JWMM

Application of Multivariate Statistical Techniques in the Assessment of Surface Water Quality in Dak Lak Province, Vietnam

Nguyen Thi Oanh and Ha Danh Duc (2026)

Dong Thap University, Cao Lanh City, Vietnam

DOI: https://doi.org/10.14796/JWMM.C574

ABSTRACT

Water quality monitoring is crucial for assessing water quality and identifying sources of water pollutants. This study assessed the surface water quality in Dak Lak province, Vietnam using multivariate statistical techniques. Pearson correlation, water quality index (WQI), principal component analysis (PCA), and cluster analysis (CA) were determined based on various physical, chemical, and biological indicators. The study found that BOD_5 concentrations exceeded permissible limits in 80% of river water samples and 36% of lake water samples. Meanwhile, all samples met the allowable limits for $N\text{-}NO_3^-$. The assessment revealed that surface water pollution in the region is due to both natural and human activities. TSS mainly originated from natural processes or mineral exploitation, while BOD_5 and COD were primarily from domestic or waste treatment. $N\text{-}NO_2^-$, $N\text{-}NO_3^-$, $N\text{-}NH_4^+$, $o\text{-}PO_4^{3-}$ were traced back to domestic and industrial wastewater, and coliforms were from domestic and non-human sources. Water samples from areas with high population density had lower WQI values than areas surrounded by agriculture fields and sparse populations. Finally, the study noted that water flow with high pollutant concentrations during the rainy season. These findings are expected to provide valuable insights into surface water pollution in Dak Lak province.

1. INTRODUCTION

Surface water plays an important role in ecosystems, providing drinking and domestic water for humans, as well as for irrigation and other purposes. However, water sources have been polluted worldwide due to human activities and natural processes (WHO 2016).

Oanh, N.T. and H.D. Duc. 2026. "Application of Multivariate Statistical Techniques in the Assessment of Surface Water Quality in Dak Lak Province, Vietnam." *Journal of Water Management Modeling* 34: C574. https://doi.org/10.14796/JWMM.C574 www.chijournal.org ISSN: 2292-6062

Anthropogenic sources of pollutants may be derived from domestic, industrial, and agricultural sectors. Moreover, land use, climate, and basin conditions are important factors affecting surface water (Anh et al. 2023). Besides, precipitation amount affects spatio-temporal variations in river water quality (Mena-Rivera et al. 2017). In addition, drought and floods alter the number of substances discharged to surface water (Peña-Guerrero et al. 2020).

Vietnam's Central Highlands were originally covered by vast tracts of primary forest. However, rapid population growth, agricultural expansion, and urbanization resulted in deforestation over the past decades. Forests have been cleared to grow food crops and coffee in this region (Meyfroidt et al. 2013). The agricultural land expansion has increased vulnerability in steep-slope regions (Frohlich et al. 2013; Lin et al. 2016). This change has resulted in poor surface water quality. For example, higher annual sediment, total nitrogen, and total phosphorus loadings were found upstream somewhere in Vietnam's Central Highlands (Tram et al. 2022). Water in rivers and lakes is an important source of human usage. Moreover, major crops grown in the region are highly dependent on irrigation, especially during the dry season (Amarasinghe et al. 2015).

Dak Lak province is in Vietnam Central Highlands with a steep-slope terrain. Most rivers here originate from mountains and flow through forests, agriculture fields, densely populated settlements, industrial zones, and other sites. The water network scattered on the watershed scale and the information on pollution sources are some of the challenges in water evaluation in mountainous areas. Therefore, only a few studies have described the water quality in the Central Highlands (Amarasinghe et al. 2015). Tram et al. (2018) showed that local land-use policies and human activities significantly affect hydrology and increase the erosion and nutrient concentrations downstream in the Central Highlands.

Multivariate statistical methods are usually applied to analyze water characteristics simultaneously. Cluster analysis (CA) and principal component analysis (PCA) are valuable to assess the quality of surface water (Anh et al. 2023; Ha et al. 2024). Pearson correlation is also a valuable tool for quantifying the strength and direction of linear relationships between different water parameters. In addition, a Water Quality Index (WQI) is a method used to assess the overall quality of water by combining multiple water indicators into a single indicator. Several studies on water quality in lowland areas have been conducted in Vietnam using these methods (Giao et al. 2022; Hong and Giao 2022; Ha et al. 2024), but there have been few studies on water quality in upland areas.

This study determined the quality of river and lake water at different positions in Dak Lak province. The water quality at different sites was evaluated using multicriteria statistical approaches, through which the sources of water pollutants were analyzed. This study aimed to provide valuable information on river and lake water status in Dak Lak province representing the highlands, mainly mountainous terrain, for future management.

2. METHODOLOGY

2.1 Study area

Dak Lak, a province in the Central Highlands in Vietnam, has an average altitude of 400–800 m above sea level. The weather is divided into two distinct seasons: the rainy season and the dry season. The rainy season usually starts from May to October or sometimes to November. The average annual rainfall in the whole province is from 1600–1800 mm. Surface water in Dak Lak includes rivers, streams, and lakes. Rivers and streams mostly originate from high mountains, flow through forests, agriculture, industry, and urban areas, and then join larger rivers. The main river is Serepok, which is a branch of the Mekong River. Majority positions of rivers were monitored in Buon Ma Thuot city with a population density and location of some industrial zones. Lakes include artificial lakes serving agriculture and natural ones.

In this study, surface water collection was conducted in main rivers and branches of rivers (considered as rivers), and lakes. The coding, sites, coordinates, and description of sites are shown in Tables 1a and 1b. A map of all sites of water collection is also shown in Figure 1. Surface water was collected from 26 locations in rivers and 13 locations in lakes to measure the concentration of water parameters and analyze QWI and principal components. The collection was conducted in March, April, June, August, October, and December of 2023. A total of 155 and 78 samples were collected from rivers and lakes, respectively.

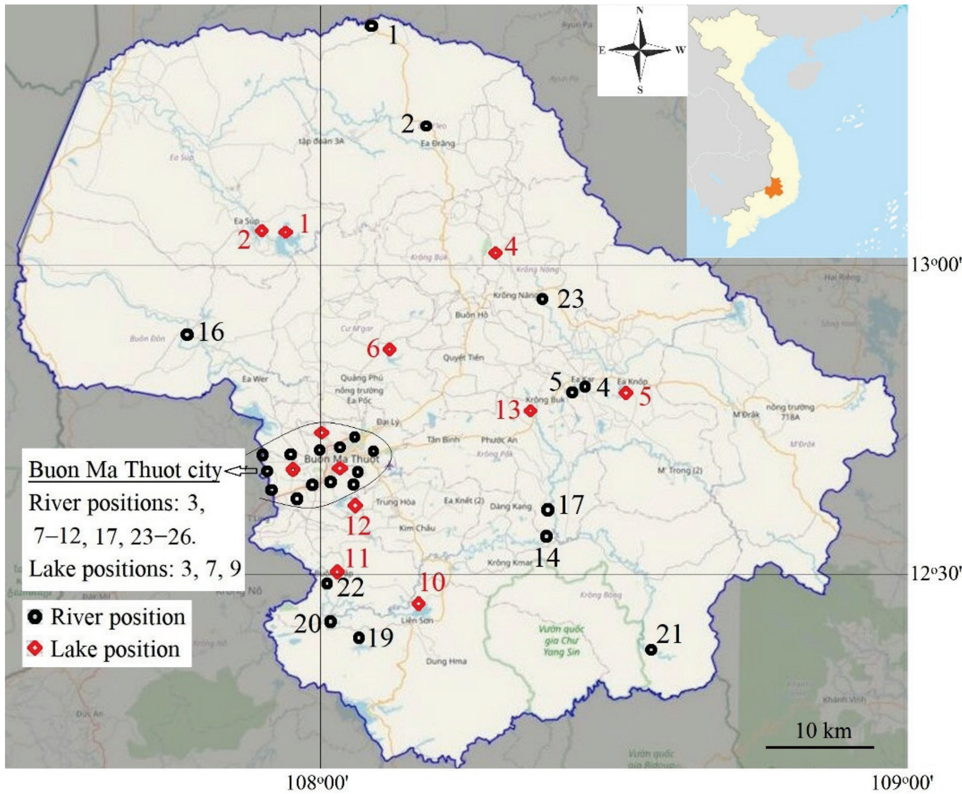

Figure 1 Sampling locations in lakes and rivers (this map was modified from the map available at https://dandautu.vn/m/ban-do-tinh-dak-lak) (*in Vietnamese*).

Table 1 (a) Location and characteristics of monitoring positions in rivers.

No.	Coding	Rivers	Coordinates		Description of site
			East	North	
1	NM 1	Ea H'leo	108°05'29"	13°24'31"	Near a residential area.
2	NM 2	Ea Khal	108°12'11"	13°11'42"	Near a residential area.
3	NM 3	Sut M'du	108°04'32"	12°44'58"	Both sides of the river are mainly industrial and agricultural areas.
4	NM 5	Krong Nang	108°31'51"	12°49'22"	Upstream of a cassava starch factory
5	NM 6	Krong Nang	108°31'56"	12°49'22"	Downstream of a cassava starch factory
6	NM 7	Ea Drueh	107°59'49"	12°40'59"	Upstream of a waste treatment site
7	NM 8	Buon Kuop hydroelectric reservoir	107°55'34"	12°31'53"	Surrounding area is mainly small trees and shrubs, far resident areas.
8	NM 9	Serepok	107°55'19"	12°31'49"	Downstream of Buon Kuop hydroelectric power plant
9	NM 11	Ea Tam	108°01'54"	12°39'38"	Both sides of the river are densely populated.

No.	Coding	Rivers	Coordinates		Description of site
			East	North	
10	NM 12	Ea Nhuol	108°02'01"	12°41'24"	Both sides of the river are densely populated.
11	NM 13	Ea Nhuol	108°01'07"	12°41'06"	Both sides of the river are densely populated.
12	NM 14	Doc Hoc	108°02'20"	12°40'51"	Both sides of the river are densely populated.
13	NM 15	Krong Kmar	108°20'33"	12°30'50"	Near agricultural areas
14	NM 16	Serepok	107°55'45"	12°36'45"	Surrounding area has dense population and it is near Phu Hoa industrial zone
15	NM 17	Serepok	107°46'53"	12°53'27"	Near Yok Đôn national forest and a residential area.
16	NM 18	Krong Buk	108°22'32"	12°46'18"	Vegetation on both banks is mainly small shrubs and weeds.
17	NM 19	Serepok	107°56'60"	12°35'44"	Near industrial zones
18	NM 20	Krong No (upstream)	108°08'14"	12°10'55"	Both banks are mainly agricultural sites and sparsely populated.
19	NM 21	Krong No (downstream)	107°59'12"	12°29'39"	Both sites of the river are mainly used for agricultural cultivation. The river is also used for mineral exploitation and aquaculture.
20	NM 22	Krong Ana (upstream)	108°22'23"	12°32'29"	Both sites of the river are mainly used for agricultural cultivation. The river is also used for mineral exploitation site and aquaculture.
21	NM 23	Krong Ana (downstream)	107°59'14"	12°29'43"	Both sites of the river are mainly used for agricultural cultivation. The river is also used for mineral exploitation site and aquaculture.
22	NM 24	Krong Buk	108°16'56"	12°55'46"	Both sides of the river are densely populated.
23	NM 25	A stream in Buon Ma Thuot City	108°00'03"	12°42'48"	Upper of a solid waste landfill
24	NM 26	A stream in Buon Ma Thuot City	107°59'38"	12°42'48"	Near a solid waste landfill
25	NM 27	Ea Mleo stream	108°03'09"	12°39'40"	Near a hospital and an agricultural area.
26	NM 28	Ea Drueh stream	107°59'49"	12°41'02"	Outlet of a wastewater treatment site.

Table 1 (b) Location and characteristics of monitoring positions in lakes.

No.	Coding	Rivers	Coordinates		Description of site
			East	North	
1	NH 1	Upper Ea Sup	107°55'59"	13°02'29"	Forrest site.
2	NH 2	Lower Ea Sup	107°89'62"	13°06'34"	Near aquaculture area, sparse population.
3	NH 3	Sứt M'dư	108°04'38"	12°44'54"	Near Tan An industrial zone.
4	NH 4	Phu Xuan	108°22'42"	12°55'09"	Surrounding is plots of grass and shrub, and an agriculture area. Business activities are next to the lake.
5	NH 5	Ea Knop	108°32'23"	12°47'36"	Surrounding area has dense population and a cane sugar factory.
6	NH 6	Krong Jing	108°45'29"	12°44'33"	Near an aquaculture area, sparse population.
7	NH 7	Ea Kao	108°02'15"	12°36'19"	Near an agricultural area, sparse population. The lake is also used for aquaculture.
8	NH 8	Ea Nhai	108°12'00"	12°43'54"	Surrounding area is agricultural
9	NH 9	Ea Chu Cap	108°08'16"	12°39'42"	Surrounding area is agricultural
10	NH 10	Lak	108°11'02"	12°24'47"	Surrounding area is mountains and business places.
11	NH 12	Sen	108°09'39"	12°46'28"	Surrounding area has dense population and business places.
12	NH 13	Ea Sim	108°09'28"	12°37'42"	Both sides of the river are densely populated.
13	NH 14	Krong Buk	108°22'23"	12°46'48"	Surrounding area is agricultural areas.

2.2 Surface water analysis

The data for this study was obtained from the monitoring process carried out by the Department of Natural Resources and Environment of Dak Lak province. Water sampling and analysis were conducted in accordance with the Vietnamese guidelines. Samples were collected, stored at 4°C, and then transferred to laboratories for analysis. Water quality parameters such as *pH*, dissolved oxygen (*DO*), chemical oxygen demand (*COD*), biological oxygen demand (BOD_5), total suspended solids (*TSS*), $N\text{-}NO_3^-$, $N\text{-}NO_2^-$, $o\text{-}PO_4^{3-}$, and coliforms were analyzed in laboratories using standard methods described in a previous study (Giao et al. 2022). Additionally, specific parameters like temperature and DO were measured using handheld devices. The permissible limits for these parameters, as used for domestic water supply (after appropriate treatment), conservation of aquatic life, and other purposes, are shown in Table 2.

Table 2 Water parameters and their permissible limits (Ha et al. 2024).

Abbreviation	Units	Permissible limits
pH		6–8.5
BOD_5	mg/L	6
COD	mg/L	15
TSS	mg/L	30
DO	mg/L	≥ 5
$N\text{-}NO_{3-}$	mg/L	10
$N\text{-}NO_{2-}$	mg/L	0.05
$N\text{-}NH_{4+}$	mg/L	0.3
$o\text{-}PO_4^{3-}$	mg/L	0.2
Coliform	MPN/100 ml	5000

2.3 Data analysis

Water quality index (WQI)

WQI is used to evaluate the complex effects of individual water quality values to reflect water quality as a whole. WQI was calculated according to the guidance of the Vietnam Environment Administration in 2019. Major parameters were used to calculate WQI. The WQI index was calculated according to Giao et al. (2021) as Equation 1:

$$WQI = \frac{WQI_{pH}}{100}\left[\frac{1}{2}\sum WQI_a \cdot \sum WQI_b\right]^{0.5} \tag{1}$$

Where:

WQI_a = WQI value for the parameters BOD_5, COD, $N\text{-}NH_4^+$, $N\text{-}NO_3^-$, $N\text{-}NO_2^-$ and $o\text{-}PO_4^{3-}$,
WQI_b = coliforms, and
WQI_{pH} = *pH*.

The calculation of WQI_a, WQI_b, and WQI_{pH} was described in a previous study (Phu 2019). The WQI parameter has a value from 0 to 100 and is divided into 6 levels used for different purposes, as presented in Table 3.

Table 3 WQI values and purpose (Nguyen and Huynh 2022; Ha et al. 2024).

Value (points)	Water quality	Purpose
>90–100	Very good	Used for domestic water supply purposes
>75–90	Good	Used for water quality, suitable for use for domestic water supply but need suitable treatment measures
>50–75	Average	Used for irrigation and other equivalent purposes
>25–50	Bad	Used for navigation and other equivalent purposes
10–25	Poor	Water is heavily tarnished, needing treatment measures
<10	Very poor	Contaminated water, and needs to be remedied and treated

Pearson correlation analysis

Pearson correlation analysis is a preliminary descriptive technique to estimate the degree of association among multiple variables. Each correlation coefficient indicates the inverse correlation between two parameters. It is calculated according to Equation 2:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (2)$$

Where:

r = Pearson r correlation coefficient between parameter X and parameter Y,
n = number of observations,
X_i = value of X (for ith observation),
Y_i = value of Y (for ith observation),
$\bar{X}$ = mean of the observed values of parameter X, and
$\bar{Y}$ = mean of the observed values of parameter Y.

Pearson correlation values vary from −1 to 1. The correlation is considered weak/low if r has an absolute value less than 0.3, and moderate when its absolute value is in a range of 0.3–0.7; in contrast, its correlation is high/strong.

Analysis of principal component and cluster

Principal component analysis (PCA) was used to determine the main water parameters in the variation of the originally correlated dataset. The PCA was analyzed using the software PRIMER6. Data of all parameters were first pretreated using normalized variables before the PCA setting. Each principal component (PC) is obtained by multiplying the original correlated variables. PCs are the result of a linear combination of the original variables. They include the eigenvalues, variation percentage, cumulative percent of the variance, and a list of coefficients called "weighing factors." Eigenvalues greater than 1 are considered significant effects. PCs explaining a higher percentage of water variation are more important. The absolute value of a weighing factor strongly correlates with PCs and parameters when it is greater than 0.75, average if it is in a range of 0.5–0.75, and weak if it is greater than 0.3 but less than 0.5.(Mukaka 2012).

Cluster analysis (CA) was also applied to analyze spatial variations. The software PRIMER6 was used to build a hierarchical dendrogram of sampling locations. CA performed groups of sample locations based on the similarity of water properties. Ward's method is used to perform hierarchical clustering, and Euclidean distance shows different similarities between the sites and is represented by a dendrogram.

2.4 Statistical analysis

Data of WQI values are presented as mean ± standard deviation. Duncan's multiple range test ($p < 0.05$) in SPSS 22.0 was used to analyze significant differences among means.

3. RESULTS AND DISCUSSION

3.1 Physiochemical and biological parameters of surface water in rivers and lakes

The *pH* and *DO* values in river and lake water showed minimal fluctuations. However, other parameters varied greatly depending on the time and location of collection. The average concentrations of *TSS*, $N\text{-}NO_2^-$, $N\text{-}NO_3^-$, $N\text{-}NH_4^+$, $o\text{-}PO_4^{3-}$, and coliform in river water were significantly higher than those in lake water (Table 4). On the other hand, *pH, DO,* BOD_5, and *COD* values did not differ significantly between the two water sources (Table 4). The number of river water samples exceeding permissible limits for domestic water supply parameters was higher than that of lake water, except for $N\text{-}NO_3^-$ which was below the limit for all samples. Among these parameters, BOD_5 had the highest number of samples above the limit, while only one sample of river water had an $o\text{-}PO_4^{3-}$ concentration above the standard.

Table 4 Average hydrochemical and microbiological characteristics of water sampled from rivers and lakes.

Parameters	River		Lake	
	Mean ± ST	Samples above permissible limit (%)	Mean ± ST	Samples above permissible limit (%)
pH	6.7±0.4	2	6.7±0.4	1
DO (mg/L)	5.7±0.6	6	5.6±0.5	2
TSS (mg/L)	27.0±29.8	7	14.5±21.7	1
BOD_5 *(mg/L)*	6.9±4.7	80	7.5±8.1	36
COD (mg/L)	12.9±8.2	40	14.1±15.0	23
$N\text{-}NO_{2-}$ *(mg/L)*	0.18±0.38	60	0.03±0.09	5
$N\text{-}NO_{3-}$ *(mg/L)*	0.44±0.47	0	0.27±0.20	0
$N\text{-}NH_{4+}$ *(mg/L)*	0.41±0.53	59	0.21±0.28	10
$o\text{-}PO_4^{3-}$ *(mg/L)*	0.055±0.041	1	0.037±0.016	0
Coliforms (*MPN/100 ml*)	5418.1±8236.1	34	2665.1±4484.9	4

A previous study in southern Vietnam reported average *pH* and *DO* values of 7.1 and 4.6 mg/L, respectively (Nguyen and Huynh 2022). In the Mekong Delta, *pH* was recorded as 7.29 ± 0.22, and the *DO* ranged from 7.15 ± 0.01 to 7.22 ± 0.11 (Hong and Giao 2022). Compared to these previous reports, the *pH* in this study was slightly lower, but the *DO* was higher. The cooler temperature in Dak Lak may have resulted in higher *DO* concentrations.

The average values of *COD* and *TSS* in southern Vietnam were 17.5 mg/L and 36 mg/L, respectively (Nguyen and Huynh 2022), which were higher than those found in both river and lake water in this study. These indicators are associated with biodegradable organic matter.

$N\text{-}NO_2^-$, $N\text{-}NO_3^-$, $N\text{-}NH_4^+$, and $o\text{-}PO_4^{3-}$ are the main nutrients found in water. Nutrient pollution in water in the Mekong Delta is mainly caused by wastewater from treatment systems and agricultural production (Hong and Giao 2022). In Son La hydropower reservoir, the concentrations of some indicators were high due to non-point sources or from the activities of domestic households, agriculture, tourism and aquaculture (Tran et al. 2024). These compounds can be converted by microorganisms. Therefore, the concentrations of these nutrients in the water depend on their sources and the activities of microorganisms.

Coliform bacteria in water are typically found in waste containing manure and due to the overuse of fertilizers. In previous studies, the average coliform density in the Mekong Delta ranged from 11,067 to 31,363 *MPN* 100 mg/L (Hong and Giao 2022) and even reached 2,500,000 *CFU* 100 mg/L (Wilbers et al. 2014). Comparatively, coliform levels in surface water in some provinces in the Mekong Delta are higher than those in Dak Lak province.

Rivers flow through high mountains, fields, residential areas, urban areas, and industrial zones where pollutants are washed away. Meanwhile, lakes and reservoirs are where rivers originate or receive water from canals or small rivers. This may be the reason that water in rivers had higher concentrations of the parameters. Pollutants in lakes have time to settle or decompose so the pollutant concentrations were lower than in the rivers.

3.2 Rivers' water parameters

The hydrochemical indicators are presented in Figure 2. Among the parameters, *pH* and *DO* were quite stable, while others greatly fluctuated at different times. All parameters had exceptionally low and high values shown as lower and upper outliers in the Figure. The exceptional lower outlier of *pH* was 5.0 recorded at the site NM 14 in August. The highest value was 8.1 mg/L found at the site NM 16 in August. Most evaluation times had values of *TSS* significantly higher than the average points in some river positions shown as upper outliers in the Figure 2. The average *TSS* concentrations were relatively low in March (13.6 mg/L) and June (22.4 mg/L). In April, the average concentrations of BOD_5 and *COD* were higher compared to other months. The highest BOD_5 and *COD* concentrations were found at the same position NM 13 in April, with 45.9 and 83.2 mg/L, respectively.

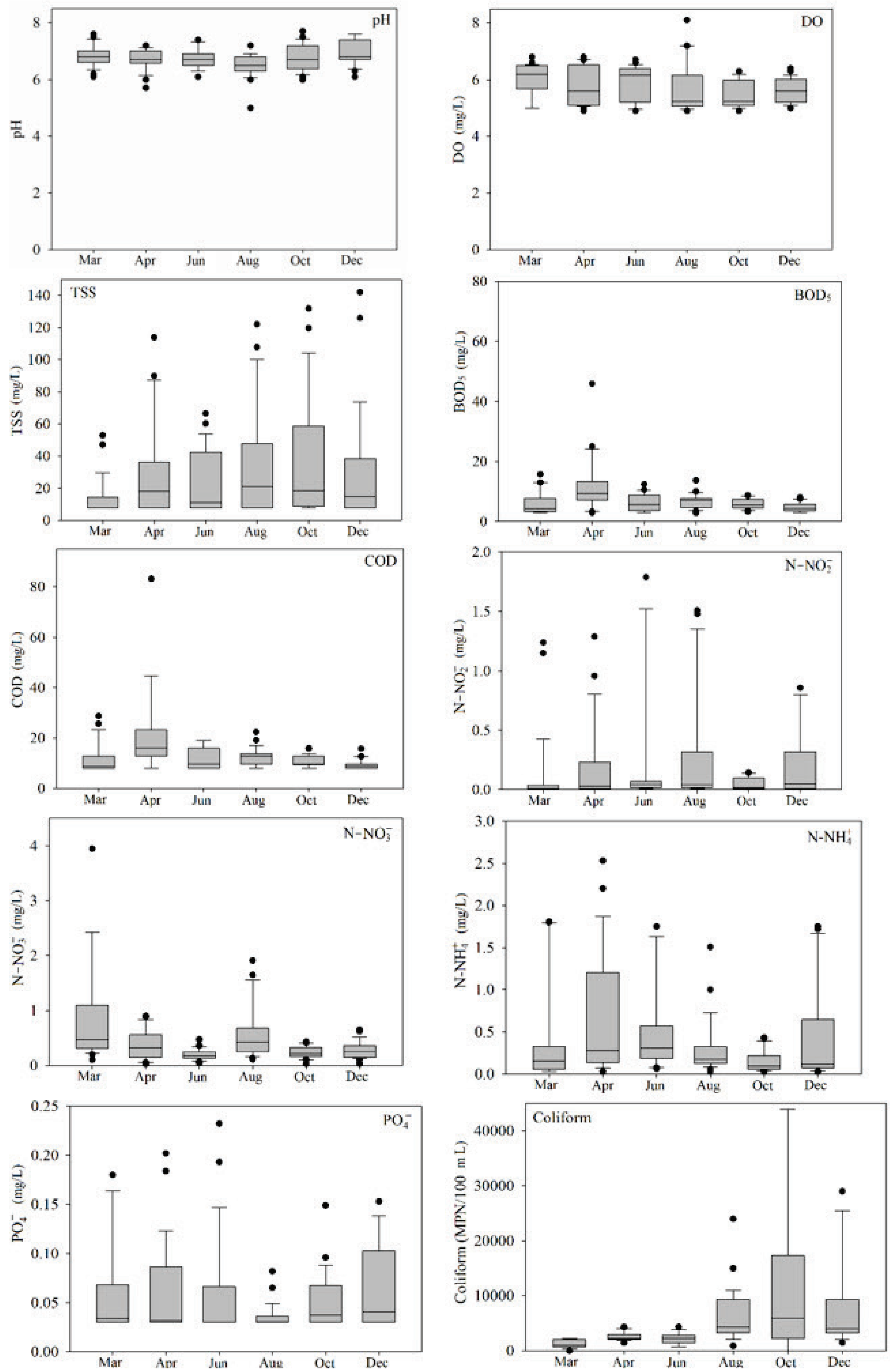

Figure 2 Hydrochemical and microbiological parameters in samples of river water.

For the components containing nitrogen, the highest concentrations of *N-*NO_2^- and *N-*NH_4^+ were found at same position NM 28 in June and April, respectively, with the corresponding concentrations of 1.79 and 2.53 mg/L. Meanwhile, highest concentration of

$N\text{-}NO_3^-$ was 6.65 mg/L at the position NM 5 in March. The $N\text{-}NO_3^-$ concentration was highest in March, followed by August. High concentrations of $N\text{-}NO_2^-$, were recorded in June and August. On the other hand, the highest concentration of $N\text{-}NH_4^+$ was found in April, averaging 0.69 mg/L. Meanwhile, low concentrations of both $N\text{-}NO_2^-$ and $N\text{-}NH_4^+$ were recorded in October. The average concentration of $o\text{-}PO_4^{3-}$ was 0.055 mg/L, and it did not exceed 0.232 mg/L. $o\text{-}PO_4^{3-}$ concentration was lowest in August. Coliforms were detected in 97.4% of river water samples, with an average density of 5,418.1 MPN/100 ml. The highest density was 46,000 MPN/100 ml found at the same position NM 12 in October. The densities of coliforms were low from March to June and higher in the later months.

Each box is drawn from its first to third quartile value, with a horizontal line drawn in the middle representing the median. Vertical lines (whiskers) extend from the least to greatest values excluding outliers. Upper and lower points are outliers which are significantly different from the rests beyond the whiskers.

3.3 Lakes' water parameters

The hydrochemical indicators in the lake water are presented in Figure 3. All parameters had exceptionally low and high values shown as lower and upper outliers in the figure. The *pH* values ranged from 6.0 to 8.1, which were not significantly different compared to river water. The *DO* data ranged from 4.9 to 6.8 mg/L. The average pH value was highest in March at 7.05. The average *DO* values remained relatively stable across all months. However, *DO* concentrations greatly fluctuated at different positions during March and April.

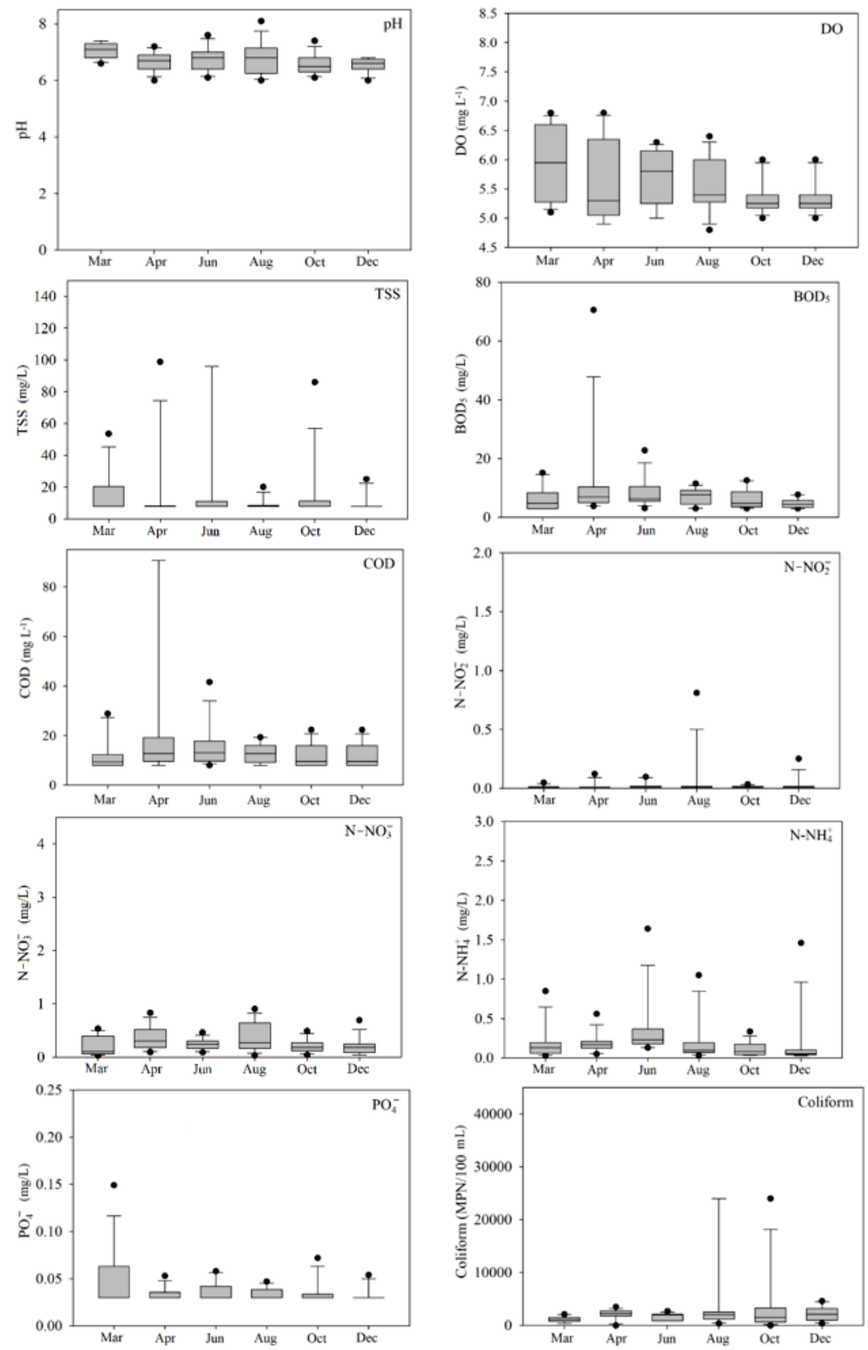

Figure 3 Hydrochemical and microbiological parameters in samples of lake water.

The highest concentrations of *TSS, BOD_5,* and *COD* were found to be 151, 70.6, and 134 mg/L, respectively. The *TSS* concentrations fluctuated significantly at different positions as indicated by high vertical lines and upper outliers in Figure 3 during March, April, June and October. The average *TSS* concentrations in the lake water were lowest in August and December. The lowest value for *BOD_5* was also found in December. Some samples collected in April exhibited exceptionally high values for these indicators. For example, the highest BOD_5 and *COD* concentrations were 70.6 and 134 mg/L, respectively, recorded in April.

$N\text{-}NO_2^-$ had the highest concentration found in August with 0.81 mg/L. $N\text{-}NH_4^+$ exhibited exceptionally high values in some samples in most evaluation times. The highest concentration of $N\text{-}NH_4^+$ was 1.64 mg/L in June. In August, the average concentrations of $N\text{-}NO_2^-$ and $N\text{-}NO_3^-$ were somewhat higher compared to other months. Meanwhile, $N\text{-}NH_4^+$ had a relatively low concentration in October and a high concentration in June.

The average concentration of $o\text{-}PO_4^{3-}$ in March was higher than in other months, and its exceptionally high value was observed during this time with 0.149 mg/L. Coliforms were detected in 97.4% of lake water samples, with the highest density being 24,000 MPN/100 ml. The upper outliers of coliform density were found in August and October resulting insignificantly higher density than in other months.

Each box is drawn from its first to third quartile value, with a horizontal line drawn in the middle representing the median. Vertical lines (whiskers) extend from the least to greatest values excluding outliers. Upper and lower points are outliers which are significantly different from the rests beyond the whiskers.

3.4 Identifying sources of water pollutants

The concentrations of water parameters in all river and lake positions are shown in Tables 5a and 5b, respectively. The *pH* and *DO* values of water from rivers and lakes were not statistically different among most river and lake positions, and consistent in both dry and rainy seasons. The river water had average *TSS* concentrations at positions 15, 19, and 20, higher than other monitoring points; however, lower average *TSS* concentrations were observed at positions 6, 16, and 26 (Table 5a). Meanwhile, the lake water had the highest *TSS* concentration at site 10 (Table 5b). In addition, *TSS* concentrations were higher during the rainy season at position 1 and 10 surrounded by or near forests. Moreover, its concentrations were significantly higher in the dry season at positions 3 and 8 when the mineral exploitation and crop cultivation were conducted in the corresponding areas. These findings indicate that *TSS* concentrations were higher in areas with forests, agriculture, and mineral exploitation, but lower in areas with wastewater treatment plants.

Table 5 (a) Parameters of water collected from rivers at different positions.

Sites	*pH*	*DO* (mg/L)	*TSS* (mg/L)	*BOD_5* (mg/L)	*COD* (mg/L)	*$N\text{-}NO_2^-$* (mg/L)	*$N\text{-}NO_3^-$* (mg/L)	*$N\text{-}NH_4^+$* (mg/L)	*PO_4^{3-}* (mg/L)	Coliform (MPN/100 ml)
1	6.9±0.3^{a}	5.6±0.5^{abcd}	26.4±21.2^{abcd}	7.5±2.7^{ab}	12.5±3.9^{ab}	0.043±0.022^{a}	0.27±0.24^{a}	0.67±0.82^{abcd}	0.055±0.023^{ab}	5200.0±3597.2^{ab}
2	6.6±0.5^{a}	5.6±0.4^{abc}	17.5±10.3^{ab}	5.5±1.7^{a}	10.1±2.1^{a}	0.055±0.029^{a}	0.25±0.25^{a}	0.85±0.79^{cd}	0.057±0.028^{ab}	7633.3±8436.5^{abc}
3	6.7±0.3^{a}	5.6±0.5^{abc}	12.7±8.5^{ab}	7.1±1.8^{ab}	12.8±2.9^{ab}	0.262±0.597^{ab}	0.32±0.20^{a}	0.80±0.48^{bcd}	0.032±0.002^{ab}	6250.0±8736.8^{abc}
4	7.2±6.9^{a}	6.5±5.1^{abc}	24.0±26.1^{abc}	8.5±4.2^{ab}	14.6±6.3^{ab}	0.021±0.018^{a}	1.37±2.59^{b}	0.11±0.12^{a}	0.033±0.004^{ab}	8283.3±8744.2^{abc}
5	6.7±0.3^{a}	5.6±0.7^{abc}	35.2±39.1^{abcde}	6.5±3.3^{ab}	12.2±4.6^{ab}	0.034±0.035^{a}	0.79±1.55^{ab}	0.13±0.09^{a}	0.032±0.004^{ab}	1701.7±1531.3^{a}
6	6.8±0.4^{a}	5.6±0.5^{abc}	8.6±1.0^{a}	8.8±7.8^{ab}	17.0±14.2^{ab}	0.780±0.584^{cd}	0.58±0.73^{ab}	1.16±0.85^{d}	0.131±0.066^{c}	1701.7±1531.3^{a}
7	6.8±0.5^{a}	6.3±0.7^{bcd}	17.1±9.9^{abc}	5.7±3.0^{ab}	11.0±4.0^{a}	0.070±0.110^{a}	0.36±0.23^{a}	0.34±0.40^{abc}	0.047±0.028^{ab}	4366.7±5323.0^{ab}
8	6.9±0.3^{a}	6.2±0.7^{abcd}	14.7±5.0^{abc}	5.7±3.6^{ab}	11.0±4.5^{a}	0.100±0.127^{a}	0.37±0.27^{a}	0.40±0.41^{abc}	0.041±0.022^{ab}	1983.3±608.0^{a}
9	6.8±0.6^{a}	5.7±0.7^{abcd}	40.5±33.6^{abcde}	6.2±1.8^{ab}	11.5±2.9^{ab}	0.072±0.089^{a}	0.28±0.15^{a}	0.42±0.49^{abc}	0.056±0.043^{ab}	5843.3±5478.2^{ab}
10	6.8±0.6^{a}	5.5±0.5^{abc}	21.5±31.8^{abc}	7.3±3.4^{ab}	13.3±5.4^{ab}	0.574±0.611^{bc}	0.70±0.75^{ab}	0.36±0.32^{abc}	0.068±0.021^{ab}	15910.0±18311.1^{cd}
11	6.9±0.6^{a}	5.4±0.4^{ab}	23.6±27.8^{abc}	12.4±16.6^{b}	22.7±29.8^{ab}	0.271±0.313^{ab}	0.69±0.70^{ab}	0.47±0.66a^{bc}	0.075±0.042^{b}	6766.7±4509.4^{abc}
12	6.3±0.9^{a}	5.4±0.6^{a}	29.4±30.6^{abcde}	6.7±3.2^{ab}	12.8±4.8^{ab}	0.751±0.655^{cd}	0.64±0.47^{ab}	0.73±0.59^{abcd}	0.112±0.051^{c}	13300.0±18204.4^{bcd}
13	6.9±0.3^{a}	5.7±0.5^{abcd}	43.6±44.7^{abcde}	11.0±7.4^{ab}	19.2±13.4^{ab}	0.012±0.003^{a}	0.24±0.11^{a}	0.15±0.12^{ab}	0.049±0.021^{ab}	2328.3±1511.3^{a}
14	6.8±0.3^{a}	6.5±1.0^{d}	12.6±4.9^{ab}	5.4±1.4^{a}	10.2±2.2^{a}	0.067±0.091^{a}	0.35±0.16^{a}	0.36±0.46^{abc}	0.051±0.017^{ab}	3816.7±2947.8^{ab}
15	6.7±0.5^{a}	5.7±0.7^{abcd}	65.2±59.2^{e}	5.3±2.1^{a}	10.3±3.2^{a}	0.018±0.010^{a}	0.38±0.21^{a}	0.12±0.08^{a}	0.035±0.006^{ab}	4816.7±4357.3^{ab}
16	6.7±0.3^{a}	5.7±0.7^{abcd}	8.2±0.6^{a}	5.1±2.3^{a}	10.1±3.0^{a}	0.010±0.000^{a}	0.22±0.17^{a}	0.16±0.10^{ab}	0.030±0.000^{a}	1183.3±854.2^{ab}
17	6.9±0.5^{a}	6.4±0.8^{cd}	18.5±11.5^{abc}	4.6±1.6^{a}	9.4±1.9^{a}	0.178±0.205^{a}	0.41±0.25^{a}	0.42±0.37^{abc}	0.058±0.029^{ab}	2416.7±1149.6^{a}
18	6.8±0.4^{a}	5.9±0.7^{abcd}	39.5±31.3^{abcde}	8.0±3.3^{ab}	14.6±5.9^{ab}	0.018±0.012^{a}	0.30±0.14^{a}	0.17±0.08^{ab}	0.043±0.019^{ab}	3385.0±2927.1^{ab}
19	6.8±0.2^{a}	5.5±0.7^{abc}	62.3±34.7^{de}	5.1±1.3^{ab}	9.4±1.8^{a}	0.051±0.048^{a}	0.27±0.15^{a}	0.14±0.09^{ab}	0.036±0.011^{ab}	3300.0±940.2^{ab}
20	6.9±0.2^{a}	5.9±0.7^{abcd}	47.3±33.2^{bcde}	9.2±3.8^{ab}	16.5±6.9^{ab}	0.056±0.067^{a}	0.28±0.15^{a}	0.16±0.13^{ab}	0.045±0.017^{ab}	3050.0±1413.9^{ab}
21	6.7±0.3^{a}	5.6±0.5^{abc}	52.3±35.8^{cde}	6.3±4.1^{ab}	12.0±6.9^{ab}	0.036±0.027^{a}	0.32±0.13^{a}	0.18±0.10^{ab}	0.032±0.004^{ab}	3700.0±748.3^{ab}
22	6.4±0.3^{a}	5.4±0.6^{ab}	20.3±14.7^{abc}	6.7±2.7^{ab}	12.6±3.8^{ab}	0.019±0.017^{a}	0.29±0.12^{a}	0.33±0.61^{abc}	0.030±0.001^{a}	2416.7±1121.5^{a}
23	6.8±0.3^{a}	6.0±0.5^{ab}	11.0±4.9^{ab}	7.7±4.1^{ab}	13.8±7.5^{ab}	0.041±0.041^{a}	0.42±0.42^{a}	0.15±0.10^{ab}	0.046±0.040^{ab}	4066.7±2915.9^{ab}
24	6.7±0.4^{a}	6.0±0.5^{abcd}	10.6±4.9^{ab}	5.8±3.7^{ab}	11.7±6.8^{ab}	0.037±0.033^{a}	0.43±0.41^{a}	0.14±0.10^{a}	0.045±0.025^{ab}	3035.0±1385.0^{ab}
25	6.6±0.4^{a}	5.4±0.7^{abcd}	42.7±41.3^{abcde}	6.9±3.9^{ab}	13.1±6.1^{ab}	0.112±0.220^{a}	0.30±0.17^{a}	0.42±0.55^{abc}	0.069±0.047^{ab}	4233.3±3141.1^{ab}
26	6.9±0.4^{a}	5.7±0.5^{abcd}	9.1±1.8^{ab}	5.0±1.6^{ab}	9.6±1.8^{a}	1.034±0.554^{d}	0.67±0.47^{ab}	1.22±0.93^{d}	0.136±0.075^{c}	20183.3±20605.0^{d}

NOTE: The values are expressed as mean ± variance. The different lowercase superscript letters (a, b, d, and e) show statistically significant differences among data within a column ($p < 0.05$, ANOVA, Duncan's multiple range test).

Table 5 (b) Parameters of water collected from lakes at different positions.

Sites	*pH*	*DO* (*mg/L*)	*TSS* (*mg/L*)	*BOD$_5$* (*mg/L*)	*COD* (*mg/L*)	*N-NO$_2^-$* (*mg/L*)	*N-NO$_3^-$* (*mg/L*)	*N-NH$_4^+$* (*mg/L*)	*PO$_4^{3-}$* (*mg/L*)	Coliform (*MPN/100 ml*)
1	6.6±0.3^{a}	5.5±0.6^{a}	10.9±5.0^{a}	6.4±2.2^{a}	12.0±3.4^{a}	0.032±0.045^{a}	0.26±0.15^{a}	0.17±0.15^{ab}	0.033±0.004^{a}	5876.7±8939.7^{a}
2	6.7±0.2^{a}	5.6±0.5^{a}	9.4±2.3^{a}	5.5±1.8^{a}	10.9±3.0^{a}	0.013±0.005^{a}	0.18±0.15^{a}	0.16±0.08^{ab}	0.032±0.004^{a}	5095.0±9284.3^{a}
3	6.6±0.5^{a}	5.5±0.5^{a}	16.8±12.6^{a}	8.5±3.7^{ab}	14.9±6.0^{a}	0.152±0.323^{b}	0.31±0.29^{a}	0.79±0.42^{c}	0.040±0.011^{a}	6300.0±8724.2^{a}
4	6.6±0.4^{a}	5.3±0.5^{a}	9.7±2.5^{a}	4.9±1.1^{a}	9.2±0.9^{a}	0.025±0.030^{a}	0.37±0.20^{a}	0.43±0.62^{b}	0.033±0.008^{a}	1660.0±931.2^{a}
5	6.8±0.4^{a}	5.5±0.7^{a}	8.0±0.0^{a}	9.8±7.0^{ab}	17.8±12.3^{ab}	0.056±0.097^{ab}	0.23±0.16^{a}	0.15±0.07^{ab}	0.034±0.010^{a}	2945.0±3429.3^{a}
6	7.0±0.4^{a}	5.9±0.6^{a}	10.1±4.1^{a}	7.2±2.7^{a}	12.7±4.9^{a}	0.011±0.002^{a}	0.28±0.21^{a}	0.11±0.07^{a}	0.035±0.007^{a}	2351.7±1403.3^{a}
7	6.8±0.4^{a}	5.7±0.5^{a}	8.0±0.0^{a}	5.8±2.9^{a}	11.5±4.9^{a}	0.019±0.012^{a}	0.19±0.08^{a}	0.13±0.08^{ab}	0.040±0.016^{a}	1361.7±718.7^{a}
8	6.9±0.7^{a}	5.5±0.6^{a}	15.9±10.7^{a}	6.1±4.4^{a}	12.6±7.1^{a}	0.012±0.005^{a}	0.35±0.25^{a}	0.16±0.13^{ab}	0.031±0.003^{a}	645.0±619.9^{a}
9	6.7±0.4^{a}	5.6±0.6^{a}	8.0±0.0^{a}	6.1±4.3^{a}	12.3±7.3^{a}	0.010±0.000^{a}	0.30±0.31^{a}	0.09±0.05^{a}	0.050±0.049^{a}	1038.3±530.2^{a}
10	6.7±0.5^{a}	5.7±0.4^{a}	67.6±55.8^{b}	17.8±26.0^{b}	33.2±49.5^{b}	0.026±0.035^{a}	0.20±0.13^{a}	0.14±0.06^{ab}	0.036±0.009^{a}	2150.0±550.5^{a}
11	6.9±0.2^{a}	5.7±0.5^{a}	8.0±0.0^{a}	5.4±2.4^{a}	10.4±3.3^{a}	0.014±0.006^{a}	0.20±0.18^{a}	0.11±0.06^{a}	0.038±0.013^{a}	1901.7±638.8^{a}
12	6.6±0.3^{a}	5.6±0.6^{a}	8.0±0.0^{a}	6.8±4.5^{a}	13.4±7.9^{a}	0.010±0.001^{a}	0.34±0.31^{a}	0.11±0.06^{a}	0.043±0.021^{a}	1685.0±605.4^{a}
13	6.7±0.4^{a}	5.7±0.7^{a}	8.0±0.0^{a}	6.7±3.5^{a}	12.8±5.6^{a}	0.010±0.000^{a}	0.28±0.10^{a}	0.14±0.11^{ab}	0.030±0.000^{a}	1636.7±623.6^{a}

NOTE: The values are expressed as mean ± variance. The different lowercase superscript letters (a, b, d, and e) show statistically significant differences among data within a column ($p < 0.05$, ANOVA, Duncan's multiple range test).

BOD_5 and *COD* concentrations were not statistically different in most positions in rivers and lakes. However, the average concentrations in some positions were higher compared to the average concentrations of all sites. The average BOD_5 and *COD* concentrations in river water were higher at positions 11 and 13, especially in the dry season. In lake water, the average BOD_5 and *COD* concentrations were highest at position 10 (Table 5b). The BOD_5 concentrations seemed to be directly proportional to the *COD* concentrations. These positions were surrounded by areas of high population density or agricultural land. This suggests that these pollutants mostly originated from domestic or agriculture processes. The elevated BOD_5 levels indicate a high amount of organic matter in certain areas of rivers and lakes. Mohamed et al. (2015) reported that the pollution from both biodegradable organic and inorganic pollutants led to high *COD* values in water.

For nitrogen-containing compounds in river water, average $N\text{-}NO_2^-$ concentrations were higher than 0.5 mg/L found at positions 6, 12, and 26, and smaller than 0.02 mg/L at positions 13, 15, 16, 18, and 22 (Table 5a). Position 4 recorded a high average $N\text{-}NO_3^-$ concentration. Meanwhile, $N\text{-}NH_4^+$ concentrations at positions 6 and 26 were significantly higher than other positions, and low average concentration at positions 4, 5, and 15 (smaller than 0.14 mg/L) (Table 5a). For lake water, $N\text{-}NO_3^-$ and $N\text{-}NH_4^+$ had the highest average concentrations at the same position 3 (Table 5b). The concentrations of $N\text{-}NO_3^-$ in lake water were not statistically different among positions. The high concentrations of compounds containing nitrogen were mostly found at sites surrounded by high population density, industrial zones, and wastewater treatment; meanwhile, their concentrations were low at sites with agricultural activity and low population density. The results proved that $N\text{-}NO_3^-$, $N\text{-}NO_2^-$, and $N\text{-}NH_4^+$ were mostly produced from human activities. Wastewater from households containing nutrients was degraded into nitrogen-containing compounds. These nitrogen compounds might be also derived from the decomposition of fertilizers used in the agriculture sector (Giao et al. 2021), but this phenomenon was observed in this study. $N\text{-}NO_3^-$ is the product of the aerobic decomposition of the organic nitrogenous compound (Mitra et al. 2018).

The water at positions 6, 12, and 26 in the river contained significantly higher levels of $o\text{-}PO_4^{3-}$ (exceeding 0.1 mg/L) than other positions. However, there was no significant difference in $o\text{-}PO_4^{3-}$ concentration among sites in all lakes. Elevated $o\text{-}PO_4^{3-}$ levels were detected in other river sites near areas with high population density and wastewater treatment facilities. A study in Vietnam also reported the presence of $o\text{-}PO_4^{3-}$ in water bodies because of effluent from residential areas (Pham et al. 2022). Additionally, the use of household detergent may contribute to the presence of $o\text{-}PO_4^{3-}$ in water bodies. The worldwide increase in nutrient concentration in some river and lake sites is attributed to the anthropogenic loading of nitrogen and phosphorus from industrial, municipal, and agricultural sources (Fianko et al. 2010).

For coliforms in river water, high densities were found at positions 10, 12, and 26, while positions 5, 6, and 16 had low average densities. Along the Ea Drueh River, upstream of a wastewater treatment plant (position 6), had low coliform density, but the outlet of the wastewater treatment (position 26) had a high coliform density. Additionally, water at position 26 had high concentrations of some pollutants. This result indicates that the wastewater treatment did not effectively remove contaminants. In lakes, the densities of coliforms were not statistically different among positions; however, higher average densities were found at sites 1, 2, and 3. In general, coliforms with high density were found in water from areas with sparse populations, forests, and agriculture except at site 8 of the lake, which is surrounded by agriculture. The average coliform density in both river and lake water was higher during the rainy season. These results suggest that coliforms might come from domestic activities and animal waste. Coliforms also naturally grow and develop in water.

3.5 Water quality assessment

Local people often use water from rivers and lakes for domestic and agricultural purposes. In addition, surface water from these locations is used for drinking after processing, even used for cooking and drinking without processing. The Water Quality Index (WQI) assesses the quality of water for multiple uses. The WQI values for river and lake water sources were calculated based on various indicators. The WQI values showed significant fluctuations depending on the collection time and sites. The average WQI value for river water was 81.6 ± 22.3, ranged from bad to very good quality. Meanwhile, the data for lake water stood at 92.0 ± 13.7 and varied from average to very good quality for domestic water supply purposes (Table 6).

Table 6 WQI values in river and lake water.

River sites	WQI of river water			Lake sites	WQI of lake water		
	Rainy season	Dry season	Mean		Rainy season	Dry season	Mean
1	66.1±20.9^{abcd}	84.2±11.1^{a}	75.2±18.0^{abcde}	1	72.6±36.7^{a}	95.3±5.6^{bc}	83.9±26.5^{ab}
2	68.2±32.5^{abcd}	78.3±23.3^{a}	73.3±25.9^{abcde}	2	74.8±39.1^{a}	99.2±0.7^{c}	87.0±28.1^{ab}
3	70.1±35.8^{abcd}	85.7±4.0^{a}	77.9±24.3^{abcde}	3	67.0±31.8^{a}	85.7±5.0^{a}	76.3±22.8^{a}
4	59.4±33.1^{abcd}	71.7±36.2^{a}	65.6±31.7^{abcd}	4	94.5±5.0^{a}	98.7±1.9^{bc}	96.2±4.3^{b}
5	93.5±7.7^{d}	96.6±3.6^{a}	95.1±5.6^{e}	5	83.3±12.7^{a}	94.4±7.9^{abc}	88.9±11.2^{ab}
6	86.8±7.4^{bcd}	85.5±14.2^{a}	86.2±10.1^{cde}	6	92.9±8.6^{a}	94.4±3.6^{abc}	93.7±6.0^{ab}
7	72.6±35.5^{abcd}	93.2±3.6^{a}	82.9±25.2^{bcde}	7	95.7±3.8^{a}	99.1±1.5^{c}	97.4±3.2^{b}
8	94.0±5.2^{d}	96.4±4.0^{a}	95.2±4.4^{e}	8	97.5±3.2^{a}	95.9±4.8^{bc}	96.7±3.8^{b}
9	68.5±34.2^{abcd}	79.3±24.5^{a}	73.9±27.3^{abcde}	9	97.9±3.6^{a}	96.4±6.1^{bc}	97.1±4.6^{b}
10	47.9±31.0^{abc}	73.3±37.6^{a}	60.6±33.8^{abc}	10	96.9±3.0^{a}	89.6±8.2^{ab}	93.2±6.8^{ab}
11	60.6±31.3^{abcd}	63.6±33.3^{a}	62.0±28.9^{abc}	11	96.2±2.7^{a}	99.5±0.9^{c}	97.8±2.6^{b}
12	40.1±38.3^{a}	64.6±31.1^{a}	52.3±34.1^{a}	12	98.5±2.6^{a}	94.2±6.1^{abc}	96.3±4.8^{b}
13	91.6±3.5^{cd}	89.1±5.2^{a}	90.4±4.2^{de}	13	94.2±5.0^{a}	98.1±2.3^{bc}	96.1±4.1^{b}
14	81.2±23.6^{abcd}	91.2±8.6^{a}	86.2±16.8^{cde}				
15	69.8±25.8^{abcd}	88.8±19.3^{a}	79.3±22.9^{bcde}				
16	99.3±0.7^{d}	97.5±3.4^{a}	98.4±2.4^{e}				
17	95.7±3.3^{d}	92.4±8.3^{a}	94.0±5.9^{e}				
18	83.8±24.4^{abcd}	91.7±5.4^{a}	87.8±16.4^{de}				
19	93.1±4.7^{d}	94.9±5.6^{a}	94.0±4.7^{e}				
20	88.4±4.6^{bcd}	90.3±2.0^{a}	89.4±3.3^{de}				
21	90.5±4.0^{bcd}	90.7±9.0^{a}	90.6±6.3^{de}				
22	90.5±4.3^{bcd}	95.2±4.2^{a}	92.9±4.6^{e}				
23	80.7±22.1^{abcd}	89.5±6.0^{a}	85.1±15.3^{cde}				
24	92.8±4.9^{cd}	93.8±3.7^{a}	93.3±3.9^{e}				
25	74.5±15.9^{abcd}	89.7±9.0^{a}	82.1±14.2^{bcde}				
26	45.9±27.0^{ab}	67.9±33.1^{a}	56.9±29.6^{ab}				

NOTE: The lowercase superscript letters show statistically significant differences among data within a column ($p < 0.05$).

For river water, the average WQI values at positions 5, 8, 13, 16, 17, 19, 21, 22, and 24 were higher than 90 (Table 6), indicating very good water quality. However, the average WQI values were lower at position 12, followed by position 26. The population density at position 12 and outlet of a wastewater treatment had the WQIs statistically lower than other positions including positions 5, 8, 16, 17, 19, and 24 which were sparsely populated or near the Phu Hoa industrial zone. The significant differences of WQIs were found in the rainy season, while the values were not statistically different in the dry season.

In river positions where the average WQI was lower than 90, the average WQI values in the rainy season were higher than in the dry season, except for positions 6 and 12. Only position

22, which was surrounded by a high population density, had very good water quality. Most other sites with WQI values lower than 90 were surrounded by residential areas and industrial zones.

Along the Serepok River, the average WQI of water in the hydroelectric reservoir (position 7) was smaller than 90. However, downstream of the reservoir (position 8) showed very good water quality. The self-purification and inaccessibility resulted in water quality improvement in some segments. Position 15, surrounded by forest and rural settlements, had a WQI higher than 90. Positions 14 and 17, along the Serepok River and surrounded by industrial zones, had an average WQI of 86.2±16.8 and 94.0±5.9, respectively.

Along the Krong Buk River, the water quality at river positions 16 and 22, as well as lake position 12, was consistently very good, despite position 22 being surrounded by a high population density. On the Krong No River, the upstream position (position 18) is surrounded by agricultural sites and has a sparse population, leading to significant fluctuation in the WQI. In contrast, the downstream position (position 19) is surrounded by agricultural sites, used for mineral exploitation and aquaculture, and consistently showed very good water quality. Meanwhile, along the Krong Ana River, the water quality did not vary significantly across all evaluation sites. The downstream outlet of the cassava starch factory had a higher average WQI value than the upstream, suggesting that the factory did not cause environmental pollution.

For lake water, the average WQI at position 3 was significantly lower compared to positions 4, 7, 8, 9, 11, 12, and 13 (Table 6). The result showed that the Tan An industrial zone at the position 3 caused water pollution. No statistical difference of WQI values was found in the rainy season, while the values among several positions in the dry season were significantly different. The result showed that the effects of seasons on WQIs in the river and lake water were different.

In some lake areas, the WQI was below 90 at positions 1, 2, 3, and 5, while other areas had very good water quality. Generally, lakes had a higher water quality than rivers, possibly because most lakes are not surrounded by densely populated areas. Site 3 had a WQI below 90 in both seasons, suggesting that industrial activities nearby were causing water pollution. Site 5, located near a sugarcane factory and a densely populated area, also had a WQI below 90, while sites 6 and 7, surrounded by sparsely populated areas, had very good water quality. These findings indicate that both natural processes and human activities impact the water quality of both rivers and lakes.

Most river and lake sites had WQI values lower than 90 due to high levels of BOD_5 and coliforms. In particular, sites 12 and 26 on the river had a low WQI due to high concentrations of COD and $N\text{-}NH_4^+$ during the rainy season. These pollutants mainly originated from domestic and industrial sources. Most river locations and all lake sites with an average WQI lower than 90 showed decreased WQI values during the rainy season compared to the dry season (see Table 6). Water runoff with high contaminant levels contributed to the reduction in water quality during the rainy season. The study indicated that water quality was influenced by both natural and anthropogenic factors.

3.6 Relationship between water indicators

The Pearson correlation analysis was used to evaluate relationships between two variables. Tables 7a and 7b showed that correlations among most parameters were weak. The *pH* and *DO* indicate a moderate correlation in both river and lake water (Table 7). *TSS* weakly correlated with BOD_5 and *COD* in river water but had moderate correlation in lake water.

Table 7 (a) Pearson correlation coefficient among the parameters of river water.

	pH	*DO*	*TSS*	BOD_5	*COD*	$N\text{-}NO_2^-$	$N\text{-}NO_3^-$	$N\text{-}NH_4^+$	PO_4^{3-}	Coliform
pH	1									
DO	0.40	1								
TSS	-0.21	-0.17	1							
BOD_5	-0.15	-0.08	0.12	1						
COD	-0.15	-0.07	0.11	0.99	1					
$N\text{-}NO_2^-$	-0.10	-0.19	-0.20	0.15	0.17	1				
$N\text{-}NO_3^-$	-0.03	0.09	-0.11	0.02	0.04	0.27	1			
$N\text{-}NH_4^+$	0.09	-0.02	-0.23	0.12	0.14	0.50	-0.03	1		
$o\text{-}PO_4^{3-}$	0.01	-0.09	-0.12	0.03	0.05	0.45	-0.02	0.54	1	
Coliform	0.19	-0.03	-0.03	-0.09	-0.09	0.12	-0.06	-0.02	0.10	1

NOTE: The correlation coefficients include positively strong (darker green), positively moderate (moderate green), positively weak (weaker green), negatively weak (weak yellow), and no correlation (no color).

Table 7 (b) Pearson correlation coefficient among the parameters of lake water.

	pH	*DO*	*TSS*	*BOD_5*	*COD*	*$N\text{-}NO_2^-$*	*$N\text{-}NO_3^-$*	*$N\text{-}NH_4^+$*	*PO_4^{3-}*	Coliform
pH	1									
DO	0.54	1								
TSS	-0.03	0.05	1							
BOD_5	-0.04	0.13	0.38	1						
COD	-0.05	0.13	0.39	0.99	1					
$N\text{-}NO_2^-$	-0.17	-0.09	0.03	0.05	0.03	1				
$N\text{-}NO_3^-$	0.00	0.02	-0.11	-0.05	-0.06	0.17	1			
$N\text{-}NH_4^+$	-0.10	-0.17	0.00	0.02	0.00	0.33	0.13	1		
$o\text{-}PO_4^{3-}$	0.02	0.17	0.02	0.18	0.18	0.07	-0.18	-0.07	1	
Coliform	0.08	0.19	0.03	0.02	0.01	0.03	-0.10	0.06	-0.03	1

NOTE: The correlation coefficients include positively strong (darker green), positively moderate (moderate green), positively weak (weaker green), negatively weak (weak yellow), and no correlation (no color).

$N\text{-}NO_2^-$, $N\text{-}NH_4^+$, and PO_4^{3-} had a moderately positive correlation in river water, and $N\text{-}NO_2^-$ showed a positive moderate correlation with $N\text{-}NH_4^+$ in lake water. PO_4^{3-} moderately correlated with $N\text{-}NH_4^+$ in river water but had weak correlation in lake water. The nitrogen and phosphorus compounds are considered nutrients and important factors for the growth of aquatic organisms. The positive correlation was likely because they were co-produced by human activities as described above. BOD_5 and *COD* had strong correlation in both water sources. The strong correlation between *BOD* and *COD* generally indicates a high proportion of biodegradable organic matter in water (Majeb et al. 2023).

3.7 Principal component analysis (PCA)

The PCA method was used to identify the indicators that affect surface water quality and to determine possible pollution sources. Principal components (PCs) with eigenvalues higher than 1 are the main components. Four major PCs were identified, explaining 69.2% of the variation, while five PCs had eigenvalues greater than 1, explaining 75.2% of the variation (Table 8).

Table 8 PCA for parameters of river and lake water.

Variable	WQI of river water				WQI of lake water				
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC5
pH	0.17	0.30	0.52	-0.27	0.03	-0.56	-0.34	0.14	-0.05
DO	0.19	0.14	0.64	0.06	0.18	-0.54	-0.38	0.02	-0.12
TSS	0.03	-0.39	-0.26	-0.21	0.38	0.08	0.07	0.06	0.32
BOD_5	-0.48	-0.38	0.31	-0.11	0.62	0.11	-0.01	0.13	0.00
COD	-0.49	-0.37	0.31	-0.09	0.62	0.11	0.01	0.14	0.00
$N\text{-}NO_2^-$	-0.43	0.33	-0.16	0.19	0.02	0.39	-0.47	-0.28	-0.23
$N\text{-}NO_3^-$	-0.09	0.08	0.12	0.76	-0.11	0.16	-0.46	0.58	-0.18
$N\text{-}NH_4^+$	-0.39	0.39	0.01	-0.13	-0.03	0.39	-0.46	-0.19	0.07
$o\text{-}PO_4^{3-}$	-0.34	0.39	-0.16	-0.18	0.21	-0.12	0.08	-0.52	-0.65
Coliforms	0.03	0.20	-0.04	-0.45	0.05	-0.16	-0.30	-0.47	0.61
Eigenvalues	2.4	2.0	1.4	1.2	2.34	1.76	1.31	1.08	1.03
% Variation	23.5	19.9	14.2	11.6	23.4	17.6	13.1	10.8	10.3
Cum% Variation	23.5	43.3	57.5	69.1	23.4	41	54.1	64.9	75.2

NOTE: The correlation coefficients include positively strong (dark green), positively and negative moderate (moderate green), positively and negatively weak (weak green), and no correlation (no color).

The analysis of water collected from rivers revealed the following findings: PC1 showed the highest variation (23.5%) and exhibited weak negative correlations with BOD_5, *COD*, $N\text{-}NO_2^-$, $N\text{-}NH_4^+$, and $o\text{-}PO_4^{3-}$. PC2 had a weak positive correlation with *pH*, $N\text{-}NH_4^+$, and $o\text{-}PO_4^{3-}$, as well as weak negative correlations with *TSS*, BOD_5, and *COD*. PC3 showed average positive correlations with *pH* and *DO*, and weak positive correlations with BOD_5 and *COD*. Lastly, PC4 demonstrated a strong correlation with $N\text{-}NO_3^-$, and a negative average correlation with coliforms.

As for water from the lakes, *pH* and *DO* were affected by PC2 and PC3 with average and weak negative correlations, respectively. Meanwhile, both BOD_5 and *COD* were affected by PC1 with average positive correlations. *TSS* was affected by PC1 and PC5 with weak positive correlations. Both $N\text{-}NO_2^-$ and $N\text{-}NH_4^+$ were affected by PC2 and PC3 with average correlations. $N\text{-}NO_3^-$ was affected by PC3 and PC4 with weak negative and average positive correlations, respectively. $o\text{-}PO_4^{3-}$ was affected by both PC4 and PC5 with average negative correlations. Coliform was affected by PC3 and PC5 with weak negative and average positive correlations, respectively.

All the parameters indicate that organic and inorganic matters/compounds could be attributed to various anthropogenic activities and natural processes or mineral exploitation as described above. The components were produced from different places along the rivers and lakes. The majority of the correlations were weak, indicating predominantly weak impacts of the parameters. Some data showed moderate correlations, while only a strong correlation was found.

3.8 Clustering water quality of the river and lake water

River and lake water can be grouped into clusters. The numbers of clusters depend on Euclid distances. Each main cluster contains several sub-clusters or groups. When considering river water, four clusters were identified at a Euclidean distance of about 3.2, while ten clusters were identified at a Euclidean distance of about 2.2 (Figure 4a).

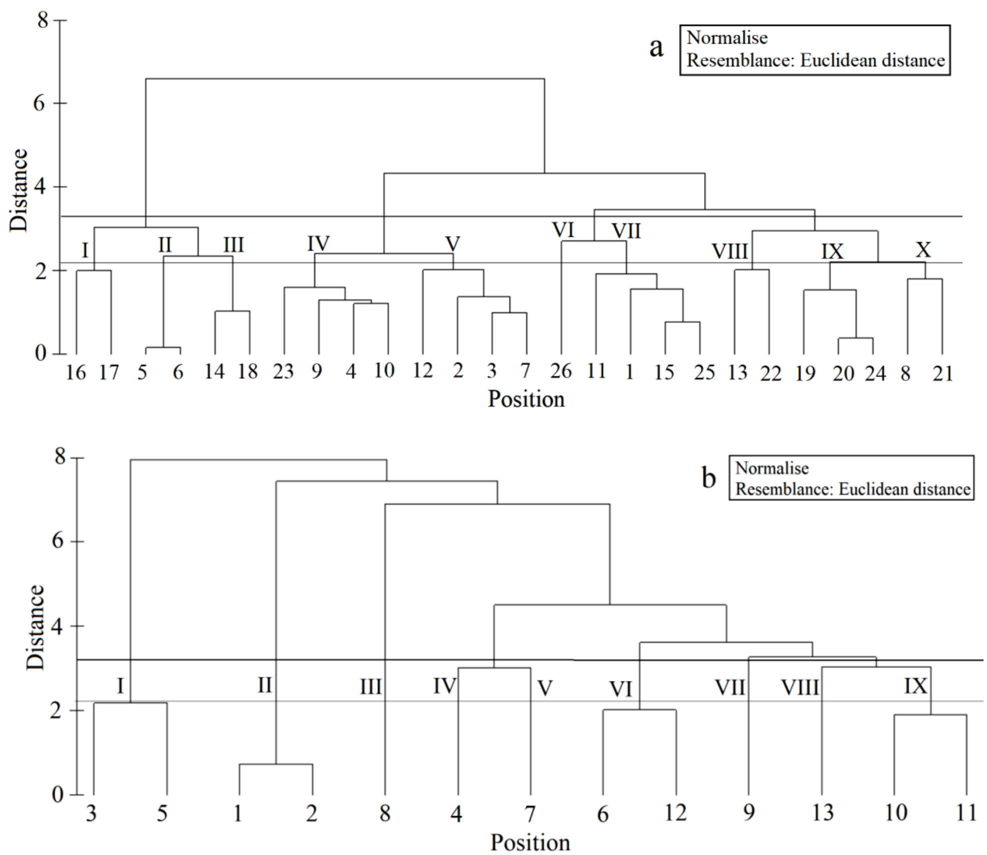

Figure 4 Clustering of water quality in (a) river water, and (b) lake water in Dak Lak Province.

For analysis of all indicators among ten clusters in Figure 4a, cluster I included positions 16 and 17. The average concentrations of $N\text{-}NO_3^-$, $N\text{-}NO_2^-$, $N\text{-}NH_4^+$, and $o\text{-}PO_4^{3-}$ and coliforms were smaller or smallest compared to other positions (Table 5a). The WQI of positions 16 and 17 was higher than 90 (Table 6). Meanwhile, cluster II included positions 5 and 6, which had no statistically significant values of *pH* and *DO*, and the same coliform density. The average density of coliform at positions 5 and 6 were quite low. Cluster III, comprised of positions 14 and 18, witnessed a great fluctuation in the concentration of most indicators, and the average WQI was smaller than 90. Cluster IV included sites 23, 9, 4, and 10, which had the high average concentration of *N-*NH_4^+ and coliforms. WQI data showed that the water quality at positions 9, 4, and 10 ranged from bad to average quality, while the quality at position 23 was good. Cluster V was composed of positions 12, 2, 3, and 7. Positions 2, 3, and 7 showed higher average concentrations of *N-*NH_4^+ and coliforms than the average of all positions except position 7. All positions in this cluster showed WQI values ranging from bad to average quality, except at WQI of position 7 in the dry season. Cluster VI had only position 26, showing the highest average concentration of *N-*NO_2^-, *N-*NH_4^+, and *o-*PO_4^{3-}, and coliform density. Moreover, water quality at this position ranged from bad to average. Cluster VII included positions 11, 1, 15, and 25, which had average WQI values showing average water quality in the rainy season. In addition, water at these positions ranged from average to good quality for domestic purposes. Cluster VIII, comprised of positions 13 and 22, witnessed very good water quality, and the highest concentration of BOD_5 and *COD*, while other parameters were at lower concentrations compared to the average of all sites. Cluster IX was composed of positions 19, 20, and 24. Positions 19 and 20 had quite high *TSS* concentrations, and most other parameters had concentrations lower than the average of all sites. The last cluster had two positions, 8 and 11, which had very good water quality.

For lake water, at a distance Euclid of about 3.2, seven clusters were identified; while at a distance Euclid of about 2.2, nine clusters were divided (Figure 4). Clusters I and II had good water quality, while the others showed very good quality. Cluster I includes positions 3 and 5. Position 3 had the highest *N-*NH_4^+ concentration, and coliform density was higher than the average of all positions. Meanwhile, average concentrations of BOD_5, *COD*, and *N-*NO_2^- at position 5 were higher than the average of all positions. Cluster II, comprising of positions 1 and 2, had a higher average coliform density, even though most other parameters had average concentrations lower than the average of all positions. Only position 8 was in cluster III, which had the lowest average coliform density, and very good water quality for domestic supply. Clusters IV and V contained one position each—4 and 7, respectively, in which most parameters were within the permissible limits. The parameter values at positions 4 and 7; however, position 4 had a higher average concentration of *N-*NO_3^- and *N-*NH_4^+, but a lower concentration of *COD* than position 7. Clusters VI (position 6 and 12), VII (position 9), and VIII (position 13) had different average values of *N-*NH_4^+, *o-*PO_4^{3-} and coliforms. Cluster IX includes positions 10 and 11. Position 10 had the higher average concentration of *TSS*, BOD_5, and *COD* than the latter, but both positions had very good water quality on average.

3.9 Limitations and potential practical applications

The study analyzed surface water quality and identified pollutant sources. Nutrients were mostly produced from anthropogenic factors, and some pollutants originated from natural processes. The limitation of this study is that it only evaluated Dak Lak province over one year. Deep investigation of certain pollutant sources as well as water quality in a long-term period is still ongoing.

However, the results shed light on positions of water pollution and the sources of water pollutants in Dak Lak province. Implementing environmentally friendly practices and treating wastewater are crucial for reducing surface water pollution. While Vietnamese environmental protection laws are in place, the management framework still needs improvement. Integrated water resource management should be used to strengthen institutions, policies, and legislation to control water pollution.

4. CONCLUSION

This study compared the water quality of rivers and lakes in Dak Lak province, Vietnam. The results indicated that the water quality in rivers and lakes varied greatly depending on the location and time of collection. The WQI values for river and lake water were 81.6±22.3 and 92.0±13.7, respectively. Water quality was adversely affected by human activities and natural factors. BOD_5 and *COD* mostly originated from domestic or waste treatment processes, while compounds containing nitrogen and phosphate were mostly found at sites surrounded by high population density, industrial zones, and wastewater treatment. Coliforms might come from domestic activities, animal waste, and naturally grow and develop in water. Moreover, natural factors were an important source of water pollutants; for example, *TSS* mainly originated from natural processes or mineral sites. Poor water quality in certain locations was primarily due to high levels of BOD_5 and coliforms, especially during the rainy season. This study provides valuable information for agency resource management, including best practices, strategies and tools to control water pollution.

ACKNOWLEDGMENTS

Data were obtained from Department of Natural Resources and Environment of Dak Lak province. The authors would like to thank the other members in the office.

REFERENCES

Amarasinghe, U.A., C.T. Hoanh, D. D’Haeze, and T.Q. Hung. 2015. "Toward sustainable coffee production in Vietnam: more coffee with less water." *Agricultural Systems* 136, 96–105. https://doi.org/10.1016/j.agsy.2015.02.008

Anh, N.T., L.D. Can, N.T. Nhan, B. Schmalz, and T.L. Luu. 2023. "Influences of key factors on river water quality in urban and rural areas: A review." *Case Studies in Chemical and Environmental Engineering* 8, 100424. https://doi.org/10.1016/j.cscee.2023.100424

Fianko, J.R., S.T. Lowor, A. Donkor, and P.O. Yeboah. 2010. "Nutrient chemistry of the Densu River in Ghana." *Environmentalist* 30 (2): 145–152. https://doi.org/10.1007/s10669-010-9254-0

Frohlich, H.L., P. Schreinemachers, K. Stahr, and G. Clemens. 2013. *Sustainable Land Use and Rural Development in Southeast Asia: Innovations and Policies for Mountainous Areas*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33377-4

Giao, N.T., P.K. Anh, and H.T.H. Nhien. 2021. "Spatiotemporal analysis of surface water quality in Dong Thap Province, Vietnam using water quality index and statistical approaches." *Water* 13 (3): 336. https://doi.org/10.3390/w13030336

Giao, N.T., H.T.H. Nhien, P.K. Anh, and P. Thuptimdang. 2022. "Combination of water quality, pollution indices, and multivariate statistical techniques for evaluating the surface water quality variation in Can Tho City, Vietnam." *Environmental Monitoring and Assessment* 194 (11): 844. https://doi.org/10.1007/s10661-022-10474-1

Ha, D.D., D.H. Nguyen, B.C. Nguyen, T.Q. Nguyen, and G.T.N. Le. 2024. "Application of multivariate statistical techniques in the assessment of long-term surface water quality in Dong Thap Province, Vietnam." *Environmental Monitoring and Assessment* 197, 14. https://doi.org/10.1007/s10661-024-13436-x

Hong, T.T.K. and N.T. Giao. 2022. "Analysis of surface water quality in upstream province of Vietnamese Mekong Delta using multivariate statistics." *Water* 14 (12): 1975. https://doi.org/10.3390/w14121975

Lin, C., R. Ma, and B. He. 2016. "Identifying watershed regions sensitive to soil erosion and contributing to lake eutrophication—A case study in the Taihu lake basin (China)." *International Journal of Environmental Research and Public Health* 13 (1): 77. https://doi.org/10.3390/ijerph13010077

Majeb, A., M. Faris, R. Ashraf, E.S. Mohamed, and N. Mohd. 2023. "Correlation between biochemical oxygen demand and chemical oxygen demand, at high salinity bioreactor-based, wastewater treatment plant in Al-Hasa Saudi Arabia." *Journal of Water* 1 (2): 35–47.

Mena-Rivera, L., V. Salgado-Silva, C. Benavides-Benavides, J. Coto-Campos, and T. Swinscoe. 2017. "Spatial and seasonal surface water quality assessment in a tropical urban catchment: Burío River, Costa Rica." *Water* 9, (8): 558. https://doi.org/10.3390/w9080558

Meyfroidt, P., V.T. Phuong, and H.V. Anh. 2013. "Trajectories of deforestation, coffee expansion and displacement of shifting cultivation in the Central Highlands of Vietnam." *Global Environmental Change* 23 (5): 1187–1198. https://doi.org/10.1016/j.gloenvcha.2013.04.005

Mitra, S., S. Ghosh, K.K. Satpathy, B.D. Bhattacharya, S.K. Sarkar, P. Mishra, and P. Raja. 2018. "Water quality assessment of the ecologically stressed Hooghly River Estuary, India: A multivariate approach." *Marine Pollution Bulletin* 126, 592–599. https://doi.org/10.1016/j.marpolbul.2017.09.053

Mohamed, I., F. Othman, A.I.N. Ibrahim, M.E. Alaa-Eldin, and R.M. Yunus. 2015. "Assessment of water quality parameters using multivariate analysis for Klang River basin, Malaysia." *Environmental Monitoring and Assessment* 187, 4182. https://doi.org/10.1007/s10661-014-4182-y

Mukaka, M.M. 2012. “Statistics corner: "A guide to appropriate use of correlation coefficient in medical research.” *Malawi Medical Journal* 24 (3): 69–71.

Nguyen, T.G. and T.H.N. Huynh. 2022. "Assessment of surface water quality and monitoring in southern Vietnam using multicriteria statistical approaches." *Sustainable Environment Research* 32, 20. https://doi.org/10.1186/s42834-022-00133-y

Peña-Guerrero, M.D., A. Nauditt, C. Muñoz-Robles, L. Ribbe, and F. Meza. 2020. "Drought impacts on water quality and potential implications for agricultural production in the Maipo River Basin, Central Chile." *Hydrological Sciences Journal* 65 (6): 1005–1021. https://doi.org/10.1080/02626667.2020.1711911

Pham, L.H., D.D. Tran, T.D.H. Le, T.D. Quang, N.K. Dao, T.T.H. Nguyen, H.A. Nguyen, T.A. Duong, and H.Q. Nguyen. 2022. "Dynamic multivariate analysis for pollution assessment and river habitat conservation in the Vietnamese La Buong watershed." *Environmental Monitoring and Assessment* 194 (Suppl 2): 774. https://doi.org/10.1007/s10661-022-10184-8

Phu, H. 2019. "Method of calculation & application of WQI index to assess the status water quality and proposal of management Luy river Binh Thuan province." *Vietnam Journal of Hydrometeorology* 2, 9–15.

Tram, V.N.Q., N.D. Liem, and N.K. Loi. 2018. "Simulating surface flow and baseflow in Poko catchment, Kon Tum province, Vietnam." *Journal of Water and Climate Change* 10 (3): 494–503. https://doi.org/10.2166/wcc.2018.185

Tram, V.N.Q., H. Somura, T. Moroizumi, and M. Maeda. 2022. "Effects of local land-use policies and anthropogenic activities on water quality in the upstream Sesan River

Basin, Vietnam." *Journal of Hydrology: Regional Studies* 44, 101225. https://doi.org/10.1016/j.ejrh.2022.101225

Tran, T.C., X-D. Do, H.T. Do, A.T. Nguyen, V.H. Dao, and M. Tran. 2024. "Temporal and spatial variation in water quality in the Son La hydropower Reservoir, Northwestern Vietnam." *Vietnam Journal of Earth Sciences* 46 (3): 411–431. https://doi.org/10.15625/2615-9783/20925

WHO. 2016. *Protecting Surface Water for Health. Identifying, assessing and managing drinking-water quality risks in surface-water catchments.* World Health Organization, Geneva, Switzerland. https://iris.who.int/server/api/core/bitstreams/262231c4-a8a7-47ec-bd07-6923d4dfa703/content

Wilbers, G.J., M. Becker, L.T. Nga, Z. Sebesvari, and F.G. Renaud. 2014. "Spatial and temporal variability of surface water pollution in the Mekong Delta, Vietnam." *Science of The Total Environment* 485-486, 653–665. https://doi.org/10.1016/j.scitotenv.2014.03.049