BIODIVERSITAS
Volume 27, Number 2, February 2026
Pages: 20

ISSN: 1412-033X
E-ISSN: 2085-4722
https://doi.org/10.13057/biodiv/d270206

Development and validation of a tropical diatom-based index for bioassessment of the Yang River, Northern Thailand

SUTTAYOT YIMPOOLSAP♥
Faculty of Science and Technology, Phranakhon Si Ayutthaya Rajabhat University. Phranakhon Si Ayutthaya Province, 13000, Thailand.
Tel.: +66-35276555, ♥email: s.yimpoolsap@gmail.com

Manuscript received: 12 October 2025. Revision accepted: 6 February 2026.

Abstract. *Yimpoolsap S. 2026. Development and validation of a tropical diatom-based index for bioassessment of the Yang River, Northern Thailand. Biodiversitas 27 (2): d270206. https://doi.org/10.13057/biodiv/d270206.* Freshwater ecosystems in tropical Asia are increasingly threatened by nutrient enrichment, organic pollution, and land-use intensification. Diatoms are widely recognised as sensitive biological indicators, yet most available diatom indices were developed for temperate regions and may not adequately reflect ecological conditions in tropical rivers. This study aimed to develop a regionally calibrated, river-specific Diatom Index using the Yang River, Northern Thailand, as a model tropical system for water quality assessment. Water sampling and benthic diatom collection were conducted across seven stations from November 2016 to April 2017. Physico-chemical parameters were analysed according to American Public Health Association (APHA 2017), and diatom assemblages were identified to the lowest possible taxonomic level. Indicator Values (IV) and Weighted Averages (WA) were computed for 29 dominant taxa, and the final Diatom Index (DI) was validated against the conventional Water Quality Index (WQI). The study addresses a key methodological limitation in tropical bioassessment by developing an index based on observed ecological responses of tropical diatom assemblages, rather than relying on tolerance values derived from temperate regions. Reproducibility was evaluated using three independent chemistry replicate datasets (R1-R3), which showed high consistency in WA/IV estimation and DI computation ($\rho = 0.860$-0.907, $p < 0.001$), indicating that the index-building process was robust and repeatable. The DI showed a significant positive relationship with WQI ($\rho = 0.549$, $r^2 = 0.3102$, $p < 0.001$), indicating that diatom assemblages effectively captured nutrient and organic pollution gradients. Spatial variation in diatom communities aligned with upstream-midstream-downstream differences in pH, DO, BOD, and nutrient levels. While developed from a single river system, the index provides a regionally appropriate and ecologically meaningful framework for diatom-based water quality assessment in tropical rivers, subject to further validation across additional catchments. The index offers a cost-effective and regionally calibrated approach for routine monitoring and supports improved management of tropical freshwater ecosystems.

Keywords: Benthic diatoms, bioindicators, ecological guilds, tropical rivers, water quality

INTRODUCTION

Freshwater ecosystems in tropical Asia are increasingly subjected to anthropogenic pressures, including nutrient enrichment, organic pollution, agricultural runoff, and expanding domestic and urban wastewater inputs. These stressors can alter community structure, reduce biodiversity, and impair ecosystem functioning, ultimately threatening the sustainability of aquatic resources. As such, robust bioassessment tools capable of capturing both chemical and biological signals are essential for effective river monitoring and management.

Among aquatic microorganisms, benthic diatoms (Bacillariophyceae) are widely recognised as sensitive indicators of water quality due to their short generation times, high species diversity, and well-characterised ecological preferences (Round et al. 1990; Kelly et al. 2007; Rimet and Bouchez 2012). Their rapid and predictable responses to gradients in nutrients, pH, oxygen, and organic matter have led to the development of numerous diatom-based indices in Europe and North America, including the Trophic Diatom Index (TDI) (Kelly and Whitton 1995) and the Pollution Sensitivity Index (IPS) (Prygiel and Coste 2000). These indices, however, are primarily calibrated using temperate datasets, where diatom species composition and pollution tolerances differ markedly from those of tropical rivers.

Consequently, applying temperate-derived indices directly to tropical systems may produce misleading or ecologically inconsistent assessments. Tropical assemblages in Southeast Asia are dominated by species with different ecological ranges, guild structures, and responses to pollution compared to their temperate counterparts. Therefore, regionally calibrated indices that incorporate local diatom autoecology and reflect national water quality standards are required for accurate assessment in tropical rivers. This need has been emphasised in recent studies, which show that temperate calibration datasets underrepresent species commonly found in tropical streams (Rimet and Bouchez 2012). Accordingly, the Diatom Index developed in this study is intended to provide an ecologically relevant assessment of water quality in tropical environments, where species-environment relationships differ from those observed in temperate regions.

In Thailand, research related to algal and diatom-based bioassessment has expanded over the past two decades. Peerapornpisal et al. (2004) conducted long-term monitoring of water quality in the Ang Kaew Reservoir, Chiang Mai University, using phytoplankton as bioindicators, demonstrating how algal communities respond to organic and nutrient pollution in tropical lentic systems. Although reservoir-based and focused on phytoplankton, this study provided an early foundation for algal bioassessment in Thailand. Subsequent riverine studies investigated spatial variation of benthic diatom communities and water quality in the Ping River (Leelahakrienkrai and Peerapornpisal 2010) and later extended these assessments across six major river basins in Thailand (Leelahakriengkrai and Peerapornpisal 2011). Prasertsin et al. (2021) examined phytoplankton and benthic diatom assemblages along environmental gradients in the Kok River, Chiang Rai, highlighting their bioindicator potential for detecting nutrient enrichment under varying flow and land-use conditions. Studies from other parts of tropical Asia, including the Philippines (Ladera et al. 2025) and Malaysia (Malek et al. 2012), similarly emphasize the need for indices that integrate local species ecology and regional water-quality characteristics. These collective findings underscore the importance of developing regionally calibrated diatom-based indices tailored to tropical river systems.

Despite these advances, there remains no unified tropical Diatom Index that combines regional species optima, ecological guilds, and quantitative relationships with water chemistry. This represents a critical research gap for Southeast Asia. Most studies describe community-environment relationships but do not translate these into a calibrated index applicable across tropical river conditions.

The Yang River, a tributary of the Nan River in Northern Thailand, provides an ideal setting for such development. The river spans forested headwaters, agricultural mid-reaches, and semi-urban downstream areas, thereby representing a gradient of environmental and anthropogenic influences. This study therefore aimed to (i) characterise spatial variation in water quality; (ii) analyse diatom community composition and ecological guilds; and (iii) develop and validate a tropical Diatom Index tailored to Northern Thailand. By integrating regionally relevant ecological information, the proposed index is expected to support water-quality assessment in tropical river systems.

MATERIALS AND METHODS

Study area

The study was conducted in the Yang River, a mountainous tributary of the Nan River located in Tha Wang Pha District, Nan Province, Thailand. The Nan River itself is one of the principal tributaries of the Chao Phraya River, forming a major river basin that sustains agriculture, livelihoods, and water resources for millions of people in central Thailand (Molle and Mollinga 2003). The Yang River catchment is characterized by steep-sloped terrain, mixed deciduous and evergreen forest in the upstream section, and agricultural land use such as paddy fields and orchards in the middle and downstream sections. The climate is tropical monsoon, with a distinct wet season (May-October) and dry season (November-April). The study was conducted during the dry season (November 2016-April 2017), which in Northern Thailand can be subdivided into a cool-dry season (approximately November-February) and a hot-dry season (approximately March-April), representing typical dry-season hydrological conditions. Seven sampling stations were established along the Yang River, covering upstream, midstream, and downstream reaches (Figure 1 and Table 1). The spacing between stations was not uniform and was determined by accessibility and the need to capture major upstream-downstream environmental transitions rather than equal-distance intervals; therefore, inter-station distance was not treated as an explicit analytical variable, but as part of the longitudinal context influencing water quality and diatom assemblages. Elevations range from 226 to 317 meters above sea level (masl), reflecting the river's gradual descent from upland headwaters to lowland agricultural areas. These stations were selected to represent a gradient of environmental conditions from relatively pristine upstream habitats to downstream areas influenced by agriculture and settlement.

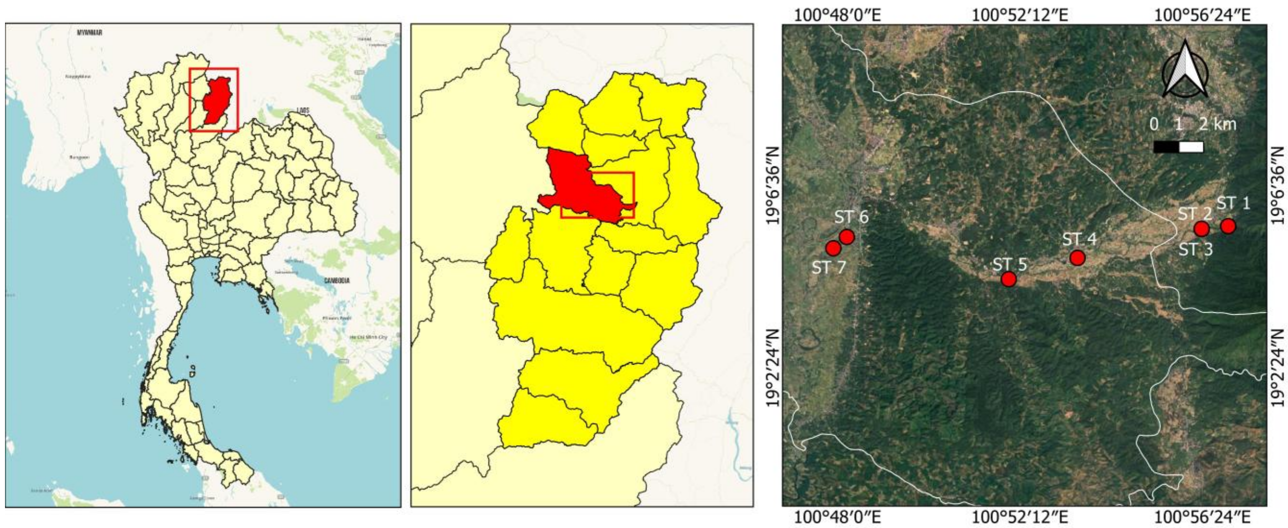

Figure 1. Map of the Yang River showing the seven sampling stations (ST1-ST7) in Nan Province, Northern Thailand**.** Upstream sites (ST1-ST2), midstream sites (ST3-ST5), and downstream sites (ST6-ST7) represent a longitudinal environmental gradient

Table 1. Sampling stations in the Yang River, Nan Province, Thailand

Code	Station name	Coordinates (Latitude/Longitude, WGS84)	Elevation (masl)	Description
ST1	Sila Phet waterfall	19°5′40.36″N, 100°56′59.95″E	316.99 m	Upstream section near the waterfall; relatively undisturbed forested area.
ST2	Ban Don Kaeo	19°5′36.33″N, 100°56′23.20″E	310.59 m	Upper catchment, higher elevation; transition between forest and agriculture.
ST3	Ban Don Mun	19°5′36.68″N, 100°56′23.20″E	265.17 m	Upstream reach with adjacent agricultural land.
ST4	Ban Kong	19°4′56.45″N, 100°53′31.87″E	257.86 m	Midstream reach; surrounded by mixed cropland.
ST5	Ban Na Fa	19°4′27.01″N, 100°51′56.86″E	254.20 m	Midstream reach with moderate anthropogenic influence.
ST6	Ban Fai Mun	19°5′25.40″N, 100°48′13.14″E	226.77 m	Downstream reach; agricultural and residential areas nearby.
ST7	Ban Ton Hang	19°5′9.68″N, 100°47′54.53″E	226.16 m	Downstream section at lowest elevation; subject to high land-use pressure.

Table 2. Sampling months at each station in the Yang River, Thailand (November 2016-April 2017)

Station	Nov 2016	Dec 2016	Jan 2017	Feb 2017	Mar 2017	Apr 2017	Total samples
ST1 - Sila Phet waterfall	-	✓	-	✓	✓	✓	4
ST2 - Ban Don Kaeo	✓	✓	-	✓	✓	✓	5
ST3 - Ban Don Mun	✓	✓	-	✓	✓	✓	5
ST4 - Ban Kong	✓	✓	-	✓	✓	✓	5
ST5 - Ban Na Fa	✓	✓	-	✓	✓	-	4
ST6 - Ban Fai Mun	✓	✓	✓	✓	✓	-	5
ST7 - Ban Ton Hang	✓	✓	✓	✓	✓	-	5

Note: ✓: Sample collected, -: Not sampled

Field sampling

Field sampling was carried out between November 2016 and April 2017, representing the cool and hot seasons. A total of 33 samples were collected from seven stations, with 4-5 samples per station depending on accessibility and seasonal conditions. Samples were collected on a monthly basis where site conditions allowed, in order to capture temporal variability in physico-chemical parameters and diatom assemblages. Each sampling event included concurrent measurements of water quality and benthic diatom collection following standardized bioassessment protocols. The detailed sampling schedule, including the months sampled at each station, is presented in Table 2.

At each visit, in situ readings for pH, temperature, electrical conductivity, and total dissolved solids were taken in field triplicate (three consecutive meter readings within 1-2 minutes) and averaged; grab-water samples for laboratory analyses were collected in triplicate per station for all physico-chemical parameters to ensure analytical replication and reliability. For benthic diatoms, 5-10 cobbles per station were composited into one homogenized sample per station per month (i.e., one composite replicate representing the station).

Water quality analysis

At each sampling station, in situ measurements of physical and chemical parameters were recorded, including water temperature (WT, °C), pH, electrical conductivity (EC, µS/cm), and dissolved solids (TDS, mg/L), using portable meters. A portable pH meter (SI Analytics, HandyLab 100) was used for measuring pH and water temperature, while a portable TDS/EC meter (Thermo Scientific, Eutech TDS6+) was used for measuring dissolved solids and electrical conductivity. Water samples were collected in polyethylene bottles for the analysis of NH_4^+-N, NO_3^--N, and Suspended Solids (SS); in BOD bottles for the determination of DO and BOD; and in glass bottles for phosphate (PO_4^{3-}-P) analysis. Laboratory analyses followed the standard methods of American Public Health Association (APHA 2017), including dissolved oxygen (DO, mg/L), biochemical oxygen demand (BOD, mg/L), ammonium (NH_4^+-N, mg/L), nitrate (NO_3^--N, mg/L), phosphate (PO_4^{3-}-P, mg/L), and suspended solids (SS, mg/L).

Quality Assurance and Quality Control (QA/QC) followed APHA guidance: meters were calibrated daily (pH three-point calibration using standard buffers at pH 4.00, 7.00, and 10.00; EC calibrated with certified conductivity standards provided in the test kit; DO air-saturation check); field blanks were included once per trip, and laboratory analyses were performed in duplicate with method blanks and certified-standard checks when applicable. Results outside control limits triggered re-analysis. Units and chemical notation were standardized across text and tables (e.g., DO, BOD in mg/L; NH_4^+-N, NO_3^--N, PO_4^{3-}-P), and all abbreviations were defined at first mention.

Diatom sampling and preparation

Benthic diatoms were collected by scraping submerged stones (5-10 cobbles per site) using a toothbrush, and the material was rinsed into polyethylene bottles. Samples were preserved with Lugol's iodine solution (final concentration 1-2%) immediately after collection (Battarbee 1986). For permanent slide preparation, organic matter was removed by oxidizing samples with 30% hydrogen peroxide (H_2O_2) in the presence of potassium dichromate ($K_2Cr_2O_7$) as a catalyst (Battarbee 1986; Taylor et al. 2007). Samples were gently heated at 70-80°C until digestion was complete, then repeatedly rinsed with distilled water by centrifugation (3,000-4,000 rpm for 10 minutes) until the supernatant was clear (Round et al. 1990). Clean diatom frustules were mounted on glass slides using Permount® resin to obtain permanent preparations for microscopic analysis. For within-sample replication, three permanent slides were prepared from each composite sample. All slides were examined by a single analyst to ensure consistency in identification and counting, as no additional diatom specialists were available at the time of analysis.

Identification and enumeration

Diatoms were identified to the lowest possible taxonomic level, primarily to species, using standard taxonomic references (Patrick and Reimer 1966; Krammer and Lange-Bertalot 1986-1991; Lange-Bertalot and Krammer 1989; John et al. 2011), supplemented with regional algal floras and identification guides from Thailand (Peerapornpisal et al. 2004). A minimum of 300 valves per slide were counted under a light microscope at 1000× magnification using oil immersion. Counting at least 300 valves is a widely accepted practice in diatom studies to ensure statistical reliability, as it provides a representative estimate of species relative abundance and reduces sampling error (Schrader and Gersonde 1978; Battarbee 1986). Taxonomic names were verified against an online taxonomic backbone (Guiry and Guiry 2025) to ensure standardized nomenclature. For quality control, approximately 10% of slides were re-examined by the same analyst after a time interval to assess intra-analyst consistency. Discrepancies exceeding 10% in dominant taxa counts were re-evaluated using the original slides.

Data analysis

One-way ANOVA was performed to test differences in water quality parameters, diatom diversity indices, and ecological guild composition among stations, and post-hoc comparisons of means were conducted using Tukey's HSD test at $p < 0.05$. Before applying parametric tests, distributional assumptions were evaluated: normality by the Shapiro-Wilk test (Shapiro and Wilk 1965) and homogeneity of variances by Levene's test (Levene 1960) (Table 3: Normality Assessment). Where necessary, data were $\log_{10}(X+1)$-transformed to stabilize variance. If the data were normally distributed but variances were heterogeneous (Levene $p < 0.05$), a Welch F-test was used instead of the standard ANOVA, followed by appropriate pairwise comparisons to identify significant differences among stations. If both assumptions of normality and homogeneity were violated, non-parametric Kruskal-Wallis tests (with Dunn's post-hoc, $\alpha = 0.05$) were applied.

Principal Component Analysis (PCA) was applied to identify major gradients in water-quality variables. All variables were standardized (z-scores) prior to ordination to remove scale effects. Because several environmental variables showed non-normal distributions and monotonic relationships, Spearman's rank correlation matrix was used to examine inter-variable relationships and reduce redundancy prior to PCA. Strongly correlated pairs ($|\rho| > 0.7$) were considered redundant, and one variable from each pair was excluded to ensure orthogonality among predictors.

Table 3. Shapiro–Wilk test for normality of water-quality parameters, diversity index, Ecological guild, Diatom Index (DI) and Water Quality Index (WQI) in the Yang River, Thailand

Variable	N	Shapiro-Wilk W	p-Value	Normality (p > 0.05)
Temp (°C)	33	0.9137	0.01	No
pH	33	0.9038	0.01	No
EC (µS/cm)	33	0.8447	<0.01	No
DS (mg/L)	33	0.3714	<0.01	No
SS (mg/L)	33	0.8257	<0.01	No
DO (mg/L)	33	0.8271	<0.01	No
BOD (mg/L)	33	0.7603	<0.01	No
NH_4^+-N (mg/L)	33	0.4553	<0.01	No
NO_3^--N (mg/L)	33	0.4142	<0.01	No
PO_4^{3-}-P (mg/L)	33	0.8929	<0.01	No
No. Of Species	33	0.9724	0.55	Yes
H′	33	0.9599	0.26	Yes
D	33	0.8855	<0.01	No
J′	33	0.9596	0.25	Yes
Low-Profile (%)	33	0.9552	0.18	Yes
High-Profile (%)	33	0.8694	<0.01	No
Motile (%)	33	0.9183	0.02	No
DI	33	0.9236	0.02	No
WQI	33	0.9119	0.01	No

Note: Normality was assessed using the Shapiro-Wilk test ($\alpha = 0.05$). Species richness (No. of species), H′, J′, and Low-profile (%) were normally distributed ($p > 0.05$), whereas all other variables were non-normally distributed. These results informed the selection of non-parametric statistics and the use of Spearman's correlations for screening prior to multivariate analyses

Canonical Correspondence Analysis (CCA) was further used to examine the relationships between diatom assemblages, ecological guilds, and environmental variables. Prior to CCA, environmental variables were screened for multicollinearity using both Spearman's rank correlations and Variance Inflation Factors (VIF). Highly correlated pairs ($|\rho| > 0.7$) or variables with VIF ≥ 10 were considered redundant and removed following Zuur et al. (2010). Consequently, NH_4^+-N was excluded, and nine independent variables (Temp, pH, EC, DS, SS, DO, BOD, NO_3^--N, and PO_4^{3-}-P) were retained for the final CCA. CCA model significance and axis significance were tested using 999-permutation Monte Carlo tests ($p < 0.05$).

Species richness, Shannon-Wiener Diversity Index (H′), evenness (J′), and Simpson's Dominance Index (D) were calculated for each sample to assess diversity patterns. Ecological guilds were classified into low-profile, high-profile, and motile guilds following Passy (2007) and Rimet and Bouchez (2012) (Table 4). Water-quality classification based on Shannon-Wiener Diversity Index values was interpreted using thresholds adapted from international standards (Table 5). All statistical analyses were carried out using PAST software (Hammer et al. 2001).

The relationship between the developed Diatom Index (DI) and the conventional Water Quality Index (WQI) was evaluated using Spearman's rank correlation (ρ) because both indices were non-normally distributed. For visualization purposes only, a scatter plot with a fitted linear trendline and corresponding r^2 value was generated in Excel as a descriptive illustration; this graphical fit was not used for statistical inference.

Development of a Diatom Index

Diatom taxa with relative abundance greater than 1% were selected for index development. For each taxon, an Indicator Value (IV) was calculated following Dufrêne and Legendre (1997), which integrates Relative Abundance (RA), representing species dominance, and Relative Frequency (RF), representing consistency of occurrence across samples, as shown in the following equation:

$$IV_{ij} = RA_{ij} \times RF_{ij}$$

Where:

IV_{ij} : The indicator value of species i in site j

RA_{ij} : The relative abundance of species i at site j

RF_{ij} : The relative frequency of species i at site j

Relative abundance and frequency were calculated according to these equations:

$$RA_{ij} = \frac{A_{ij}}{A_i}$$

$$RF_{ij} = \frac{S_{ij}}{S_i}$$

Where:

A_{ij} : The mean abundance of species i at site j

A_i : The sum of mean abundances of species i across all sites

S_{ij} : The number of samples where species i was present at site j

S_j : The total number of samples collected at site j

To estimate ecological optima and tolerance values of each selected taxon, a Weighted Average (WA) approach (Lepš and Šmilauer 2003) was applied, as shown in the following equation:

$$WA_{(Sp)} = \frac{\sum_{i=1}^{n}(Env_i \times Abund_i \times IV_i)}{\sum_{i=1}^{n} Abund_i \times IV_i)}$$

Where:

Env_i : The measured environmental variable (e.g., BOD, NO_3^--N, NH_4^+-N, PO_4^{3-}-P) in sample i

$Abund_i$: The abundance of species i

IV_i : The indicator value of species i

Table 5. Classification of water quality based on Shannon-Wiener Diversity Index (H′)

H′ range	Water quality class	Pollution status
0.00 - 0.90	Very poor	Very heavily polluted
1.00 - 1.90	Poor	Heavily polluted
2.00 - 2.90	Fair	Moderately polluted
3.00 - 3.90	Good	Lightly polluted
≥4.00	Excellent	Unpolluted

Note: Classification thresholds adapted from Jørgensen et al. (2010) and applied in tropical river studies for ecological assessment

Table 4. Ecological guilds of benthic diatom genera used for classification

Ecological guilds	Genus
Low - profile	*Achnanthidium, Achnanthes, Amphora, Aneumastus, Brachysira, Cyclotella, Cymbopleura, Cocconeis, Cyclostephanos, Delicata, Diploneis, Discostella, Encyonema, Encyonopsis, Eucocconeis, Karayevia, Kolbesia, Meridion, Placoneis, Planothidium, Platessa, Psammothidium, Rhoicosphenia, Reimeria, Stephanodiscus*
High - profile	*Aulacoseira, Cymbella, Diadesmis, Diatoma, Didymosphenia, Eunotia, Fragilaria, Frustulia, Gomphonella, Gomphonema, Gomphoneis, Gomphosphenia, Melosira, Pleurosira, Pseudostaurosira, Staurosira, Staurosirella, Tabularia, Ulnaria*
Motile	*Adlafia, Bacillaria, Caloneis, Craticula, Delicata, Denticula, Eolimna, Epithemia, Fallacia, Fistulifera, Geissleria, Gyrosigma, Hantzschia, Hippodonta, Luticola, Mayamaea, Navicula, Naviculadicta, Nitzschia, Nitzschiella, Nupela, Pinnularia, Sellaphora, Simonsenia, Stauroneis, Surirella, Tryblionella*

Note: Classification follows Passy (2007) and Rimet and Bouchez (2012). Low-profile taxa are small adnate species tolerant of disturbance; high-profile taxa are upright forms adapted to stable habitats; motile taxa are mobile species tolerant of organic enrichment

Table 6. Water quality classification based on BOD, ammonium-nitrogen (NH_4^+-N), nitrate-nitrogen (NO_3^--N), and phosphate-phosphorus (PO_4^{3-}-P) used for calculating indicator values

Parameter	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
BOD (mg/L)	<0.5	0.5-1.0	1.1-2.0	2.1-4.0	4.1-10.0	10.1-20.0	>20.0
NH_4^+-N (mg/L)	<0.01	0.01-0.19	0.20-0.39	0.40-0.59	0.6-0.99	1.0-5.0	>5.0
NO_3^--N (mg/L)	<0.01	0.01-0.19	0.20-0.39	0.40-0.79	0.8-1.90	2.0-10.0	>10.0
PO_4^{3-}-P (mg/L)	<0.01	0.02-0.04	0.05-0.06	0.07-0.19	0.2-0.99	1.0-3.0	>3.0
Nutrient level	Ultra-oligotrophic	Oligotrophic	Oligo-mesotrophic	Mesotrophic	Meso-eutrophic	Eutrophic	Hypertrophic

Note: Classes 1-7 represent trophic categories ranging from ultra-oligotrophic (Class 1) to hypertrophic (Class 7). Source: Adapted from Lorraine and Vollenweider (1981), Wetzel (2001), Peerapornpisal et al. (2004), Jones and Medrano (2006), and Pollution Control Department (PCD 2010)

In this approach, the environmental variable (e.g., BOD, NH_4^+-N, NO_3^--N, or PO_4^{3-}-P) measured in each sample was multiplied by the abundance of a given species in that sample. The results were summed and divided by the total abundance to obtain the ecological optimum for that species. After obtaining the Indicator Value (IV) and Weighted Average (WA) for each taxon, the final Diatom Index (DI) was calculated for each sample by integrating the Relative Abundance (RA) and indicator values of the selected taxa. The relationship is expressed in the Equation below, which computes the Diatom Index for each sample as the weighted mean of the indicator values of all species present, proportional to their relative abundances.

$$Diatom\ index = \frac{\sum[Relative\ Abundant\ (\%) \times WA(Sp)_{(Sp)}]}{\sum Relative\ Abundant\ (\%)}$$

In this formulation, IV represents the indicator strength of each species, while RA reflects its proportional contribution to the assemblage at a given site. Higher DI values therefore correspond to assemblages associated with higher nutrient classes (mesotrophic to hypertrophic conditions), whereas lower DI values indicate assemblages typical of oligotrophic or low-nutrient conditions.

The resulting DI values were interpreted in relation to established water-quality and trophic classification schemes based on BOD, NH_4^+-N, NO_3^--N, and PO_4^{3-}-P (Lorraine and Vollenweider 1981; Wetzel 2001; Peerapornpisal et al. 2004; Jones and Medrano 2006; PCD 2010), as summarized in Table 6.

To evaluate consistency with conventional water-quality assessment, the developed Diatom Index (DI) was compared with the Water Quality Index (WQI) using Spearman's rank correlation (ρ), as both indices were non-normally distributed. A simple linear trendline and corresponding r^2 value were used for descriptive visualization only and were not treated as inferential regression statistics.

RESULTS AND DISCUSSIONS

Water quality assessment

Water quality characteristics

The mean values of key physico-chemical parameters are summarized in Table 7. Water temperature ranged from 23.0-26.7°C, showing no significant difference among stations ($H_6 = 5.19$, $p = 0.52$). Electrical conductivity (EC: 102.9-143.5 μS/cm), dissolved solids (DS: 57.3-138.1 mg/L), and suspended solids (SS: 3.5-7.1 mg/L) also did not differ significantly among sites ($p > 0.05$). Although pH and Dissolved Oxygen (DO) showed spatial variability, non-parametric tests indicated that these differences were not statistically significant ($p > 0.05$), except for Biochemical Oxygen Demand (BOD), which remained significantly different among stations ($H_6 = 18.93$, $p < 0.01$). Nutrient concentrations were generally low (NH_4^+-N = 0.22-0.92 mg/L; NO_3^--N = 0.43-1.73 mg/L; PO_4^{3-}-P = 0.07-0.11 mg/L) and showed no significant differences among stations ($p > 0.05$), consistent with oligotrophic-mesotrophic conditions. Although most parameters did not differ statistically, gradual upstream-midstream-downstream patterns were evident in several variables (e.g., higher SS and BOD at downstream sites). These patterns are interpreted as descriptive ecological tendencies rather than statistically supported differences, reflecting longitudinal changes in hydrology and land-use influence along the river.

The present study highlights the spatial heterogeneity of physico-chemical water quality in the Yang River and provides insights into how local conditions influence ecological status. Overall, the river exhibits characteristics of a relatively healthy upland stream in Northern Thailand, with low conductivity, low nutrient concentrations, and high Dissolved Oxygen (DO). Because statistical tests indicated that differences in pH and DO among stations were not significant, interpretations of their spatial patterns are made cautiously and emphasize ecological tendencies rather than statistically supported differences. The slightly alkaline pH values (7.2-8.3) are typical of carbonate-buffered rivers in upland areas (Allan and Castillo 2007). The higher pH at ST1 may reflect upstream geology dominated by limestone and carbonate rocks, which enhance buffering capacity. Conversely, lower pH values at ST6-ST7 represent a decreasing tendency rather than a statistically supported decline, and may be ecologically linked to localized organic inputs and microbial decomposition. Similar upstream-downstream pH patterns have been reported in tributaries of the Ping River (Leelahakriengkrai and Peerapornpisal 2010).

Table 7. Mean values of physico-chemical water quality parameters measured at seven sampling stations in the Yang River, Thailand (November 2016-April 2017)

Stations	Parameters									
	Temp (°C)	pH	EC (µS/cm)	DS (mg/L)	SS (mg/L)	DO (mg/L)	BOD (mg/L)	NH_4^+-N (mg/L)	NO_3^--N (mg/L)	PO_4^{3-}-P (mg/L)
ST1	23.0^{a} (±3.9)	8.30^{a} (±0.14)	128.13^{a} (±19.21)	57.32^{a} (±5.29)	3.45^{a} (±2.55)	9.12^{ab} (±1.12)	2.01^{abc} (±0.22)	0.34^{a} (±0.19)	0.51^{a} (±0.34)	0.08^{a} (±0.04)
ST2	23.3^{a} (±2.6)	7.94^{ab} (±0.51)	117.25^{a} (±19.46)	58.15^{a} (±11.33)	4.93^{a} (±1.72)	8.70^{ab} (±0.57)	1.28^{c} (±0.59)	0.30^{a} (±0.25)	0.43^{a} (±0.48)	0.11^{a} (±0.11)
ST3	24.7^{a} (±2.7)	7.69^{abc} (±0.62)	116.09^{a} (±14.82)	61.29^{a} (±2.81)	6.05^{a} (±3.01)	10.04^{ab} (±1.04)	2.20^{ab} (±0.43)	0.26^{a} (±0.24)	0.43^{a} (±0.38)	0.08^{a} (±0.08)
ST4	26.2^{a} (±1.9)	7.40^{bc} (±0.60)	113.09^{a} (±10.59)	59.59^{a} (±13.62)	7.06^{a} (±3.78)	6.55^{c} (±1.51)	3.39^{a} (±2.00)	0.92^{a} (±1.64)	1.73^{a} (±3.30)	0.10^{a} (±0.10)
ST5	26.7^{a} (±1.2)	7.24^{c} (±0.31)	102.94^{a} (±10.26)	65.43^{a} (±6.73)	3.73^{a} (±1.15)	8.85^{ab} (±0.58)	1.55^{bc} (±0.55)	0.22^{a} (±0.19)	0.55^{a} (±0.46)	0.07^{a} (±0.07)
ST6	25.1^{a} (±2.8)	7.17^{c} (±0.24)	141.98^{a} (±44.38)	138.05^{a} (±136.94)	7.01^{a} (±7.21)	13.28^{a} (±3.63)	4.98^{a} (±2.23)	0.40^{a} (±0.26)	0.62^{a} (±0.51)	0.07^{a} (±0.07)
ST7	24.9^{a} (±2.6)	7.24^{bc} (±0.19)	143.53^{a} (±45.64)	79.85^{a} (±18.64)	7.07^{a} (±3.82)	8.57^{bc} (±0.43)	1.56^{bc} (±0.32)	0.43^{a} (±0.26)	0.67^{a} (±0.50)	0.08^{a} (±0.08)

Note: Values are mean ± SD (n = 3). Different letters indicate significant differences among stations ($p < 0.05$, one-way ANOVA, Kruskal-Wallis followed by Dunn's post hoc); H and p values were obtained from one-way ANOVA, Kruskal-Wallis (Dunn's post hoc, $p < 0.05$)

Table 8. Nutrient classification of water quality at seven sampling stations in the Yang River, Thailand (November 2016-April 2017), derived from BOD, NH_4^+-N, NO_3^--N, and PO_4^{3-}-P according to Table 6

Stations	Nov 2016	Dec 2016	Jan 2017	Feb 2017	Mar 2017	Api 2017
ST1	-	3	-	4	3	3
ST2	3	3	-	4	3	3
ST3	3	4	-	4	4	4
ST4	3	4	-	4	4	4
ST5	4	4	-	4	4	-
ST6	4	4	4	4	4	-
ST7	4	4	4	4	4	-

Note: Numbers represent trophic classes (see Table 6), ranging from 1: Ultra-oligotrophic to 7: Hypertrophic. "-" indicates no sample collected

DO concentrations (6.6-13.3 mg/L) were generally high throughout the study area. Although DO differences among stations were not statistically significant, higher DO at upstream and midstream sites is consistent with hydrological features such as turbulence or riffles, while lower DO at ST4 may reflect enhanced organic decomposition. Comparable localized decreases in DO have been observed in the Yom River (Yana et al. 2013). The most striking and statistically supported difference among stations was the significantly higher BOD at ST6 (4.98 mg/L). According to the PCD (2010), BOD values exceeding 4 mg/L indicate organic pollution. Elevated BOD at ST6 suggests the influence of anthropogenic discharges, consistent with increases in dissolved and suspended solids observed at the same site. Elevated organic loading not only reduces water quality but also creates conditions favourable for eutrophication, potentially altering diatom assemblages and favouring tolerant taxa (Kelly et al. 1998). Comparable findings were reported in the Lower Mekong Basin, where agricultural intensification increased nutrient loads and eutrophication risks (Sor et al. 2021). Nutrient concentrations (NH_4^+-N, NO_3^--N, PO_4^{3-}-P) were low overall, placing the Yang River within oligotrophic to mesotrophic conditions. Although spatial differences were not significant, the relatively higher NO_3^--N at ST4 suggests potential localized fertilizer inputs. Even though these values remain below eutrophication thresholds, they highlight future risks if land use intensifies (Carpenter et al. 1998).

Taken together, the statistical analysis indicates that BOD was the only parameter showing significant spatial variation, whereas other water-quality variables varied in ecologically meaningful but statistically non-significant ways. These parameters nevertheless remain important for interpreting diatom-based indices, as diatom assemblages respond sensitively to pH, oxygen availability, and organic pollution (van Dam et al. 1994; Stevenson et al. 2010). Thus, ecological interpretations are informative for understanding longitudinal patterns in the Yang River, while statistical conclusions are limited to BOD as the sole parameter with significant spatial differences. The lack of statistically significant differences among stations for most variables may partly reflect the relatively short longitudinal distances between some sampling sites, which can result in similar physico-chemical conditions, as reported in other riverine studies. In this context, the observed patterns emphasize gradual longitudinal transitions rather than abrupt spatial contrasts.

Nutrient classification of water quality

Based on nutrient thresholds (Table 6, the trophic classification of the Yang River is summarized in Table 8. Most stations were consistently assigned to Class 3 (oligo-mesotrophic) or Class 4 (mesotrophic) throughout the study period. Upstream sites (ST1-ST3) were predominantly classified as Class 3, with occasional shifts to Class 4 during December and March. Midstream and downstream sites (ST5-ST7) were consistently classified as Class 4 across most sampling months. ST6 and ST7 in particular remained mesotrophic (Class 4) throughout the monitoring period. Although the trophic classifications suggest a longitudinal tendency from upstream to downstream, no nutrient parameter (NH_4^+-N, NO_3^--N, PO_4^{3-}-P) differed significantly among stations ($p > 0.05$). Therefore, the observed trophic patterns represent descriptive ecological tendencies rather than statistically distinct nutrient regimes.

The nutrient classification indicates that the Yang River generally falls within oligo-mesotrophic to mesotrophic conditions, reflecting relatively low nutrient enrichment typical of upland rivers in Northern Thailand. The upstream stations (ST1-ST3) maintained comparatively better water quality (Class 3), which is consistent with reduced human activities and greater forest cover in headwater areas. In contrast, downstream stations (ST5-ST7) were consistently classified as mesotrophic (Class 4), likely influenced by agricultural runoff and local settlements. Because nutrient concentrations did not differ significantly among stations, these upstream-downstream shifts in trophic class should be interpreted as ecological gradients rather than statistically confirmed differences. Nevertheless, the observed pattern reflects typical longitudinal transitions reported in tropical and subtropical rivers, where nutrient availability tends to increase downstream as land-use intensity rises (PCD 2010). Similar longitudinal tendencies have been reported in tributaries of the Mekong Basin affected by agricultural expansion (Sor et al. 2021), tropical urban and peri-urban waters influenced by runoff in Malaysia (Malek et al. 2012), and Southeast Asian rivers where regionally calibrated diatom-based indices have been applied for water-quality assessment (Ladera et al. 2025). These consistent regional observations support the ecological interpretation that gradual increases in nutrient availability, even when not statistically distinct at the site scale, can contribute to downstream changes in river condition. Such trophic gradients provide important contextual information for subsequent multivariate analyses (e.g., PCA and CCA), in which nutrient-related variables often contribute to major environmental gradients. Furthermore, because nutrient availability is a key determinant of algal and diatom community structure, the observed trophic patterns offer a useful framework for interpreting biological responses, even in the absence of statistically significant nutrient differences among stations.

Principal Component Analysis (PCA) of water quality

Principal Component Analysis (PCA) was conducted using nine water-quality variables (Temp, pH, EC, DS, SS, DO, BOD, NO_3^--N, and PO_4^{3-}-P) after removing NH_4^+-N due to its strong redundancy with NO_3^--N. All variables were standardized (z-scores) prior to ordination to eliminate scale effects. The first two principal components explained 47.30% of the total variance (PC1 = 26.44%, PC2 = 20.86%), capturing the major environmental gradients observed in the Yang River (Table 9 and Figure 2). PC1 represented a nutrient-organic matter gradient, with strong positive loadings for EC, BOD, and NO_3^--N, and negative loadings for pH and PO_4^{3-}-P. Stations located on the positive side of PC1 (e.g., ST4 Feb 2017; ST6 Feb-Mar 2017) corresponded to periods associated with higher nutrient concentrations and organic enrichment, whereas upstream stations (ST1-ST2) consistently plotted on the negative side, reflecting low conductivity, low BOD, and oligotrophic conditions. PC2 represented a secondary environmental gradient related primarily to dissolved and suspended solids and oxygen availability, driven by positive loadings of DS, SS, and DO. Upstream samples tended to plot higher on PC2, whereas downstream samples (ST6-ST7) were more frequently located toward lower PC2 values during certain months. Although no formal statistical tests were applied to evaluate group separation, the ordination indicated a gradual environmental transition along the river rather than distinct clustering among stations. In general, upstream stations (ST1-ST2) were positioned toward low-nutrient, low-BOD conditions; midstream stations (ST3-ST5) occupied intermediate positions; and downstream stations (ST6-ST7) showed shifts toward higher nutrient and BOD influence along PC1, particularly during the dry season when dilution is reduced. Episodic enrichment events were evident at ST4 and ST6 during February-March 2017, where samples shifted toward the positive side of PC1, indicating increased nutrient accumulation and organic loading. The exclusion of NH_4^+-N did not alter the overall PCA configuration, as the dominant gradients remained driven by EC, BOD, NO_3^--N, SS, and DO, indicating consistency of the ordination after multicollinearity screening.

Table 9. PCA loadings of water-quality variables on the first two principal components (PC1 and PC2), with eigenvalues and percentage variance explained

Variable	PC1	PC2
Temp (°C)	-0.008	0.249
pH	-0.407	0.064
EC (µS/cm)	0.440	-0.061
DS (mg/L)	0.145	0.597
SS (mg/L)	-0.083	0.541
DO (mg/L)	0.393	0.342
BOD (mg/L)	0.478	0.079
NO_3^--N (mg/L)	0.255	-0.362
PO_4^{3-}-P (mg/L)	-0.406	0.165
Eigenvalue	2.380	1.877
Variance (%)	26.44	20.86

Note: Loadings with an absolute value greater than 0.4 are shown in bold to indicate strong contributions to each axis. PC1 and PC2 together explained 47.3% of the total variance

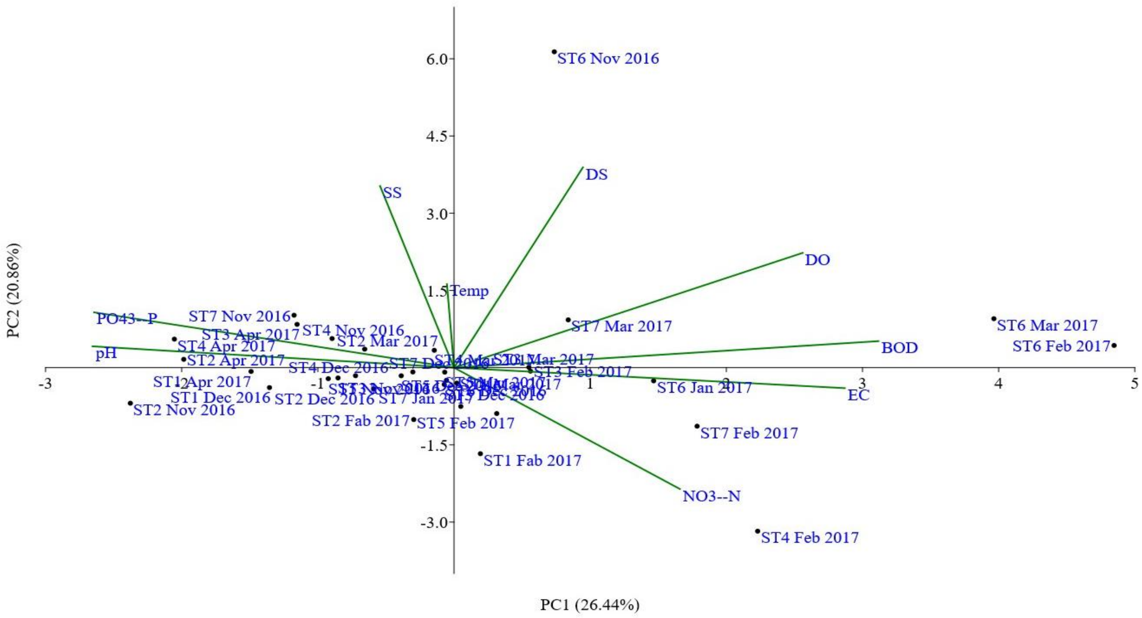

Figure 2. PCA biplot of water-quality variables at the seven sampling stations in the Yang River, Thailand (Nov 2016-Apr 2017). Vectors represent environmental variables including temperature (Temp), pH, Electrical Conductivity (EC), Dissolved Solids (DS), Suspended Solids (SS), Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), and nutrient parameters (NO_3^--N, PO_4^{3-}-P)

The PCA ordination revealed ecologically interpretable gradients in the Yang River, reflecting both longitudinal variation and seasonal hydrological influences. PC1, explaining 26.44% of the total variance, represented a nutrient-organic enrichment gradient dominated by EC, BOD, and NO_3^--N, while PC2 (20.86%) captured a secondary gradient associated with solids and oxygen-related variables (DS, SS, and DO). Because differences among stations were not statistically significant, interpretations of PCA axes are considered descriptive patterns rather than evidence of distinct group separation. Nevertheless, the observed ordination patterns align with expected upstream-downstream transitions in tropical river systems. Upstream stations (ST1-ST2) plotted toward the negative side of PC1 and higher values of PC2, indicating relatively oligotrophic and well-oxygenated conditions typical of upland headwaters (Vannote et al. 1980; Allan and Castillo 2007). Midstream stations (ST3-ST5) occupied intermediate positions along both axes, reflecting transitional physicochemical conditions influenced by mixed land use. Downstream stations (ST6-ST7) tended toward the nutrient- and BOD-enriched side of PC1, particularly during February-March, indicating periods of increased nutrient accumulation and organic loading. Although these downstream tendencies were not supported by statistically significant differences, they resemble longitudinal patterns reported in other tropical river systems influenced by agricultural intensification, such as the Lower Mekong Basin (Sor et al. 2021). The stability of the PCA structure following removal of NH_4^+-N supports the interpretation that the ordination is driven by broader nutrient and organic gradients rather than by redundancy among nitrogen species. Overall, the PCA illustrates a gradual longitudinal gradient from oligotrophic headwaters to more nutrient-influenced downstream reaches. These tendencies are consistent with the River Continuum Concept (Vannote et al. 1980), while seasonal shifts particularly at ST4 and ST6 during the dry season highlight the sensitivity of tropical rivers to low-flow conditions that can enhance nutrient and organic matter accumulation.

Structure and ecological patterns of benthic diatom assemblages

Diatom assemblage composition

A total of 154 diatom species were identified from the Yang River, belonging to 3 classes, 14 orders, 26 families, and 47 genera. Among these, 29 taxa were common, with relative abundances exceeding 1% across samples (Table 10).

Table 10. List of benthic diatom taxa with relative abundance >1% across all sampling stations and months in the Yang River, Thailand, including species abbreviations, mean relative abundance, and frequency of occurrence

Species name	Abbreviation	Relative abundance (%)	Frequency (%)
Achnanthes oblongella Østrup, 1902	AchObl	3.28	66.67
Achnanthes suchlandtii Hustedt, 1933	AchSuc	2.05	57.58
Achnanthidium caravelense M.E.Novais & L.Ector, 2011	AchCar	3.23	75.76
Achnanthidium minutissimum (Kützing) Czarnecki, 1994	AchMin	4.83	78.79
Achnanthidium pusillum Czarnecki & Edlund, 1995	AchPus	1.38	63.64
Achnanthidium straubianum (Lange-Bertalot) Lange-Bertalot, 1999	AchStr	2.72	72.73
Cocconeis placentula Ehrenberg, 1838	CocPla	5.19	90.91
Cymbella affinis Kützing, 1844	CymAff	1.05	57.58
Cymbella minuta Hilse, 1862	CymMin	1.62	66.67
Cymbella parva (W.Smith) Kirchner, 1878	CymPar	1.21	63.64
Cymbella tumida (Brébisson) Van Heurck, 1880	CymTum	1.71	69.70
Cymbella turgidula Grunow, 1875	CymTur	4.87	100.00
Diadesmis confervacea Kützing, 1844	DiaCon	5.90	66.67
Encyonema minutum (Hilse) D.G.Mann, 1990	EncMin	1.13	63.64
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller, 1997	EolMin	3.69	96.97
Gomphonema clevei Fricke, 1902	GomCle	1.55	72.73
Gomphonema lagenula Kützing, 1844	GomLag	2.47	93.94
Mayamaea atomus (Kützing) Lange-Bertalot, 1997	MayAto	2.98	75.76
Navicula capitatoradiata H.Germain ex Gasse	NavCap	1.56	69.70
Navicula cryptocephala Kützing	NavCry	3.09	90.91
Navicula cryptotenella Lange-Bertalot	NavCrypnel	2.43	90.91
Navicula suprinii Gerd Moser, Lange-Bertalot & Metzeltin	NavSup	1.26	63.64
Nitzschia palea (Kützing) W.Smith, 1856	NitPal	3.25	87.88
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot, 1999	PlaFre	3.27	90.91
Planothidium lanceolatum (Brébisson ex Kützing) Lange-Bertalot	PlaLan	1.95	81.82
Planothidium rostratum (Østrup) Lange-Bertalot, 1999	PlaRos	1.24	72.73
Rhopalodia gibberula (Ehrenberg) O.Müller, 1895	RhoGib	1.68	27.27
Synedra ulna var. *aequalis* (Kützing) Brun, 1880	SynUln	2.73	63.64
Ulnaria lancleolata (Kützing) P.Compère, 2001	UlnLan	1.05	57.58

Note: Relative abundance is expressed as the percentage of total frustules counted. Frequency of occurrence is expressed as the percentage of samples (n = 33) in which the taxon was present

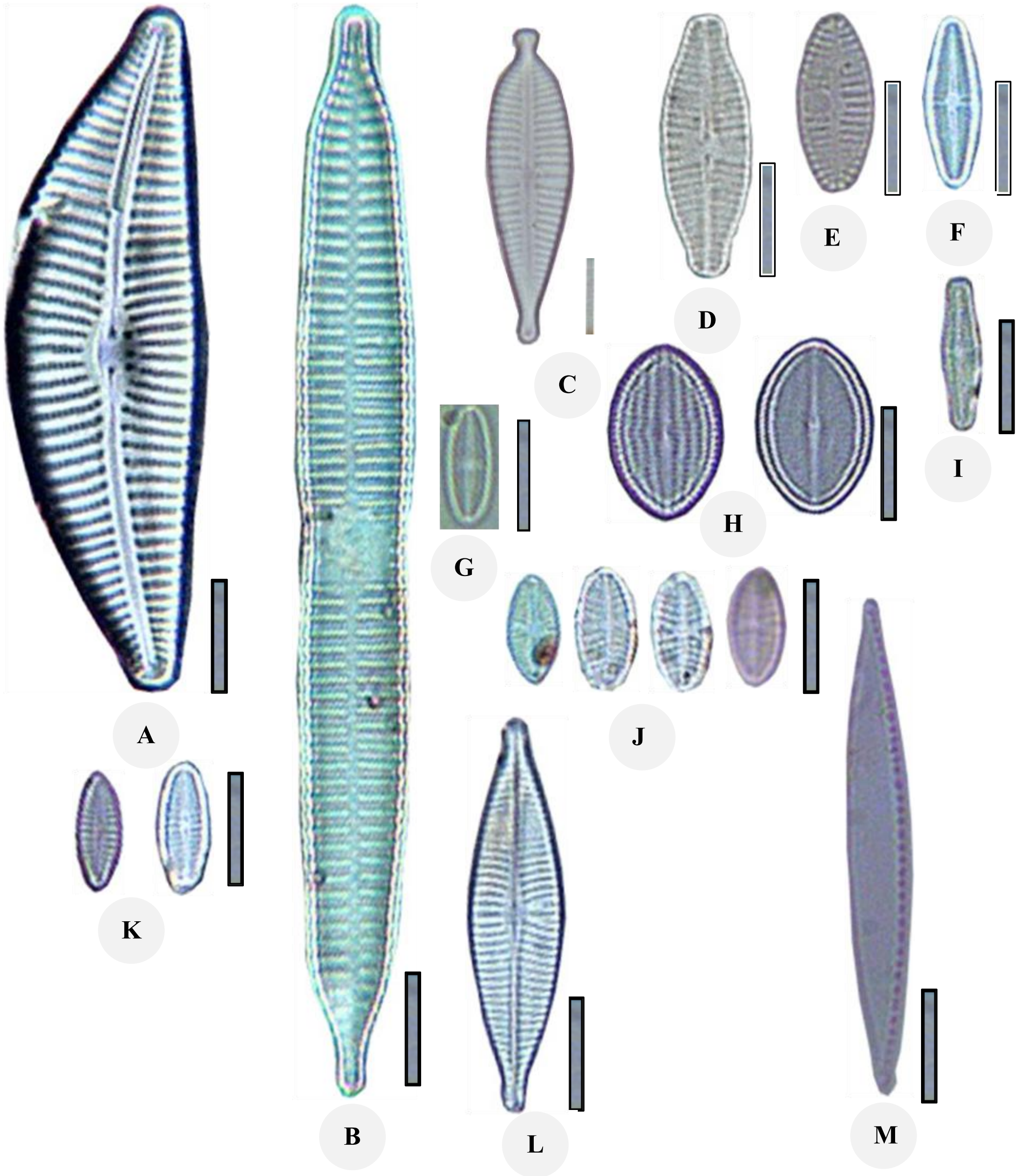

Figure 3. Light micrographs of selected dominant benthic diatom taxa observed in the Yang River, Thailand. Scale bar = 10 µm. A. *Cymbella turgidula*, B. *Synedra ulna* var. *aequalis*, C. *Gomphonema lagenula*, D. *Planothidium frequentissimum*, E. *Planothidium lanceolatum*, F. *Achnanthidium caravelense*, G. *Achnanthidium straubianum*, H. *Cocconeis placentula*, I. *Achnanthidium minutissimum*, J. *Mayamaea atomus*, K. *Eolimna minima*, L. *Navicula capitatoradiata*, M. *Nitzschia palea*

The most abundant species included *Cocconeis placentula* Ehrenberg, 1838 (5.19%), *Diadesmis confervacea* Kützing, 1844 (5.90%), and *Cymbella turgidula* Grunow, 1875 (4.87%). Other frequently occurring taxa were *Achnanthidium minutissimum* (Kützing) Czarnecki, 1994 (4.83%), *Planothidium frequentissimum* (Lange-Bertalot) Lange-Bertalot, 1999 (3.27%), *Nitzschia palea* (Kützing) W.Smith, 1856 (3.25%), *Achnanthes oblongella* Østrup, 1902 (3.28%), and *Eolimna minima* (Grunow) Lange-Bertalot & W.Schiller, 1997 (3.69%). Species such as *C. turgidula*, *Navicula cryptocephala* Kützing, and *P. frequentissimum* exhibited both relatively high abundance and broad frequency of occurrence across stations, indicating ecological relevance along the upstream-downstream gradient. Representative light micrographs of selected dominant taxa are presented in Figure 3.

The diatom assemblages of the Yang River were dominated by benthic taxa such as *A. minutissimum*, *P. frequentissimum*, *E. minima*, and *C. turgidula*. These taxa are well adapted to lotic environments with stable substrates and moderate nutrient availability, and their dominance is consistent with reports from upland streams in Northern Thailand (Yana et al. 2013) and other tropical Asian rivers, including studies in the Philippines (Ladera et al. 2025) and the Mekong River basin in Thailand (Moonsin et al. 2013). In such systems, small adnate or low-profile taxa commonly dominate under oligo-mesotrophic to mesotrophic conditions, reflecting relatively high current velocity and limited fine sediment accumulation. Similar assemblage structures have also been reported from temperate European rivers, where adnate or low-profile diatoms are characteristic of moderately enriched but well-oxygenated environments (van Dam et al. 1994; Kelly et al. 2007). The prevalence of *C. placentula* and *D. confervacea* further supports this interpretation, as both species are common epilithic forms tolerant of moderate nutrient enrichment and physical disturbance (Taylor et al. 2007). The presence of more tolerant taxa, such as *N. palea* and *N. cryptocephala*, occurring at moderate abundance, suggests localized anthropogenic influence in midstream and downstream sections. These species are widely recognized as indicators of organic enrichment and are frequently associated with elevated biochemical oxygen demand and nutrient inputs (van Dam et al. 1994). Their occurrence in the Yang River is consistent with episodic nutrient enrichment patterns reported for other Northern Thailand rivers, including the Yom River (Yana et al. 2013). Similar assemblage patterns where sensitive oligotrophic taxa co-occur with more tolerant meso-eutrophic indicators have been documented in tropical rivers in the Philippines (Ladera et al. 2025) and in Malaysian urban and peri-urban waters affected by turbidity and organic runoff (Malek et al. 2012). The coexistence of taxa indicative of both low and moderate nutrient conditions in the Yang River is consistent with its physicochemical classification as oligo-mesotrophic to mesotrophic. The relatively high species richness (154 taxa) reflects considerable habitat heterogeneity along the river continuum, with upstream reaches supporting assemblages dominated by oligotrophic and low-profile taxa, while downstream reaches include a higher proportion of tolerant species associated with increased nutrient availability. These patterns align with broader stream ecology concepts emphasizing longitudinal gradients of enrichment and disturbance in structuring benthic algal communities (Allan and Castillo 2007). Overall, the assemblage composition suggests a river system experiencing moderate nutrient pressure without evidence of severe ecological degradation, providing a suitable biological basis for the development of a regionally calibrated Diatom Index for tropical rivers.

Diversity indices of diatom assemblages

Species richness and diversity indices of diatom assemblages are summarized in Table 11. The number of species ranged from 42.5±6.0 taxa at ST1 to 53.0±4.5 taxa at ST7, indicating moderate richness across all stations. Shannon-Wiener Diversity Index (H′) values varied between 2.88±0.15 at ST5 and 3.26±0.33 at ST4. According to the classification criteria (Table 5), most stations fell within the range of 3.00-3.90, indicating lightly polluted conditions, whereas ST5 (H′ < 3.00) was classified as moderately polluted. Simpson's Dominance Index (D) showed relatively low values (0.07-0.12), reflecting moderate dominance patterns, whereas evenness (J′) was consistently high (0.77-0.84), suggesting a relatively balanced distribution of taxa among stations.

Statistically, the variation in diversity indices among stations was relatively limited. Species richness and Shannon-Wiener diversity showed overlapping ranges across sites, indicating no clear spatial differentiation. Although some stations exhibited slightly lower H′ values, these differences were not sufficiently large to indicate statistically distinct diversity patterns among stations. Overall, the diversity indices suggest broadly comparable community structure along the river, despite localized differences in taxonomic composition.

Spearman's rank correlation analysis (Table 12) indicated that none of the diversity indices (species richness, H′, J′, or D) exhibited statistically significant correlations with any measured water-quality parameter ($p > 0.05$). Although several relationships showed positive or negative tendencies (e.g., species richness with NO_3^--N, dominance with BOD and DO), these trends were weak and not statistically significant.

Table 11. Diversity indices of benthic diatom assemblages at the seven sampling stations in the Yang River, Thailand, including number of species (S), Shannon-Wiener Diversity Index (H′), Simpson's Dominance Index (D), And Pielou's Evenness Index (J′)

Stations	Diversity indices			
	Number of species	H′	D	J′
ST1	42.5^a(±6.0)	3.19^a(±0.14)	0.07^a(±0.02)	0.84^a(±0.02)
ST2	43.8^a(±5.9)	3.04^a(±0.27)	0.08^a(±0.03)	0.80^a(±0.05)
ST3	50.8^a(±6.3)	3.02^a(±0.20)	0.11^a(±0.06)	0.77^a(±0.06)
ST4	52.0^a(±13.0)	3.26^a(±0.33)	0.07^a(±0.03)	0.83^a(±0.04)
ST5	43.0^a(±3.6)	2.88^a(±0.15)	0.12^a(±0.02)	0.77^a(±0.03)
ST6	47.6^a(±11.5)	3.11^a(±0.52)	0.09^a(±0.06)	0.81^a(±0.09)
ST7	53.0^a(±4.5)	3.17^a(±0.33)	0.09^a(±0.04)	0.80^a(±0.07)

Note: Values are expressed as mean ± SD. Different letters indicate significant differences at $p < 0.05$. Water quality classes based on Shannon-Wiener Diversity Index (H′): 0.00-0.90 = Very heavily polluted; 1.00-1.90 = Heavily polluted; 2.00-2.90 = Moderately polluted; 3.00-3.90 = Lightly polluted; ≥4.00 = Unpolluted; Values are mean ± SD (n = 3). Different letters indicate significant differences among stations ($p < 0.05$, one-way ANOVA for Number of species, H′ and D, and $p < 0.05$, one-way ANOVA, Kruskal-Wallis for J′); F and p values were obtained from one-way ANOVA (Tukey HSD, $p < 0.05$ for Number of species, H′ and D, and Dunn's post hoc, $p < 0.05$ for J′)

Table 12. Spearman's rho coefficients (ρ) between water quality parameters and diatom diversity indices in Yang River, Thailand

Variable	No. of species	H′	1-D	J′
Temp (°C)	0.10^{ns}	0.10^{ns}	0.12^{ns}	0.13^{ns}
pH	-0.25^{ns}	0.03^{ns}	0.26^{ns}	0.20^{ns}
EC (µS/cm)	0.10^{ns}	0.23^{ns}	0.24^{ns}	0.24^{ns}
DS (mg/L)	0.33^{ns}	0.28^{ns}	0.21^{ns}	0.16^{ns}
SS (mg/L)	0.10^{ns}	0.06^{ns}	0.09^{ns}	0.07^{ns}
DO (mg/L)	-0.13^{ns}	-0.20^{ns}	-0.11^{ns}	-0.18^{ns}
BOD (mg/L)	0.03^{ns}	-0.14^{ns}	-0.14^{ns}	-0.17^{ns}
NH_4^+-N (mg/L)	0.29^{ns}	0.32^{ns}	0.20^{ns}	0.22^{ns}
NO_3^--N (mg/L)	0.26^{ns}	0.23^{ns}	0.11^{ns}	0.13^{ns}
PO_4^{3-}-P (mg/L)	-0.12^{ns}	-0.18^{ns}	-0.11^{ns}	-0.16^{ns}

Note: Values are Spearman's ρ; significance indicated at * $p < 0.05$, ** $p < 0.01$, ns: Not significant (Bonferroni corrected)

Table 13. Relative abundance (%) of ecological guilds of benthic diatoms at seven sampling stations in the Yang River, Thailand

Stations	Ecological guilds of diatom		
	Low-profile (%)	High-profile (%)	Motile (%)
ST1	60.84^{a}(±3.46)	12.23^{b}(±6.40)	26.93^{abc}(±5.89)
ST2	65.44^{a}(±10.77)	16.02^{b}(±7.02)	18.54^{bc}(±5.05)
ST3	27.93^{b}(±18.60)	56.10^{a}(±21.28)	15.98^{c}(±6.82)
ST4	27.02^{b}(±11.36)	34.79^{a}(±5.26)	38.19^{a}(±11.21
ST5	27.96^{b}(±21.42)	46.81^{a}(±28.19)	25.24^{abc}(±11.26)
ST6	29.54^{b}(±4.68)	29.81^{a}(±17.19)	40.65^{a}(±13.86)
ST7	49.46^{a}(±9.19)	17.96^{b}(±9.08)	32.58^{ab}(±10.82)

Note: Values are mean ± SD (n = 3). Different letters indicate significant differences among stations ($p < 0.05$, one-way ANOVA followed by Whitney pairwise for Low-profile, and $p < 0.05$, one-way ANOVA, Kruskal-Wallis followed by Dunn's post hoc for High-profile and Motile); Welch F and p values were obtained from one-way ANOVA (Whitney pairwise, $p < 0.05$) for Low-profile; H and p values were obtained from one-way ANOVA, Kruskal-Wallis (Dunn's post hoc, $p < 0.05$) for High-profile and Motile

The diversity indices of diatom assemblages in the Yang River indicate an ecosystem experiencing moderate nutrient influence while remaining ecologically functional. Shannon-Wiener Diversity Index (H′) values (approximately 2.9-3.3) generally correspond to lightly polluted conditions based on established classification criteria (Jørgensen et al. 2010), with one station (ST5) falling within the moderately polluted category. Similar H′ values have been reported from other rivers in Northern Thailand, such as the Yom and Ping Rivers (Yana et al. 2013; Itayama et al. 2015), and from tropical rivers elsewhere in Southeast Asia. Globally, Shannon-Wiener values between approximately 2.5 and 3.5 are often associated with mesotrophic conditions and moderate anthropogenic disturbance (Kelly et al. 2007). The relatively high evenness values (J′ = 0.77-0.84) observed in the Yang River further indicate that diatom assemblages were not strongly dominated by a few taxa, while low Simpson's Dominance Index (D) values reflect a balanced community structure. Despite these spatial differences in diversity indices, correlation analysis showed that none of the measured physicochemical parameters were significantly associated with species richness or diversity metrics (Table 12). This suggests that variation in diatom diversity along the Yang River is not driven by single environmental variables such as nutrients, pH, or BOD alone. Instead, diversity patterns likely reflect the combined influence of multiple interacting factors, including moderate nutrient availability, hydrological variability, substrate heterogeneity, and local habitat conditions. The absence of statistically significant correlations may also be explained by the relatively narrow range of water-quality conditions observed across stations, where parameters such as BOD, NH_4^+-N, and PO_4^{3-}-P remained at low to moderate levels. Under such conditions, diversity indices may respond only weakly to small environmental fluctuations. Importantly, the lack of significant correlations does not negate the ecological relevance of the observed diversity patterns, as diatom communities integrate environmental influences over time and respond to cumulative rather than instantaneous conditions. Taken together, these findings indicate that diatom assemblages in the Yang River exhibit subtle but ecologically meaningful variation in diversity structure. While no single water-quality parameter exerted a strong statistical influence on diversity indices, the overall pattern of moderate diversity and high evenness is consistent with rivers experiencing mild nutrient inputs and limited organic pollution. This further supports the use of diatom-based metrics as integrative indicators of ecological condition in tropical river systems, where biological responses may not align directly with individual physicochemical variables.

Ecological guilds of diatoms

The distribution of diatom ecological guilds varied significantly among stations in the Yang River (Table 13). Low-profile taxa dominated upstream stations, comprising 60.84±3.46% at ST1 and 65.44±10.77% at ST2, and were significantly higher than values observed at several mid- and downstream stations ($p < 0.05$). High-profile taxa were most abundant at midstream sites, reaching 56.10±21.28% at ST3 and 46.81±28.19% at ST5, with significantly higher proportions compared to upstream stations ($p < 0.05$). Motile taxa showed the highest relative abundance at downstream stations, particularly at ST6 (40.65±13.86%), where values were significantly higher than those recorded upstream and at some midstream sites ($p < 0.05$). However, Spearman's rank correlation analysis (Table 14) indicated that none of the ecological guilds (low-profile, high-profile, or motile) exhibited statistically significant correlations with any measured physicochemical parameter ($p > 0.05$). Although weak tendencies were observed for certain variables (e.g., low-profile taxa with pH; motile taxa with SS and BOD), these relationships were not statistically significant and are therefore interpreted as descriptive ecological trends rather than quantitative associations.

The dominance of low-profile taxa at upstream stations (ST1-ST2) reflects adaptation to oligotrophic and fast-flowing conditions, where adnate species such as *A. minutissimum* and *P. frequentissimum* efficiently exploit light and nutrients on stable substrates. This pattern is consistent with observations from upland rivers in Northern Thailand (Yana et al. 2013) and aligns with the ecological guild framework proposed by Passy (2007), in which low-profile taxa are characteristic of nutrient-poor and high-disturbance habitats. Although no statistically significant correlations were detected between low-profile taxa and individual physicochemical parameters (Table 14), their upstream dominance is ecologically consistent with the combined influence of low nutrient availability, higher current velocity, and carbonate-buffered conditions typical of headwater systems. The higher abundance of high-profile taxa at midstream stations (ST3-ST5) suggests moderate nutrient availability and increased habitat stability. High-profile genera such as *Cymbella* and *Gomphonema* possess upright growth forms that allow them to exploit light and nutrient gradients in less turbulent waters (Rimet and Bouchez 2012). Similar increases in high-profile or attached taxa have been reported in mesotrophic tropical rivers of the Philippines (Ladera et al. 2025) and in European stream ecosystems under moderate nutrient enrichment and stable flow conditions (Rimet and Bouchez 2012; Rimet et al. 2019). Although high-profile taxa did not show significant correlations with nutrients in Table 14, their midstream prevalence is still consistent with guild theory, reflecting ecological responses to moderate enrichment and stable substrata. Motile genera such as *Nitzschia* and *Navicula* are widely recognized for their tolerance to organic enrichment and their ability to migrate across substrates to exploit localized resource patches (Kelly et al. 2007). While motile taxa did not exhibit statistically significant correlations with BOD or SS in Table 14, their downstream prevalence remains ecologically informative and consistent with guild-based interpretations of river condition. Comparable patterns of motile-guild dominance under organic loading have been documented in tropical urban waters of Malaysia (Malek et al. 2012) and in lowland rivers of Europe (Rimet and Bouchez 2012). Taken together, the distribution of ecological guilds suggests a longitudinal gradient in the Yang River, characterized by upstream dominance of low-profile taxa, midstream enrichment of high-profile taxa, and increased representation of motile taxa downstream. Although these patterns are not driven by statistically significant correlations with single physicochemical variables, they reflect the integrated response of diatom assemblages to multiple interacting environmental factors, including nutrient availability, flow regime, substrate stability, and organic inputs. This finding reinforces the utility of ecological guilds as robust indicators for detecting broad-scale patterns of enrichment and disturbance in tropical river systems.

Canonical Correspondence Analysis (CCA)

Canonical Correspondence Analysis (CCA) was conducted to evaluate the relationships between benthic diatom assemblages and environmental gradients in the Yang River. Following multicollinearity screening using Spearman's correlations and Variance Inflation Factor (VIF) values, ammonium (NH_4^+-N) was excluded, and nine independent variables (Temp, pH, EC, DS, SS, DO, BOD, NO_3^--N, and PO_4^{3-}-P) were retained for the final CCA. The first two canonical axes explained 32.97% and 21.72% of the variation, respectively (cumulative = 54.69%), representing the main environmental gradients structuring diatom assemblages in the Yang River (Figure 4). CCA axis 1 was primarily associated with organic pollution and nutrient enrichment, showing strong relationships with BOD, NO_3^--N, PO_4^{3-}-P, and EC, whereas CCA axis 2 was related mainly to oxygen availability and suspended solids, as indicated by DO and SS vectors. Diatom taxa showed distinct positioning along these gradients: (i) Taxa associated with organic enrichment and nitrate loading included *N. palea* (NitPal), *N. cryptocephala* (NavCry), *E. minima* (EolMin), and *Rhopalodia gibberula* (Ehrenberg) O.Müller (RhoGib), which were oriented toward higher BOD, NO_3^--N, and PO_4^{3-}-P values; (ii) Taxa associated with well-oxygenated upstream conditions included low-profile species such as *A. oblongella* (AchObl), *A. minutissimum* (AchMin), *C. placentula* (CocPla), and *Planothidium lanceolatum* (Brébisson ex Kützing) Lange-Bertalot (PlaLan), which were positioned toward higher DO and lower BOD; (iii) Taxa associated with moderate nutrient availability included high-profile species such as *Cymbella tumida* (Brébisson) Van Heurck,1880 (CymTum), *C. turgidula* (CymTur), *Cymbella affinis* Kützing, 1844 (CymAff), and *Gomphonema clevei* Fricke, 1902 (GomCle), located near EC and PO_4^{3-}-P vectors.

Table 14. Spearman's rho coefficients (ρ) between ecological guilds of benthic diatoms and water quality parameters in the Yang River, Thailand

Variable	Low-profile (%)	High-profile (%)	Motile (%)
Temp (°C)	-0.31^{ns}	0.22^{ns}	0.11^{ns}
pH	0.31^{ns}	-0.15^{ns}	-0.29^{ns}
EC (µS/cm)	0.23^{ns}	-0.22^{ns}	0.04^{ns}
DS (mg/L)	-0.09^{ns}	0.07^{ns}	0.16^{ns}
SS (mg/L)	-0.20^{ns}	0.15^{ns}	0.16^{ns}
DO (mg/L)	0.09^{ns}	-0.16^{ns}	-0.19^{ns}
BOD (mg/L)	-0.49^{ns}	0.27^{ns}	0.26^{ns}
NH_4^+-N (mg/L)	0.11^{ns}	-0.06^{ns}	-0.00^{ns}
NO_3^--N (mg/L)	0.03^{ns}	-0.05^{ns}	0.08^{ns}
PO_4^{3-}-P (mg/L)	0.22^{ns}	-0.06^{ns}	-0.19^{ns}

Note: $p < 0.05$ indicates significant correlations, ns: Not significant

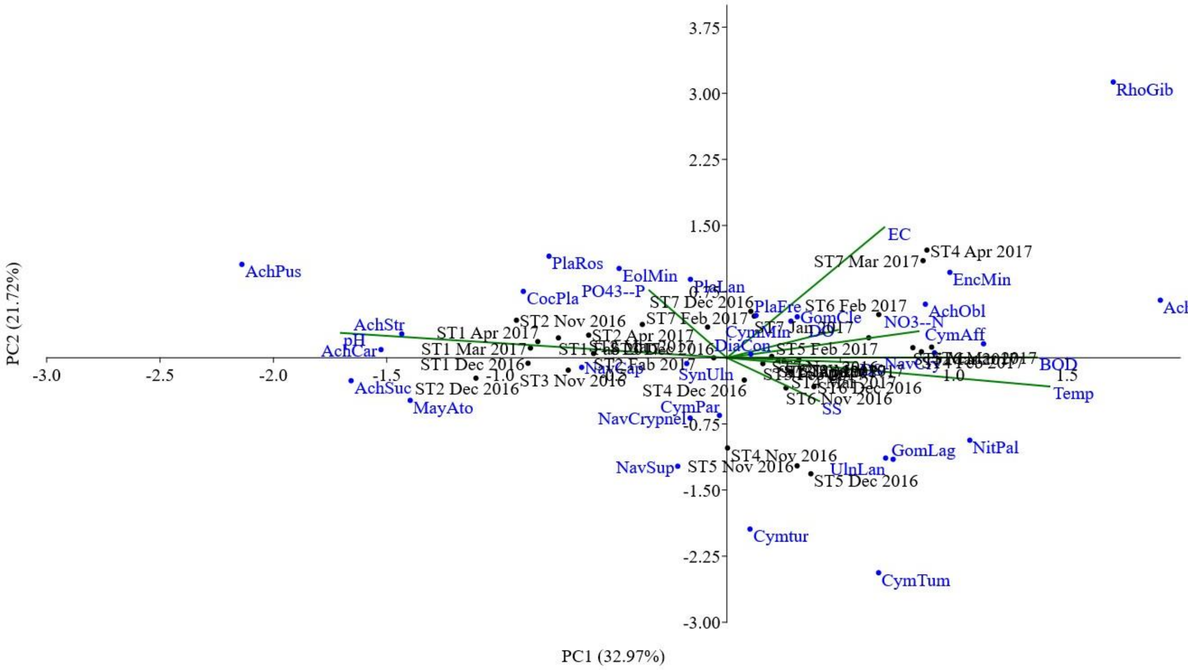

Figure 4. Canonical Correspondence Analysis (CCA) biplot showing relationships between benthic diatom assemblages and environmental variables in the Yang River. Vectors represent environmental variables including water temperature (Temp), pH, Electrical Conductivity (EC), Dissolved Solids (DS), Suspended Solids (SS), Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD) and nutrient parameters (NO_3^--N, and PO_4^{3-}-P)

Spatially, station scores followed a longitudinal tendency: (i) Upstream stations (ST1-ST2) were positioned near DO and low BOD; (ii) Midstream stations (ST3-ST5) were associated with EC and PO_4^{3-}—P; (iii) Downstream stations (ST6-ST7) were oriented toward higher BOD, NO_3^--N, and SS, corresponding to increased organic and nutrient influence.

The CCA results indicate that benthic diatom assemblages in the Yang River are structured primarily along gradients of organic loading and nutrient availability, with additional influence from conductivity, oxygen availability, and suspended solids. The first canonical axis reflects a combined gradient of organic pollution and nutrient enrichment, while the second axis represents variation related to oxygenation and particulate matter. Motile taxa such as *N. palea*, *N. cryptocephala*, and *E. minima* were closely associated with higher BOD, NO_3^--N, and PO_4^{3-} -P values, consistent with their documented tolerance to organic pollution and nutrient-enriched conditions. Similar associations have been reported from rivers in Northern Thailand (Yana et al. 2013) and from temperate systems (Rimet and Bouchez 2012), supporting their use as indicators of degraded conditions. Conversely, low-profile taxa including *A. minutissimum*, *C. placentula*, and *P. lanceolatum* were associated with higher DO and lower BOD, reflecting preference for well-oxygenated and less impacted habitats. Their positioning opposite to the BOD vector is consistent with previous reports identifying these taxa as indicators of relatively good water quality (Kelly et al. 2007; Rimet and Bouchez 2012). High-profile taxa (*C. tumida*, *C. turgidula*, *G. clevei*) were positioned near EC and PO_4^{3-} -P, suggesting an association with moderately enriched and hydrologically stable habitats. This pattern aligns with Passy's (2007) guild concept, which predicts that upright forms are favored under conditions of moderate nutrient availability and reduced physical disturbance. Comparable responses of *Cymbella* species to mesotrophic conditions have been documented in the Yom River (Yana et al. 2013). Although CCA results do not represent formal statistical testing of species-environment relationships, the ordination reveals coherent ecological patterns along the river continuum. Upstream reaches were characterized by oxygen-rich, low-BOD conditions and dominance of low-profile taxa; midstream reaches showed associations with moderate nutrient availability and high-profile taxa; and downstream reaches were linked to increased organic loading and motile taxa. This longitudinal organization is consistent with the River Continuum Concept (Vannote et al. 1980) and parallels patterns reported for the Ping River basin (Itayama et al. 2015). Overall, the CCA highlights the sensitivity of benthic diatom assemblages to combined gradients of organic pollution, nutrient enrichment, and oxygen availability in tropical rivers. These results support the need for regionally calibrated diatom indices, as tropical taxa may respond differently to environmental stressors compared to temperate assemblages, reinforcing the rationale for developing a tropical Diatom Index for the Yang River.

Diatom Index development and validation

Development of the Diatom Index

The Diatom Index was developed using 29 dominant benthic diatom taxa with relative abundances ≥1%. Weighted Averages (WAs) and Indicator Values (IVs) calculated for BOD, NH_4^+ -N, NO_3^--N, and PO_4^{3-}-P revealed clear gradients in species-specific responses to organic matter and nutrient enrichment (Table 15). Among the selected taxa, *A. minutissimum* exhibited the highest Mean IV (5.00), indicating a strong association with meso-eutrophic conditions rather than strictly oligotrophic environments. Other taxa with relatively high Mean IVs included *G. clevei* (4.50), *C. affinis* (4.25), and *N. cryptocephala* (4.25), as well as several species with Mean IV = 4.00 such as *Cymbella parva* (W.Smith) Kirchner, 1878, *Encyonema minutum* (Hilse) D.G.Mann, 1990, *Synedra ulna* var. *aequalis* (Kützing) Brun, 1880, and *Ulnaria lancleolata* (Kützing) P.Compère, 2001. These taxa predominantly corresponded to mesotrophic to meso-eutrophic conditions. In contrast, taxa with lower Mean IVs (approximately 3.00-3.25), including *Achnanthidium pusillum* Czarnecki & Edlund, 1995, *Achnanthidium caravelense* M.E.Novais & L.Ector, 2011, *Achnanthes suchlandtii* Hustedt, 1933, *Mayamaea atomus* (Kützing) Lange-Bertalot, 1997, and *Navicula capitatoradiata* H.Germain ex Gasse, were associated with oligo-mesotrophic conditions, reflecting preference for low-nutrient or transitional environments. Intermediate Mean IV values (approximately 3.50-3.75), observed in taxa such as *C. placentula*, *D. confervacea*, *Gomphonema lagenula* Kützing, 1844, *P. frequentissimum*, *N. palea*, and *R. gibberula*, characterized typical mesotrophic taxa that were common across mid- and downstream stations. Overall, trophic classifications derived from Mean IV values were consistent with physicochemical assessments, indicating that the Yang River predominantly exhibits oligo-mesotrophic to mesotrophic conditions. To evaluate reproducibility, three independent replicate chemistry datasets (R1-R3) for BOD, NH_4^+-N, NO_3^--N, and PO_4^{3-}-P were used to compute replicate WAs, IVs, and Diatom Index (DI) for all taxa. Pairwise Spearman's rank correlations among the resulting Diatom Index datasets (DI_R1, DI_R2, DI_R3) showed very high agreement (ρ = 0.860-0.907, $p < 0.001$), demonstrating strong reproducibility of the index calculations. Relationships between the Diatom Index and the conventional Water Quality Index (WQI) were positive but weak when based on individual replicate datasets (r^2 = 0.06-0.15). However, the relationship improved substantially when mean chemistry values were used (r^2 = 0.3102; Figure 5), indicating that averaging across replicates reduced analytical variability while retaining the underlying ecological signal.

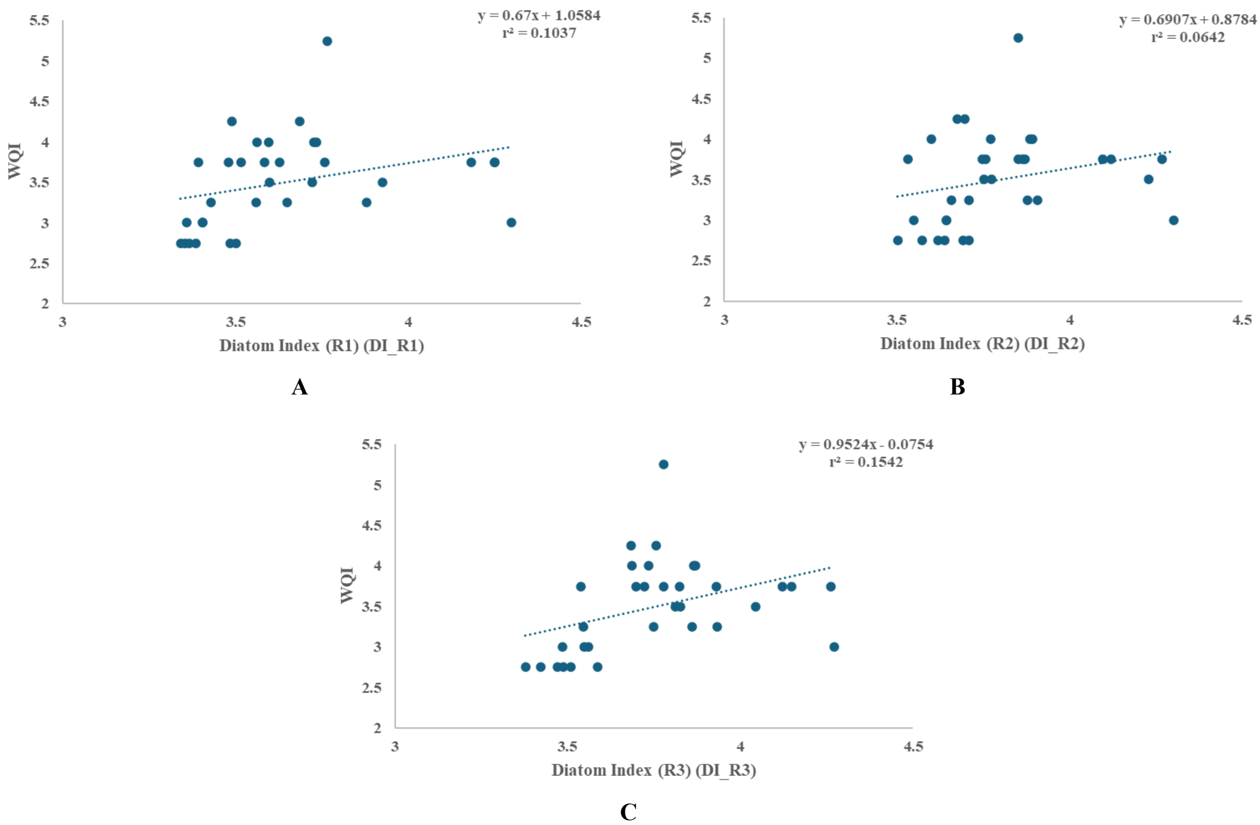

Figure 5. Relationships between A. DI_R1 and WQI, B. DI_R2 and WQI and C. DI_R3 and WQI

Table 15. Weighted Averages (WAs) and Indicator Values (IVs) of 29 dominant benthic diatom taxa (≥1% relative abundance) used for the development of the Yang River, Thailand Diatom Index

Species	WA BOD	WA NH_4^+	WA NO_3^-	WA PO_4^{3-}	IV BOD	IV NH_4^+	IV NO_3^-	IV PO_4^{3-}	Mean Ivs	Trophic status
Achnanthes oblongella Østrup, 1902	1.96	0.48	0.77	0.09	3	4	4	4	3.75	Mesotrophic
Achnanthes suchlandtii Hustedt, 1933	1.68	0.22	0.29	0.08	3	3	3	4	3.25	Oligo-mesotrophic
Achnanthidium caravelense M.E.Novais & L.Ector, 2011	1.71	0.25	0.27	0.08	3	3	3	4	3.25	Oligo-mesotrophic
Achnanthidium minutissimum (Kützing) Czarnecki, 1994	3.84	1.28	2.39	0.08	4	6	6	4	5.00	Meso-eutrophic
Achnanthidium pusillum Czarnecki & Edlund, 1995	1.86	0.29	0.13	0.13	3	3	2	4	3.00	Oligo-mesotrophic
Achnanthidium straubianum (Lange-Bertalot) Lange-Bertalot, 1999	1.55	0.26	0.28	0.11	3	3	3	4	3.25	Oligo-mesotrophic
Cocconeis placentula Ehrenberg, 1838	1.72	0.38	0.45	0.10	3	3	4	4	3.50	Mesotrophic
Cymbella affinis Kützing, 1844	2.86	0.50	0.80	0.16	4	4	5	4	4.25	Mesotrophic
Cymbella minuta Hilse, 1862	2.32	0.36	0.52	0.08	4	3	4	4	3.75	Mesotrophic
Cymbella parva (W.Smith) Kirchner, 1878	2.07	0.47	0.76	0.07	4	4	4	4	4.00	Mesotrophic
Cymbella tumida (Brébisson) Van Heurck, 1880	2.26	0.39	0.63	0.08	4	3	4	4	3.75	Mesotrophic
Cymbella turgidula Grunow, 1875	2.18	0.17	0.26	0.08	4	2	3	4	3.25	Oligo-mesotrophic
Diadesmis confervacea Kützing, 1844	2.22	0.24	0.46	0.07	4	3	4	4	3.75	Mesotrophic
Encyonema minutum (Hilse) D.G.Mann, 1990	4.27	0.38	0.56	0.09	5	3	4	4	4.00	Mesotrophic
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller, 1997	1.57	0.27	0.35	0.13	3	3	3	4	3.25	Oligo-mesotrophic
Gomphonema clevei Fricke, 1902	2.19	0.60	1.00	0.09	4	5	5	4	4.50	Meso-eutrophic
Gomphonema lagenula Kützing, 1844	2.69	0.27	0.40	0.09	4	3	4	4	3.75	Mesotrophic
Mayamaea atomus (Kützing) Lange-Bertalot, 1997	1.99	0.19	0.23	0.09	3	2	3	4	3.00	Oligo-mesotrophic
Navicula capitatoradiata H.Germain ex Gasse	1.83	0.24	0.25	0.08	3	3	3	4	3.25	Oligo-mesotrophic
Navicula cryptocephala Kützing	4.54	0.40	0.59	0.07	5	4	4	4	4.25	Mesotrophic
Navicula cryptotenella Lange-Bertalot	3.03	0.23	0.23	0.10	4	3	3	4	3.50	Mesotrophic
Navicula suprinii Gerd Moser, Lange-Bertalot & Metzeltin	2.50	0.35	0.51	0.08	4	3	4	4	3.75	Mesotrophic
Nitzschia palea (Kützing) W.Smith, 1856	2.65	0.32	0.47	0.08	4	3	4	4	3.75	Mesotrophic
Planothidium frequentissimum (Lange-Bertalot) Lange-Bertalot, 1999	1.86	0.31	0.47	0.10	3	3	4	4	3.50	Mesotrophic
Planothidium lanceolatum (Brébisson ex Kützing) Lange-Bertalot	1.64	0.35	0.52	0.11	3	3	4	4	3.50	Mesotrophic
Planothidium rostratum (Østrup) Lange-Bertalot, 1999	1.73	0.37	0.56	0.11	3	3	4	4	3.50	Mesotrophic
Rhopalodia gibberula (Ehrenberg) O.Müller, 1895	2.42	0.34	0.47	0.14	4	3	4	4	3.75	Mesotrophic
Synedra ulna var. *aequalis* (Kützing) Brun, 1880	4.08	0.38	0.55	0.08	5	3	4	4	4.00	Mesotrophic
Ulnaria lancleolata (Kützing) P.Compère, 2001	2.26	0.46	0.79	0.08	4	4	4	4	4.00	Mesotrophic

Note: WA: Weighted Average, indicating the mean environmental value at which each taxon occurs; IV: Indicator value based on species response to individual environmental variables. Mean IVs represents the average across four variables (BOD, NH_4^+-N, NO_3^--N, and PO_4^{3-}-P). Indicator levels were categorized into seven classes, and trophic status was assigned following international and Thai standards (Lorraine and Vollenweider 1981; Wetzel 2001; Peerapornpisal et al. 2004; PCD 2010)

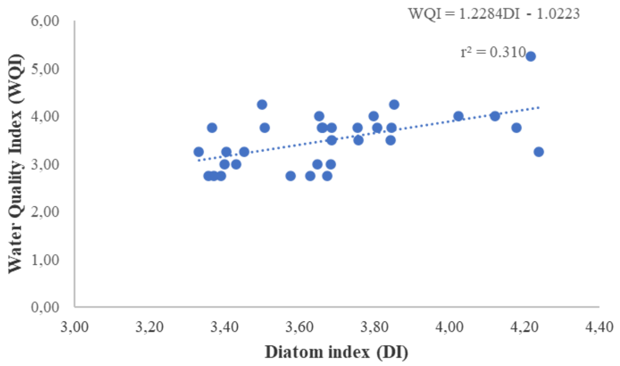

Figure 6. Relationship between the Diatom Index and the conventional Water Quality Index (WQI) with regression line

The indicator patterns derived from Table 15 demonstrate that benthic diatom taxa in the Yang River span a broad ecological gradient from oligo-mesotrophic to meso-eutrophic conditions. Species with high Mean IV values (≥ 4.00), including *A. minutissimum*, *G. clevei*, *N. cryptocephala*, *C. affinis*, and *E. minutum*, showed strong associations with nutrient-enriched or organically influenced conditions. Similar ecological affinities have been reported for these taxa in rivers of Northern Thailand and other regions (Kelly et al. 2007; Rimet and Bouchez 2012; Moonsin et al. 2013; Yana et al. 2013). Conversely, taxa with lower Mean IV values (approximately 3.00-3.25), such as *A. pusillum*, *A. caravelense*, *M. atomus*, and *A. suchlandtii*, were associated with oligo-mesotrophic conditions, reflecting upstream or less disturbed habitats. These species are commonly reported from low-nutrient streams in mountainous or upland regions and are consistent indicators of relatively good water quality in tropical systems. Taxa with intermediate IV values (approximately 3.50-3.75), including *N. palea*, *G. lagenula*, and *P. frequentissimum*, represented mesotrophic communities typically found in mid-catchment environments. Their widespread occurrence suggests moderate nutrient enrichment linked to mixed land-use influences along the river continuum. The close correspondence between IV-derived trophic classifications and physicochemical assessments confirms that the Yang River is moderately enriched but not severely degraded, in agreement with earlier water-quality evaluations. The high reproducibility of WAs, IVs, and Diatom Index (DI) across replicate chemistry datasets further demonstrates the robustness of the index calculation framework. Although DI-WQI relationships were modest for individual replicates, the improved relationship obtained using averaged chemistry values indicates that the Diatom Index captures the same underlying environmental gradient as conventional chemical indicators, while integrating biological responses over time. Importantly, differences between the ecological optima derived for Yang River taxa and those embedded in temperate-based indices such as the Trophic Diatom Index (TDI) or the Specific Pollution Sensitivity Index (IPS) highlight the necessity of region-specific calibration. Tropical diatom assemblages often exhibit broader ecological tolerances and distinct responses to nutrients and organic pollution compared to temperate counterparts. Therefore, the development of a locally calibrated Diatom Index provides a more ecologically meaningful and reliable tool for assessing water quality in tropical river systems.

Validation of the Diatom Index against water quality standards

The relationship between the Diatom Index (DI), developed from 29 dominant taxa, and the conventional Water Quality Index (WQI) was examined to evaluate the capacity of the DI to reflect overall water-quality conditions. Because both DI and WQI values deviated from normality, Spearman's rank correlation (ρ) was applied. The analysis revealed a moderate and statistically significant positive correlation between DI and WQI ($\rho = 0.549$, $p < 0.001$), indicating that higher DI values were associated with increased nutrient and organic enrichment as reflected by WQI scores. For visualization purposes, a linear trendline was fitted in Excel (WQI = 1.2284(DI) - 1.0223; $r^2 = 0.3102$; Figure 6). This regression was used solely as a descriptive illustration, not as a parametric inference, due to the non-parametric nature of the data. The coefficient of determination (r^2) indicates that approximately 31% of the variation in WQI could be described by changes in DI, suggesting that the Diatom Index captures a substantial proportion of the chemical water-quality gradient.

The moderate but significant relationship between the Diatom Index (DI) and WQI demonstrates that the DI effectively reflects major gradients of nutrient enrichment and organic pollution in the Yang River, while not being a direct surrogate for chemical indices. The proportion of unexplained variance (approximately 69%) likely reflects environmental drivers not explicitly included in the WQI framework, such as hydrological variability, substrate heterogeneity, light availability, diel oxygen dynamics, and biotic interactions (e.g., grazing), all of which are known to strongly influence benthic diatom assemblages in tropical river systems (Stevenson and Pan 1999; Rimet and Bouchez 2012). Despite the moderate strength of the DI-WQI relationship, the ecological coherence of the index was supported by consistent spatial patterns along the river continuum. Upstream stations, characterized by higher dissolved oxygen and lower nutrient concentrations, generally exhibited lower DI values, whereas mid- and downstream reaches influenced by agricultural activities and local settlements showed higher DI values. These patterns align with previous diatom-based assessments in Northern Thailand, including studies from the Yom and Ping Rivers, where diatom assemblages responded predictably to gradients of organic loading and nutrient enrichment (Leelahakriengkrai and Peerapornpisal 2011; Yana et al. 2013). Comparable DI-WQI relationships have also been reported in other tropical and subtropical systems, where biological indices typically explain only part of the variance in chemical water-quality metrics due to the integrative nature of biological responses (Kelly et al. 2007; Malek et al. 2012; Ladera et al. 2025). In this context, the explanatory power observed in the present study ($r^2 = 0.31$) is considered ecologically meaningful and consistent with expectations for bioassessment tools based on benthic communities.

Limitations and future research: The present study was restricted to the cool and hot seasons; therefore, wet-season hydrology, flood pulses, and storm-driven nutrient inputs were not captured. In addition, hydrological variables (e.g., flow velocity), substrate characteristics, and light availability were not measured, although these factors strongly influence diatom growth and assemblage structure. Future studies should incorporate these variables and expand seasonal coverage to further refine the predictive performance and robustness of the Diatom Index.

Overall, the Diatom Index developed for the Yang River provides a practical, ecologically meaningful, and regionally calibrated bioassessment tool for tropical rivers. While it does not replace conventional chemical indices, it complements them by integrating biological responses over time. The results further emphasize the necessity of region-specific calibration, as tropical diatom assemblages exhibit ecological optima that differ substantially from those embedded in temperate-derived indices.

The present study developed and validated a tropical Diatom Index (DI) for the Yang River in Northern Thailand, demonstrating that benthic diatom assemblages respond consistently to gradients of nutrient enrichment and organic loading in this tropical lotic system. The DI showed a moderate but statistically significant relationship with the conventional Water Quality Index (WQI), indicating that the index reflects major water-quality patterns and can complement physicochemical assessments. However, the proportion of unexplained variation suggests that additional environmental factors, such as hydrology, light availability, and habitat heterogeneity, may also influence diatom assemblages. Overall, the Yang River exhibited oligo-mesotrophic to mesotrophic conditions during the study period, although localized increases in biochemical oxygen demand and nutrients were observed in downstream sections. These results emphasize the importance of developing region-specific diatom indices for tropical rivers, where species composition and ecological tolerances differ from those underlying temperate-based indices. Future refinement of the index would benefit from inclusion of wet-season sampling, incorporation of hydrological variables, and validation across a wider range of tropical river systems. Despite these limitations, the Diatom Index developed in this study provides a practical and ecologically relevant tool for water-quality assessment in tropical rivers and contributes to addressing the current lack of calibrated diatom-based bioassessment methods for Southeast Asia.

REFERENCES

Allan JD, Castillo MM. 2007. Stream Ecology: Structure and Function of Running Waters. Springer, Dordrecht.

American Public Health Association (APHA). 2017. Standard Methods for the Examination of Water and Wastewater. 23rd ed American Public Health Association, American Water Works Association, and Water Environment Federation, Washington, DC.

Battarbee RW. 1986. Diatom analysis. In: Berglund BE (eds.). Handbook of Holocene Palaeoecology and Palaeohydrology. John Wiley & Sons, Chichester.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol Appl 8 (3): 559-568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2.

Dufrêne M, Legendre P. 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. Ecol Monogr 67 (3): 345-366. https://doi.org/10.2307/2963459.

Guiry MD, Guiry GM. 2025. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org. [Accessed November 2, 2025].

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological Statistics software package for education and data analysis. Palaeontol Electronica 4 (1): 9.

Itayama T, Hawkins PR, Leelahakriengkrai P, Kullasoot S, Whangchai N, Chitmanat C, Peerapornpisal Y, Kawabata Z. 2015. Bioassessment of dry season water quality in the Ping River around Chiang Mai City, Thailand. Chiang Mai J Sci 42 (2): 349-366.

John DM, Whitton BA, Brook AJ. 2011. The Freshwater Algal Flora of the British Isles: An Identification Guide to Freshwater and Terrestrial Algae. 2nd ed. Cambridge University Press, Cambridge.

Jones W, Medrano S. 2006. Indiana Lake Water Quality Assessment Report for 1999-2003. School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana, USA.

Jørgensen SE, Xu L, Costanza R. 2010. Handbook of Ecological Indicators for Assessment of Ecosystem Health. 2nd ed. CRC Press, Boca Raton. https://doi.org/10.1201/EBK1439809365.

Kelly MG, Cazaubon A, Coring E, Dell'Uomo A, Ector L, Goldsmith B, Guasch H, Hürlimann J, Jarlman A, Kawecka B, Kwandrans J, Laugaste R, Lindstrøm EA, Leitao M, Marvan P, Padisák J, Pipp E, Prygiel J, Rott E, Sabater S, van Dam H, Vizinet J. 1998. Recommendations for the routine sampling of diatoms for water quality assessments in Europe. J Appl Phycol 10: 215-224. https://doi.org/10.1023/A:1008033201227.

Kelly MG, Juggins S, Guthrie R, Pritchard S, Jamieson J, Rippey B, Hirst H, Yallop M. 2007. Assessment of ecological status in UK rivers using diatoms. Freshw Biol 53 (2): 403-422. https://doi.org/10.1111/j.1365-2427.2007.01903.x.

Kelly MG, Whitton BA. 1995. The trophic Diatom Index: A new index for monitoring eutrophication in rivers. J Appl Phycol 7 (4): 433-444. https://doi.org/10.1007/BF00003802.

Krammer K, Lange-Bertalot H. 1986-1991. Bacillariophyceae. Süsswasserflora von Mitteleuropa (Vols. 1-4). Gustav Fischer Verlag, Stuttgart. [German]

Ladera RA, Sinco AL, Sendaydiego JP, Saab LL, Hallazgo CIJS, Raagas EL. 2025. Diatom-based Index: A tool for assessing water quality in A Southeast Asian tropical river basin. Pollution 11 (2): 298-311. https://doi.org/10.22059/poll.2024.378540.2437.

Lange-Bertalot H, Krammer K. 1989. *Achnanthes*: Eine Monographie der Gattung. Gebrüder Borntraeger Verlagsbuchhandlung, Berlin. [German]

Leelahakriengkrai P, Peerapornpisal Y. 2011. Diversity of benthic diatoms in six main rivers of Thailand. Thai J Bot 3 (2): 125-138.

Leelahakrienkrai P, Peerapornpisal Y. 2010. Diversity of benthic diatoms and water quality of the Ping River, Northern Thailand. Environ Asia 3 (1): 82-94.

Lepš J, Šmilauer P. 2003. Multivariate analysis of ecological data using CANOCO. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511615146.

Levene H. 1960. Robust tests for equality of variances. In: Olkin I (eds.). Contributions to Probability and Statistics. Stanford University Press, Palo Alto.

Lorraine R, Vollenweider RA. 1981. Eutrophication of Waters: Monitoring, Assessment and Control. OECD, Paris.

Malek S, Salleh A, Milow P, Baba MS, Sharifah SA. 2012. Applying artificial neural network theory to exploring diatom abundance at tropical Putrajaya Lake, Malaysia. J Freshw Ecol 27 (2): 211-227. https://doi.org/10.1080/02705060.2011.635883.

Molle F, Mollinga PP. 2003. Water poverty indicators: Conceptual problems and policy issues. Water Policy 5 (5-6): 529-544. https://doi.org/10.2166/wp.2003.0034.

Moonsin P, Peerapornpisal Y, Rott E, Pipp E, Pimmongkol A. 2013. Diversity of benthic diatoms and water quality of the Mekong River passing through Ubon Ratchathani Province, Thailand. Res J Environ Earth Sci 5 (2): 58-66. https://doi.org/10.19026/rjees.5.5639.

Passy SI. 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. Aquat Bot 86 (2): 171-178. https://doi.org/10.1016/j.aquabot.2006.09.018.

Patrick R, Reimer CW. 1966. The Diatoms of the United States, Vol. 1. Academy of Natural Sciences, Philadelphia.

Peerapornpisal Y, Chaiubol C, Pekkoh J, Kraibut H, Chorum M, Wannathong P, Ngearnpat N, Jusakul K, Thammathiwat A, Chuanunta J, Inthasotti T. 2004. Monitoring of water quality in Ang Kaew Reservoir of Chiang Mai University using phytoplankton as bioindicator from 1995-2002. Chiang Mai J Sci 31 (1): 85-94.

Pollution Control Department (PCD). 2010. Surface Water Quality Standards of Thailand. Bangkok. Ministry of Natural Resources and Environment. https://www.pcd.go.th. [Accessed June 12, 2024].

Prasertsin T, Suk-ueng K, Phinyo K, Yana E. 2021. The diversity and abundance of phytoplankton and benthic diatoms in varying environmental conditions in Kok River, Chiang Rai, Thailand as bio-indicators of water quality. Biodiversitas 22 (4): 1853-1862. https://doi.10.13057/biodiv/d220431.

Prygiel J, Coste M. 2000. Guide Méthodologique pour l'application de l'indice Biologique Diatomées (IBD). Agence de l'Eau Artois-Picardie, France. [French]

Rimet F, Bouchez A. 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowledge Manag Aquat Ecosyst 406: 01. https://doi.org/10.1051/kmae/2012018.

Rimet F, Gusev E, Kahlert M, Kelly MG, Kulikovskiy M, Maltsev Y, Mann DG, Pfannkuchen M, Trobajo R, Vasselon V, Zimmermann J, Bouchez A. 2019. Diat.barcode, an open-access curated barcode library for diatoms. Sci Rep 9 (1): 15116. https://doi.org/10.1038/s41598-019-51500-6.

Round FE, Crawford RM, Mann DG. 1990. The Diatoms: Biology and Morphology of the Genera. Cambridge University Press, Cambridge.

Schrader HJ, Gersonde R. 1978. Diatoms and silicoflagellates. In: Zachariasse WS (eds.). Micropaleontological Counting Methods and Techniques: An Exercise on an Eight Metres Section of the Lower Pliocene of Capo Rossello, Sicily (Utrecht Micropaleontological Bulletins 17, pp. 129-176). Micropaleontological Press, Utrecht.

Shapiro SS, Wilk MB. 1965. An analysis of variance test for normality (complete samples). Biometrika 52 (3/4): 591-611. https://doi.org/10.1093/biomet/52.3-4.591.

Sor R, Ngor PB, Soum S, Chandra S, Hogan ZS, Null SE. 2021. Water quality degradation in the Lower Mekong Basin. Water 13 (11): 1555. https://doi.org/10.3390/w13111555.

Stevenson RJ, Pan Y, van Dam H. 2010. Assessing environmental conditions in rivers and streams with diatoms. In: Smol JP, Stoermer EF (eds.). The diatoms: Applications for the Environmental and Earth Sciences 2nd ed. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511763175.005.

Stevenson RJ, Pan Y. 1999. Assessing environmental conditions in rivers and streams with diatoms. In: Stoermer EF, Smol JP (eds.). The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511613005.003.

Taylor JC, Harding WR, Archibald CGM. 2007. An Illustrated Guide to Some Common Diatom Species from South Africa. WRC Report No. TT 282/07. Water Research Commission, Pretoria.

van Dam H, Mertens A, Sinkeldam J. 1994. A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Netherlands J Aquat Ecol 28: 117-133. https://doi.org/10.1007/BF02334251.

Vannote RL, Minshall GW, Cummins KW, Sedell JR, Cushing CE. 1980. The river continuum concept. Can J Fish Aquat Sci 37 (1): 130-137. https://doi.org/10.1139/f80-017.

Wetzel RG. 2001. Limnology: Lake and River Ecosystems 3rd ed. Academic Press, San Diego.

Yana E, Peerapornpisal Y, Mayama S. 2013. Diversity of benthic diatoms and correlation with water quality of Yom River, Thailand. Intl J Appl Environ Sci 8 (15): 1935-1948.

Zuur AF, Ieno EN, Elphick CS. 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol 1 (1): 3-14. https://doi.org/10.1111/j.2041-210X.2009.00001.x.