Environmental Science and Ecotechnology 19 (2024) 100326

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Optical measurements of dissolved organic matter as proxies for COD_{Mn} and BOD_5 in plateau lakes

Xuan Yang [a, b, c], Yongqiang Zhou [d, **], Xiaoying Yang [a, b], Yunlin Zhang [d],
Robert G.M. Spencer [e], Justin D. Brookes [f], Erik Jeppesen [a, g, h, i, j], Hucai Zhang [a],
Qichao Zhou [a, k, *]

[a] *Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Sciences, Yunnan University, Kunming, 650500, China*
[b] *Institute of International River and Eco-security, Yunnan University, Kunming, 650500, China*
[c] *Zhejiang College of Security Technology, Wenzhou, 325016, China*
[d] *Taihu Laboratory for Lake Ecosystem Research, State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, 210008, China*
[e] *Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, 32306, United States*
[f] *Water Research Centre, School of Biological Science, The University of Adelaide, Adelaide, 5005, Australia*
[g] *Department of Ecoscience, Aarhus University, Aarhus, 8000, Denmark*
[h] *Sino-Danish Centre for Education and Research, Chinese Academy of Sciences, Beijing, 100101, China*
[i] *Limnology Laboratory, Department of Biological Sciences and Centre for Ecosystem Research and Implementation, Middle East Technical University, Ankara, 06800, Turkey*
[j] *Institute of Marine Sciences, Middle East Technical University, Erdemli, Mersin, 33731, Turkey*
[k] *Yunnan Key Laboratory of Pollution Process and Management of Plateau Lake-Watershed, Yunnan Research Academy of Eco-environmental Sciences, Kunming, 650034, China*

ARTICLE INFO

Article history:
Received 18 January 2023
Received in revised form
25 September 2023
Accepted 27 September 2023

Keywords:
Dissolved organic matter
Chemical oxygen demand
Biochemical oxygen demand
Water quality
Yungui plateau lakes

ABSTRACT

The presence of organic matter in lakes profoundly impacts drinking water supplies, yet treatment processes involving coagulants and disinfectants can yield carcinogenic disinfection by-products. Traditional assessments of organic matter, such as chemical oxygen demand (COD_{Mn}) and biochemical oxygen demand (BOD_5), are often time-consuming. Alternatively, optical measurements of dissolved organic matter (DOM) offer a rapid and reliable means of obtaining organic matter composition data. Here we employed DOM optical measurements in conjunction with parallel factor analysis to scrutinize COD_{Mn} and BOD_5 variability. Validation was performed using an independent dataset encompassing six lakes on the Yungui Plateau from 2014 to 2016 ($n = 256$). Leveraging multiple linear regressions (MLRs) applied to DOM absorbance at 254 nm (a_{254}) and fluorescence components C1−C5, we successfully traced COD_{Mn} and BOD_5 variations across the entire plateau (68 lakes, $n = 271$, $R^2 > 0.8$, $P < 0.0001$). Notably, DOM optical indices yielded superior estimates (higher R^2) of COD_{Mn} and BOD_5 during the rainy season compared to the dry season and demonstrated increased accuracy ($R^2 > 0.9$) in mesotrophic lakes compared to oligotrophic and eutrophic lakes. This study underscores the utility of MLR-based DOM indices for inferring COD_{Mn} and BOD_5 variability in plateau lakes and highlights the potential of integrating *in situ* and remote sensing platforms for water pollution early warning.

1. Introduction

Lakes and reservoirs are important resources for drinking water for large- and mega-cities, making it imperative to prioritize the preservation of water quality to safeguard human health. Nevertheless, the water quality of lakes and reservoirs is increasingly

---

* Corresponding author. Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Sciences, Yunnan University, Kunming, 650500, China.
** Corresponding author.
E-mail addresses: yqzhou@niglas.ac.cn (Y. Zhou), qchzhou@ynu.edu.cn (Q. Zhou).

https://doi.org/10.1016/j.ese.2023.100326

threatened by the dual pressures of climate change and urbanization. Organic matter (OM) in human-impacted lakes is closely associated with the toxicity and cycling of heavy metals and micropollutants, such as perfluoroalkyl substances, and its presence directly threatens the safety of surrounding ecosystems and drinking water supplies [1–3]. Climate change and anthropogenic disturbances are both key drivers that influence the chemical composition and biogeochemical processes of OM [4–8]. More specifically, dissolved OM (DOM) constitutes the vast majority of OM in aquatic ecosystems and is responsible for the unpleasant odor and taste of drinking water [9], the formation of carcinogenic disinfection by-products, and the fouling of the filtration membranes in water treatment plants [10,11]. The excessive production and degradation of DOM are increasingly influencing the anoxia of water columns, the bio-lability of micro-pollutants, and the formation of black water blooms in various environments. Thus, they may threaten the safety of water supplies [1,12]. Therefore, there is an urgent and pressing need for research to swiftly detect and facilitate the safe and effective removal of OM during water treatment.

During wastewater treatment and the evaluation of surface water quality, chemical oxygen demand (COD) and five-day biochemical oxygen demand (BOD_5) are often used to trace the dynamics of OM in aquatic ecosystems [13]. However, these two indices have some limitations. For COD detection, conventional chemical reagents such as potassium dichromate and permanganate are often employed, but their use can lead to the generation of chemical residues, thereby potentially causing secondary pollution. The digestion process for COD measurements is time-consuming, and the chemical titration method involved in the measurement often contains prediction errors [14]. The BOD_5 method requires five days of laboratory bio-incubation, which naturally restricts rapid monitoring of OM, especially during contamination events. Moreover, it exhibits limited reproducibility due to the necessity of conducting initial and final measurements of different subsamples [14]. Additionally, not all bio-labile OM can be utilized by site-specific microorganisms, potentially leading to errors in the final BOD_5 results. Therefore, operationally straightforward, time-saving, economical, environmentally friendly, and accurate monitoring technologies are needed.

Researchers have been looking for new monitoring technologies (photometric or spectral methods, optical or microbial sensors, etc.) as potential replacements for established detection technologies for some time (e.g., TiO_2 thin film electrode, 5-sulfosalicylic acid, and titanium dioxides, etc.) [15–18]. Among these technologies, optical measurements of DOM are perhaps the most rapid, simple, and convenient. DOM optical measurements can also be easily adapted into *in situ*, real-time monitoring sensor platforms [19,20]. In addition, DOM absorbance spectra can usually be used to characterize the abundance and terrestrial input signals of DOM [15,21,22]. Similarly, fluorescent spectra can characterize the optical composition of DOM, detect subtle changes in the DOM pool, and distinguish between bio-labile and bio-stable organic components [13,21–23].

Previous studies have used the absorbance and fluorescent spectra of DOM as surrogates for water quality indices, including total nitrogen (TN), total phosphorus (TP), and biological toxicity in various aquatic environments [24–26]. However, the regression results between DOM optical indices and TN or TP have been inconsistent between different water bodies [13]. These inconsistencies can be attributed to the DOM optical differences due to the distinct characteristics of catchment landscapes among different seasons, the trophic status when they were taken, and the differences in source waters [27,28]. A few studies have applied DOM optical measurements for tracing the variability of COD_{Mn} and BOD_5 together, but these studies were limited to single rivers or lakes [10,29,30].

In this study, we aimed to explore the potential application of DOM optical measurements as surrogates for COD_{Mn} and BOD_5 in a diverse set of plateau lakes and to compare the optical measurements prediction capabilities across sampling seasons and trophic states. To achieve this, we conducted field sampling campaigns in 68 Yungui Plateau lakes in southwestern China from 2017 to 2018 to trace the variability in COD_{Mn} and BOD_5 using DOM optical measurements. We validated the linkages between DOM optical indices and COD_{Mn} and BOD_5 using a series of laboratory dilution experiments on different DOM sources (household effluents and soil leachates). Further validation of our prediction models was performed using an additional dataset from the six largest lakes on the plateau seasonally from 2014 to 2016, totaling 12 sampling events. Using DOM optical measurements to trace the variability in COD_{Mn} and BOD_5, our results hold significant implications for the advancement of water resource management schemes for the plateau lakes included in this study and possibly in an even wider environment. We hypothesized that employing multiple models instead of the single linear model of DOM optical indices would yield better results in tracing COD_{Mn} and BOD_5 in these plateau lakes. Furthermore, we anticipated that the predictive capability could be largely improved by modeling the results separately for different seasons and trophic status rather than aggregating all data.

2. Materials and methods

2.1. Field sampling and laboratory experiments

The Yungui Plateau in southwestern China exhibits distinct dry (November–April) and rainy seasons (May–October, contributing more than 80% of the annual precipitation). Numerous lakes at different elevations can be found on the plateau. The surface area of the lakes on the plateau ranges from <1 to 310 km^2, the mean water depth of the lakes ranges from <1 to 89.6 m, and the trophic states vary from oligotrophic to highly eutrophic [31,32]. The lakes of the Yungui Plateau are important for supplying water to the surrounding residential areas, controlling floods and drainage, providing agricultural irrigation water, and providing habitat for biota. However, the water quality of the Yungui Plateau lakes has been increasingly threatened by the dual pressures of enhanced human activities and climate change.

We collected samples from a total of 68 Yungui Plateau lakes during the period from 2017 to 2018 (Fig. S1). According to lake size and ecological characteristics, the samples were collected from 1 to 4 sites per lake, and the number of sampling sites was determined based on the surface area of the sampled lakes, yielding a total of 271 samples. Additional field sampling campaigns were carried out in the six largest lakes on the plateau, with states varying from oligotrophic to highly eutrophic to validate the multiple linear regressions (MLRs) obtained from the Yungui Plateau lakes. These included Lake Fuxian (FX), Lake Dianchi (DC), Lake Erhai (EH), Lake Yangzong (YZ), Lake Lugu (LG), and Lake Chenghai (CH) [32]. Water samples were collected 0.5 m below the water surface and transported to the nearest laboratory (e.g., Dali Erhai Lake Research Institute) following collection. These samples were then filtered immediately and stored in the dark at 4 °C. A fraction of the water samples was filtered through 0.22 μm Millipore membranes for DOM absorbance and fluorescence spectroscopy. COD_{Mn}, BOD_5, and other water quality indices were then determined for the remaining water samples.

To determine whether the variability in COD_{Mn} and BOD_5 can be traced using optical measurements of DOM from different sources, we conducted a series of laboratory dilution experiments on different sources of DOM in 2021. Briefly, soil leachate samples (500 g topsoil, see Fig. S1 for sampling location, and 2.5 L Milli-Q water were mixed and shaken vigorously at 300 g min^{-1} for 2 h), and domestic raw sewage (collected from the sewers in the residential areas of Kunming City) was also collected. The samples were then diluted in the following ratios of the raw solution to Milli-Q water 1:0, 1:1, 1:2, 1:4, 1:9, 1:29, and 1:99 before COD_{Mn}, BOD_5, and DOM optical measurements were made.

2.2. Physicochemical indices and DOM optical measurement

Detailed descriptions of the analytical methods of COD_{Mn} and BOD_5 can be found in the research conducted by Jiao et al. [14]. The field and laboratory measurements for other physicochemical indices, including water depth (WD), water temperature (WT), Secchi disk depth (SDD), TN, TP, and chlorophyll *a* (Chl. *a*), of the 68 lakes sampled from 2017 to 2018 are available in Zhou et al. [31]. The measurement, calibration of DOM absorbance and fluorescent spectra, and parallel factor analysis (PARAFAC) can also be found in Refs. [33,34]. Specifically, excitation emission matrixes (EEMs) at excitation wavelengths below 225 nm and emission wavelengths below 300 nm and above 550 nm were removed due to deteriorating signal-to-noise ratios in these regions [33]. The optical spectra of the obtained five components were compared to the fluorophores identified earlier in other ecosystems using the OpenFluor optical library [35].

The five components (C1−C5) identified by PARAFAC are shown in Fig. S2. C1 (230 (325)/441 nm) represents terrestrial humic-like materials and is often closely associated with terrestrial-derived OM and composed of highly aromatic materials [36,37]. C2 displays two excitation maxima (at ≤260 and 350 nm) that correspond to a single emission maximum (at 481 nm), similar to a humic-like fluorophore. C3 (240 (300)/302 nm) shares spectral characteristics similar to a tyrosine-like substance [38], and C4 exhibits an excitation and emission spectrum (Ex./Em. maxima at 230 (275)/339 nm) similar to a tryptophan-like substance. Lastly, C5 exhibits a single excitation/emission maximum (at 275/<300 nm), consistent with a tyrosine-like fluorophore [38].

Details about other DOM absorbance indices (including the absorbance ratio, a_{250}/a_{365}, and DOM spectral slope, $S_{275-295}$) and other fluorescence indices (including the humification index, HIX; fluorescence index, FI; autochthonous index, BIX; and freshness index) can be found in earlier studies [33,39,40]. According to the correlation matrices among our COD_{Mn}, BOD_5, and DOM optical indices, we selected a_{254} and C1−C5 as independent variables to predict the variabilities in COD_{Mn} and BOD_5 in plateau lakes.

2.3. Statistical analysis

We used Primer 6.0 to perform cluster analysis on all lake results in the rainy season ($n = 150$) based on water quality indices (Table S1). The 150 sampling sites were divided into three groups: oligotrophic, mesotrophic, and eutrophic lakes according to the SDD, COD_{Mn}, TN, TP, and Chl. *a* (Fig. S3). Additionally, we further divided our results into dry and rainy seasons relative to the sampling season to explore the influence of precipitation. Since the data points of measured and predicted COD_{Mn} and BOD_5 were not normally distributed, we used log-log relationships to test the validity of using DOM optical measurements to trace the variability in COD_{Mn} and BOD_5, which have been widely used in the field of remote sensing [41,42].

Statistical analysis in this study was performed using R version 3.6.2, encompassing computations such as ranges, mean values, standard deviations, and nonparametric tests to assess the statistical significance of differences in independent sample mean. A significance threshold of $P < 0.05$ was applied for statistical significance for *t*-tests, linear, and nonlinear regressions. To examine the predictive capability of DOM spectra in tracing the variability of COD_{Mn} and BOD_5, single linear regressions (SLRs) were carried out using the R basicTrendline package. MLRs were conducted using Origin 9.1. COD_{Mn} and BOD_5 were chosen as dependent variables, and the six DOM spectral indices (a_{254} and C1−C5) were selected as independent variables according to the results of redundancy analysis using the R vegan package in R version 3.6.2 (Fig. S4). The R ggplot2 package was also used during linear fittings for visualization.

3. Results

3.1. Linear regression models for all sampled Yungui Plateau lakes

The trophic states of the lakes studied on the Yungui Plateau ranged from clear oligotrophic to turbid and highly eutrophic (Table S1). In our study, all selected DOM spectral indices, a_{254} and humic-like C1−C2, in particular ($R^2 > 0.5$, Fig. S5), were closely correlated with COD_{Mn} and BOD_5 ($P < 0.0001$). Furthermore, compared to the SLR models, the MLR models showed markedly increased prediction capabilities for COD_{Mn} and BOD_5 using DOM optical indices ($R^2 > 0.7$, Fig. 1).

3.2. Linear regression models for lakes varying in season and trophic state

The mean DOM-related indices, including C1, C2, and C4, as well as COD_{Mn} and BOD_5 concentrations, were significantly higher in the rainy season than in the dry season (Table S1), and the predictive capability of DOM spectra used to trace the variability of COD_{Mn} and BOD_5 varied seasonally. Both SLR and MLR models demonstrated better prediction capabilities for COD_{Mn} and BOD_5 in the rainy season than in the dry season (Fig. S6 and Fig. 2). Considering the seasonal variability in DOM-related indices, we divided our MLR models into rainy or dry season models, and the predictive capability was improved by modeling the results separately in the dry or rainy season before they were compiled, compared to pooling all results (Fig. 2).

The DOM-related indices could also be used to trace variability in COD_{Mn} and BOD_5 in lakes with varying trophic states. The SLR predictive capability of tyrosine-like C3 for COD_{Mn} and BOD_5 in the mesotrophic lakes was better than in the oligotrophic and eutrophic lakes (Fig. S7); and, the MLR model had a better predictive capability for COD_{Mn} and BOD_5 in the mesotrophic lakes than in the

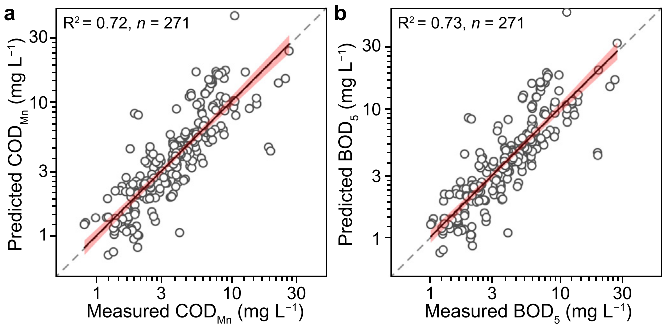

Fig. 1. Scatter plots of logarithmically-transformed measured and predicted values of chemical oxygen demand (COD_{Mn}, **a**) and biochemical oxygen demand (BOD_5, **b**) based on multiple linear regression models with all samples collected from the Yungui Plateau lakes in 2017−2018. Grey dot lines: 1:1 lines.

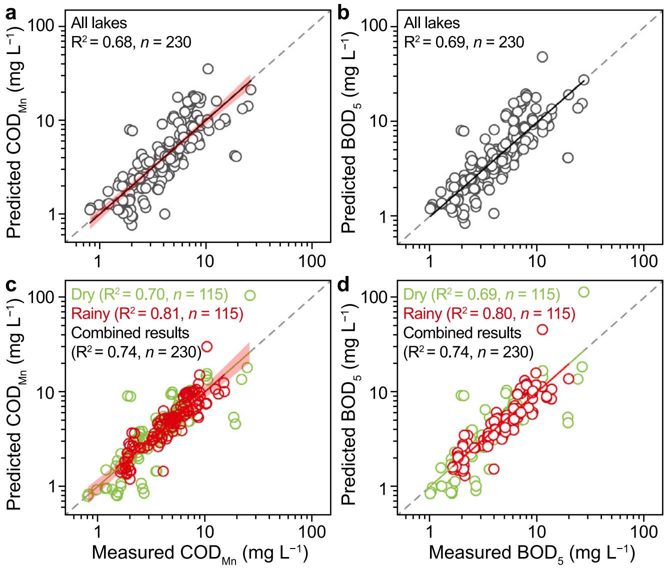

Fig. 2. Scatter plots of logarithmically-transformed measured and predicted values of COD_{Mn} and BOD_5 based on multiple linear regression models with all data pooled (samples varying by season) (**a**–**b**) and with the sampling results being categorized into rainy or dry seasons before they were compiled (**c**–**d**). Grey dot lines: 1:1 lines.

oligotrophic and eutrophic lakes (Fig. 3c–d). We further found that the trophic state classification resulted in the elevated predictive capability of the MLR models, with R^2 increasing from 0.84 (pooling all data for direct modeling) to 0.89 (the classification before compilation) for COD_{Mn} and from 0.83 to 0.88 for BOD_5 (Fig. 3).

3.3. Validation results of dilution experiments

The predictive capability of DOM-related indices for tracing the variability of COD_{Mn} and BOD_5 was differentiated from DOM sources. For the laboratory dilution experiment, a_{254} and both humic-like and protein-like components could be used to predict COD_{Mn} and BOD_5 for the domestic sewage samples, but only the a_{254} and the humic-like components instead of the protein-like components could be used satisfactorily to predict COD_{Mn} and BOD_5 for the soil leachate samples (Fig. S8). DOM optical measurements for domestic sewage had a high capability to trace both COD_{Mn} and BOD_5, but the predictive capability of DOM optical indices for BOD_5 in soil leachate samples was low (Fig. 4).

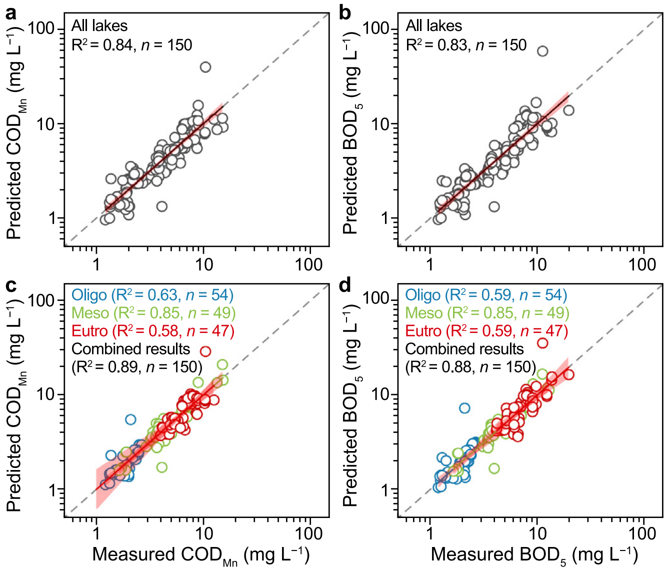

Fig. 3. Scatter plots of logarithmically-transformed measured and predicted values of COD_{Mn} and BOD_5 based on multiple linear regression models with all data pooled (samples varying by trophic state) (**a**–**b**) and with the sampling results being categorized into oligotrophic, mesotrophic, or eutrophic states before they were compiled (**c**–**d**). Grey dot lines: 1:1 lines.

3.4. Field validation results

We also validated our modeling results using an additional data set that included samples from the six largest lakes on the Yungui Plateau and varied from oligotrophic to eutrophic, collected seasonally from 2014 to 2016. Interestingly, we found that the obtained MLR models of all lakes could be used to trace the dynamics of COD_{Mn} and BOD_5 when all data collected from the six lakes were pooled together (Fig. 5; Table S2).

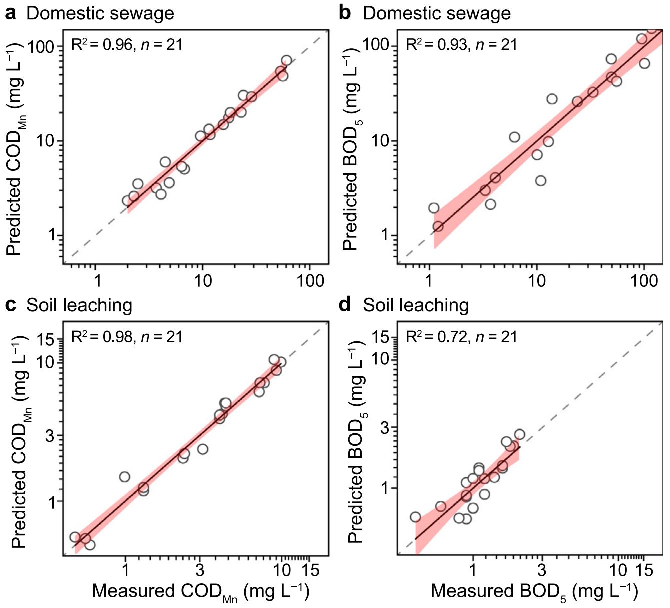

Fig. 4. Scatter plots of logarithmically-transformed measured and predicted values of COD_{Mn} and BOD_5 based on multiple linear regression models developed from the dilution experiments of domestic sewage samples (**a**–**b**) and soil leachate samples (**c**–**d**). Grey dot lines: 1:1 lines.

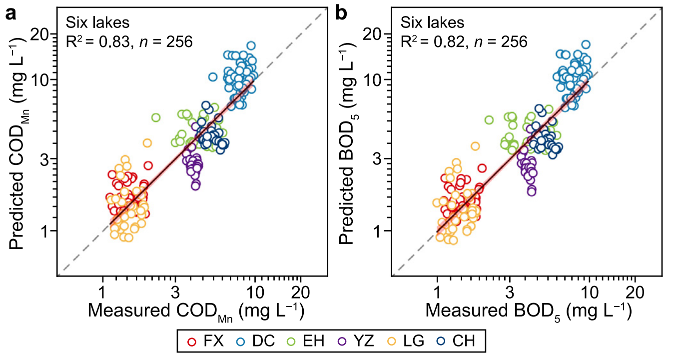

Fig. 5. Validation of the multiple linear regression models to predict the variabilities in COD_{Mn} (**a**) and BOD_5 (**b**) obtained from additional data sets on samples collected from the six largest lakes on the Yungui Plateau. FX: Lake Fuxian; DC: Lake Dianchi; EH: Lake Erhai; YZ: Lake Yangzong; LG: Lake Lugu; CH: Lake Chenghai. Grey dot lines: 1:1 lines.

4. Discussion

COD_{Mn} and BOD_5 have been widely used to trace the dynamics of OM in aquatic ecosystems, and DOM typically represents most of the variability in OM in many different environments. Thus, it is not surprising that the two were correlated in our study. However, few studies have used DOM optical measurements to examine the variability in COD_{Mn} and BOD_5 in lakes or reservoirs across a large and diverse region. Our study stands as one of the initial assessments aiming to track COD_{Mn} and BOD_5 dynamics using DOM optical indices in a wide range of lakes and across seasons and trophic states.

Plateau lakes in deep valleys typically have long water residence time and are key drinking water sources for medium and large cities. However, they are highly vulnerable to anthropogenic disturbances. Our MLR models based on multiple spectral indices performed better than the SLRs in all cases, which is consistent with our first hypothesis, likely reflecting the complexity and heterogeneity of water quality in these plateau lakes (Table S1). Ultraviolet–visible (UV–vis) absorbance has been used in wastewater quality monitoring programs worldwide [13,43], and compared to UV–vis absorbance spectroscopy, fluorescence spectroscopy offers a marked increase in optical resolution by enabling differentiation of optically sensitive DOM fractions that absorb light at the same wavelength but emit light in different spectral regions [38]. Due to its high sensitivity and affordability, the single fluorescent peaks of DOM and even the ratios between them have been used to trace water quality, including the TN and TP of surface water and municipal wastewater [22,24,44]. However, a single DOM spectral index cannot provide a complete overview with respect to DOM complexity and heterogeneity of water quality. Constructing MLR models using DOM optical indices (including UV–vis absorbance and several fluorescent peaks) across a range of trophic states and seasons allowed us to better trace the migration and transformation of DOM across a range of hydrologic conditions (see Table S1).

The long water residence time of the lakes on the Yungui Plateau can cause high spatiotemporal heterogeneity of DOM. Moreover, bioavailable fractions of DOM are utilized as substrate, and there can also be photochemical degradation of DOM near the surface [45]. There is general variability in the hydrologic conditions of lakes in the dry and rainy seasons. The input of allochthonous substances, water temperature, underwater ultraviolet radiation, and biological activity caused by seasonal changes can all affect the various optical properties of DOM [21,46,47]. During the rainy season, more OM (compared to the dry season) enters lakes due to enhanced water runoff through the litter and surface soil profiles that leach dissolved organic carbon. DOM also enters the lakes from the erosion of soil OM from farmland and the flushing of trash and organic debris from impervious landscapes [21,48,49].

We found that the fluorescence intensity and contribution percentage of terrestrial humic-like C1 and C2 increased with increasing rainwater flushing (Table S1), and the R^2 of linear fittings between humic-like components and COD_{Mn} or BOD_5 consequently increased during the rainy season [48]. Furthermore, higher water temperatures in rainy summers may promote the accumulation and degradation of algal cells and hence the release of algal protein-like C3 (Fig. S9) that often has high bio-lability [38] that leads to better predictions from of COD_{Mn} and BOD_5 in the rainy season than in the dry season. Tyrosine-like C3 and C5 are mainly derived from biological pyrolysis and release (such as algal pyrolysis and microbial release), and these two components can be relatively stable during their biogeochemical cycling in lakes. In comparison, tryptophan-like C4 has been attributed to mainly amino acids and linked to microorganisms from domestic sewage, landfill leachate, and other polluted water bodies [50]. Tryptophan-like C4 traces the variability in COD_{Mn} and BOD_5 better in the dry season compared to the rainy season as domestic effluent signals can be significantly diluted by rainwater during the rainy season. This could also be attributed to a higher microbial activity during the lower water (higher wastewater percentage) period than in the higher water period [10,21,49]. In addition, the strong penetration of underwater ultraviolet radiation and long water retention time of the plateau lakes [32,51] can promote photochemical and microbial degradation of bio-labile DOM and result in stable DOM exported downstream [52–54]. The predictive capability can be improved by modeling the results separately according to the sampling seasons of the sites than when all of the data were pooled (Fig. 2).

The quantity and quality of DOM typically vary between trophic states, and our results showed that the protein-like component C3 of mesotrophic lakes traced the variability of COD_{Mn} and BOD_5 better than other components and that MLR was a better predictor of COD_{Mn} and BOD_5 in mesotrophic lakes than in oligotrophic and eutrophic lakes. In oligotrophic lakes, the bio-stable OM (Table S1) is the main reason for the relatively weak correlation between the protein-like components and COD_{Mn} or BOD_5 [27]. However, the chemical composition of OM was also strongly affected by the ratio of humic-like to protein-like components and the proportion of fluorescent substances to nonfluorescent substances in lakes with low OM concentrations [13]. Human activities significantly affect mesotrophic lakes, and allochthonous nutrients and organic loading increase with increasing anthropogenic disturbances, which can promote algal proliferation [55]. Hence, high levels of protein-like OM derived from algal degradation (Fig. S9) can result in a significant association between protein-like C3 and COD_{Mn} or BOD_5. Similar to the seasonal results, the predictive capability was improved by modeling the results separately according to sites with different trophic status than when all data were pooled (Fig. 3), which largely reflects the heterogeneity of lake DOM with different trophic states.

The greater variability observed in both DOM fluorescence and the concentration of organic carbon within mesotrophic lakes, as compared to oligotrophic and eutrophic lakes, offers a plausible explanation for the heightened predictive efficacy of the MLR model in the former [10]. For eutrophic lakes, the residual errors of COD_{Mn} and BOD_5 caused by organic suspended matter, including algae, particulate algal secretions, and some particulate substances derived from terrestrial inputs and sediment re-suspension, may have interfered with the predictive value of the MLRs. Since DOM optical measurements cannot capture the particulate OM included in COD_{Mn} and BOD_5, the predictive capability of the MLR method in eutrophic lakes may be enhanced by adding organic suspended matter as an independent variable [30,56–58].

The spectral characteristics of the different sources of DOM also vary. The organic compounds in our soil leaching solution were mainly bio-stable humic-like components not easily utilized by microorganisms, as demonstrated by the low BOD_5 concentration (<2 mg L^{-1}). Therefore, a_{254}, humic-like compounds, protein-like compounds, and even MLR had a better predictive ability for COD_{Mn} than BOD_5 in the soil-leaching solution. Domestic sewage is rich in protein- and humic-like components, which can be used as indicators of BOD_5 in municipal waters [13,24]. Thus, the SLRs and the MLR using a_{254}, humic-like components C1–C2, and protein-like components C3–C5 of domestic sewage showed good predictive ability for COD_{Mn} and BOD_5.

Our investigation showed that optical measurements of DOM are a promising tool for both COD_{Mn} and BOD_5 monitoring in Yungui Plateau lakes with a wide range of trophic states. The predictive models we formulated were validated based on an external data set on the six largest lakes sampled seasonally from 2014 to

2016. Monitoring and managing over large geographical scales require new techniques and methods to assess water quality. Our results are applicable in water quality evaluation for at least the plateau lakes included here and can be easily adapted into *in situ* absorbance and fluorescence sensors. Currently, many fluorescent online monitoring methods and remote sensing technology have been applied in real-time and large-scale water monitoring programs [46,56,59,60]. Furthermore, we found that the pluralistic parameter acquisition, including the integration of multiple variables on sensors, should be considered in the actual application. With advances in artificial intelligence and machine learning technologies, these predictions may become even more precise, especially with the accumulation of data from additional lakes. Our results support our initial hypothesis, indicating that surrogate models of DOM spectra for water quality vary concerning seasons, trophic states, and different sources of DOM. Our results strongly advocate for the improved predictive capability achieved by developing separate models according to the sampling seasons and trophic status of the sampling sites rather than pooling all data. The application of DOM spectral indices thus needs to be adjusted and optimized according to different sampling times and trophic states. For example, field observations may be necessary for individual lakes to enhance water quality modeling [6]. Future research might benefit from considering other variables, including suspended particulate matter and the underwater climate, when modeling the variability of COD_{Mn} and BOD_5 in clear oligotrophic and turbid eutrophic lakes.

5. Conclusion

Optical measurements of DOM can be used to trace the variability of COD_{Mn} and BOD_5. However, inter-lake and seasonal differences also need to be considered due to the complexity of DOM in the studied plateau lakes in different regions and during hydrological periods. More specifically, the MLR framework outperforms the SLR model at tracing COD_{Mn} and BOD_5. Furthermore, the predictive capability of MLRs for COD_{Mn} and BOD_5, using DOM optical indicators, exhibits variations between seasons, with better performance observed during the rainy season compared to the dry season. Additionally, MLRs demonstrate better predictive capability in mesotrophic lakes than oligotrophic and eutrophic lakes. Our validation results also showed that regional-scale prediction models may not always be applicable to individual lakes. Hence, the MLR prediction models of DOM spectra for COD_{Mn} and BOD_5 need to be adjusted according to different sampling times and trophic states.

CRediT authorship contribution statement

Xuan Yang: Conceptualization, Investigation, Formal Analysis, Writing - Original Draft. **Yongqiang Zhou:** Conceptualization, Funding Acquisition, Methodology, Writing - Review & Editing. **Xiaoying Yang:** Data Curation, Investigation. **Yunlin Zhang:** Conceptualization, Writing - Review & Editing. **Robert G. M. Spencer:** Writing - Review & Editing. **Justin D. Brookes:** Writing - Review & Editing. **Erik Jeppesen:** Funding Acquisition, Writing - Review & Editing. **Hucai Zhang:** Resources, Supervision. **Qichao Zhou:** Conceptualization, Data Curation, Funding Acquisition, Writing - Review & Editing, Validation.

Data availability

Data used in this study is available at https://figshare.com/10.6084/m9.figshare.23744565.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the support of the Science and Technology Project of Yunnan Province (No. 202103AC100001), the National Natural Science Foundation of China (Nos. 42322104 and 42071131), the National Key Research and Development Program of China (No. 2017YFA0605202), the Youth Innovation Promotion Association CAS (No. 2021312), the Provincial Natural Science Foundation of Jiangsu (No. BK20220162), and the TÜBITAK program BIDEB2232 (project 118C250). We thank Licheng Huang, Jiang Qin, Jufen Nie, Jingyi Li, and Yuanrui Li for their assistance with the fieldwork and laboratory measurement, and Anne Mette Poulsen from Aarhus University for proof reading the draft. We are grateful to the editor and two anonymous reviewers for their constructive comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2023.100326.

References

[1] D.N. Kothawala, S.J. Köhler, A. Östlund, K. Wiberg, L. Ahrens, Influence of dissolved organic matter concentration and composition on the removal efficiency of perfluoroalkyl substances (PFASs) during drinking water treatment, Water Res. 121 (2017) 320–328.

[2] Q. Zhou, W. Wang, L. Huang, Y. Zhang, J. Qin, K. Li, L. Chen, Spatial and temporal variability in water transparency in Yunnan Plateau lakes, China, Aquat. Sci. 81 (2) (2019) 36.

[3] S. Wang, Y. Gao, J. Jia, Y. Lu, K. Sun, X. Ha, Z. Li, W. Deng, Vertically stratified water source characteristics and associated driving mechanisms of particulate organic carbon in a large floodplain lake system, Water Res. 209 (2022) 117963.

[4] Y. Zhang, J. Wang, J. Tao, Y. Zhou, H. Yang, X. Yang, Y. Li, Q. Zhou, E. Jeppesen, Concentrations of dissolved organic matter and methane in lakes in Southwest China: different roles of external factors and in-lake biota, Water Res. 225 (2022) 119190.

[5] S. Liu, J. Hou, C. Suo, J. Chen, X. Liu, R. Fu, F. Wu, Molecular-level composition of dissolved organic matter in distinct trophic states in Chinese lakes: implications for eutrophic lake management and the global carbon cycle, Water Res. 217 (2022) 118438.

[6] T. Dittmar, S.T. Lennartz, H. Buck-Wiese, D.A. Hansell, C. Santinelli, C. Vanni, B. Blasius, J.-H. Hehemann, Enigmatic persistence of dissolved organic matter in the ocean, Nat. Rev. Earth Environ. 2 (8) (2021) 570–583.

[7] R. Benner, K. Kaiser, Biological and photochemical transformations of amino acids and lignin phenols in riverine dissolved organic matter, Biogeochemistry 102 (1) (2011) 209–222.

[8] X. Zhu, Z. Wang, Y. Luo, Y. Ma, Z. Xu, L. Wang, F. Peng, Q. Pang, Y. Li, F. He, B. Xu, Overlying water fluoride concentrations influence dissolved organic matter composition and migration from pore water in sediment via bacterial mechanisms, Environ. Sci. Ecotechnol. 10 (2022) 100163.

[9] C.A. Stedmon, B. Seredyńska-Sobecka, R. Boe-Hansen, N. Le Tallec, C.K. Waul, E. Arvin, A potential approach for monitoring drinking water quality from groundwater systems using organic matter fluorescence as an early warning for contamination events, Water Res. 45 (18) (2011) 6030–6038.

[10] J.H. Goldman, S.A. Rounds, J.A. Needoba, Applications of fluorescence spectroscopy for predicting percent wastewater in an urban stream, Environ. Sci. Technol. 46 (8) (2012) 4374–4381.

[11] M. Sgroi, P. Roccaro, G.V. Korshin, F.G.A. Vagliasindi, Monitoring the behavior of emerging contaminants in wastewater-impacted rivers based on the use of fluorescence excitation emission matrixes (EEM), Environ. Sci. Technol. 51 (8) (2017) 4306–4316.

[12] Y. Zhou, E. Jeppesen, Y. Zhang, C. Niu, K. Shi, X. Liu, G. Zhu, B. Qin, Chromophoric dissolved organic matter of black waters in a highly eutrophic Chinese lake: freshly produced from algal scums? J. Hazard Mater. 299 (2015) 222–230.

[13] R.K. Henderson, A. Baker, K.R. Murphy, A. Hambly, R.M. Stuetz, S.J. Khan, Fluorescence as a potential monitoring tool for recycled water systems: a review, Water Res. 43 (4) (2009) 863–881.

[14] N. Jiao, J. Liu, B. Edwards, Z. Lv, R. Cai, Y. Liu, X. Xiao, J. Wang, F. Jiao, R. Wang, X. Huang, B. Guo, J. Sun, R. Zhang, Y. Zhang, K. Tang, Q. Zheng, F. Azam, J. Batt, W.-J. Cai, C. He, G.J. Herndl, P. Hill, D. Hutchins, J. LaRoche, M. Lewis, H. MacIntyre, L. Polimene, C. Robinson, Q. Shi, C.A. Suttle, H. Thomas, D. Wallace, L. Legendre, Correcting a major error in assessing organic carbon pollution in natural waters, Sci. Adv. 7 (16) (2021) eabc7318.

[15] J. Hur, D.S. Kong, Use of synchronous fluorescence spectra to estimate biochemical oxygen demand (BOD) of urban rivers affected by treated sewage, Environ. Technol. 29 (4) (2008) 435–444.

[16] T. Fang, C. Yang, L. Liao, Photoelectrocatalytic degradation of high COD dipterex pesticide by using TiO2/Ni photo electrode, J. Environ. Sci. 24 (6) (2012) 1149–1156.

[17] S. Li, F. Zheng, S. Cai, W. Liang, Y. Li, A visible light assisted photocatalytic system for determination of chemical oxygen demand using 5-sulfosalicylic acid *in situ* surface modified titanium dioxide, Sensor. Actuator. B Chem. 188 (2013) 280–285.

[18] Y. Zhang, G. Zhu, B. Qin, Y. Zhou, K. Shi, Application feasibility of chromophoric dissolved organic matter (CDOM) absorption coefficient as the substitute for chemical oxygen demand concentration in lakes and reservoirs, J. Lake Sci. 32 (6) (2020) 1575–1584.

[19] D.P. Hamilton, C.C. Carey, L. Arvola, P. Arzberger, C. Brewer, J.J. Cole, E. Gaiser, P.C. Hanson, B.W. Ibelings, E. Jennings, T.K. Kratz, F.-P. Lin, C.G. McBride, M. David de Marques, K. Muraoka, A. Nishri, B. Qin, J.S. Read, K.C. Rose, E. Ryder, K.C. Weathers, G. Zhu, D. Trolle, J.D. Brookes, A Global Lake Ecological Observatory Network (GLEON) for synthesising high-frequency sensor data for validation of deterministic ecological models, Inland Waters 5 (1) (2015) 49–56.

[20] A.S. Wymore, J. Potter, B. Rodríguez-Cardona, W.H. McDowell, Using in-situ optical sensors to understand the biogeochemistry of dissolved organic matter across a stream network, Water Resour. Res. 54 (4) (2018) 2949–2958.

[21] P.G. Coble, Marine optical biogeochemistry: the chemistry of ocean color, Chem. Rev. 107 (2) (2007) 402–418.

[22] D.P. Mesquita, C. Quintelas, A.L. Amaral, E.C. Ferreira, Monitoring biological wastewater treatment processes: recent advances in spectroscopy applications, Rev. Environ. Sci. Biotechnol. 16 (3) (2017) 395–424.

[23] D.M. Reynolds, The differentiation of biodegradable and non-biodegradable dissolved organic matter in wastewaters using fluorescence spectroscopy, J. Chem. Technol. Biotechnol. 77 (8) (2002) 965–972.

[24] J. Hur, J. Cho, Prediction of BOD, COD, and total nitrogen concentrations in a typical urban river using a fluorescence excitation-emission matrix with PARAFAC and UV absorption indices, Sensors 12 (1) (2012) 972–986.

[25] X. Wang, F. Zhang, H.-t. Kung, A. Ghulam, A.L. Trumbo, J. Yang, Y. Ren, Y. Jing, Evaluation and estimation of surface water quality in an arid region based on EEM-PARAFAC and 3D fluorescence spectral index: a case study of the Ebinur Lake Watershed, China, Catena 155 (2017) 62–74.

[26] L. Fu, L. Bin, Z. Luo, Z. Huang, P. Li, S. Huang, D. Nyobe, F. Fu, B. Tang, Spectral change of dissolved organic matter after extracted by solid-phase extraction and its feasibility in predicting the acute toxicity of polar organic pollutants in textile wastewater, J. Hazard Mater. 443 (2023) 130344.

[27] Y. Zhang, Y. Yin, L. Feng, G. Zhu, Z. Shi, X. Liu, Y. Zhang, Characterizing chromophoric dissolved organic matter in Lake Tianmuhu and its catchment basin using excitation-emission matrix fluorescence and parallel factor analysis, Water Res. 45 (16) (2011) 5110–5122.

[28] W.-E. Zhuang, W. Chen, Q. Cheng, L. Yang, Assessing the priming effect of dissolved organic matter from typical sources using fluorescence EEMs-PARAFAC, Chemosphere 264 (2021) 128600.

[29] N. Hudson, A. Baker, D. Ward, D.M. Reynolds, C. Brunsdon, C. Carliell-Marquet, S. Browning, Can fluorescence spectrometry be used as a surrogate for the Biochemical Oxygen Demand (BOD) test in water quality assessment? An example from South West England, Sci. Total Environ. 391 (1) (2008) 149–158.

[30] A. Goffin, S. Guérin-Rechdaoui, V. Rocher, G. Varrault, An environmentally friendly surrogate method for measuring the soluble chemical oxygen demand in wastewater: use of three-dimensional excitation and emission matrix fluorescence spectroscopy in wastewater treatment monitoring, Environ. Monit. Assess. 191 (7) (2019) 421.

[31] Q. Zhou, Y. Zhang, J. Tao, L. Ye, H. Wang, K. Shan, E. Jeppesen, L. Song, Water depth and land-use intensity indirectly determine phytoplankton functional diversity and further regulate resource use efficiency at a multi-lake scale, Sci. Total Environ. 834 (2022) 155303.

[32] W. Wang, X. Yang, L. Huang, J. Qin, Q. Zhou, Attenuation of ultraviolet radiation and photosynthetically active radiation in six Yunnan Plateau lakes of China based on seasonal field investigations, J. Limnol. 79 (2) (2020) 151–163.

[33] L. Zhou, Y. Zhou, Y. Hu, J. Cai, C. Bai, K. Shao, G. Gao, Y. Zhang, E. Jeppesen, X. Tang, Hydraulic connectivity and evaporation control the water quality and sources of chromophoric dissolved organic matter in Lake Bosten in arid northwest China, Chemosphere 188 (2017) 608–617.

[34] C.A. Stedmon, S. Markager, R. Bro, Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy, Mar. Chem. 82 (3) (2003) 239–254.

[35] K.R. Murphy, C.A. Stedmon, P. Wenig, R. Bro, OpenFluor– an online spectral library of auto-fluorescence by organic compounds in the environment, Anal. Methods 6 (3) (2014) 658–661.

[36] C.L. Osburn, L.T. Handsel, M.P. Mikan, H.W. Paerl, M.T. Montgomery, Fluorescence tracking of dissolved and particulate organic matter quality in a river-dominated estuary, Environ. Sci. Technol. 46 (16) (2012) 8628–8636.

[37] C.A. Stedmon, S. Markager, Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis, Limnol. Oceanogr. 50 (2) (2005) 686–697.

[38] K.R. Murphy, C.A. Stedmon, T.D. Waite, G.M. Ruiz, Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy, Mar. Chem. 108 (1) (2008) 40–58.

[39] J.E. Birdwell, A.S. Engel, Characterization of dissolved organic matter in cave and spring waters using UV–Vis absorbance and fluorescence spectroscopy, Org. Geochem. 41 (3) (2010) 270–280.

[40] M. Sepp, T. Kõiv, P. Nõges, T. Nõges, The role of catchment soils and land cover on dissolved organic matter (DOM) properties in temperate lakes, J. Hydrol. 570 (2019) 281–291.

[41] C. Le, C. Hu, D. English, J. Cannizzaro, C. Kovach, Climate-driven chlorophyll-a changes in a turbid estuary: observations from satellites and implications for management, Rem. Sens. Environ. 130 (2013) 11–24.

[42] S.V. Balasubramanian, N. Pahlevan, B. Smith, C. Binding, J. Schalles, H. Loisel, D. Gurlin, S. Greb, K. Alikas, M. Randla, M. Bunkei, W. Moses, H. Nguyễn, M.K. Lehmann, D. O'Donnell, M. Ondrusek, T.-H. Han, C.G. Fichot, T. Moore, E. Boss, Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters, Rem. Sens. Environ. 246 (2020) 111768.

[43] G. Langergraber, N. Fleischmann, F. Hofstädter, A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater, Water Sci. Technol. 47 (2) (2003) 63–71.

[44] M. Philibert, S. Luo, L. Moussanas, Q. Yuan, E. Filloux, F. Zraick, K.R. Murphy, Drinking water aromaticity and treatability is predicted by dissolved organic matter fluorescence, Water Res. 220 (2022) 118592.

[45] A.M. Kellerman, T. Dittmar, D.N. Kothawala, L.J. Tranvik, Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology, Nat. Commun. 5 (1) (2014) 3804.

[46] C. Niu, Y. Zhang, Y. Zhou, K. Shi, X. Liu, B. Qin, The potential applications of real-time monitoring of water quality in a large shallow lake (Lake Taihu, China) using a chromophoric dissolved organic matter fluorescence sensor, Sensors 14 (7) (2014) 11580–11594.

[47] G.J.C. Underwood, C. Michel, G. Meisterhans, A. Niemi, C. Belzile, M. Witt, A.J. Dumbrell, B.P. Koch, Organic matter from Arctic sea-ice loss alters bacterial community structure and function, Nat. Clim. Change 9 (2) (2019) 170–176.

[48] Y. Su, E. Hu, M. Feng, Y. Zhang, F. Chen, Z. Liu, Comparison of bacterial growth in response to photodegraded terrestrial chromophoric dissolved organic matter in two lakes, Sci. Total Environ. 579 (2017) 1203–1214.

[49] L.M. Lynch, N.A. Sutfin, T.S. Fegel, C.M. Boot, T.P. Covino, M.D. Wallenstein, River channel connectivity shifts metabolite composition and dissolved organic matter chemistry, Nat. Commun. 10 (1) (2019) 459.

[50] A. Baker, Fluorescence excitation–emission matrix characterization of some sewage-impacted rivers, Environ. Sci. Technol. 35 (5) (2001) 948–953.

[51] Y. Du, F. Chen, K. Xiao, C. Song, H. He, Q. Zhang, Y. Zhou, K.-S. Jang, Y. Zhang, P. Xing, Z. Liu, Y. Zhang, Y. Lu, Water residence time and temperature drive the dynamics of dissolved organic matter in alpine lakes in the Tibetan Plateau, Global Biogeochem. Cycles 35 (11) (2021) e2020GB006908.

[52] C.P. Ward, S.G. Nalven, B.C. Crump, G.W. Kling, R.M. Cory, Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic pathways and stimulates respiration, Nat. Commun. 8 (1) (2017) 772.

[53] X. Yang, Y. Zhang, Q. Liu, J. Guo, Q. Zhou, Progress in the interaction of dissolved organic matter and microbes (1991–2020): a bibliometric review, Environ. Sci. Pollut. Res. 29 (12) (2022) 16817–16829.

[54] E.A. Pugh, D. Olefeldt, S.N. Leader, K.J. Hokanson, K.J. Devito, Characteristics of dissolved organic carbon in boreal lakes: high spatial and inter-annual variability controlled by landscape attributes and wet-dry periods, Water Resour. Res. 57 (11) (2021) e2021WR030021.

[55] O.E. Senar, I.F. Creed, C.G. Trick, Lake browning may fuel phytoplankton biomass and trigger shifts in phytoplankton communities in temperate lakes, Aquat. Sci. 83 (2) (2021) 21.

[56] J. Hur, B.-M. Lee, T.-H. Lee, D.-H. Park, Estimation of biological oxygen demand and chemical oxygen demand for combined sewer systems using synchronous fluorescence spectra, Sensors 10 (4) (2010) 2460–2471.

[57] K. Khamis, C. Bradley, R. Stevens, D.M. Hannah, Continuous field estimation of dissolved organic carbon concentration and biochemical oxygen demand using dual-wavelength fluorescence, turbidity and temperature, Hydrol. Process. 31 (3) (2017) 540–555.

[58] K. Khamis, C. Bradley, H.J. Gunter, G. Basevi, R. Stevens, D.M. Hannah, Calibration of an *in-situ* fluorescence-based sensor platform for reliable BOD_5 measurement in wastewater, Water Sci. Technol. 83 (12) (2021) 3075–3091.

[59] X. Liu, Y. Zhang, K. Shi, G. Zhu, H. Xu, M. Zhu, Absorption and fluorescence properties of chromophoric dissolved organic matter: implications for the monitoring of water quality in a large subtropical reservoir, Environ. Sci. Pollut. Res. 21 (24) (2014) 14078–14090.

[60] N. Mladenov, A. Bigelow, B. Pietruschka, M. Palomo, C. Buckley, Using submersible fluorescence sensors to track the removal of organic matter in decentralized wastewater treatment systems (DEWATS) in real time, Water Sci. Technol. 77 (3) (2017) 819–828.