Supporting information

Metabolite-derived fluorescence signals track decay of biochemical oxygen demand throughout dissolved organic matter biodegradation

Hehe Wang[1,2], Yinghui Mo (✉)[1], Kang Xiao (✉)[2,3], Yichen Tian[2], Yizhe Lai[2], Xin Zhao[4], Jianqing Du[2,3], Jihua Tan[2], Xia Huang[5]

1 State Key Laboratory of Advanced Separation Membrane Materials, School of Environmental Science and Engineering, Tiangong University, Tianjin 300387, China

2 Beijing Yanshan Earth Critical Zone National Research Station, College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101408, China

3 State Key Laboratory of Earth System Numerical Modeling and Application, University of Chinese Academy of Sciences, Beijing 101408, China

4 School of Environmental Science and Engineering, Tianjin University, Tianjin 300350, China

5 State Key Laboratory of Regional Environment and Sustainability, School of Environment, Tsinghua University, Beijing 100084, China

✉ Corresponding authors

E-mail: moyinghui@tiangong.edu.cn (Y. Mo); kxiao@ucas.ac.cn (K. Xiao)

Section S1. Sampling and pretreatment of the real water sample and its dissolved organic matter (DOM) compositions.

The real water sample was the membrane bioreactor (MBR) supernatant. It was collected in sterilized bottles, transported to the laboratory in an icebox, and filtered through a 0.45 μm membrane within 6 hours of collection. The MBR supernatant was then stored in the dark at 4 °C until further analysis.

The concentrations of different DOM components were: polysaccharides (PS), 6.1 mg/L; proteins (PN), 3.3 mg/L; and humic acids (HA), 4.8 mg/L. These concentrations were comparable to those of the synthetic mixed DOM sample of this study. PS concentration was determined by the sulfuric acid-phenol method (Dubious, 1956), using sodium alginate (SA; 90%, Macklin, China) as the standard. PN and HA concentrations were determined by a modified Lowry method (Shen et al., 2013; Xiao et al., 2022), with bovine serum albumin (BSA; ≥98%, Macklin, China) and a reference HA (≥90%, Aladdin, China) serving as standards for PN and HA, respectively. For all three components, these standards were used to establish calibration curves relating concentration to absorbance.

Section S2. Bacterial strain selection and inoculum preparation.

Pseudomonas fluorescens (P17) and *Spirillum sp.* (NOX) were selected as inoculum microorganisms. These two standardized model strains are commonly used for determining assimilable organic carbon in various water matrices.(Van der Kooij, 1977; Van der Kooij and Hijnen, 1984) Their combination ensures that a wide range

of organic substrates found in aquatic environments can be degraded. P17 is capable of effectively degrading various readily biodegradable organic matter, including amino acids, simple sugars, and organic acids.(Van der Kooij, 1977) NOX can utilize specific organic acids (e.g., oxalate, often derived from the oxidation of complex organic matter) that P17 cannot effectively utilize.(Van der Kooij and Hijnen, 1984)

For inoculum preparation,(Administration, 2001) 2 mL of the P17 or NOX glycerol stock (provided by a local university), stored at −80 °C, was added into the sterilized nutrient broth of 2 L (5 g/L peptone, 3 g/L beef extract, and 5 g/L NaCl). The nutrient broth was incubated at 25 ℃ with shaking at 180 rpm until the optical density at 600 nm (OD_{600}) reached 0.8−0.9 (approximately 24 hours). Subsequently, 1 mL of the culture was washed three times with 9 mL of sterile phosphate-buffered saline through successive cycles of centrifugation, vortexing, and resuspension. The washed bacterial culture was added (100 μL) into 100 mL of a sterile 5 mg/L sodium acetate solution. The resulting suspension was incubated statically in the dark for 7 days until it reached the stationary phase with a final concentration of approximately 10^7 CFU/mL, and this final suspension served as the inoculum for subsequent experiments.

Section S3. Biodegradation experimental design.

Each DOM sample (i.e., synthetic BSA, SA, HA, MIX samples, and MBR supernatant) was transferred to pre-sterilized dissolved oxygen (DO) bottles (100 mL). Seven to eight bottles were prepared for each DOM sample. The samples were inoculated with 1 mL of the P17 inoculum and 1 mL of the NOX inoculum, resulting in an initial concentration of approximately 10^5 CFU/mL for each strain. The samples

were then supplemented with 100 μL of each mineral salt stock solution to achieve a final conductivity of approximately 720 μS/cm. The stock solutions included a phosphate buffer (8.5 g/L KH_2PO_4, 21.8 g/L K_2HPO_4, 33.4 g/L $Na_2HPO_4 \cdot 7H_2O$, 1.7 g/L NH_4Cl), a $MgSO_4$ solution (11.0 g/L), a $CaCl_2$ solution (27.6 g/L), and a $FeCl_3$ solution (0.15 g/L). The bottles were filled to capacity with the DOM samples.

All bottles were incubated statically at 25 ℃ ± 1 ℃ in the dark, with the bottles being water-sealed. Daily measurement was performed on only one bottle for each DOM, of which the solution was discarded after the measurement (Fig. S1). The entire experiment last approximately 7−8 days. During the initial 5 days, the following indicators were monitored at regular intervals (24 ± 2 h). Upon opening the bottle, DO was measured immediately using a portable dissolved oxygen meter (HQ40d, Hach). Samples were then filtered (0.45 μm) and the pH was adjusted to 7.0 ± 0.1 prior to the analysis of fluorescence excitation-emission matrix (EEM) spectra, ultraviolet-visible (UV-vis) absorption spectra, and total organic carbon (TOC). After the 5th day, only DO was daily monitored until the values remained stable for several days (2−3 days), indicating that the biodegradation process had almost ceased. The stable DO value was defined as the final DO concentration (DO_{end}). By subtracting DO_{end} from DO_t (which was the DO at Day *t*), we obtained the value of biochemical oxygen demand (BOD). To highlight that we investigated the BOD decay during DOM biodegradation, we defined the as-calculated BOD on a certain day as the residual BOD, using the symbol of BOD_r, i.e., $BOD_r(t) = DO_t - DO_{end}$.

Section S4. EEM & UV-vis spectra measurements.

UV-vis absorption spectra were measured using a UV-vis spectrophotometer (Youke, T3200, China). The wavelength scanning range was 200−500 nm, with a scanning interval of 5 nm. A quartz glass cuvette with a 1 cm pathlength was used. Ultrapure water served as the reference solution for baseline correction.

EEM spectra were measured using a fluorescence spectrophotometer (Agilent, Cary Eclipse, USA). Prior to measurement, all samples were first filtered (0.45 μm) and their pH was adjusted to 7.0 ± 0.1 with dilute HCl/NaOH. Given that measurements at lower excitation (Ex) wavelengths may yield significant deviations and noise, some researchers suggest measuring fluorescence intensity at Ex ≥ 220 nm (Yu et al., 2020). Therefore, in this study, the scanned Ex and emission (Em) ranges were 220−400 nm and 250−500 nm, respectively, both with a step size of 5 nm. The photomultiplier tube (PMT) voltage was set to 700 V. Slit widths were set to 5 nm. The scan speed was set to 2400 nm/min.

Section S5. EEM spectra normalization.

The fluorescence matrix of ultrapure water was subtracted from the EEM of the samples. First-order and second-order Rayleigh and Raman scattering signals were eliminated using an interpolation method (Yu et al., 2020). Inner filter effects were corrected using the UV-vis absorbance of the same sample in the 200−500 nm wavelength range (a maximum absorbance less than 1.5 was ensured to validate the correction effectiveness) (Yu et al., 2020). Finally, fluorescence intensity units were normalized by referencing the area of the water's primary Raman peak, resulting in intensities expressed in Raman units (R.U.) (Yu et al., 2020).

Section S6. Traditional spectral parameter calculation.

$$UV_{ave} = \frac{\sum_{\lambda=200}^{400} A_{\lambda}}{N} \quad (S1)$$

$$SUV = \frac{UV_{ave}}{TOC} \quad (S2)$$

$$SUV_{254} = \frac{A_{254}}{TOC} \quad (S3)$$

$$SUV_{280} = \frac{A_{280}}{TOC} \quad (S4)$$

$$Fl_{ave} = \frac{1}{N}\sum_{Ex}\sum_{Em} I \quad (S5)$$

, $FI\text{-}i = \frac{1}{N}\sum_{Ex}\sum_{Em} I_{ith\ regions}$, (S6)

(i = I, II, ..., V)

$$Peak\ A = \underset{Em \in [400,460]}{Max}\ I_{260,\ Em} \quad (S7)$$

$$Peak\ B = I_{275, 305} \quad (S8)$$

$$Peak\ T = I_{275, 340} \quad (S9)$$

$$Peak\ M = \underset{Ex \in [290, 310],\ Em \in [370, 410]}{Max}\ I_{Ex,Em} \quad (S10)$$

$$Peak\ C = \underset{Ex \in [320, 360],\ Em \in [420, 460]}{Max}\ I_{Ex,Em} \quad (S11)$$

$$Peak\ N = I_{280, 370} \quad (S12)$$

$$SFI = \frac{Fl_{ave}}{TOC} \quad (S13)$$

$$FRI\text{-}i = \frac{FI\text{-}i}{\sum_i FI\text{-}i}, (i = I, II, ..., V) \quad (S14)$$

$$HIX_{syn} = \frac{I_{390, 408}}{I_{355, 373}} \quad (S15)$$

$$HIX_{em} = \frac{\langle I \rangle_{254, 435-480}}{\langle I \rangle_{254, 300-345}} \quad (S16)$$

$$BIX = \frac{I_{310, 380}}{I_{310, 430}} \quad (S17)$$

$$fi = \frac{I_{370, 450}}{I_{370, 500}} \quad (S18)$$

$$Peak\ TC = \frac{I_{275, 350}}{I_{330, 420}} \quad (S19)$$

$$QY = \frac{FI_{ave}}{UV_{ave}} \quad (S20)$$

where A represents the ultraviolet (UV) absorbance; N is the number of valid data points within the specified region; Ex represents the excitation wavelength (nm); Em represents the emission wavelength (nm); I is the fluorescence intensity at a specific Ex and Em wavelength pair (Ex is indicated by the first number in the subscript of I and Em is indicated by the second number); represents the averaged fluorescence intensity within a corresponding defined spectral region, calculated as = $\frac{1}{N}\sum_{Ex}\sum_{Em} I$;

and TOC represents the total organic carbon.

The EEM spectra are traditionally divided into five distinct regions based on Ex and Em wavelength ranges: Region I (Ex ≤ 250 nm, Em ≤ 330 nm); Region II (Ex ≤ 250 nm, Em = 330−380 nm); Region III (Ex < 250 nm, Em > 380 nm); Region IV (Ex > 250 nm, Em < 380 nm); and Region V (Ex > 250 nm, Em > 380 nm).

Section S7. Molecular analysis.

Ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS) was used to analyze the organic compositions at the molecular level, which was performed on a Vanquish UHPLC system (Thermo Fisher, Germany) coupled with a Q Exactive^{TM} HF-X hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher, Germany). The UHPLC system was equipped with a reversed-phase Hyperil GOLD^{TM} C18 column (Thermo Fisher, Germany). The mass spectrometer was operated with the positive/negative polarity mode at a spray voltage of 3.2 kV, a capillary temperature of 320 °C, a sheath gas flow rate of 40 arb, and an aux gas flow rate of 10 arb. The $|m/z|$ range of 100–1050 was collected to identify the molecular compositions. Molecular formulas based on additive ions, molecular ion peaks, and fragment ions were identified with the mzCloud database.

Table S1. Traditional UV-vis parameters.

UV-vis parameters	Definition	Significance	Classification
A_{254}	Absorbance at 254 nm	Content of aromatic compounds and humic-like substances	Absolute parameters[a]
A_{280}	Absorbance at 280 nm	Content of protein and compounds with double bonds	
UV_{ave}	Equation (S1) in Section S6	Overall content of organic matter absorbing ultraviolet light	
SUV	Equation (S2) in Section S6	Averaged UV absorbance per TOC	Relative parameters[b]
SUV_{254}	Equation (S3) in Section S6	Content of aromatic compounds and humic-like substances per TOC	
SUV_{280}	Equation (S4) in Section S6	Content of protein components and compounds with double bonds per	
A_{250}/A_{365}	The ratio of absorbance at 250 nm to that at 365 nm	Molecular weight	
A_{210}/A_{254}	The ratio of absorbance at 210 nm to that at 254 nm	Aromaticity	
A_{220}/A_{254}	The ratio of absorbance at 220 nm to that at 254 nm	Polarity	

a: Absolute parameters refer to absorbances; b: relative parameters refer to absorbance ratios or absorbances per TOC.

Table S2. Traditional fluorescence parameters (absolute parameters[a]).

Fluorescence parameters	Definition	Significance
FI_{ave}	Equation (S5) in Section S6	Content of fluorescent substances
FI-I & FI-II	Equation (S6) in Section S6	Tyrosine and tryptophan-like aromatic proteins
FI-III		Fulvic acid-like substances
FI-IV		Dissolved microbial metabolic product-like substances
FI-V		Humic acid-like substances
Peak A & Peak B	Equations (S7) and (S8) in Section S6	Tyrosine-like substances
Peak T	Equation (S9) in Section S6	Tryptophan-like substances
Peak M	Equation (S10) in Section S6	Marine humic-like substances
Peak C	Equation (S11) in Section S6	Humic acid-like substances
Peak N	Equation (S12) in Section S6	Phytoplankton-derived DOM

a: Absolute parameters refer to fluorescence intensities.

Table S3. Traditional fluorescence parameters (relative parameters[a]).

Fluorescence parameters	Definition	Significance
SFI	Equation (S13) in Section S6	Average fluorescence contribution per TOC
FRI-I, II, III, IV, V	Equation (S14) in Section S6	Same as FI-I, II, III, IV, V in Table S2
HIX_{syn}	Equation (S15) in Section S6	Humification degree related to C/N ratio
HIX_{em}	Equation (S16) in Section S6	Humification degree related to H/C ratio
BIX	Equation (S17) in Section S6	Degree of microbial involvement in DOM formation
fi	Equation (S18) in Section S6	Origin of DOM (microbial versus terrestrial)
Peak TC	Equation (S19) in Section S6	Biodegradability of DOM
QY	Equation (S20) in Section S6	Relative efficiency of absorbed UV light converted to fluorescence

a: Relative parameters refer to intensity ratios or intensities per TOC.

Table S4. Possible metabolites with known structures for the BSA sample.

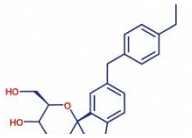
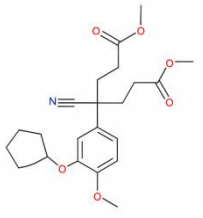
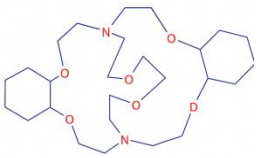
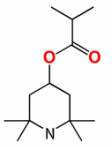
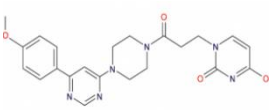
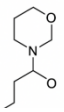
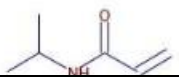
Name	Molecular formula	Structure
Tofogliflozin	C22H26O6	
Dimethyl 4-cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl]heptanedioate	C22H29NO6	
11, 17, 24, 29, 32-Hexaoxa-1, 14-diazatetracyclo [12.12.8.0~5, 10~.0~18, 23~]tetratriacontane	C26H48N2O6	
2, 2, 6, 6-Tetramethyl-4-piperidyl methacrylate	C13H23NO2	
1-(3-{4-[6-(4-Methoxyphenyl)-4-pyrimidinyl]-1-piperazinyl}-3-oxopropyl)-2, 4(1H, 3H)-pyrimidinedione	C22H24N6O4	
N-Nonanoylmorpholine	C13H25NO2	
Nipam	C6H11NO	

Table S5. Possible metabolites with known structures for the HA sample.

Name	Molecular formula	Structure

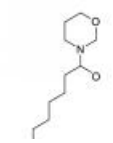
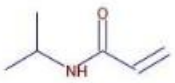
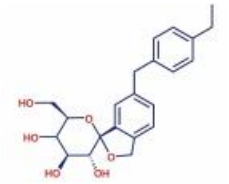
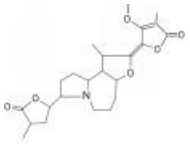
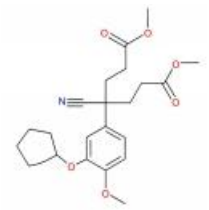
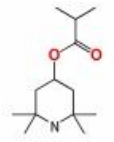
N-Nonanoylmorpholine	$C_{13}H_{25}NO_2$	
Nipam	$C_6H_{11}NO$	
Tofogliflozin	$C_{22}H_{26}O_6$	
Protostemonine	$C_{23}H_{31}NO_6$	
Dimethyl 4-cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl]heptanedioate	$C_{22}H_{29}NO_6$	
2, 2, 6, 6-Tetramethyl-4-piperidyl Methacrylate	$C_{13}H_{23}NO_2$	

Table S6. Possible metabolites with known structures for the SA sample.

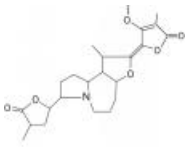
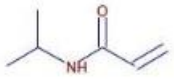
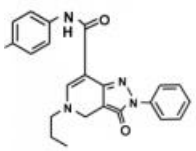
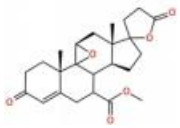
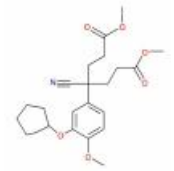
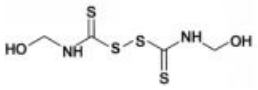
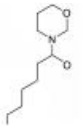
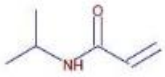
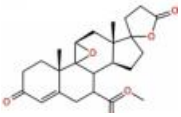
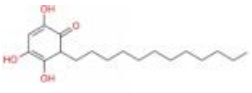
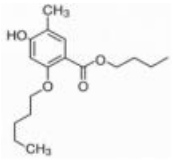
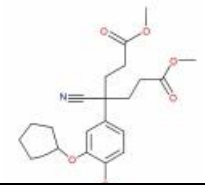
Name	Molecular formula	Structure
Protostemonine	$C_{23}H_{31}NO_6$	
Nipam	$C_6H_{11}NO$	
N-(4-Methylphenyl)-3-oxo-2-phenyl-5-propyl-3, 5-dihydro-2H-pyrazolo[4, 3-c]pyridine-7-carboxamide	$C_{23}H_{22}N_4O_2$	
Eplerenone	$C_{24}H_{30}O_6$	
Dimethyl 4-cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl]heptanedioate	$C_{22}H_{29}NO_6$	
11-Aminoundecanoic acid	$C_{11}H_{23}NO_2$	
[Disulfanediylbis(carbonothioylimino)] bis(hydroxymethane)	$C_4H_8N_2O_2S_4$	

Table S7. Possible metabolites with known structures for the MIX sample.

Name	Molecular formula	Structure
N-Nonanoylmorpholine	$C_{13}H_{25}NO_2$	
Nipam	$C_6H_{11}NO$	
Eplerenone	$C_{24}H_{30}O_6$	
Embelin	$C_{17}H_{26}O_4$	
Diethylene glycol n-butyl ether	$C_8H_{18}O_3$	
Diethylene glycol	$C_4H_{10}O_3$	
Vanillyl nonanoate	$C_{17}H_{22}O_4$	
Dimethyl 4-cyano-4-[3-(cyclopentyloxy)-4-methoxyphenyl]heptanedioate	$C_{22}H_{29}NO_6$	

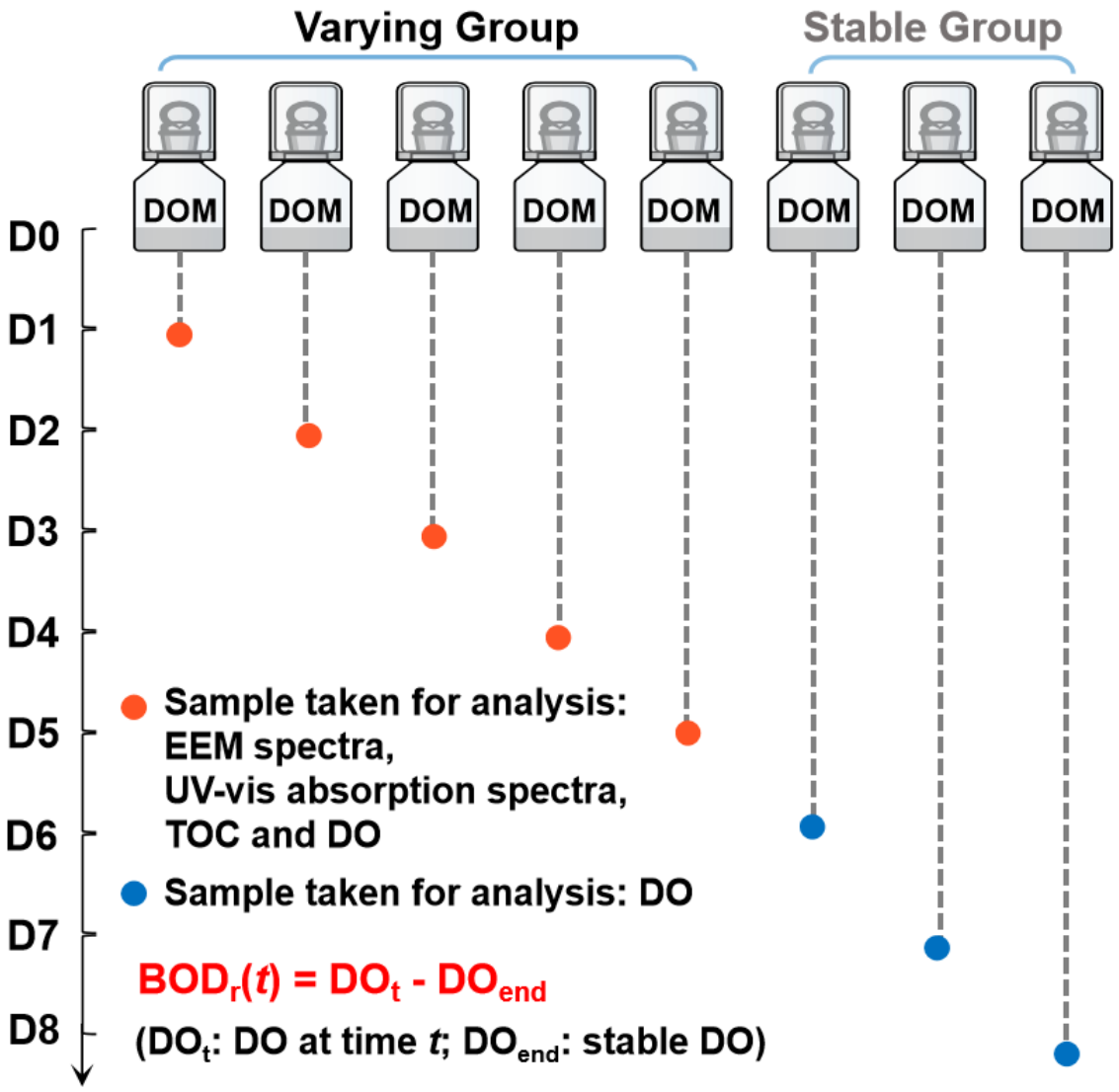


Fig. S1 Schematic diagram of the experimental design (BOD_r refers to residual BOD; the subscript 'r' indicates 'residual', which was added in order to highlight that BOD decay during DOM biodegradation was investigated; D refers to 'day').

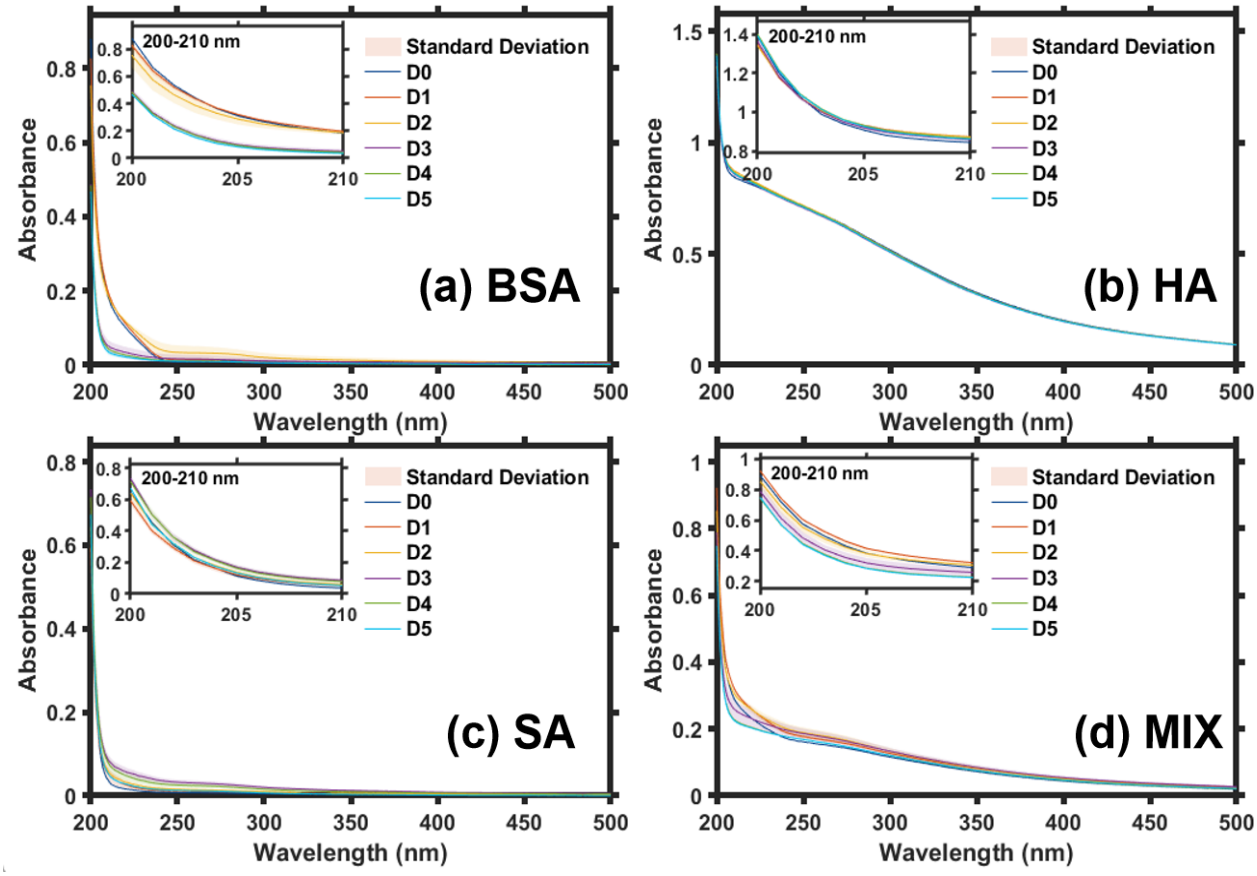


Fig. S2 Temporal evolution of UV-vis absorption spectra for four DOM samples (a: BSA; b: HA; c: SA; d: MIX) over a 5 d's biodegradation period (D0~D5 denote Day 0 to Day 5).

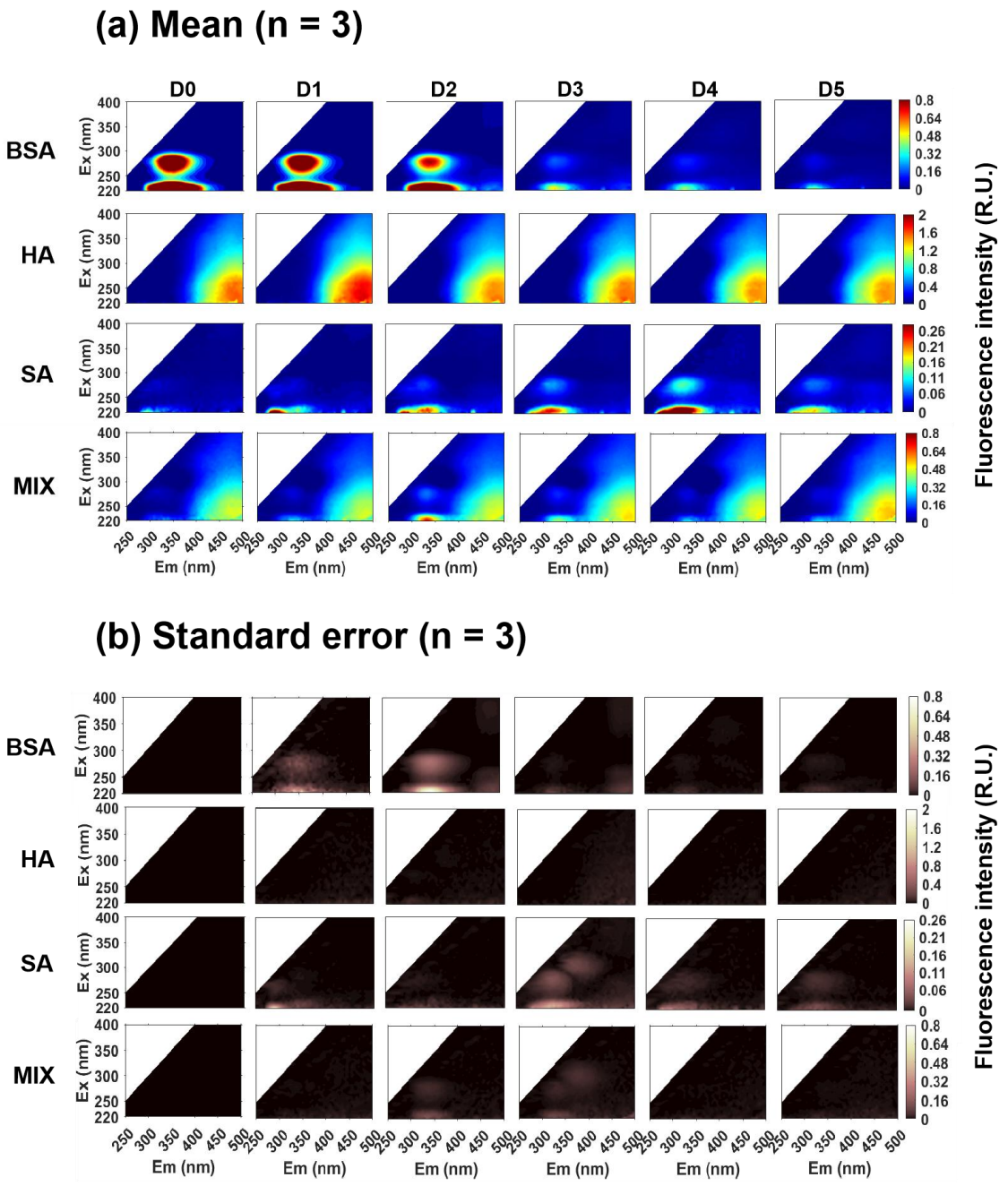


Fig. S3 Temporal evolution of fluorescence EEM spectra for four DOM samples (BSA, HA, SA, MIX) over a 5 d's biodegradation period (D0~D5 denote Day 0 to Day 5): (a) Mean fluorescence intensity (in Raman units, R.U.); (b) Standard deviation of fluorescence intensity (sample number was 3).

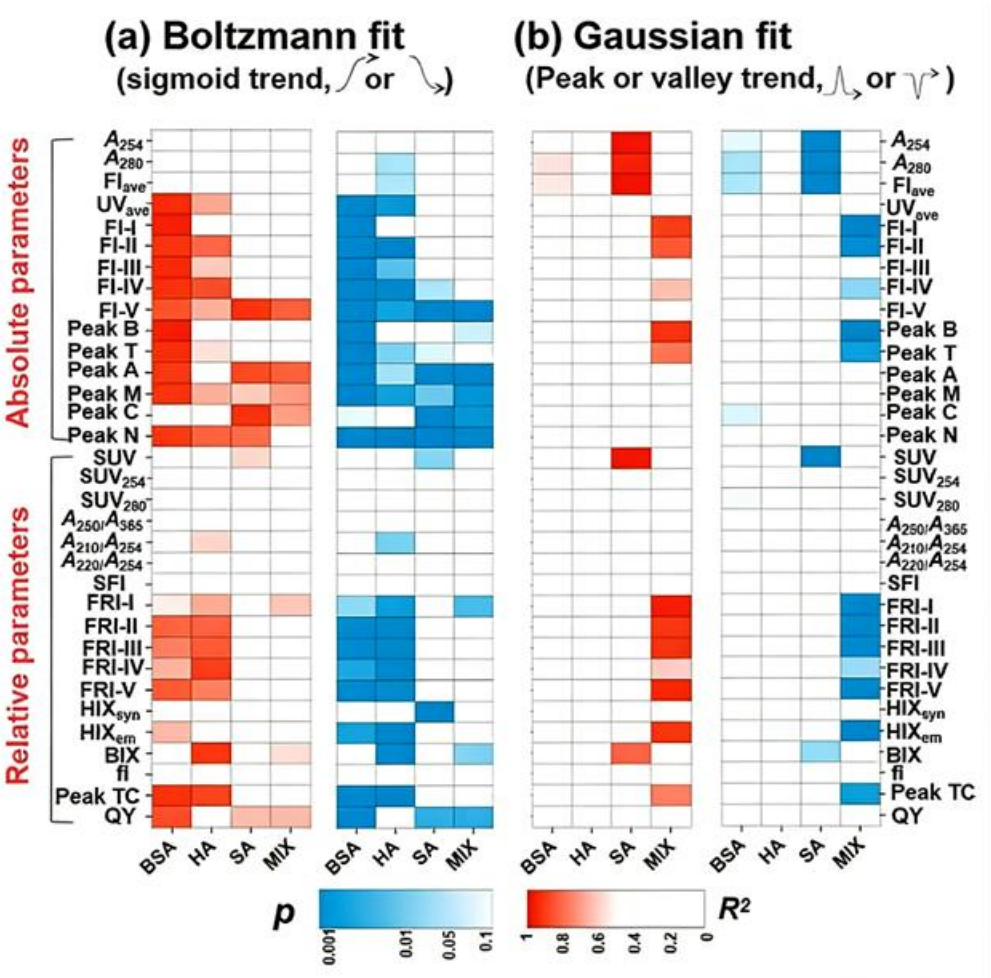


Fig. S4 Coefficient of determination (R^2) and statistical significance (p) obtained by fitting the traditional spectral parameter data with time using the Boltzmann and Gaussian models for the four DOM samples: (a) Fitting with the Boltzmann model; (b) Fitting with the Gaussian model. Absolute parameters refer to intensities or absorbances, and relative parameters refer to intensity or absorbance ratios, or intensities or absorbances per TOC.

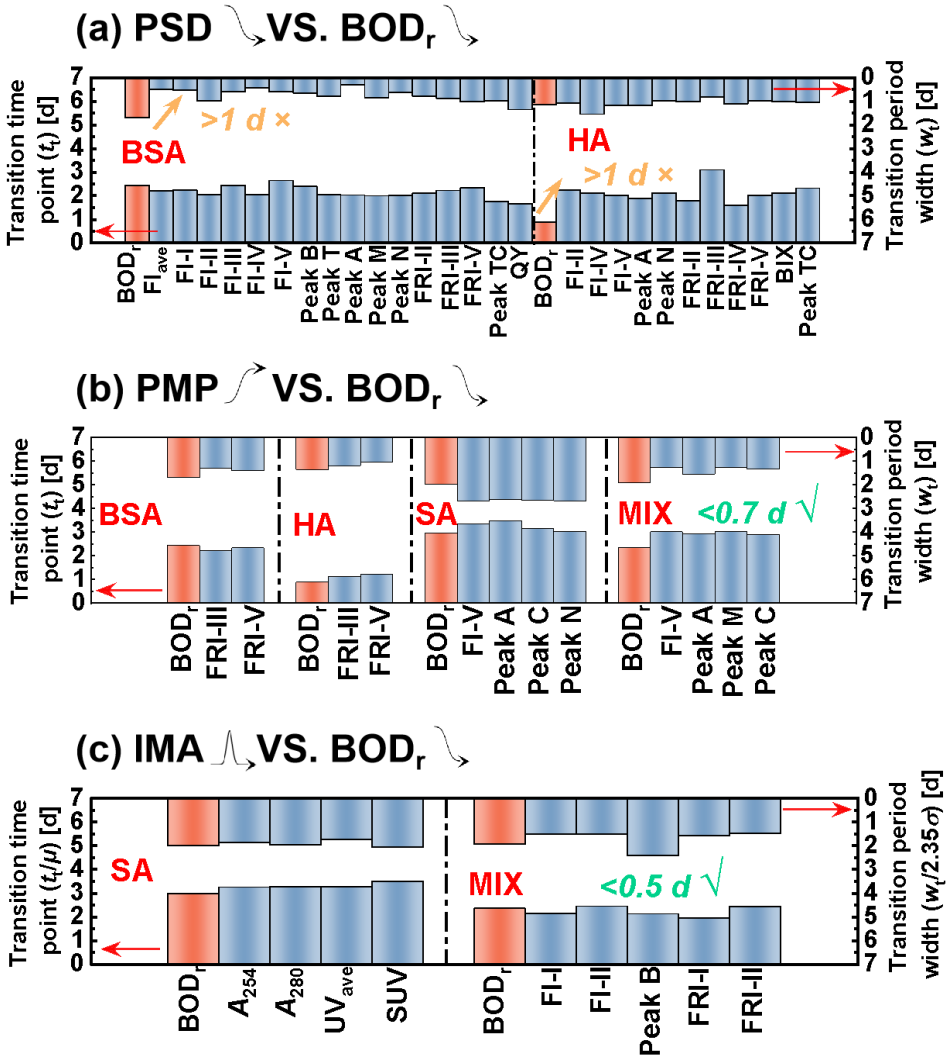


Fig. S5 Comparison of model parameters derived from the temporal variation of traditional spectral parameters with those derived from the temporal variation of BOD_r (a: for those spectral parameters following the PSD pattern; b: for those spectral parameters following the PMP pattern; c: for those spectral parameters following the IMA pattern). Note that t_t and w_t are parameters of the Boltzmann model, and and 2.35σ are parameters of the Gaussian model. t_t and share analogous significance, and w_t and 2.35σ share analogous significance. The explanation of the PSD, PMP, and IMA patterns can be found in Table 1 in the main text.

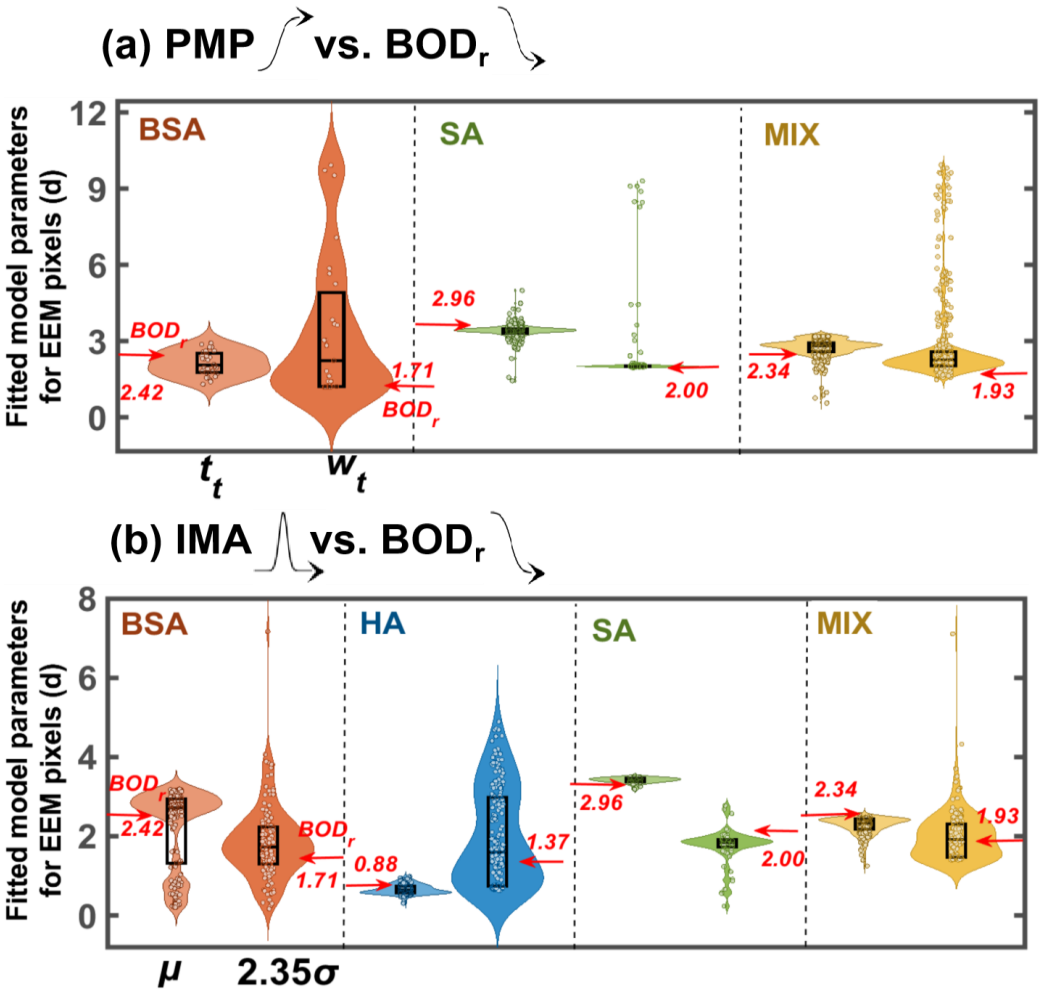


Fig. S6 Comparison of model parameters derived from the temporal variation of EEM image pixel intensities with those derived from BOD_r (a: for those EEM image pixel intensities following the PMP pattern; b: for those EEM image pixel intensities following the IMA pattern). The models used to fit the EEM image pixel intensity data were the Boltzmann and Gaussian models, and the model used to fit the BOD_r data was the Boltzmann model. Note that the central line within the box represents the median of the data, and the upper and lower edges of the box represent the 75th and 25th percentiles of the data, respectively. The violin shape illustrates the data distribution density. The arrows indicate the model parameters derived from the BOD_r data.

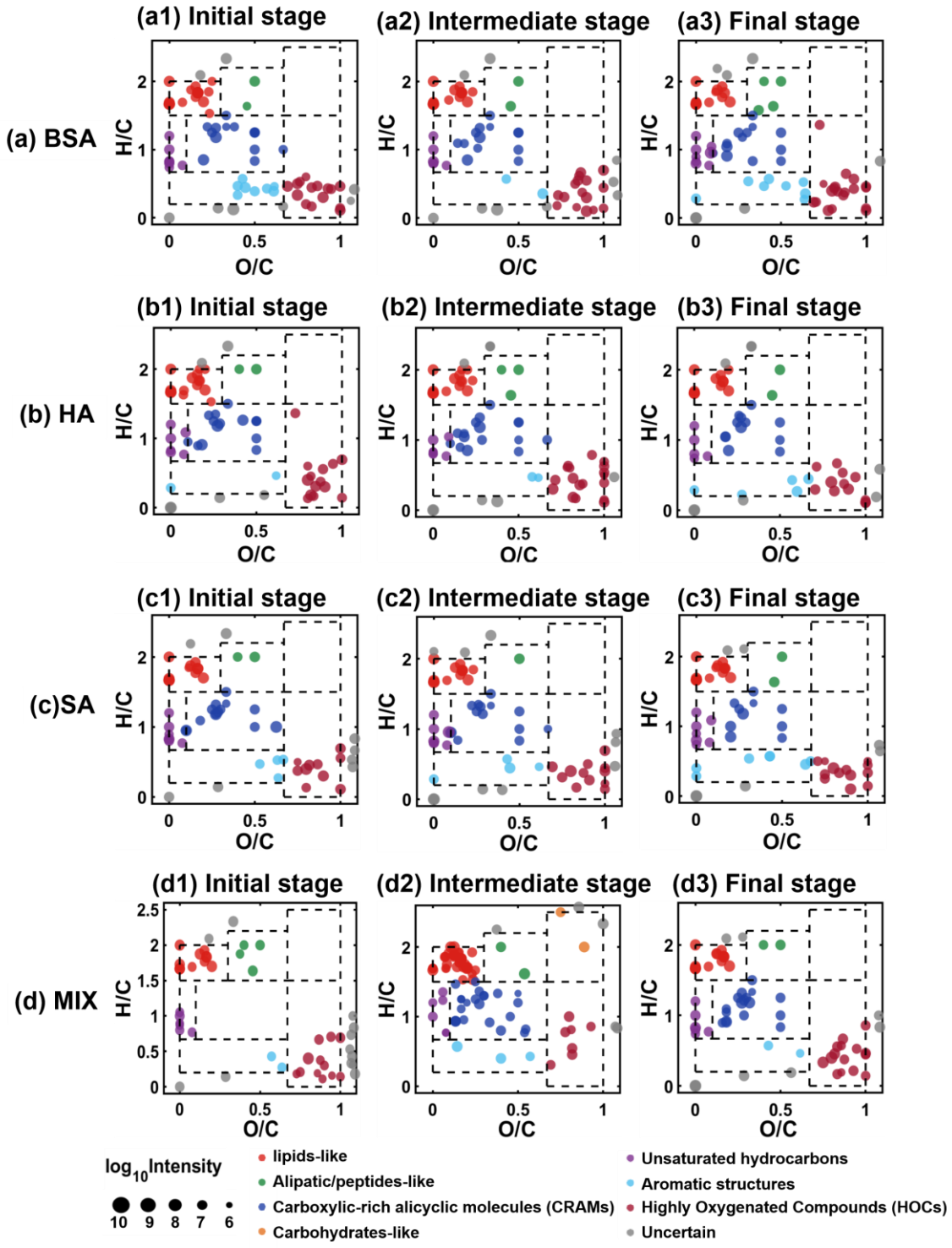


Fig. S7 Molecular compositions during the biodegradation of the four DOM samples (a: BSA; b: HA; c: SA; d: MIX) at key stages (initial, intermediate, and final stages) revealed by intensity-based van Krevelen (VK) diagrams derived from UHPLC-MS data. The initial and final stages correspond to Day 0 and Day 5, respectively. The intermediate stage was selected based on

derived from the Gaussian model, which were Day 3, Day 1, Day 3, and Day 2 for the four samples, respectively.

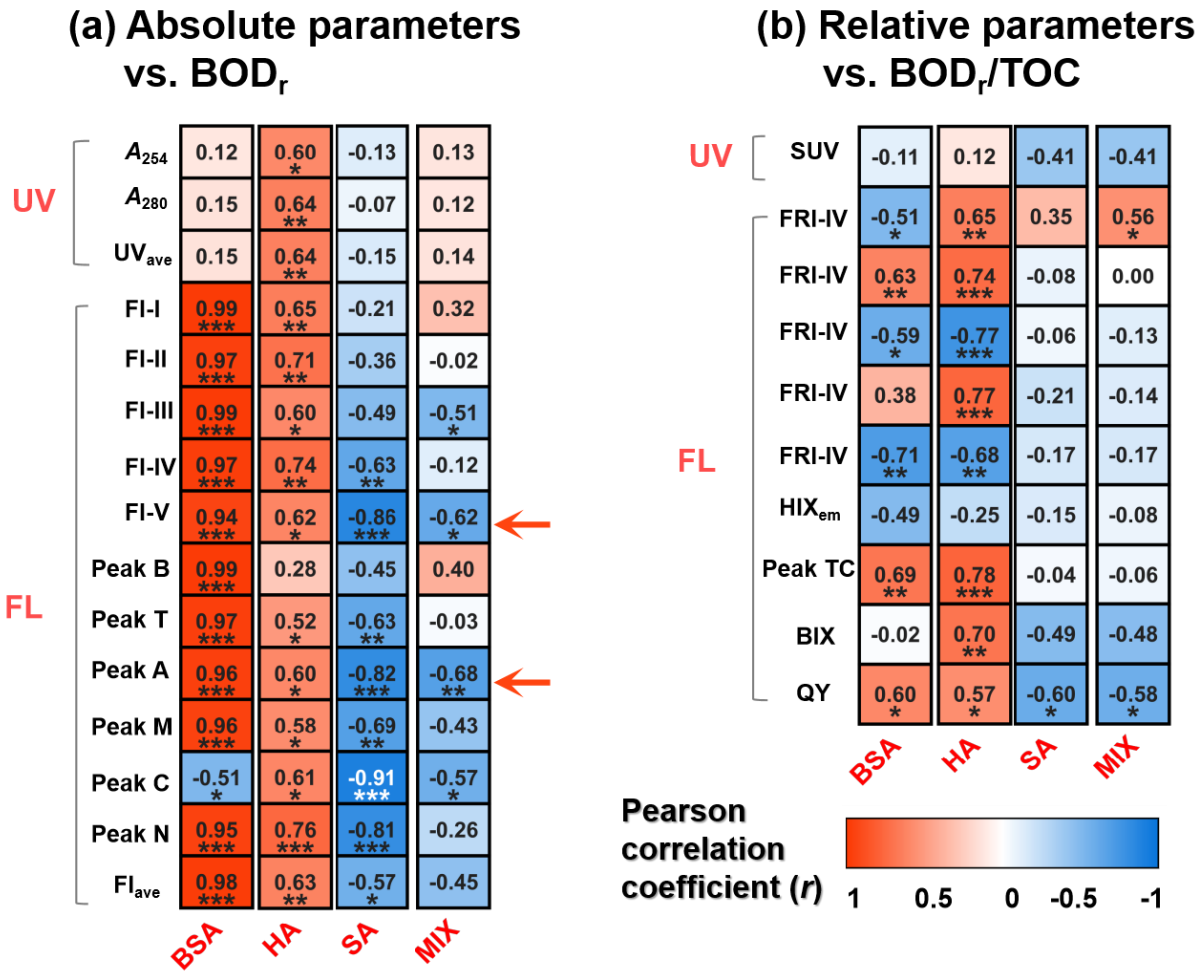


Fig. S8 Pearson correlation analysis between the critical traditional spectral parameters and BOD_r (a: Absolute parameters vs. BOD_r; b: Relative parameters vs. BOD_r/TOC). Pearson correlation coefficients (r) are color-coded, and statistical significance levels are indicated by asterisks (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). Absolute parameters refer to intensities or absorbances, and relative parameters refer to intensity or absorbance ratios, or intensities or absorbances per TOC.

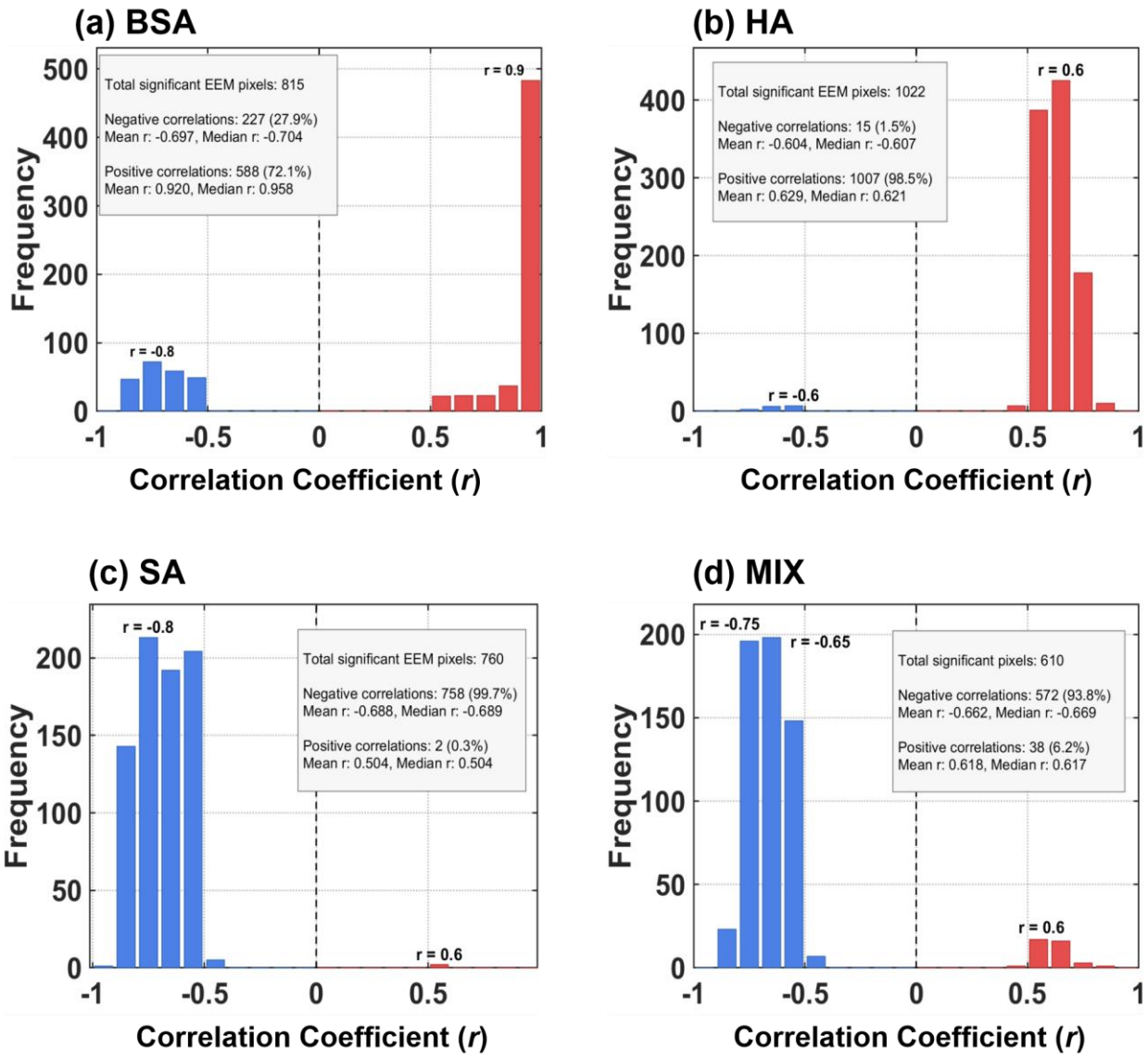


Fig. S9 Pearson correlation coefficients (*r*) between EEM image pixel intensities and BOD_r (a: BSA; b: HA; c: SA; d: MIX). Red columns indicate significantly positive correlations with BOD_r, while blue columns indicate significantly negative correlations (statistical significance threshold: $p < 0.05$).

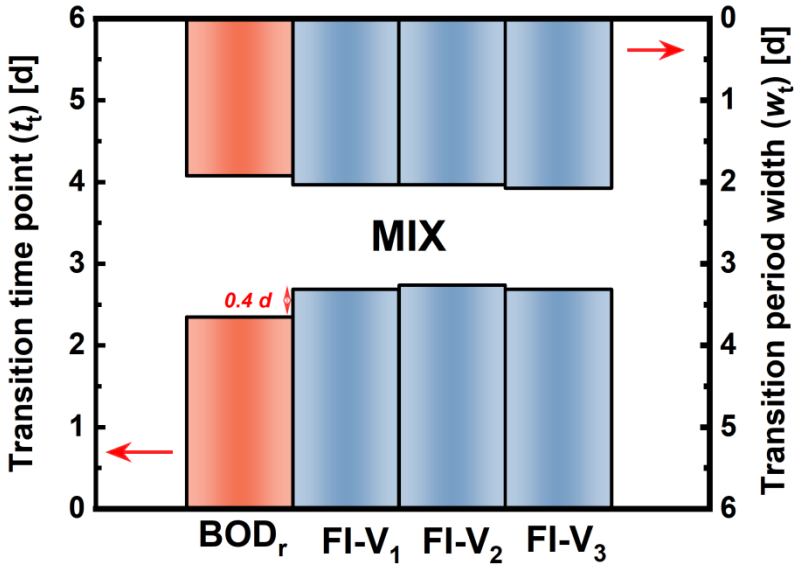


Fig. S10 Comparison of model parameters derived from the temporal variation of the novel fluorescence parameters with those derived from the temporal variation of BOD_r. The temporal data were fitted using the Boltzmann model. The sample investigated was the MIX sample.

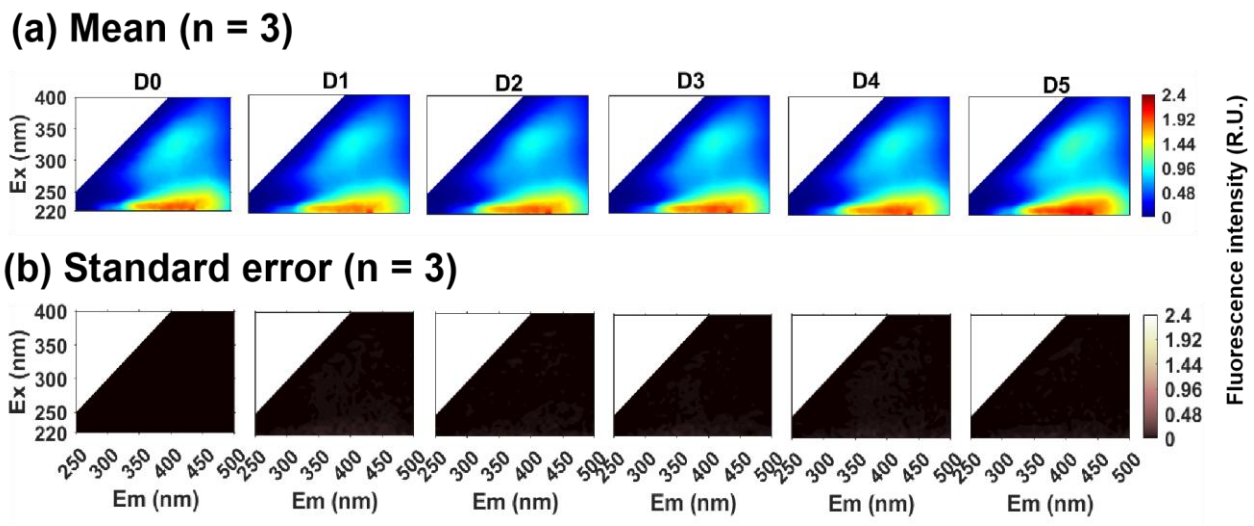


Fig. S11 Temporal evolution of fluorescence EEM spectra for the MBR supernatant over a 5 d's biodegradation period (D0~D5 denote Day 0 to Day 5): (a) Mean fluorescence intensity (in Raman units, R.U.); (b) Standard deviation of fluorescence intensity (sample number for calculation was 3).

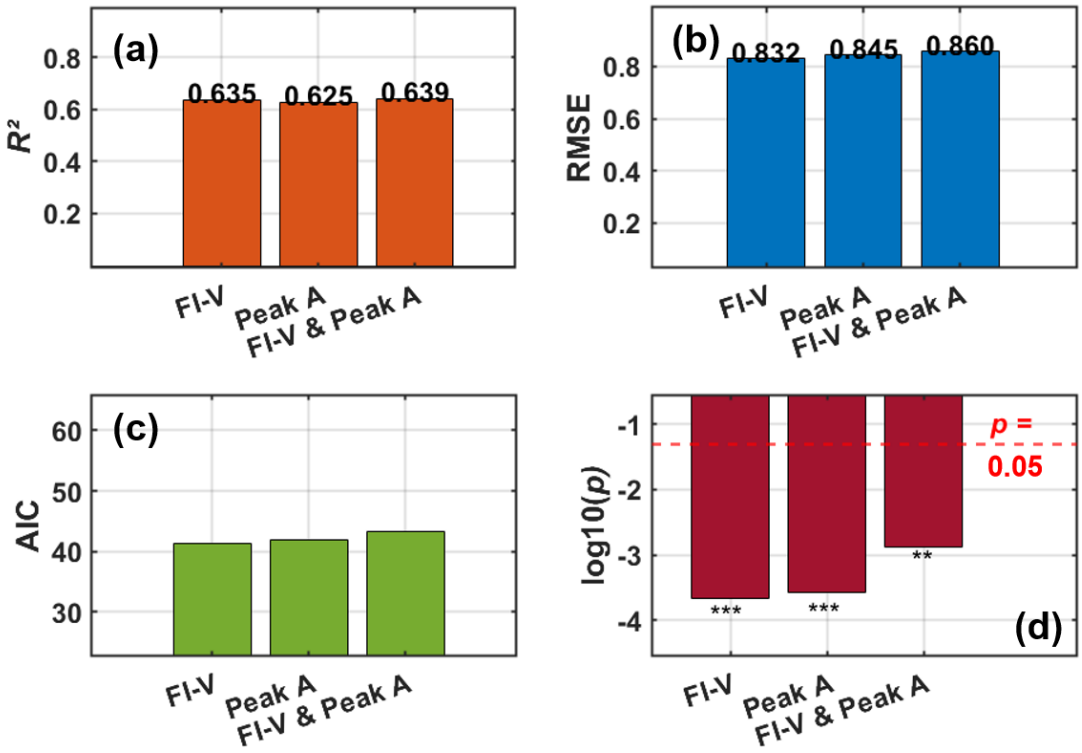


Fig. S12 Key parameters of the linear regression models correlating BOD_r with the two traditional fluorescence parameters using the MBR supernatant as the sample (a: Coefficients of determination (R^2); b: Root mean square error (RMSE); c: Akaike information criterion (AIC), for which lower values indicate a better model; d: Statistical significance (p, presented as $\log_{10}(p)$ for better comparison)). FI-V and Peak A indicate the univariate linear regression. FI-V & Peak A indicates the bivariate linear regression.

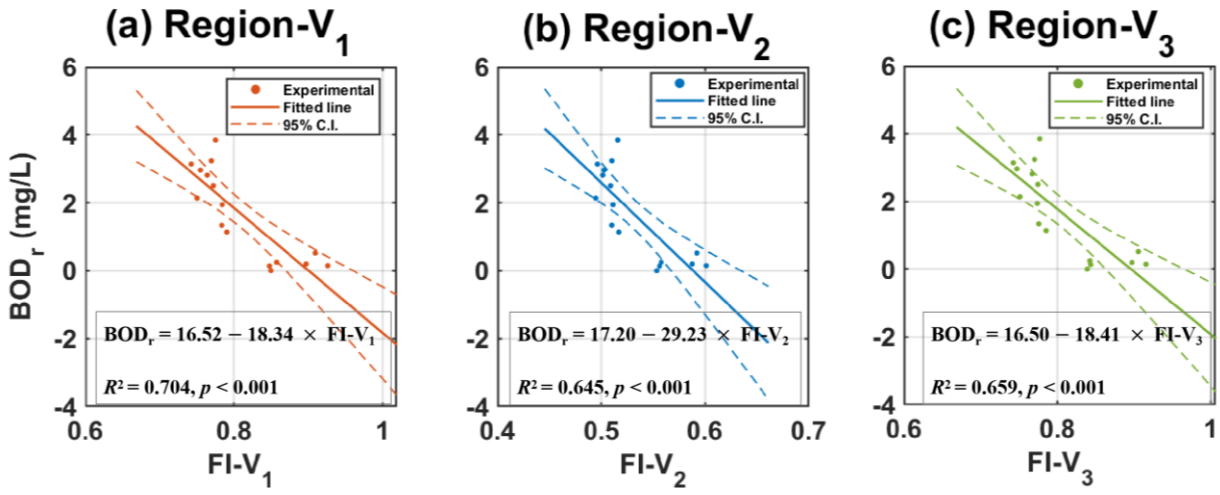


Fig. S13 Univariate linear regression models correlating BOD_r and the three newly developed fluorescence parameters (a: FI-V_1; b: FI-V_2; c: FI-V_3) using the MBR supernatant as the sample. The BOD_r, FI-V_1, FI-V_2, and FI-V_3 were tested/calculated during 5 d's biodegradation of the MBR supernatant sample. The dashed lines represent the 95% confidence intervals (C.I.).

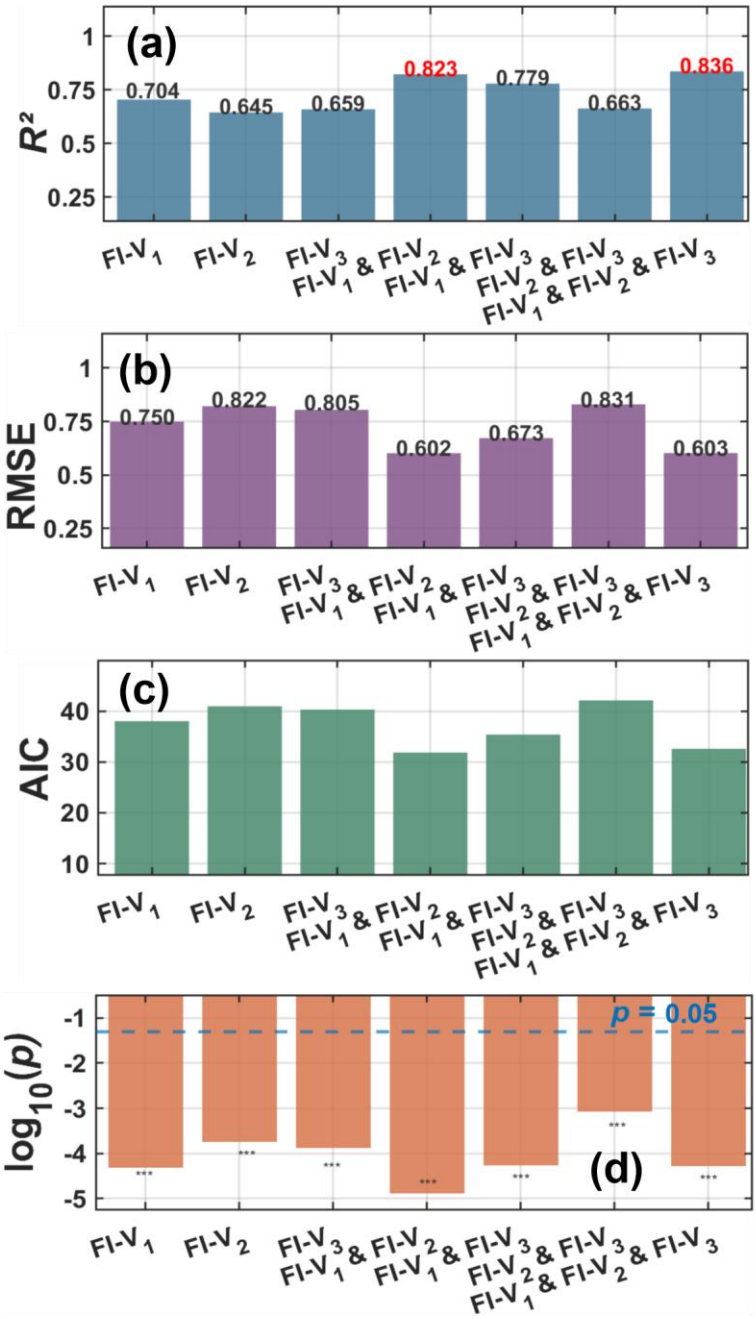


Fig. S14 Key parameters of the linear regression models correlating BOD_r with the three newly developed fluorescence parameters using the MBR supernatant as the sample (a: Coefficients of determination (R^2); b: Root mean square error (RMSE); c: Akaike information criterion (AIC), for which lower values indicate a better model; d: Statistical significance (p, presented as $\log_{10}(p)$ for better comparison)). FI-V_i indicates the univariate linear regression. FI-V_i & FI-V_j indicates the bivariate linear regression. FI-V_1 & FI-V_2 & FI-V_3 indicates the trivariate linear regression.

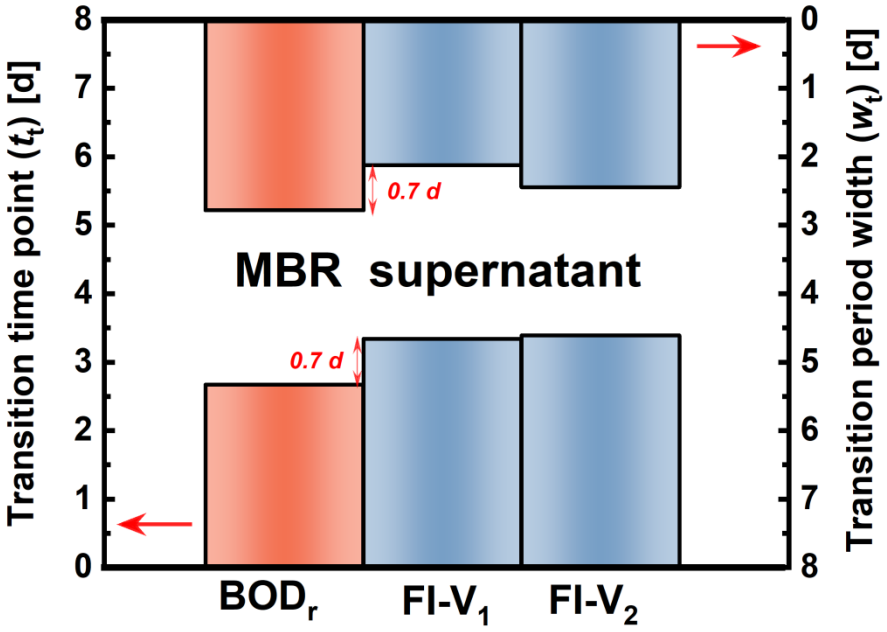


Fig. S15 Comparison of model parameters derived from the temporal variation of the novel fluorescence parameters with those derived from the temporal variation of BOD_r. The temporal data were fitted using the Boltzmann model. The sample investigated was the MBR supernatant.

Abbreviations

AIC: Akaike Information Criterion; **BIX**: Biological Index; **BOD:** Biochemical Oxygen Demand; **BOD_r**: Residual Biochemical Oxygen Demand; **BSA**: Bovine Serum Albumin; **CRAMs**: Carboxylic-rich alicyclic molecules; **DO**: Dissolved Oxygen; **DOM**: Dissolved Organic Matter; **EEM**: Excitation-Emission Matrix; **Em**: Emission; **Ex:** Excitation; **FI**: Fluorescence Intensity; **FRI**: Fluorescence Regional Integration; **FWHM**: Full Width at Half Maximum; **fi**: fluorescence index; **HA**: Humic Acid; **HIX_{em}**: Humification Index (based on emission wavelength); **HIX_{syn}**: Humification Index (based on synchronous scan); **HOCs**: Highly Oxygenated Compounds; **IMA**: Intermediate Metabolite Accumulation; **MBR**: Membrane Bioreactor; **PMP**: Progressive Metabolite Production; **PSD**: Progressive Substrate Depletion; **QY**: Quantum Yield; **RMSE**: Root Mean Square Error; **SA**: Sodium Alginate; **SFI**: Specific Fluorescence Intensity; **SUV**: Specific Ultraviolet Absorbance; **TOC**: Total Organic Carbon; **UHPLC-MS**: Ultra-High-Performance Liquid Chromatography-Mass Spectrometry; **UV-vis**: Ultraviolet-visible.

References

Administration, U.S.F.a.D. (2001) Bacteriological Analytical Manual, Online, AOAC International, Gaithersburg, Md.

Dubious, M, 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. 28, 350-366.

Shen, Y.-x., Xiao, K., Liang, P., Ma, Y.-w. and Huang, X, 2013. Improvement on the modified Lowry method against interference of divalent cations in soluble protein measurement. Applied Microbiology and Biotechnology 97(9), 4167-4178.

Van der Kooij, D, 1977. The occurrence of Pseudomonas spp. in surface water and in tap water as determined on citrate media. Antonie van Leeuwenhoek 43(2), 187-197.

Van der Kooij, D. and Hijnen, W, 1984. Substrate utilization by an oxalate-consuming Spirillum species in relation to its growth in ozonated water. Applied and Environmental Microbiology 47(3), 551-559.

Xiao, K., Horn, H. and Abbt-Braun, G, 2022. “Humic substances” measurement in sludge dissolved organic matter: A critical assessment of current methods. Chemosphere 293, 133608.

Yu, J., Xiao, K., Xue, W., Shen, Y.-x., Tan, J., Liang, S., Wang, Y. and Huang, X, 2020. Excitation-emission matrix (EEM) fluorescence spectroscopy for characterization of organic matter in membrane bioreactors: Principles, methods and applications. Frontiers of Environmental Science & Engineering 14(2), 31.