

RESEARCH

Open Access



Insight into temporal–spatial variations of DOM fractions and tracing potential factors in a brackish-water lake using second derivative synchronous fluorescence spectroscopy and canonical correlation analysis

Kuotian Lu^{1†}, Weining Xu^{1†}, Huibin Yu^{1*}, Hongjie Gao^{1*}, Xiaobo Gao² and Ningmei Zhu^{1,3}

Abstract

Background: Insight into temporal–spatial variations of dissolved organic matter (DOM) fractions were undertaken to trace potential factors toward a further understanding aquatic environment in Lake Shahu, a brackish-water lake in northwest China, using synchronous fluorescence spectroscopy (SFS) combined with principal component analysis (PCA), second derivative and canonical correlation analysis (CCA).

Result: Five fluorescence peaks were extracted from SFS by PCA, including tyrosine-like fluorescence (TYLF), tryptophan-like fluorescence (TRLF), microbial humic-like fluorescence (MHLF), fulvic-like fluorescence (FLF), and humic-like fluorescence (HLF), whose relative contents were obtained by second derivative synchronous fluorescence spectroscopy. The increasing order of total fluorescence components contents was July ($11,789.38 \pm 12,752.61$) < April ($12,667.58 \pm 15,246.91$) < November ($19,748.87 \pm 17,192.13$), which was attributed to tremendous enhancement in TYLF content from April (1615.56 ± 258.56) to November (5631.96 ± 634.82). The PLF (the sum of TYLF and TRLF) dominated the fluorescence components, whose proportion was 40.55, 37.09, or 46.91% in April, July, or November. DOM fractions in November were distinguished from April and July, which could be attributed to that water of the Yellow River was continuously loaded into the lake as water replenishment from April to September. From the replenishment period to non-replenishment, the contents of the five components gradually changed from low in the middle and high around the lake to high throughout entire lake. Based on the CCA results, the potential factors included TYLF, TRLF, MHLF, SD, and BOD₅ in April, which were relative to organic matter pollution. The potential factors contained TYLF, TRLF, FLF, Chl-a, TP, COD_{Cr}, and DO in July, indicating the enrichment of TP lead algae and plants growth. The potential factors in November consisted of TYLF, TRLF, COD_{Cr}, SD, TN, and FLF, representing residue of the algae and plants have been deeply degraded.

Conclusion: The replenishment of water led to enrichment of TP, resulting in growth of algae and plants, and was the key factor of water quality fluctuations. This work provided a workflow from perspective of DOM to reveal causes of water quality fluctuations in a brackish-water lake and may be applied to other types of waterbodies.

*Correspondence: yhbxbx@163.com; gaohj@craes.org.cn

†Kuotian Lu and Weining Xu co-first author

¹ State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, 100012 Beijing, People's Republic of China

Full list of author information is available at the end of the article

Keywords: Dissolved organic matter, Synchronous fluorescence spectroscopy, Principal component analysis, Second derivative synchronous fluorescence spectra, Canonical correlation analysis

Background

Dissolved organic matter (DOM) is a complex mixture consisting of proteins, polysaccharides, and humic substances [1, 2]. DOM has been defined as organic matter in solution that can pass through a 0.45- μm membrane filter [3]. It exists ubiquitously in natural and engineered aquatic systems, is an important carrier for pollutants, which is associated with the retention and release of nutrients, biological availability, and the migration and transformation of contaminants [4]. Therefore, the study of the variation of DOM composition and distribution is significant for understanding aquatic environments and evaluating water quality.

Various characteristic techniques, such as HPLC, FTIR, UV-vis, and fluorescence spectroscopy, have been applied to determine the structure, composition, and functionalities of DOM [5–8]. Fluorescence spectroscopy techniques, including excitation–emission matrix spectroscopy (EEMs) and synchronous fluorescence spectroscopy (SFS), are non-destructive techniques characterized by rapid analysis, high sensitivity, and simple operation [8, 9]. Thus, fluorescence spectroscopy techniques are emerging as available tools that have been widely employed for investigating DOM. EEMs provide a whole range of incremented excitation wavelengths and the corresponding emission data with large and complicated datasets, including various peaks and their specific locations. Recently, EEMs combined with parallel factor analysis (PARAFAC) was widely applied to decompose EEMs into fluorescence components [10, 11]. However, PARAFAC requires large amounts of samples to ensure that the extracted fluorescence fractions are correct. Specifically, SFS, a method that scans both excitation and emission monochromators with a selected constant wavelength, provides simpler spectra, which could be easier to interpret without losing important information [12, 13]. SFS also provides better structure and resolved peaks, which can be easily analyzed and differentiate the fluorescence spectra of samples of various origins and is suitable for a small number of samples. Statistical methods such as principal component analysis (PCA) could be used in SFS in order to acquire more information to assist in further analysis. SFS combined with PCA can decompose complex synchronous fluorescence spectra and reveal the similarity and dissimilarity between the samples [14, 15]. Moreover, derivatives are applied in SFS to reduce extensive spectroscopic overlap and eliminate matrix interference [16].

Lake Shahu (106°18'E, 38°45'N), a terminal lake, is a typical brackish-water lake located in an arid region that frequently experiences dropping water quality levels [17]. Its level of water quality was highly affected by water replenishment, which input into the lake with large amounts of contaminants introducing eutrophication [18, 19]. Lake Shahu was eutrophic during the period of replenishment, especially in the July, which contributed to high TP and TN concentrations in the replenishment water inflowing from the Yellow River [18]. Organic pollution also plagued Lake Shahu reflecting in the high concentration of COD_{Cr} . DOM can alter nutrient availability, and that can strongly affect algae abundance and community structure across lakes [20]. Thus, it is urgent to investigate the variation of DOM composition and distribution in Lake Shahu in order to understand the dramatic fluctuations of water quality.

The objectives of this study were (a) to extract fluorescence components of DOM from Lake Shahu and characterize their temporal–spatial variations by SFS combined with PCA and second derivative; (b) to seek potential factors among water quality parameters and fluorescence components and identify pollution sources using canonical correlation analysis (CCA).

Materials and methods

Study area and sample collection

Lake Shahu covered an area of 8.2 km^2 and an average depth of 2.2 m [21]. It had an arid and semi-arid continental climate with annual average temperature of 9.74 °C, annual average precipitation of 172.5 mm, and annual average evaporation of 1755.1 mm [22]. The south side of Lake Shahu is a sandy beach, the east side is adjacent to a wetland, and the large areas of farmland are locating on the west and north sides of Lake Shahu. As a terminal lake in an arid region with no natural surface runoff or outflow, Lake Shahu is particularly sensitive to an extremely high evaporation proportion [23]. For maintaining the ecological water storage, replenishment water from the Yellow River is loaded into the lake from April to September since 2013 [24] (Fig. 1). The volume of replenishment water from Donggan and Bayi channels is 23.40 and 16.13 million m^3 , respectively, in 2020. Because of the intense evaporation and the water recharge, the lake has undergone dramatic changes in water quality and accumulated contamination from various sources.

Eleven sampling sites were selected based on recharge water access, potential sources of pollution, and

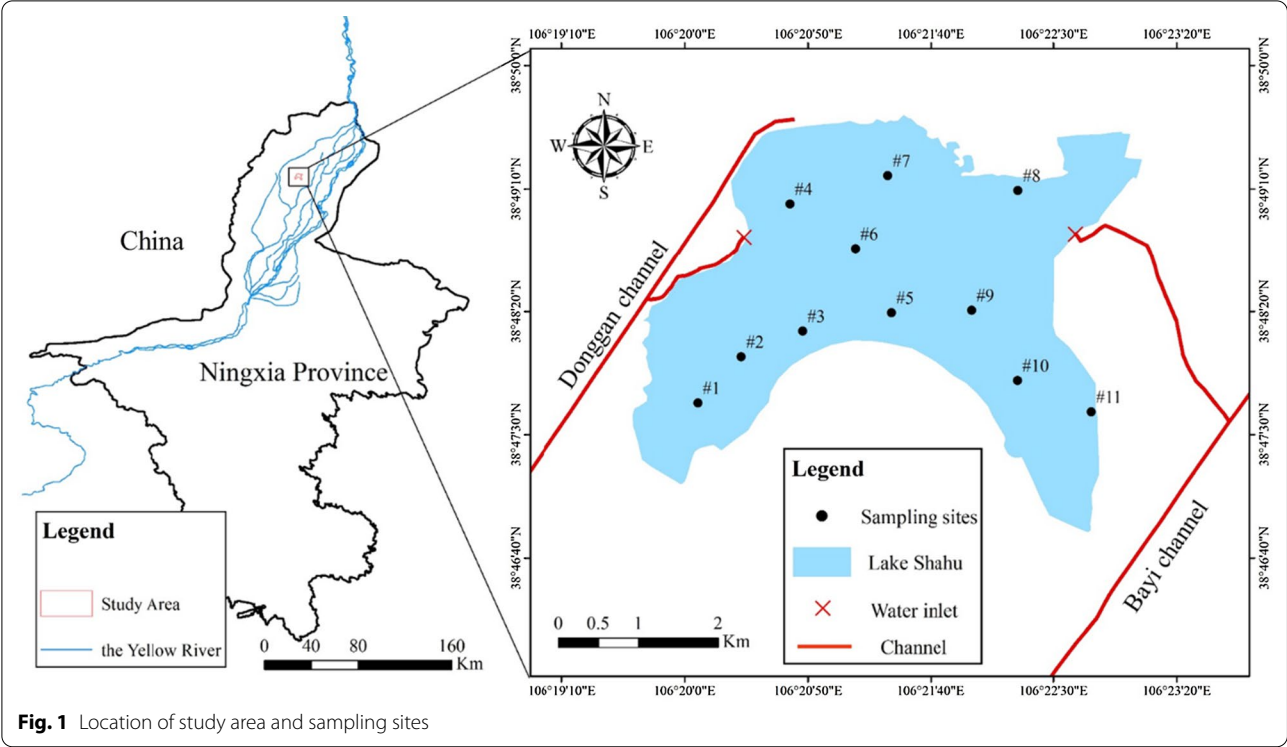


Fig. 1 Location of study area and sampling sites

geographical proximity (Fig. 1). Sampling sites #1–3 were located on the southwest of Lake Shahu. Sampling sites #4–6 were in the central region of the lake, especially #6 close to the water inlet. Sampling sites #7 and #8 adjoined a resort, and the sampling sites #9–11 were in Niaodao island. Water samples from Lake Shahu were collected in April, July, and November 2020, which were the early water replenishment period, mid-replenishment period, and non-replenishment period, respectively. At each selected sampling site, water sample collection was carried out with a 5 L Van Dom water sampler. Three water samples were collected from different depths (10, 20, and 30 cm) and completely mixed with the same volume of water. The samples were shipped to the lab in a cooled container for analysis.

Data preparation

Measurements of physico-chemical parameters

Temperature (TEMP), electrical conductivity (EC), and dissolved oxygen (DO) were measured using a YSI portable multiparameter water quality tester in situ. The Secchi depth (SD) was measured using a standard Secchi disk with black and white quarters. The samples were transported to the laboratory in pre-cleaned polyethylene bottles, which were used to measure chemical oxygen demand (COD_{Cr}), total phosphorus (TP), total nitrogen (TN), ammonia nitrogen (NH₃-N), chlorophyll a (Chl-a), and biochemical oxygen demand (BOD₅). The

standard analytical methods for those parameters are presented in Table 1.

Measurements of synchronous fluorescence spectroscopy

Water samples were filtered using glass fiber filters (Millipore, 0.45 μm fiber Ø) before the fluorescence determination. The SFS was measured using a Hitachi Fluorescence Spectrophotometer (F-7000) with a 1 cm quartz cuvette, which equipped with the fluorescence solution 1.00.000 (FL-solution software) for data processing. PMT voltage was set at 700 V, and scan speed was fixed at 240 nm min⁻¹. The SFS was obtained by a constant wavelength difference ($\Delta\lambda = \lambda_{em} - \lambda_{ex} = 55$ nm) with the excitation wavelength range from 260 to

Table 1 Analytical methods for selected parameters of water samples

Parameters	Methods	Units
COD _{Cr}	Potassium dichromate method	mg/L
TP	Ammonium molybdate method	mg/L
TN	Ultraviolet spectrophotometry	mg/L
NH ₃ -N	Nessler reagent method	mg/L
Chl-a	Spectrophotometry	mg/m ³
BOD ₅	5-Day BOD test	mg/L

550 nm [9]. Before further analysis, the spectrum of blank was subtracted from all spectra.

Analysis methods

Statistical analysis

PCA was performed for the SFS of DOM at 11 sampling sites by SPSS 25.0 software to identify the variations of DOM fractions and to trace dominated fluorescence components in different replenishment period. The sampling sites were set as variables and fluorescence intensity of spectroscopic wavelengths was set as cases, when performed PCA. Based on the score plots for spectral wavelengths, spectroscopic waveform and dominated fluorescence of each principal component (PC) was characterized. And the variations of DOM fractions were investigated by loading plots for the 11 sites. The potential factors between water quality parameters and DOM fractions were traced using the CCA, which was carried out by Canoco 4.5, with multivariate direct gradient analysis [25].

Second derivative method

The second derivative method was used to reduce extensive spectroscopic overlap and identify accurate wavelength range of each fluorescence peak of SFS by Origin 2021 software. After the second derivative, the fluorescent peaks were transformed into valleys, while the valleys were transformed into peaks. Therefore, the second derivative synchronous fluorescence spectroscopy (SDSFS) should be normalized with the intensities of fluorescence spectra multiplied by negative one [26]. For removing excess noise, the Savitzky–Golay method with 10 points of windows was applied to smooth the SFS after second derivative. The interval in the process was 2 nm.

Results and discussion

Spatio-temporal characteristics

Parameters of water quality

The spatio-temporal variation of water quality was displayed through the matrices of water quality parameters (Fig. 2). The highest average TEMP occurred in July (24.2 ± 1.03 °C). Sampling site #7 in July showed the highest TEMP (26.1 °C), and the lowest TEMP (8.0 °C) was observed in November at sampling sites #1 and #6 (Fig. 2a). The EC mean values decreased in the order of April (206.09 ± 18.11 $\mu\text{S cm}^{-1}$) > July (192.10 ± 16.61 $\mu\text{S cm}^{-1}$) > November (161.7 ± 32.74 $\mu\text{S cm}^{-1}$) (Fig. 2b). In April, the EC values at sampling sites #3 (172 $\mu\text{S cm}^{-1}$) and #4 (173 $\mu\text{S cm}^{-1}$) were the lowest. In July and November, site #6 showed the lowest EC values at 146 $\mu\text{S cm}^{-1}$ and 82 $\mu\text{S cm}^{-1}$, respectively. The DO mean value in July was higher than those in April and November (Fig. 2c), especially at sampling sites #4 (13.5 mg L^{-1}) and

#6 (13.4 mg L^{-1}), which showed the highest DO values in July. The SD mean value increased in the order of July (52.4 ± 8.09 cm) < November (55.9 ± 19.99 cm) < April (64.09 ± 9.93 cm) (Fig. 2d).

All sampling sites, except site #6, presented considerably higher concentration of the COD_{Cr} in November than those in July and April (Fig. 2e). The highest COD_{Cr} value was observed at site #1 (30 mg L^{-1}) in November, and site #4 (12 mg L^{-1}) showed the lowest concentration of COD_{Cr} in April. The highest mean TP value was observed in November (0.038 mg L^{-1}), followed by July (0.033 mg L^{-1}) and April (0.026 mg L^{-1}) (Fig. 2f). An evidence promotion of concentration of TP occurred from April to July, especially sampling sites adjacent to water inlets (Figs. 1, 2f). This could be associated with the water recharge with higher amounts of TP entered the lake in July, the period of intensive agricultural activities, which will cause the growth of algae and aquatic plants. Moreover, sampling sites #9–11 formed a zone with a higher concentration, and the highest TP value (0.06 mg L^{-1}) was observed at sampling site #9 in July TP (Fig. 2f). The TN mean values increased in the order of July (0.69 ± 0.037 mg L^{-1}) < November (0.93 ± 0.28 mg L^{-1}) < April (1.01 ± 0.20 mg L^{-1}) (Fig. 2g). The TN value at sampling site #6 (1.51 mg L^{-1}) in November was the highest, and the lowest value (0.64 mg L^{-1}) was obtained at site #5 in July. Sampling site #6 presented the highest TN value in each month. In contrast, the $\text{NH}_3\text{-N}$ at site #6 showed the lowest value in each month. The $\text{NH}_3\text{-N}$ mean values increased in the order of November (0.09 ± 0.03 mg L^{-1}) < July (0.12 ± 0.03 mg L^{-1}) < April (0.13 ± 0.05 mg L^{-1}) (Fig. 2h). The highest $\text{NH}_3\text{-N}$ value was obtained in April at site #5. The highest Chl-a mean value was presented in November (11.45 ± 5.01 $\mu\text{g L}^{-1}$), followed by July (6.36 ± 3.70 $\mu\text{g L}^{-1}$) and April (3.73 ± 3.00 $\mu\text{g L}^{-1}$) (Fig. 2i). The sampling sites #9–11 showed higher concentrations of Chl-a (Fig. 2i), in which the trend was similar to the TP, but the highest Chl-a value (19 $\mu\text{g L}^{-1}$) was observed at sampling site #5 in November. The BOD_5 mean values increased in the order of April (1.66 ± 0.38 mg L^{-1}) < November (2.28 ± 0.24 mg L^{-1}) < July (2.37 ± 0.18 mg L^{-1}) (Fig. 2j). Sampling site #9 in July showed the lowest BOD_5 (1.2 mg L^{-1}), and the highest BOD_5 (2.7 mg L^{-1}) was observed at sampling sites #8 and #3 in November and July, respectively.

Noticeably, the lower concentration of COD_{Cr} , $\text{NH}_3\text{-N}$, Chl-a, and BOD_5 were exhibited at site #6 were lower than the concentration of other sites, representing a higher level of water quality in central of the lake (Fig. 1). In addition, the values of COD_{Cr} , TP, Chl-a, and BOD_5 in November and July were higher than those in April. This indirectly indicated that the phenomena of algae and aquatic plants growth, intense anthropogenic

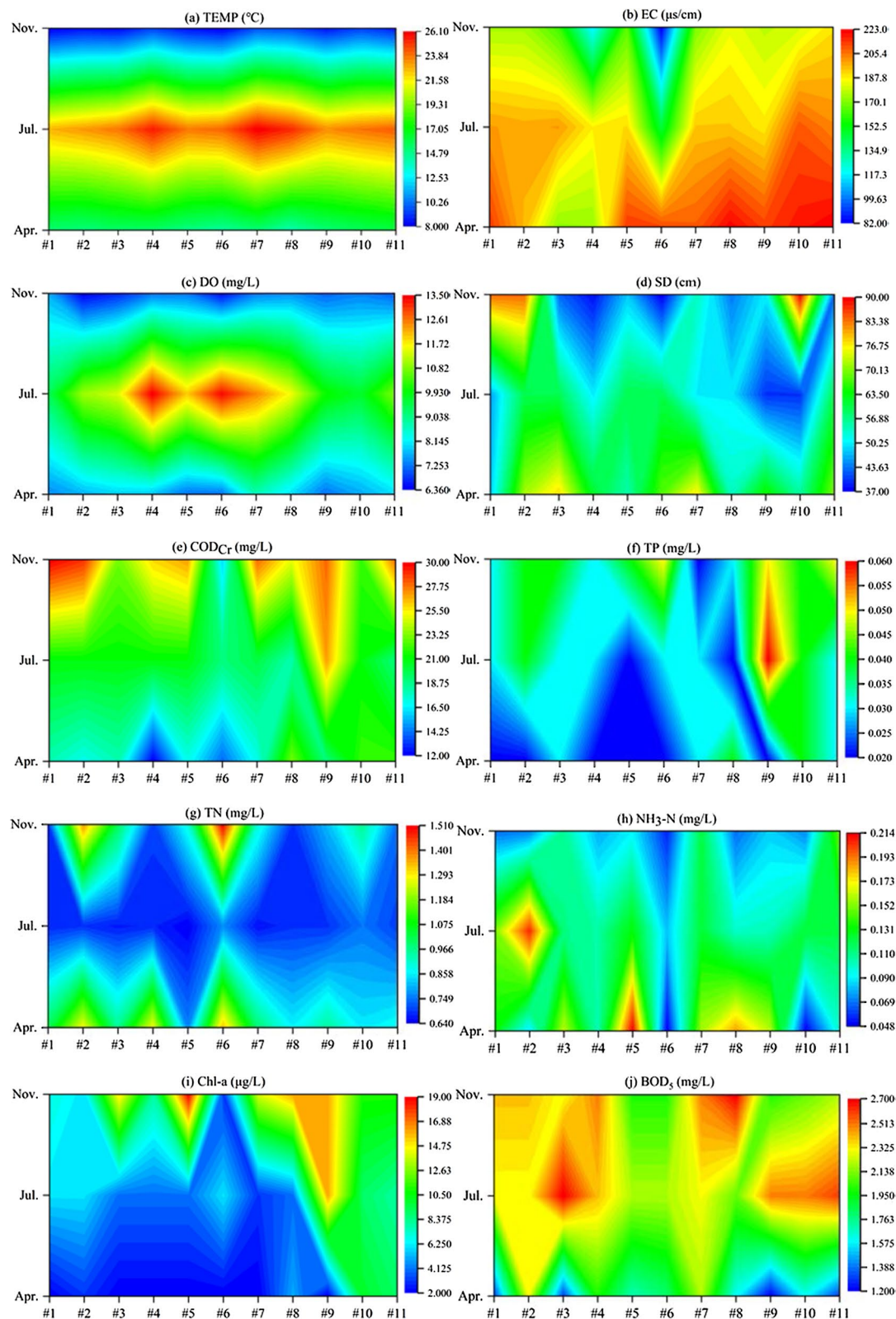


Fig. 2 Spatio-temporal distributions of bio-chemical (c, e, f, g, h) and physical (a, b, d) parameters in Lake Shahu

activities, by-product of algae and decomposition of plants occurred in July and November. According to the spatio-temporal variation of water quality, Lake Shahu is mainly polluted by organic matter in July and November.

Synchronous fluorescence spectroscopy

SFS of DOM from Lake Shahu exhibited a prominent peak, a relatively weak peak, and two broad shoulders (Fig. 3). The prominent peak at the wavelengths of 260–310 or 260–310 nm was denoted as the protein-like fluorescence (PLF) component, containing the tyrosine-like (TYLF) and tryptophan-like fluorescence (TRLF) [27]. The PLF in surface waters was at lower contents than humic substances [28]. Thus, a higher level of PLF could be associated with exogenous pollution. And the TRLF in natural water was mainly impacted by anthropogenic activities [29], while the TYLF was plant-derived [30] and biodegradation-derived DOM [31]. In addition, the TRLF is one of the significance nutrients for aquatic plants and microorganisms [31, 32]. The first shoulder presented at the wavelength range from 300 to 345 or 310 to 355 nm and was associated with the microbial humic-like fluorescence (MHLF) component, which concerned microbial activities [9]. The weak peak at 345–420 or 355–420 nm was assigned to the fulvic-like fluorescence (FLF) component from lignin and other terrestrial plant-derived

precursor material [33]. The second shoulder was observed at the wavelength range of 420–500 nm, which was related to the humic-like fluorescence (HLF) component [34]. Obviously, the peaks with minor fluorescence intensity were covered up by the preponderant one.

Principal component analysis

PCA was employed on the SFS of Lake Shahu in order to decompose the overlaps of spectrums and qualitatively investigate the spectral composition of datasets from diverse sampling periods. The Kaiser–Meyer–Olkin (KMO) and Bartlett sphericity test were first performed to test the adequacy and applicability of factor analysis/principal component analysis [35]. The KMO values of the four datasets were 0.94, 0.905, 0.932, and 0.843, respectively, and the significance levels of the Bartlett sphericity test were less than 0.001, indicating that the SFS was suitable for PCA. According to the loadings and scores of principal components (PCs), the characteristics of sampling sites and dominated fluorescence components of DOM could be identified (Fig. 4).

PCA on the each SFS datasets yielded two PCs, which accounted for 99.37, 99.856, 99.883, and 99.820% of total variables, respectively. Five fluorescence peaks were extracted by PCA (Fig. 4b). As shown in the map of PC

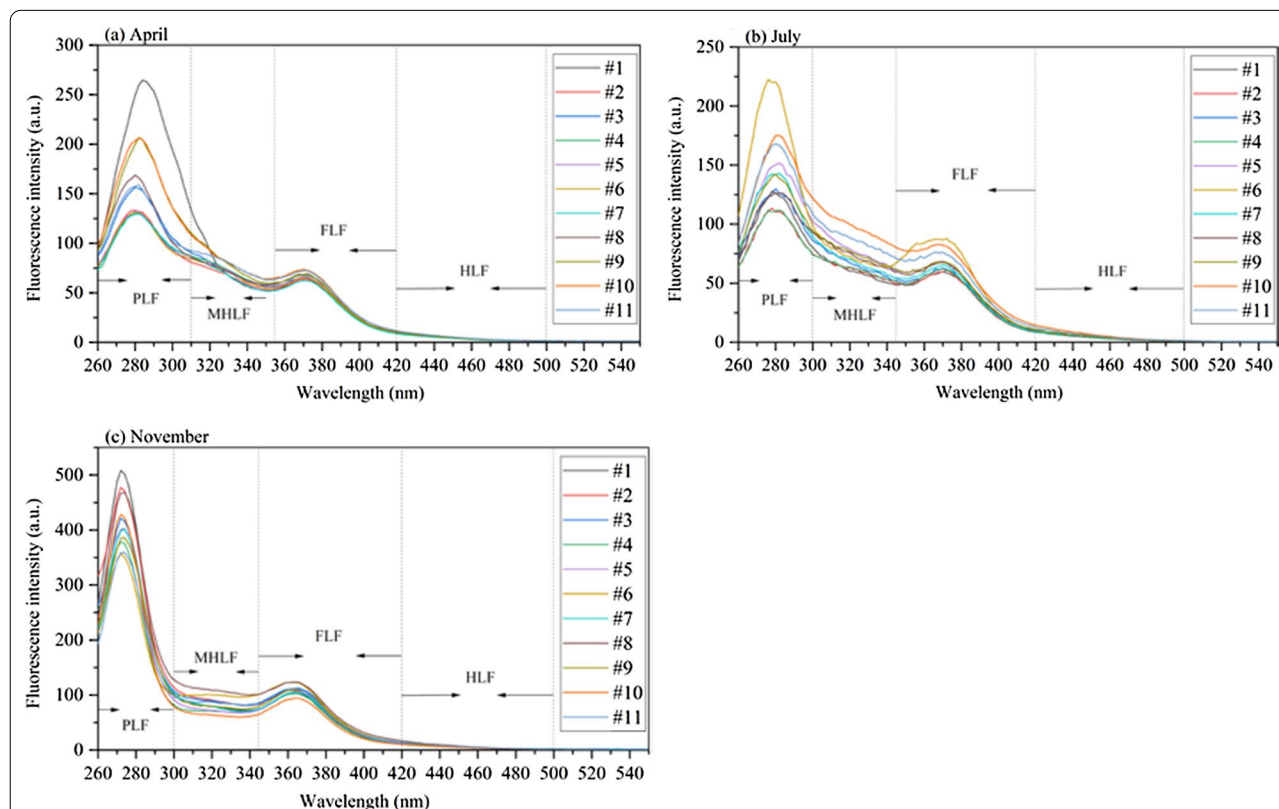


Fig. 3 Synchronous fluorescence spectra of the dissolved organic matter from Lake Shahu in April (a), July (b), and November (c)

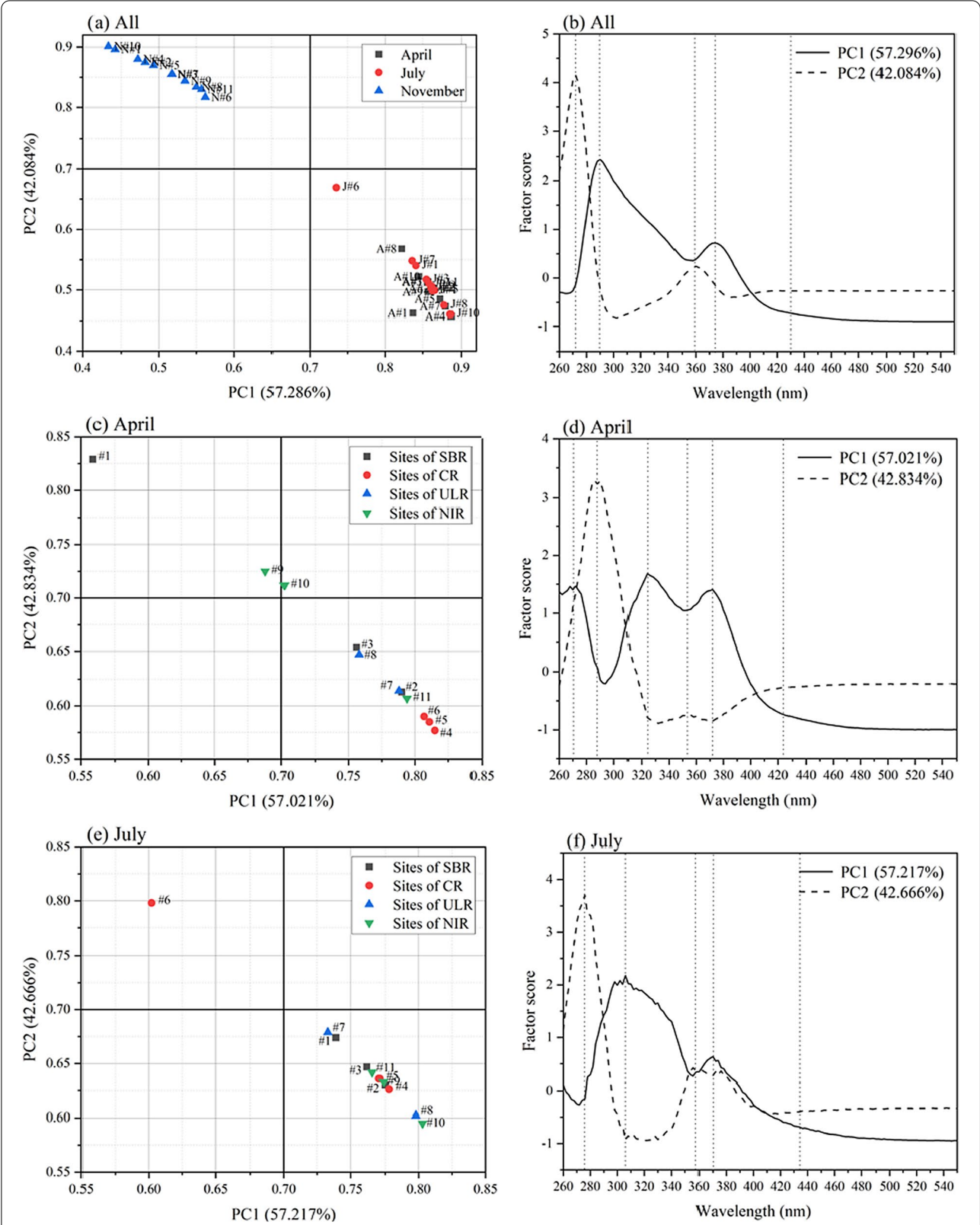


Fig. 4 Loadings plots of All-period (a), April (c), July (e), and November (g) and scores plots of All-period (b), April (d), July (f), and November (h) for spectral wavelengths in Lake Shahu

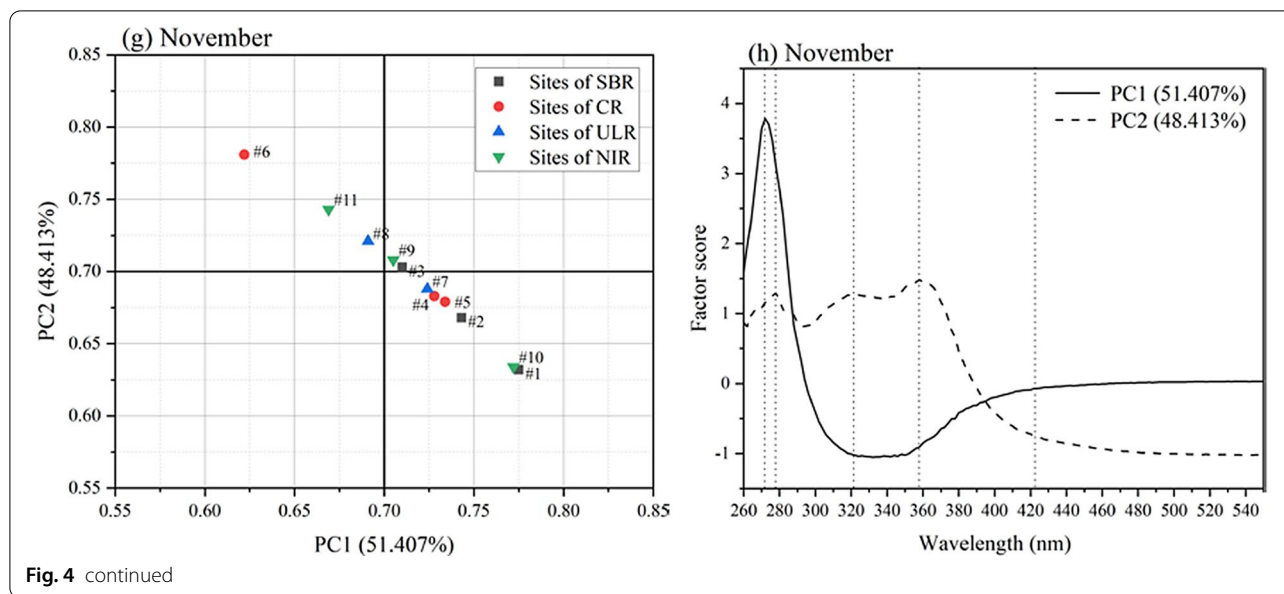


Fig. 4 continued

loading plots (Fig. 4a), the sampling sites of the entire study period were divided into two groups: group A, with the sites during November (water non-replenishment period), and group B, with the sites during April and July (water replenishment period), which could attribute that water of the Yellow River was continuously loaded into the lake as water replenishment from April to September. Group A had higher second (>0.8) and lower first (<0.6) PC loadings, indicating that the TYLF was dominant in the water non-replenishment period (Fig. 4b). Therefore, the metabolism of native organisms and decomposition of aquatic plants were the sources of pollution in November. Furthermore, in the water replenishment period (April and July), the TRLF dominated group B, because it presented higher first (>0.7) and lower second (<0.7) PC loadings (Fig. 4b). Thus, the most prominent pollution sources in April and July were anthropogenic activities, including agricultural cultivation, irrigation, domestic wastewater, and tourism.

The two PCs explained 99.856% of the total variances in the early water replenishment period (April), including 57.021% for PC1 and 42.834% for PC2. The score plots for the SFS wavelengths distinguished the fluorophores from DOM (Fig. 4d). PC1 showed three peaks with similar factor scores (Fig. 4d), which could be related to the components of TYLF, MHLE, and FLF. A prominent peak and two shoulders were obtained in PC2 (Fig. 4d). The prominent peak was associated with the TRLF. The first shoulder was referred to as the MHLE, presenting a blue-shift of 30 nm compared with PC1, and the second shoulder was involved in the HLE. With the exception of sites #1 and #9, all sites had a higher loading (>0.7) of

PC1 (Fig. 4c), illustrating that the TYLF, MHLE, and FLF were the preponderant components. This demonstrated that the pollution in April was mainly derived from the native organisms and contaminants accumulation during the freeze-up period. Sites #1, #9, and #10, especially site #1, with higher second PCA loadings, represented the characteristics of the replenishment period, which were dominated by the TRLF. This indirectly indicated that the sampling sites that were close to the shore, where are more easily to be affected by anthropogenic activities. The map of loading plots in the early water replenishment period showed a high degree of dispersion between the sampling sites (Fig. 4c), which could be associated with the differences in DOM components due to insufficient water replenishment. The preponderant component is the TYLF rather than the TRLF due to insufficient supply of water recharge.

Two PCs in July were extracted by PCA that explained 57.217 and 42.666% of the total variables, respectively. Interestingly, the trend of score plots in July (Fig. 4f) were similar to the full period (Fig. 4b), i.e., PC1 contained the TRLF and FLF and PC2 involved the TYLF and MHLE. Most of the sites with more than 0.7 of PC1 loading and less than 0.70 of PC2 loading (Fig. 4e) presented characteristics of the water replenishment period, i.e., the TRLF was the representative component, which was similar to the water replenishment period (Fig. 4b). This indirectly indicated that July was the primary period of water recharge. In addition, the results demonstrated that anthropogenic activities were the dominant source of pollution, during the tourist season and the intense agricultural irrigation season. Site #6, with more than 0.75 of

PC2 loading and less than 0.65 of PC1 loading, was dominated by the TYLF (Fig. 4e), which was consistent with its results in April. This indirectly indicated that there was consistently a better level of water quality at site #6 (Fig. 2). In addition, compared to April, the degree of dispersion decreased with a rise in the similarities of DOM components between the sampling sites, which could have been due to the sufficient water fluidity caused by water replenishment.

In November, PC1 (51.407% of the total variance) exhibited higher positive loadings (>0.7) at sites #1, #10, #2, #5, #4, #7, #3, and #9. This indicated that the component of these sites mainly was the TYLF (Fig. 6h), which was associated with biodegradation and decomposition of lake aquatic plants. PC2 (48.413% of the total variance) showed better positive loadings (>0.7) at sites #3, #9, #8, #11, and #6, indirectly verifying that the FLF was dominant, followed by the TRLF and MHLF (Fig. 6h). This could be associated with degradation of aquatic plants and algae. Based on the loading map, the degrees of dispersion in November were lower than those in April and higher than those in July, which could be attributed to the water replenishment.

Specially, the fluorophore characteristics of site #6 in July were similar to those during April, and the characteristics of site #6 in November were consistent with all the sampling sites, except for site #6, in July. This suggested that site #6 exhibited hysteresis in the variation of fluorophore constitution. In other words, there was consistently a better level of water quality at site #6 as indicated by the water quality parameters in “Parameters of water quality” Section.

Derivative fluorescence spectroscopy

The peak at around 280 nm was so strong that extensive overlaps at the bands of PLF and MHLF were found (Fig. 3), and the overlaps presented obstacles that hindered the discernment of fluorescence peaks. Although five fluorophore datasets were extracted by PCA, it could not identify the characteristics of each sample. Thus, in order to reduce extensive spectroscopic overlaps, the accurate wave band of each peak is sought, and the composition of DOM fractions at each site is investigated, and the derivative method was employed. The process of the derivative method is shown in Fig. 5. There were five peaks corresponding to five regions, which were defined as TYLF (I), TRLF (II), MHLF (III), FLF (IV), and HLF (V) (Fig. 6). The integrated area in each region was calculated to indicate relative content of corresponding fluorescence component [9, 36].

The mean values of total content of fluorescence components (TFC) increased in the order of July ($11,789.38 \pm 12,752.61$) < April

($12,667.58 \pm 15,246.91$) < November ($19,748.87 \pm 17,192.13$). A sharp rise of DOM content occurred in November, which indirectly indicated that the degree of organic pollution in the lake is the highest in November (Fig. 2). Among the five fluorescence components, the content of TYLF performed the most significant elevation (Fig. 7c, Additional file 1: Fig. S1), which was 17,771.18 in April, 16,640.09 in July, and 61,951.50 in November, indicating large amounts of algae-derived and plant biodegradation-derived DOM release into the lake during the non-replenishment period. The content of the PLF (the sum of TYLF and TRLF) among all sampling sites almost dominated the fluorescence components, whose average percentages among the DOM fractions were 40.55% in April, 37.09% in July, and 46.91% in November (Fig. 7b, d, f). And the TRLF had a larger share of PLF in April (68.84%) or July (65.78%), while the TYLF became the component with a larger share of PLF in November (61%). This indicated intense anthropogenic activities and release by-products of the lake organism in water replenishment period. The proportion of the MHLF decreased in the order of April (36.82%) > July (35.92%) > November (27%). The FLF and HLF showed a relatively consistent status and lower proportion in all periods, which varied from 15.64 (April) to 19.91% (November) and 6.09 (November) to 7.45% (July), respectively. From replenishment period to non-replenishment, the contents of the five components gradually changed from low in the middle and high around the lake to high throughout lake (Additional file 1: Fig. S1).

In April, the decreasing order of mean content value of each DOM fraction in the lake was MHLF (4665.63 ± 474.53) > TRLF (3569.28 ± 951.37) > HLF (1962.51 ± 106.07) > TYLF (1615.56 ± 258.56) > FLF (854.61 ± 75.42). The order of each DOM fraction in July was similar to April at 4231.01 ± 661.79 , 2907.56 ± 566.31 , 2258.48 ± 329.16 , 1512.74 ± 312.51 , and 879.60 ± 132.78 for the MHLF, TRLF, HLF, TYLF, and FLF, respectively. In the November, the TYLF showed the highest mean value (5631.96 ± 634.92) among the DOM fractions (Fig. 7e), followed by the MHLF (5318.51 ± 708.98), HLF (3920.16 ± 354.24), TRLF (3652.68 ± 407.31), and FLF (1225.57 ± 139.27), which approximately 3.48 times more than that in April (Fig. 7a) and about 3.72 times more than that in July (Fig. 7c). The decreasing order of the TRLF was November (3652.68 ± 407.31) > April (3569.28 ± 951.37) > July (2907.56 ± 566.3081), as was the MHLF. Compared to the HLF, a relatively obvious increasing order was obtained in FLF, i.e., April (1962.51 ± 106.07) < July (2258.48 ± 320.16) < November (3920.16 ± 354.24) (Additional file 1: Fig. S1). It could be associated with degradation of aquatic plants and algae.

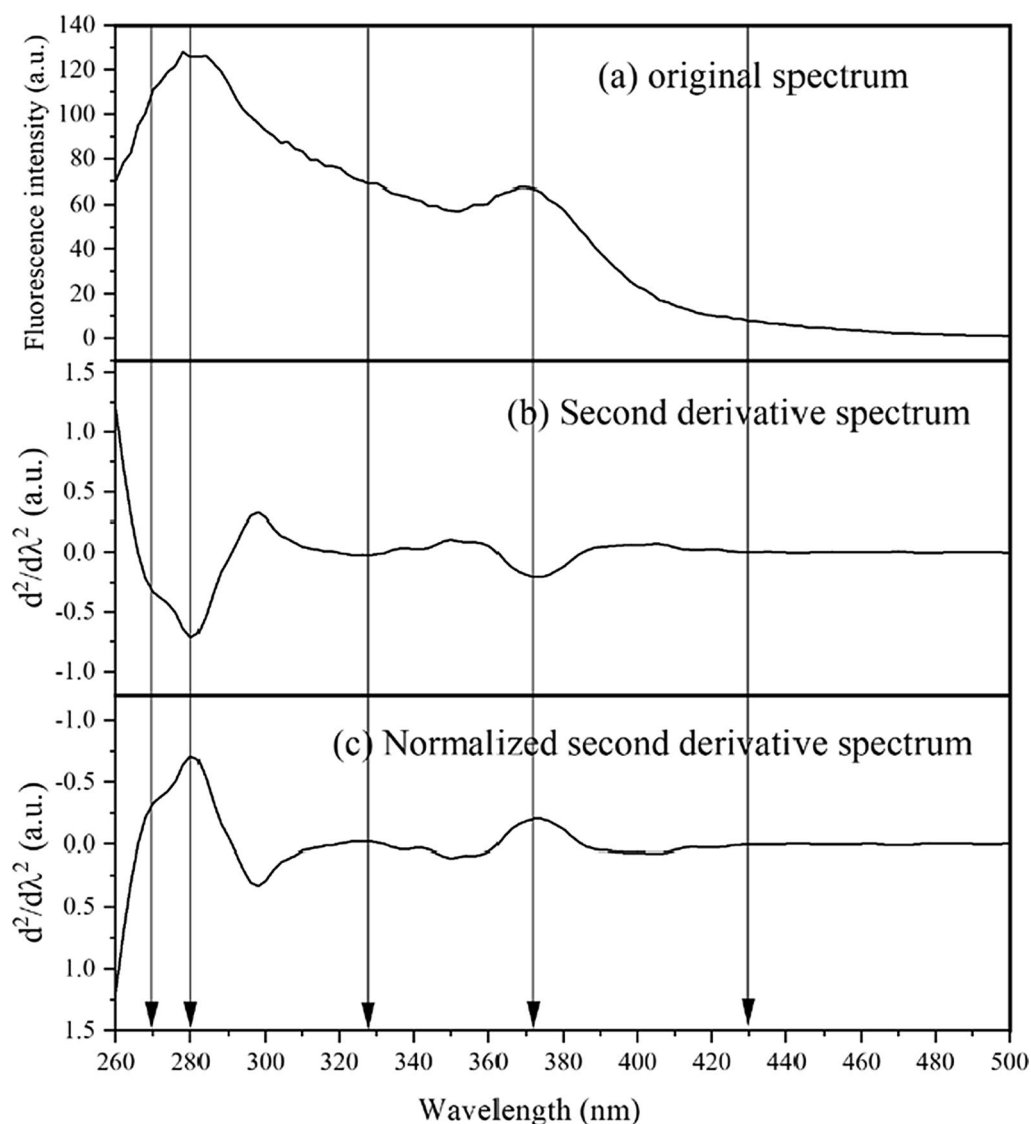


Fig. 5 Synchronous fluorescence spectroscopy (a), second derivative fluorescence spectroscopy (b), and normalized second derivative spectroscopy (c) of DOM at site #8 in April

In a word, DOM fractions were mainly derived from by-products of the lake organisms and anthropogenic activities during the replenishment period. In the non-replenishment period, the TYLF and FLF were derived from aquatic plants residue and biodegradation of the algae dominated DOM fractions, which resulted in an increase in DOM contents and an increase in degree of organic pollution (Fig. 2, 4h, Additional file 1: Fig. S1).

Canonical correlation analysis

The CCA could be applied to visualize the comprehensive correlations between water quality, environmental factors, and sampling sites, and to identify the potential

factors of the sites [25, 37]. The results of CCA visualized by CANOCO 5, the longer arrow indicates the greater influence of the factor, the smaller angle between the arrow and the coordinate axis implicates the higher correlation, and the smaller distance between the sampling sites and the arrow presents the stronger effect of the factor on the sampling sites. In the CCA ordination biplot of April, the TYLF, TRLF, MHLF, SD, and BOD₅ were relative to organic matter pollution (Fig. 8a). The arrows of the TYLF, TRLF, and MHLF pointed toward the positive direction of AX1 with small angles (<45°) and much longer arrows, which illustrated that the positive half of AX1 could be related to the organic matter from water

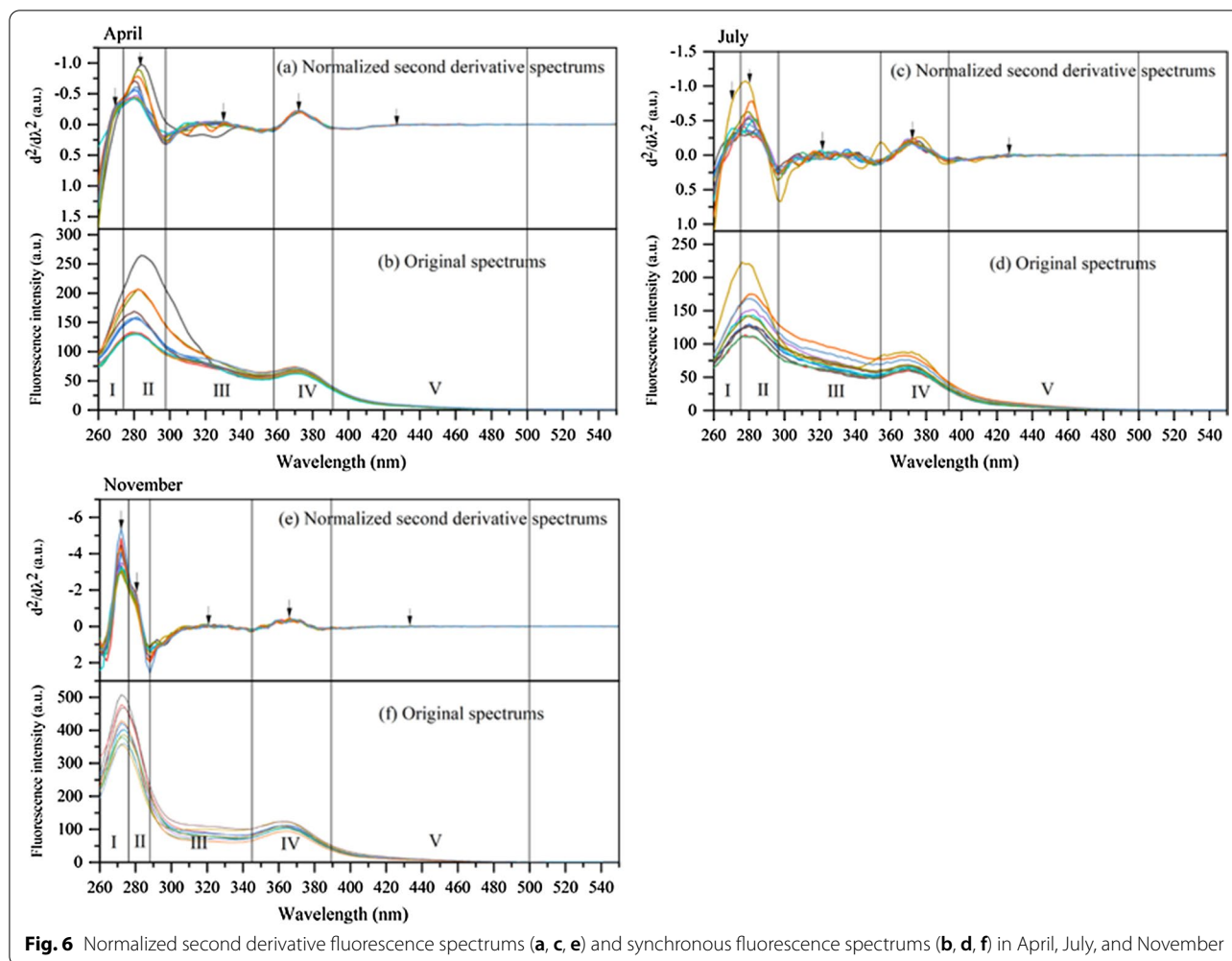


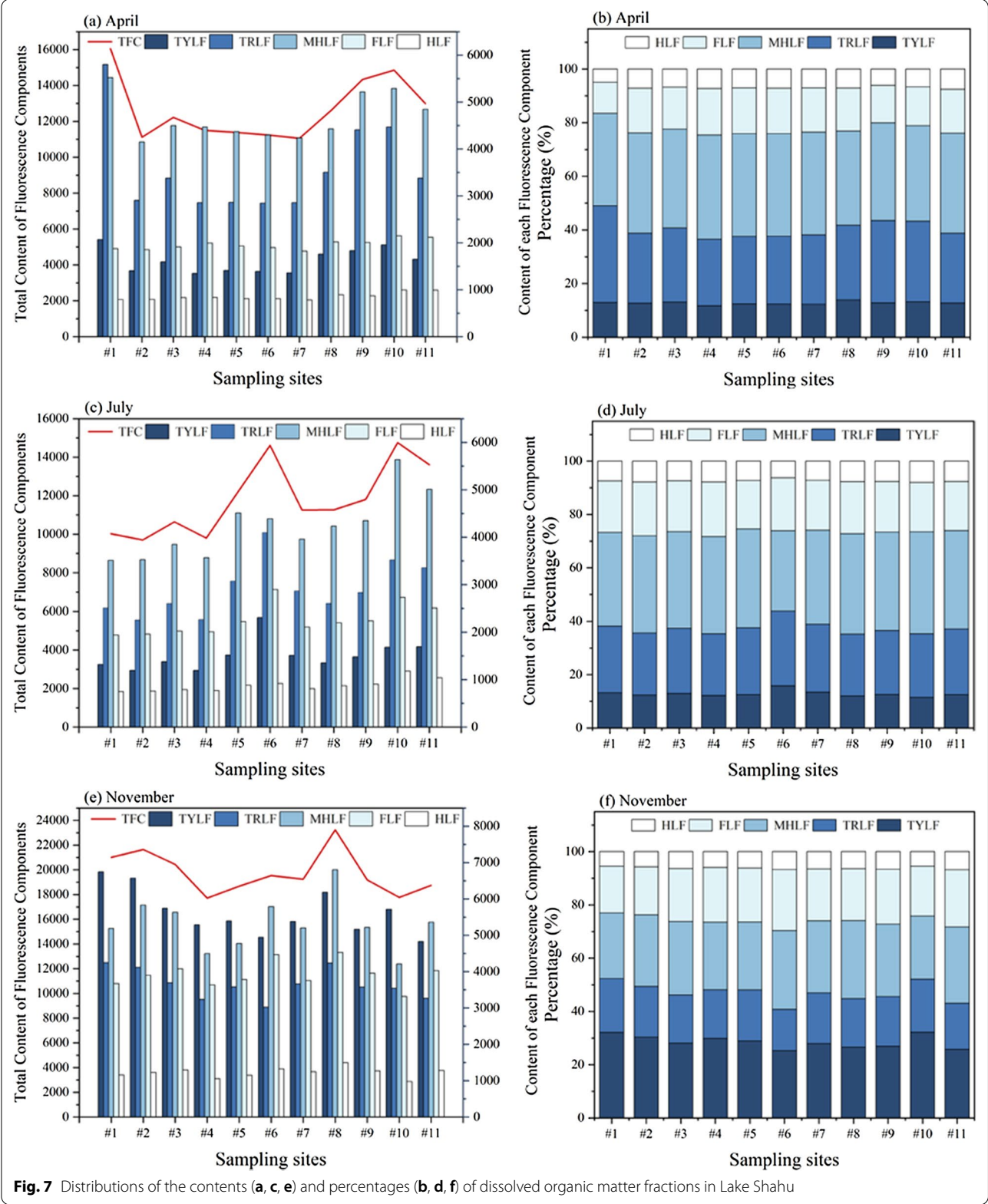
Fig. 6 Normalized second derivative fluorescence spectrums (a, c, e) and synchronous fluorescence spectrums (b, d, f) in April, July, and November

recharge. Sites #1, #3, #9, and #10, with positive loadings in AX1 (Fig. 8a), were affected by organic matter pollution. This indirectly indicated that these sites had much greater TYLF, TRLF, and MHLF content (Fig. 7a, b). The BOD₅ was the potential factor of the sites located in the negative region of AX1. Moreover, in the positive region of AX1, the loadings decreased in the order of #1 > #9 > #3 > #10 > #8, which was practically consistent with the PCA result (Fig. 4c).

As shown in Fig. 8b, in July, the potential factors included TRLF, TYLF, FLE, Chl-a, TP, COD_{Cr}, and DO. This indicated that an amount of recharge water with a high TP, COD_{Cr}, and TRLF load entered the lake resulted in algae and aquatic plants blooms. Site #6 obtained highest positive loading of AX2, followed by #5, #11, #7, and #10, which was associated with the potential factors of TYLF, TRLF, and FLE. This indirectly proved that the highest contents of TYLF, TRLF, and FLE were present at site #6 (Fig. 7c, d). The potential factors of site #9 were Chl-a, TP, and COD_{Cr}, which illustrated that the higher

values of those parameters at site #9 in July (Fig. 2e, f, i). The arrows of MHLF, HLF, Chl-a, and TP with smaller angles were the potential factors of the positive half of AX1. Moreover, the arrows of DO, SD, and TEMP pointed toward the negative direction of AX1, which represented the potential factors of sites #3, #4, #5, #7, and #8. The negative half of AX1 may be concerned with natural pollution.

The potential factors in November were COD_{Cr}, TYLF, TRLF, SD, TN, and FLF (Fig. 8c), representing that residue of the algae and aquatic plants have been deeply degraded into TRLF, TYLF, and FLF during the non-replenishment period. As we expected, the sampling sites #1, #2, and #10, with the highest loadings in PC1 during November (Fig. 4g), were extracted by CCA (Fig. 8c), proving that the leading potential factor of the positive half of AX1 was TYLF. And FLF was the main potential factor of negative half of AX1, the sites #5, #7, #8, and #9 were deeply affected by it (Fig. 8c). On the positive half of AX2, COD_{Cr} was the unique potential factor, associated with sites #4, #1, #11,



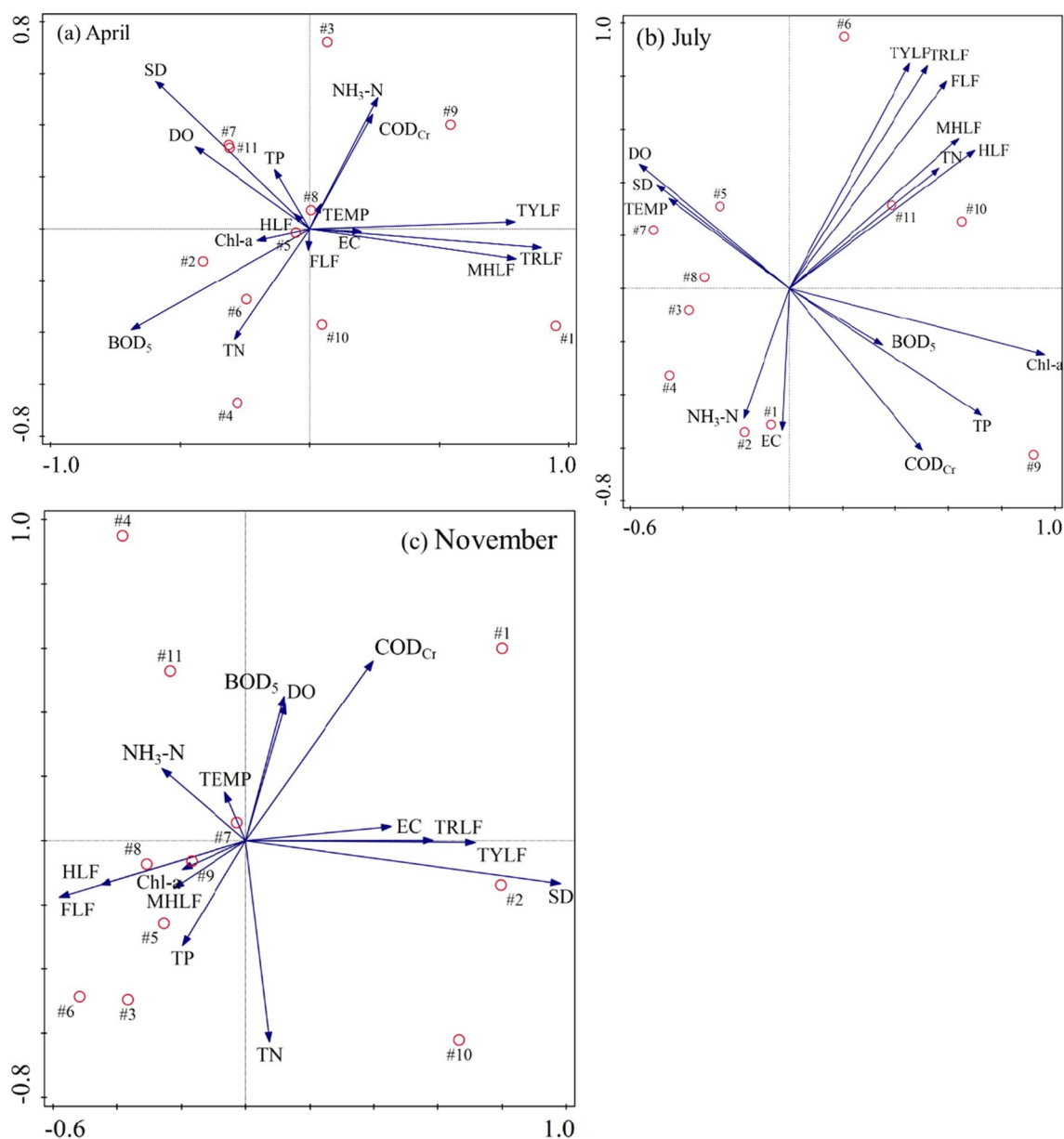


Fig. 8 The ordination biplot based on the canonical correlation analysis of the relationship among the water quality parameters and dissolved organic matter fractions of the sampling sites in April (a), July (b), and November (c)

and #7 (Fig. 8c). The smaller angle was obtained between the negative half of AX2 and the TN arrow, indicating that the TN was the potential factor. Therefore, sites #10, #3, and #6, with higher negative loadings, presented higher values of TN in November (Fig. 2g).

Conclusion

The water quality level in the lake was the highest in April, followed by July and November. The water quality of the central region in the lake was better than those of

other regions in the above-mentioned months. DOM from the lake contained five fluorescence components: TYLF, TRLF, MHLF, FLF, and HLF, which in November were distinguished from April and July. It is attributed to the water of the Yellow River being continuously loaded into the lake as water replenishment from April to September. The TRLF dominated in DOM fractions during the period of water replenishment, while TYLF was the predominant component during the period of water non-replenishment. The potential factors

included TYLF, TRLF, MHLF, SD and BOD₅ in April, which were associated with organic matter pollution. The potential factors contained TYLF, TRLF, FLF, Chl-a, TP, COD_{Cr} and DO in July, suggesting the enrichment of TP lead algae and plants growth. The potential factors in November consisted of TYLF, TRLF, COD_{Cr}, SD, TN and FLF, representing residue of the algae and plants have been deeply degraded. The replenishment water was the main factor controlling the water level as well as potential factor effecting the variations of DOM fractions, resulting in fluctuations of the water level. According to the results of this study, monitoring and controlling of nutrients input from replenishment water, as well as removing of overgrown aquatic plants from the lake during non-replenishment period, can be one of the methods to control pollution of Lake Shahu and brackish-water lakes or other types of waterbodies with similar pollution status.

Abbreviations

DOM: Dissolved organic matter; SFS: Synchronous fluorescence spectroscopy; PCA: Principal component analysis; SDSFS: Second derivative synchronous fluorescence spectra; CCA: Canonical correlation analysis; PLF: Protein-like fluorescence; TYLF: Tyrosine-like fluorescence; TRLF: Tryptophan-like fluorescence; MHLF: Microbial humic-like fluorescence; FLF: Fulvic-like fluorescence; HLF: Humic-like fluorescence; EEMs: Excitation–emission matrix spectroscopy; TEMP: Temperature; EC: Electrical conductivity; DO: Dissolved oxygen; SD: Secchi depth; COD_{Cr}: Chemical oxygen demand; TP: Total phosphorus; TN: Total nitrogen; NH₃-N: Ammonia nitrogen; Chl-a: Chlorophyll a; BOD₅: Biochemical oxygen demand; KMO: Kaiser–Meyer–Olkin; PCs: Principal components; TFC: The total content of fluorescence components.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12302-021-00538-3>.

Additional file 1: Fig. S1. Spatio-temporal distributions of contents of fluorescence components in Lake Shahu.

Acknowledgements

We would like to express our sincere thanks to the anonymous reviewers. Their insightful comments were helpful for improving the manuscript.

Authors' contributions

KL: data curation and analysis, visualization, writing—original draft. WX: investigation, writing—review and editing. HY: methodology, writing—review and editing. HG: supervision, writing—review and editing. XG: supervision, project administration, funding acquisition. NZ: investigation, data analysis. All authors read and approved the final manuscript.

Funding

This work was financially supported by the Ningxia Hui Autonomous Region Science and Technology Research Project for Environmental Protection (Grant No. 2019-04, P.R. China).

Availability of data and materials

The datasets obtained and analyzed in the research is available from the first author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interest.

Author details

¹State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, 100012 Beijing, People's Republic of China. ²Ningxia Research Academy of Environmental Sciences (Limited Liability Company), Yinchuan 750021, People's Republic of China. ³College of Geography and Environment, Shandong Normal University, Jinan, Shandong 20358, People's Republic of China.

Received: 18 May 2021 Accepted: 31 July 2021

Published online: 13 August 2021

References

- Xu JL, Tan WF, Xiong J, Wang MX, Fang LC, Koopal LK (2016) Copper binding to soil fulvic and humic acids: NICA–Donnan modeling and conditional affinity spectra. *J Colloid Interface Sci* 473:141–151
- Guo XJ, Yuan DH, Jiang JY, Zhang H, Deng Y (2013) Detection of dissolved organic matter in saline-alkali soils using synchronous fluorescence spectroscopy and principal component analysis. *Spectrochim Acta A Mol Biomol Spectrosc* 104:280–286
- Cilenti A, Provenzano MR, Senesi N (2005) Characterization of dissolved organic matter from saline soils by fluorescence spectroscopy. *Environ Chem Lett* 3(2):53–56
- Leenheer JA, Croue JP (2003) Characterizing aquatic dissolved organic matter. *Environ Sci Technol* 37(1):18A–26A
- Daoud AB, Tremblay L (2019) HPLC–SEC–FTIR characterization of the dissolved organic matter produced by the microbial carbon pump. *Mar Chem*. <https://doi.org/10.1016/j.marchem.2019.103668>
- Zhang F, Li X, Duan L, Zhang H, Gu W, Yang X, Li J, He S, Yu J, Ren M (2021) Effect of different DOM components on arsenate complexation in natural water. *Environ Pollut* 270:116221
- Guo XJ, He XS, Li CW, Li NX (2019) The binding properties of copper and lead onto compost-derived DOM using Fourier-transform infrared, UV-vis and fluorescence spectra combined with two-dimensional correlation analysis. *J Hazard Mater* 365:457–466
- Landry C, Tremblay L (2012) Compositional differences between size classes of dissolved organic matter from freshwater and seawater revealed by an HPLC–FTIR system. *Environ Sci Technol* 46(3):1700–1707
- Yu H, Song Y, Tu X, Du E, Liu R, Peng J (2013) Assessing removal efficiency of dissolved organic matter in wastewater treatment using fluorescence excitation emission matrices with parallel factor analysis and second derivative synchronous fluorescence. *Bioresour Technol* 144:595–601
- Catalan N, Pastor A, Borrego CM, Casas-Ruiz JP, Hawkes JA, Gutierrez C, von Schiller D, Marce R (2021) The relevance of environment vs. composition on dissolved organic matter degradation in freshwaters. *Limnol Oceanogr* 66(2):306–320
- Yang CH, Liu YZ, Sun XM, Miao SC, Guo YM, Li TJ (2019) Characterization of fluorescent dissolved organic matter from green macroalgae (*Ulva prolifera*)-derived biochar by excitation–emission matrix combined with parallel factor and self-organizing maps analyses. *Bioresour Technol* 287:6
- Zhi EQ, Yu HB, Duan L, Han L, Liu L, Song YH (2015) Characterization of the composition of water DOM in a surface flow constructed wetland using fluorescence spectroscopy coupled with derivative and PARAFAC. *Environ Earth Sci* 73(9):5153–5161
- Patra D, Mishra AK (2002) Recent developments in multi-component synchronous fluorescence scan analysis. *Trends Anal Chem* 21(12):787–798

14. Barker JD, Sharp MJ, Turner RJ (2009) Using synchronous fluorescence spectroscopy and principal components analysis to monitor dissolved organic matter dynamics in a glacier system. *Hydrol Process* 23(10):1487–1500
15. Chen YY, Zheng ZP, Yang F, Bai Y, Yu HB (2020) The composition and structure of dissolved organic matter in saline soil were studied by synchronous fluorescence spectroscopy combined with principal components and two-dimensional correlation. *Spectrosc Spect Anal* 40(2):489–493
16. Yu H, Song Y, Xi B, Zhang M, He X (2012) Application of derivative synchronous fluorescence spectroscopy (DSFS) to indicate salinisation processes of saline soils in semi-arid region. *Ecol Indic* 18:532–539
17. Wu J, Xue C, Tian R, Wang S (2017) Lake water quality assessment: a case study of Shahu Lake in the semiarid loess area of northwest China. *Environ Earth Sci*. <https://doi.org/10.1007/s12665-017-6516-x>
18. Tian L, Zhu X, Wang L, Du P, Peng F, Pang Q (2020) Long-term trends in water quality and influence of water recharge and climate on the water quality of brackish-water lakes: a case study of Shahu Lake. *J Environ Manage* 276:111290
19. Ren L, Rabalais NN, Turner RE, Morrison W, Mendenhall W (2009) Nutrient Limitation on Phytoplankton Growth in the Upper Barataria Basin, Louisiana: Microcosm Bioassays. *Estuaries Coasts* 32(5):958–974
20. Daggett CT, Saros JE, Lafrancois BM, Simon KS, Amirbahman A (2015) Effects of increased concentrations of inorganic nitrogen and dissolved organic matter on phytoplankton in boreal lakes with differing nutrient limitation patterns. *Aquat Sci* 77(3):511–521
21. Li J-Y, Zhang Y-F, Yang Z, Wang M (2017) Bacterial diversity in Shahu lake, northwest China is significantly affected by nutrient composition rather than location. *Ann Microbiol* 67(7):469–478
22. Li P, Feng W, Xue C, Tian R, Wang S (2016) Spatiotemporal Variability of contaminants in lake water and their risks to human health: a case study of the Shahu Lake Tourist Area, Northwest China. *Expo Health* 9(3):213–225
23. Qian H, Wu JH, Zhou YH, Li PY (2014) Stable oxygen and hydrogen isotopes as indicators of lake water recharge and evaporation in the lakes of the Yinchuan Plain. *Hydrol Process* 28(10):3554–3562
24. Chen J, Qian H, Gao Y, Wang H, Zhang M (2020) Insights into hydrological and hydrochemical processes in response to water replenishment for lakes in arid regions. *J Hydrol*. <https://doi.org/10.1016/j.jhydrol.2019.124386>
25. He W, Lee JH, Hur J (2016) Anthropogenic signature of sediment organic matter probed by UV-Visible and fluorescence spectroscopy and the association with heavy metal enrichment. *Chemosphere* 150:184–193
26. Guo X, Yu H, Yan Z, Gao H, Zhang Y (2018) Tracking variations of fluorescent dissolved organic matter during wastewater treatment by accumulative fluorescence emission spectroscopy combined with principal component, second derivative and canonical correlation analyses. *Chemosphere* 194:463–470
27. Hur J, Jung KY, Jung YM (2011) Characterization of spectral responses of humic substances upon UV irradiation using two-dimensional correlation spectroscopy. *Water Res* 45(9):2965–2974
28. Mangaliri KP, Timko SA, Gonsior M, Blaney L (2017) PARAFAC modeling of irradiation- and oxidation-induced changes in fluorescent dissolved organic matter extracted from poultry litter. *Environ Sci Technol* 51(14):8036–8047
29. Yin HL, Wang Y, Yang Y, Huang JS, Xu ZX (2020) Tryptophan-like fluorescence as a fingerprint of dry-weather misconnections into storm drainage system. *Environ Sci Eur*. <https://doi.org/10.1186/s12302-020-00336-3>
30. Zhuang WE, Chen W, Cheng Q, Yang LY (2021) Assessing the priming effect of dissolved organic matter from typical sources using fluorescence EEMs-PARAFAC. *Chemosphere* 264:10
31. Zhang YL, Liu XH, Wang MZ, Qin BQ (2013) Compositional differences of chromophoric dissolved organic matter derived from phytoplankton and macrophytes. *Org Geochem* 55:26–37
32. Wang SH, Wang WW, Chen JY, Zhang B, Zhao L, Jiang X (2020) Characteristics of dissolved organic matter and its role in lake eutrophication at the early stage of algal blooms—a case study of Lake Taihu, China. *Water* 12(8):17
33. Yu HB, Xi BD, Ma WC, Li DL, He XS (2011) Fluorescence spectroscopic properties of dissolved fulvic acids from salined fluvo-aquic soils around wuliangshuai in hetao irrigation District, China. *Soil Sci Soc Am J* 75(4):1385–1393
34. Kida M, Kojima T, Tanabe Y, Hayashi K, Kudoh S, Maie N, Fujitake N (2019) Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams. *Water Res*. <https://doi.org/10.1016/j.watres.2019.114901>
35. Varol M (2020) Use of water quality index and multivariate statistical methods for the evaluation of water quality of a stream affected by multiple stressors: a case study. *Environ Pollut* 266(Pt 3):115417
36. Liu D, Yu H, Gao H, Feng H, Zhang G (2021) Applying synchronous fluorescence and UV-vis spectra combined with two-dimensional correlation to characterize structural composition of DOM from urban black and stinky rivers. *Environ Sci Pollut Res Int* 28(15):19400–19411
37. Wei H, Yu H, Zhang G, Pan H, Lv C, Meng F (2018) Revealing the correlations between heavy metals and water quality, with insight into the potential factors and variations through canonical correlation analysis in an upstream tributary. *Ecol Indic* 90:485–493

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)