Original Research

Spectral Characteristic Changes of Dissolved Organic Matter in Aquatic Systems Under the Influences of Agriculture and Coal Mining

Xudong Chen¹, Liangmin Gao^{1*}, Xiaoqing Chen¹, Zhendong Pang², Juan Ge¹, Haiqiang Zhang¹, Zhen Zhang¹, Ming Mu¹, Yuhui Qiu¹, Xinglan Zhao¹

¹School of Earth and Environment, Anhui University of Science and Technology, Huainan, China ²Environmental Protection Monitoring Station of Huainan City, Huainan, China

> Received: 2 September 2021 Accepted: 7 December 2021

Abstract

Under the combined influence of various factors, the dissolved organic matter (DOM) in the drainage basin presents complex variation characteristics. In this study, we investigated the spectral characteristics of DOM in the middle reach of the Huai River aquatic systems under the influence of complex conditions (i.e., agricultural nonpoint pollution, coal mining activities, photovoltaics on the water surface, and rural sewage). In addition to the general water quality indices, ultraviolet-visible (UV-Vis) absorption spectroscopy, the excitation-emission matrix spectroscopy (EEMs) combined with the parallel factor analysis (PARAFAC), the fluorescence regional integration (FRI), and the fluorescence indices were used to describe the spectral characteristics of DOM. The results show that agricultural non-point source input from agricultural activities deteriorates water quality, but the coal mining activities weaken the agricultural non-point pollution intensity (i.e., occupy farmland, and subsidence area collected agricultural runoff). The PARAFAC model was composed of the following three chemical components: fulvic-like (C1), humic-like (C2), and tryptophan-like (C3). The results of PRAFAC and FRI indicated that the photovoltaic industry significantly influenced the properties of DOM in the aquatic systems by slowing the photochemical reactions. The principal component analysis (PCA) and Adonis test showed that under different influencing conditions, the aquatic systems within the four study areas showed significant differences. Spectral characteristics of DOM can sensitively indicate water distinction. It is of practical significance to monitor the change characteristics of DOM in a minor watershed under complex influencing factors by spectroscopy.

Keywords: dissolved organic matter, excitation-emission matrix spectroscopy, parallel factor analysis, fluorescence regional integration, principal component analysis

Introduction

Under the comprehensive influence of various factors, the water quality in the basin presents complex variation characteristics, and it is difficult to effectively monitor and forecast by a single method [1]. When point source pollution emissions are effectively controlled, nonpoint source pollution from agriculture is an important factor contributing to pollution in the watershed environment [2-3]. The agricultural soil erosion and runoff caused by rainfall make the organic matter in the soil enter the water body of the basin, which has a nonnegligible impact on the aquatic systems [4]. Since ancient times, the middle reach of the Huai River in China has served as an area of intensive agricultural production activities which is a major source of nonpoint source pollutants [5]. In addition, a large proportion of rural villages within these watersheds still lack the capacity for efficient sewage collection and treatment facilities, leading to direct discharges of sewage into the receiving rivers [6]. In recent decades, several problems caused by coal mining activities have also emerged such as large areas of subsidence formed as a result of coal mining [7]. Mining activity is usually associated with heavy industry, the degradation of space and areas with a clearly industrial character. In fact, mining can be associated with other functions [8]. In order to use these collapsed lands and create economic value, solar photovoltaic panels were paved on the water surface of the collapsed pond formed by the coal mining activities. Due to these complex surrounding factors, the local aquatic system was affected differently, showing diverse characteristics.

The dissolved organic matter (DOM) is considered the one of the largest pools of organic matter in natural waters, accounting for more than 90% of the total organic matter [9], the regulated metabolic and biogeochemical processes in the water bodies ultimately determine the aquatic ecosystems' contribution to the global carbon cycle [10]. The characteristics of DOM are difficult to elucidate because of its complex chemical structures (i.e., carboxyl, hydroxyl, amino, glycosyl phenol, quinonyl, aldehyde, ester, ketone, and other functional groups) [11]. The quantity and composition of the temporal variation and spatial distribution of DOM is a good indicator of the ecological condition of the water [10, 12]. For a better health assessment and management of aquatic ecosystems under the influence of complex conditions, it is critical to understand the characteristics of DOM [10]. Ultraviolet-visible (UV-Vis) absorption spectroscopy is used to characterize the composition and structural characteristics of DOM [13]. The excitation-emission matrix spectroscopy (EEMs) with high sensitivity has been widely used to detect the chemical composition and biogeochemical cycle of DOM [14, 15]. The parallel factor analysis (PARAFAC) decomposed the fluorescence of DOM into several chemically meaningful components based on the mathematical trilinear model. PARAFAC has

become the most popular algorithm to analyze EEMs data [16, 17]. Fluorescence regional integration (FRI) disintegrates the EEMs of DOM into specific regional volumes and composes the volume beneath each region. FRI realizes the semi-quantitative analysis of the configuration and heterogeneity of DOM [18]. Agricultural nonpoint pollution has significant impacts on dissolved organic matter in aquatic systems, as indicated from previous research. However, there are few studies about the characteristics of DOM on a more complex environmental situation. In this study, we selected watersheds that were influenced typically by agriculture nonpoint pollution and coal mining activities. The UV-Vis, EEMs-PARAFAC, and EEMs-FRI methods were used to trace the characteristics and the transformation of DOM in the aquatic systems within the studied areas.

Experimental

Watershed and Sample Collection

Watershed Information

As shown in Fig. 1, The study area is in the Panji District of the Huainan City in the Anhui Province of China, which is located at 32°35'-32°50'N, 116°45'-117°5'E. The Ni River is the natural system of surface water around the Panji mining area; its primary functions are farmland irrigation, aquaculture, and drainage. The Ni River passes through the Panji coal mining areas (i.e., the Panji No.1 coal mine, the Panji No.2 coal mine, and the East Panji coal mine). There are many collapses around the Ni River due to the longterm coal mining activities. The Ni River was divided into upstream (NRU) and downstream (NRD). There is a large area of a lake formed by coal mining subsidence in the NRU. In recent years, the regional government has introduced water photovoltaic power generation projects, utilizing a large number of solar panels to cover the water surface. The most of water surfaces of the NRD are covered by photovoltaic panels. The Hei River and the Ni River converge and flow into the mainstream of the Huai River. The large agricultural nonpoint source emissions (e.g., fisheries, poultry, livestock, and farmland) exist in the Hei River Basin. As shown in Fig. 1, the NRD, NRU, HER, and Huai River (HR) were selected as the study areas. The Huai River Basin has been a concentrated area of agricultural production, and these study areas are affected by both agricultural nonpoint sources and coal mining activities.

Sample Collection

In January of 2021, the sampling process was carried out, and during this time, there was no obvious precipitation process; the river water was less disturbed, and the water quality was stable. The climate before



Fig. 1. Location and sampling sites in the middle reach of the Huai River.

sampling has a major impact on agricultural non-point pollution intensity and the characteristics of DOM on aquatic systems because of the overland runoff formed by precipitation [19]. The sampling process lasted for a week, and twenty-seven water samples were collected in the four research areas. Water samples were collected using an acrylic glass water collector at 0.5 m depth, and 1 L of water samples were collected at each sampling site. All samples were stored at 4°C and protected from the light. The test of all the indexes was completed within three days after sample collection. When collecting the samples, we utilized the GPS (GPSMAP 621sc, Garmin, USA) to record the coordinates of the sampling points. The distribution and geographical location of sampling points were shown in Fig. 1.

Analytical Methods

General Water Quality Indices

The temperature (T) and total dissolved solids (TDS) were measured by a portable conductivity meter (EC300, YSI, USA), and a pH meter (PhSJ-4F, INESA, China) was used to measure the pH immediately during sample collection in the field. The collected samples were filtered by 0.45µm glass fiber filters which had been pretreated at 400°C before any data analysis was performed. Dissolved organic carbon (DOC) was determined by a total organic carbon analyzer (TOC-VCPH, Shimadzu, Japan). The other water quality indices included the following ammonium (NH⁺, using Nessler's reagent spectrophotometry), total nitrogen (TN, using the alkaline potassium persulfate digestion UV spectrophotometric method), total phosphorous (TP, using the ammonium molybdate spectrophotometric method), nitrate-nitrogen (NO,-, using ultraviolet spectrophotometry), and chemical

oxygen demand (COD, using the dichromate method) were determined according to the methods proposed by the State Environmental Protection Administration of China [20].

UV-Vis Absorption Spectrum

The temperature of the sample was kept at 25°C before performing the DOM spectral measurement. The UV-Vis absorption spectrums were collected from 200-800 nm in 1 nm intervals using a UV-Vis absorption spectrophotometer (N5000PLUS, YOKE, China) with a 1 cm quartz cell. In addition, we utilized the Milli-Q water as a blank correction. The specific UV absorbance was at 254 nm (SUVA₂₅₄), which is a significant, positive correlation with the aromaticity of DOM; this result is defined as the UV absorption at 254 nm, which was normalized for the DOC concentration, and recorded as L mg⁻¹ m⁻¹ [21]. The slope ratio (S_R) was defined as $S_{275-295}/S_{350-400}$ (i.e., the ratio of the spectral slope at the range of 275-295 nm and 350-400 nm) [22]. $S_{\rm \scriptscriptstyle R}$ and E2/E3 (i.e., the absorption ratio between 254 nm and 365 nm) are opposite with the molecular size of DOM [23, 24]. E3/E4 (i.e., the absorption ratio between 300 nm and 400 nm) indicates the source of humus in water. When E3/E4 is greater than 3.5, the humus substance mainly consists of a fulvic-like composition; otherwise, it demonstrates humic-like qualities [25].

Fluorescence Spectrum

Fluorescence excitation-emission matrices spectrum (EEMs) of all samples were detected by a fluorescence spectrometer (F4600, Hitachi, Japan). The machine utilizes a 150 W ozone-free xenon arc lamp as artificial light, the length of the absorption cell is 1 cm, and the material is made of quartz. The slit width of the excitation and emission was set at 5 nm. The excitation range of the EEMs was 200-500 nm in 5 nm increments and the emission range was 200-550 nm in 5 nm increments. The integration time was set to auto, and the scan speed was set at 1200 nm/min. Additionally, we utilized the Milli-Q water as the blank. However, before the PARAFAC modeling and FRI were performed, it is necessary to calibrate the system, using the package called spectroscopic analysis of DOM ("staRdom", 1.1.21) in R (4.0.5), to correct the EEMs and to acquire the precise results for our study [26]. A 7-step correction procedure was applied for EEMs correction: (1) absorbance baseline correction, (2) spectral correction, (3) blank subtraction, (4) inner-filter effect correction, (5) Raman normalization (producing corrected fluorescence in Raman Units), (6) remove and interpolate scattering, and (7) correction for the sample dilution [26].

PARAFAC Modeling and FRI

PARAFAC Modeling

The PARAFAC model was run and validated following by the execution of the package "staRdom" (1.1.21) in R (4.0.5). The EEMs of all the samples involved in the modeling were corrected perfectly. The signal at excitation wavelength below 250 nm was removed from all EEMs because the high levels of noise are negative for PARAFAC modeling. For the PARAFAC modeling, the non-negativity constraint and normalization were set, because the negative concentrations and fluorescence intensities are impossible to determine, and the components are highly correlated [27]. After multiple parameter selection and model establishment processes, the three-component model was determined, ultimately. The three-component model explained 99.5% of the variance within all EEMs, and all the Tucker congruence coefficients of the split-half analysis were greater than 0.99, as shown in Fig. 2. The core consistency of the model is 96.45 [28]. The four to seven component models do not reach the threshold set for model validation and therefore are not considered. Three components' excitation and emission loadings were searched by an online spectral library, the OpenFluor website (https://openfluor.lablicate.com/). Both emission and excitation spectral matches were restricted to the Tucker congruence coefficients of greater than 0.95 [29].

FRI Calculating Process

Through the FRI method, the whole EEMs dataset can be decomposed into five specific regions, and the volume of each region enables the semi-quantitative analysis of the configuration and heterogeneity of DOM. As Chen's report indicates [18], the region in which the excitation wavelengths are shorter than 250 nm and the emission wavelengths are shorter than 330 nm is related to the aromatic proteins I (tyrosine), as shown in region I of Fig. 3. The region at short excitation wavelengths (<250 nm) with emission wavelength ranges of 330-380 nm and greater than 380 nm is related to the aromatic proteins II (tryptophan) and the fulvic acid-like materials (Fig. 3, regions II and III). The region located at the long excitation of more than 250 nm with an emission range from 250-380 nm is assigned to the protein-like and soluble microbial product compounds (Fig. 3, region IV). When region V is at an excitation and emission of >250/>380 nm, this region represents the humic-like acid component (Fig. 3, region V). After calculating the volume of each region, the integral volume was corrected to the standard volume by using the area integral, and the percentage fluorescence response (P_{in}) of each region was calculated [30-32]. Fig. 3 shown the FRI results of the selected Specific samples from the four areas of the study.

Fluorescence Indices

Fluorescence index (FI), biological index (BIX), and humification index (HIX) are typically used to describe the fluorescence properties of aquatic systems. FI is the ratio of the emission wavelength at 450 nm and 500 nm with an excitation of 370 nm. FI is used to distinguish the source of humus with autochthonous (microbial,



Fig. 2. The split-half validation results for each identified component of the PARAFAC model.

FI>1.9) and allochthonous (terrigenous, FI<1.4) [33]. BIX is described as the ratio of the fluorescence intensities of the emission wavelength at 380 nm and at 500 nm when the excitation wavelength is at 310 nm. The high BIX indicates that DOM in the aquatic systems contains more autochthonous components [34]. HIX equals the fluorescence intensity integration at the emission wavelength of 435-480 nm divided by the sum of the fluorescence intensity integration at 300-445 nm and 435-480 nm when the excitation wavelength is at 254 nm. The higher value of HIX implies a higher level of humification of the water [35]. We computed the fluorescence indices after the spectrum correction by using the package "staRdom" (1.1.21) in R (4.0.5).

Statistical Methods

Statistical analysis was performed using R (4.0.5) (P<0.05, P<0.01, and P<0.001 indicated that there were significant differences at different statistical levels, respectively). The one-way analysis of the variance (ANOVA) and the Least Significance Difference (LSD) post hoc test were obtained via the software package of "agricolae" (1.3-5) in R (4.0.5). Additionally, we utilized the software package of "factoextra" (1.0.7) and "FactoMineR" (2.4) in R (4.0.5) to complete the principal component analysis (PCA). All data were presented as the mean \pm the standard deviation (SD).

Results and Discussion

General Water Quality Indices Analysis

Table 1 shows the statistical analysis results of the water quality indices (i.e., T, TDS, pH, DOC, TN, NH_4^+ , NO_3^- , TP, and COD) of the four aquatic areas (i.e., HR, NRU, NRD, and HER). The concentration of TN in HR (2.36±0.89 mg/L) is higher than in other areas. Likewise, HR owns the highest concentration of NO_{2}^{-} (1.59±0.64 mg/L) and is significantly more than the other areas (P<0.001, LSD). But the concentration of NH_4^+ in HER (0.86±0.49 mg/L) is the highest. The above result shows that the TN content in HR is high, and NO_3^{-1} is the main contribution index, which is related to a large number of agricultural nonpoint sources in the study area. Rainstorms can flush large amounts of agricultural nonpoint source pollution from land sources into water bodies, threatening water supply [36]. The NO² migrates to HR through the surface runoff and groundwater. Due to the existence of a large number of fisheries and livestock in the HER area, the content of NH⁺ in HER increased. Peng's report shows that anthropogenic activities (i.e., agricultural cultivation and sewage emissions) contributed to nitrate pollution of aquatic systems [37]. Meanwhile, the concentration of TP in HER (0.54±0.76 mg/L) is higher than in other areas but is not as significant (P>0.05, LSD). As shown in Table 1, a similar distribution



Fig. 3. The FRI regions division of EEMs dataset from the four studied areas.

Area	HR	NRU	NRD	HER	ANOVA
Parameters	Mean±SD	Mean±SD	Mean±SD	Mean±SD	
T (°C)	4.44±0.65 b	5.28±0.48 a	4.58±0.69 ab	4.56±0.57 ab	NA
TDS (mg/L)	0.35±0.06 ab	0.38±0.06 ab	0.28±0.05 b	0.42±0.10 a	*
pH	8.08±0.09 a	8.28±0.31 a	7.51±0.17 b	7.67±0.17 b	***
DOC (mg/L)	11.04±4.30 b	5.41±1.56 b	7.61±1.28 b	16.57±7.89 a	**
TN (mg/L)	2.36±0.89 a	1.53±0.53 a	1.76±0.67 a	1.73±0.51 a	NA
NH_4^+ (mg/L)	0.59±0.27 ab	0.32±0.04 b	0.63±0.52 ab	0.86±0.49 a	NA
NO ₃ ⁻ (mg/L)	1.59±0.64 a	0.80±0.46 b	0.93±0.28 b	0.54±0.25 b	***
TP (mg/L)	0.17±0.07 a	0.09±0.03 a	0.39±0.25 a	0.54±0.76 a	NA
COD (mg/L)	21.23±9.02 b	12.41±3.15 b	18.42±8.60 b	46.66±23.45 a	**

Table 1. Statistical analyses of the water quality index.

Note: The bold lowercase letters indicate significant differences (P<0.05, LSD) among the study areas; The results of ANOVA represent as "NA", "*", "**", and "***" (P>0.05, P<0.05, P<0.01, and P<0.001).

of DOC and COD is revealed. The contents of DOC ($16.57\pm7.89 \text{ mg/L}$) and COD ($46.66\pm23.45 \text{ mg/L}$) in HER are significantly the highest among the four areas (P<0.01, LSD). There are lots of villages around HER with no sewage treatment facilities, and the sewage directly discharged into the water caused the HER to own the highest value of DOC and COD. The intensive coal mining activities around the Ni River weakened the input of agricultural nonpoint sources. At the same time, Long-term coal mining caused a large area of land subsidence, gathering the surrounding surface runoff

and weakening the nonpoint input process to the NRU. The value of COD and DOC in NRU and NRD is lower than in other areas as shown in Table 1.

Characteristics of UV-Vis Indices

Fig. 4 (a-d) presents the distribution of UV-Vis indices (i.e., SUVA₂₅₄, S_R, E2/E3, and E3/E4) in HR, NRU, NRD, and HER. As shown in Fig. 4a), the SUVA₂₅₄ value of samples in NRU ($5.08\pm1.56 \text{ L}\cdot\text{mg}^{-1}\cdot\text{m}^{-1}$) and NRD ($4.19\pm1.20 \text{ L}\cdot\text{mg}^{-1}\cdot\text{m}^{-1}$) is significantly



Fig. 4. The distribution of UV-Vis spectral indices of four study areas. The different lowercase letters indicate significant differences (P<0.05, LSD). The results of ANOVA represent as "NA", "*", "**", and "***" (P>0.05, P<0.05, P<0.01, and P<0.001).

greater than HER $(2.02\pm1.03 \text{ L}\cdot\text{mg}^{-1}\cdot\text{m}^{-1})$ and HR $(1.73\pm0.82 \text{ L}\cdot\text{mg}^{-1}\cdot\text{m}^{-1})$ (P<0.001, LSD). The humification degree of the aquatic systems in NRD and NRU is more significant than HER and HR. The high value of the humification in the water means that more conjugated benzene ring structures, hydrophobic components, and quinone groups, and the $\pi \to \pi$ * transition of unsaturated organic molecules is more active in DOM [38]. The distribution of S_{p} and E2/E3 among the study areas indicates no significant differences (P>0.05, ANOVA). The rank of SR is NRU (1.01±0.03)>NRD (0.99±0.09)>HER $(0.94 \pm 0.06) > HR$ $(0.93 \pm 0.07),$ and the rank of E2/E3 is NRD (7.01±1.11)>NRU (6.47±0.24)>HER (6.45±1.14)>HR (6.34±1.10). The result indicates that the average molecular size of the four areas is not significantly different (Fig. 4 (b-c)). The average value of E3/E4 in NRD (8.37±1.30) is significantly higher than HER (5.74±1.62), HR (5.50±2.16), and NRU (4.94±0.50) (P<0.05, LSD). The mean of E3/E4 in all regions is greater than 3.5, which shows that the humus substance in the aquatic system mainly has a microbial fulvic-like composition. Water leaching is the primary mode of transport for non-point pollutants from soils, sediments, and other sources to adjacent water bodies [39]. All the samples were collected in the dry season with no precipitation process, which causes a decrease in the input of the aquatic systems' allochthonous materials with high molecular weight. Previous studies showed greater aromaticity and molecular weight of DOM in allochthonous systems than in most aquatic ecosystems [40]. As Shang's reports, strong soil-stream hydrological connectivity can increase stream water DOM concentration and the proportion of humic-like substances [41].

PARAFAC and FRI Results

PARAFAC Components

The EEMs dataset of all samples acquires three components by the PARAFAC modeling in R (4.0.5). As shown in Fig. 5, the EEMs and spectral loadings



Fig. 5. The fluorescence spectrum of the three PARAFAC components separated from the EEMs datasets of the middle reach of the Huai River in China is noted in a), c), and e). The line plots in the right of each contour plot represents the excitation (red) and emission (black) loadings to the corresponding PARAFAC component, as illustrated in b), d), and f).

of the three components were presented. The peak of component 1 (C1) occurred where the excitation wavelength (λ_{Ex}) = 250 nm and the emission wavelength $(\lambda_{\rm Em}) = 405$ nm (Fig. 5 (a-b)). The peak of component 2 (C2) was observed at $\lambda_{Ex} = 270$ nm with a secondary peak at $\lambda_{Ex} = 365$ nm, the two peaks $\lambda_{Em} = 450$ nm (Fig. 5 (c-d)). The peak of component 3 (C3) was observed at $\lambda_{_{Ex}}=$ 280 nm and $\lambda_{_{Em}}=$ 335 nm (Fig. 5 (e-f)). The spectral loadings were compared with the OpenFluor database [29], and the three components' chemistry information was determined. C1 was classified as containing a fulvic-like substance, and the position of the peak near "Peak A" and "Peak M" [42, 43]. C2 was identified as terrestrial humic-like composition and own the long excitation and emission wavelengths [44, 45]. The two peaks' positions of C2 were at "Peak A" and "Peak C," respectively. C3 was categorized as containing a tryptophan-like substance and near the "Peak T" [46]. C1 and C2 were classified as humus-like substances and have been identified at the bulk of the environmental studies (e.g., forest streams, agricultural runoff, and wetlands) [9]; this is a comprehensive product under the influence of terrestrial input, microbial activities, and photochemical processes [47]. C3 is mainly produced by biological activities in the aquatic systems and can indicate the autochthonous process to a high biological degree. Meanwhile, there were some studies that have pointed out that human sewage also contains tryptophan-like components [48, 49].

PARAFAC and FRI Results Analysis

Fig. 6 (a-b) shows the average of the three PARAFAC component fluorescence loadings in the four study areas (i.e., HR, NRU, NRD, and HER) and the percentage of each component. From Fig. 6a), the fluorescence loadings of the C1 in NRD (1.06 ± 0.56 R.U.) is significantly higher than HER (0.74 ± 0.05 R.U.), NRU (0.53 ± 0.07 R.U.), and HR (0.47 ± 0.01 R.U.) (P<0.01, LSD). Similar

to C1, the NRD owned the highest fluorescence loadings in C2 (P<0.001, LSD). The distribution of fluorescence loadings in C3 is not a significant difference among the four study areas (P>0.05, LSD). The total fluorescence loadings of C1, C2, and C3 were computed as the sum of each component fluorescence loadings. Because of the contribution of the C1 and C2 loadings, the total fluorescence loadings of the NRD (4.28±2.50 R.U.) is significantly far beyond the other areas (i.e., HER (1.89±0.38 R.U.), NRU (1.51±0.61 R.U.), and HR (1.14±0.05 R.U.)) (P<0.001, LSD). The long-term coal mining activities around the Ni River have forced the coal's organic matter via the precipitation process and the atmospheric dust accumulation to continuously enter the aquatic systems. Meanwhile, in recent years, the power photovoltaic generation projects have blocked the sunlight on the water surface, greatly slowing down the photochemical oxidation process of the DOM [50, 51]. In a study on the long-term fate of DOM, Vahatalo et al. found that 96 % of DOM in aquatic systems was destroyed by solar radiation, and the decomposition time in the dark is 70 times more than in the sunlight [52].

As shown in Fig. 6b), the average percentage of each PARAFAC component was 42±2% (fulvic-like, C1), 21±1 % (humic-like, C2), and 38±3% (tryptophanlike, C3) in the HR respectively. It indicated that fulvic-like and tryptophan-like components consisted of the majority of the DOM in the HR aquatic system. The average percentage of three components in the NRU was: 37±8% of the fulvic-like component (C1), $30\pm15\%$ of the humic-like component (C2), and $33\pm8\%$ of the tryptophan-like component (C3), and the three components were distributed evenly in the NRU aquatic systems. The humic-like component (C2) takes an average of 57±8% in the NRD aquatic systems, and the humic-like component is the main contributor to the fluorescent DOM. The rank of the components' average percentage was C1 (40±6%)>C3 (31±8 %)>C2 (29±12%) in the HER.



Fig. 6. Comparisons of PARAFAC components among the four areas. a) the PARAFAC component loadings and b) the percentage of each component. The error bars represent the standard deviation (SD). Different lowercase letters indicate significant differences (P<0.05, LSD) among the study areas. The results of ANOVA represent as "NA", "**", and "***" (P>0.05, P<0.05, P<0.01, and P<0.001).



Fig. 7. The distribution of FRI results of four studied areas. a) the sum percentage of autochthonous regions. b) the sum percentage of allochthonous regions. The different lowercase letters indicate significant differences (P<0.05, LSD) among the study areas. The results of ANOVA represent as "NA", "*", "**", and "***" (P>0.05, P<0.05, P<0.01, and P<0.001).

The five regions of EEMs were classified into two parts. The regions of I, II, and IV were called the autochthonous region, and the other regions (i.e., III and V) were categorized as the allochthonous region. The P_{in} of each region was calculated, and Fig. 7 (a-b) shows the result of the FRI. The percentages of autochthonous and allochthonous regions distribute differently in the four areas (P<0.001, LSD). The NRD owns the highest percentage of the allochthonous region (71±10%) and the lowest percentage of the autochthonous region $(29\pm10\%)$ in four study areas. The percentage of the autochthonous region in HR (56±2%) is significantly greater than HER (50±6%) (P<0.05, LSD). DOM is affected not only by natural lake attributes but also by human activities in a lake's watershed [53]. There is a lot of agricultural nonpoint source pollution and sewage inputting around the HER, and the NRD is connected with the HER and the water is exchanged between two areas in large numbers. In addition, the sampling with low temperature resulted in the weakening biological activity in aquatic systems [54]. Totally, the results of FRI were similar to the PARAFAC model.

Fluorescence Indices Analysis

Fluorescence indices (i.e., BIX, FI, and HIX) can indicate the characteristic of DOM in aquatic systems well. As shown in Fig. 8 (a-b), the average of BIX in NRD (0.66±0.18) is significantly lower than HR (1.01±0.04), HER (0.95±0.08), and NRU (0.98±0.11) (P<0.001, LSD). A low average BIX value means that there is little autochthonous production in the aquatic systems. HR, NRU, and HER contain a relatively high level of autochthonous components in the DOM except for the NRD. In contrast to the BIX, the HIX reflects the concentration of the allochthonous ingredient of the DOM in water environments. The distribution of the HIX is opposite to the BIX, and the NRD owns the highest HIX mean (0.91±0.04) among the four study areas (P<0.001, LSD). The average of FI in NRD (2.30 ± 0.18) is greater than 1.9, and the mean of FI in other areas (i.e., NRU (1.86±0.25), HER (1.82±0.21), and HR (1.73±0.09)) were between 1.4 and 1.9. The humus in the NRD was generally generated by the microbial processes. The other study areas' FI value



Fig. 8. Comparisons of the fluorescence indices (BIX, FI, and HIX) among the four studied areas a). The error bars represent the Standard Deviation (SD). The different lowercase letters indicate significant differences (P<0.05, LSD) among the study areas. The BIX and HIX distribution of samples in the study areas b). The results of ANOVA represent as "NA", "*", "**", and "***" (P>0.05, P<0.05, P<0.01, and P<0.001).

was near 1.9 with more microbial humus than the terrigenous sources' humus. Collectively, these facts and speculations suggest that the main part of humus in the aquatic systems was produced by the microbial process when the surface runoff was weak. The process of microbial activities to convert protein-like components into fulvic-like and humic-like components, or higher levels of stable substances [55]. The photochemical oxidation process of DOM in the water was slowed down by the surface cover (solar cell panel) and the humus substances accumulated year by year in the NRD watershed [47]. Solar radiation is an important factor for the photodegradation of DOM and organic contaminants in aquatic systems [14]. Previous studies have shown that increase in the fluxes of UV radiation can substantially increase the quantity of reactive free radicals (i.e., HO· and H_2O_2) in waters [56].

Correlations between the General Water Quality Indices and Fluorescence Characteristic of DOM

The general water quality indices (i.e., COD, NH_4^+ , NO_2^- , TN, and TP) can evaluate and predict the change of the aquatic systems. These traditional indices contributed to the water quality class evaluation of the surface waters in China [57]. Depending on the conventional indices, it is difficult to accurately acquire the source analysis of the water pollution. PARAFAC components and FRI can provide information related to the sources of DOM in the water because the DOM owns the complex biogeochemical properties. To reveal the comprehensive water qualities and the distribution of the DOM characteristics of each sampling site, a principal component analysis (PCA) based on DOM fluorescence characteristics (i.e., PARAFAC component percentage, FRI results, and fluorescence indices) and general water indices were shown in Fig. 9 [58]. By using the package "psych" (2.1.6) in R (4.0.5) computing, the result of the Bartlett test was P < 0.001, and the value of the KMO was 0.65, indicating that the selected indices were suitable for the PCA analysis. Furthermore, the nonparametric multivariate analysis of the variance (Adonis) was used to verify the significance of the difference among the group of the four research areas, utilizing the software package "Vegan" (2.5-7) in R (4.0.5) [59]. The result presented that there was a significant difference among the four groups (P < 0.01, Adonis).

As shown in Fig. 9, the first and the second principal components (i.e., PC1 and PC2) can explain 54.7 % and 20.0 % of the total variance respectively. The first two principal components explain most of the variance for all data (74.7 %). The positive direction of PC1 is characterized by parameters representing the terrestrial DOM with high aromaticity such as C2 %, HIX, and III + V%. Conversely, the autochthonous indices (i.e., BIX, C3 %, and I + II + IV%) in the negative direction of PC1 represent the biological DOM contribution in the aquatic systems. The positive direction of PC2 is represented

by the water eutrophication and contamination level; this direction includes the indices of COD, DOC, NH⁺, NO_3^- , TN, and TP. It can be found that the samples of HR, NRU, and HER mainly posited at the left side of the PCA plot, demonstrating that these watersheds are featured by higher autochthonous substances. The percentage of fulvic-like components (C1 %) observed in the negative direction of PC1 illustrated that the biological activities generated not only the protein-like substances of DOM, but also the humus substances in the aquatic systems. The NRD samples located on the left of the PCA plot demonstrated high aromaticity and molecular size of DOM in the aquatic systems. The humic-like component (C2) contributed the most fluorescence loadings because the total fluorescence loadings (C1 + C2 + C3) were in the positive direction of PC1 and near the C2%. Most of the samples in HER were observed in at the positive direction of PC2 with high loadings of nutrients and COD. Agricultural nonpoint sources, aquaculture, and sewage caused tremendous water pollution, resulting in dire environmental circumstances in the HER. The NRD interconnected with HER and was influenced. At the same time, the lack of light contributes to the photochemical oxidation process of the DOM was delayed in the NRD because of surface covering. The NRU near the three coal mining areas had many subsidence areas. The position of the NRU samples in the negative direction of PC2 exhibits the characteristics of low nutrients and pollution. Watershed DOM and nutrient material dynamics change can be influenced by multiple interacting spatial scales and dimensions of hydrologic connectivity [60]. The subsidence areas around the NRU formed the buffer of agriculture nonpoint source and coal mining activities concurrently occupied a lot of the farmland; it is for these reasons agriculture nonpoint source pollution has



Fig. 9. The principal components analysis (PCA) of the four studied watersheds. The percentages in the axis titles indicate the percent variance explained by each principal component of the PCA. The different colors represent the various sampling areas respectively.

been effectively suppressed in the NRD study area. The spectral characteristics of DOM is also a sensitive indicator reflecting the potential impacts of land use [20]. The result of the PCA suggests that regional characteristics can substantially affect the DOM's composition and water qualities. At the same time, the Adonis test significantly proved the differences among the four aquatic regions.

Conclusions

The Hei River (HER) owned the poorest water quality because of agriculture nonpoint source pollution input. In contrast, the Ni River upstream (NRU) owned the best water quality due to the coal mining activities of the occupied farmland and subsidence ponds weaken the input of nonpoint pollution. In order to investigate the complex influencing factors and the heterogeneity of DOM in the basin, we have chosen spectroscopy to analyze the characteristics of DOM in the watershed. The UV-Vis spectrum indicated that there was no significant difference in the average molecule size of DOM (P>0.05, ANOVA) among the four studied areas, but the aromaticity of DOM in the NRU and the Ni River downstream (NRD) was significantly higher than in other areas (P<0.05, LSD). The PARAFAC model was used to analyze the EEMs of DOM, and the following three components were obtained (i.e., fulvic-like (C1), humic-like (C2), and tryptophan-like (C3)). The results of PARAFAC and FRI show that the allochthonous ingredients of DOM in the NRD were greater than the other study locations. The photovoltaic industry significantly influenced the properties of DOM in aquatic systems, and photochemical oxidation is an important factor affecting DOM content and composition. The principal component analysis (PCA) and Adonis test show that the HER is directly influenced by agriculture nonpoint source pollution and human activities because of the high nutrient concentration. The main ingredients of DOM in the NRD were terrestrial humus, but Huai River (HR), NRU, and HER was autochthonous substances. There was a significant difference (P<0.01, Adonis) of the DOM constitution and water qualities between the four research regions.

Although spectral methods on the characteristics of fast, cheap, and high sensitivity, it is vulnerable to the interference of inorganic ions in water. How to eliminate the influence of interfering substances on spectral accuracy is an urgent problem to be solved. DOM in the natural aquatic system is affected by multiple factors, and even complex interactions exist between factors. In this study, we considered the effects of agricultural nonpoint source input, sewage, coal mining activities, and water surface photovoltaic power generation on DOM properties and content. But there are still many factors not considered, such as temperature, precipitation, microbial activity, mineralization, adsorption, and desorption. At present, most studies on the distribution characteristics of watershed DOM are based on the surrounding land use, and this method is proved to be effective for the large drainage basin scale. However, the land use on a small watershed scale is difficult to meet the required accuracy requirements, and the conclusions are also questionable. Therefore, how to achieve accurate DOM analysis on the minor drainage basin scale is meaningful.

Acknowledgments

This work was supported by the Ecological Environment Special Project of the Key Research and Development Program of Anhui Province (202004i07020012).

Conflict of Interest

The authors declare no conflict of interest.

References

- CHOW M.I., LUNDIN J.I., MITCHELL C.J., DAVIS J.W., YOUNG G., SCHOLZ N.L., MCINTYRE J.K. An urban stormwater runoff mortality syndrome in juvenile coho salmon. Aquatic Toxicology, 214, 10, 2019.
- TREDWAY J.C., HAVLICK D.G. Assessing the potential of low-impact development techniques on runoff and streamflow in the Templeton Gap Watershed, Colorado. The Professional Geographer, 69 (3), 372, 2017.
- XU B., HUANG Q., LI J., LI P., XIANG Y., TAKAHASHI, J. Effects of land use on the amount and composition of dissolved organic matter in a Chinese Headwater Stream watershed. Polish Journal of Environmental Studies, 25 (1), 385, 2016.
- KRISNAYANTI D.S., BUNGANAEN W., FRANS J.H., SERAN Y.A., LEGONO D. Curve number estimation for ungauged watershed in semiarid region. Civil Engineering Journal, 7 (6), 1070, 2021.
- XIA J., WANG L.F., YU J.J., ZHAN C.S., ZHANG Y.Y., QIAO Y.F., WANG Y.L. Impact of environmental factors on water quality at multiple spatial scales and its spatial variation in Huai River Basin, China. Science China-Earth Sciences, 61 (1), 82, 2018.
- WANG J., LIU P. Study on environmental pollution and governance in rural development. Fresenius Environmental Bulletin, 28 (12), 9218, 2019.
- HU J., CHEN X., CHEN Y., LI C., REN M., JIANG C., ZHENG L. Nitrate sources and transformations in surface water of a mining area due to intensive mining activities: Emphasis on effects on distinct subsidence waters. Journal of environmental management, 298, 113451, 2021.
- 8. LANGER P. Mining of groundwater in contemporary urban development on the example of selected German spas. **2019**.
- COBLE P.G. Marine optical biogeochemistry: The chemistry of ocean color. Chemical Reviews, 107 (2), 402, 2007.
- LIU D., DU Y.X., YU S.J., LUO J.H., DUAN H.T. Human activities determine quantity and composition of dissolved

organic matter in lakes along the Yangtze River. Water Research, 168, 11, 2020.

- NEBBIOSO A., PICCOLO A. Molecular characterization of dissolved organic matter (DOM): A critical review. Analytical & Bioanalytical Chemistry, 405 (1), 109, 2013.
- HE Y.H., SONG N., JIANG H.L. Effects of dissolved organic matter leaching from macrophyte litter on black water events in shallow lakes. Environmental Science and Pollution Research, 25 (10), 9928, 2018.
- LI P.H., HUR J. Utilization of UV-Vis spectroscopy and related data analyses for dissolved organic matter (DOM) studies: A review. Critical Reviews in Environmental Science and Technology, 47 (3), 131, 2017.
- CARSTEA E.M., POPA C.L., BAKER A., BRIDGEMAN J. In situ fluorescence measurements of dissolved organic matter: A review. Science of the Total Environment, 699, 16, 2020.
- 15. LI L., WANG Y., ZHANG W.J., YU S.L., WANG X.Y., GAO N.Y. New advances in fluorescence excitationemission matrix spectroscopy for the characterization of dissolved organic matter in drinking water treatment: A review. Chemical Engineering Journal, 381, 12, 2020.
- 16. WANG D., WANG C., JI M., SUN L., QIU C., LO S. Characterization of dissolved organic matter removal during biological treatment of commingled chemical industrial wastewater: Relationship with fluorescent dissolved organic matter transformation. Polish Journal of Environmental Studies, 29 (1), 307, 2019.
- COHEN E., LEVY G.J., BORISOVER M. Fluorescent components of organic matter in wastewater: Efficacy and selectivity of the water treatment. Water Research, 55, 323, 2014.
- CHEN W., WESTERHOFF P., LEENHEER J.A., BOOKSH K. Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. Environmental Science & Technology, 37 (24), 5701, 2003.
- EKWUEME B. N., AGUNWAMBA J. C. Modeling the influence of meteorological variables on runoff in a tropical watershed. Civil Engineering Journal, 6 (12), 2344, 2020.
- 20. YANG X.L., YU X.B., CHENG J.R., ZHENG R.Y., WANG K., DAI Y.X., TONG N.J., CHOW A.T. Impacts of land-use on surface waters at the watershed scale in southeastern China: Insight from fluorescence excitationemission matrix and PARAFAC. The Science of the total environment, 627, 647, 2018.
- WEISHAAR J.L., AIKEN G.R., BERGAMASCHI B.A., FRAM M.S., FUJII R., MOPPER K. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. Environmental Science & Technology, 37 (20), 4702, 2003.
- HELMS J.R., STUBBINS A., RITCHIE J.D., MINOR E.C., KIEBER D.J., MOPPER K. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnology and Oceanography, 53 (3), 955, 2008.
- TENG C.Y., ZHOU K.G., ZHANG Z., PENG C.H., CHEN W. Elucidating the structural variation of membrane concentrated landfill leachate during Fenton oxidation process using spectroscopic analyses. Environmental Pollution, 256, 9, 2020.
- ZHANG Z., TENG C.Y., ZHOU K.G., PENG C.H., CHEN W. Degradation characteristics of dissolved organic matter in nanofiltration concentrated landfill leachate during electrocatalytic oxidation. Chemosphere, 255, 9, 2020.

- STEDMON C.A., BRO R. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial. Limnology and Oceanography: Methods, 6 (11), 572, 2008.
- PUCHER M., WÜNSCH U., WEIGELHOFER G., MURPHY K., HEIN T., GRAEBER D. staRdom: Versatile software for analyzing spectroscopic data of dissolved organic matter in R. Water, 11 (11), 2366, 2019.
- STEDMON C.A., BRO R. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial. Limnology and Oceanography: Methods, 6 (11), 572, 2008.
- MURPHY K.R., STEDMON C.A., GRAEBER D., BRO R. Fluorescence spectroscopy and multi-way techniques PARAFAC. Analytical Methods, 5 (23), 6557, 2013.
- MURPHY K.R., STEDMON C.A., WENIG P., BRO R. OpenFluor – an online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods, 6 (3), 658. 2014.
- CHEN W., YU H.Q. Advances in the characterization and monitoring of natural organic matter using spectroscopic approaches. Water Research, 190, 116759, 2021.
- XU H., YAN M., LI W., JIANG H., GUO L. Dissolved organic matter binding with Pb (II) as characterized by differential spectra and 2D UV–FTIR heterospectral correlation analysis. Water Research, 144, 435, 2018.
- 32. ZHANG D.L., CHEN W., CHEN H., YU H.Q., KASSAB G., CHENG J.X. Chemical imaging of fresh vascular smooth muscle cell response by epi-detected stimulated Raman scattering. Journal of Biophotonics, 11 (2), 7, 2018.
- MAIE N., PARISH K.J., WATANABE A., KNICKER H., BENNER R., ABE T., JAFFÉ R. Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem. Geochimica et Cosmochimica Acta, 70 (17), 4491, 2006.
- HUGUET A., VACHER L., RELEXANS S., SAUBUSSE S., FROIDEFOND J.M., PARLANTI E. Properties of fluorescent dissolved organic matter in the Gironde Estuary. Organic Geochemistry, 40 (6), 706, 2009.
- OHNO T. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. Environmental Science & Technology, 36 (4), 742, 2002.
- 36. ZHANG S., WANG Z., ZHANG B., SONG K., LIU D., REN C. Characteristics of nonpoint source pollution from an agricultural watershed during rainfall events. International Conference on Mechanic Automation and Control Engineering, 2010.
- 37. PENG C., HE J.T., WANG M., ZHANG Z., WANG L. Identifying and assessing human activity impacts on groundwater quality through hydrogeochemical anomalies and NO_3^- , NH_4^+ , and COD contamination: a case study of the Liujiang River Basin, Hebei Province, P.R. China. Environmental Science and Pollution Research, **25** (4), 3539, **2018**.
- LIAO Z., CHU J., LUO C., CHEN H. Revealing the characteristics of dissolved organic matter in urban runoff at three typical regions via optical indices and molecular composition. Journal of Environmental Sciences, 108, 8, 2021.
- 39. MA Y., HAO S., ZHAO H., FANG J., ZHAO J., LI X. Pollutant transport analysis and source apportionment of the entire non-point source pollution process in separate sewer systems. Chemosphere, **211**, 557, **2018**.
- 40. LIU C., LI Z.W., BERHE A.A., XIAO H.B., LIU L., WANG D.Y., ZENG G.M. Characterizing dissolved

organic matter in eroded sediments from a loess hilly catchment using fluorescence EEM-PARAFAC and UV-Visible absorption: Insights from source identification and carbon cycling. Geoderma, **334**, 37, **2019**.

- SHANG P., LU Y.H., DU Y.X., JAFFÉ R., FINDLAY R.H., WYNN A. Climatic and watershed controls of dissolved organic matter variation in streams across a gradient of agricultural land use. Science of the Total Environment, 612, 1442, 2018.
- 42. KULKARNI H., MLADENOV N., DATTA S. Effects of acidification on the optical properties of dissolved organic matter from high and low arsenic groundwater and surface water. Science of the Total Environment, **653**, 1326, **2019**.
- 43. WU H.H., XU X.K., FU P.Q., CHENG W.G., FU C.S. Responses of soil WEOM quantity and quality to freezethaw and litter manipulation with contrasting soil water content: A laboratory experiment. Catena, **198**, 13, **2021**.
- 44. LEE D., KWON M., AHN Y., JUNG Y., NAM S.N., CHOI I.H., KANG J.W. Characteristics of intracellular algogenic organic matter and its reactivity with hydroxyl radicals. Water Research, 144, 13, 2018.
- 45. HEIBATI M., STEDMON C.A., STENROTH K., RAUCH S., TOLJANDER J., SAVE-SODERBERGH M., MURPHY K.R. Assessment of drinking water quality at the tap using fluorescence spectroscopy. Water Research, 125, 1, 2017.
- 46. ISHII, S.K.L., BOYER T.H. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineered systems: A critical review. Environmental Science & Technology, 46 (4), 2006, 2012.
- MURPHY K.R., TIMKO S.A., GONSIOR M., POWERS L.C., WÜNSCH U.J., STEDMON C.A. Photochemistry illuminates ubiquitous organic matter fluorescence spectra. Environmental Science & Technology, 52 (19), 11243, 2018.
- MURPHY K.R., HAMBLY A., SINGH S., HENDERSON R.K., BAKER A., STUETZ R., KHAN S.J. Organic matter fluorescence in municipal water recycling schemes: Toward a unified PARAFAC model. Environmental Science & Technology, 45 (7), 2909, 2011.
- 49. BORISOVER M., LAOR Y., PARPAROV A., BUKHANOVSKY N., LADO M. Spatial and seasonal patterns of fluorescent organic matter in Lake Kinneret (Sea of Galilee) and its catchment basin. Water Research, 43 (12), 3104, 2009.

- SONG F., WU F., FENG W., LIU S., HE J., LI T., BAI Y. Depth-dependent variations of dissolved organic matter composition and humification in a plateau lake using fluorescence spectroscopy. Chemosphere, 225, 507, 2019.
- WANG S., QIN J., SUN H., CHEN W., XIE B. Photodegradation characteristics of dissolved organic matter (DOM) in alpine soils of the eastern margin of Tibetan Plateau. Polish Journal of Environmental Studies, 30 (3), 2815, 2021.
- 52. VAEHAETALO A.V., WETZEL R.G. Photochemical and microbial decomposition of chromophoric dissolved organic matter during long (months+years) exposures. Marine Chemistry, **89**, 313, **2004**.
- 53. ZHANG Y. L., ZHANG E., YIN Y., VAN DIJK M. A., FENG L., SHI Z., LIU M., QIN B. Characteristics and sources of chromophoric dissolved organic matter in lakes of the Yungui Plateau, China, differing in trophic state and altitude. Limnology and Oceanography, 55 (6), 2645, 2010.
- LARSEN S., ANDERSEN T., HESSEN D.O. Climate change predicted to cause severe increase of organic carbon in lakes. Global Change Biology, 17 (2), 1186, 2011.
- 55. ZHANG K., GAO J., MEN D., ZHAO X., WU S. Insight into the heavy metal binding properties of dissolved organic matter in mine water affected by water-rock interaction of coal seam goaf. Chemosphere, 265, 129134, 2021.
- 56. QIAN J., MOPPER K., KIEBER D. J. Photochemical production of the hydroxyl radical in Antarctic waters. Deep-Sea Research Part I: Oceanographic Research Papers, 48 (3), 741, 2001.
- 57. LU X., ZHOU B., VOGT R.D., SEIP H.M., XIN Z., EKENGREN O. Rethinking China's water policy: the worst water quality despite the most stringent standards. Water International, **41** (7), 1044, **2016**.
- 58. YU H., GUI H., JIANG Y., LI Z., WANG M., FANG H., ZHANG Y. Characteristics of dissolved organic matter content in urban rivers under different environmental impact zones: A case study of China's Tuo River. Polish Journal of Environmental Studies, **29** (5), 3891, **2020**.
- SYLVAIN I.A., ADAMS R.I., TAYLOR J.W. A different suite: The assemblage of distinct fungal communities in water-damaged units of a poorly-maintained public housing building. PLoS One, 14 (3), 19, 2019.
- COVINO T. Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. Geomorphology, 277, 133, 2017.