Supplementary Information

Optical and Molecular Characteristics of Urban Wastewater Dissolved Organic Matter: Insights into Their Correlations

Jiangyong Chu,[ab] Zhenliang Liao*[abc]

[a] College of Environmental Science and Engineering, Tongji University, Shanghai, 200092, China

[b] UNEP-Tongji Institute of Environment for Sustainable Development, State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai 200092, China

[c] College of Civil Engineering and Architecture, Xinjiang University, Urumqi 830047, China

* **Corresponding author:** Zhenliang Liao; 04150@tongji.edu.cn

Supplementary Information content

S1 Sampling area

Urban domestic wastewater samples were collected in Shanghai, China (Yangpu District, Area: 60.61 km^2). Meteorological data were recorded, including rainfall (average 48.6 mm), relative humidity (about 94%), atmospheric pressure (average 29.40 mmHg), and surface temperature (average 23 °C). The detailed information on the sampling environment is similar to the stormwater samples, which have been described in detail [1].

S2 TOC-VCPH analyzer detection settings

Instrument settings for TDN (Total dissolved organic nitrogen, TDN) and DOC (Dissolved organic carbon, DOC) detection are as follows: injection frequency of 1, cleaning frequency of 2, sample injection volume of 150 μL, acid addition of 1.5% of the sample injection volume, injection airflow of 80 ml, injection time of 1.5 min, automatic dilution of 1-fold, and maximum integration time of 4 min 50s.

S3 CDOM parameters calculation

CDOM spectral parameters analysis, the absorption at 350 nm (a_{350}) is used to quantify the CDOM pool. The absorption coefficient (a_{CDOM}) [2] is obtained using the following equation:

$$a_{CDOM}(\lambda) = 2.303 * a_{CDOM}(\lambda)/L \quad (1)$$

$a_{CDOM}(\lambda)$ is the absorbance of DOM containing chromophoric groups at wavelength λ, L is the

path length of the light pool in meters, and 2.303 is the conversion factor from common logarithm to natural logarithm. The equation for calculating the spectral slope coefficient (S, μm) [3] is:

$$a(\lambda)= a(\lambda_0)\, e^{S(\lambda_0-\lambda)} + K \quad (2)$$

Where $a(\lambda)$ is the absorption coefficient, λ_0 is the reference wavelength, λ is the selected wavelength, $a(\lambda_0)$ is the absorption coefficient at the reference wavelength, and K is the background parameter.

In addition, the spectral slope ratio (S_R) [4] is calculated as follows.

$$S_R = S_{275-295} / S_{350-400} \quad (3)$$

$S_{275–295}$ and $S_{350–400}$ are linear fittings of the logarithmic transformation absorption coefficients within the wavelength ranges of 275-295 nm and 350-400 nm, respectively, with units of nm-1. SR is commonly used for the characterization of molecular weight and aromaticity [4]. Specific UV-vis absorbance at 254 and 280 nm ($SUVA_{254}$ and $SUVA_{280}$) are calculated by dividing the absorbance at 254 and 280 nm by the concentration of dissolved organic carbon (DOC) and are often used as indicators of aromaticity [5].

$$SUVA_{254} = a_{254} / \rho_{DOC} \quad (4)$$

$SUVA_{254}$ is the specific UV absorbance at 254 nm, with simplified units of $L \cdot mg^{-1}\ m^{-1}$; a_{254} is the absorption coefficient at a wavelength of 254 nm, with units of m^{-1}; ρ_{DOC} is the concentration of dissolved organic carbon (DOC) in the water sample, with units of $mg \cdot L^{-1}$.

S4 FDOM parameters calculation

The specific formulas for calculating the three fluorescence optical parameters are as follows:

$$FI = I_{470nm} / I_{520nm},\ \lambda_{Ex}=370\ nm \quad (5)$$

$$BIX = I_{380nm} / I_{430nm},\ \lambda_{Ex}=310\ nm \quad (6)$$

$$HIX = \sum I_{435-480nm} / \sum I_{300-345nm},\ \lambda_{Ex}=254\ nm \quad (7)$$

The fluorescence index (FI) is the ratio of fluorescence intensity at emission wavelengths of 470 nm and 520 nm, respectively, to the excitation wavelength of 370 nm [6,7]. FI is used to indicate the source of dissolved organic matter (DOM), ranging from terrestrial sources (plant and soil organic matter, approximately 1.2) to microbial sources (bacterial and algal byproducts, approximately 1.8). The biological index (BIX) is the ratio of fluorescence intensity at emission wavelengths of 380 nm and 430 nm to the excitation wavelength of 310 nm [8]. BIX increases with the freshness of DOM [8]. The humification index (HIX) is the ratio of the integrated fluorescence intensity at emission wavelengths of 435-480 nm and 300-345 nm to the excitation wavelength of 254 nm (or 255 nm in this study due to the 5 nm excitation interval of the spectrofluorometer), representing the degree of humification of DOM, which generally increases with increasing aromaticity [9,10].

S5 EEM-PARAFAC modeling analysis

The combination EEM dataset of all collected samples was modeled using PARAFAC with drEEM in MATLAB [11,12]. The model was tested for repeated convergence using least squares and non-negative constraints on the data. To avoid the disproportionate impact of low signal-to-noise ratios on the modeling, excitation wavelengths below 250 nm and emission wavelengths above 600 nm were removed [11]. Prior to PRAAFAC analysis, spectral regions without fluorescence and affected by first- and second-order Raman scattering peaks were removed and replaced with missing values. The PARAFAC model was validated through residual analysis, split-half analysis, random initialization, sum of squares error, explained variance, and core consistency tests, explaining over 99.7% of the total fluorescence variance. The content of each component was determined based on the Fmax of identified components. The emission and excitation load of the identified components were compared with those in the open fluorescence database, OpenFluor (https://openfluor.lablicate.com). The PARAFAC model was further compared and validated using Tucker's congruence coefficient (TCC, TCC Ex/Em > 0.95). PARAFAC modeling was performed using the DOM-Fluor toolbox in MATLAB R2022a (www.models.kvl.dk/source) [13].

S6 Molecular parameters calculation

In addition, the double bond equivalent (DBE), the modified aromaticity index (AI_{mod}), and the nominal oxidation state of carbon (NOSC) were calculated as characterization parameters for DOM molecular features [14,15]. DBE represents the unsaturation of DOM molecules [14]. A higher DBE value indicates greater diversity and more double bonds and aromatic rings in the compounds [14]. AI_{mod} represents the modified aromaticity index of DOM molecules [16], estimating the

proportion of aromatic and condensed aromatic structures, and considering the abundance of carboxylic groups in natural organic matter [14]. NOSC represents the nominal oxidation state of carbon, which is used to describe the oxidation state or degree of carbon atoms in organic compounds. The nominal oxidation state of carbon is estimated for each carbon atom in the compound [15]. Weight average (wa) represents the weighted average based on the mass spectral ion intensities (Schmidt et al., 2017).

The specific calculation formulas are as follows:

$$DBE = 1 + \frac{2C - H + N + P}{2} \tag{8}$$

$$AI_{mod} = \frac{1 + C - 0.5O - S - 0.5H}{2C - 0.5O - S - N - P} \tag{9}$$

$$(NOSC - O = 4 - \frac{4C + H - 3N - 2O + 5P - 2S}{C} \tag{10}$$

C, H, N, P, and S refer to the number of Carbon, Hydrogen, Nitrogen, Phosphorus, and Sulfur atoms in the molecular formula.

$$X = \frac{\sum_{i=1}^{n} X_i * I_i}{\sum_{i=1}^{n} I_i} \tag{11}$$

X_i represents the corresponding value in each molecule, and I_i represents the peak intensity of the molecule in the mass spectrum.

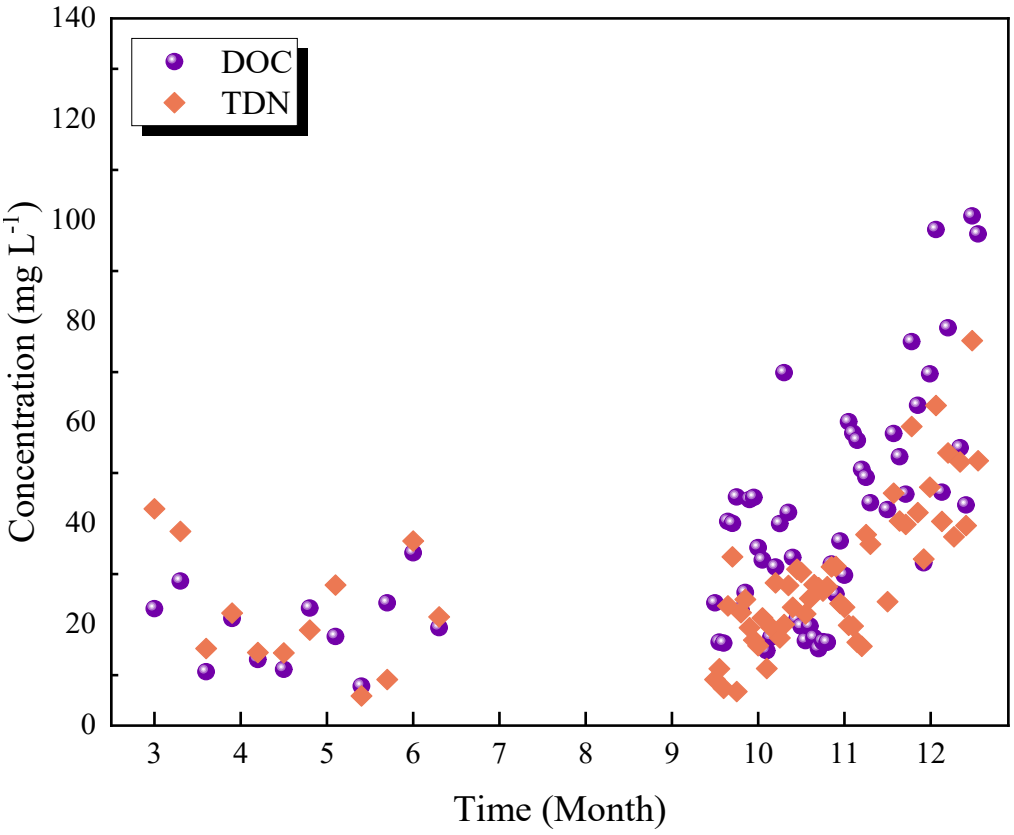

Fig. S1 The content of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) in urban domestic wastewater at different periods.

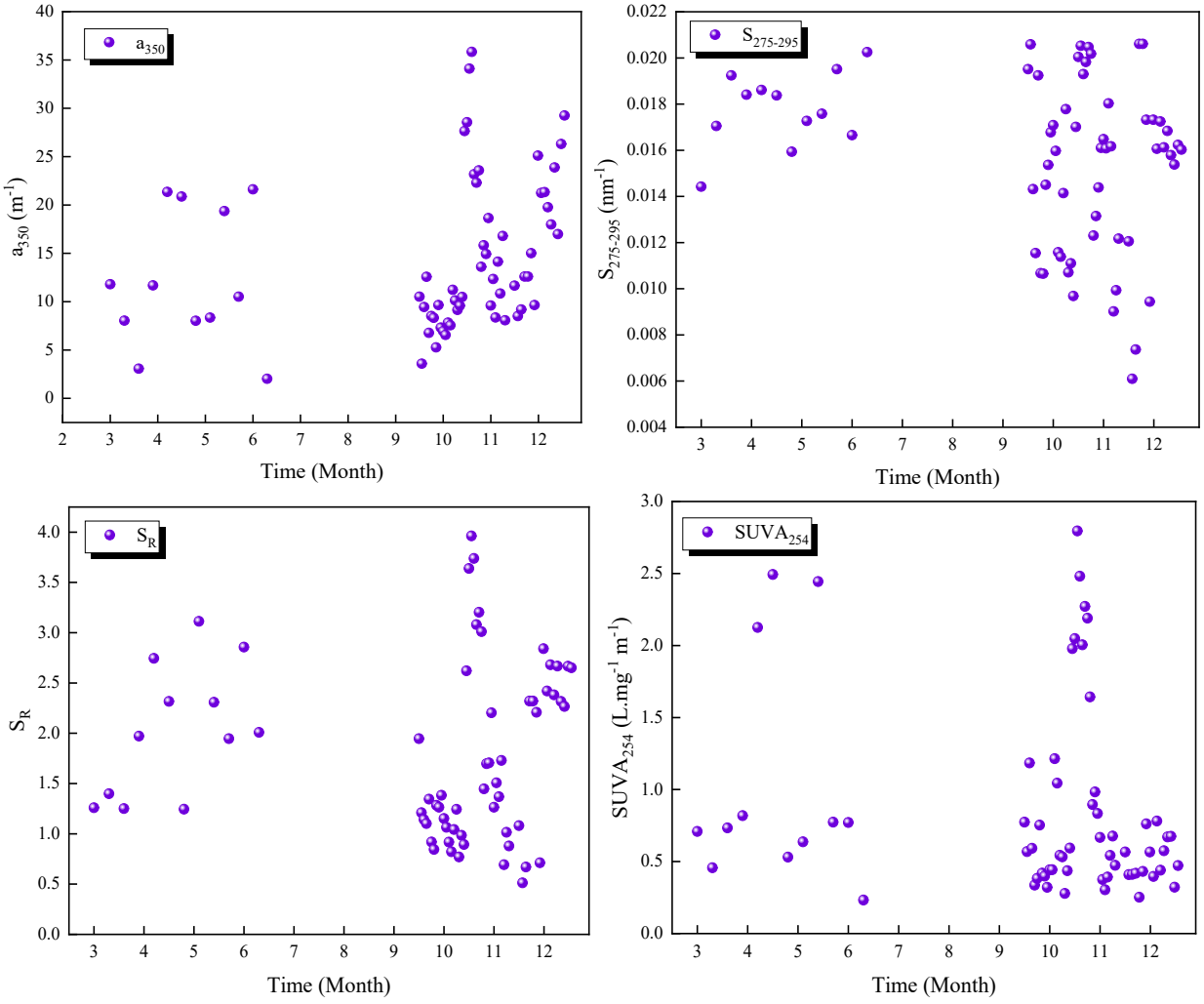

Fig. S2 The CDOM optical parameters (a_{350}, $S_{275\text{-}295}$, $S_{R,}$ and $SUVA_{254}$) of urban domestic wastewater at different periods.

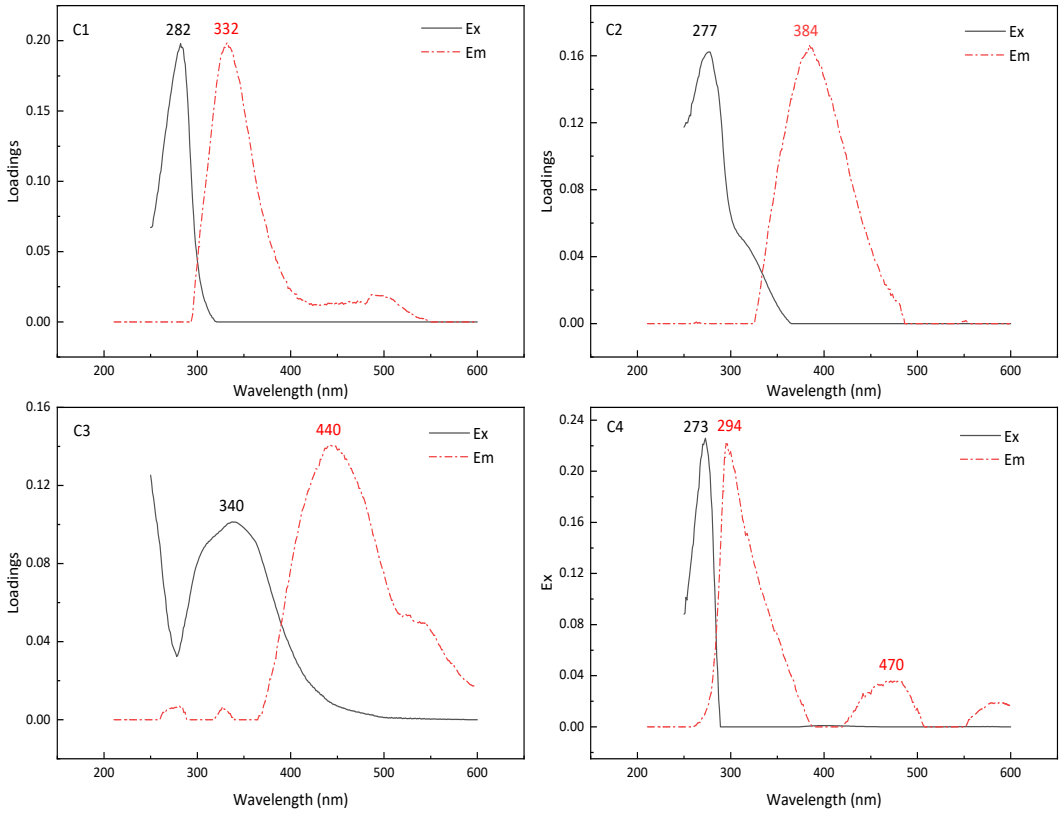

Fig. S3 The excitation and emission spectra of fluorescent components C1–C4 identified by EEM-PARAFAC. The numbers indicate the peak wavelengths.

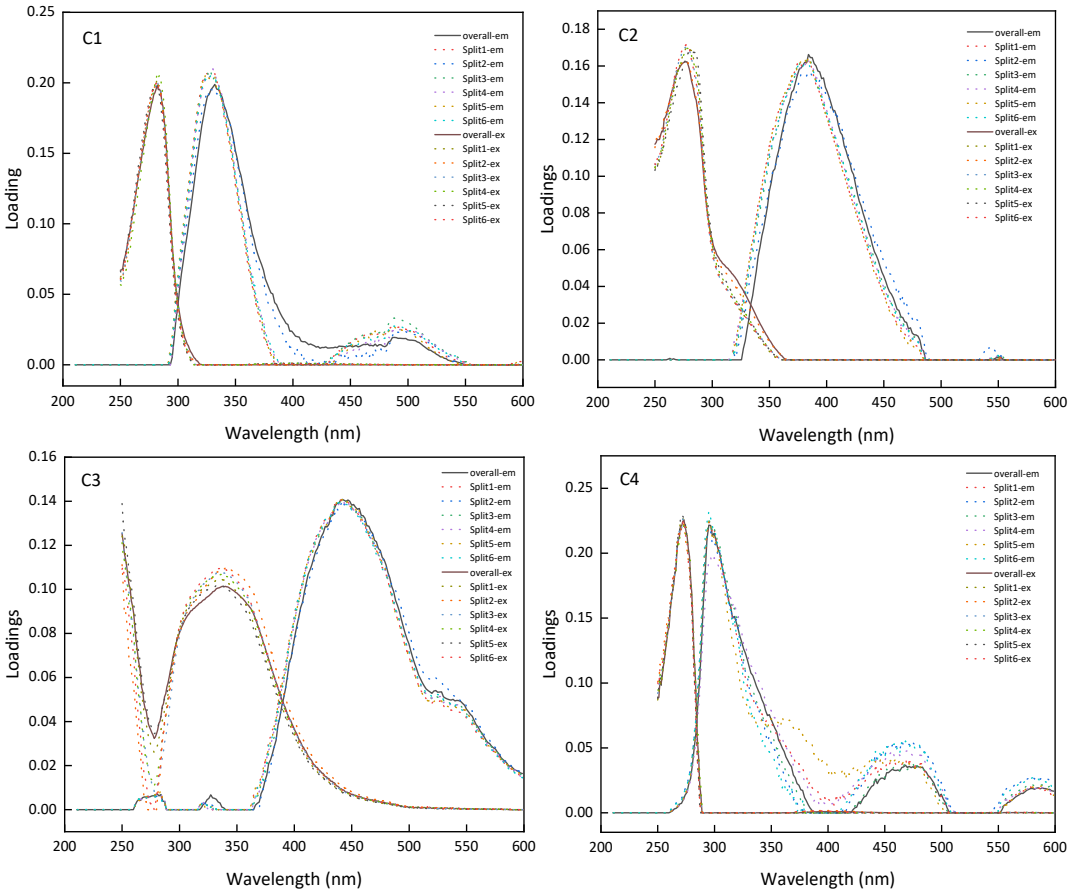

Fig. S4 Half Split-validation of the four fluorescent components of the urban domestic wastewater.

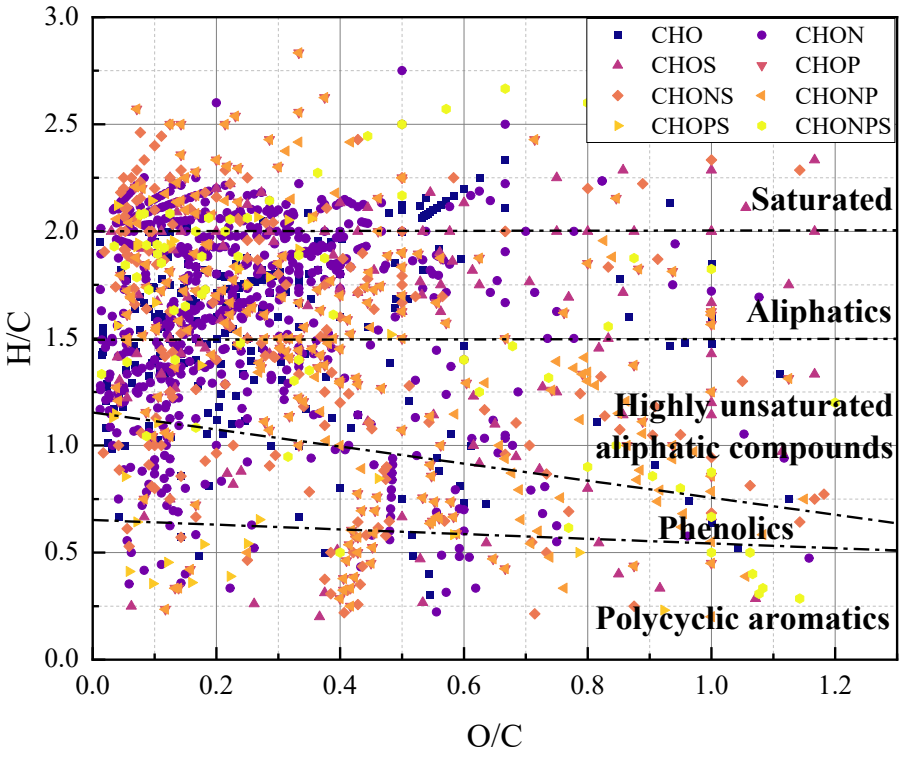

Fig. S5 The Van Krevelen diagram of compound classes in DOM from urban domestic wastewater. Lines separating compound categories are for visualization and exact categorization may differ.

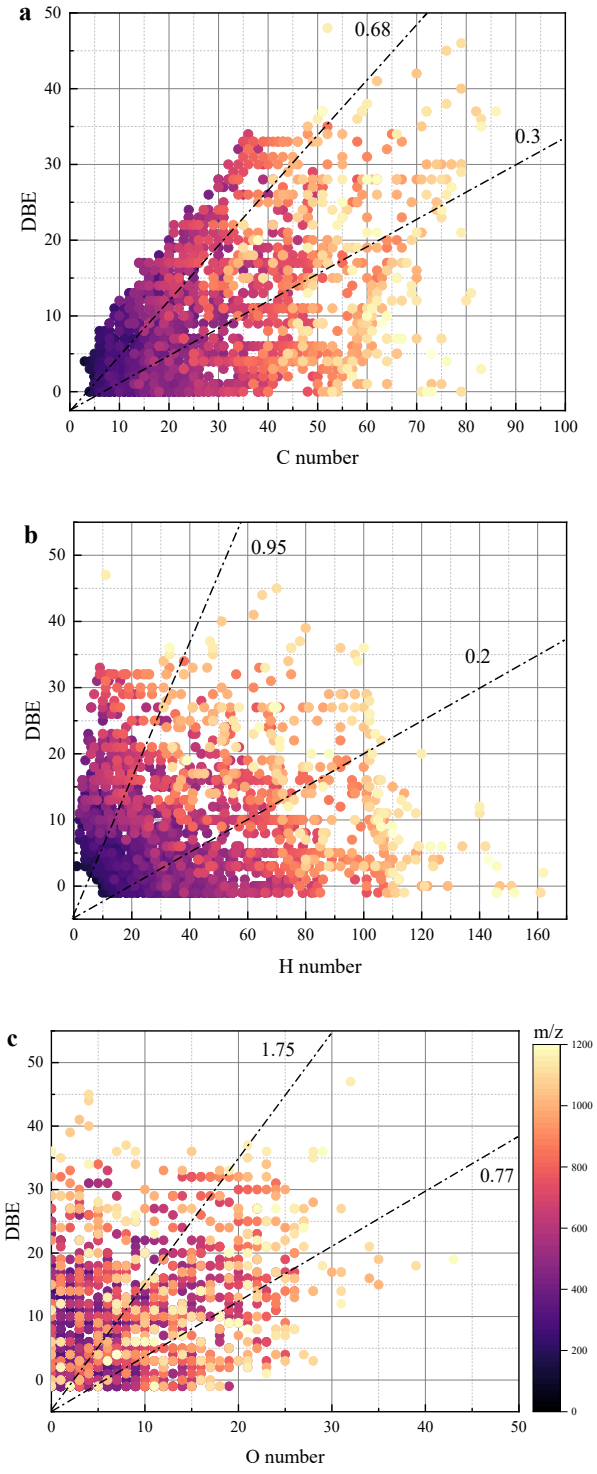

Fig. S6 Molecular DBE vs. C (a), H (b), and O (c) for all molecular formulae. The color scale represents the m/z ratio.

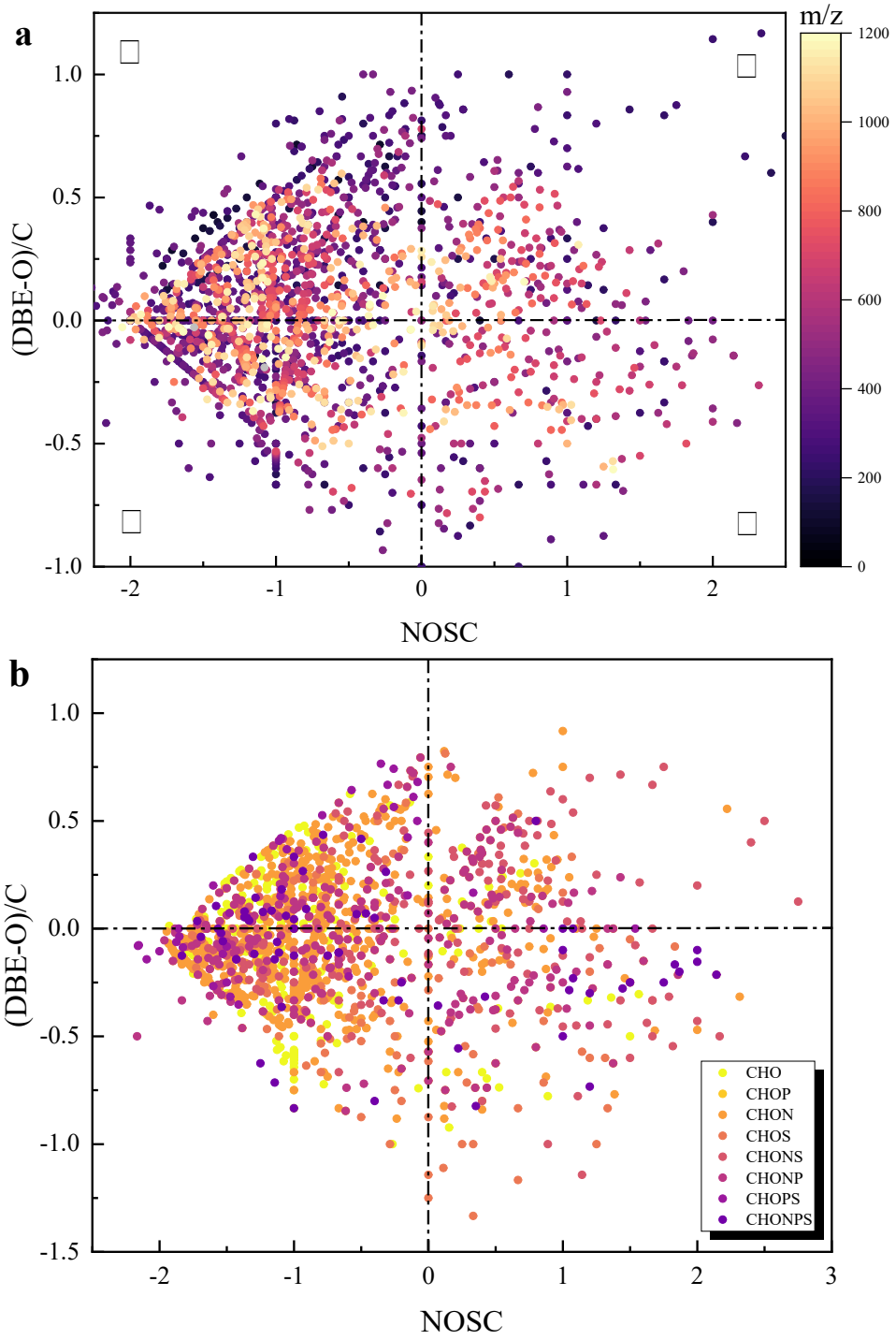

Fig. S7 NOSC vs (DBE-O)/C diagrams of urban domestic wastewater-derived DOM (the third dimension is respectively (a) m/z and (b) molecular composition). The four regions represent different compound classes: ①unsaturated and oxidized compounds, ②unsaturated and reduced compounds, ③saturated and reduced compounds, and ④saturated and oxidized compounds.

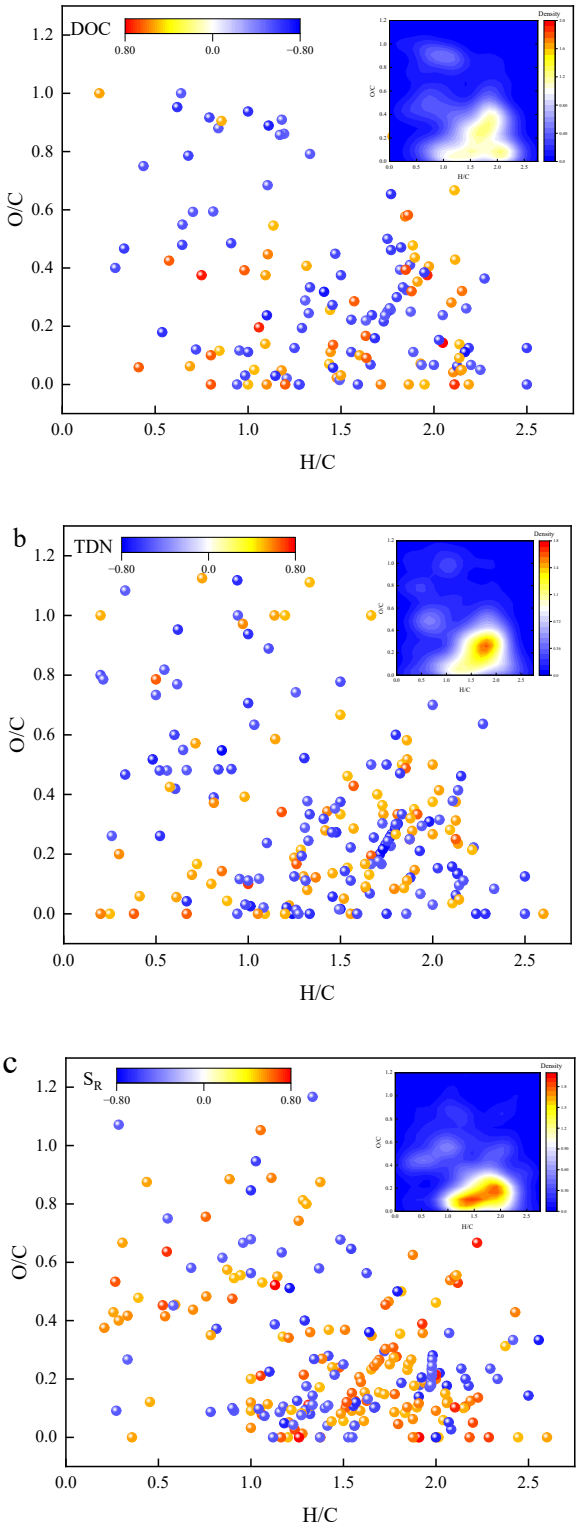

Fig. S8 The correlation between photochemical parameters (DOC, TDN, and S_R) and mass spectrometry molecules ($p<0.05$), as well as the distribution density of correlated molecules.

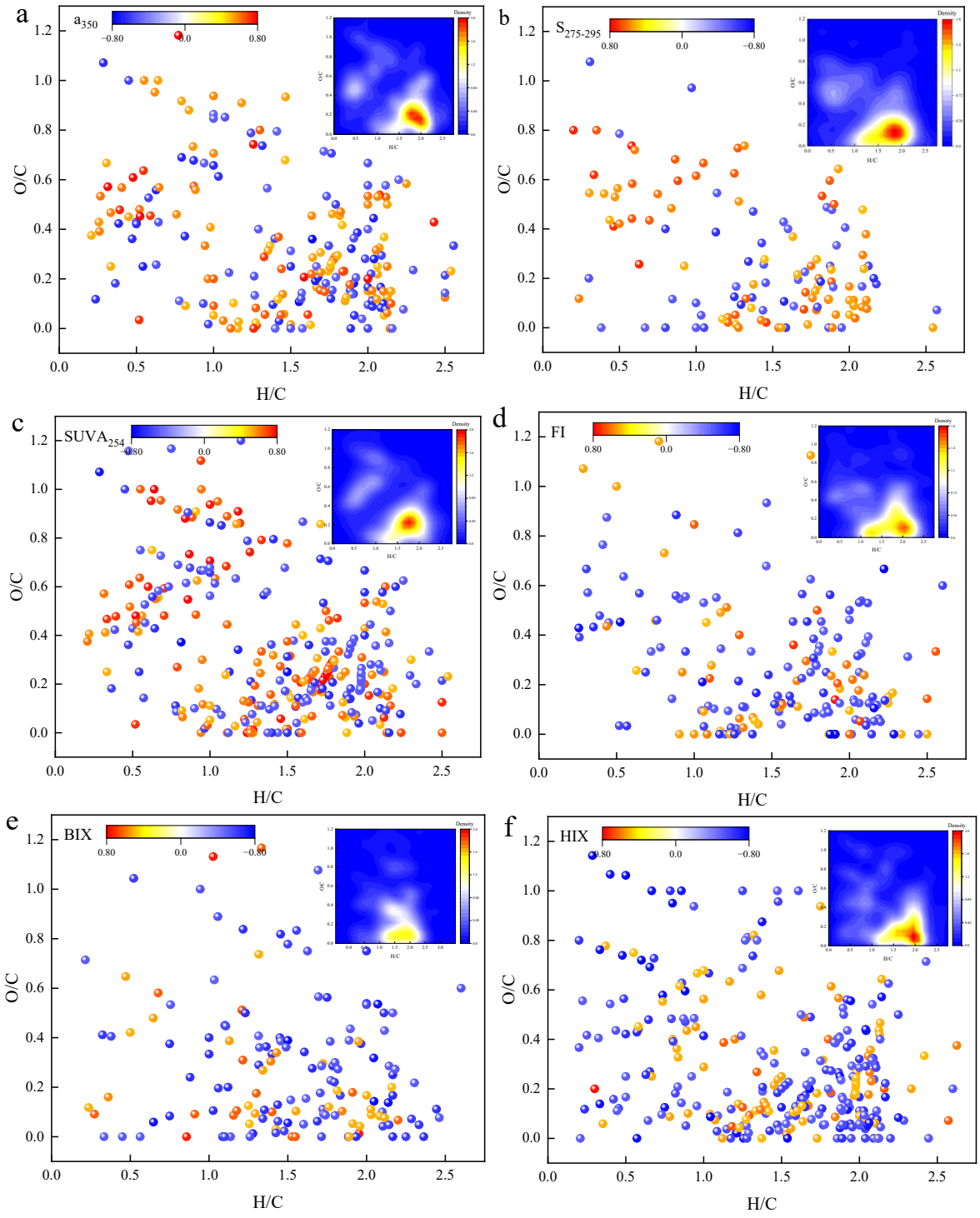

Fig. S9 The correlation between optical parameters (a: a_{350}, b: $S_{275\text{-}295}$, c: $SUVA_{254}$, d: FI, e: BIX, f: HIX) and molecules detected by UPLC-Q-TOF-MS ($p<0.05$), as well as the distribution density of correlated molecules.

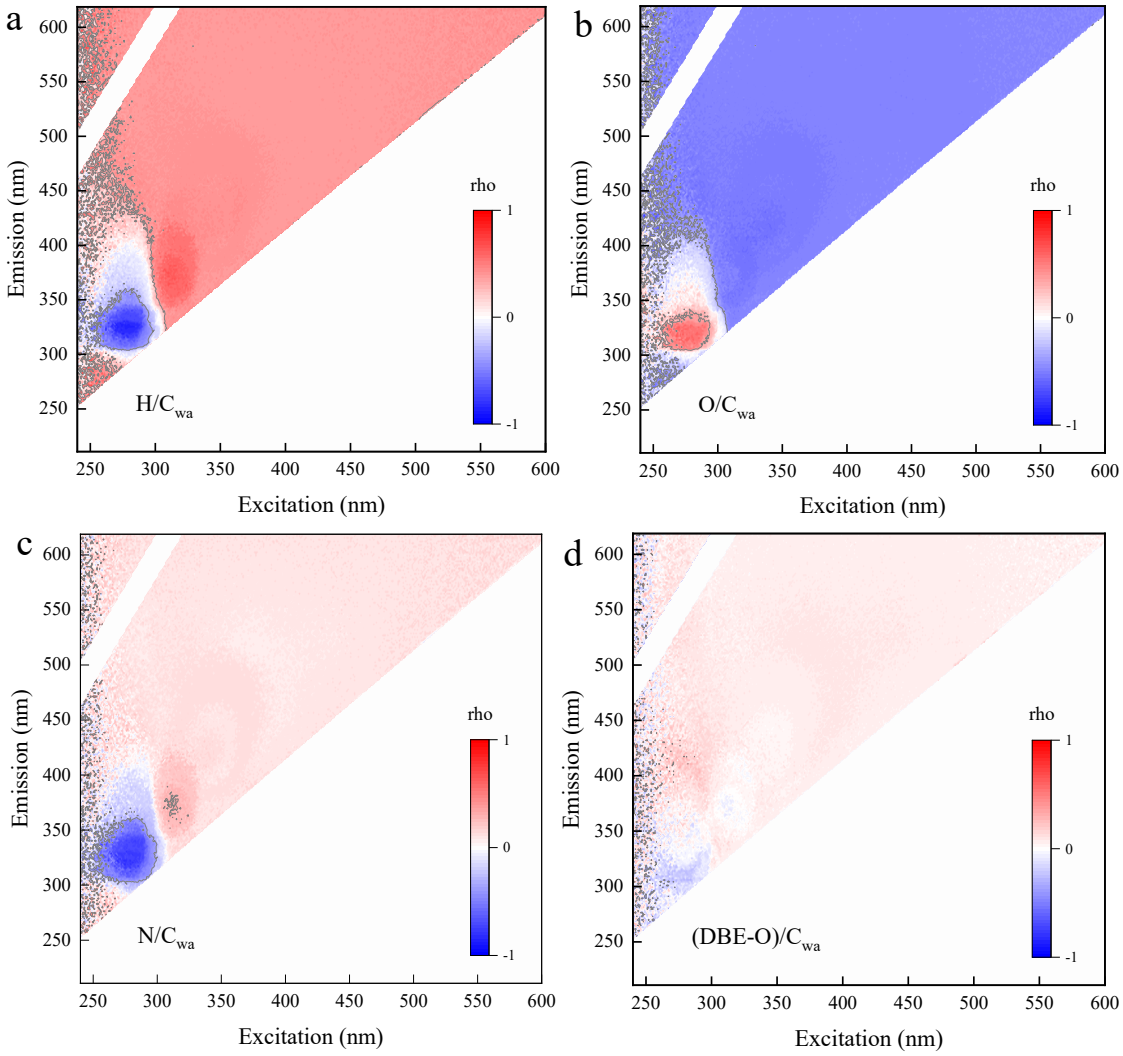

Fig. S10 The correlation between molecular indices (a: H/C_{wa}, b: O/C_{wa}, c: N/C_{wa} and d: $(DBE-O)/C_{wa}$) of DOM in domestic wastewater detected by UPLC-Q-TOF-MS and emission-excitation matrices (wa means weight average based on the molecules intensity).

References

1 Z. Liao, J. Chu, C. Luo and H. Chen, Revealing the characteristics of dissolved organic matter in urban runoff at three typical regions via optical indices and molecular composition, *J. Environ. Sci.*, 2021, **108**, 8–21.

2 C. Hu, F. E. Muller-Karger and R. G. Zepp, Absorbance, absorption coefficient, and apparent quantum yield: A comment on common ambiguity in the use of these optical concepts. *Limnol. Oceanogr.*, 2002, **47**, 1261–1267.

3 C. A. Stedmon, S. Markager and H. Kaas, Optical properties and signatures of chromophoric dissolved organic matter (CDOM) in Danish coastal waters. *Estuar. Coast. Shelf Sci.*, 2000, **51**, 267–278.

4 J. R. Helms, A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber and K. Mopper, Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.*, 2008, **53**, 955–969.

5 J. L. Weishaar, G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii and K. Mopper, Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon, *Environ. Sci. Technol.*, 2003, **37**, 4702–4708.

6 R. M. Cory, M. P. Miller, D. M. Mcknight, J. J. Guerard and P. L. Miller, Effect of instrument-specific response on the analysis of fulvic acid fluorescence spectra, *Limnol. Oceanogr. Methods*, 2010, **8**, 67–78.

7 D. M. McKnight, E. W. Boyer, P. K. Westerhoff, P. T. Doran, T. Kulbe and D. T. Andersen, Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity, *Limnol. Oceanogr.*, 2001, **46**, 38–48.

8 A. Huguet, L. Vacher, S. Relexans, S. Saubusse, J. M. Froidefond and E. Parlanti, Properties of fluorescent dissolved organic matter in the Gironde Estuary, *Org. Geochem.*, 2009, **40**, 706–719.

9 T. Ohno, Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter, *Environ. Sci. Technol.*, 2002, **36**, 742–746.

10 A. Zsolnay, E. Baigar, M. Jimenez, B. Steinweg and F. Saccomandi, Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying, *Chemosphere*, 1999, **38**, 45–50.

11 K. R. Murphy, C. A. Stedmon, D. Graeber and R. Bro, Fluorescence spectroscopy and multi-way techniques. PARAFAC, *Anal. Methods*, 2013, **5**, 6557–6566.

12 C. A. Stedmon and R. Bro, Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial, *Limnol. Oceanogr. Methods*, 2008, **6**, 572–579.

13 C. J. Williams, Y. Yamashita, H. F. Wilson, R. Jaffe and M. A. Xenopoulos, Unraveling the role of land use and microbial activity in shaping dissolved organic matter characteristics in stream ecosystems, *Limnol. Oceanogr.*, 2010, **55**, 1159–1171.

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

15 D. E. LaRowe and P. Van Cappellen, Degradation of natural organic matter: A thermodynamic analysis, *Geochim. Cosmochim. Acta*, 2011, **75**, 2030–2042.

16 M. Seidel, M. Beck, T. Riedel, H. Waska, I. G. N. A. Suryaputra, B. Schnetger, J. Niggemann, M. Simon and T. Dittmar, Biogeochemistry of dissolved organic matter in an anoxic intertidal creek bank, *Geochim. Cosmochim. Acta*, 2014, **140**, 418–434.