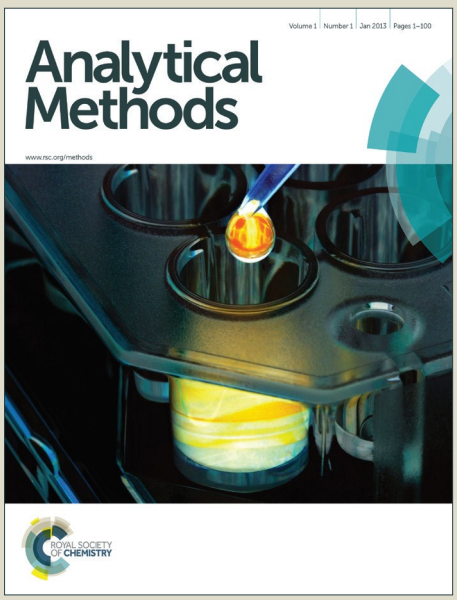
Analytical Methods

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

New insight into DOC and DON in a drinking water Biological Aerated Filter (BAF) by multimethod and correlation analysis of 3D-EEM

Jia Kang[a], Tengfei Ma[a], Qiuhong Zhou[a], Xu Gao[b,*], Youpeng Chen[c], Jianming Wu[a], Jiahao Chen[a], Yu Xiang[a]

[a] Faculty of Urban Construction and Environmental Engineering, Chongqing University, Chongqing 400045, China

[b] Chongqing Water Group Co., Ltd., Chongqing 400045, China.

[c] Key Laboratory of Reservoir Aquatic Environment of CAS, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China

[*]Corresponding author

Prof. Xu Gao

Fax: +86–23–63860939; Tel: +86–23–63860805;

E–mail: hughgao@outlook.com

Abstract: To get insight into the components and variations of dissolved organic carbon (DOC) and dissolved organic nitrogen (DON) in a lab–scale drinking water biological aerated filter (BAF), the concentrations and three–dimensional excitation–emission matrix (3D–EEM) spectrum of DOC and DON were determined and analyzed, peak identification, fluorescence regional integration (FRI) analysis and parallel factor (PARAFAC) model were applied to analyze the 3D–EEM spectrum. Fluorescent DOC in the BAF mainly existed in the form of humic acids and fulvic acids, fluorescent DON mainly existed in tryptophan protein form. Protein–like substances accounted for more than 60% of fluorescent organic matters through FRI analysis. Tyrosine and tryptophan–like proteins, and fulvic acid–like substances were effectively removed through BAF process, while humic acids kept almost unchanged. The results of peak identification, FRI analysis, and PARAFAC model were consistent with each other, especially for FRI and PARAFAC, it is manifest that FRI technique and PARAFAC model are effective tools for 3D–EEM spectrum analysis, FRI technique proved to be more suitable to characterize DOC and DON in this study than peak identification and PARAFAC model.

Key words: dissolved organic carbon; dissolved organic nitrogen; three–dimensional excitation–emission matrix; fluorescence regional integration; parallel factor model

1. Introduction

Biological aerated filter (BAF) is a widely applied process for micro–polluted source water treatment because of its effective removal of dissolved organic matter (DOM) [1-3]. The exploration of dissolved organic carbon (DOC) and dissolved organic nitrogen (DON) (major components of DOM) in the BAF is conducive to optimize process operation and improve removal efficiency. Fluorescence three–dimensional excitation–emission matrix (3D–EEM) spectroscopy, which is based on the presence of fluorophores about humic acids, fulvic acids and protein–like compounds, is a rapid characterization tool with high sensitivity and selectivity towards fluorescent DOM, could offer insights into compositions, variations, and characteristics of DOC and DON in the BAF[4-6].

The present studies on fluorescence 3D–EEM spectrum commonly adopt visual methods such as peak picking[7]and fluorescence index[8-10]to analyze the complex spectrum information qualitatively and interpret the variation of DOM. These visual identifications lack a quantitative analysis of fluorescence 3D–EEM spectrum and could not capture the heterogeneity of DOC and DON.

Fluorescence regional integration (FRI) is a quantitative analytical approach of 3D–EEM spectrum proposed by Chen[11]. FRI defines the fluorescence spectrum into five regions that are related to tyrosine protein–like fraction, tryptophan protein–like fraction, fulvic acids–like fraction, humic acids–like fraction and soluble microbial products (SMPs)–like fraction. By integrating volume beneath

each region, quantitative analysis of 3D–EEM spectrum is achieved and the composition and proportion of DOC and DON are determined. The quantitative assessment on the configuration and heterogeneity of DOC and DON makes FRI a valuable research tool for 3D–EEM. The overlapped spectra emitted by fluorophores in the 3D–EEM spectrum always prevent the comprehensive and accurate characterization of DOC and DON. Parallel factor (PARAFAC) analysis overcomes this problem through decomposing the complex 3D–EEM into independent but spectrally overlapping fluorescence trilinear components[12]. Thus, individual component is extracted and attributed to corresponding substance such as protein–like substances, fulvic acids–like, or humic acids–like substances[13,14].

A limited number of studies have reported the application of FRI and PARAFAC approaches for fluorescence 3D–EEM spectrum analysis in drinking water treatment processes[15]. It is just found that both PARAFAC loadings and FRI volumes exhibit agreements with the total overall fluorescence intensity on a long–term monitoring before and after coagulation–filtration treatment in a drinking water treatment plant[6].

In this paper, peak identification, FRI analysis and PARAFAC model were adopted to analyze the fluorescence 3D–EEM spectrum in order to explore the compositions and variations of DOC and DON in a lab–scale BAF reactor. Correlations and comparative analysis of the three 3D–EEM spectrum analysis approaches were also conducted to clarify the methods' characteristics and

differences, and to determine the applicable condition of each approach, this methodology could provide reference for the reader to select appropriate method for fluorescence 3D-EEM spectra analysis.

2. Materials and methods

2.1 Reactor setup

All the experiments in this study were carried out based on an upward flow lab–scale BAF made of plexiglass with an inner diameter of 50 mm, a height of 1.4 m, and an effective working volume of 12.17 L (Fig. 1). The supporting layer was filled with pebbles, and ceramsite with diameter of 2–5 mm was selected as filter media and filled in 800 mm high filter layer. Six sampling outlets equally distributed along the filter layer at 160 mm intervals for effluents and filter media collection, the interval was determined according to Yu[16].

The BAF reactor was fed with synthetic micro–polluted source water with DOC at around 20 mg/L, NH_4^+–N concentration at around 10 mg/L. The exact composition of synthetic water was as follows: CH_3COONa, 62mg/L; NH_4Cl, 38 mg/L; KH_2PO_4, 0.44 mg/L; K_2HPO_4, 0.56 mg/L; $CaCl_2$, 0.67 mg/L; $MgSO_4$, 0.67 mg/L; $FeCl_3$, 0.42 mg/L; $CuSO_4$, 0.0083 mg/L; KI, 0.05 mg/L; $MnSO_4$, 0.0333 mg/L; $ZnSO_4$, 0.0333 mg/L; $CoCl_2$, 0.0417 mg/L; $NaMoO_4$, 0.0067 mg/L. All the reagents used were of analytical grade. The influent pH was adjusted to 7-8 by dosing $NaHCO_3$ and monitored by a pH meter.

A peristaltic pump (BT–100EA/153Yx×2PPS, Jieheng, China) was applied to

maintain the filtration velocity of BAF reactor at around 2 m/h, and the gas–water ratio was kept in the range of 1:1–2:1 by an air compressor (OLF–2530, Jiebao, China). Backwashing was performed every three or four days to prevent BAF from clogging and to maintain biofilm activity. The BAF reactor had run stably for three months before sampling to guarantee the reactor was in a steady state throughout the experiment.

2.2 Sampling

Sampling program was carried out for 7 times from May 21th, 2014 to June 17th, 2014. Water samples were collected in brown glass bottles on the third day after backwashing. 7 samples, including influent, and samples from six different media depths, were obtained once. The locations of sampling outlets in Fig. 1 showed that there existed no filter media above the depth of 80 cm, thus water samples from the depth of 80 cm were effluents. The collected water samples were filtered with pre–washed 0.45 μm cellulose acetate membranes (Anpel Co. Ltd., China), preserved in 4 ℃, and measured within 24 h.

2.3 Analytical methods

DOC was measured using an Elementar Vario TOC analyzer (high–temperature combustion at 850 ℃, nondispersive infrared detection; Elementar Analysensysteme Gmbh, Germany). Total dissolved nitrogen (TDN) was measured using an alkaline potassium persulfate digestion UV spectrophotometry; NO_3^-–N was measured using a UV spectrophotometry; NO_2^-–N was measured

using an N–(1–naphthyl) –ethylenediamine photometric method; and NH_4^+–N was measured using a salicylate–hypochlorite spectrophotometry. These experiments were carried out according to the Chinese National Standard Methods (SEPA of China, 2002). DON was quantified as the difference between TDN and dissolved inorganic nitrogen: DON = TDN − NO_3^- − NO_2^- − NH_3/NH_4^+.

2.4 Oxygen uptake rate determination

The oxygen uptake rates (OURs) of ammonia–oxidizing bacteria (AOB), nitrite–oxidizing bacteria (NOB), and heterotrophic bacteria were determined using a respirometric method proposed by Surmaczgorska[17]and Urfer[18]. The media samples from BAF reactor was wetly weighted, and put into a closed batch respirometer, then fill the respirometer up with dechlorinated pure water. Dissolved oxygen (DO) was measured with a DO probe. Detailed measurement procedure is as follows. First, the total OUR is determined. When the DO concentration has decrease by about 3 mg/L, $NaClO_3$, which is a selective inhibitor of Nitrobacter, is added to the mixed liquor sample (final $NaClO_3$ concentration is 20 mmol/L) and the OUR is determined. The difference between the total OUR and the OUR measured in the presence of $NaClO_3$ is considered as the oxygen uptake due to NOB. Finally, after the DO has decreased by another 2 mg/L, allylthiourea (ATU) is added to the mixed liquor sample (final concentration is 5 mg/L) to inhibit the activity of NOB and the remaining OUR is measured. The difference between the OUR with $NaClO_3$ and both inhibitors,

$NaClO_3$ and ATU, represents the oxygen uptake due to AOB. The OUR measured in presence of both inhibitors reflects the oxygen consumption of the heterotrophs.

2.5 Fluorescence 3D–EEM spectroscopy

Fluorescence 3D–EEM spectra of water samples on May 30th, June 5th and June 8th were determined. The 3D–EEM spectrum of influent on May 30th was failed to get for insufficient volume of water sample. Considering that the BAF reactor was fed with synthetic water, thus influent water quality was stable and the fluorescence spectra of influent on June 5th and June 8th were representative, 3D–EEM fluorescence spectra of effluent water samples on May 30th were still valid. The total number of water samples for 3D–EEM spectrum measurement is 20.

Fluorescence 3D–EEM spectrum was measured using an F–7000 fluorescence spectrophotometer (Hitachi, Japan). The excitation and emission slits were both set to a 5 nm band–pass, the scanning speed was set at 60000 nm/ min, and the photomultiplier detector voltage was fixed at 700 V. The emission wavelengths were increased from 220 to 600 nm at 5 nm steps by varying the excitation wavelengths incrementally from 200 to 450 nm at 5 nm steps.

Raman and Raleigh scatter were removed for FRI technique and PARAFAC model analysis by subtracting the response of fluorometer to a blank solution from the 3D–EEM spectra of water samples. The blank solution was prepared from Super–Q water with a resistivity of 18.25 Ω·cm.

2.6 FRI Data processing

FRI technique proposed by Chen[11] could identify and characterize 3D–EEM spectrum quantitatively. The FRI technique was developed to divided fluorescence 3D–EEM spectrum into five regions, each region represents specific component of DOM (Table 1).

The volume ($\emptyset_i$) beneath region "i" of the 3D–EEM spectrum was used to represent the cumulative fluorescence response of DOM with similar properties, and can be calculated by a Riemann sum algorithm as follows:

$$\emptyset_i = \sum_{ex}\sum_{em} I(\lambda_{ex}\lambda_{em})\Delta\lambda_{ex}\Delta\lambda_{em} \quad (1)$$

where $\Delta\lambda_{ex}$ is the excitation wavelength interval (5 nm in this study), $\Delta\lambda_{em}$ is the emission wavelength interval (5 nm in this study), and $I(\lambda_{ex}\lambda_{em})$ is the fluorescence intensity (AU) at each excitation–emission wavelength pair. The total volume ($\emptyset_T$) beneath the 3D–EEM spectrum was calculated as:

$$\emptyset_T = \sum \emptyset_i \quad (2)$$

$\emptyset_T$ and $\emptyset_i$ values were normalized to a DOC concentration of 1 mg/L as follows for comparison of 3D–EEMs from different sources:

$$\emptyset_{i,n} = MF_i\emptyset_i \quad (3)$$

MF_i is a multiplication factor for each region, equal to the inverse of fractional projected excitation–emission area. The percent fluorescence response in a specific region (P_i) was calculated as:

$$P_{i,n} = \Phi_{i,n}/\Phi_{T,n} \times 100\% \quad (4)$$

2.7 PARAFAC model

The 3D–EEM spectrum dataset of 20 water samples obtained from the BAF reactor was modeled with PARAFAC, which uses an alternating least squares algorithm to minimize the sum of squared residuals in a trilinear model, and reduces a 3D–EEM dataset into a set of trilinear terms and a residual array[13]:

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}; \quad i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K \quad (5)$$

where x_{ijk} is the fluorescence intensity of the ith sample at the kth excitation and jth emission wavelength; a_{if} is directly proportional to the concentration of the fth fluorophore in the ith sample (defined as scores); b_{jf} and c_{kf} are estimates of the emission and excitation spectra of the fth fluorophore at wavelength j and k, respectively (defined as loadings); F is the number of fluorophores (components); ε_{ijk} is the residual matrix representing the unexplained variation in the model[14].

PARAFAC models were conducted with N–way v.3.20 Toolbox (http://www.models.life.ku.dk/nwaytoolbox/download) based on MATLAB software. The raw fluorescence 3D–EEM spectra were corrected before analysis by subtracting the fluorescence 3D–EEM spectrum of Super–Q water, then setting the 3D–EEM data close to Rayleigh scattering line as zero to eliminate the interference of Rayleigh scattering on PARAFAC analysis. Outlier samples were found by leverage comparison. Split half analysis[14] was used to determine the appropriate number of components for PARAFAC model analysis.

2.8 Statistical analysis

Pearson correlation was carried out using the software SPSS version 13.0 for Windows (SPSS, Chicago, IL).

3. Results and discussion

3.1 BAF reactor performance

The profiles of DOC and DON concentrations in the BAF are illustrated in Fig. 2(a). DOC in the BAF presented a significant decline first, and then gradually stabilized; this result was consistent with the DOC variation in the biofilter of relevant literatures[19,20]. At the media depth of 32 cm, DOC decreased to 7.60 mg/L from 20.41 mg/L in the influent, the DOC removal efficiency at depth of 32 cm was 62.78%. Above the depth of 32 cm, DOC concentration became stable, and the removal efficiency was not significant. DOC concentration at the depth of 80 cm was 6.62 mg/L, and the removal efficiency was 67.58%. In general, the removal of DOC in the BAF mainly depended on the biomass below media depth of 32 cm. The DOC removal efficiency upon BAF process was 67.58%, thus DOC can be removed effectively upon BAF treatment. The removal efficiency of DOC is used as a main parameter to assess the BAF running condition[21].

There are adequate nutrients and DO (>6 mg/L) in the lower part of the media (below depth of 32 cm). Heterotrophic biomass gained the upper hand in the competition with autotrophic biomass (Fig. 2(b)), resulting in a good performance on DOC removal (Fig. 2(a)). As nutrients were consumed and microbial metabolites were cumulated, autotrophic biomass took a dominant position in the

biofilm instead of heterotrophic biomass (above depth of 32 cm) (Fig. 2(b)). This led to DOC concentration stop decreasing.

The average DON concentration at different media depth in the BAF fluctuated in the range of 0.97 mg/L – 1.56 mg/L, and no significant trend was observed. DON has not been removed effectively upon BAF treatment. The average effluent concentration of DON was 0.97 mg/L, significantly higher than that in surface water, shallow and deep groundwater and source water. Westerhoff and Mash [22] reported that DON comprise only 0.5%–10% of natural organic matter in surface waters by weight. Based upon the analysis of the United States Geologic Survey National Water Quality Assessment databases (a survey of 23,000 water samples), in surface waters, DON concentrations range from less than 0.1 to larger than 10 mg N/L, with a median concentration of approximately 0.3 mg N/L. In shallow and deep groundwater, the average concentration is about 0.24 and 0.18 mg N/L, respectively[22]. According to the analysis on water samples from 28 drinking water treatment plants in United States, Lee[23] demonstrated the average DON concentrations were 0.186 mg N/L for raw waters and 0.148 mg N/L for finished waters. Therefore subsequent effective DON removal process was essential.

3.2 Fluorescence 3D–EEM spectra analysis

3.2.1 Peak identification

The 3D–EEM spectra of water samples from different media depths of BAF were determined, and the peak positions are summarized in Table 2. Three

fluorescence peaks could be identified from the 3D–EEM spectra of samples in the BAF. Peak A was located at the Ex/Em wavelength of 225–235 nm/335–350 nm and was attributed to aromatic tryptophan proteins[4,24]. Peak B at Ex/Em of 275 nm/320–325 nm was also described as tryptophan and proteins[11]. The peak at Ex/Em of 225–245 nm/375–410 nm (Peak C) was described as the fluorescence of fulvic acids[11,25], and a red shift was observed for Peak C with the increase of media depth. The red shift was associated with the later produced microbial metabolites, indicated that the composition of fulvic acids had changed. No obvious variation was observed in the fluorescence intensities of samples from different media depths. Peak identification of 3D–EEM spectra showed that in the BAF the main fluorescent component of DOC was fulvic acids, while tryptophan–like protein was the main fluorescent component of DON.

3.2.2 FRI quantitative analysis

The FRI technique divides the 3D–EEM spectrum into five regions that are related to protein–like substances (Regions I and II), fulvic acid–like substances (Region III), soluble microbial products (SMPs) –like substances (Region IV) and humic acid–like substances (Region V). The normalized integration volumes (NIV) ($\Phi_{i,n}$) of each region for samples from different media depths of BAF are listed in Table 3. Fulvic acids in Region III and humic acids in Region V are typical fluorescent DOC. The NIV of humic acids after BAF treatment kept almost unchanged, while a significant decrease in NIV of fulvic acids from 1.02×10^{7}

AU-nm^2-(mg/L)$^{-1}$ to 0.77×10^7 AU-nm^2-(mg/L)$^{-1}$ was observed. The removal efficiency of fulvic acids upon BAF was 24.51% and that of fluorescent DOC was 11.96%. SMPs in Region IV are mainly composed of tryptophan and protein [11]. Therefore, the compounds in Regions I, II and IV are all fluorescent DON. The NIV of Regions I decreased from 0.91×10^7 AU-nm^2-(mg/L)$^{-1}$ in the influent to 0.46 AU-nm^2-(mg/L)$^{-1}$ in the effluent and 49.46% of tyrosine was removed by BAF. A decrease was also observed in the NIV of Region II from 1.81×10^7 AU-nm^2-(mg/L)$^{-1}$ to 1.21 and 33.52% of tryptophan was removed. The NIV of Region IV remained basically unchanged after BAF treatment. 25.67% of the fluorescent DON was removed in total. Among the compounds in DON, tyrosine and tryptophan–like proteins were effectively removed through BAF process. Simon [26] drew the similar conclusion in the study of a packed–bed biofilter with an EBCT = 6–11 min for seawater treatment. They found that the highest reduction (21%) was observed in Region I, which is primarily composed of aromatic protein–like substances. Protein–like substances are susceptible to be biodegradable [27], leading to the significant DON removal efficiency.

The percentage of NIV for each region was shown in Fig. 3. After BAF treatment, the proportion of tyrosine dropped to 9.16% from 14.40% in the influent, and that of tryptophan dropped from 28.72% to 24.30%. Increases were observed in the percentages of SMPs and humic acids, from 27.72% to 33.54%, and from 13.00% to 17.32%, respectively. BAF treatment did not change the

percentage of fulvic acids. Overall, protein (Regions I, II and IV) accounted for more than 60% of the total FRI and were the main fluorescent substances in the BAF. The percentage of each region in the effluent was in the order of: Region IV > Region II > Region V > Region III > Region I.

3.2.3 PARAFAC model analysis

PARAFAC analysis was applied to quantitatively characterize DOC and DON in the BAF. With an initial exploratory analysis, one outlier (Sample on May 30th at depth of 32 cm) was identified and further removed from the dataset. PARAFAC models with two to seven components were computed, and based on split half analysis, a two component model was appropriate for explaining this specific dataset. The contour plots for the two components as well as the excitation and emission loadings, which reflect the fluorescence character of the component, are plotted in Fig. 4. Component 1, whose peaks were located at Ex/Em of 245 nm/410 nm and 315 nm/410 nm, were associated with humic acids–like substances and represented Region V in FRI. Component 2 was related to tryptophan–like proteins and soluble microbial byproduct–like substances (Ex/Em: 225 nm/335 nm and 275 nm/335 nm), which stood for Regions II and IV, respectively. The fluorescence intensity of Component 2 at the peaks were higher than that of Component 1, this indicated that component 2 got a higher proportion than component 1, which is consist with the conclusion of FRI analysis.

The fluorescence intensity scores of each component for samples at various

BAF depths could also be obtained by PARAFAC model (Fig. 5). The scores of Component 1 kept almost unchanged in the filter media, whereas scores of Component 2 presented a continued decrease. And scores at the bottom of BAF were higher than that of influent. For Component 2, below depth of 32 cm, the score declined slowly from 7484 at the bottom to 7080 at depth of 32 cm, then the score decreased significantly to 4561 when reached depth of 80 cm. The variations of the fluorescence intensity score for both Components 1 and 2 were consistent with the trends of FRI in corresponding regions (Region V for Component 1; Regions II and IV for Component 2). Component 2 decreased by 39.06% in the BAF process by fluorescence intensity score. Through FRI analysis, soluble microbial byproduct–like substances maintained a stable level in the filter media, thus the decrease of Component 2 is primarily due to the degradation of tryptophan–like protein.

The main components of DON and DOC in BAF were identified from the 3D–EEM spectra; the results of peak identification, FRI analysis and PARAFAC model were not in complete accord. Through the peaks identified from 3D–EEM spectra, it is found that fluorescent DOC mainly existed in the form of fulvic acids, fluorescent DON mainly existed in tryptophan protein form. The results obtained by FRI analysis and PARAFAC model are consistent, humic acids–like substance was the main existing form of DOC, and tryptophan–like protein was the main existing form of fluorescent DON.

Due to the low concentration of fluorescent organic matter in source water and the limited sensitivity of peak identification, it is difficult to provide an insight into the variation of peak fluorescence intensity with the increase of media depth. The study of Liu[9] confirmed this through a comprehensive investigation of a full–scale drinking water treatment plant with a treatment train of coagulation/sedimentation, biofiltration and disinfection; they showed that only a slight variation was observed on the peak fluorescence intensity of water samples after a series of treatment process. FRI analysis and PARAFAC model could analyze 3D–EEM spectrum quantitatively, and more details were analyzed. Tyrosine and tryptophan–like proteins were effectively removed through BAF process, while humic acids kept almost unchanged after BAF treatment. Therefore subsequent processes are necessary for the effective removal of humic acids. Peldszus[28] also concluded that direct biofiltration without prior coagulant addition reduced the protein–like content of the membrane feed water which in turn reduced the irreversible fouling potential for UF membranes.

3.3 Correlations between peak fluorescence intensity, FRI and fluorescence intensity score of 3D–EEM spectrum

The correlations between peak fluorescence intensity (PFI), FRI, and fluorescence intensity score (FIS) from 3D–EEM spectrum was analyzed via linear regression by exploring the linear relevance of corresponding parameters. Results are summarized in Figs. 6 (a), (b), and (c). Fig. 6 (a) exhibits the relation

of PFI and FRI. Significant correlations were observed between fluorescence intensity of peak A (PFI–A) and $\Phi_{II,n}$ as well as between PFI–C and $\Phi_{IV,n}$, with R^2=0.86 and 0.86, respectively. The correlation between PFI–B and $\Phi_{III,n}$ was not so significant with R^2=0.49. For PFI and FIS, there were fairly good correlations between PFI–A, PFI–B and FIS–2 (R^2=0.64), as shown in Fig. 6(b).The relevance of FRI and FIS are presented in Fig. 6(c). FIS–1 was significantly correlated with $\Phi_{V,n}$ with R^2=0.89. A strong correlation was also observed between FIS–2 and $\Phi_{II+IV,n}$ with R^2 of 0.97.

The linear relevance analysis demonstrated that the results of three analysis approaches for 3D–EEM spectrum were consistent with each other. Among these correlations, the relevance of FRI and FIS were most significant. Because the peak identification was not sensitive enough for the analysis of 3D–EEM spectrum and the existence of overlapping spectra, the relevance of PFI and FRI as well as PFI and FIS were not as significant as FRI and FIS.

3.4 Comparison of FRI analysis and PARAFAC model

As two distinct quantitative analysis approaches for fluorescence 3D–EEM spectrum, both FRI analysis and PARAFAC model have their own characteristics.

FRI analysis defines fluorescence 3D–EEM spectrum into five regions. The volume for each region (Φ_i) is calculated by a Riemann sum algorithm and then normalized to a DOC concentration of 1 mg/L to calculate a normalized integration volume ($\Phi_{i,n}$) for comparison of 3D–EEM spectra from different

samples[6,11]. The percent fluorescence response ($P_{i,n}$) for each region was also obtained. In the process of PARAFAC analysis, following by determining the appropriate number of component for 3D–EEM spectrum through CORCONDIA or split half technique, PARAFAC model decompose the fluorescence 3D–EEM spectrum into two loading matrixes and one score matrix, which is proportional to the concentration of corresponding fluorophore[6].

FRI analysis require the volume beneath each region of the 3D–EEM spectrum to be integrated one by one, whereas PARAFAC model overcome the problem of spectral overlap and process the 3D–EEM spectrum in batch mode based upon the existing software (MATLAB) and models (e.g., N–way Toolbox, DOMFluor toolbox). Therefore, PARAFAC model could rapidly explain the characteristic fluorescence signal of samples from different sources for a large number of fluorescence 3D–EEM spectra. For a small amount of 3D–EEM spectra, FRI analysis could determine the percentage of each component and offer a comprehensive evaluation of the fluorophore distribution of samples; thereby overcome the non–quantitative comparison of 3D–EEM spectrum. Volumetric analysis is also conducive in studying the relevance of NIV and DOC, DON, SMPs, and so on. The analysis of FRI on substance species was more comprehensive and the presence of tyrosine–like proteins was detected in this study, comparing with PARAFAC model. Therefore, FRI analysis is more suitable to characterize DOC and DON in the BAF than the other two analysis techniques,

i.e., peak identification and PARAFAC model.

4. Conclusion

This study was based on a lab–scale BAF fed with simulated micro–polluted source water. Three approaches, peak identification, FRI analysis and PARAFAC model, were applied to analyze the 3D–EEM spectrum and to characterize the components and characteristics of DOC and DON in the BAF.

Fluorescent DOC mainly existed in the form of humic acids and fulvic acids, fluorescent DON mainly existed in tryptophan protein form. Protein–like substance accounted for more than 60% of the fluorescent DOM.

Tyrosine and tryptophan–like proteins were effectively removed through BAF process, while humic acids kept almost unchanged after BAF treatment.

The results of peak identification, FRI analysis, and PARAFAC model were consistent with each other, and the relevance of FRI and FIS presented to be most significant.

FRI technique is more suitable to characterize DOC and DON in the BAF in this study than peak identification and PARAFAC model.

Acknowledgments

The authors gratefully acknowledge the National High Technology Research and Development Program (863 Program) (2012AA062607) for funding this research.

References:

[1] Bouwer, E. J., Crowe, P. B. Biological processes in drinking water treatment. Journal–American Water Works Association, 1988, 80(9): 82–93.

[2] Hasan, H. A., Abdullah, S. R. S., Kamarudin, S. K., et al. Response surface methodology for optimization of simultaneous COD, NH_4^+–N and Mn_2^+ removal from drinking water by biological aerated filter. Desalination, 2011, 275(1–3): 50–61.

[3] Yu Xin, Qi Zhihua, Zhang Xiaojian et al. Nitrogen loss and oxygen paradox in full–scale biofiltration for drinking water treatment. Water Research, 2007, 41(7): 1455–1464.

[4] Baker, A. Fluorescence excitation–emission matrix characterization of some sewage–impacted rivers. Environmental Science & Technology, 2001, 35(5): 948–953.

[5] Carstea, E. M., Baker, A., Bieroza, M., et al. Continuous fluorescence excitation–emission matrix monitoring of river organic matter. Water Research, 2010, 44(18): 5356–5366.

[6] Sanchez, N. P., Skeriotis, A. T., Miller, C. M. Assessment of dissolved organic matter fluorescence PARAFAC components before and after coagulation–filtration in a full scale water treatment plant. Water Research, 2013, 47(4): 1679–1690.

[7] Coble, P. G. Characterization of marine and terrestrial DOM in seawater using

excitation–emission matrix spectroscopy. Marine Chemistry, 1996, 51(4): 325–346.

[8] Liu, B., Gu, L., Yu, X., et al. Dissolved organic nitrogen (DON) profile during backwashing cycle of drinking water biofiltration. Science of the Total Environment, 2012, 414(1): 508–514.

[9] Liu, B., Gu, L., Yu, X.,et al. Dissolved organic nitrogen (DON) in a full–scale drinking water treatment plant. Journal of Water Supply: Research and Technology–AQUA, 2012, 61(1): 41–49.

[10] Gu, L., Liu, B., Yu, X. Dissolved organic nitrogen (DON) in the processes of polluted source water treatment. Chinese Science Bulletin, 2010, 55(27–28): 3098–3101.

[11] Chen, W., Westerhoff, P., Leenheer, J. A., et al. Fluorescence excitation–emission matrix regional integration to quantify spectra for dissolved organic matter. Environmental Science & Technology, 2003, 37(24): 5701–5710.

[12] Bro, R. PARAFAC: tutorial and applications. Chemometr. Intell. Lab. Syst. 1997, 38(2): 149–71.

[13] Bro, R., Kiers, H. A. A new efficient method for determining the number of components in PARAFAC models. Journal of Chemometrics, 2003, 17(5): 274–286.

[14] Stedmon, C. A., Markager, S., Bro, R. Tracing dissolved organic matter in

aquatic environments using a new approach to fluorescence spectroscopy. Marine Chemistry, 2003, 82(3–4): 239–254.

[15] Ishii, S. K., Boyer, T. H. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineered systems: a critical review. Environmental Science & Technology, 2012, 46(4): 2006–2017.

[16] Yu, Xin, Ye, Lin, Wei, Gu, et al. Modeling the formation of soluble microbial products(SMP) in drinking water biofiltration. Water Science & Engineering, 2008, 1(3): 93-101.

[17] Surmaczgorska, J., Gernaey, K., Demuynck, C., et al. Nitrification monitoring in activated sludge by oxygen uptake rate (OUR) measurements. Water Research, 1996, 30(5): 1228–1236.

[18] Urfer D, Huck PM. Measurement of biomass activity in drinking water biofilters using a respirometric method. Water Research, 2001, 35(6): 1469–1477.

[19] Carlson, K. H., Amy, G. L. The importance of soluble microbial products (SMPs) in biological drinking water treatment. Water Research, 2000, 34(4): 1386–1396.

[20] Velten, S., Boller, M., Köster, O., et al. Development of biomass in a drinking water granular active carbon (GAC) filter. Water Research, 2011, 45(19): 6347–6354.

[21] Jing, Z. Q., Li, Y. Y., Cao, S. W., et al. Performance of double-layer biofilter packed with coal fly ash ceramic granules in treating highly polluted river water.

Bioresource Technology, 2012, 120(3): 212-217.

[22] Westerhoff, P., Mash, H. Dissolved organic nitrogen in drinking water supplies: a review. Journal of Water Supply: Research and Technology–AQUA, 2002, 51(8): 415–448.

[23] Lee W., Westerhoff, P., Esparza-Soto, M. Occurrence and removal of dissolved organic nitrogen in US water treatment plants. Journal–American Water Works Association. 2006, 98(10): 102–110.

[24] Yamashita, Y., Tanoue, E. Chemical characterization of protein–like fluorophores in DOM in relation to aromatic amino acids. Marine Chemistry, 2003, 82: 255–271.

[25] Aryal, R., Lebegue, J., Vigneswaran, S., et al. Identification and characterisation of biofilm formed on membrane bio–reactor. Separation and Purification Technology, 2009, 67(1): 86–94.

[26] Simon, F. Xavier, Rudé, E., Llorens, J., et al. Study on the removal of biodegradable NOM from seawater using biofiltration. Desalination, 2013, 316(5): 8–16.

[27] Salinas, Rodriguez, S. G., Kennedy, M. D., Schippers, J. C., et al. Organic foulants in estuarine and bay sources for seawater reverse osmosis–comparing pretreatment processes with respect to foulant reductions. Desalin. Water Treat., 2009, 9 (1–3): 155–164.

[28] Peldszus, S., Halle, C., Peiris, R. H., et al. Reversible and irreversible low–pressure membrane foulants in drinking water treatment: Identification by principal component analysis of fluorescence EEM and mitigation by biofiltration pretreatment. Water Research, 2011, 45(16): 5161–5170.

Figures and Tables:

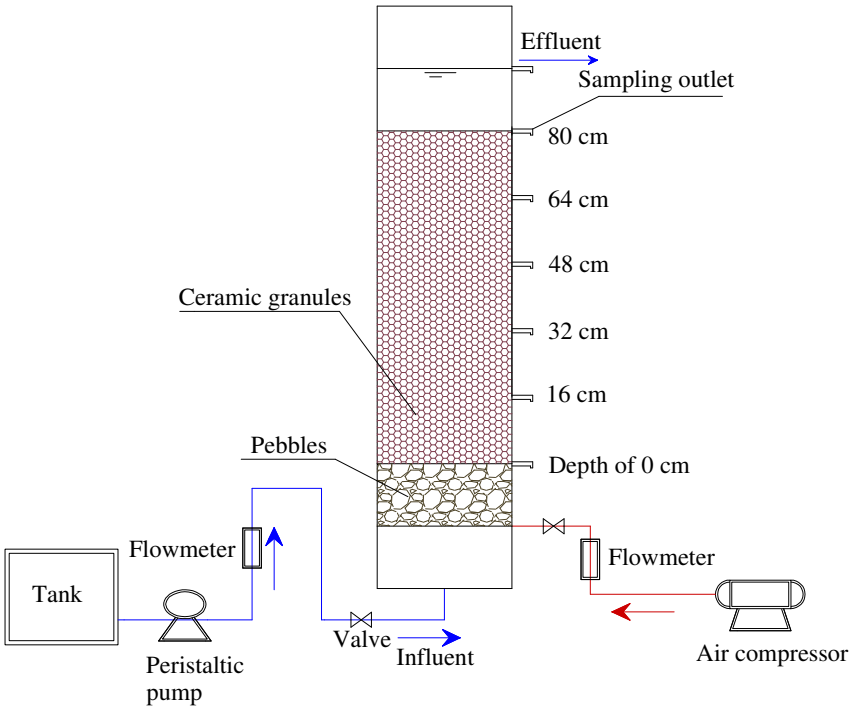
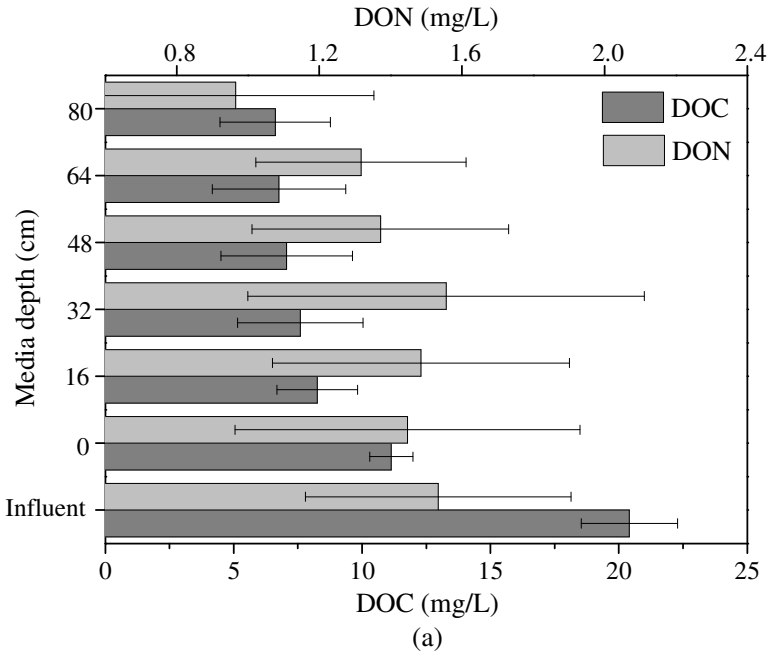


Fig. 1 Schematic diagram of the BAF reactor



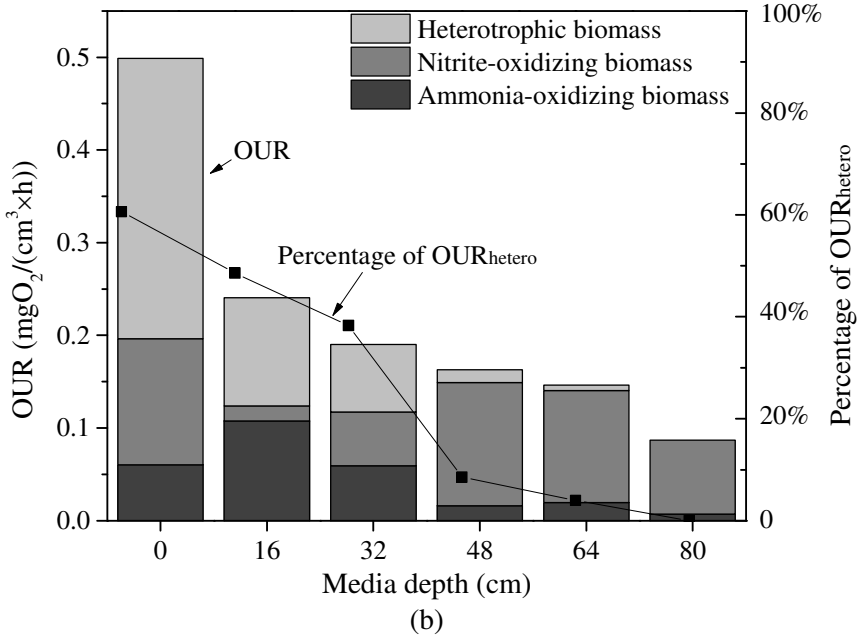


Fig. 2 Profiles of (a) DOC and DON; (b) OUR in the BAF media layer. Error bars are standard deviations of DOC and DON, respectively, calculated from seven times measurements.

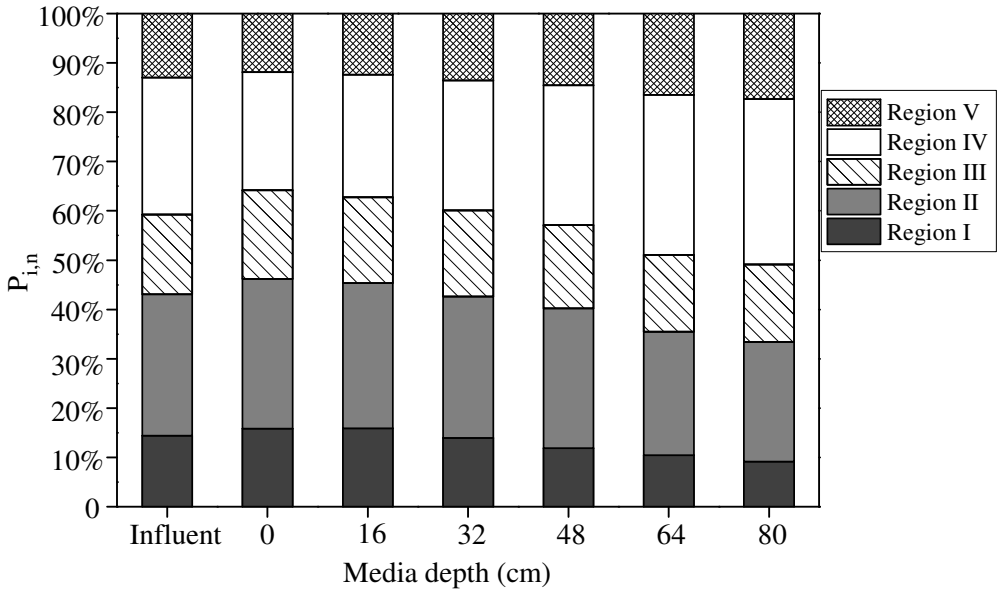


Fig. 3 Percentages ($P_{i,n}$) distribution of FRI at different media depths in the BAF

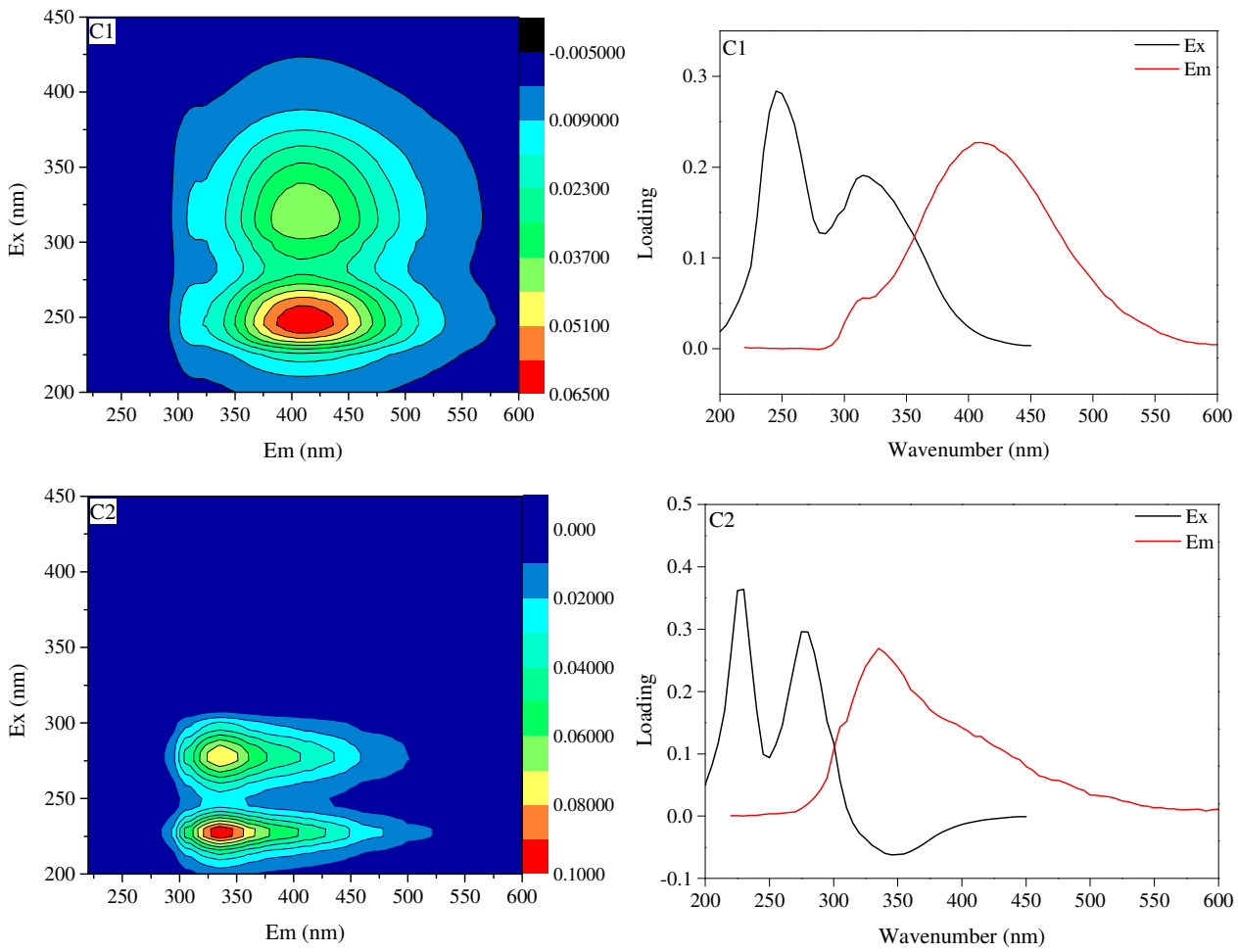


Fig. 4 Two components decomposed by PARAFAC model: C1: Component 1; C2: Component 2

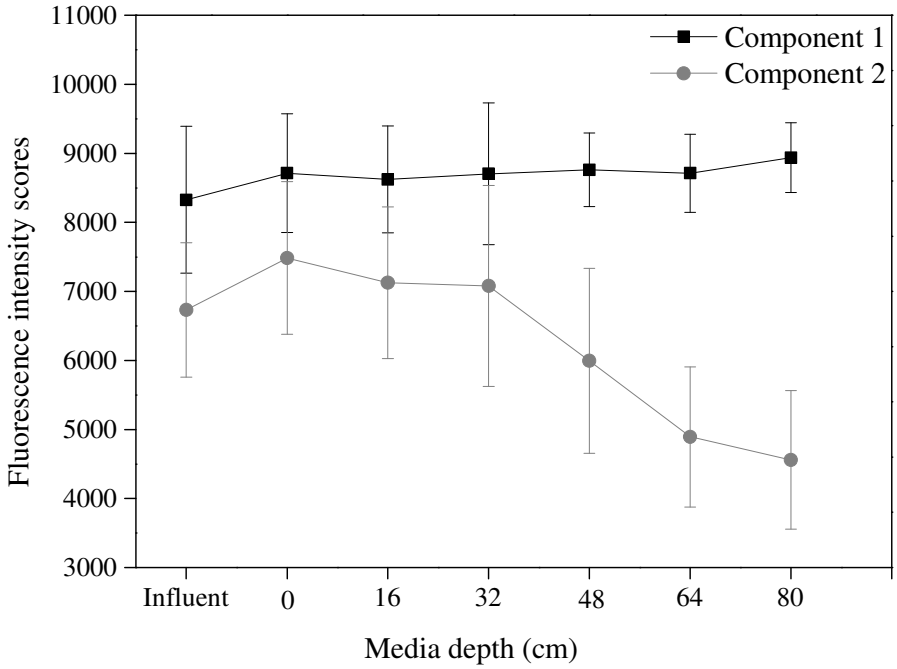


Fig. 5 Fluorescence intensity scores of each component by PARAFAC model

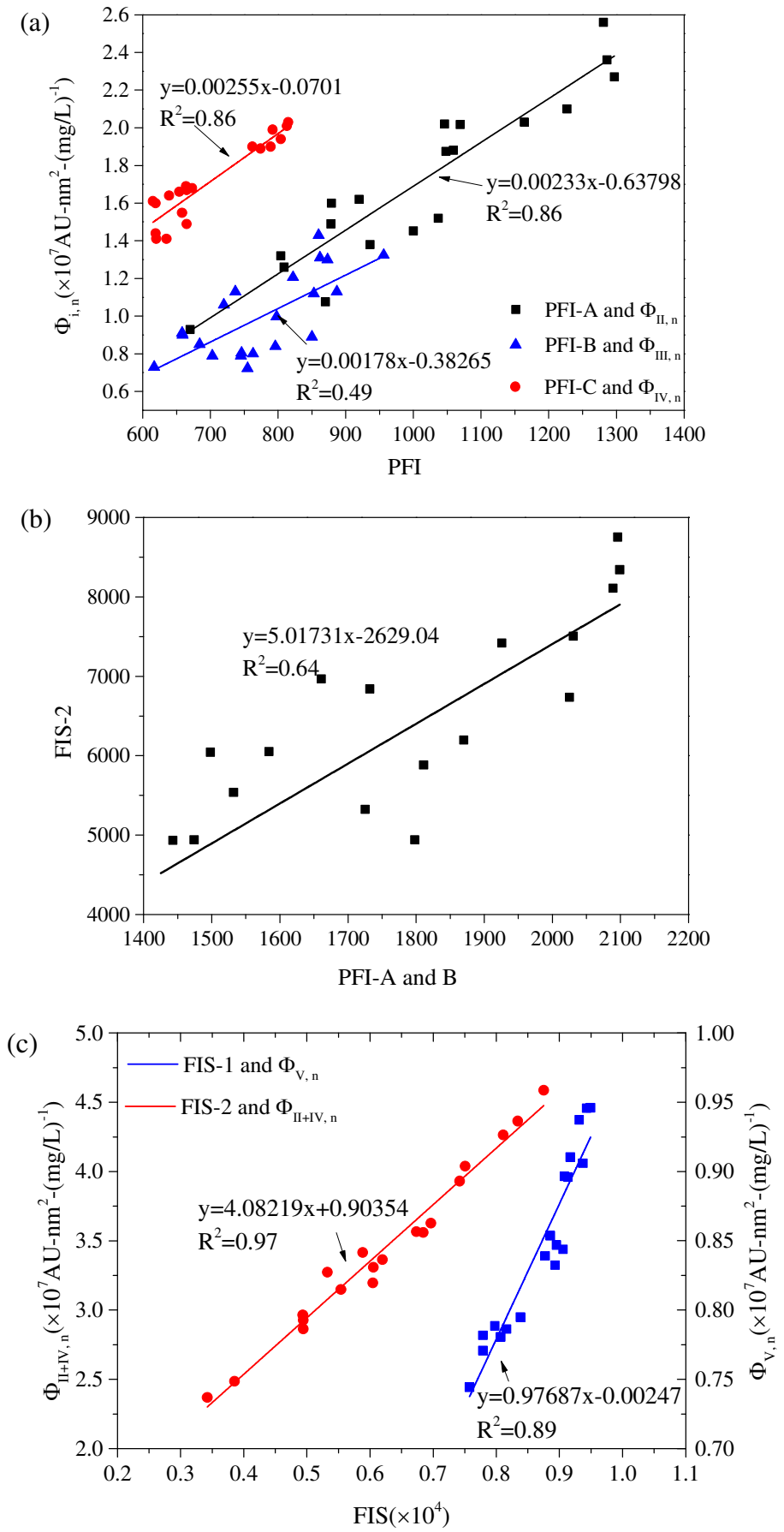


Fig. 6 Linear correlation between (a) PFI and FRI; (b) PFI and FIS; (c) FIS and FRI

Table 1 Excitation and emission wavelength boundaries for five 3D-EEM regions in FRI

Region	Description	Ex/Em wavelengths (nm)
I	Tyrosine - like protein	200-250/280-330
II	Tryptophan - like protein	200-250/330-380
III	Fulvic acid - like substances	200-250/380-550
IV	Soluble microbial by - products	250-400/280-380
V	Humic acid -like substances	250-400/380-550

Table 2 Fluorescence peak positions and intensities of 3D-EEM spectra from different media depths of BAF

Media depth	Peak A		Peak B		Peak C	
(cm)	Ex/Em(nm)	Intensity	Ex/Em(nm)	Intensity	Ex/Em(nm)	Intensity
0 (30^{th})	225/335	1069	235/375	956	275/325	658
16 (30^{th})	225/340	1048	235/385	822	275/325	665
32 (30^{th})	230/345	1050	230/385	885	275/320	600
48 (30^{th})	230/345	1000	235/385	798	275/335	620
64 (30^{th})	230/345	870	240/390	746	275/320	635
80 (30^{th})	235/345	670	245/400	755	275/320	619
Influent (5^{th})	230/340	1164	230/395	887	275/325	762
0 (5^{th})	230/345	1281	225/405	860	275/325	815
16 (5^{th})	230/340	1286	230/390	862	275/330	813
32 (5^{th})	230/340	1297	230/395	873	275/320	792
48 (5^{th})	230/335	1227	230/395	853	275/320	804
64 (5^{th})	230/345	1037	240/395	796	275/325	774
80 (5^{th})	230/340	936	240/400	763	275/320	789
Influent (8^{th})	230/335	879	230/400	658	275/320	619
0 (8^{th})	225/335	1046	230/395	737	275/320	615
16 (8^{th})	225/335	1059	235/385	720	275/320	673
32 (8^{th})	230/335	920	240/400	659	275/320	664
48 (8^{th})	230/340	878	240/400	684	275/320	654
64 (8^{th})	230/340	809	245/400	617	275/320	665
80 (8^{th})	230/350	804	240/400	745	275/320	639

* (30^{th}) - represents samples taken on May 30^{th}; (5^{th}) - represents samples taken on June 5^{th}; (8^{th}) - represents samples taken on June 8^{th}.

Table 3 Distribution of $\Phi_{i,n}$ ($\times 10^7$, AU-nm^2-(mg/L)$^{-1}$) based on FRI analysis

Media Depth (cm)	$\Phi_{1,n}$	$\Phi_{2,n}$	$\Phi_{3,n}$	$\Phi_{4,n}$	$\Phi_{5,n}$	Summations
Influent	0.91	1.82	1.02	1.75	0.82	6.32
0	1.15	2.20	1.29	1.73	0.86	7.23
16	1.11	2.04	1.19	1.73	0.85	6.92
32	0.93	1.87	1.14	1.72	0.88	6.54
48	0.72	1.68	0.99	1.67	0.85	5.91
64	0.54	1.29	0.79	1.66	0.84	5.11
80	0.46	1.21	0.77	1.66	0.85	4.95

A graphical and textual abstract for the contents

Peak identification, FRI and PARAFAC were adopted and compared to analyze the 3D–EEM spectra.