Water Practice & Technology

Water Practice & Technology Vol 18 No 5, 1209 doi: 10.2166/wpt.2023.058

A multivariate approach to the water quality environment of a tropical lake surrounded by agricultural land

Tivany Edwin[a], Puti Sri Komala[a,*], Mas Mera[b], Denny Helard[a] and Vera Surtia Bachtiar[a]

[a] Environmental Engineering Department, University of Andalas, Kampus Unand Limau Manis, Padang 25163, Indonesia
[b] Civil Engineering Department, University of Andalas, Kampus Unand Limau Manis, Padang 25163, Indonesia
*Corresponding author. E-mail: putisrikomala@eng.unand.ac.id

TE, 0000-0003-0129-3469

ABSTRACT

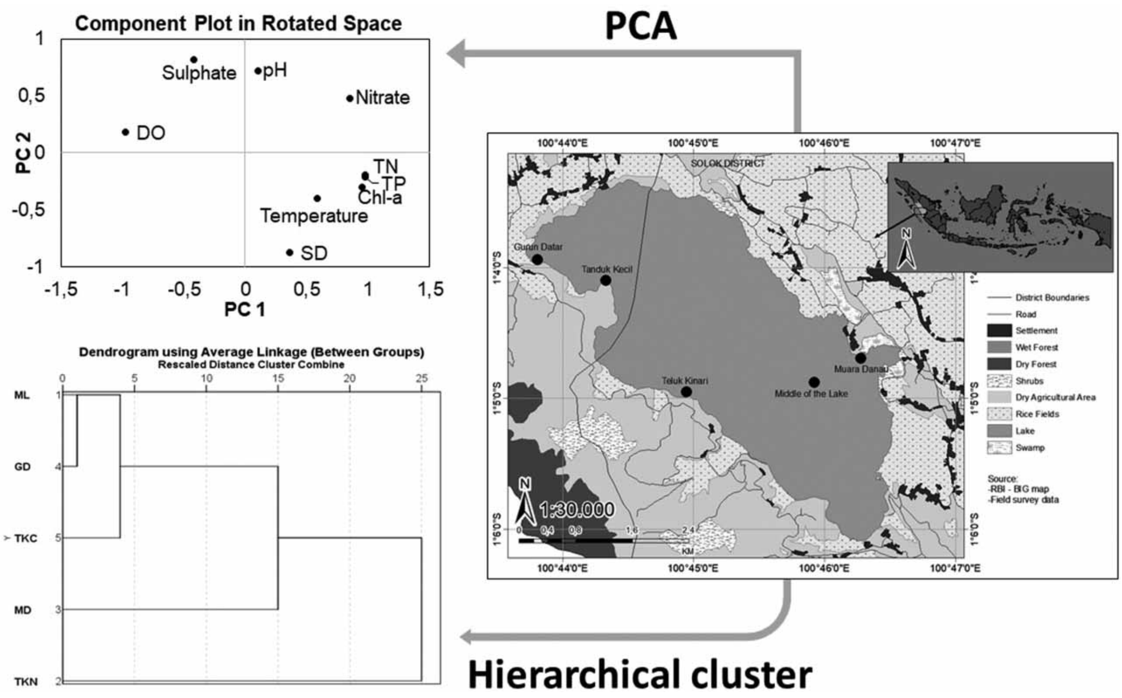
Lake Diatas is a tropical lake surrounded by agricultural land prone to pollution. A multivariate approach to lake water quality will be useful for lake management. The study's objectives are to describe fertiliser use by farmers around the lake, the lake's quality and trophic state spatially, and to use a multivariate approach to lake water quality. The results of a survey showed that most farmers use synthetic fertilisers because of low fertility soil, with self-estimated doses applied. The levels of total nitrogen (TN), total phosphorus (TP), and Secchi depth (SD) marginally exceeded the regulatory standard. However, the trophic status of the lake indicates an intermediate level of nutrients. Principal component analysis (PCA) showed the presence of two main factors with a variance of 85.46% – which showed the important drivers for lake water quality is mainly affected by agricultural activities around the lake. Cluster analysis showed three groups with the same water quality characteristics: the first group consists of locations with higher sulphate levels, the second with low SD, and the third with higher levels of TN and TP. Clearly, agricultural activities affect lake water quality and management regarding land use around the lake is important to prevent pollution.

Key words: agricultural land, multivariate approach, tropical lake, water quality

HIGHLIGHTS

- Lake Diatas is a tropical lake surrounded by agricultural land prone to pollution.
- The levels of TN, TP, and Secchi depth marginally exceeded the regulatory standard.
- The lake indicates mesotrophic status.
- PCA showed the important drivers for lake water quality are mainly affected by agricultural activities around the lake.

GRAPHICAL ABSTRACT



INTRODUCTION

Lakes are vulnerable to pollution because the waterbodies are stationary. Lake water quality must be maintained within the environmental standards to prevent the disturbance of aquatic life and the environment around the lake. Water quality monitoring is important for studying the impact of human activities on aquatic environments, as well as predicting the natural processes occurring in the water (Medallon & Garcia 2021). Lake water quality can be measured using various parameters, for instance, dissolved oxygen (DO), pH, temperature, Secchi depth (SD), chlorophyll-a (Chl-a), total phosphorus (TP), total nitrogen (TN), and other species and parameters such as nitrate (NO_3^-) and sulphate (SO_4^{2-}).

Lake Diatas is formed by magmatic and/or tectonic processes. The main water source is rainfall and it has no major inlet, only runoff from small channels around it (De Maisonneuve *et al.* 2019). The lake is important in West Sumatra province, Indonesia, because it is also a tourist destination. The lake is surrounded by land used for a variety of purposes, mainly rice fields and agricultural land for vegetable crops such as tomatoes and leeks. Lake water is also used as a clean resource for residents (Edwin *et al.* 2021) and receives wastewater from residential areas, which can cause pollution (Nakano *et al.* 2008).

Lake management should be carried out properly to minimise pollution and its harmful impact on aquatic biota. For this reason, a study of the spatial quality of the lake waters and agricultural activity around it are necessary. To support lake water management, it is also important to conduct multivariate statistical analysis, which helps in interpreting information from the water quality parameters measured, as well as estimating the contributors to pollution. The multivariate techniques commonly used are principal component analysis (PCA) and hierarchical clusters (Gradilla-Hernández *et al.* 2020; Horvat *et al.* 2021). Therefore, the main objective of this study is to provide information on fertilisation behaviour by farmers in general on agricultural land around the lake, monitor the lake's water quality, measure physicochemical parameters including temperature, pH, DO, TP, TN, Chl-a, nitrate (NO_3^-), sulphate (SO_4^{2-}), and SD, using both PCA and hierarchical clusters, to determine the multivariate statistics of lake water quality variables and determine the lake's trophic state.

METHODS

Survey analysis

In this study, 30 farmers who own and cultivate properties near the lake were interviewed to obtain an overview of farmers' knowledge and use of fertilisers around the lake. The questions were previously validated by experts and included the type of fertiliser used, the method and time of application, the reasons for using fertiliser, knowledge of cropping patterns, and recommendations for fertiliser dosage. The data obtained are presented in the tabular form containing frequency and percentage information.

Water quality analysis

Water sampling was carried out monthly from October 2019 to February 2020, inclusive. Three samples were collected at the same point each time to ensure consistency of the measured parameter levels. The determination of sample points follows Indonesian National Standard SNI 6989.57:2008 concerning methods for surface water sampling (Indonesian National Standard 2018). Sampling locations were determined according to the national standard, which requires lake inflow; a location that represents the area affected by the lake's use (settlement and rice field runoff at Gurun Datar, dry agricultural area runoff to Tanduk Kecil, and an irrigation channel to Teluk Kinari); the lake's deepest point (the middle of the lake); and its outflow (Muara Danau). Sampling must be carried out at two levels for depths of less than 10 m, and three (surface, middle, and lake bottom) for depths of 10–30 m (Indonesian National Standard 2018). The sampling points used are shown in Figure 1. The sampling point explanations on the graph use the abbreviations in Table 1.

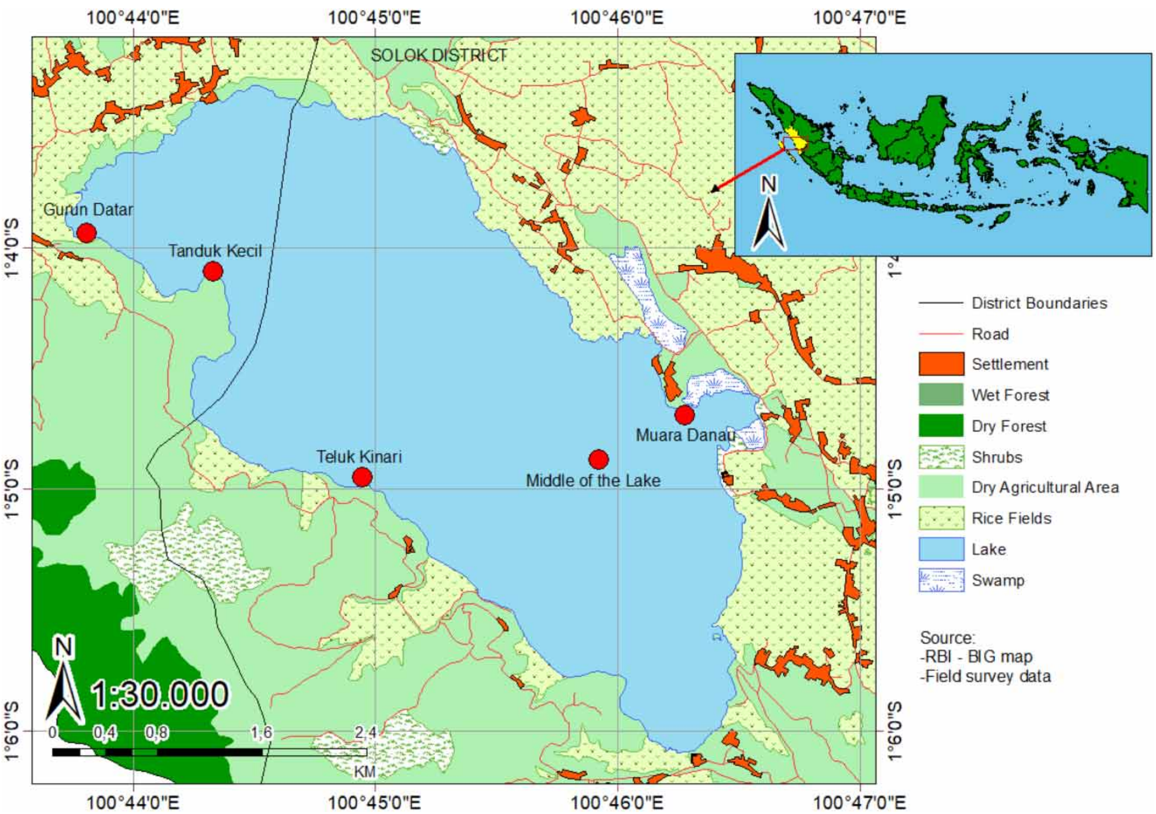


Figure 1 | Sampling points on Lake Diatas.

Samples were collected with a vertical water sampler, put into 1 L containers, and stored in a cooler before being measured in the laboratory. Physicochemical parameters such as DO, temperature, and pH were measured directly in the field with a pH meter and a DO meter EUTECH 1403157. The lake water's transparency was measured with a Secchi disk.

Meanwhile, TP, TN, nitrate, sulphate, and Chl-a were measured in the laboratory using the standard method for the examination of water and wastewater (APHA 2017). The method used for each parameter is shown in Table 2. Samples for Chl-a measurement were collected 1 m below surface and filtered with a nitrocellulose membrane with a vacuum pump, before $MgCO_3$ and distilled water were added to minimise damage to Chl-a. The filter paper containing Chl-a was put in aluminium foil in a cool box and stored at 4 °C until ready to be analysed (APHA 2017).

Table 1 | Sampling point depths

Sampling point	Sampling depth (m)	Longitude Latitude	Description
Middle of the lake (ML)	0; 13; 26; 36	100°45′52.95″ E 1°4′51.68″ S	ML is the deepest point in the lake (max = 47 m below the surface).
Teluk Kinari (TKN)	0; 9; 18	100°44′55.94″ E 1°4′56.59″ S	TKN is close to irrigation canal drains and dry agricultural areas.
Muara Danau (MD)	0; 10.5	100°46′17″ E 1°4′4″ S	MD is the lake's only outlet. It is near major tourism spots. Water is withdrawn here for a community water supply.
Gurun Datar (GD)	0; 6; 13	100°43′48.306″ E 1°3′55.962″	GD is close to rice fields and settlement areas.
Tanduk Kecil (TKC)	0; 6; 13	100°44′19.194″ E 1°4′5.172″ S	TKC is close to a forest now converted to dry agriculture for vegetable crops.

Table 2 | Water quality analysis methods

Parameter	Method	Unit
DO	O_2 probe	mg/L
Temperature	Temperature probe	°C
pH	pH probe	pH units
TP	Ascorbic acid method*	mg/L
TN	Kjeldahl method*	mg/L
Chl-a	Spectrophotometric*	mg/m^3
Nitrate	Colorimetric, Brucine*	mg/L
Sulphate	Turbidimetric*	mg/L
Secchi depth (SD)	Secchi disk	m

*APHA (2017).

The concentration of TN, nitrate, and sulphate in lake sediment was also measured in the last 2 months of sampling. Sediment sampling was carried out at all points except in the middle of the lake because no equipment was available for sediment collection at 47 m depth. Sediment samples were collected by coring to 15 cm depth of sediment, and the samples were stored in bottles in a cooling container before transfer to the laboratory for determination.

Statistical analysis

Statistical analysis was carried out using IBM's SPSS Statistics 22 application. Univariate statistical analysis was performed to determine the range, mean, and standard deviation of each variable, and correlation analysis.

Factor analysis with PCA is used to determine the most significant variables related to a set (Heland *et al.* 2012; Indah *et al.* 2018; Gradilla-Hernández *et al.* 2020). In this case, highly correlated variables were removed to obtain significant variables by recognising the variance in the number of correlated variables to create a smaller group of uncorrelated variables – the principal components (PCs), which are weighted linear combinations of the new variables (Sergeant *et al.* 2016). The Bartlett test is carried out to verify the correlation among the variables, while the Kaiser-Meyer-Olkin (KMO) to ensure the suitability of the correlation matrix. KMO expresses the measure of sample adequacy (MSA). KMO/MSA should be equal to or more than 0.5 to be suitable for factor analysis. Communalities should be examined after the correlation matrix checks. A communality of 100% (or 1) is obtained if the variable determined by all factors has equal variances. The factor loading matrix must be used to evaluate how well the factors describe the items, the factor loading showing how much the factors affect the items. The term 'eigenvalue' refers to the total squared factor loadings for given factors. To make interpretation easier, the factor matrix is typically rotated. It is often rotated orthogonally, which is known as varimax rotation, the principle of which is to rotate the coordinate system until the total of the squared loading variances is maximised. This usually makes the interpretation easier to understand (Cleff 2019).

Sampling site grouping is conducted using hierarchical cluster analysis according to water quality characteristics. A cluster is a group of individuals or objects that have similar characteristics. The results of the cluster analysis also display data that make it easier to understand the source of pollution, so that it can provide information to optimise lake water quality monitoring (Indah *et al.* 2018). Hierarchical clustering is calculated on the basis of data that have been normalised using the Ward method with Euclidian distances (Cleff 2019).

Lake's trophic state

The lake's trophic state was determined based on Lamparelli's trophic state index (TSI), which is suited to tropical lakes (Lamparelli 2004). The Chl-a and TP concentrations were employed as criteria to determine the lake's trophic state.

$$\mathrm{TSI(Chl-a)} = 10 \times \{6 - [0.92 - 0.34\ln(\mathrm{Chl-a})/\ln(2)]\} \quad (1)$$

$$\mathrm{TSI(TP)} = 10 \times \{6 - [1.77 - 0.42\ln(\mathrm{TP})/\ln(2)]\} \quad (2)$$

$$\mathrm{TSI} = \frac{\mathrm{TSI(Chl-a)} + (\mathrm{TSI(TP)}}{2} \quad (3)$$

The lake's trophic status was further divided into six categories within TSI: ultraoligotrophic ($\mathrm{TSI} \leq 47$), oligotrophic ($47 < \mathrm{TSI} < 52$), mesotrophic ($52 < \mathrm{TSI} < 59$), eutrophic ($59 < \mathrm{TSI} < 63$), super eutrophic ($63 < \mathrm{TSI} < 67$), and hyper-eutrophic ($\mathrm{TSI} > 67$).

RESULTS AND DISCUSSION

Survey analysis

The results of the survey analysis are presented in Table 3. Most farmers around the lake use synthetic fertilisers, typically in one to two forms for agricultural land. Some 83.33% use synthetic SP fertiliser, which contains phosphate and sulphur, and all use Nitrogen Phosphor Kalium, which contains nitrogen, phosphate, and potassium. Excessive use of synthetic fertilisers can cause severe environmental impacts; for example, runoff containing fertiliser residue can cause eutrophication, as well as air pollution from the release of gases containing nitrogen and sulphur (Malhi *et al.* 2021; Savci 2012).

Mostly, farmers apply fertiliser after planting for reasons of low fertility soil, and a small number to increase income. Most farmers grow only one type of crop (monoculture). Monoculture is known to have several weaknesses, including water and soil contamination, increased water demand, and reduced biological control (Bourke *et al.* 2021). Most farmers use fertiliser based on their own experience, but do not get information on conservation agriculture that benefits the environment. Socialisation from government to farmers regarding fertilisers and dosage recommendations is very important, so that the community is more directed towards environmentally friendly farming practices, as well as more profitable production.

Physicochemical water quality

The water quality analytical data at each sampling point is presented in Table 4, while the vertical distribution of lake water quality is presented in Figure 2. Based on Indonesian government regulatory standards for water bodies class II, namely those that can be used for water recreation infrastructure/facilities, freshwater fish cultivation, animal husbandry, plants irrigation, and other uses requiring the same water quality, the pH must be in the range 6–9, and the minimum DO concentration 4 mg/L, while the temperature at a standard deviation of less than 3 °C. The concentrations of TP and TN cannot exceed 0.03 and 0.75 mg/L, respectively. The SD of the lake water should not be less than 4 m, the concentration of Chl-a should not be more than 50 mg/m^3, and the maximum sulphate concentration allowed is 300 mg/L (Indonesian Government Regulation 2021). Nitrate is not specified in Indonesian government regulatory standards.

Depending on the sampling location, the average temperature ranged from 21.3 to 22.7 °C, a very small variation. In Figure 2, temperature levels at the midpoint of the lake vary very little with increasing depth, and the lake bottom temperature is higher than that above. However, the temperature's standard deviation is still within the regulatory standard, indicating weak stratification. Drastic temperature changes can be fatal for aquatic organisms. The rates of biological and chemical processes depend on temperature, and aquatic organisms such as microbes and fish depend on a certain temperature range for optimal health. Temperature affects the oxygen content of the water with oxygen concentrations falling with increasing temperature. Temperature also

Table 3 | Behavioural analysis of fertiliser uses for farmers around Lake Diatas

No	Item	*n*	%
1.	Type of fertiliser		
	Synthetic	27	90
	Compost	0	0
	Mixed	3	10
2.	Application method		
	Sown	27	90
	Cast	0	0
	Sown and cast	3	10
3	Application time		
	After planting	27	90
	Manure before planting, synthetic fertiliser after planting	3	10
4	Reasons for using fertilisers		
	Low fertility soil	24	80
	To increase production	2	7
	High vegetable prices	0	0
	Cheap fertiliser prices	0	0
	To increase revenue	3	10
	Other	1	3
5	Cropping pattern		
	Monoculture	22	73
	Intercropping	1	3
	Rotating intercropping	7	23
6	Counselling on conservation techniques in a year		
	Yes	0	0
	No	30	100
7	Knowledge regarding the recommended dosage		
	Understand	27	90
	Do not understand	3	10
8	Recommended dosage source		
	Self-estimated	20	69
	Other farmers	9	31
	Agricultural instructor	0	0
	Other	0	0

Table 4 | Summary of Lake Diatas' water quality statistics

Parameter	ML	TKN	MD	GD	TKC
Temperature (°C)	Range: 20.9–25.9 Mean(SD) = 22.0(1.1)	20.2–22.0 21.3(0.6)	21.1–22.9 21.8(0.6)	20–24.2 22.2(1.2)	20.1–26.9 22.7(1.7)
DO (mg/L)	6.2–8.7 7.5(0.6)	6.6–7.9 7.4(0.4)	6.7–9.0 7.8(0.7)	5.1–8.2 7.1(0.9)	5.8–7.8 7.2(0.6)
pH (unitless)	6.6–8.7 7.3(0.5)	7.4–8.4 7.8(0.4)	7.1–8.6 7.8(0.6)	6.6–8.3 7.6(0.4)	6.4–8.3 7.5(0.5)
TP (mg/L)	0.004–0.036 0.016(0.0001)	0.039–0.050 0.046(0.004)	0.032–0.036 0.017(0.001)	0.038–0.063 0.031(0.028)	0.039–0.040 0.025(0.016)
TN (mg/L)	0.82–5.49 2.00(1.10)	1.96–5.61 2.74(1.43)	0.82–2.03 1.62(0.46)	1.12–2.00 1.67(0.35)	0.82–3.67 2.03(0.99)
Chl-a (mg/m^3)	0.16–0.55 0.40(0.20)	0.19–0.33 0.29(0.07)	0.21–0.38 0.30(0.09)	0.19–0.38 0.31(0.08)	0.30–0.59 0.52(0.14)
Nitrate (mg N/L)	0.82–2.90 1.46(0.96)	1.11–3.66 2.04(1.07)	0.67–3.42 1.52(1.05)	1.02–3.08 1.52(0.87)	0.81–3.33 1.67(1.05)
Sulphate (mg/L)	1.29–24.57 16.47(11.94)	0–18.54 7.33(7.95)	0–14.95 6.15(7.00)	0.06–48.04 19.76(16.77)	0–47.45 20.38(18.94)
SD (m)	2.29–5.8 3.74(1.63)	2.8–4.2 3.33(0.76)	1.6–3.6 2.30(0.78)	1.9–3.82 3.05(0.82)	1.5–3.8 2.39(1.02)

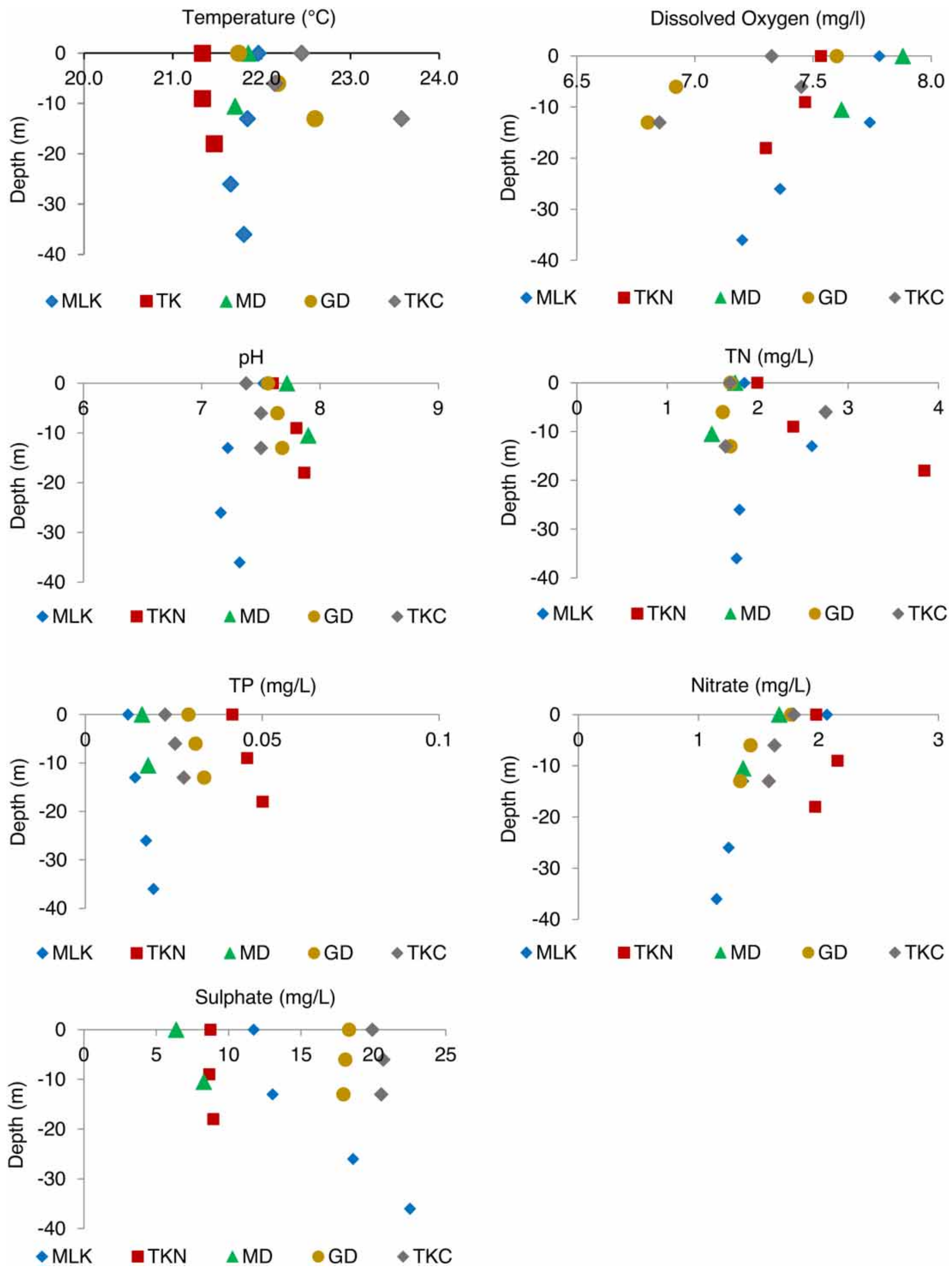


Figure 2 | Vertical distribution of parameters in the lake.

affects the rate of photosynthesis by aquatic plants, the metabolic rate of aquatic organisms, and the sensitivity of organisms to threats (Bhateria & Jain 2016). Causes of water temperature change include weather, shaded lakeside vegetation, and entry of water, including groundwater, with different temperatures into the lake.

Spatially, DO ranges from 7.1 to 7.8 mg/L, which indicates that the lake is still in good condition in this respect. The highest DO was at Muara Danau the lake's outlet, while the lowest was adjacent to the agricultural land near Gurun Datar. DO concentrations tend to decrease with increasing water depth and in locations affected by agriculture. However, the concentrations are still within the government's required range.

The pH of the lake water is in the neutral range, between 7.3 and 7.8. In general, locations near the shore showed higher pH than those in the middle. The trend of the vertical pH profile shows that water depth also affects pH to some extent. This is similar to the results of Wang's (2009) study of the vertical profile in Lake Nam Co (Wang *et al.* 2009). pH is the most important parameter in determining the corrosive properties of water, with low pH correlating with high corrosivity. The decrease in the rate of photosynthetic activity, as well as carbon dioxide and bicarbonate assimilation at higher temperatures and lower oxygen concentrations contributed to the increase in pH, indicating that the balance of carbon dioxide, and carbonates–bicarbonates is influenced mostly by changes in physicochemical conditions (Bhateria & Jain 2016).

In this study, the SD of the lake water in all sampling locations is below 4 m. Low SD is an indicator of eutrophication. SD, which shows how easily light penetrates the water, indicates, in deeper waters that the lake is in good condition. Some species of lake plants and algae require light to grow, thereby providing oxygen to other aquatic organisms (Angradi *et al.* 2018). Low SD in the water is caused by the presence of particles in the water, including algae and sediment.

The Chl-a concentration at the sampling locations was quite low, ranging from 0.29 to 0.52 mg/m^3. Chl-a is a photosynthetic pigment indicating the presence of primary producers or algal biomass (Bennett & Lee 2019), and plays an important role in photosynthesis in the natural environment. The concentration of Chl-a in water depends on the availability of nutrients and the intensity of sunlight, as well as the concentration of species carrying chlorophyll. Algae and other aquatic plants play an important role in lakes because they add oxygen to the water as a by-product of photosynthesis. On the other hand, too much algal mass can reduce the water's transparency, allowing less light to enter. Only the lake's upper layer has enough light for photosynthesis. In eutrophic lakes with lots of algae, their decomposition requires a lot of oxygen so other aquatic organisms are threatened.

Concentrations of TN at all sampling points slightly exceeded the quality standard. There are different trends of TN in each location, TN ranged from 1.62 to 2.74 mg/L, with the lowest concentration at the lake outlet, and the highest at TKN, close to the agricultural runoff inlet to the lake. High levels of TN are found at sampling points of lakes affected by runoff from agricultural activities that may contain fertiliser residues. The vertical profile of TN concentration in Figure 2 shows the same trend as in Zeng & Yang (2018) study in Spring Lake.

Meanwhile, TP exceeded government standards at sampling points affected by agricultural activities, such as TKN and GD, while, at others, they were within the standards. The TP content in lakes ranges from 0.016 to 0.046 mg/L. For the vertical distribution shown in Figure 2, TP concentrations differ at any sampling point. In nature, phosphorus usually exists as part of the phosphate molecule. Phosphorus in aquatic systems can be in the form of either organic or inorganic phosphate. Organic phosphates consist of carbon-based phosphate molecules, such as in plant or animal tissues. Inorganic phosphorus is needed by plants, but animals can use both organic and inorganic phosphates. Organic and inorganic phosphorus can be dissolved in water or suspended attached to particles in the water column (Bhateria & Jain 2016).

Nitrate concentrations ranged from 1.46 to 2.04 mg/L in lake waters, with the highest concentration at the sampling location close to agricultural runoff channels, and the lowest in the middle of the lake. Vertically, the nitrate concentration tends to decrease with increasing depth. The nitrification process produces nitrate under aerobic environmental conditions, which explains the vertical distribution of nitrate was in line with DO in this study (Bhateria & Jain 2016).

Sulphate concentrations ranged from 6.15 to 20.38 mg/L, indicating significantly different concentrations by sampling location. The lowest sulphate concentration is in the middle of the lake, followed by Teluk Kinari. In other areas, the sulphate is much higher. Sulphate concentration increased in the deeper part of the lake, but still met the quality standards set by the government. The sulphate concentration in the lake is influenced by the biogeochemical reactions. When the oxygen concentration in the lake's water decreases, nitrate and sulphate will become the alternative acceptors used by anaerobic bacteria to decompose the remaining organic material. When the nitrate has been used up, sulphate will be substituted to meet oxygen needs, leaving sulphide ions that bond with hydrogen to form hydrogen sulphide, which is toxic (Holmer & Storkholm 2001). The concentration of DO at the bottom of the lake, which still exceeds 7 mg/L, as well as the neutral pH there, indicate that the lake's condition was very good in the sampling period.

The results of the measurement of TN, sulphate, and nitrate in the lake sediments are shown in Figure 3. The sulphate concentration in lake sediments is in the range of 25.21–30.59 mg/g. In sediments, sulphate is available naturally or as a result of agricultural activities. TN concentration was in the range 1.42–2.33 mg/g, and nitrate in the range 0.88–1.71 mg/g. In this study, the levels of TN, sulphate, and nitrate found in the sediment differ little from those in the water column.

Statistical analysis

The correlation results between water quality variables are shown in Table 5. TN and TP have a strong, positive correlation to nitrate, showing that nitrate, the main nitrogen form in organic fertilisers, is almost directly

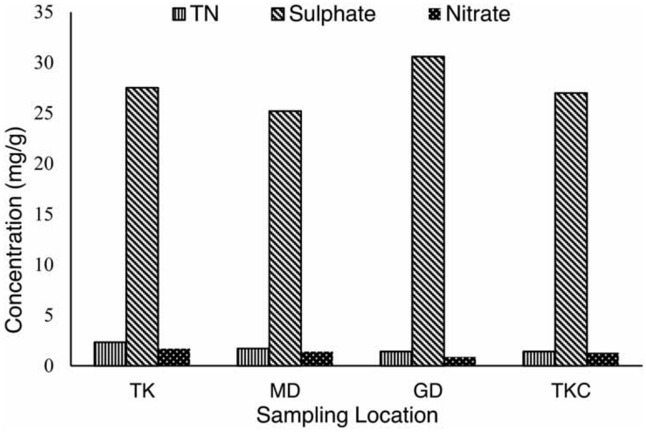


Figure 3 | Parameter concentrations in the lake sediment.

Table 5 | *p*-value comparison between water quality parameters

	Temperature	DO	pH	TP	TN	Chl-a	SO_4^{2-}	SD	NO_3^-
Temperature	R^2: 1 *p*-value: –	−0.507 0.383	−0.572 0.313	−0.471 0.423	−0.528 0.360	0.811 0.096	0.808 0.192	−0.401 0.503	−0.534 0.354
DO		1	0.331 0.586	−0.450 0.447	−0.209 0.736	−0.436 0.463	−0.709 0.291	−0.147 0.813	−0.205 0.740
pH			1	0.509 0.381	0.182 0.770	−0.652 0.233	−0.732 0.268	−0.450 0.447	0.535 0.353
TP				1	0.750 0.144	−0.353 0.560	−0.173 0.827	0.217 0.726	0.886 0.046
TN					1	−0.045 0.942	−0.365 0.635	0.434 0.466	0.906 0.034
Chl-a						1	0.517 0.483	−0.212 0.733	−0.187 0.763
SO_4							1	0.496 0.504	−0.529 0.471

proportional to the TN content in the lake. Phosphate is a fraction of TP which is also the main component in synthetic fertiliser. Even though TN and TP are both major components in the natural environment, the data also indicate that the TN and TP within the lake are related to the contribution of runoff from agricultural activities around it, most of which use synthetic fertilisers.

Table 6 explains the total variance of the PCA analysis results from this study, which comprises two main components. This is supported by the scree plot in Figure 4 which shows that there are two components with eigenvalues exceeding 1. The water quality variables were reduced into two independent main components based on the PCA results.

PC1 as the first factor, accounted for 62.78% of the total variance. There are several water quality variables that show a strong and positive contribution of nutrients including Chl-a, TN, TP, and also temperature. The ranges of Chl-a, TN, TP, and temperature during the study period were 0.29–0.52 mg/m^3, 1.62–2.74 mg/L, 0.016–0.046 mg/L, and 21.3–22.7 °C, respectively. PC1 shows the water quality group that is influenced by agricultural activities because it is dominated by nutrient variables.

PC2 accounted for 22.69% of the total variance, with two variables – nitrate and pH– showing strong contributions. The ranges for nitrate and pH in this study were 1.46–2.04 mg/L and 7.3–7.8, respectively. These data show that the important drivers for Lake Diatas water quality were Chl-a, TN, TP, temperature, nitrate, and pH, which may arise from agricultural activity around the lake.

The hierarchical cluster analysis – see Figure 5 – shows the grouping of sampling locations based on water quality variables. From the results obtained, monitoring of lake water quality can be more efficient at a location that is

Table 6 | Principal component analysis

	Initial Eigenvalues		
Component	**Total**	**% of variance**	**Cumulative %**
1	5.649	62.768	62.768
2	2.042	22.688	85.456
3	0.850	9.443	94.900
4	0.459	5.100	100
5	1.265×10^{-16}	1.406×10^{-15}	100
6	5.146×10^{-17}	5.718×10^{-16}	100
7	3.499×10^{-19}	3.888×10^{-18}	100
8	-1.820×10^{-16}	-2.022×10^{-15}	100
9	-2.714×10^{-16}	-3.016×10^{-15}	100
	Component		
Variable	**PC1**	**PC2**	
Chl-a	0.992	0.000	
TP	0.984	0.178	
TN	0.978	0.188	
DO	−0.971	−0.200	
Sulphate	−0.700	0.604	
Temperature	0.689	−0.147	
Nitrate	0.603	0.766	
pH	−0.172	0.706	
SD	0.666	−0.675	

Extraction method: principal component analysis.
Two components extracted.

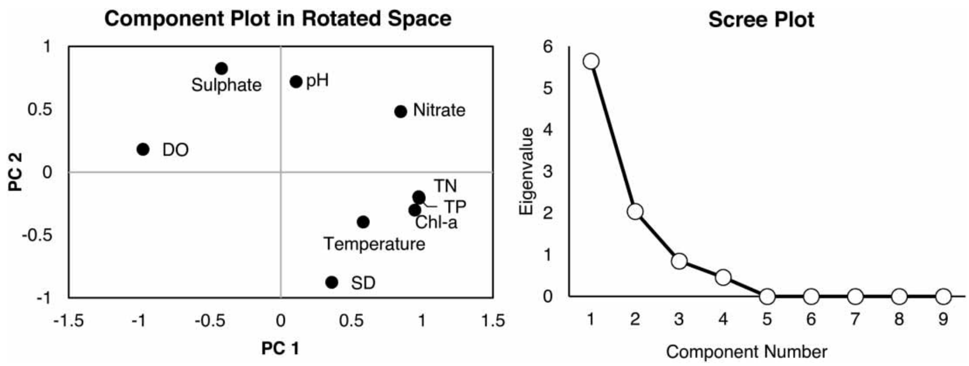


Figure 4 | Eigenvalue component and scree plots for PCA analysis.

representative of each group (Indah *et al.* 2018). Three clusters of sampling locations were obtained from the analysis. The first group consists of three locations: ML, GD, and TKC, which share similar characteristics of temperature, DO, TN, Chl-a, nitrate, and sulphate. These locations contain higher sulphate concentrations than the other two. Sulphate can occur naturally in water, but can also reach lake waters in agricultural area runoff that contains sulphate in fertiliser residues.

Another group with characteristics closes to those of the first group is MD, the lake's outlet. This area has lower concentrations of TN and sulphate than others. However, the SD here is also lower than the others, even though the sampling point is the shallowest. This area is close to the tourism zone, and there is also a boat rental service, which has the potential to cause turbidity, resulting in low SD.

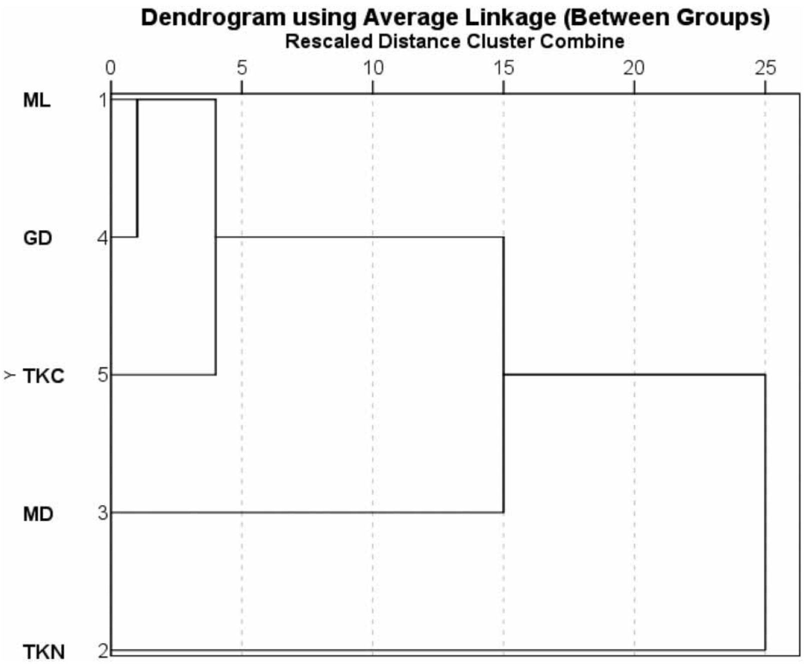


Figure 5 | Dendrogram of sampling location cluster in Lake Diatas.

The group whose characteristics are farthest from the others is TKN, which receives agricultural runoff for dry land agriculture. Dry land agriculture is treated differently from paddy fields, which is why the TN and TP concentrations were higher here than elsewhere.

The trophic state of the lake

The trophic state of the lake was determined by TSI (Lamparelli 2004), calculated using the concentrations of Chl-a and TP in the lake water. The results indicated that the lake of mesotrophic status with TSI of about 54. Mesotrophic lakes are those with medium level productivity. This lake possesses medium nutrient levels, along with submerged aquatic plants. Chl-a provides an estimate of the phytoplankton biomass concentration. (Phytoplankton are microscopic, photosynthetic organisms found in lakes and oceans, and include cyanobacteria and algae). Phytoplankton can quickly develop into extremely productive algal blooms under ideal situations, and can cover water body surfaces (Bennett & Lee 2019). Lake Diatas, surrounded by agricultural areas, cannot be separated from the use of chemical or organic fertilisers. Good management is important to prevent high nutrient concentrations in the lake that may lead to plant growth and have the potential to endanger lake aquatic biota.

CONCLUSIONS

The results of the study show that farmers around Lake Diatas use synthetic fertilisers on their agricultural land. The TSI reveals that Lake Diatas still reports mesotrophic status, indicating a moderate level of nutrients in the lake. The PCA analysis shows that the water quality variables can be reduced to two main factors, the first showing that the important drivers for lake water quality are Chl-a, TP, TN, temperature, nitrate, and pH. The use of fertilisers around the lake may also be important in relation to lake water quality.

Hierarchical cluster results show three groups of sampling locations based on water quality characteristics. The first group consists of locations close to paddy fields, the second is near the lake's outlet, and the third is close to dry land agriculture. All the sampling locations in the first group have high sulphate concentrations, the second has better water quality than other locations but very low SD values, while the third has higher TN and TP concentrations than the others. These findings emphasise that land use near sampling locations affects lake water quality.

ACKNOWLEDGEMENT

The authors would acknowledge Research and Community service Institutions of University of Andalas for the financial Support.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Angradi, T. R., Ringold, P. L. & Hall, K. 2018 Water clarity measures as indicators of recreational benefits provided by U.S. lakes: swimming and aesthetics. *Ecological Indicators* **93**, 1005–1019.

APHA 2017 *Standard Methods for the Examination of Water and Wastewater*, 23rd edn. American Public Health Association/American Water Works Association, Washington, DC, USA.

Bennett, M. G. & Lee, S. S. 2019 Measuring lotic ecosystem responses to nutrients: a mismatch that limits the synthesis and application of experimental studies to management. *Limnology and Oceanography Bulletin* **28**(1), 26–30.

Bhateria, R. & Jain, D. 2016 Water quality assessment of lake water: a review. *Sustainable Water Resources Management* **2**, 161–173.

Bourke, P. M., Evers, J. B., Bijma, P., van Apeldoorn, D. F., Smulders, M. J. M., Kuyper, T. W., Mommer, L. & Bonnema, G. 2021 Breeding beyond monoculture: putting the 'Intercrop' into crops. *Frontiers in Plant Science* **12**, 734167.

Cleff, T. 2019 *Applied Statistics and Multivariate Data Analysis for Business and Economics*. Springer International Publishing, Cham, Switzerland.

De Maisonneuve, C. B., Eisele, S., Forni, F., Hamdi, Park, E., Phua, M. & Putra, R. 2019 Bathymetric survey of lakes Maninjau and Diatas (West Sumatra), and Lake Kerinci (Jambi). *Journal of Physics: Conference Series* **1185**, 012001.

Edwin, T., Bachtiar, V. S., Helard, D., Komala, P. S. & Kusuma, R. N. 2021 Case study: Water resources, knowledge and safety practices of pesticides among farmers around lake Diatas of West Sumatera. *Journal of Physics: Conference Series* **757**, 012047.

Gradilla-Hernández, M. S., de Anda, J., Garcia-Gonzalez, A., Meza-Rodríguez, D., Yebra Montes, C. & Perfecto-Avalos, Y. 2020 Multivariate water quality analysis of Lake Cajititlán, Mexico. *Environmental Monitoring and Assessment* **192**(1), 5.

Helard, D., Fajri, J. A., Setiyawan, A. S., Li, F., Yamada, T., Horio, A., Huang, M. & Kawaguchi, T. 2012 Formation and role of bacterial community in the sediment bed of open channel receiving Johkasou effluent: multivariate statistical analysis interpretation. *Journal of Japan Society of Civil Engineers, Series G (Environmental Research)* **68**, III_1–III_11.

Holmer, M. & Storkholm, P. 2001 Sulphate reduction and sulphur cycling in lake sediments: a review. *Freshwater Biology* **46**, 431–451.

Horvat, M., Horvat, Z. & Pastor, K. 2021 Multivariate analysis of water quality parameters in Lake Palic, Serbia. *Environmental Monitoring and Assessment* **193**(7), 410.

Indah, S., Helard, D., Herfi, M. A. & Hamid, H. 2018 Spatial variation of metals in the Batang Arau River, West Sumatera, Indonesia. *Water Environment Research* **90**, 234–243.

Indonesian Government Regulation 2021 *Indonesian Government Regulation No. 22 Year 2021 about Implementation of Environmental Protection and Management*. Environmental Protection and Management Regulation, Indonesia.

Indonesian National Standard 2018 *SNI 6989.57:2008 about Methods for Surface-Water Sampling*. Indonesia.

Lamparelli, M. C. 2004 *Graus de trofia em corpos d'água do estado de São Paulo: avaliação dos métodos de monitoramento (Degrees of Trophicity in Water Bodies in the State of São Paulo: Evaluation of Monitoring Methods)*. Doctoral Thesis, São Paulo University, São Paulo, Brazil (accessed 19 February 2023).

Malhi, G. S., Kaur, M. & Kaushik, P. 2021 Impact of climate change on agriculture and its mitigation strategies: a review. *Sustainability* **13**, 1–21.

Medallon, M. C. & Garcia, E. 2021 Physical and chemical properties of water: case of Taal Lake, Philippines. *International Journal of GEOMATE* **21**, 11–19.

Nakano, T., Tayasu, I., Yamada, Y., Hosono, T., Igeta, A., Hyodo, F., Ando, A., Saitoh, Y., Tanaka, T., Wada, E. & Yachi, S. 2008 Effect of agriculture on water quality of Lake Biwa tributaries, Japan. *Science of the Total Environment* **389**, 132–148.

Savci, S. 2012 An agricultural pollutant: chemical fertilizer. *International Journal of Environmental Science and Development* **3**, 73–80.

Sergeant, C. J., Starkey, E. N., Bartz, K. K., Wilson, M. H. & Mueter, F. J. 2016 A practitioner's guide for exploring water quality patterns using principal components analysis and procrustes. *Environmental Monitoring and Assessment* **188**, 1–15.

Wang, J., Zhu, L., Daut, G., Ju, J., Lin, X., Wang, Y. & Zhen, X. 2009 Investigation of bathymetry and water quality of Lake Nam Co, the largest lake on the central Tibetan Plateau, China. *Limnology* **10**, 149–158.

Zeng, Y. & Yang, C. 2018 Vertical distribution of total carbon, nitrogen and phosphorus in sediments of Drug Spring Lake, Wudalianchi. *IOP Conference Series: Earth and Environmental Science* **113**(1), 0.012196.

First received 6 January 2023; accepted in revised form 2 April 2023. Available online 13 April 2023