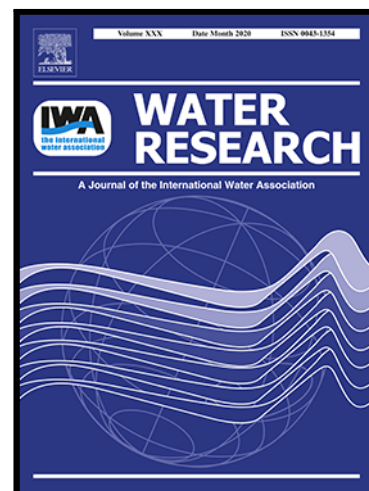# Journal Pre-proof

Natural versus anthropogenic controls on the dissolved organic matter chemistry in lakes across China: insights from optical and molecular level analyses

Yingxin Shang , Zhidan Wen , Kaishan Song , Ge Liu , Fengfa Lai , Lili Lyu , Sijia Li , Hui Tao , Junbin Hou , Chong Fang , Chen He , Quan Shi , Ding He

Please cite this article as: Yingxin Shang , Zhidan Wen , Kaishan Song , Ge Liu , Fengfa Lai , Lili Lyu , Sijia Li , Hui Tao , Junbin Hou , Chong Fang , Chen He , Quan Shi , Ding He , Natural versus anthropogenic controls on the dissolved organic matter chemistry in lakes across China: insights from optical and molecular level analyses, *Water Research* (2022), doi: https://doi.org/10.1016/j.watres.2022.118779

Natural versus anthropogenic controls on the dissolved organic matter chemistry in lakes across China: insights from optical and molecular level analyses

Yingxin Shang[a], Zhidan Wen[a], Kaishan Song*[a,c], Ge Liu[a], , Fengfa Lai[a] , Lili Lyu[a], Sijia Li[a], Hui Tao[a], Junbin Hou[a], Chong Fang[d], Chen He[e], Quan Shi[e], Ding He*[b]

[a] Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China

[b] Department of Ocean Science and Hong Kong Branch of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

[c] School of Environment and Planning, Liaocheng University, Liaocheng 252000, China

[d] Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China

[e] State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Changping District, Beijing 102249, China

* Corresponding authors at: Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China (Kaishan Song); Department of Ocean Science and the Hong Kong Branch of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Hong Kong, China (Ding He)

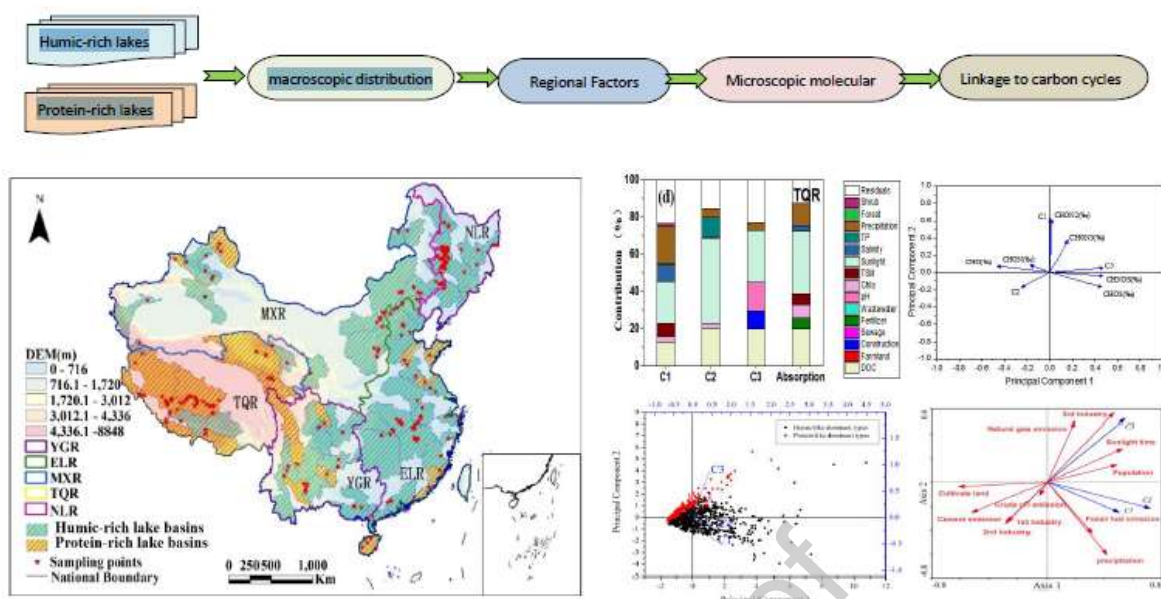E-mail addresses: songks@neigae.ac.cn (K. Song); dinghe@ust.hk (D. He).

Graphical abstract

Highlights

- Optical properties and molecular diversity of DOM across Chinese lakes were analyzed.
- Diverse DOM sources for humic-rich and protein-rich lakes across large-scale regions were firstly revealed.
- Natural and anthropogenic factors affecting regional DOM chemistry were evaluated quantitatively.
- Linkages between DOM optical properties and aquatic carbon parameters in Chinese lakes were established.

Abstract:

Dissolved organic matter (DOM) plays an essential role in the global carbon biogeochemical cycle for aquatic ecosystems. The complexity of DOM compounds contributes to the accurate monitoring of its sources and compositions from large-scale patterns to microscopic molecular groups. Here, this

study demonstrates the diverse sources and compositions for humic-rich lakes and protein-rich lakes for large-scale regions across China with the linkage to optical components and molecular high-resolution mass spectrometry properties. The total fluorescence intensity of colored DOM (CDOM) for humic-rich lake regions (0.176 Raman unit; R.U.) is significantly ($p<0.05$) higher than that of the protein-rich lake region (0.084 R.U.). The combined percentages of CDOM absorption variance explained by the anthropogenic and climatic variables across the five lake regions of Northeastern lake region (NLR), Yungui Plateau lake region (YGR), Inner Mongolia-Xinjiang lake region (MXR), Eastern lake region (ELR), and Tibetan-Qinghai Plateau lake region (TQR) were 86.25%, 82.57%, 80.23%, 88.55%, and 87.72% respectively. The averaged relative intensity percentages of CHOS and CHONS formulas from humic-rich lakes (90.831‰, 10.561‰) were significantly higher than that from the protein-like lakes (47.484‰, 5.638‰), respectively. The more complex molecular composition with higher aromaticity occurred in the humic-rich lakes than in the protein-rich lakes. The increasing anthropogenic effects would significantly enhance the sources, transformation, and biodegradation of terrestrial DOM and link to the greenhouse gas emission and the carbon cycle in inland waters.

1. Introduction

Dissolved organic matter (DOM) with autochthonous and allochthonous origins, as one of the largest pools of carbon cycling in freshwater ecosystems, plays a ubiquitous role in biogeochemical processes and ultimately affects the global carbon cycles (Jaffé et al., 2013; Meng et al., 2013;

Huang et al., 2017). The chromophoric dissolved organic matter (CDOM), a significant fraction of

DOM with optical functions (absorption of the ultraviolet and visible region) and parts with

fluorescent properties, could be used as a potential proxy to assess the organic pollution and human

activities effects at large spatial scales (Chen et al., 2003; Stedmon and Bro, 2008; Shang et al.,

2021). Many regions across China have been undergoing rapid social development and urbanization

(Han et al., 2015; Li et al., 2020; Shang et al., 2021), which would increase the pollutants from waste

discharges and fuel carbon emissions. In addition, CDOM could produce an unpleasant smell to

influence the water quality and form carcinogenic by-products to increase the risk of causing cancer

(Zhou et al., 2015; Shang et al., 2021). However, the complexity of CDOM compounds limits

understanding of the accurate monitoring of CDOM sources and compositions for large-scale

patterns (Ishii and Boyer, 2012; Wang et al., 2015).

CDOM is closely linked to the relative concentration of DOM and is generally used to assess the

relative levels of DOM (especially dissolved organic carbon; DOC) and trace the sources, cycling,

biogeochemical processes via the optical properties of aquatic environments (Yu et al., 2010; Hestir

et al., 2015). Generally, the allochthonous CDOM primarily consists of humic substances, and

autochthonous sources includes more protein-like and tryptophan-like components through terrestrial

input, in-situ primary production, photo-degradation and biological transformation processes in

natural waters (Mostofa et al., 2009). The utilization of excitation-emission matrices (EEMs) of

CDOM with the techniques of parallel factor (PARAFAC) analysis is widely considered to be a

sensitive way to elucidate the qualitative fraction variations of fluorescent CDOM (FDOM)

components for inland waters (e.g., Lyu et al., 2020). In recent decades, Fourier transform ion

cyclotron resonance mass spectrometry (FT-ICR MS) with ultra-high resolving power and accuracy is becoming a powerful tool to study the molecular composition for DOM samples of inland waters (e.g., He et al., 2020a). With the help of the aromaticity index (AI), molecular lability index (MLB), and the classification of multiple molecular groups, including polycyclic aromatics, highly aromatic and unsaturated compounds and saturated compounds, FT-ICR MS is helpful to demonstrate the bioavailability and degree of refractory degradation for aquatic DOM import and degradation processes for carbon biogeochemical cycles (Koch and Dittmar, 2006; He et al., 2020b). Studies have found that the contributions of humic substances account for 14–90 % of total DOM in lakes (e.g., Mostofa et al., 2009). Therefore, many studies have started to address the significant variations in the spatial and temporal CDOM concentration and sources in reservoirs and lakes through the absorption and fluorescence properties (Zhou et al., 2017; Shang et al., 2019). Both humic-like and protein-like components were widely detected in FDOM from various aquatic ecosystems (e.g., lakes, reservoirs, and rivers). However, little is known about the detailed differences in composition sources and factors between humic-like matters and protein-like matters of DOM for large-scale regions. Thus, studies are needed for characterizing the linkage between DOM components and their factors to quantify the role of inland waters in the global carbon cycle from both large-scale regional and molecular perspectives (Wang et al., 2015).

CDOM is one of the most critical optically-active components of DOM. Its chemical composition, sources, and photochemical or microbial processes are affected by natural and anthropogenic stressors (Shank et al., 2010; Shang et al., 2019; Liu et al., 2020; Du et al., 2021). The climate conditions and landscape properties are highly structured in space over regional to global scales, and

a strong influence of these factors should induce spatially structured patterns in lake DOM at large

spatial scales (Hijmans et al., 2005; Lapierre et al., 2016). The significant variations of regional

patterns such as climate factors and geographical locations along with the land use and land cover

change (LULC) would bring terrestrial input of organic pollutants to aquatic systems and also affect

its photochemical transformation (Tzortziou et al., 2015; Yates et al., 2019; Shao et al., 2020).

Moreover, the increase of urbanization development with the LULC would influence the variations

of terrestrial organic composition and microbial degradation processes (Aufdenkampe et al., 2011).

Studies have shown that the land-use types contribute to DOM chemical and compositional

complexity (Wilson and Xenopoulos, 2009; Piirsoo et al., 2012; Lambert et al., 2015). Meanwhile,

human disturbance such as sewage and industrial wastewater discharge also affects the DOM

loadings significantly (Graeber et al., 2015; Yates et al., 2019). Even though these concepts have

been well proposed for past decades, which and to what extent these landscape and environmental

drivers would affect DOM composition and sources in large-scale regions remain elusive. The

unaddressed gap would affect the large-scale spatial structuring of lake carbon cycling (Lapierre et

al., 2016). In the current stage, the study of CDOM distribution of lakes across China generally

centered on regional classical lakes or rivers; the potential factors for the large-scale regional

variations have not been recognized comprehensively. Moreover, the linkages among the CDOM

sources, greenhouse gas emissions (GHGs), and sediment organic carbon burial have not been

investigated in previous studies, which would be critical for considering the regional carbon cycling

models.

In this study, the EEMs with PARAFAC analysis and FT-ICR MS are combined to understand DOM compositions, sources, and the spatial distribution for lakes across China. The main objectives are to (a) evaluate the variations in DOM composition and distribution across national-scale regions, (b) determine the primary climate, landscape, and anthropogenic drivers over large spatial regions across China (c) explore the main factors (natural versus anthropogenic) on the lake DOM chemistry across China.

2. Methods and materials

2.1 Study sites and field sampling

The lakes of China are well classified into five spatial region patterns due to the significant variations in land cover types and climate conditions (Ma et al., 2011; Shang et al., 2019). They are Northeastern lake region (NLR), Inner Mongolia-Xinjiang lake region (MXR), Tibetan-Qinghai Plateau lake region (TQR), Yungui Plateau lake region (YGR), and Eastern lake region (ELR). Most lakes in NLR are located in plain areas, and salinity and agricultural activities would affect the lakes significantly (Song et al., 2019). The MXR region has low precipitation and arid climate conditions in which the sunlight is intensive. The ELR is the most developed region of China, and it is also with a large density of lakes and human population. The YGR is in a high-altitude area and there are a significant proportion of tectonic lakes. The TQR is in the zone of extreme climate with an altitude of over 4km and intense sunlight radiation. The representative lakes were selected based on both natural and anthropogenic conditions as well as watershed characteristics. 1165 samples were

collected from 256 lakes between 2015 and 2019. The in-situ parameter measurement, including GPS, pH, salinity, secchi disk depth (SDD), is carried out using a GPS receiver, YSI 600 XLM Sondes (YSI Inc., Yellow Springs, OH), and Secchi disk. The water samples were filtered as soon as possible at the nights of each sampling date in the hotel immediately for each campaign. The water samples of CDOM and the filter membranes for other water quality parameters (TSM, Chl-a) were stored in the car refrigerator at 4 °C and transported to the laboratory as soon as possible. The sampling sites of lakes across China are shown in Fig. 1.

2.2   Laboratory measurement

2.2.1 Water parameter measurement

The dissolved organic carbon (DOC) concentration was determined with Shimadzu TOC-5000 Analyzer after being filtered with a 0.4 μm microporous membrane filter in the fieldtrips. The Chlorophyll-a concentration (Chla) was measured with the extraction of 90% acetone solution. The total suspended matter (TSM) was measured with gravimetrical methods with the pre-combusted 47 mm Whatman GF/F glass fiber filters. The concentration of total phosphorus (TP) was determined spectrophotometrically after digestion of the samples with alkaline potassium persulfate (Clesceri et al., 1998; Shang et al., 2019). The detailed procedure for water chemical parameter measurement is shown in the document of supplementary information (SI).

2.2.2 CDOM analyses

The water samples were filtered through 0.2 μm polycarbonate membrane (Millipore) filters under a gentle vacuum, and the absorption spectra were measured with a Shimadzu UV-Vis

spectrophotometer from 200 to 800 nm. Meanwhile, the fluorescence excitation-emission matrices (EEMs) were determined using a Hitachi F-7000 fluorescence spectrometer. The detailed analysis procedure for absorption coefficient $a_{CDOM}(254)$, $a_{CDOM}(250):a_{CDOM}(365)$ (M), and the PARAFAC analysis of fluorescence EEMs could be found in SI. The calculation processes of corresponding fluorescence indices, including humification index (HIX), recent biological autochthonous index (BIX) and freshness index ($\beta:\alpha$) are also shown in SI. The SPE-DOM extracts of studied lake samples were analyzed using a 9.4 T Bruker Apex-ultra FT-ICR MS (Bruker Daltonics) equipped with a negative electron spray ionization source. The total molecular numbers, mass range (m/z), CHO(‰), CHON(‰), CHOS(‰), CHONS(‰), molecular lability index (MLB), double bond equivalence (DBE), and modified aromaticity index (AI) were analyzed (Dittmar and Koch, 2006, 2016), in addition to molecular formulae groups described by He et al. (2020). The detailed description is shown in SI.

2.3 Land use, climate condition, and environmental data

To assist with the analysis, a suite of datasets, including land-use data, meteorological data (precipitation and sunlight time), and socio-economic data (wastewater discharge, sewage discharge, and fertilizer usage), were assembled within the sampling catchment boundary. The land-use data were downloaded from GlobeLand 30 (http://www.globallandcover.com/Chinese/GLC30Download/index.aspx) (Fig.S4) and the meteorological data (average monthly precipitation and daily sunlight time) for the sampling months (May to October) from 2015 to 2019 were obtained from the National Meteorological Information Center (http://data.cma.cn/). In addition, the socio-economic data (wastewater discharge, sewage

discharge, and fertilizer usage) at the city (country) level were obtained from China Statistical Yearbook and China Statistical Yearbook on Environment (http://data.cnki.net/Yearbook), and the datasets were transformed to watershed scale level according to the population distribution. The more specific information is provided in the SI.

2.3 Statistical analysis

Spatial mapping of sampling sites and the types of DOM sources are conducted with ArcGIS 10.1. The PARAFAC analysis for EEMs was carried out in MATLAB R2016a (Mathworks, USA), and the three components were compared online through the software OpenChrom (Certain coefficient > 0.9) (Stedmon and Bro, 2008). Principal components analysis (PCA) was performed via Origin 9.0 (Microcal Software, Inc., MA) using the three PARAFAC components. Redundancy analysis (RDA) (CANOCO program; Ter Braak and Šmilauer 2002) was used to visualize relationships between CDOM indices and natural and anthropogenic conditions. Analysis of the contribution of driving factors to trends in CDOM and fluorescence components of FDOM for different regions is conducted through random forest regression with R studio. In addition, a spearman correlation analysis was conducted using SPSS Statistics 22.0 to explore relationships between humic-rich sources and protein-rich sources of lake DOM across China. Meanwhile, results with p values lower than 0.05 were reported as significant in the linear regression and one-way ANOVA analyses.

[Insert Fig.1 about here]

3. Results

3.1 The distribution of CDOM absorption and fluorescence intensities

In this study, the EEM-PARAFAC analysis of the lakes across China is carried out, and three DOM components are identified, including C1 (humic-like component, Ex/E_M: 250nm/438nm), C2 (humic-like component; Ex/E_M: 230nm/430nm), and C3 (protein-like component; Ex/E_M: 280nm/325nm). C1 component is similar to humic-like C peak, which is widely observed in natural waters (Coble et al., 1996, 2007; Stedmon et al., 2011). C2 component is considered as the fulvic origin or the terrestrially derived humic-like products (Stedmon and Markager, 2005; Massicotte and Frenette, 2011; Kothawala et al., 2015). C3 component in the UVA region overlaps the peak T for tryptophan protein-like sources (Coble, 1996; Walker et al., 2013). The lakes are operationally defined as humic-rich lakes and protein-rich lakes according to the PARAFAC result based on decision tree method and the consideration of land use% and the Heihe-Tengchong Line. The detailed method is shown in the SI. The total fluorescence intensity (Ft) for humic-rich lake regions (0.176 ± 0.032 R.U.) is significantly higher than that of the protein-rich lake region (0.084 ± 0.011 R.U.) ($p<0.05$). The comparison of fluorescence intensity for C1, C2, and C3 among regions is shown in Fig.2. Significant differences between humic-rich and protein-rich lakes are observed among and within the regions ($p<0.05$). The mean $a_{CDOM}(254)$ ranged from 13.95 to 44.71 m^{-1} among five lake regions. The mean $a_{CDOM}(254)$ in humic-rich lakes (26.54 m^{-1}) is significantly higher than that of protein-rich lakes (16.32 m^{-1}) in the same regions (Fig.2b). The significant differences in $a_{CDOM}(254)$ in humic-rich lakes are observed among regions. In terms of seasons, humic-rich lakes are dominant for spring, summer, and autumn (Figure S5). In terms of the individual lake, we have done the result comparisons for eight lakes for four seasons (Figure S6). It is

found that the fluorescence components of DOM of a single lake changed for different

seasons, which is likely caused by the changing hydrological and climate conditions for

small-scale regions. And this study would focus on the external annual natural and

anthropogenic factors for average DOM components across four seasons in large-scale

regions.

[Insert Fig.2 about here]

3.2   The PCA of CDOM fluorescence properties of lakes across China

The fluorescent components (C1, C2, C3) for all lake water samples in five lake regions across

China, which help determine the degree of separation between the humic-rich lakes and

protein-rich lakes, were further investigated by PCA. The two PCA axes (PC1 and PC2)

explained 93.73% of the total variance in the dataset, with component 1 and component 2

accounting for 67.03% and 26.70%, respectively (Fig.3a). The protein-like component C3 shows

the positive loadings on PC2 with more contribution of sunlight. In contrast, the humic-like

components C1 and C2 show the negative loadings on PC2 with more allochothonous sources

and human disturbance, which could be used to address the apparent separation of the variances.

Regarding the classification of lake types, it is evident that the protein-rich lakes are centered in

the same direction as C3, while the humic-rich lakes are focused on the same trend of C2 and C1.

The PC1 and PC2 show a general clustering of humic-rich types and protein-rich lakes, with

component 1 scores in the range of 0.75 to 4.5 and component 2 scores ranging from -1.0 to 1.5

(Fig.3). In terms of lake regions, the protein-rich samples from TQR lakes were generally

clustered along the positive axis of component 1, whereas the water samples from the other four

region lakes were generally scattered, with high negative component 2 scores (Fig.3b).

[Insert Fig.3 about here]

3.3   The composition and sources of CDOM in lakes across China

[Insert Table 1 about here]

The composition and sources of DOM would be reflected by the CDOM absorption and fluorescence

properties (M, BIX_{310}, HIX), including the overall aromaticity, the proportion of autochthonous and

allochthonous DOM, and aquatic biological activity. The significant differences for averaged M,

autochthonous index (BIX_{310}), and HIX are respectively observed between the humic-rich lakes and

protein-rich lakes ($p<0.05$), and the values are shown in Table 1. Meanwhile, for the lake regions, the

comparison of humic-rich types and protein-rich types is shown in Table 2. The averaged HIX, M, and

BIX_{310} in TQR are significantly different from that of other lake regions (NLR, ELR, MXR, and YGR)

($p<0.05$). The significant variances of averaged HIX, M, and BIX_{310} between humic-like types and

protein-like types in the same regions are shown as well ($p<0.05$). For both humic-rich types and

protein-rich types, the fluorescence components C1, C2, and C3 show an important relation to CDOM

absorption, DOC concentration, and M ($p<0.01$). The M value is negatively correlated to the C1

component but is positively correlated to the C3 component. The humic-like C1 is significantly related to

fluorescent indices in humic-like lakes, while protein-like component C3 is more related to HIX and

BIX_{310} in protein-rich lakes ($p<0.01$, Table 3).

[Insert Table.2 about here]

[Insert Table.3 about here]

3.4 The main factors contributing to CDOM composition

The natural factors (precipitation, sunlight), anthropogenic factors (wastewater discharge, sewage discharge, and fertilizer usage), land use proportion, and aquatic water quality parameters (*Chla*, TSM, pH, salinity, etc.) would mutually contribute to the variance of DOM composition and sources that could be reflected on the changes of CDOM absorption and fluorescence intensities. The variability of CDOM absorption and fluorescence in various regions can result from multiple socio-economic and climatic factors (Fig.4). The combined percentages of CDOM absorption variance explained by the anthropogenic and climatic variables across the five lake regions of NLR, YGR, MXR, ELR, and TQR were 86.25%, 82.57%, 80.23%, 88.55%, and 87.72%, respectively. Meanwhile, the percentages of humic-like and protein-like variances (C1, C2, and C3) explained by the anthropogenic and climatic variables in different regions are higher than 65%. For the NLR region, CDOM absorption and fluorescence variances are mostly affected by DOC concentration, farmland%, fertilizer usage, and precipitation. For YGR, except DOC, precipitation, and sunlight contribute to the fluorescent component of DOM significantly. In terms of MXR, DOC concentration, *Chla* concentration, construction%, and sewage discharge affect the CDOM component primarily. In the region of ELR, the percentages of the construction area and sewage discharge are critical factors to protein-like components. However, in the high-altitude area of TQR, sunlight time, salinity, and precipitation contribute to the humic-like components and protein-like components with a large

proportion significantly.

[Insert Fig.4 about here]

3.5 Characterization of Molecular DOM complexity by FT-ICR MS

The FT-ICR MS analyses for inland lakes across China showed that 10128 and 9071 formulas (on average) were detected as significantly different from the humic-rich lakes and protein-rich lakes, respectively (Table S1, $p<0.05$). The CHO formulae were most abundant regarding peak count, followed by CHON, CHOS, and CHONS formulae for all lakes (Table S1). The averaged relative intensity percentage of CHOS, CHONS formula from humic-rich lakes (90.831‰, 10.561‰) was significantly higher than that from the protein-rich lakes (47.484‰, 5.638‰) respectively (Fig.5). The comparison for molecular parameters (Table S1) demonstrated that higher averaged m/z, AI, and low MLB occurred in the humic-rich lakes. Meanwhile, in terms of the compound-grouped fractions, higher averaged abundances of black carbon (‰), polyphenols (‰), highly unsaturated (‰), and CRAM (‰) occurred in humic-rich lakes while higher unsaturated aliphatics compounds (‰), peptides (‰), sugars (‰), and saturated fatty acids (‰) in protein-rich lakes. The PCA results explained 73.51% of the total variances of the proportion of intensity-weighted parameters (CHO, CHON, CHOS, etc.), while 60.16% of the total variances of the formulae groups could be explained with PC1 and PC2 (Fig.6).

[Insert Fig.5 about here]

[Insert Fig.6 about here]

4. Discussion

4.1 The natural conditions contributed to spatial variations in DOM sources

The effects of geographical domains and meteorology on the DOM composition and absorption coefficients in inland waters have been recognized (Fig.4). The significant correlation between C2 and HIX in humic-rich lakes and the similar correlation between C3 and BIX_{310} show the variability of DOM sources in different types of lakes (Table 3, $p<0.05$). In the same regions, the humic-rich lakes trends to contain more aromatic organic compounds and higher humification levels compared with the local protein-rich lakes, which illustrates that the protein-rich lakes have more newly-produced DOM with high biological activities from aquatic biodegradation (Tables 1 and 2) (Helms et al., 2009; Zhou et al., 2018; Shi et al., 2020). The surrounding soil type and soil nutrient would influence the CDOM composition of lakes through allochthonous humic-like substances (Butman et al., 2014). In terms of climatic conditions, the mean monthly precipitation value is more correlated to C2 and M in humic-rich lakes (r_{C2}=0.675, r_M=-0.643, $p<0.01$) than that of protein-rich lakes (r_{C1}=0.305, r_M=-0.411, $p<0.05$), which demonstrate the humification and export of humic-like materials would increase with the rainfall than that of protein-like materials especially in humic-rich lakes during the rainfall events (Song et al., 2019; Lyu et al., 2021). The significant contribution of precipitation to humic-like C1 components in five regions illustrates that surface runoff likely increases the export of terrestrial DOM from the surface soil layer (Shi et al., 2020; Lee et al., 2021). Generally, the mean monthly precipitation of humic-rich lakes (99.5 ± 13.1 mm) is significantly

higher than that of protein-rich lakes (83.7 ± 35.6 mm), which would demonstrate that the fast flushing after intense precipitation would deliver more humic-like matters with the increasing water residence time (declined groundwater level) and microbial degradation activities from soils (Lyu et al., 2021).

Most protein-rich lakes are located in high-altitude areas or with strong sunlight intensity. The significant contribution of sunlight duration to the decreasing C1 components (rC1=0.532, p<0.05) in protein-rich lakes indicates that solar radiation would decrease the humic-like fluorescent components of CDOM significantly and increase the relative content of protein-like components (Shang et al., 2019). The correlation coefficient between sunlight duration and decreasing HIX in protein-rich lakes (slope=-0.662, p<0.05) is higher than that of humic-rich lakes (slope=-0.452, p<0.05), which illustrate the humification level tended to decrease sharply in the protein-rich lakes for strong sunlight with the preferred degradation processes and the release of small molecular weight CDOM (Mostofa et al., 2009; Shao et al., 2020).

4.2 The influences of anthropogenic factors on DOM variations in lake regions

The DOM in lakes is generally affected by anthropogenic factors, including the land use of watersheds and human activities besides the natural conditions (Sobek et al., 2007; Xia and Zhang, 2011). In this study, the DOM absorption and fluorescent components are affected by various anthropogenic activities and land use types in different regions. The percentage of land use for humic-rich lakes is significantly higher than that of protein-rich lakes (p<0.05). The protein-rich lakes are distributed in the regions with more bare land area %, and the C3 component is

significantly related to bare land area% (r=0.67, p<0.05). Also, according to the Heihe-Tengchong population Line (Figure S4), the protein-rich lakes are generally distributed in the northwest of the line (87%) and the humic-rich lakes are generally distributed in the southeast of the line (83%). In the eastern regions of China, increasing anthropogenic land-use %farmland and fertilizer usage contributes significantly to the terrestrial humic-like C1 component and CDOM concentration, which indicate the agricultural activities as the dominant land use type to export more terrestrial humic-like DOM and nutrients to lakes across these regions (Antony et al., 2017; Morling et al., 2017; Roebuck et al., 2020; Sankar et al., 2020). Moreover, our result that the %construction and sewage discharge has a higher contribution to the protein-like DOM component across different regions indicates human activities are essential sources of importing FDOM to lakes with rapid social development (Meng et al., 2013; Cohen et al., 2014; Liu et al., 2020).

Human activities likely contribute both allochthonous (terrestrial) and autochthonous DOM to all lakes for the observed tight correlation between C1 (C3) and sewage (wastewater discharge) in all sampled lakes (p<0.05). The decreasing M with increasing sewage discharge for protein-rich lakes (p<0.05) indicated that the DOM in protein-rich lakes is contributed by urbanization significantly besides the natural sunlight conditions (Liu et al., 2020). Therefore, anthropogenic factors and natural factors such as land use, human activities, precipitation, and sunlight contribute to variations for DOM components between humic-rich lakes and protein-rich lakes comprehensively for the overall trend. However, the detailed seasonal variations in DOM components should consider seasonal climate and hydrological conditions for different locations.

4.3 The linkage between molecular composition and DOM sources

With the optical and molecular composition obtained in the same study, the linkage between optical

and molecular composition is explored. The high S-contained formulae (CHOS and CHONS) and

N-contained formulae (CHON2 and CHON3) for humic-rich lakes (Table S1) also suggest the high

abundances of heteroatoms organic matter input from agricultural activities and urbanization in the

effluent (Gonsior et al., 2017; He et al., 2019). Meanwhile, the correlations between C3 and

S-contained formulae (N-contained formulae) (Fig. 6a) indicate the strong coupling between the

production of newly-produced autochthonous organic matter and the anthropogenic derived sources

in protein-rich lakes. The humic-like components C1 and C2 are highly correlated to formulae

groups usually with higher molecular weight, including black carbon-like compounds, polyphenols,

and highly unsaturated compounds. The protein-rich lakes are characterized by high percentages of

smaller molecular peptides, sugars, and saturated fatty acids, which are possibly photo-degraded

products with higher biolability and strong autochthonous algal or phytoplankton activities and

microbial processes (Stubbins and Dittmar, 2015; D'Andrilli et al., 2015). The strong autochthonous

signature from the molecular level in protein-rich lakes is also consistent with higher BIX values

observed. The significant negative correlation between C3 and AI ($p<0.05$) suggests that the

newly-produced molecular formulas from protein-like lakes tend to be more bioavailable for

microbial activity. In contrast, the humic-rich lakes contain more aromatic formulas with a higher

degree of aromaticity and humification. The terrestrial humic-like component C2 is closed related to

the percentage of black C and polyphenols (Fig. 6b), which indicates the high aromatic levels for

anthropogenic input from human sewage-related organic matter or vascular plant-derived terrestrial

organic matters (Jaffé et al., 2013; Seidel et al., 2015; He et al., 2019, 2020b) for humic-rich lakes.

The significant correlation ($p<0.05$) between C2 and AI (HIX) indicates that the humic-like organic

matter with a high mass range was dominant in DOM compositional compounds. Meanwhile, higher percentages of O_3S and O_5S compounds, and unsaturated aliphatic compounds in protein-rich lakes demonstrated that the sewage and crop runoff or surfactant-like compounds from human activities likely contributed significantly to protein-like CDOM in addition to the freshwater phytoplankton or algae-derived autochthonous DOM (He et al., 2019; Wang et al., 2019). Variations of complex DOM sources with severe anthropogenic input for inland lakes across China are characterized on both fluorescence and molecular levels. Therefore, combining multiple complementary methods could provide a better perspective of DOM cycling in inland water at the regional scale.

4.4 Implications for carbon cycling in inland waters and further considerations

[Insert Fig.7 about here]

Freshwaters are considered as an important source of greenhouse gases (GHGs) (Li et al., 2018), generally released from soils to surface waters and rapidly evaded into the atmosphere as physical gas exchange (Öquist et al., 2009). The changes in CDOM sources and compositions are related to the biogeochemical cycle and energy flow of ecosystems (Williams et al., 2016) and are involved in the carbon budget, potentially affecting GHGs emissions (Zhou et al., 2018; Emilson et al., 2018). However, the variations of CDOM sources within regions due to the various anthropogenic and natural stressors would affect the spatiotemporal variability of aqueous carbon processing (Li et al., 2020). In our study, the carbon emissions for administrative regions of humic-rich lakes are significantly higher than those of the protein-rich lakes (Fig. 7, $p<0.05$), which indicates that fuel

carbon emissions (CO_2 emissions for fossil fuel raw coal, crude oil, natural gas, cement) along with economic development would affect the CDOM fluorescence as the fossil fuel compounds and secondary organic aerosol from terrestrial imports are detected in waters (Mitra et al., 2017). As the humic-rich lakes are generally located in the east of China (40% of the population from The seventh Census of China) with rapid urbanization, this result indicates that humic-rich lakes trends to be formed in the regions with more anthropogenic activities, while protein-rich lakes are located in the western regions (27% of the population from The seventh Census of China) with a low density of population and few human activities with extreme climate conditions.

Based on the RDA analysis of natural and anthropogenic conditions for CDOM fluorescence components, the RDA analysis explains 69.2% of the variations. The humic-like CDOM are positively correlated to fossil fuel gas emission and precipitation, while sunlight contributes to the protein-like C3 components (Fig.8a). The significant correlation between humic-like components and fossil fuel gas emission ($p<0.05$) indicates that the anthropogenic carbon emission is a path to import organic matter to water surface through the land-atmospheric and water-atmospheric exchange (Yu et al., 2018). The humic-like components are positively related to the primary and secondary industries that generally are the dominant industry structure in less developed regions with more usage of fossil fuel, which demonstrates that the composition of DOM could be a potential proxy to reflect the intensity of regional industry structure and levels of economic development (Shang et al., 2021).

To explore the relationships between CDOM fluorescent components and the aqueous carbon cycles, the datasets for DOC storage, annual CO2 equivalents, carbon burial capacity, and organic carbon

accumulation rate of five lake regions of China are collected from previous studies (Table 4). The RDA analysis explains 65.1% of the variations (Fig.8). The humic-like components are positively correlated to the organic carbon accumulation rate (OCAR), and the OCAR is previously demonstrated to be related to socio-economics and temperature (Yu et al., 2018). We suggest that DOM derived from human activities is important for carbon burial in lakes (Stackpoole et al., 2014). For instances, the DOM compounds with carboxylic and hydroxylic groups (usually biologically recalcitrant) are preferred to be adsorbed on mineral surface of particles and finally deposited in sediments (Kothawala et al., 2012). Moreover, previous research showed that there was a positive relationship between the contents of recalcitrant DOC and total OC in sediments (Schmidt et al., 2009, Wang et al., 2021c). Therefore, our result indicates that the humic-rich lakes trend to become increasing carbon pools along with increasing anthropogenic activities and precipitation (Heathcote et al., 2015). Meanwhile, the humic-like component C1(or C2) is related to annual CO_2 equivalents, indicating that humic substances with high terrestrial influence promote more heterotrophic CO_2 production than in-lake production, acting as a nutrient subsidy to suppressing primary production due to light attenuation (Allesson et al., 2020). The highest OCAR for ELR and lowest OCAR for TQR observed here (Table 4) and in a previous study (Zhang et al., 2017) are likely caused by that the developed region of ELR has gone through more soil erosion, river damming, and eutrophication (Kastowski et al., 2011), while the TQR is located at high altitude (with low annual temperature) and has low primary productivity. The significant amount of DOC storage and annual CO_2 equivalents in TQR is likely related to the terrestrial import based on climate change (Ran et al., 2021). Therefore, the carbon cycles for Chinese inland waters must consider the significant regional variations observed in this study. According to previous studies (Table 4), the sum of DOC storage and

carbon burial capacity for five lake regions exceed the annual CO_2 equivalents, which may demonstrate the importance of Chinese lakes for carbon sink (Keller et al., 2021) and more accurate measurements of organic carbon sources, gas emission, and burial capacity are needed for better constraining the carbon cycles for both regional and global scales.

[Insert Fig.8 about here]

[Insert Table.4 about here]

5. Conclusions

This study demonstrates the variations of DOM sources and compositions for lakes across China from large-scale distribution to molecular-scale analysis. Significant variations in optical properties and molecular composition, including CDOM absorption properties, fluorescent parameters, and molecular complexity, are observed between protein-rich lakes and humic-rich lakes. The significant spatial variation of factors influencing CDOM components in different regions is mainly caused by natural conditions and anthropogenic activities. However, the dominant influencing factors differ across five regions and between protein-like and humic-rich lakes. The complex molecular variations for humic-rich lakes and protein-rich lakes are observed, and the alterations of spatial sources and molecular composition of DOM in inland lakes are linked to protein-like components and humic-like components. These spatial variations are essential to consider to better quantify the role of Chinese inland waters in the global carbon cycle. The combination of multiple complementary methods in this study provides a comprehensive perspective of DOM cycling in inland water at the regional scale, and helps predict the changes in DOM sources with the increasing anthropogenic input at

larger globalization scales. Nevertheless, further studies need to explore the source-specific markers and the degradation process (photochemical or microbial degradation) related to molecular abundances at the regional scale to better understand the DOM signatures with the complex linkage of natural and anthropogenic influences.

Declaration of interests

☒ 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.

Reference:

Allesson, L., Koehler, B., Thrane, J., Andersen, T., Hessen, D., 2020. The role of photomineralization for CO_2 emissions in boreal lakes along a gradient of dissolved organic matter. Limnol. Oceanogr.158-170

Antony, R., S Willoughby, A., Grannas, A., Catanzano, V., L Sleighter, R., Thamban, M., Hatcher, P., Nair, S., 2017. Molecular insights on dissolved organic matter transformation by supraglacial microbial communities. Environ. Sci. Technol. 51(8), 4328-4337.

Aufdenkampe, A.K., Mayorga, E., Raymond, P.A., 2011. Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. Front. Ecol. Environ. 9, 53–60.

Butman, D.E., Wilson, H.F., Barnes, R.T., Xenopoulos, M.A., Raymond, P.A., 2014. Increased mobilization of aged carbon to rivers by human disturbance. Nat. Geosci. 8(2), 112-116.

Stedmon, C.A., Thomas, D.N., Granskog, M., Kaartokallio, H. Papadimitriou, S., Kuosa, H., 2007. Characteristics of dissolved organic matter in Baltic coastal sea ice: allochthonous or autochthonous origins? Environ. Sci. Technol. 41, 7273-7279.

Chen, W., Westerhoff, P., Leenheer, J.A., Booksh, K., 2003. Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. Environ. Sci. Technol. 37(24), 5701-5710.

Clesceri, L.S., Greenberg, A.E., Eaton, A.D., 1998. Standard Methods for the Examination of Water and Wastewater, 20th ed. APHA, AWWA, and WEF, Washington, DC.

Coble, P.G., 1996. Characterisation of marine and terrestrial DOM in seawater using excitation eemission matrix spectroscopy. Mar. Chem. 51, 325-346.

Coble, P.G., 2007. Marine optical biogeochemistry: the chemistry of ocean color. Chemical Reviews 107, 402-418.

Cohen, E., Levy, G.J., Borisover, M., 2014. Fluorescent components of organic matter in wastewater: efficacy and selectivity of the water treatment. Water Res. 55, 323-334.

D'Andrilli, J., Cooper, W.T., Foreman, C.M., Marshall., A.G., 2015. An ultrahigh‑resolution mass spectrometry index to estimate natural organic matter lability. Rapid Commun. Mass Spectrom. 29:2385-2401.

Du, Y.X., Lu, Y.H., Roebuck, J.A., Liu, D., Chen, F.Z., Zeng, Q.F., Xiao, K., He, H., Liu, Z.W., Zhang, Y.L. and Jaffe, R., 2021. Direct versus indirect effects of human activities on dissolved organic matter in highly impacted lakes. Sci. Total Environ. 752, 141839.

Emilson, E.J., Carson, M.A., Yakimovich, K.M., Osterholz, H., Dittmar, T., Gunn, J., Mykytczuk, N., Basiliko, N., Tanentzap, A.J., 2018. Climate-driven shifts in sediment chemistry enhance methane production in northern lakes. Nat. Commun. 9 (1), 1801.

Meng F., Huang, G., Yang, X., Li, Z., Jian, L., Jing, C., 2013. Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers. Water Res. 47(14), 5027-5039.

Fellman, J.B., Hood, E., Spencer, R.G.M., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: A review. Limnol. Oceanogr. 55(6), 2452-2462.

Fellman, J.B., Hood, E., D'Amore, D.V., Edwards, R.T., White, D., 2009. Seasonal changes in the chemical quality and biodegradability of dissolved organic matter exported from soils to streams in coastal temperate rainforest watersheds. Biogeochem. 95 (2), 277-293.

Gonsior, M., Luek, J., Schmitt-Kopplin, P., Grebmeier, J.M., Cooper, L.W., 2017. Optical properties and molecular diversity of dissolved organic matter in the Bering Strait and Chukchi Sea. Deep Sea Research Part II: Topical Studies in Oceanogr. 144, 104-111.

Graeber, D., Boëchat, I.G., Encina-Montoya, F., 2015. Global effects of agriculture on fluvial dissolved organic matter. Sci. Rep. 5, 16328.

Han, N., Yu, W.Y., 2015. A quantitative analysis of the impact of China's industrial structure on environmental pollution. Stat. Decis. 440 (20), 133-136.

He, D., He, C., Li, P., Zhang, X., Shi, Q., Sun, Y., 2019. Optical and molecular signatures of dissolved organic matter reflect anthropogenic influence in a coastal river, Northeast China. J. Environ. Qual. 48(3), 603-613.

He, C., Zhang, Y., Li, Y., Zhuo, X., Li, Y., Zhang, C., Shi, Q., 2020a. In-House Standard Method for Molecular Characterization of Dissolved Organic Matter by FT-ICR Mass Spectrometry. ACS Omega 5(20), 11730-11736.

He, D., Wang, K., Pang, Y., He, C., Li, P., Li, Y., Xiao, S., Shi, Q., Sun, Y., 2020b. Hydrological management constraints on the chemistry of dissolved organic matter in the Three Gorges Reservoir. Water Res. 187, 116413.

Heathcote, A.J., Anderson, N.J., Prairie, Y.T., Engstrom, D.R., Del Giorgio, P.A., 2015. Large increases in carbon burial in northern lakes during the Anthropocene. Nat. Commun. 6, 10016.

Helms, J.R., Stubbins, A., Ritchie, J.D., Minor, E.C., Kieber, D.J., Mopper, K., 2009. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr. 54 (3), 1023

Hestir, E.L., Brando, V., Campbell, G., Dekker, A., Malthus, T., 2015. The relationship between

dissolved organic matter absorption and dissolved organic carbon in reservoirs along a temperate to tropical gradient. Remote Sens. Environ. 156, 395-402.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25, 1965-1978.

Hu, Q., Wang, H., Li, J., Hu, C., Zhang, Y., Yu, T., Zhang, Y., 2015. Various inflows to Taihu lake in autumn: spectroscopy characteristics and DOM flux. Environ. Sci. Technol. 38, 152-158.

Hu, Y., Lu, Y., Edmonds, J., Liu, C., Wang, S., Das, O., Liu, J., Zheng, C., 2016. Hydrological and land use control of watershed exports of DOM in a large Arid River basin in northwestern China. J. Geophys Res.-Biogeo. 121 (2).

Huang, C., Zhang, L., Li, Y., Lin, C., Huang, T., Zhang, M., Zhu, A.X., Yang, H., Wang, X., 2017. Carbon and nitrogen burial in a plateau lake during eutrophication and phytoplankton blooms. Sci. Total Environ. 296–304.

Ishii, S.K.L., Boyer, T.H., 2012. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineered systems: a critical review. Environ. Sci. Technol. 46 (4), 2006-2017.

Jaffé, R., Cawley, K.M., Yamashita, Y., 2014. Applications of Excitation Emission Matrix Fluorescence with Parallel Factor Analysis (EEM-PARAFAC) in Assessing Environmental Dynamics of Natural Dissolved Organic Matter (DOM) in Aquatic Environments: A Review. 1160, 27-73.

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

Kastowski, M., Hinderer, M., Vecsei, A., 2011. Long-term carbon burial in European lakes: analysis and estimate. Glob. Biogeochem. Cycles. 25, GB3019.

Keller, P S, Marcé, R, Obrador, B., 2021, Global carbon budget of reservoirs is overturned by the quantification of drawdown areas. Nat. Geosci., 14: 402-408.

Koch, B.P., Dittmar, T., 2006. From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun. Mass Spectrom. 20, 926–932.

Kothawala, D.N., von Wachenfeldt, E., Koehler, B., Tranvik, L.J., 2012b. Selective loss and preservation of lake water dissolved organic matter fluorescence during long-term dark incubations. Sci. Total Environ. 433, 238–246.

Kothawala, D.N., Ji, X., Laudon, H., Ågren, A., Futter, M.N., Köhler, S.J., Tranvik, L.J., 2015. The relative influence of land cover, hydrology, and in-stream processing on the composition of dissolved organic matter in boreal streams, J. Geophys. Res. Biogeo. 120, 1491-1505,

Lambert, T., Darchambeau, F., Bouillon, S., et al., 2015. Landscape control on the spatial and temporal variability of chromophoric dissolved organic matter and dissolved organic carbon in large African rivers. Ecosys. 18, 1224–1239.

Lapierre, J.F., Seekell, D.A., Giorgio, P.D., 2016. Climate and landscape influence on indicators of lake carbon cycling through spatial patterns in dissolved organic carbon. Global Change Biol. 21(12), 4425-4435.

Lapierre, J.F., Guillemette, F., Berggren, M., Del Giorgio, P.A., 2013. Increases in terrestrially derived carbon stimulate organic carbon processing and CO_2 emissions in boreal aquatic ecosystems. Nat. Commun. 4 (1), 2972.

Lee, M.H., Osburn, C.L., Shin, K.H., 2018. New insight into the applicability of spectroscopic

indices for dissolved organic matter (DOM) source discrimination in aquatic systems affected by

biogeochemical process, Water Res., 147,164-176.

Lee, M.H., Lee, Y.K., Derrien, M., Choi, K., Shin, K.H., Jang, K.S., 2019. Evaluating the

contributions of different organic matter sources to urban river water during a storm event via optical

indices and molecular composition. Water Res., 165(15), 115006.1-115006.10.

Li, S., Bush, R. T., Santos, I. R., Zhang, Q., Song, K., Mao, R., 2018. Large greenhouse gases

emissions from china's lakes and reservoirs. Water Res. 147, 13-24.

Liu, D., Du, Y.X., Yu, S.J., Luo, J.H., Duan, H.T., 2020. Human activities determine quantity and

composition of dissolved organic matter in lakes along the Yangtze River. Water Res. 168, 115132.

Lyu, L., Liu, G., Shang, Y., Wen, Z., Hou, J., Song, K., 2021. Characterization of dissolved organic

matter (DOM) in an urbanized watershed using spectroscopic analysis. Chemosphere 277, 130210.

Lyu, L., Wen, Z., Jacinthe, P.-A., Shang, Y., Zhang, N., Liu, G., Fang, C., Hou, J., Song, K., 2020.

Absorption characteristics of CDOM in treated and non-treated urban lakes in Changchun, China.

Environ. Res. 182, 109084.

Zhu M., Zhu G., Zhao L., Yao X., Zhang Y., Gao G., Qin B., 2013. Influence of algal bloom

degradation on nutrient release at the sediment–water interface in Lake Taihu, China, Environ. Sci.

Pollut. Res. 20, 1803-1811.

Ma, R., Yang, G., Duan, H., Jiang, J., Wang, S., Feng, X., Li, A., Kong, F., Xue, B., Wu, J., Li, S.,

2011. China's lakes at present: Number, area and spatial distribution. Sci. China-Earth Sci. 54(2),

283-289.

Massicotte, P., Frenette, J.-J., 2011. Spatial connectivity in a large river system: resolving the sources

and fate of dissolved organic matter, Ecol. Appl. 21, 2600-2617.

Meng, F., Huang, G., Yang, X., Li, Z., Li, J., Cao, J., Wang, Z., Sun, L., 2013. Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers. Water Res. 47 (14), 5027-5039.

Morling, K., Herzsprung, P., Kamjunke, N., 2017. Discharge determines production of decomposition of and quality changes in dissolved organic carbon in pre-dams of drinking water reservoirs. Sci. Total Environ. 577 (15), 329–339.

Mostofa, K.M.G., Wu, F.C., Yoshioka, T., Sakugawa, H., Tanoue, E., 2009. Dissolved organic matter in the aquatic environments. In: Wu, F.C., Xing B (Ed), Natural Organic Matter and its Significance in the Environment. Science Press, Beijing, 3-66

Oquist, M. G., Wallin, M., Seibert, J., Bishop, K., Laudon, H., 2009. Dissolved inorganic carbon export across the soil/stream interface and its fate in a boreal headwater stream. Environ. Sci. Technol. Lett. 43, 7364-7369.

Piirsoo, K., Viik, M., Kõiv, T., 2012. Characteristics of dissolved organic matter in the inflows and in the outflow of Lake Võrtsjärv, Estonia. J. Hydrol. 475, 306-313.

Ran, L., Butman, D.E., Battin, T.J., 2021. Substantial decrease in CO2 emissions from Chinese inland waters due to global change. Nat. Commun. 12, 1730.

Roebuck, J.A., Seidel, M., Dittmar, T., Jaffé, R., 2020. Controls of land use and the river continuum concept on dissolved organic matter composition in an anthropogenically disturbed subtropical watershed. Environ. Sci. Technol. 54 (1), 195-206.

Sankar, M.S., Dash, P., Lu, Y., Mercer, A.E., Turnage, G., Shoemaker, C.M., Chen, S., Moorhead, R.J., 2020. Land use and land cover control on the spatial variation of dissolved organic matter across 41 lakes in Mississippi, USA. Hydrobiologia 847 (4), 1159-1176.

Seidel, M., Yager, P.L., Ward, N.D., Carpenter, E.J., Gomes, H.R., Krusche, A.V., Richey, J.E., Ditmar, T., Medeiros, P.M., 2015. Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. Mar. Chem. 177, 218-231.

Shang, Y., Song, K., Jacinthe, P.A., Wen, Z., Lyu, L., Fang, C., Liu, G., 2019. Characterization of CDOM in reservoirs and its linkage to trophic status assessment across China using spectroscopic analysis. J. Hydrol. 576, 1-11.

Shang, Y., Song, K., Jacinthe, P.-A., Wen, Z., Zhao, Y., Lyu, L., Fang, C., Li, S., Liu, G., Hou, J., Zhang, N., 2021. Fluorescence spectroscopy of CDOM in urbanized waters across gradients of development/industrialization of China. J. Hazard. Mater. 415, 125630.

Shank, G.C., Lee, R., Vähätalo, A., Zepp, R.G., Bartels, E., 2010. Production of chromophoric dissolved organic matter from mangrove leaf litter and floating Sargassum colonies. Mar. Chem. 119(1), 172-181.

Shao, T., Wang, T., 2020. Effects of land use on the characteristics and composition of fluvial chromophoric dissolved organic matter (CDOM) in the yiluo river watershed, China. Ecol. Indic. 114, 106332.

Schmidt, F., Elvert, M., Koch, B.P., Witt, M., Hinrichs, K.U., 2009. Molecular characterization of dissolved organic matter in pore water of continental shelf sediments. Geochim. Cosmochim. Acta 73 (11), 3337−3358.

Shi, Y., Zhang, L.Q., Li, Y.P., Zhou, L., Zhou, Y.Q., Zhang, Y.L., Huang, C.C., Li, H.P., Zhu, G.W., 2020. Influence of land use and rainfall on the optical properties of dissolved organic matter in a key drinking water reservoir in China. Sci. Total Environ. 699, 134301.

Siddhartha Mitra, Christopher L. Osburn, Andrew S. Wozniak, 2017, A Preliminary Assessment of

Fossil Fuel and Terrigenous Influences to Rainwater Organic Matter in Summertime in the Northern Gulf of Mexico, Aquat Geochem.

Sobek, S., Tranvik, L.J., Prairie, Y. T., Kortelainen, P., Cole, J.J., 2007. Patterns and regulation of dissolved organic carbon: an analysis of 7500 widely distributed lakes. Limnol. Oceanogr. 52, 1208-1219.

Song, K., Shang, Y.X., Wen, Z.D., Jacinthe, P.A., Liu, G., Lyu, L.L., 2019. Characterization of CDOM in saline and freshwater lakes across china using spectroscopic analysis. Water Res. 150(3), 403-417.

Song, K., Wen, Z., Shang, Y., 2018. Quantification of dissolved organic carbon (DOC) storage in lakes and reservoirs of mainland China. Journal of Environmental Management, 217, 391-402.

Stackpoole, S.M., Stets, E.G., Striegl, R.G., 2014. The impact of climate and reservoirs on longitudinal riverine carbon fluxes from two major watersheds in the central and intermontane west. J. Geophys. Res. Biogeosciences 119 (5), 848–863.

Stedmon, C.A., Markager, S., 2005. Resolving the variability in DOM fluorescence in a temperate estuary and its catchment using PARAFAC, Limnol. Oceanogr., 50, 686-697.

Stedmon, C.A., Bro, R., 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnology and Oceanography-Methods 6, 572-579.

Stedmon, C.A., Seredyńska-Sobecka, B., Boe-Hansen, R., Le Tallec, N., Waul, C.K., Arvin, E., 2011. 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), 6030-6038.

Stedmon, C.A., Thomas, D.N., Papadimitriou, S., Granskog, M.A., Dieckmann, G.S., 2011. Using fluorescence to characterize dissolved organic matter in Antarctic sea ice brines. Journal of

Geophysical Research 116(G3).

Stubbins, A., Dittmar, T., 2015. Illuminating the deep: molecular signatures of photo- chemical alteration of dissolved organic matter from North Atlantic Deep Water. Mar. Chem. 177, 318-324.

Ter Braak CJF, Šmilauer P. 2002. CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, NY, USA.

Tzortziou, M., Zeri, C., Dimitriou, E., Ding, Y., Jaffé, R., Anagnostou, E., Pitta, E., Mentzafou, A., 2015. Colored dissolved organic matter dynamics and anthropogenic influences in a major transboundary river and its coastal wetland: CDOM dynamics in a transboundary river. Limnol. Oceanogr. 60 (4), 1222-1240.

Walker, S.A., Amon, R. Stedmon, C.A., 2013. Variations in high latitude riverine fluorescent dissolved organic matter: a comparison of large arctic rivers. J. Geophys. Res.-Biogeo.118(4).

Wallin, M. B. et al. Evasion of CO_2 from streams—the dominant component of the carbon export through the aquatic conduit in a boreal landscape. Glob. Change Biol. 19, 785-797.

Wang Z., Cao J., Meng F., 2015. Interactions between protein-like and humic-like components in dissolved organic matter revealed by fluorescence quenching. Water Res. 68(1), 404-413.

Wang, K., Pang, Y., He, C., Li, P., Xiao, S., Sun, Y., Pan, Q., Zhang, Y., Shi, Q., He, D., 2019. Optical and molecular signatures of dissolved organic matter in Xiangxi Bay and mainstream of Three Gorges Reservoir, China: Spatial variations and environmental implications. Sci. Total Environ. 657, 1274-1284.

Wang, K., Li, P., He, C., Shi, Q., He, D., 2021a. Hydrologic heterogeneity induced variability of dissolved organic matter chemistry among tributaries of the Three Gorges Reservoir, Water Research,

201, 117358.

Wang, K., Li, P., He, C., Shi, Q., He, D., 2021b. Density currents affect the vertical evolution of dissolved organic matter chemistry in a large tributary of the Three Gorges Reservoir during the water-level rising period, Water Res. 204, 117609.

Wang, K., Pang, Y., Gao, C., Chen, L., Jiang, X., Li, P., He, C., Shi, Q., He, D., 2021c. Hydrological management affected dissolved organic matter chemistry and organic carbon burial in the Three Gorges Reservoir, Water Res. 199, 117195.

Williams, C.J., Frost, P.C., Morales-Williams, A.M., Larson, J.H., Richardson, W.B., Chiandet, A.S., Xenopoulos, M.A., 2016. Human activities cause distinct dissolved organic matter composition across freshwater ecosystems. Glob. Chang. Biol. 22 (2), 613-626.

Wilson, H.F., Xenopoulos, M.A., 2009. Effects of agricultural land use on the composition of fluvial dissolved organic matter. Nat. Geosci. 2, 37-41.

Xia, B., Zhang, L., 2011. Carbon distribution and fluxes of 16 rivers discharging into the Bohai Sea in summer. Acta Oceanol. Sin. 30, 43-54.

Xiao M., Wu F.C., Zhang R., Wang L., Li X., Huang R., 2011. Temporal and spatial variations of low-molecular-weight organic acids in Dianchi Lake, China. J. Environ. Sci. 23:1249-1256

Yates, C.A., Johnes, P.J., Owen, A.T., 2019. Variation in dissolved organic matter (DOM) stoichiometry in UK freshwaters: assessing the influence of land cover and soil C: N ratio on DOM composition. Limnol. Oceanogr. 64, 2328-2340.

Yu, Q., Wang, F., Yan, W., Zhang, F., Lv, S., Li, Y., 2018. Carbon and nitrogen burial and response to climate change and anthropogenic disturbance in Chaohu Lake, China, Int. J. Env. Res. Pub. He. 15(12), 2734-2752.

Zhang, F., Yao, S., Xue, B., Lu, X., Gui, Z., 2017. Organic carbon burial in Chinese lakes over the past 150 years. Quaternary International. 438, 94-103.

Zhang, F., Harir, M., Moritz, F., Zhang, J., Witting, M., Wu, Y., Schmitt-Kopplin, P., Fekete, A., Gaspar, A., Hertkorn, N., 2014. Molecular and structural characterization of dissolved organic matter during and post cyanobacterial bloom in Taihu by combination of NMR spectroscopy and FTICR mass spectrometry. Water Res. 57, 280-294.

Zhang, Y.L., Zhang, E.L., Yan, Y., Dijk, M.A.V., Feng, L.Q., Shi, Z.Q., Liu, M.L., Qin, B.Q., 2010. Characteristics and sources of chromophoric dissolved organic matter in lakes of the Yungui Plateau, China, differing in trophic state and altitude. Limnol. Oceanogr. 55 (6), 2645–2659.

Zhou Y.P., 2020. Spatial changes in molecular composition of dissolved organic matter in the Yangtze River Estuary: Implications for the seaward transport of estuarine DOM. 759,143531-143540.

Zhou, Y., Jeppesen, E., Zhang, Y., Cheng, N., Shi, K., Liu, X., Zhu, G., Qin, B., 2015. Chromophoric dissolved organic matter of black waters in a highly eutrophic Chinese lake: freshly produced from algal scums. J. Hazard. Mater. 299 (6), 222-230.

Zhou, Y., Shi, K., Zhang, Y., 2017. Fluorescence peak integration ratio IC: IT as a new potential indicator tracing the compositional changes in chromophoric dissolved organic matter. Sci. Total Environ. 574, 1588-1598.

Zhou, Y., Xiao, Q., Yao, X., Zhang, Y., Zhang, M., Shi, K., Lee, X., Podgorski, D.C., Qin, B., Spencer, R.G.M., Jeppesen, E., 2018. Accumulation of terrestrial dissolved organic matter potentially enhances dissolved methane levels in eutrophic Lake Taihu, China. Environ. Sci. Technol. 52 (18), 10297-10306.

Table 1 The comparison of averaged absorption and fluorescent property indices for humic-rich and protein-rich lakes

Types	M	BIX$_{310}$	HIX
Humic-rich lakes	7.612	0.683	10.756
Protein-rich lakes	12.617	1.460	4.117

Table 2 The comparison of averaged absorption and fluorescent property indices for different regions

Regions/Types	HIX	M	BIX$_{310}$
NLR	**14.701**	**6.314**	**0.652**
Humic-rich	17.649	5.985	0.548
Protein-rich	13.515	10.397	0.767
ELR	**8.653**	**7.260**	**0.776**
Humic-rich	10.105	5.693	0.509
Protein-rich	4.395	9.349	1.227
MXR	**9.081**	**9.544**	**0.907**
Humic-rich	11.513	7.145	0.806
Protein-rich	4.994	11.409	1.266
TQR	**4.947**	**16.695**	**1.263**
Humic-rich	7.006	11.595	0.922
Protein-rich	3.723	18.155	1.307
YGR	**9.146**	**8.999**	**0.726**
Humic-rich	11.334	7.580	0.625
Protein-rich	5.612	13.288	0.833

Table 3 The correlation between fluorescent components and CDOM property indices among different types of lakes

Indices	Humic-rich lakes			Protein-rich lakes		
	C1	C2	C3	C1	C2	C3
a$_{CDOM}$(254)	.486**	.373**	.220**	.149**	.270**	.441**
DOC	.484**	.298**	.255**	.278**	.250**	.499**
M	-.197**	0.034	0.013	0.000	0.025	.315**
C1	1	.762**	.350**	1	.313**	.585**
C2	.762**	1	.546**	.313**	1	.697**
C3	.350**	.546**	1	.585**	.697**	1
BIX$_{310}$	-.422**	-.218**	.208**	-.469**	-.257**	.693**
HIX	.649**	.236**	-.223**	.420**	-.208**	-.587**

** denote statistically significant level at 0.01

Table 4 The literature estimates of DOC storage, annual CO_2 equivalents, OCAR, and carbon burial capacity of different regions of lakes across China

Types	NLR	ELR	MXR	YGR	TQR	China	Reference
DOC storage (Tg C)	0.19	0.43	1.75	0.13	13.39	15.89	Song et al., 2018
Annual CO_2 equivalents (Tg C)	13.56	18.11	16.96	0.37	79.36	128.36	Li et al., 2018
Organic carbon accumulation rate $(g/m^2\ y)$	25.40	30.60	30.40	24.30	14.30	22.00	Zhang et al., 2017
Carbon burial capacity (Tg C)	17.70	98.80	59.40	4.50	89.10	269.50	Zhang et al., 2017
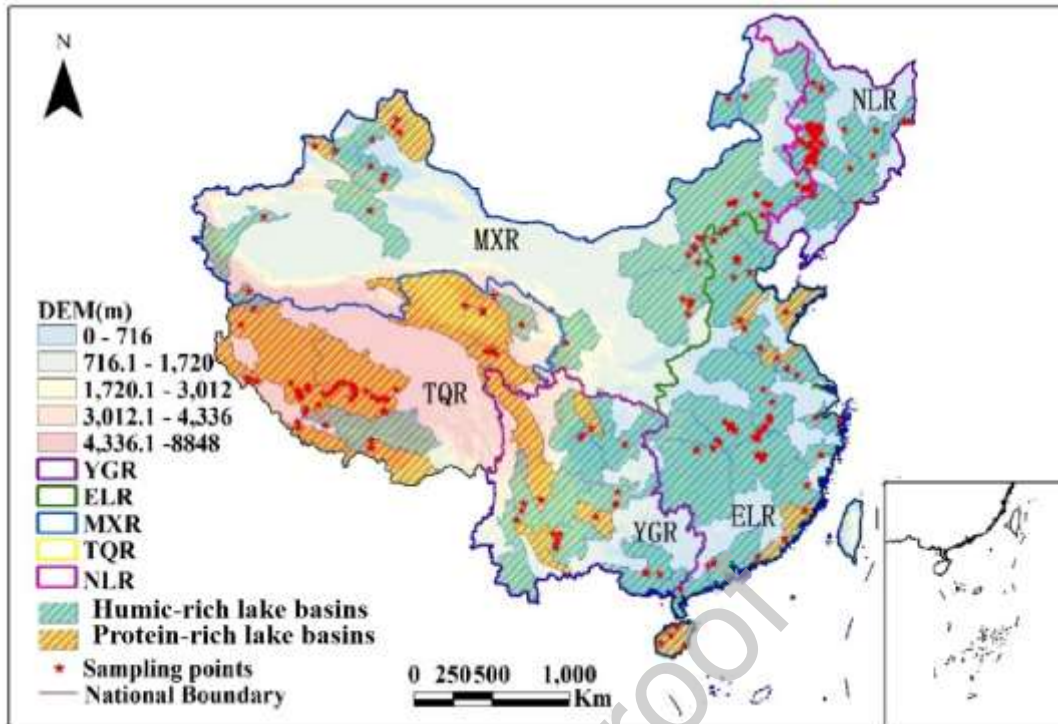
Fig.1 The distribution of humic-rich lakes and protein-rich lakes in different regions across China.
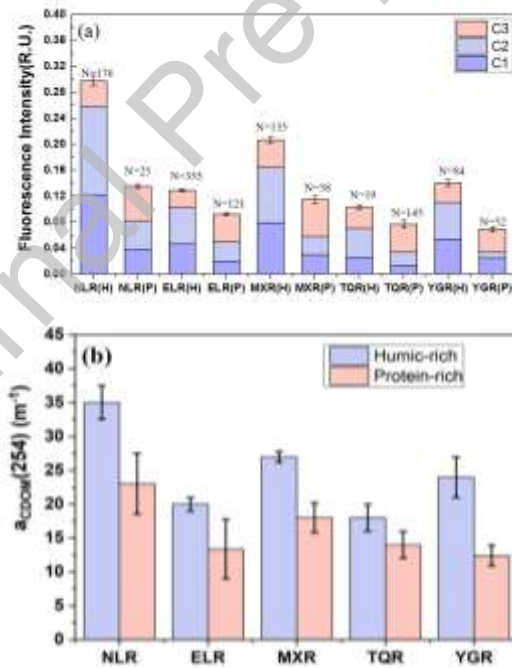
Fig.2 The variation of FDOM (a) and $a_{CDOM}(254)$ (b) for humic-rich and protein-rich lakes across China.
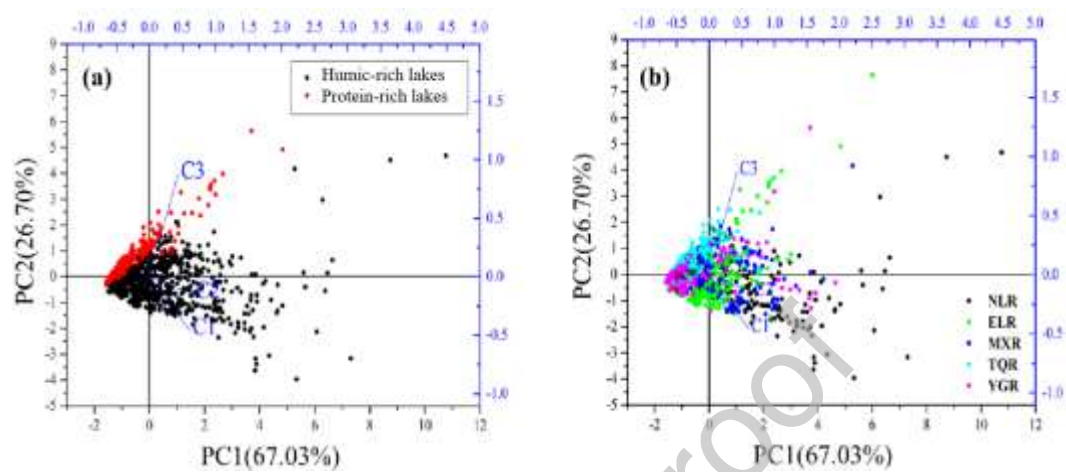
Fig.3 The principal component analysis (PCA) results of PARAFAC-based EEMs (a) for different water types and (b) for different regions.
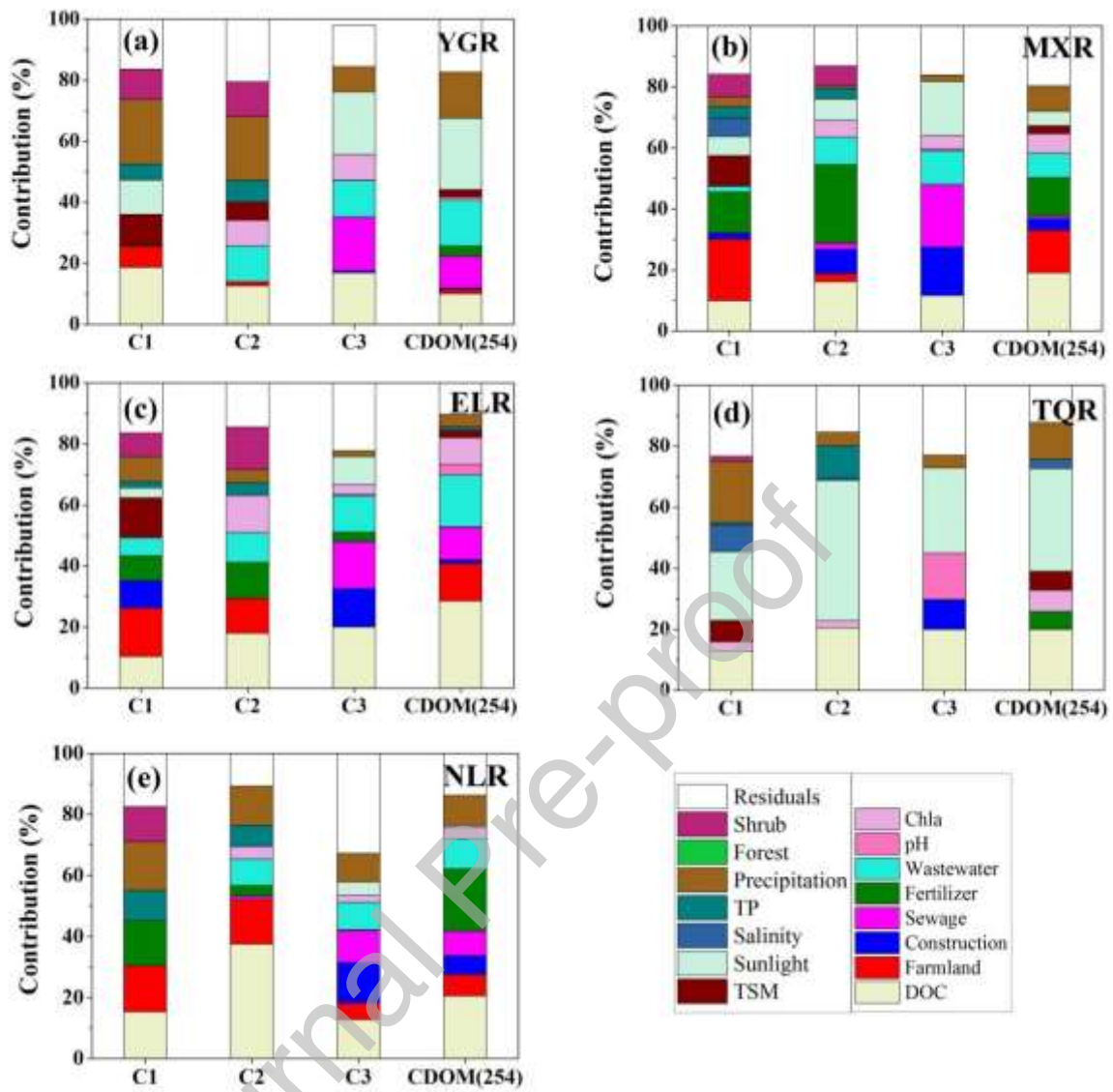
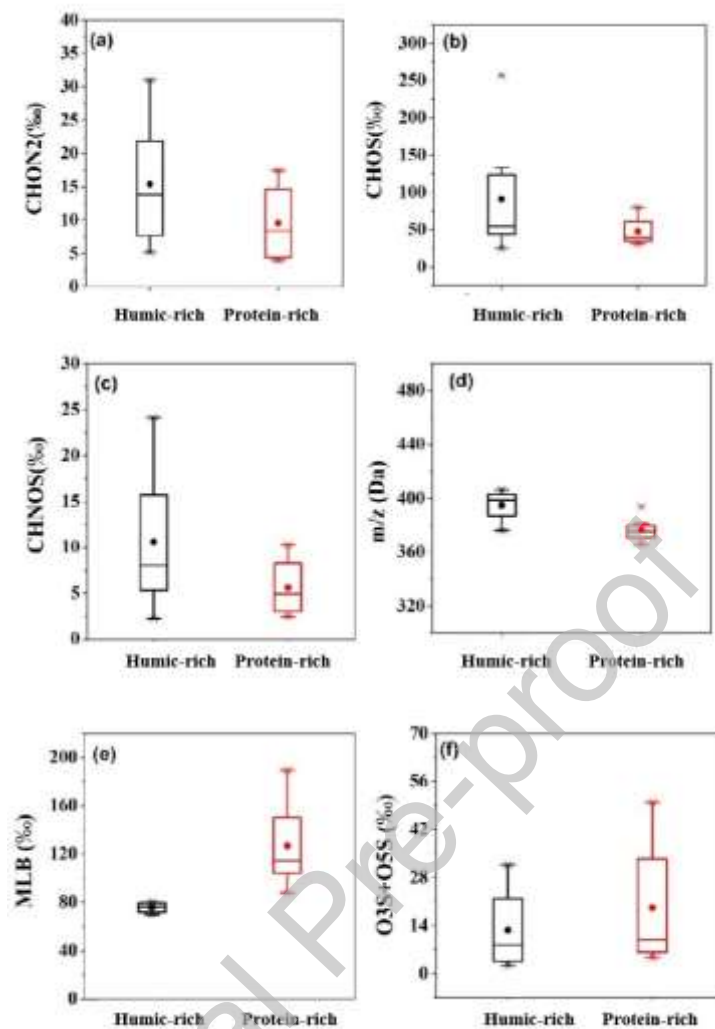Fig.4 The controlling factors of CDOM components for lakes in regions across China

Fig. 5 The property variations of DOM in lakes. (a) CHON2(‰), (b) CHOS(‰), (c) CHNOS(‰) , (d)

m/z (Da) , (e) MLB(‰), (f) O3S+O5S(‰) (p<0.05; The square represents the mean value, and the

filled circles represent the mean. The horizontal edges of boxes denote the 25th and 75th percentiles,
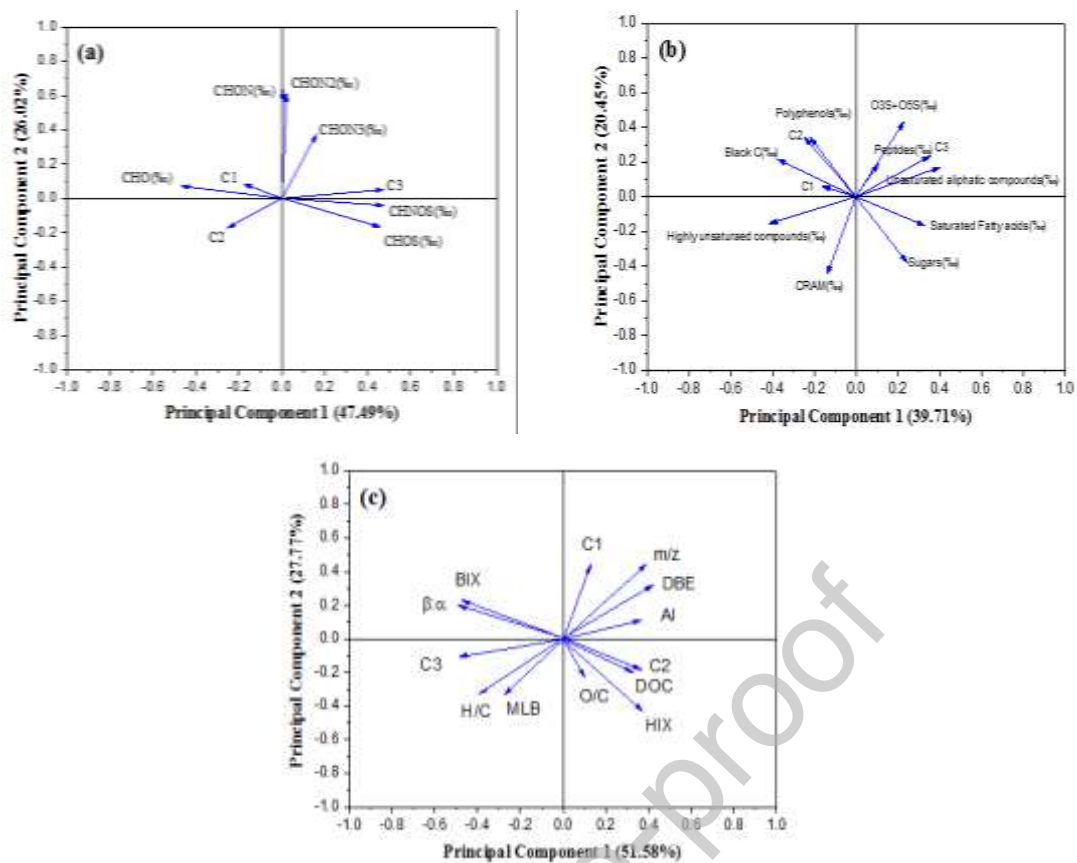
and the cross represent outliers.)

Fig.6 The PCA results of molecular complexity and optical parameters (a) for intensity-weighted parameter ‰ (b) for molecular formulae groups ‰ (c) molecular indices
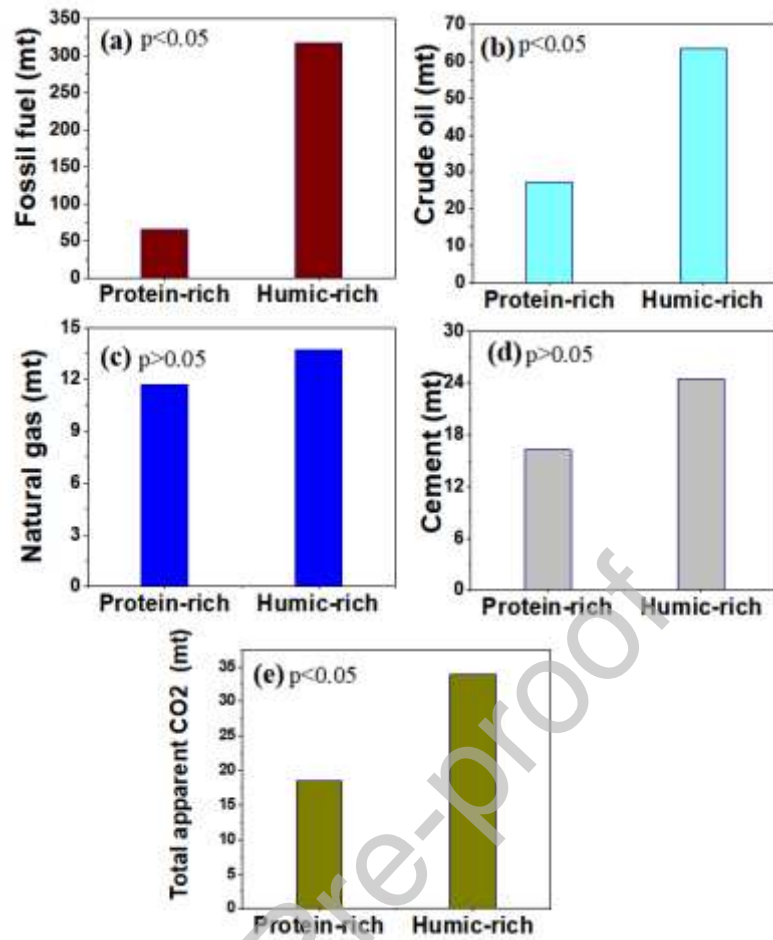
Fig.7 The comparison of average annual carbon emissions (2015 to 2019) between the administrative regions of humic-rich lakes and protein-rich lakes. Note: mt denotes million tons.
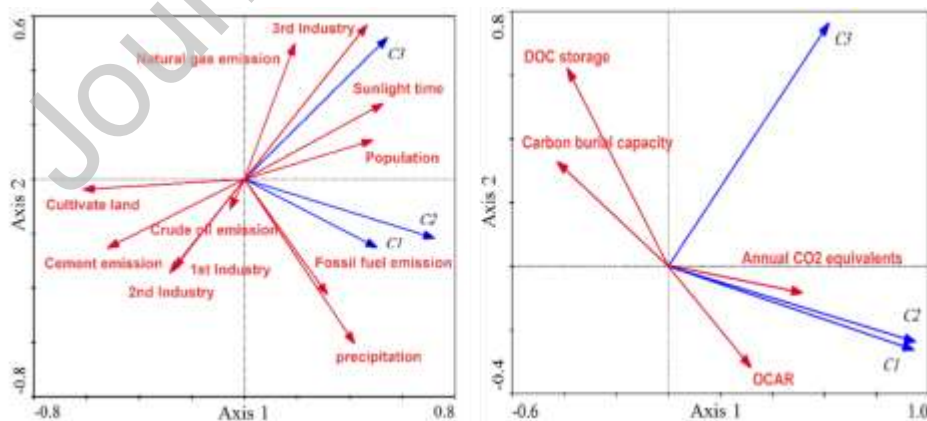

Fig. 8 The analysis of RDA for CDOM fluorescent components with a) natural and anthropogenic factors; b) aquatic carbon cycles.