38], and Serbia [39].

XLSTAT 2016 (Addinsoft, New York, USA), Microsoft Excel 2016, and “Statistical Package for the Social Sciences Software, IBM SPSS 25 for Windows” were employed to perform statistical analysis integrally.

3.2. Data Treatment and Multivariate Statistical Methods

PCA is sensitive to outliers, missing data, and poor linear correlation among variables due to insufficient assigned variables. Thus, the data treatment needs to be performed for missing data and outliers in the monitored water quality data before executing multivariate statistical analysis. There might be a real shift in the value of an observation that arises from non-random causes. In this study, outliers were detected according to Grubbs [40] test method using XLSTAT 2016. On the other hand, data collection and analysis were conducted with great prudence to minimize the amount of missing data. However, the incidence of missing data is inevitable and was handled by the multiple imputation of missing values technique using Markov Chain Monte Carlo (MCMC) [41].

The raw water quality parameters were standardized to a mean of 0 and variance of 1 using Z-scale transformation to examine the normality of the distribution of data sets and to ensure that the different variables were equally weighted in the statistical analyses [36]. The data were further checked for normality using Kaiser–Meyer–Olkin (KMO) and Bartlett’s sphericity tests to determine if our measured variables may be factorized efficiently. KMO is the degree of sampling adequacy, which shows the percentage of variance that is likely attributable to the underlying factors. Generally, the KMO index ought to be greater than 0.5 for satisfactory factor analysis. When the KMO index is close to 1, the PCA of the variables is suitable; however, when it is close to 0, the PCA is not relevant. In this study, the KMO had a value of 0.68. Bartlett’s test of sphericity shows whether the correlation matrix is an identity with variables that are unrelated. The significance level, which is 0 in this study (less than 0.05), indicates that there are significant relationships among the variables.

3.2.1. Principal Component (PCs)/Factor Analysis (FA)

PCA reduces the dimensionality of the data set by explaining the correlations amongst a large number of variables in terms of a smaller number of underlying factors without losing much information [42,43]. The original variables of PCs produce loadings that have correlation coefficients with PCs. The PCs’ formula was taken from [33,36]:

$$Y_{mn} = Z_{m1}X_{1n} + Z_{m2}X_{2n} + Z_{m3}X_{3n} \dots + Z_{mi}X_{in} \tag{3}$$

where z is the component loading, y is the component score, x is the measured value of a variable, m is the component number, n is the sample number, and m is the total number of variables.

Meanwhile, FA attempts to extract a lower-dimensional linear structure from the data set and extracts the new group of variables known as varifactors (VFs) via rotation along the PCA axis. In FA, the basic concept is borrowed from [33,36]:

$$Y_{mn} = Z_{p1}P_{1m} + Z_{p2}P_{2m} + Z_{p3}P_{3m} + \dots + Z_{pr}P_{rm} + e_{pm} \tag{4}$$

where y is the measured value of the variable, z refers to the factor loading, p is the factor score, m is the sample number, n is the variable number, r is the total number of factors, and e is the residual term accounting for errors or other sources of variation.

In this study, PCA was employed for qualitative determination of pollution sources.

3.2.2. Discriminant Analysis

DA was used for discriminating between and among groups by applying discriminating variables. These variables measure characteristics regarding which the groups are expected to differ [44]. DA applies a linear equation of a regression analysis on raw data with prior knowledge of membership of objects to particular clusters and provides statistical classification of samples, expressed in the following equation [43,45]:

$$f(G_i) = K_i + \sum_{i=1}^n (W_{ij} * P_{ij}) \tag{5}$$

where K_i is a constant specific to each particular group, i is the number of groups (G), n is the number of parameters used in group classification, and W_{ij} is the weight coefficient designated by DA for the specific parameter (P_{ij}).

Independent variables are entered into DA either all together or stepwise, using both backward and forward approaches. In the first approach of variable entry, the discriminant function is calculated by engaging all the independent variables at once. This approach is used when there are a limited number of independent variables in the interest of discovering how well certain variables perform as discriminants in the absence of others. The stepwise method, on the other hand, involves entering the independent variables into the discriminant function (DF) one at a time. This stepwise input is based on the fact that variables with relative importance to the cluster variables with greater discriminant weights were entered first [46].

In this study, standard, forward, and backward stepwise approaches of DA were applied to each matrix of the primary data. In the forward stepwise mode, discriminant function analysis (DFA) variables were added stepwise until no significant change occurred, while in the backward stepwise mode, variables were removed starting from least significant until a significant change occurred. For this purpose, two groups obtained from CA were selected for spatial evaluations [35].

3.2.3. Pollution Index (PI)

Pollution index (PI) is a simple technique to examine surface water quality and was applied by Tiwan EPA. The parameters such as DO, BOD, SS, and $\text{NH}_3\text{-N}$ employed to determine PI were classified into four index scores (Table 3) and computed using the equation formulated by [47,48]. In particular, PI refers to the arithmetic mean of the index values with respect to the water quality.

$$PI = \frac{1}{4} \sum_{K=1}^4 S_i \quad (6)$$

Table 3. Classification system for pollution index.

Item	Rank			
	Non-Polluted (Good)	Slightly Polluted (LP)	Moderately Polluted (MP)	Highly Polluted (HP)
DO (mg/L)	>6.5	4.6–6.5	2.0–4.5	<2.0
BOD ₅ (mg/L)	<3	3.0–4.9	5.0–15.0	>15
SS (mg/L)	<20	20–49	50–100	>100
$\text{NH}_3\text{-N}$ (mg/L)	<0.5	0.5–0.9	1.0–3.0	>3.0
Index score	1	3	6	10

PI classifies water quality into four categories: (0–2) for good or non-polluted, (2–3) for slightly polluted, (3–6) for moderately polluted, and (>6) for highly polluted. Anthropogenic activities have been associated with water quality degradation [47,49].

3.3. Cluster Analysis

Hierarchical agglomerative CA was carried out on the normalized data set using Ward's approach, where Euclidean distances were used as the degree of similarity among samples, and a distance was represented by the distinction among analytical values. In hierarchical clustering, sequentially higher clusters formed [23,45,50–52]. In cluster analysis, cases are classified into classes based on similarities between two samples, which are usually

given by the Euclidean distance between analytical values of the two samples. The squared Euclidean distance can be calculated by [53]:

$$\text{Distance} (Q_i, Q_j) = \sum_{j=1}^n (X_{1i} - X_{2j})^2 \tag{7}$$

where Q_i is the i th object, and X_{ij} is the value of the j th variable of the i th object.

The dendrogram provides a visual summary of the clustering process to classify a sample of entities into a smaller number of mutually exclusive groups on the basis of multivariate similarities among entities [33].

Therefore, CA, DA, PCA, and pollution index were applied in this study to identify the underlying interrelationship among the parameters and monitoring stations. CA was applied based on prior knowledge of monitoring stations and the results of DA and pollution index to accurately cluster monitoring stations. PCA was employed to qualitatively identify pollution sources and the type of contaminants contributing to pollution.

4. Results and Discussion

4.1. Correlation Matrix Evaluation and Seasonal Variation

Correlation coefficients are established to portray a correlation among variables and measure statistical significance between pairs of water quality variables [54,55]. Correlation analysis measures the proximity between the identified dependent and independent variables. Correlation coefficients that are close to -1 or $+1$ demonstrates a strong correlation between x and y , which have a linear correlation. The correlation between the parameters is referred to as strong from $(+0.8$ to $1.0)$ or $(-0.8$ to $-1.0)$, moderate from $(+0.5$ to $0.8)$ or $(-0.5$ to $-0.8)$ and weak from $(+0.0$ to $0.5)$ or $(-0.0$ to $-0.5)$ [56]. In cases where the correlation coefficient between variables is zero, there could be no correlation with a degree of $p < 0.05$ between the two variables [57]. In this study, a correlation matrix was constructed for each dry and wet season using the physicochemical parameters. Pearson’s correlation coefficient (r) is determined using correlation matrix to identify the highly correlated and interrelated water quality parameters. To test the significance of the pair of parameters, the p -value is determined.

In the wet season, strong positive correlations were observed between TDS values and EC, temperature, TP, TN, and Na^+ values ($r = 0.992, r = 0.874, r = 0.850, r = 0.836; p < 0.05$), and strong negative correlations between TDS and DO with -0.825 at $p < 0.05$. Moderate positive correlations were found between TDS and $PO_4-P, BOD, COD,$ and K^+ values ($r = 0.797, r = 0.698, r = 0.695, r = 0.523; p < 0.05$), and low positive correlation between TDS and pH with $r = 0.26; p < 0.05$ (Table 4). Strong negative correlations were found between DO and EC, TDS, TP, and TN ($r = -0.825, r = -0.850, r = -0.851, r = -0.806; p < 0.05$), and moderate negative correlations were observed between DO and temperature, BOD, COD, and Na^+ values ($r = -0.526, r = -0.544, r = -0.692, r = -0.599; p < 0.05$).

Table 4. Correlation matrix Pearson (r) and alpha (p) values for the wet season.

Parameters	TDS	EC	NH ₃ -N	NO ₃ -N	PO ₄ -P	DO	BOD	COD	TN	TP	Temp	Mg ²⁺	Ca ²⁺	Na ⁺	K ⁺
TDS	1														
EC	0.992	1													
NH ₃ -N	0.446	0.379	1												
NO ₃ -N	0.183	0.172	-0.030	1											
PO ₄ -P	0.797	0.824	0.416	-0.116	1										
DO	-0.825	-0.850	-0.216	-0.275	-0.793	1									
BOD	0.698	0.719	0.106	-0.173	0.712	-0.526	1								
COD	0.695	0.714	0.204	-0.111	0.730	-0.544	0.965	1							
TN	0.874	0.855	0.481	0.059	0.825	-0.851	0.587	0.602	1						
TP	0.850	0.871	0.249	0.255	0.602	-0.806	0.485	0.482	0.736	1					
Temperature	0.860	0.864	0.331	0.410	0.594	-0.692	0.454	0.447	0.669	0.82	1				
Mg ²⁺	-0.005	0.029	-0.317	0.070	-0.013	-0.085	0.224	0.159	0.046	0.09	-0.020	1			
Ca ²⁺	0.375	0.397	-0.085	-0.080	0.350	-0.394	0.523	0.528	0.429	0.24	0.137	0.401	1		
Na ⁺	0.836	0.853	0.314	0.268	0.709	-0.599	0.619	0.632	0.572	0.68	0.849	-0.062	0.19	1	
K ⁺	0.523	0.431	0.531	0.155	0.290	-0.429	0.149	0.190	0.700	0.34	0.320	-0.080	0.20	0.19	1

Values in bold are different from 0 with a significance level alpha = 0.05.

Strong positive correlations were observed between temperature and the values of EC, TDS, Na⁺ and TP ($r = 0.86, r = 0.864, r = 0.849, r = 0.821; p < 0.05$), and a moderate positive correlation was observed between temperature and the values of TN and PO₄-P ($r = 0.525, r = 0.669, r = 0.594; p < 0.05$). There was also a moderate negative correlation between temperature and DO, with $r = -0.692$ at $p < 0.005$. There was a weak correlation between temperature and the values of COD and BOD ($r = 0.447, r = 0.454; p < 0.05$).

NH₃-N had a moderate positive correlation with K⁺, with $r = 0.531$ at $p < 0.005$, and weak positive correlations with TN and temperature ($r = 0.331, r = 0.481$ at $p < 0.05$). NO₂-N correlated moderately positively with BOD and COD ($r = 0.721, r = 0.664$ at $p < 0.05$) and weakly positively with PO₄-P and Ca²⁺ ($r = 0.449, r = 0.404$ at $p < 0.05$).

A strong positive correlation was found between PO₄-P and TN, with $r = 0.825$ at $p < 0.005$, moderate positive correlations were found between PO₄-P and COD, BOD, TP, and temperature ($r = 0.712, r = 0.709, r = 0.730, r = 0.602, r = 0.594; p < 0.05$), and a moderate negative correlation was observed between PO₄-P and DO values ($r = -0.793; p < 0.05$). No statistically significant difference was found between pH and NO₃-N and the rest of the parameters of LHW ($p > 0.05$).

In the dry season, strong positive correlations were observed between TDS values and EC, TP, Na⁺, PO₄-P, and temperature values ($r = 0.999, r = 0.814, r = 0.899, r = 0.839, r = 0.933; p < 0.05$), moderate positive correlations were observed between TDS and BOD, COD, K⁺, and TN values ($r = 0.686, r = 0.561, r = 0.645, r = 0.534; p < 0.05$), and a strong negative correlation was found between TDS values and DO ($r = -0.819$ at $p < 0.05$).

Strong negative correlations were observed between the values of DO and TDS, EC, and Na⁺ ($r = -0.819, r = 0.817, r = -0.826; p < 0.05$), moderate negative correlations were observed between DO and TN, TP, BOD, K⁺, and temperature values ($r = -0.577, r = -0.568, r = -0.687, r = -0.639, r = -0.729; p < 0.05$), and a moderate negative correlation was observed between DO and NO₃-N, with $r = -0.464$ at $p < 0.005$.

Strong positive correlations were found between temperature and EC and TDS ($r = 0.839, r = 0.842; p < 0.05$), and moderate positive correlations were found for temperature with TP and PO₄-P ($r = 0.730, r = 0.532; p < 0.05$). There was also a moderate negative correlation observed between temperature and DO, with $r = -0.729$ at $p < 0.005$. NH₃-N had a moderate positive correlation with COD, TP, temperature, and Na⁺ ($r = 0.476, r = 0.484, r = 0.550, r = 0.343; p < 0.005$).

A strong positive correlation was found between PO₄-P and TP, with $r = 0.921$ at $p < 0.005$, moderate positive correlations were found for PO₄-P with BOD, COD, TP, Na⁺, and temperature ($r = 0.749, r = 0.647, r = 0.680, r = 0.76; p < 0.05$), and a moderate negative correlation was found between PO₄-P and DO values $r = -0.626; p < 0.05$) (Table 5).

Table 5. Correlation matrix Pearson (r) and alpha (p) values for dry season.

Parameters	TDS	EC	NH ₃ -N	NO ₃ -N	NPO ₄ -P	DO	BOD	COD	TN	TP	Tem	Mg ²⁺	Ca ²⁺	Na ⁺	K ⁺
TDS	1														
EC	0.999	1													
NH ₃ -N	0.433	0.419	1												
NO ₃ -N	0.208	0.212	-0.10	1											
PO ₄ -P	0.814	0.815	0.383	-0.04	1										
DO	-0.82	-0.82	-0.31	-0.46	-0.63	1									
BOD	0.686	0.686	0.450	-0.12	0.749	-0.58	1								
COD	0.561	0.564	0.476	-0.19	0.647	-0.41	0.871	1							
TN	0.645	0.642	0.410	0.184	0.680	-0.57	0.520	0.619	1						
TP	0.899	0.899	0.484	-0.03	0.921	-0.69	0.804	0.683	0.535	1					
Temperature	0.839	0.842	0.343	0.237	0.532	-0.73	0.436	0.344	0.291	0.730	1				
Mg ²⁺	-0.27	-0.27	-0.25	-0.13	-0.04	0.305	-0.13	-0.20	-0.16	-0.13	-0.42	1			
Ca ²⁺	0.385	0.392	-0.19	0.398	0.091	-0.33	0.235	0.324	0.208	0.17	0.455	-0.33	1		
Na ⁺	0.933	0.931	0.550	0.173	0.760	-0.83	0.813	0.694	0.601	0.881	0.788	-0.38	0.37	1	
K ⁺	0.534	0.531	0.182	0.419	0.261	-0.64	0.237	0.240	0.701	0.197	0.335	-0.39	0.42	0.53	1

Values in bold are different from 0 with a significance level alpha = 0.05.

The pH of rivers was 7.4 (7.1 to 7.6) in the dry season and 8.2 (7.5 to 8.7) in the wet season, and the pH of lake was 8.2 (7.3 to 8.9) in the dry season and 8.5 (7.5 to 9) in the wet season. The pH of point sources was 8.3 (7.1 to 9) in the dry season and 8.3 (8.1 to 8.7) in the wet season. The recommended pH as per the standard for drinking, irrigation, and aquatic life is 6.5–8.6, and the pH of LHW was within the accepted limit (Table 6). The EC (TDS) of rivers was 148mg/L (297 μ S/cm) in dry seasons and 89 mg/L (179 μ S/cm) in wet seasons, and EC (TDS) of lakes was 453 mg/L (877 μ S/cm) in dry season and 421 mg/L (829 μ S/cm) in wet seasons. The EC (TDS) of point sources was 1655 mg/L (3509 μ S/cm) in dry season and 1395 mg/L (2809 μ S/cm) in wet seasons. This shows that the EC (TDS) of rivers, lakes, and point sources increases significantly with increasing temperature (Table 6). The NO₃–N concentration of rivers was 0.5 mg/L, NO₃–N concentration of Lake Hawassa was 1.4 mg/L, and that of point sources was 1.5 mg/L for the dry season. In the wet season, the NO₃–N concentration was 0.7, 1.9, and 1.9 for rivers, Lake Hawassa, and point sources, respectively. The value of NO₃–N increases in the rainy season due to the contribution of agricultural runoff and use of fertilizers. The PO₄–P concentration of rivers was 6.5 mg/L, PO₄–P of Lake Hawassa was 3.3 mg/L, and that of point sources was 43.8 mg/L in dry season. In the wet season, the PO₄–P concentration was 7.4, 2.9, and 25.7 for rivers, Lake Hawassa, and point sources, respectively (Table 6). Similarly, Gebre-Mariam [58] reported that Ethiopian Rift Valley lakes generally have lower EC values in the rainy season than in the dry season, due to dilution by rain coupled with minimal evaporation rates during the rainy season.

Table 6. Descriptive statistics (mean and standard deviation) of the physicochemical characteristics of LHW collected during dry season.

Codes	SS	TDS	EC	pH	NH ₃ –N	NO ₃ –N	PO ₄ –P	DO	BOD	COD	TN	TP	Mg ²⁺	Ca ²⁺	Na ⁺	K ⁺	Temperature
MS1	17.3 (1.6)	89 (4)	178 (7)	7.1 (0.2)	0.04 (0.01)	0.6 (0.01)	3.6 (2)	4.1 (0.7)	13.8 (1.5)	88.3 (26.9)	5.8 (1.5)	0.001 (0)	7.2 (2.1)	20 (7.4)	32.5 (1.5)	6.7 (0.6)	19.2 (0.8)
MS2	27.3 (5.8)	100 (15)	200 (30)	7.6 (0.5)	0.16 (0.07)	0.4 (0.04)	10.2 (6.7)	3.5 (1)	23.7 (7.2)	107.5 (32.5)	7.2 (1.8)	0.5 (0.5)	54.0 (16.4)	9 (8.4)	26.2 (3)	8.1 (0.6)	17.7 (1)
MS3	54.5 (3)	87 (6)	175 (12)	7.7 (0.3)	0.10 (0.01)	0.6 (0.1)	5.9 (0.1)	4.4 (0.4)	69.0 (20.5)	313.8 (93.3)	7.5 (2.5)	0.001 (0)	153.4 (50.2)	4.6 (4.2)	25.8 (3)	6.8 (0.6)	18.1 (0.7)
MS4	58.0 (10.4)	1575 (59)	3825 (108)	7.1 (0.4)	7.60 (1.49)	2.8 (0.5)	18.7 (2.9)	1.5 (0.1)	63.3 (10.8)	263.7 (84.9)	23.8 (5.9)	15 (3)	11.4 (0.1)	50.4 (5.6)	501.1 (83)	19.8 (0.3)	33.8 (0.3)
MS5	27.7 (1.5)	2349 (193)	4698 (385)	9.5 (0.6)	12.35 (5.15)	0.6 (0.05)	118.3 (40)	0.9 (0.1)	190 (1.3)	600 (241)	41.3 (16.7)	6.5 (1.6)	2.9 (1.8)	15.0 (10.6)	1078.1 (178)	19.3 (0.8)	29 (0.6)
MS6	23.3 (0.1)	317 (63)	635 (126)	7.6 (0.1)	0.06 (0.03)	1 (0.2)	6.3 (0.6)	4 (0.5)	5.3 (1.8)	26.3 (5.8)	11.3 (3.7)	0.001 (0)	5.7 (1)	18.7 (0.7)	111.8 (24)	9.4 (0.9)	24.5 (0.4)
MS7	10.6 (1.4)	388 (7)	765 (25)	8.8 (0.003)	0.37 (0.08)	0.9 (0.02)	2.5 (0.5)	4.5 (0.5)	5.9 (0.3)	116 (88)	6.8 (1.2)	0.8 (0.2)	5.1 (0.7)	16.2 (1.1)	221.9 (15.9)	20.1 (0.3)	22.8 (1)
MS8	13.6 (0.1)	518 (26)	851 (32)	8.9 (0.02)	11.75 (3.9)	1 (0.02)	4.3 (0.1)	5.3 (0.3)	9.5 (1.5)	135 (5)	4.5 (1.5)	0.4 (0.1)	3.9 (0.1)	24.2 (1.8)	255.0 (30.1)	22.2 (0.8)	22.8 (1)
MS9	9.0 (1.2)	392 (3)	748 (24)	8.7 (0.1)	0.38 (0.11)	2 (1)	3.0 (0.1)	4 (0)	9.9 (0)	45 (0)	3 (0)	0.001 (0)	12.8 (1)	22.2 (1.9)	191.9 (5.4)	20.0 (0.3)	22.7 (1.8)
MS10	10.8 (0.4)	473 (8)	955 (5)	8.5 (0.04)	0.12 (0.04)	0.6 (0.1)	2.5 (0.7)	4.3 (0.3)	71.8 (22.8)	326 (104)	1.1 (0.1)	0.1 (0.1)	5.2 (0.5)	18.4 (1.4)	224.2 (7.9)	20.0 (0.1)	21.3 (1.4)
MS11	13.5 (0.2)	463 (3)	880 (20)	8.6 (0.04)	3.71 (1.23)	3.1 (1.8)	2.0 (0.3)	3.3 (0.1)	9.0 (1)	96 (20.5)	4.5 (1.5)	0.001 (0)	5.4 (0.1)	20.9 (0.1)	205.1 (4.9)	20.7 (0.3)	23.1 (1.3)
MS12	10.3 (1.8)	460 (18)	921 (35)	8.6 (0.2)	1.34 (0.56)	1 (0)	2.3 (0.4)	4.5 (0)	10.1 (0.4)	46 (1.8)	4.0 (1.8)	0.001 (0)	10.1 (1.9)	26 (1.9)	225.0 (10.9)	23.4 (2.7)	22.6 (1.1)
MS13	12.5 (3)	411 (9)	807 (33)	8.5 (0.2)	0.15 (0.03)	3.1 (2.1)	3.1 (0.6)	4.0 (0.5)	47.3 (8.8)	255 (55)	6.9 (2.1)	0.5 (0.1)	13.5 (2.5)	40.9 (7.1)	280.8 (29.2)	19.0 (0.7)	23.2 (1.3)
MS14	9.3 (1.3)	358 (82)	714 (166)	7.3 (0.1)	1.19 (0.38)	1.3 (0.1)	3.8 (0.8)	3.5 (0.5)	20.2 (4)	134 (23.8)	3.8 (1.2)	0.001 (0)	6.3 (0.3)	16.9 (0.2)	150.8 (37.3)	16.7 (3.3)	20.8 (0.9)
MS15	24.2 (0.9)	1632 (39)	3266 (78)	8.3 (0.005)	24.97 (7.06)	1.6 (0.8)	36.7 (6.8)	1.5 (0.03)	63.5 (9.1)	290 (40)	49.5 (15.5)	5.6 (1.9)	13.7 (2.1)	33.7 (2.9)	420.2 (41.3)	44.7 (3.3)	23.9 (0.8)
MS16	16.6 (0.8)	483 (8)	935 (45)	8.6 (0.1)	0.96 (0.78)	1.0 (0.1)	3 (0.3)	4.2 (0.1)	22.6 (3.1)	75.5 (10.5)	6.3 (0.8)	3.8 (1.2)	3.2 (0.3)	8.8 (1.3)	197.2 (13.7)	17.8 (2.3)	21.5 (0.3)
MS17	14.3 (0.2)	479 (1)	935 (25)	8.6 (0.01)	3.17 (0.04)	1 (0.01)	2.7 (0.1)	4.2 (0.1)	48 (3)	160 (10)	5.3 (0.2)	0.001 (0)	14.1 (1.7)	33.8 (1.3)	159.0 (16.3)	18.0 (2.3)	22.0 (0.5)
MS18	96.4 (3.0)	561 (34)	1133 (48)	8.7 (0.1)	0.86 (0.21)	1 (0.01)	7.8 (0.7)	4.3 (0.3)	55.5 (1.5)	185 (5)	12.3 (2.7)	0.8 (0.2)	16 (1.7)	34 (2)	243.7 (11.2)	19.1 (1.2)	23 (0.3)
MS19	7.2 (0.8)	1065 (215)	2246 (469)	8.4 (0.1)	0.05 (0.02)	0.8 (0.04)	9.8 (1.4)	4.2 (0.1)	126 (3)	420 (10)	12.8 (2.3)	2 (0.3)	18.3 (3.8)	53.7 (4.9)	301.4 (25.5)	20.9 (0.05)	21.2 (0.4)

All units in mg/L except pH (Dimensionless), Temperature (°C), EC (μ S/cm) and Turbidity (NTU).

The TN (TP) of rivers was 8 (0.12) mg/L in dry seasons and 5(0.26) mg/L in wet season, and TN (TP) of lakes was 5.3 (0.2) mg/L in dry season and 5.2 (0.6) mg/L in wet season. Hence, there is an obvious increase of TN in rivers and Lake Hawassa when temperature increases due to lower dilution and greater agricultural contribution from the upper stream by irrigation, whereas TP in rivers and Lake Hawassa increases in wet seasons due to greater agricultural, rural, and urban runoff. The TN (TP) from point sources was 31.8 (7.2) mg/L in dry season and 13.9 (5.4) mg/L in wet season. This shows that TN (TP) of point sources increases significantly with increasing temperature due to lower dilution. The $\text{NH}_3\text{-N}$ of rivers was 0.2 mg/L, $\text{NH}_3\text{-N}$ of Lake Hawassa was 0.83 mg/L, and that of point sources was 4.72 mg/L in dry season. In the wet season, the $\text{NH}_3\text{-N}$ values were 0.03, 0.71, and 3.6 for rivers, Lake Hawassa, and point sources, respectively. The decreases in $\text{NH}_3\text{-N}$ level in the rainy season might be due to dilution effect (Table 6).

The positive correlation between temperature and TN, TP, EC, TDS, $\text{NH}_3\text{-N}$, and $\text{PO}_4\text{-P}$ indicates the increase in the concentration of nutrients as the temperature increases (dry period). It also confirms the major contributors of nutrients were the point sources that are releasing a relatively higher amount of pollutants than the agricultural and other sources, as this value lowers during the wet season due to dilution effect. However, the increase in nutrient ($\text{NO}_3\text{-N}$) concentration in rivers and Lake Hawassa in the wet season might be due to the increased contribution of agricultural runoff and use of fertilizers.

Sodium, calcium, magnesium, and potassium concentrations of the rivers were 49.1, 13.06, 55.1, and 7.74 mg/L in dry season and 28.9, 32.7, 10.1, and 5.7 mg/L in wet seasons. Sodium, calcium, magnesium and potassium concentrations of the lake were 214, 23.8, 8.7, and 19.7 mg/L in dry season and 178.9, 25.1, 7.3, and 17.2 mg/L in wet season. The sodium, calcium, magnesium, and potassium concentrations of the point sources were 575.2, 38.2, 11.5, and 26.2 mg/L, respectively, in the dry season and 375.2, 38.2, 9.5, and 50.1 mg/L, respectively in the wet season (Table 6). There was an observed decrease in ions when the temperature decreased in the study area. This can be ascribed to the discharge of industrial and domestic effluents, which contribute large amounts of alkaline ions to the river system, as the conductivity depends mainly on the ion concentration in surface water [52]. The natural range of sodium ions in water and soil is so low that their existence can show river pollution caused by human activities. Calcium is added to water from soil, industrial wastes, and natural resources. Magnesium is an essential nutrient required for numerous biochemical and physiological functions [59].

The TDS of water generally increases with the level of dissolved pollutants (such as nitrate, ammonium, and phosphate). Conductivity of ions in water depends on water temperature, and ions move faster when water is warm. Hence, conductivity apparently increases when water has a higher temperature [60]. In addition, Taylor et al. [61] pointed out a strong relationship between these variables or ions, such as nitrate, ammonium, and phosphate, and stated that high concentrations of EC indicate high concentrations of soluble salts. There are strong correlations between EC/TDS, as evidenced by an increase in conductivity as the concentration of all dissolved constituents increases [62] Table 6.

The BOD (COD) of rivers was 19.7 (96.5) mg/L in dry seasons and 6.9 (89.4) mg/L in the wet season, and the BOD and COD of lakes was 28.1 (133.3) mg/L in dry season and was 19.1 (112.9) mg/L in wet season. The BOD and COD concentrations for point sources were 116.2 (398.6) mg/L in dry season and 111.6 (353.7) mg/L in wet season (Table 6). The DO of rivers was 3.5 mg/L in dry season and 6 mg/L in wet season, and the DO of lakes was 4.2 mg/L in dry season and 4.4 mg/L in wet season. The DO of point sources was 2 mg/L in dry season and 2.3 mg/L in the wet season (Table 6).

The DO of the rivers in the dry seasons and Lake Hawassa were well below the standard value. This indicates that the discharge of industrial and domestic effluents has resulted in serious organic pollution of these rivers, as the decrease of DO was mainly caused by the decomposition of organic compounds. Moreover, an extremely low DO content usually indicates the degradation of an aquatic system [63].

The DO showed a negative correlation with most parameters in both dry and rainy seasons, revealing the value of DO decreases with the increase in other water quality parameters. This could explain the temporal variations, as more oxygen was available for reaction with the pollutants, especially metals and organic pollutants, during dry seasons. Additionally, the characteristics of temporal variation in water quality of LHW were affected by DO. DO was strongly correlated with organic matters, nutrients, and metals, and thus seasonal variation should be considered when DO is used as an indicator to evaluate surface water quality. Low dissolved oxygen (DO) is primarily the result of excessive algal growth caused by nutrients. As the algae die and decompose, this process consumes dissolved oxygen. This may result in insufficient dissolved oxygen for fish and other aquatic life. Temperature was significantly correlated with water quality parameters such as EC, TDS, TP, PO₄-P, and DO in both seasons. Temperature had significant negative correlation with DO in the dry and wet seasons, indicating that when water temperature increases, the metabolic rate of microorganisms also increases, and the amount of DO in the water decreases. This might be because faster biodegradation of organic matter during dry seasons can effectively improve water quality. The solubility of oxygen was inversely related to temperature, as the water becomes warmer and more easily saturated with oxygen, hence holds less DO during the dry season. Singh et al. [32] observed the inverse relationship between temperature and DO in natural processes, as water can hold less DO with increasing temperature.

4.2. Pollution Index (PI)

The mean pollution index of the rivers in the lake watershed was 4.5 in dry and 3.3 in wet season, indicating a moderately polluted condition of rivers. Lake Hawassa PI was 5 in both dry and rainy season, indicating that the quality of the lake was moderately polluted. Anthropogenic activities were causing deterioration of the water quality of the rivers and Lake Hawassa, and the overall status of the water quality is moderately polluted. The PI for the point sources was measured for comparison purposes, and it was found to be highly polluted, having a PI index of 6.8 and 7.3 for the wet and dry seasons, respectively (Table 7).

Table 7. Average concentrations of monitoring stations for rivers, Lake Hawassa, and point sources (PS) observed in both dry and wet seasons.

Parameters	Seasons	Rivers	Lake Hawassa	PS
DO (mg/L)	dry seasons	4.2	4.2	1.7
	wet seasons	6	4.3	2.1
BOD ₅ (mg/L)	dry seasons	19.7	28.1	116.2
	wet seasons	6.9	19.1	111.6
SS (mg/L)	dry seasons	30.6	19.7	29.3
	wet seasons	51.1	20.9	28.1
NH ₃ -N (mg/L)	dry seasons	0.2	0.8	1.2
	wet seasons	0.002	0.71	14.4
PI	dry seasons	4.5	5	7.3
	wet seasons	3.3	5	6.8
Rank	dry seasons	MP	MP	HP
	wet seasons	MP	MP	HP

4.3. Cluster Analysis

Spatial and Temporal Similarities

Cluster analysis was applied to find out if the monitoring stations had similar characteristics in terms of water quality parameters. It was implemented with the water quality data set to group comparable monitoring sites (spatial variability) spread over the watershed. Results from CA display high homogeneity within clusters and high heterogeneity between clusters [64]. Hierarchical agglomerative CA was carried out with the normalized data set employing Ward's method, using Euclidean distances as a measure of similarity. In this approach, the analysis of variance method is used to evaluate the distances between

clusters, attempting to reduce the sum of squares of all clusters that can be made at each step. In this method, the clusters are grouped sequentially, beginning with the most comparable pair of objects and establishing better clusters one after the other, demonstrated through a dendrogram [2,65].

The dendrogram presents a visual summary of the clustering processes and provides the map of the groups with a dramatic reduction in the dimensionality of the original records [2,5,32,43,44]. The CA grouped all 19 monitoring stations into two statistically significant clusters for the dry and wet seasons in LHW, and the dendrogram displays the grouping of stations for the wet and dry seasons, as demonstrated in Figure 2. Regarding the clustering for the dry and wet seasons, monitoring stations from most of the watershed upstream, from the eastern and western sides of the lake, and from the center of Lake Hawassa have been grouped in Cluster 1. Stations in these clusters typically consist of rivers and Lake Hawassa and are categorized as moderately polluted. The monitoring stations in these clusters are MS1-MS3, MS6-MS14, and MS16-MS18, which can be labeled as “moderate anthropogenic effect”. This cluster received pollution from point sources and non-point sources, consisting of animal waste and runoff. It is characterized by moderate anthropogenic impact and labelled as moderately polluted.

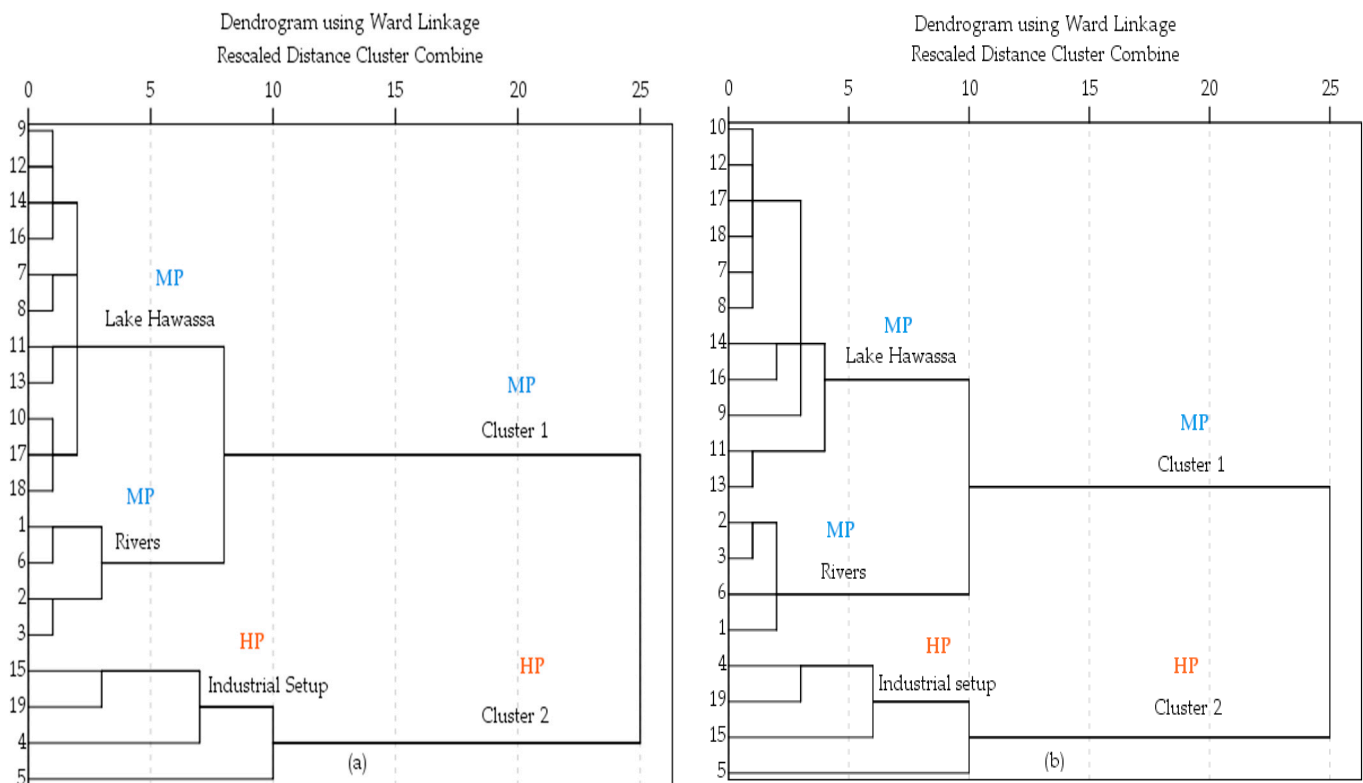


Figure 2. Dendrogram for LHW based on Ward’s method showing the clustering of 19 monitoring stations into two significant clusters for both dry (a) and wet (b) seasons.

The pollution sources for monitoring stations MS1-MS3 were mainly anthropogenic activities from non-point pollution sources such as agricultural and sewage pollution, whereas pollution sources for monitoring stations MS6 (Tikur-Wuha river) and Lake Hawassa (MS7–MS14, MS16–MS18) were mainly industrial pollution, dispersed point sources, agricultural pollution, urban runoff, and sewage pollution.

Owing to their relative sources, all stations in this cluster were rivers and lakes, suggesting that clustering is reasonable for both dry and wet seasons.

The spatial trend of water quality was generally driven by anthropogenic activities from point and non-point sources of pollution, especially anthropogenic activities with respect to pollutant loading and land use.

Cluster 2 includes four monitoring stations in the middle part of the LHW and groups monitoring stations in this cluster as MS4, MS5, MS15, and MS19. Four point sources, specifically BGI, Pepsi Factory, Referral Hospital, and Industrial Park monitoring stations, were assigned to this cluster. Consequently, this cluster is characterized by comparatively heavy pollution.

4.4. Discriminant Analysis

Discriminant analysis (DA) was used to evaluate the spatial variations in water quality and to distinguish the most critical parameters in relation to variations between clusters. Both the standard and stepwise modes were applied to the primary data by dividing them into wet and dry seasons, and the two spatial groups resulting from CA were used in DA. In this case, the WQ parameters were treated as independent variables, while the clusters were considered as dependent variables. The confusion matrixes (CM) showed that 100%, 100%, and 100% of the data points were correctly classified in the standard, forward stepwise, and backward stepwise modes for both dry and wet seasons, respectively (Table 8).

Table 8. Classification matrix for standard, forward stepwise, and backward stepwise DA of spatial variation in LHW for both dry and wet seasons, showing percentage of correct assignment for discriminating parameters.

Monitoring Stations	% Correct	Stations Assigned by DA	
		C1	C2
Standard DA mode for dry season			
C1	100	15	0
C2	100	0	4
Total	100	15	4
Standard DA mode for wet season			
C1	100	15	0
C2	100	0	4
Total	100	15	4
Forward stepwise DA mode for dry season			
C1	100	15	0
C2	100	0	4
Total	100	15	4
Forward stepwise DA mode for wet season			
C1	94	15	0
C2	85	0	4
Total	84.5	15	4
Backward stepwise DA mode for dry season			
C1	100	15	0
C1	100	0	4
Total	100	15	4
Backward stepwise DA mode for wet season			
C1	100	15	0
C2	75	0	4
Total	87.5	15	4

C1: Includes stations (MS1-MS3, MS6-MS14, and MS16-MS18). C2: Includes stations (MS4, MS5, MS15, and MS19).

The standard DA method builds DFs using eighteen parameters, while only three and seven parameters were the critical parameters useful to make distinction within the two pollution groups for both the forward stepwise modes and backward stepwise modes, respectively, for both dry and wet seasons. In forward stepwise mode, most of the parameters such as turbidity, TDS, pH, NH₃-N, NO₃-N, PO₄-P, DO, COD, NO₂-N, TN, TP, temperature, Mg²⁺, Ca²⁺, and K⁺ were insignificant variables leading to less variation, and they were deleted in the further process. However, in the forward stepwise DA

mode, the three significant variables that were useful to make distinctions within the two pollution groups with 100% correct assignment were EC, BOD, and Na⁺. The backward stepwise mode deleted the least significant and identified seven significant variables: EC, DO, COD, TN, TP, Na⁺ and K⁺. These seven parameters, which were 100% correctly assigned, were the critical parameters useful to make distinctions within the two pollution groups. This implies that the expected spatial variation in water quality can be explained sufficiently using variables EC, DO, COD, TN, TP, Na⁺, and K⁺. Wilks' lambda shows that the discriminant distribution is skewed towards high concentrations.

On the other hand, the standard DA functions was constructed using eighteen parameters, of which three and four parameters were used for forward stepwise mode and backward stepwise mode, respectively, for wet season. In forward stepwise mode, the pollutants that were found to be insignificant variables and had less variation in terms of their spatial distribution were deleted in the further process. However, in the backward stepwise DA mode, the three significant variables that were useful to make distinctions within the two pollution groups with 84.5% correct assignment were EC, Na⁺, and COD. The backward stepwise mode deleted the least significant and identified two significant variables: EC and Ca²⁺. These two parameters were the critical parameters useful to make distinctions within the two pollution groups with 87.5% correct assignment (Table 8). This implies the spatial water quality variation can be sufficiently explained by using variables EC, Na⁺, COD, and Ca²⁺, with Wilks' lambda value showing discriminatory distribution is skewed toward high concentration, as shown in Figure 3.

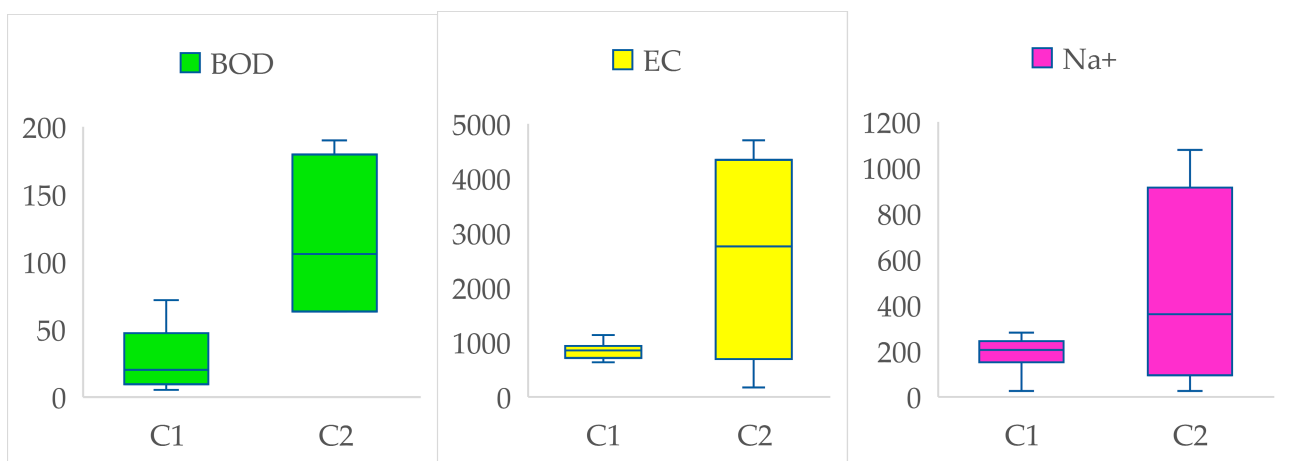


Figure 3. Box plot of the most discriminating parameters, BOD (mg/L), EC (µS/cm) and Na⁺ (mg/L) and Wilks' lambda showing skewedness of discriminatory distribution toward high concentration.

4.5. Pollution Source Identification of Monitored Variables

Principal Component Analysis

PCA was applied to the normalized data and was able to identify three principal components (PCs) using the Kaiser criterion [66] based on loading higher than 0.5. The scree plot graphs are used widely to identify the number of PCs to be retained to understand the underlying data structure [26]. Based on the scree plot and the eigenvalues >1 criterion, three factors were chosen as principal factors. The variables with eigenvalues lower than 1 were removed due to their low significance [67].

In this study, the scree plot (Figure 4) shows the sorted eigenvalues from large to small as a function of the number of PCs. This figure shows a pronounced change in slope after the third eigenvalue; three components were retained (Table 9). After the third PC (Figure 4a,b), beginning with the upward curve, the remaining components were circumvented. It was used to classify the number of PCs to be retained in order to figure out the underlying data

structure [25]. Consequently, a new set of data is obtained that may explain the variation of data set having fewer variables.

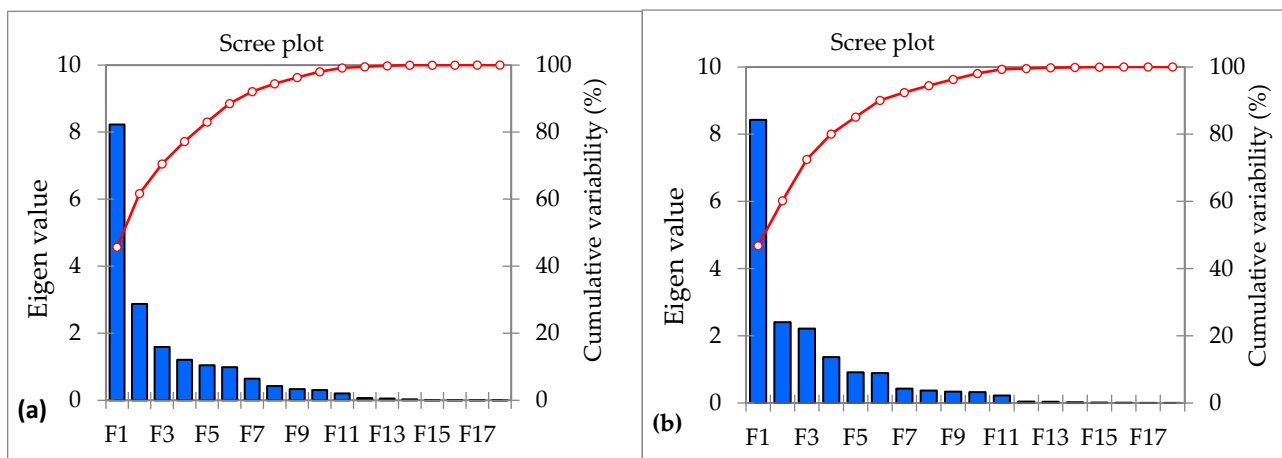


Figure 4. Factor loadings derived from scree plot and eigenvalue for LHW and three factors are retained for dry (a) and wet (b) seasons.

Table 9. Matrix of factor loadings calculated based on water quality parameters measured in the period from May to January in the Lake Hawassa Watershed and factor loadings of variables on the first three PCs extracted by using eigenvalue for both wet (a) and dry (b) seasons.

Parameters	F1 (a)	F2 (a)	F3 (a)	F1 (b)	F2 (b)	F3 (b)
Turbidity	0.282	−0.420 ^c	0.452^c	−0.032	− 0.781^a	−0.320 ^c
TDS	0.974^a	0.136	0.044	0.962^a	0.020	−0.084
EC	0.978^a	0.078	0.079	0.961^a	0.018	−0.098
pH	0.285	0.324 ^c	− 0.710^b	0.056	−0.178	0.775^a
NH ₃ –N	0.416 ^c	0.516^b	−0.313 ^c	0.521 ^b	−0.244	0.700^c
NO ₂ –N	0.428 ^c	−0.475 ^c	− 0.620^b	−0.088	− 0.531^b	−0.064
NO ₃ –N	0.131	0.398 ^c	0.507^b	0.195	0.599^b	−0.168
PO ₄ –P	0.871^a	−0.035	−0.174	0.830^a	−0.414 ^c	−0.200
DO	− 0.842^a	−0.055	−0.365 ^c	− 0.847^a	−0.246	0.186
BOD	0.784^a	−0.461 ^c	−0.297	0.796^a	−0.394 ^c	0.015
COD	0.793^a	−0.388 ^c	−0.302 ^c	0.721^b	−0.320 ^c	0.135
TN	0.898^a	0.064	0.101	0.724^b	−0.015	0.047
TP	0.812^a	0.139	0.436 ^c	0.897^a	−0.333 ^c	−0.105
Temp	0.825^a	0.290	0.194	0.783^a	0.246	−0.143
Mg ²⁺	0.077	− 0.654^b	0.389 ^c	−0.350 ^c	− 0.567^b	−0.380 ^c
Ca ²⁺	0.449 ^c	− 0.627^b	0.103	0.401 ^c	0.524^b	−0.246
Na ⁺	0.832^a	0.205	−0.116	0.973^a	0.001	0.076
K ⁺	0.477^c	0.335 ^c	0.035	0.572^b	0.522 ^b	0.106
Eigenvalue	8.4	2.4	2.2	8.2	2.9	1.6
Variability (%)	46.8	13.4	12.3	45.7	16	8.8
Cumulative %	46.8	60.2	72.5	45.7	61.7	70.5

^a strongly correlated factor loadings, ^b moderately correlated factor loadings, ^c weakly correlated factor loadings.

Moreover, scree plots are used to visually evaluate which components or factors elucidate the maximum variability in the data.

The PCA results, which include the loadings (participation of the original variable in the new one), are summarized in Table 9. The FA in LHW extracted three factors by retaining the PCs through varimax rotation that explained 72.5% of the total variance for the wet season. An eigenvalue offers a degree of the importance of the factor, and factors having the highest eigenvalues are the most significant. Eigenvalues of 1.0 or more are considered significant. Liu et al. [26] additionally categorized the factor loadings as ‘strong’,

'moderate', and 'weak', corresponding to absolute loading values of >0.75 , $0.75-0.50$, and $0.50-0.30$, respectively.

The first factor (F1), accounting for 46.8% of the total variance, showed strong positive loadings of TDS, EC, $\text{PO}_4\text{-P}$, BOD, COD, TP, TN, Na^+ , and temperature with factor loadings of 0.974, 0.978, 0.871, 0.811, 0.784, 0.793, 0.898, 0.812, 0.825, and 0.832, respectively; a weak positive loading of K^+ (0.477); and strong negative loading of DO (-0.842) (Table 9). High positive loadings of temperature and high negative loading of DO might suggest the impact of seasonal variation, and temperature is inversely related to DO. The strong and moderate positive loading of BOD and COD signify biodegradation of organic matter and are negatively affected by DO of water bodies. F1 stands clearly for pollution by BOD or COD, and nutrients and oxygen depletion is a consequence. When the temperature of water bodies decreases, the biodegradation of organic matter decreases, and the solubility of oxygen in the water increases. Similar reports of high concentrations of BOD and COD exist elsewhere [42,44,45]. Similarly, the strong negative DO loading indicates the utilization of DO under anaerobic conditions in rivers and lakes for the degradation of organic matter. F1 showed strongly positive loadings for both COD and BOD, while the loading for DO was strongly negative. This indicates a group of purely organic pollution indicator parameters from industrial effluents, domestic discharges, and livestock affecting water bodies [23,27,51].

High nutrient loadings of factors such as TN and TP represent pollution from point and non-point sources from industrial setup, agriculture areas, domestic sewage, and urban runoff. The high loading of metals demonstrates the influences of industrial effluents and agriculture activities. Phosphorus and nitrogen can originate from point sources such as sewage pollution, industrial facilities and livestock, as well as from non-point sources, mainly from agricultural activities, runoff from rural and urban areas, soil erosion, and livestock. These results are consistent with findings of other reports elsewhere [27,68]. Consequently, the component is more likely to be explained by the combination of domestic pollution and industrial factors. These factors are characteristic of the monitoring stations in the upper catchment (MS1 and MS2), in the middle section including point sources (MS5 and MS15), along Tikur-Wuha River (MS6), and on the eastern side of Lake Hawassa (MS7, MS9, MS12, MS13, and MS14), where domestic and industrial effluents and agricultural runoff are predominant.

The strongly positive loadings of Na^+ and weak positive loadings of K^+ are likely due to industrial effluents discharged into the river Tikur-Wuha and Lake Hawassa. Reports also indicate that the sources of Na^+ and K^+ might be domestic sources, fertilizers, and residential waste in addition to industrial effluents [69]. During field observation, it was found that the major industries are discharging their treated and untreated effluents directly into the Tikur-Wuha River and the lake during the rainy period when the flow rate is high, resulting in high dilution, but during the dry period, the dilution effect is lower and consequent pollution is higher.

On the other hand, the strong loadings of TN and TP in F1 suggest higher contribution from point sources in industry and non-point sources such as agricultural land use, urban drainage, and residential areas during the rainy season. In general, these factors are symbolic of a blended source of contamination, encompassing industrial discharges, urban runoff, and agricultural land use. The results are in agreement with those of other studies [5,24,67,69]. Hence, they can be considered as the contamination index for surface water [44,45].

The second factor (F2) explained 13.4% of the total variance. It had a moderately negative loading of Mg^{2+} and Ca^{2+} (-0.654 , -0.627) and a moderately positive loading of $\text{NH}_3\text{-N}$ (0.516). This factor's moderately negative loading of Mg^{2+} and Ca^{2+} is likely to originate from industrial wastewater discharged into the Tikur-Wuha River and Lake Hawassa, usually from carbonate minerals, which are naturally present in the soils of the Lake Hawassa watershed. This factor is more pronounced at monitoring stations affected by point sources, agricultural lands, and rural and urban runoff, such as MS3 in the upper

catchment, MS19 in the middle section (point source), and MS8, MS11, MS16, and MS18 monitoring stations on both eastern and western sides of Lake Hawassa.

A moderately positive loading of $\text{NH}_3\text{-N}$ (0.7) indicates biodegradation of organic matter. This variable is primarily from runoff, with high loading of solids and wastes from point sources of pollution from domestic and industrial areas. Furthermore, $\text{NH}_3\text{-N}$ is triggered by the decomposition of organic matter, indicating the discharge of domestic sewage to surface water. Studies elsewhere have showed comparable results [42,44,45,69,70].

The third factor (F3), explaining 12.3% of the total variance, had a moderately negative loading for pH (-0.710), suggesting the dominance of physical reactions by aquatic plants and natural weathering of the basin, possibly due to industrial impact from different sources [22]. It had weak positive loading of turbidity (0.452), moderate negative loading of $\text{NO}_2\text{-N}$ (-0.620), and moderate positive loading of $\text{NO}_3\text{-N}$ (0.507). $\text{NO}_3\text{-N}$ may additionally have derived from agricultural areas in the region, where inorganic nitrogen fertilizers are in common use and the role of domestic waste is strong, and hence, this component can be best explained by a “nutrient” factor representing influences from non-point sources such as agricultural runoff and the domestic pollution factor. The reports of Yilma et al. [35] in Ethiopia and Zhang et al. [27] elsewhere were comparable with this result. This factor is typical of the monitoring stations in the middle section including point sources and eastern and western sides of Lake Hawassa (MS4, MS10, and MS17), where domestic sewage, industrial effluents, and agricultural runoff are predominant.

The FA in LHW extracted three factors by retaining the PCs through varimax rotation that explained 70.5% of the total variance for the dry season. The first factor (F1), accounting for 45.7% of the total variance, showed strong positive loadings of TDS, EC, $\text{PO}_4\text{-P}$, BOD, DO, TP, Na^+ , and temperature, having factor loadings of 0.962, 0.961, 0.830, 0.796, 0.897, 0.783, and 0.973, respectively; moderate positive loadings of K^+ , COD, and TN (0.572, 0.721, 0.724); and strong negative loadings of DO (-0.847). Strong positive loadings of temperature and strong negative loadings of DO might suggest the impact of seasonal variations. The strong and moderate positive loading of BOD and COD signify biodegradation of organic matters and negatively affect DO of water bodies. F1 stands clearly for pollution by BOD or COD, and nutrients and oxygen depletion is a consequence. High temperature increases biodegradation and reduces solubility of oxygen in the water. This PC was correlated with COD and BOD_5 , indicating a group of purely organic pollution indicator parameters from uncontrolled domestic discharges caused by rapid urbanization and industrial effluents. Biodegradation of organic matter causes concentrations of BOD and dissolved oxygen in water [23,27,51].

A high loading of nutrients represents pollution from industrial setup and domestic wastewater. High loading of metals demonstrates the influences of industrial discharges. Phosphorus and nitrogen may originate from point sources such as sewage pollution, agricultural runoff in the upper stream due to irrigation, industrial facilities, and livestock. Consequently, this component is more likely to be explained by the combination of domestic pollution factors and industrial factors. Strongly positive loading of Na^+ and moderate positive loadings of K^+ are likely to originate from industrial effluents discharged directly into the Tikur-Wuha River and Lake Hawassa. These results are also supported by similar findings obtained elsewhere [27,69].

This factor is more pronounced at monitoring stations in the upper catchment (MS1 and MS3), monitoring stations in the middle section including point sources (MS4, MS5, MS15 and MS19), Tikur-Wuha River (MS6), and monitoring stations from both eastern and western sides of Lake Hawassa (MS9, MS10, MS14, MS16, and MS17), where domestic sewage, industrial effluents, and agricultural activities are predominant. The major industries discharge their treated and untreated effluents directly into Tikur-Wuha River and the lake during the dry period when the flow is low, which might lead to higher pollution. On the other hand, the strong loadings of TN and TP at F1 suggest a higher contribution of point sources from industrial facilities and agricultural runoff in the upper stream due to irrigation. Generally, these factors suggest a blended source of contamination encompass-

ing municipal and industrial point source and livestock. This result is also confirmed by other studies [5,23,33,67,69]. Hence, it can be considered to be the contamination index for surface water [44,45].

The second factor (F2) explained 16% of the total variance and had a strong negative loading of turbidity (-0.781), a moderate negative loading of $\text{NO}_2\text{-N}$ and Mg^{+2} (-0.567 , -0.531), and a moderate positive loading of $\text{NO}_3\text{-N}$ and Ca^{+2} (0.599 , 0.524). $\text{NO}_3\text{-N}$ could be mainly from point sources, and the role of domestic waste is also strong. Hence, this component can be explained by the “nutrient” factor, which represents influences from non-point sources such as the domestic pollution factor [24,27,32,35,66,69]. A moderately positive loading of K^+ and a moderately negative loading of Mg^{2+} in this factor likely originate from industrial discharges into the Tikur–Wuha River and Lake Hawassa. This PC is more influenced by industrial discharges, and monitoring stations from the LHW, where industry is predominant, are more pronounced. This factor is more pronounced in monitoring stations in the upper catchment (MS2) and the monitoring stations in the eastern and western sides of Lake Hawassa (MS11, MS12, MS13, and MS18), where domestic, industrial, and agricultural activities are predominant in the upper stream due to irrigation.

The third factor (F3), explaining 8.8% of the total variance, had a strong positive loading of pH (0.775), suggesting the dominance of physical reactions by aquatic plants and natural weathering of the basin, and attributed to industrial impact from different sources [22]. A moderate positive loading of $\text{NH}_3\text{-N}$ (0.7) indicates the biodegradation of organic matter causing concentrations of waterborne factors such as $\text{NH}_3\text{-N}$. This variable originated primarily from wastes from point sources of pollution from domestic and industrial areas. Furthermore, $\text{NH}_3\text{-N}$ is triggered by organic matter decomposition, indicating the discharge of domestic sewage to surface water. Reports elsewhere support the findings of this study [42,44,45,70]. This factor is more pronounced in monitoring stations on the eastern side of Lake Hawassa (MS7 and MS8), where domestic sewage, industrial effluents, and agricultural activities are prevalent.

The bi-plot of PCs on key parameters TDS, EC, $\text{PO}_4\text{-P}$, DO, BOD, COD, TN, TP, temperature, Na^+ , K^+ , Turbidity, $\text{NO}_2\text{-N}$, $\text{NO}_3\text{-N}$, Mg^{2+} , and Ca^{2+} that characterize monitoring stations from rivers in the upper and middle catchment, point sources in the middle catchment, and the eastern and western sides of Lake Hawassa are presented in Figure 5a,b for dry and wet seasons. In fact, the average values of EC, TDS, BOD, COD, Na^+ , K^+ , Mg^{2+} , Ca^{2+} , and $\text{NH}_3\text{-N}$ of point sources were exceedingly higher than that of rivers in the upper and middle catchment (MS1–MS3, and MS6) and Lake Hawassa (MS7–MS14, MS16 and MS18) in Table 6. In addition, $\text{NO}_3\text{-N}$, $\text{NO}_2\text{-N}$, TN, TP, and $\text{PO}_4\text{-P}$ were the main parameters characterizing the stated monitoring sites in both seasons. These stations predominantly include rural areas, urban and peri-urban areas, and industrial sites from which domestic sewage, urban runoff, and effluents are discharged into the lake. Furthermore, the influence of agricultural activities in the upper catchment and Tikur-Wuha River feeding the lake was evident. The results of this investigation were comparable to the findings of the studies conducted by Tibebe et al. [71] and Meshesha et al. [72] on Lake Ziway. In particular, higher EC and TDS values were recorded for similar monitoring stations in both seasons (Table 6). In an aquatic environment, EC is used to categorize the pollution status of surface waters, and an increase in conductivity indicates the presence of dissolved ions that can affect aquatic life and water quality [73].

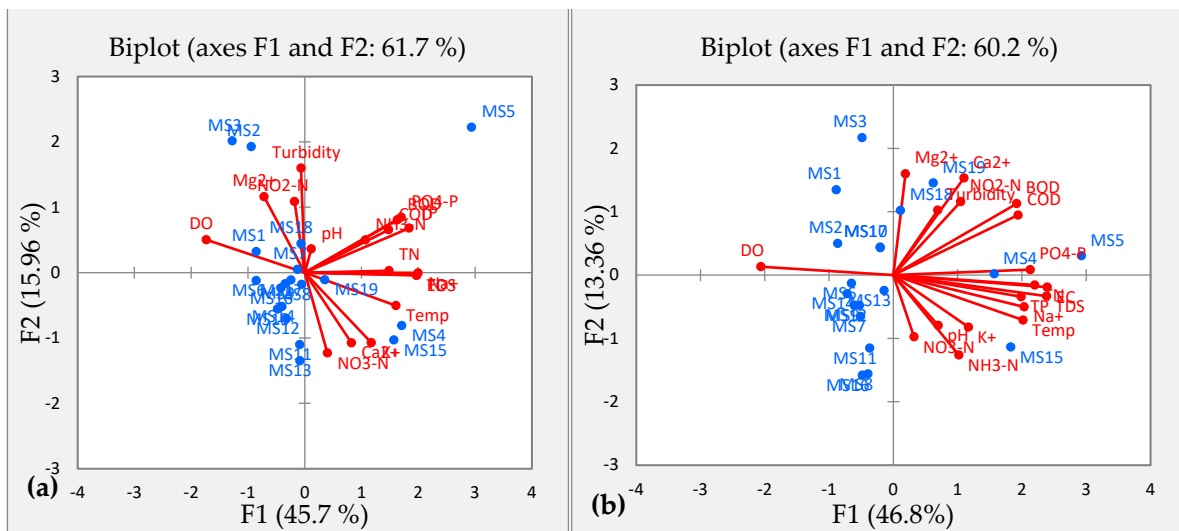


Figure 5. PCA biplots (a,b) suggest the projection of the monitoring sites (blue dots) and the variable loadings of the primary components (F1 and F2). The biplots additionally display the relationship between highly correlated variables and monitoring stations for dry (a) and wet (b) seasons. High and low values indicate strong positive and negative correlation, respectively, while values close to 0 imply weak correlation between F1 and F2 and the respective parameter.

4.6. Total Nitrogen to Total Phosphorus (TN:TP) Ratio

The TN:TP ratio in lakes and reservoirs is a key element, as it gives an idea of which of these nutrients is either in excess or limiting to growth, and it was used to estimate the nutrient limitation in the lake. According to Smith [74], blue-green algae (cyanobacteria) has a capacity to dominate in the lake section when the TN:TP ratio was less than 29, and it tends to be rare in the lake when $TN:TP > 29$. On the other hand, Fisher et al. [75] used a more conservative ratio of TN:TP. According to them, the ratio > 20 is designated as the phosphorus limitation and nitrogen limitation when the ratio is < 10 , while a TN:TP ratio of 10 to 16 demonstrates either phosphorus or nitrogen (or both) are limiting for growth. The estimated ratio for Lake Hawassa was 31, which is higher than 20 and 30, revealing cyanobacteria dominance in the lake section, which is rare. The $TN:TP > 20$ in Lake Hawassa indicated that phytoplankton growth in the lake might be phosphorous deficient.

5. Conclusions

Multivariate statistical techniques help researchers to scrutinize the relationships between parameters in a broader fashion by applying different approaches such as cluster analysis, correlation, factor analysis, discriminant analysis, and multiple regressions to determine the association between dependent and independent variables. They reduce the dimensionality of data so that the whole picture can be visualized more easily than looking at specific cases allows. Furthermore, multivariate techniques provide powerful significance testing compared to univariate techniques. Despite their various merits, the results of multivariate statistical modeling are not easy to interpret and require a large data set to get meaningful results due to the high standard errors. In particular, PCA/FA is likely to lose information if PCs or factors are not chosen judiciously.

This study was conducted to evaluate seasonal and spatial variations in water quality and to identify potential sources of pollution using multivariate statistical techniques for the Lake Hawassa Watershed. The results of this study show that the condition of Lake Hawassa Watershed was classified into moderately and highly polluted categories in both dry and wet seasons. In data-limited developing countries such as Ethiopia, it is especially clumsy to identify possible sources of pollution due to certain contaminants, as this requires frequently monitored water quality data, which are often not available. To address this serious problem, this study applied MVST. Multivariate statistics were used to

perform temporal and spatial assessment of surface water quality to reduce the number of monitoring stations and chemical parameters in LHW. In this study, we used Pearson correlation, PCA/FA, CA, and DA to evaluate spatial and temporal variance in surface water quality.

CA grouped the monitoring stations into two statistically significant clusters for the dry and wet seasons, labelled MP and HP, using PI. Accordingly, this resulted in a dendrogram with two clusters for the dry and wet seasons. The findings of the study revealed that rivers in the upstream and middle portion of the lake watershed and Lake Hawassa were moderately polluted (MP), while point sources (industries, hospitals, and hotels) in the middle of the LHW were found to be highly polluted (HP).

DA was used to identify the most critical parameters to investigate the spatial variations and extracted seven significant parameters: EC, DO, COD, TN, TP, Na⁺, and K⁺, with spatial variance to distinguish the pollution statuses of the groups obtained using CA.

PCA/FA techniques helped to identify the potential sources of water quality degradation. This study comprehensively analyzed the water quality of LHW and identified three significant sources responsible for pollution of Lake Hawassa Watershed in dry and wet seasons affecting the water quality. Accordingly, the pollution is due to mixed sources including point sources such as municipal and industrial effluents, natural processes, livestock, urban runoff, and non-point sources from agricultural activities.

Poor industrial effluent management combined with non-point sources from agriculture and urban runoff contribute significantly to the pollution of Lake Hawassa. Discharge of industrial effluents into the surface water system is the largest point source of anthropogenic pollution. Diffuse sources that contribute enormously to LHW come from agricultural activities, i.e., intensive farming and livestock (F1, F2, and F3).

We conclude that effective management of point and non-point source pollution is imperative to improve domestic, industrial, livestock, and agricultural runoff to reduce pollutant inputs into the lake. A stringent management that requires a comprehensive application of technologies such as fertilizer management, ecological ditches, constructed wetlands, and buffer strips should complement proper municipal and industrial wastewater treatment set-up.

Furthermore, application of indigenous aeration practices such as the use of drop structures at critical locations would help improve water quality in the lake watershed.

Author Contributions: Conceptualization and improvement of the methodology, S.M.L., M.D.U. and A.M.; Data collection, analysis, and interpretation, S.M.L.; Writing of the original manuscript, S.M.L.; Supervision, follow-up of the work, reviewing and modifying of the manuscript, A.M. and M.D.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research was part of the DAAD-EECBP Home Grown PhD Scholarship Program at EECBP Homegrown PhD Program, 2019 (57472170). The Open Access Department, University of Rostock, has funded the APC.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the first author. The data are not publicly available, as they are experimental.

Acknowledgments: We are grateful to the German Academic Exchange Service (DAAD) for offering a stipend for the first author in the course of the study.

Conflicts of Interest: The authors claim no conflict of interest in connection with the work submitted.

References

1. Cherif, E.K.; Salmoun, F.; Mesas-Carrascosa, F.J. Determination of Bathing Water Quality Using Thermal Images Landsat 8 on the West Coast of Tangier: Preliminary Results. *Remote Sens.* **2019**, *11*, 972. [[CrossRef](#)]
2. Fan, X.; Cui, B.; Zhao, H.; Zhang, Z.; Zhang, H. Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. *Procedia Environ. Sci.* **2010**, *2*, 1220–1234. [[CrossRef](#)]

3. Noori, R.; Sabahi, M.; Karbassi, A.; Baghvand, A.; Zadeh, H.T. Multivariate statistical analysis of surface water quality based on correlations and variations in the data set. *Desalination* **2010**, *260*, 129–136. [[CrossRef](#)]
4. Wang, Q.; Li, S.; Jia, P.; Qi, C.; Ding, F. A Review of Surface Water Quality Models. *Sci. World J.* **2013**, *2013*, 1–7. [[CrossRef](#)]
5. Simeonov, V.; Stratis, J.; Samara, C.; Zachariadis, G.; Voutsas, D.; Anthemidis, A.; Sofoniou, M.; Kouimtzis, T. Assessment of the surface water quality in Northern Greece. *Water Res.* **2003**, *37*, 4119–4124. [[CrossRef](#)]
6. Massoud, M.A.; El-Fadel, M.; Scrimshaw, M.D.; Lester, J.N. Factors influencing development of management strategies for the Abou Ali River in Lebanon. I: Spatial variation and land use. *Sci. Total Environ.* **2006**, *362*, 15–30. [[CrossRef](#)]
7. Wang, X.-L.; Lu, Y.-L.; Han, J.-Y.; He, G.-Z.; Wang, T.-Y. Identification of anthropogenic influences on water quality of rivers in Taihu watershed. *J. Environ. Sci.* **2007**, *19*, 475–481. [[CrossRef](#)]
8. Abebe, Y.D.; Geheb, K. Wetlands of Ethiopia. In *Proceedings of a Seminar on the Resources and Status of Ethiopia's Wetlands*; IUCN, The World Conservation Union: Gland, Switzerland, 2003.
9. Cherif, E.K.; Salmoun, F.; El Yemlahi, A.; Magalhaes, J.M. Monitoring Tangier (Morocco) coastal waters for As, Fe and P concentrations using ESA Sentinels-2 and 3 data: An exploratory study. *Reg. Stud. Mar. Sci.* **2019**, *32*, 100882. [[CrossRef](#)]
10. Amare, T.A.; Yimer, G.T.; Workagegn, K.B. Assessment of Metal concentration in water, sediment and macrophyte plant collected from Lake Hawassa, Ethiopia. *Environ. Anal. Toxicol.* **2014**, *05*, 1–7. [[CrossRef](#)]
11. Teshome, F.B. Seasonal water quality index and suitability of the water body to designated uses at the eastern catchment of Lake Hawassa. *Environ. Sci. Pollut. Res.* **2019**, *27*, 279–290. [[CrossRef](#)]
12. Zigde, H.; Tsegaye, M.E. Evaluation of the current water quality of Lake Hawassa, Ethiopia. *Int. J. Water Resour. Environ. Eng.* **2019**, *11*, 120–128.
13. Amare, D. *Assessment of Non-Point Source Pollution in Lake Awassa Watershed Using the Annualized Agricultural Non-Point Source (AnnAGNPS) Model*; Addis Ababa University: Addis Ababa, Ethiopia, 2008.
14. Drevnick, P.E.; Engstrom, D.; Driscoll, C.T.; Swain, E.; Balogh, S.J.; Kamman, N.C.; Long, D.T.; Muir, D.; Parsons, M.J.; Rolffhus, K.R.; et al. Spatial and temporal patterns of mercury accumulation in lacustrine sediments across the Laurentian Great Lakes region. *Environ. Pollut.* **2012**, *161*, 252–260. [[CrossRef](#)] [[PubMed](#)]
15. Kebede, W.; Tefera, M.; Habitamu, T.; Alemayehu, T. Impact of Land Cover Change on Water Quality and Stream Flow in Lake Hawassa Watershed of Ethiopia. *Agric. Sci.* **2014**, *5*, 647–659. [[CrossRef](#)]
16. Lencha, S.M.; Tränckner, J.; Dananto, M. 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* **2021**, *18*, 8904. [[CrossRef](#)]
17. Nigussie, K.; Chandravanshi, B.S.; Wondimu, T. Correlation among trace metals in Tilapia (*Oreochromis niloticus*), sediment and water samples of lakes Awassa and Ziway, Ethiopia. *Int. J. Biol. Chem. Sci.* **2011**, *4*. [[CrossRef](#)]
18. Wondrade, N.; Dick, Ø.B.; Tveite, H. GIS based mapping of land cover changes utilizing multi-temporal remotely sensed image data in Lake Hawassa Watershed, Ethiopia. *Environ. Monit. Assess.* **2014**, *186*, 1765–1780. [[CrossRef](#)]
19. Abiye, T.A. Environmental resources and recent impacts in the Awassa collapsed caldera, Main Ethiopian Rift. *Quat. Int.* **2008**, *189*, 152–162. [[CrossRef](#)]
20. *APHA Standard Methods for the Examination of Water and Wastewater*, 23rd ed.; Baird, R.B.; Eaton, A.D.; Rice, E.W.; Brigewater, L.L. (Eds.) American Public Health Association (APHA): Washington, DC, USA; American Water Works Association (AWWA): Washington, DC, USA; Water Environment Federati, Water Environment Federation: Washington, DC, USA, 2017; ISBN 9780123821652.
21. Emerson, K.; Russo, R.C.; Lund, R.E.; Thurston, R.V. Aqueous Ammonia Equilibrium Calculations: Effect of pH and Temperature. *J. Fish. Res. Board Can.* **1975**, *32*, 2379–2383. [[CrossRef](#)]
22. Angello, Z.A.; Tränckner, J.; Behailu, B.M. Spatio-Temporal Evaluation and Quantification of Pollutant Source Contribution in Little Akaki River, Ethiopia: Conjunctive Application of Factor Analysis and Multivariate Receptor Model. *Pol. J. Environ. Stud.* **2021**, *30*, 23–34. [[CrossRef](#)]
23. Singh, K.P.; Malik, A.; Sinha, S. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—A case study. *Anal. Chim. Acta* **2005**, *538*, 355–374. [[CrossRef](#)]
24. Sharma, S.; Reddy, A.S.; Dalwani, R.R. Ecological water quality index development and evaluation of water quality of the Satluj river. *Indian J. Environ. Prot.* **2015**, *35*, 477–489.
25. Palma, P.; Alvarenga, P.; Palma, V.L.; Fernandes, R.M.; Soares, V.M.; Amadeu, M.A.; Barbosa, I.R. Assessment of anthropogenic sources of water pollution using multivariate statistical techniques: A case study of the Alqueva's reservoir, Portugal. *Environ. Monit. Assess.* **2010**, *165*, 539–552. [[CrossRef](#)]
26. Liu, C.-W.; Lin, K.-H.; Kuo, Y.-M. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Sci. Total Environ.* **2003**, *313*, 77–89. [[CrossRef](#)]
27. Zhang, Q.; Li, Z.; Zeng, G.; Li, J.; Fang, Y.; Yuan, Q.; Wang, Y.; Ye, F. Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: A case study of Xiangjiang watershed, China. *Environ. Monit. Assess.* **2009**, *152*, 123–131. [[CrossRef](#)] [[PubMed](#)]
28. 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. *J. Environ. Sci.* **2011**, *23*, 1460–1471. [[CrossRef](#)]
29. Chen, P.; Li, L.; Zhang, H. Spatio-Temporal Variations and Source Apportionment of Water Pollution in Danjiangkou Reservoir Basin, Central China. *Water* **2015**, *7*, 2591–2611. [[CrossRef](#)]

30. Ma, X.; Wang, L.; Yang, H.; Li, N.; Gong, C. Spatiotemporal Analysis of Water Quality Using Multivariate Statistical Techniques and the Water Quality Identification Index for the Qinhuai River Basin, East China. *Water* **2020**, *12*, 2764. [[CrossRef](#)]
31. Mahmud, R.; Inoue, N.; Sen, R. Assessment of Irrigation Water Quality by Using Principal Component Analysis in an Arsenic Affected Area of Bangladesh. *J. Soil Nat.* **2007**, *1*, 8–17.
32. Singh, K.P.; Malik, A.; Mohan, D.; Sinha, S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Res.* **2004**, *38*, 3980–3992. [[CrossRef](#)]
33. Sharma, M.; Kansal, A.; Jain, S.; Sharma, P. Application of Multivariate Statistical Techniques in Determining the Spatial Temporal Water Quality Variation of Ganga and Yamuna Rivers Present in Uttarakhand State, India. *Water Qual. Expos. Health* **2015**, *7*, 567–581. [[CrossRef](#)]
34. Banda, T.D.; Kumarasamy, M. Application of Multivariate Statistical Analysis in the Development of a Surrogate Water Quality Index (WQI) for South African Watersheds. *Water* **2020**, *12*, 1584. [[CrossRef](#)]
35. Yilma, M.; Kiflie, Z.; Windsperger, A.; Gessese, N. Assessment and interpretation of river water quality in Little Akaki River using multivariate statistical techniques. *Int. J. Environ. Sci. Technol.* **2018**, *16*, 3707–3720. [[CrossRef](#)]
36. Mohd, N.F.; Samsudin, M.S.; Mohamad, I.; Awaluddin, M.R.; Mansor, A.; Hafizan, J.; Ramli, N. River water quality modeling using combined principle component analysis (PCA) and multiple linear regressions (MLR): A case study at Klang River, Malaysia. *World Appl. Sci. J.* **2011**, *14*, 73–82.
37. Kazi, T.; Arain, M.; Jamali, M.; Jalbani, N.; Afridi, H.; Sarfraz, R.; Baig, J.A.; Shah, A.Q. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicol. Environ. Saf.* **2009**, *72*, 301–309. [[CrossRef](#)] [[PubMed](#)]
38. Helena, B.; Pardo, R.; Vega, M.; Barrado, E.; Fernandez, J.M.; Fernand, L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res.* **2000**, *34*, 807–816. [[CrossRef](#)]
39. Sojka, M.; Siepak, M.; Ziola-frankowska, A.; Frankowski, M. Application of multivariate statistical techniques to evaluation of water quality in the Mala Welnia River (Western Poland). *Environ. Monit. Assess.* **2008**, *147*, 159–170. [[CrossRef](#)]
40. Grubbs, F.E.; Beck, G. Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations. *Technometrics* **1972**, *14*, 847–854. [[CrossRef](#)]
41. Suhaimi, N.; Ghazali, N.A.; Nasir, M.Y.; Mokhtar, M.I.; Ramli, N.A. Markov Chain Monte Carlo Method for Handling Missing Data in Air Quality Datasets. *Malays. J. Anal. Sci.* **2017**, *21*, 552–559.
42. Vega, M.; Pardo, R.; Barrado, E.; Debán, L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res.* **1998**, *32*, 3581–3592. [[CrossRef](#)]
43. Alberto, W.D.; del Pilar, D.M.; Valeria, A.M.; Fabiana, P.S.; Cecilia, H.A.; Ángeles, B.M.d.l. Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A Case Study: Suquia River Basin (Córdoba–Argentina). *Water Res.* **2001**, *35*, 2881–2894. [[CrossRef](#)]
44. Shrestha, S.; Kazama, F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environ. Model. Softw.* **2007**, *22*, 464–475. [[CrossRef](#)]
45. Wu, E.M.-Y.; Kuo, S.-L. Applying a Multivariate Statistical Analysis Model to Evaluate the Water Quality of a Watershed. *Water Environ. Res.* **2012**, *84*, 2075–2085. [[CrossRef](#)]
46. Swanson, R.A.; Holton, E.F. *Research in Organizations Foundations and Methods of Inquiry*; Holton, E.F., Ed.; Berrett-Koehler Organization: San Francisco, CA, USA, 2005; pp. 115–142.
47. Liou, S.-M.; Lo, S.-L.; Wang, S.-H. A Generalized Water Quality Index for Taiwan. *Environ. Monit. Assess.* **2004**, *96*, 35–52. [[CrossRef](#)]
48. Chen, Y.-C.; Yeh, H.-C.; Wei, C. Estimation of River Pollution Index in a Tidal Stream Using Kriging Analysis. *Int. J. Environ. Res. Public Health* **2012**, *9*, 3085–3100. [[CrossRef](#)] [[PubMed](#)]
49. Mahesh Kumar, M.K.; Mahesh, M.K.; Sushmitha, B.R. CCME Water Quality Index and Assessment of Physico-Chemical Parameters of Chikkakere, Periyapatna, Mysore District, Karnataka State, India. *Int. J. Innov. Res. Sci. Eng. Technol.* **2014**, *3*. [[CrossRef](#)]
50. Varol, M.; Şen, B. Assessment of surface water quality using multivariate statistical techniques: A case study of Behrimaz Stream, Turkey. *Environ. Monit. Assess.* **2008**, *159*, 543–553. [[CrossRef](#)]
51. Zhou, F.; Liu, Y.; Guo, H. Application of Multivariate Statistical Methods to Water Quality Assessment of the Watercourses in Northwestern New Territories, Hong Kong. *Environ. Monit. Assess.* **2006**, *132*, 1–13. [[CrossRef](#)] [[PubMed](#)]
52. Parinet, B.; Lhote, A.; Legube, B. Principal component analysis: An appropriate tool for water quality evaluation and management—Application to a tropical lake system. *Ecol. Model.* **2004**, *178*, 295–311. [[CrossRef](#)]
53. Rencher, A. *Methods of Multivariate Analysis Second Edition*, 2nd ed.; A Wiley-Interscience Publication: New York, NY, USA, 2002; pp. 156–504. ISBN 0471418897.
54. Gummadi, S.; Swarnalatha, G.; Vishnuvardhan, Z.; Harika, D. Statistical Analysis of the Groundwater Samples from Bapatla Mandal, Guntur District, Andhra Pradesh, India. *IOSR J. Environ. Sci. Toxicol. Food Technol.* **2014**, *8*, 27–32. [[CrossRef](#)]
55. Karakuş, C.B. Evaluation of water quality of Kızılırmak River (Sivas/Turkey) using geo-statistical and multivariable statistical approaches [Internet]. *Environ. Dev. Sustain.* **2020**, *22*, 4735–4769. [[CrossRef](#)]
56. Shroff, P.; Vashi, R.T.; Champa neri, V.A.; Patel, K.K. Correlation Study among water quality parameters of groundwater of Valsad District of south Gujarat (India). *J. Fundam. Appl. Sci.* **2015**, *151*, 1–10. [[CrossRef](#)]

57. Kumar, M.; Ramanathan, A.; Rao, M.S.; Kumar, B. Identification and evaluation of hydrogeochemical processes in the groundwater environment of Delhi, India. *Environ. Earth Sci.* **2006**, *50*, 1025–1039. [[CrossRef](#)]
58. Gebre-mariam, Z. The effects of wet and dry seasons on concentrations of solutes and phytoplankton biomass in seven Ethiopian rift-valley lakes. *Limnologia* **2002**, *179*, 169–179. Available online: <http://www.sciencedirect.com/science/article/B7GX1-4GWP957-6/2/4d9fcc9b986796cd8cfb4a1b3193ad03> (accessed on 1 December 2002).
59. Yilmaz, E.; Koç, C. Physically and Chemically Evaluation for the Water Quality Criteria in a Farm on Akçay. *J. Water Resour. Prot.* **2014**, *06*, 63–67. [[CrossRef](#)]
60. Poisson, A. Conductivity/salinity/temperature relationships of diluted and concentrated standard seawater. *Mar. Geodesy* **1982**, *5*, 359–361. [[CrossRef](#)]
61. Taylor, M.; Elliott, H.A.; Navitsky, L.O. Relationship between total dissolved solids and electrical conductivity. *Water Sci. Technol.* **2018**, *1*–7. [[CrossRef](#)]
62. Dhanasekarapandian, M.; Chandran, S.; Devi, D.S.; Kumar, V. Spatial and temporal variation of groundwater quality and its suitability for irrigation and drinking purpose using GIS and WQI in an urban fringe. *J. Afr. Earth Sci.* **2016**, *124*, 270–288. [[CrossRef](#)]
63. Boyle, T.P.; Fraleigh, H.D. Natural and anthropogenic factors affecting the structure of the benthic macroinvertebrate community in an effluent-dominated reach of the Santa Cruz River, AZ. *Ecol. Indic.* **2003**, *3*, 93–117. [[CrossRef](#)]
64. Johnson, R.A. *Applied Multivariate Statistical Analysis*, 6th ed.; Recker, P., Ryan, D., Behrens, L.M., Eds.; Pearson Education, Inc.: San Antonio, TX, USA, 2007; ISBN 0131877151.
65. McKenna, J. An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environ. Model. Softw.* **2003**, *18*, 205–220. [[CrossRef](#)]
66. Kasier, H.F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [[CrossRef](#)]
67. 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. *Int. Soil Water Conserv. Res.* **2016**, *4*, 284–292. [[CrossRef](#)]
68. Boyacioglu, H.; Boyacioglu, H. Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey. *Environ. Geol.* **2008**, *54*, 275–282. [[CrossRef](#)]
69. Su, S.; Zhi, J.; Lou, L.; Huang, F.; Chen, X.; Wu, J. Spatio-temporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques. *Phys. Chem. Earth Parts A/B/C* **2011**, *36*, 379–386. [[CrossRef](#)]
70. Gazzaz, N.M.; Yusoff, M.K.; Ramli, M.F.; Aris, A.Z.; Juahir, H. Characterization of spatial patterns in river water quality using chemometric pattern recognition techniques. *Mar. Pollut. Bull.* **2012**, *64*, 688–698. [[CrossRef](#)] [[PubMed](#)]
71. Tibebe, D.; Beshah, F.Z.; Lemma, B.; Kassa, Y.; Bhaskarwar, A.N. External Nutrient Load and Determination of the Trophic Status of Lake Ziway. *CSVTU Int. J. Biotechnol. Bioinform. Biomed.* **2018**, *3*, 1–16. [[CrossRef](#)]
72. Meshesha, D.T.; Tsunekawa, A.; Tsubo, M. Continuing land degradation: Cause-effect in Ethiopia's Central Rift Valley. *Land Degrad. Dev.* **2012**, *23*, 130–143. [[CrossRef](#)]
73. Badillo-Camacho, J.; Reynaga-Delgado, I.; Barcelo-Quintal, I.; Valle, P.F. Water Quality Assessment of a Tropical Mexican Lake Using Multivariate Statistical Techniques. *J. Environ. Prot.* **2015**, *6*, 215–224. [[CrossRef](#)]
74. Smith, V.H. Low Nitrogen to Phosphorus Ratios Favor Dominance by Blue-Green Algae in Lake Phytoplankton. *Science* **1983**, *221*, 669–672. [[CrossRef](#)] [[PubMed](#)]
75. Fisher, M.M.; Reddy, K.R.; James, R.T. Internal Nutrient Loads from Sediments in a Shallow, Subtropical Lake. *Lake Reserv. Manag.* **2005**, *21*, 338–349. [[CrossRef](#)]