Electronic Supplementary Material (ESI) for Environmental Science: Processes & Impacts.

Brown carbon aerosols in the Indo-Gangetic Plain outflow: insights from excitation emission matrix (EEM) fluorescence spectroscopy

Supriya Dey[1], Arya Mukherjee[2], Anuraag J. Polana[1], Archita Rana[1], Jingying Mao[3], Shiguo Jia[4,5], Amit K. Yadav[6], Pandit. S. Khillare[6], Sayantan Sarkar[1,7*]

[1]Department of Earth Sciences, Indian Institute of Science Education and Research (IISER) Kolkata, Mohanpur741246, Nadia, India.

[2]Department of Chemical Sciences, Indian Institute of Science Education and Research (IISER) Kolkata, Mohanpur741246, Nadia, India.

[3]Institute for Environmental and Climate Research, Jinan University, Guangzhou 510632, PR China

[4]Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University, Guangzhou 510275, PR China.

[5]School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, 510275, PR China.

[6]School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110067, India.

[7]School of Engineering, Indian Institute of Technology (IIT) Mandi, Kamand, Himachal Pradesh 175075, India.

*Corresponding author at: Room No. F8, Building A8, Indian Institute of Technology (IIT) Mandi, Kamand, Himachal Pradesh 175075, India; Phone: +91-1905-267829; Email: sayantan@iitmandi.ac.in

Supplementary Information

Contains 18 pages, 2 tables, 8 figures

1. Methodological details

1.1 Pre-treatment of fluorescence data

Fluorescence EEM data were preprocessed by instrumental bias correction, inner filter correction, blank subtraction and Raman normalization according to Murphy et al.[1]

(a) Instrumental bias correction: The fluorescence light intensity varies with time, instrumental configuration, air temperature, etc., which causes spectral biases in the measured fluorescence data. To reduce this effect, instrument-specific correction factors offered by the instrument manufacturer are automatically applied to the raw fluorescence data. The final data are reported in the S1c/R1c (corrected signal/corrected reference) mode.

(b) Inner filter correction: At shorter wavelengths, some fluorophores absorb light, constraining them to be fully excited. They could also absorb the light from emission, which underestimates the final fluorescence data. These effects are referred to as inner filter effects (IFE).[1,2] To avoid this effect, the following standard method is implemented based on the measured light absorption data of fluorophores:

$$F^{IFE} = F^{Ori}.10^{0.5(A_{\lambda ex} + A_{Em})}$$

Where F^{IFE} and F^{ori} refer to the IFE corrected and original fluorescence data obtained from the spectrophotometer, respectively. The value of 0.5 is half of the optical path length of the cuvette (usually 1 cm). The parameters $A_{\lambda ex}$ and $A_{\lambda em}$ correspond to the absorbance of excitation and emission light at a certain wavelength (λ), which can be obtained by a UV-visible spectrometer.

(c) Raman normalization: Raman scattering is an inelastic scattering due to light absorption and re-emission from vibrational states,[3] which disturbs the detected fluorescence signal. It appears in the excitation-emission matrix as a diagonal line. To eliminate the Raman effect, the fluorescence

data for samples and blanks are normalized to the Raman data collected on the same day, and the corrected fluorescence data (F^{Raman}) is reported in Raman Units (*RU*). For BrC_{aq} in this study, fluorescence data (F^{IFE}) was calibrated by the Raman peak area (A_{rp}^{350}), which was derived from the integrated water Raman ($WR_{350,\lambda em}$) between wavelengths of 381 and 426 nm at an excitation wavelength of 350 nm:

$$F^{Raman} = \frac{1}{A_{rp}^{350}} \times F^{IFE}$$

$$A_{rp}^{350} = \int_{381}^{426} WR_{350,\ \lambda em} d\lambda_{em}$$

For methanol-soluble BrC, the IFE corrected fluorescence data (F^{IFE}) was normalized by the Raman peak value (a_{rp}^{350}) of the methanol blank sample at the excitation wavelength of 350 nm:

$$F^{Raman} = \frac{1}{a_{rp}^{350}} \times F^{IFE}$$

(d) Blank subtraction: Blank samples were pre-treated following the same above, and were subtracted from the fluorescence data for each sample.

1.2 PARAFAC analysis

After correcting the fluorescence data, parallel factor analysis was performed to decompose the EEM into their underlying chemical composition as follows:[4]

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}$$

where, i = 1,… I; j = 1,… J; k = 1,… K.

Here, x_{ijk} represents the fluorescence intensity of the i^{th} sample of emission wavelength j and excitation wavelength k; a_{if} is proportional to the concentration of the f^{th} analyte of the i^{th} sample; b_{jf} and c_{kf} are the

scaled estimates of the emission and excitation spectra at wavelengths *j* and *k* for the f^{th} analyte; and e_{ijk} is the residual error for the element x_{ijk}. As BrC is a complex mixture of fluorophores with variable efficiencies at absorbing and converting incident radiation to fluorescence, it is usual to track the fluorescence intensity at the maximum (F_{max}; Raman units, R.U.) for each component. F_{max} is calculated by multiplying the maximum excitation loading and maximum emission loading for each component by its score.

PARAFAC model was run using the drEEM (ver.0.1.0) and N-way toolbox (https://www.Models.life.ku.dk/drEEM) in MATLAB 2018a. To remove first order Rayleigh and Raman scatter, an interpolation scheme was applied.[4,5] To reduce concentration-related collinearity and to avoid extremely different leverages across the datasets, normalization to unit norm was performed. However, after validating the model, normalization was reversed to its unscaled fluorescence intensity (R.U.).[4]

Two individual PARAFAC models were run for BrC_{aq} and BrC_{me}. For BrC_{aq} (BrC_{me}), a 3 (2)-component model was adopted from 2- to 5-component PARAFAC models in the exploratory phase by evaluating the shape of spectral loading (Fig. S2), core consistency (Fig. S3), residual analysis (Fig. S4), and split half analysis (Fig. S5). The core consistency diagnostic test is used to determine the appropriate number of components. A high core consistency value (~100%) indicates that the model is under-fitting while an abrupt drop to low values (<50%) suggests over-fitting. In our study, the core consistency value was 63.3% (99.8%) of BrC_{aq} (BrC_{me}) with an explained variance of 99.7% (98.7%) for the 3 (2)-component model (Fig. S3). The residual error calculated from the difference between measured and modeled EEM was less than 5% of the measured EEM (Fig. S4). The total sum of squared errors (SSE) of each component (2 to 5) along with iteration, explained variance, core consistency and component size for the BrC_{aq} and BrC_{me} models are presented in Table S2. The 2- to 5-component models were run individually 3 (5) times for BrC_{aq} (BrC_{me}) with random initializations in the non-negativity constrained mode with a convergence criterion of 10^{-8},

and the solution with the minimum error was considered. Finally, the optimum number of PARAFAC components was determined by split-half analysis, which is a very powerful tool to evaluate the robustness of the model. In this method, the EEM data are first divided into four splits (A, B, C, D) alternatingly, and later assembled into six combined splits (AB, AC, AD, BC, BD and CD). Each combined split includes half samples of the EEM data and three paired validation tests can be produced (AB vs. CD, BC vs. AD, AC vs. BD). Each combined split is independently decomposed using the PARAFAC model, from which the resolved excitation and emission spectra are compared for each of the three paired validation tests. If near identical split-half results are obtained, the PARAFAC model is considered to be robust.[4,6] In this study, the split-half results were almost identical with each other (Fig. S7) which indicated the robustness of model results.

1.2 Calculation of fluorescence indices

Three types of indices: humification index (HIX), biological index (BIX) and fluorescence index (FI) were calculated to comment on potential BrC sources.[7,8] HIX was calculated as the ratio of the integrated fluorescence emission intensity in the range of 435-480 nm and 300-345 nm with the excitation wavelength of 250 nm. BIX was calculated as the ratio of emission intensity at 380 and 430 nm at the excitation wavelength of 310 nm. FI was determined by the ratio of the emission intensity at 450-500 nm at the excitation wavelength of 370 nm.

1.3 Calculation of mass absorption efficiency (MAE) and complex refractive index (k)

We have calculated MAE and k according to Liu *et al*.[9]

$$MAE(\lambda) = \frac{b_{abs\,(\lambda)}}{M}$$

$$k(\lambda)_{aq/me} = \frac{\rho\lambda.\, MAE_{aq/me}}{4\pi}$$

Where, M is the concentration of WSOC and OC for BrC_{aq} and BrC_{me} respectively.

1.4 Estimation of the water-insoluble BrC fraction

To estimate the relative contributions of water-soluble and water-insoluble fractions to total BrC, we used the approach outlined by Zhang *et al*.[10]and Satish and Rastogi:[11]

$$\%\ water-insoluble\ BrC\ (\lambda) = \frac{b_{abs_{methanol(\lambda)}} - b_{abs_{water(\lambda)}}}{b_{abs_{methanol(\lambda)}}} \times 100 \quad (1)$$

1.5 Relative radiative forcing contribution of BrC

The absorption-based relative radiative forcing contribution of BrC_{aq} and BrC_{me} with respect to EC were calculated for wintertime based on the method outlined in Kirillova *et al.*[12]

$$f = \frac{\int I_0(\lambda)\left\{1 - e^{-(MAE_{365} * \left(\frac{365}{\lambda}\right)^{AAE} * C_{WSOC/OC} * H_{ABL})}\right\} d\lambda}{\int I_0(\lambda)\left\{1 - e^{-(MAE_{550} * (\frac{550}{\lambda}) * C_{EC} * H_{ABL})}\right\} d\lambda} \quad (3)$$

Here, MAE is the mass absorption efficiency at reference wavelengths of 365 nm for BrC_{aq} and BrC_{me}, and 550 nm for EC; AAE is the absorption Ångstrom exponent; $C_{WSOC/OC}$ is the concentration of WSOC or OC; C_{EC} is the concentration of elemental carbon (EC); H_{ABL} is the height of the atmospheric boundary layer. For EC, MAE_{550} was considered to be 7.5 $m^2\ g^{-1}$ and AE was set as 1 as per existing literature.[12–14] The solar emission spectrum (i.e., $I_0(\lambda)$) was calculated using the Air Mass 1 Global Horizontal (AM1GH) irradiance model while H_{ABL} was considered to be 1000 m.[15] The equation was solved for three wavelength ranges: 300-2500 nm, 300-700 nm, and 300-400 nm.

Table S1. Descriptive statistics (mean ± σ) of carbonaceous aerosols, ionic species, and UV-Vis and fluorescence parameters for the overall study period (n=27) in the eastern IGP during the winter of 2018-2019. Here, AE, MAE and k stand for the Angstrom exponent, mass absorption efficiency, and the complex component of the refractive index, respectively. HIX, BIX and FI stand for humification, biological, and fluorescence indices respectively.

Carbonaceous aerosol	**Overall (n=27)**
OC (μg m^{-3})	38.8 ± 10.3
EC (μg m^{-3})	16.0 ± 7.7
WSOC (μg m^{-3})	18.8 ± 5.0
WSOC/OC	0.52 ± 0.12
UV-Vis parameters	
$b_{abs\_365\_aq}$ (Mm^{-1})	17.9 ± 4.2
$b_{abs\_365\_me}$ (Mm^{-1})	35.9 ± 7.1
AE_{BrC_aq} (300-700 nm)	6.1 ± 0.1
AE_{BrC_aq} (300-700 nm)	5.7 ±0.1
MAE_{365_aq} (m^2 g^{-1})	0.93 ± 0.14
MAE_{365_me} (m^2 g^{-1})	0.97 ± 0.2
k_{365_aq}	0.041 ± 0.006
k_{365_me}	0.043 ± 0.009
$(E_2/E_3)_{aq}$	6.4 ± 0.4
$(E_2/E_3)_{me}$	6.8 ± 0.3
Fluorescence parameters	
$HIX_{\_aq}$	4.8 ± 0.3
$HIX_{\_me}$	3.1 ± 0.3
$BIX_{\_aq}$	0.93 ± 0.03
$BIX_{\_me}$	0.95 ± 0.05
$FI_{\_aq}$	1.4 ± 0.03
$FI_{\_me}$	1.18 ± 0.04

Table S2. The total sum of squared errors (SSE) of each component (2 to 5) along with iteration, explained variance, core consistency and component size for the BrC_{aq} and BrC_{me} PARAFAC models.

	No. of components	SSE	Iteration	Explained variance	Core	Component size
PARAFAC details of BrC_{aq}	2	2477	46	98.73	82.77	78.4, 46.9
	3	609	274	99.68	63.30	80.3, 34.8, 20.5
	4	392	1114	99.80	16.75	49.4, 45.7, 29.7, 20.6
	5	236	756	99.87	0.22	68, 37.8, 16.5, 13.1, 14.1
PARAFAC details of BrC_{me}	2	2498	98	98.71	99.83	74.6, 53.9
	3	2127	110	98.90	84.67	70.4, 61.1, 2.4
	4	1808	74	99.07	43.21	77.6, 37.8, 28.2, 1.9
	5	1492	108	99.23	3.41	76.4, 40.8, 26.7, 1.9, 1.8

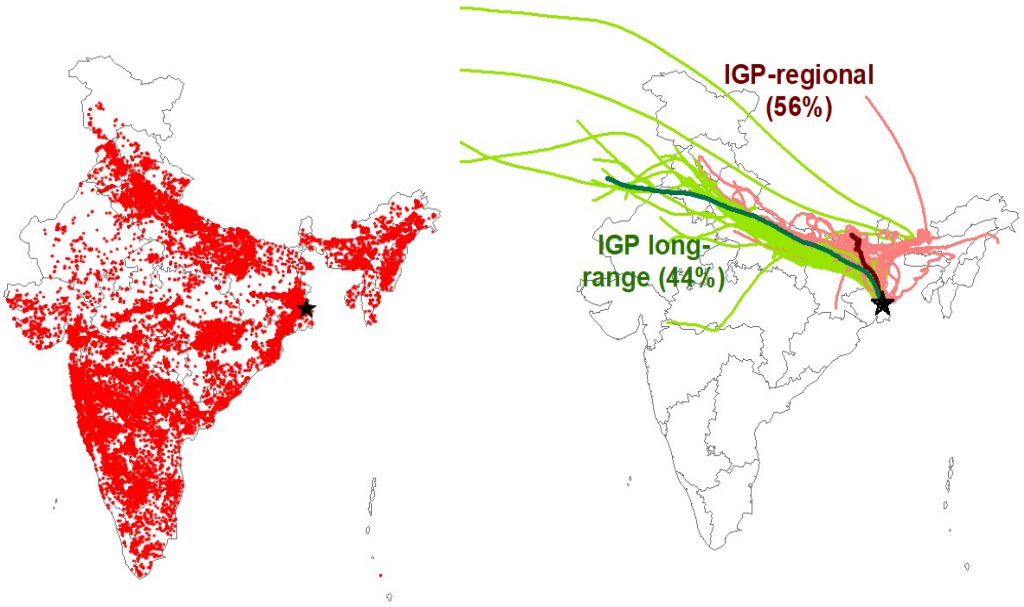

Fig. S1. Visible infrared imaging radiometer suite (VIIRS)-derived fire hotspots (375 m resolution) (http://viirsfire.geog.umd.edu/pages/mapsData.php) in India during 4th December, 2018 –14th January, 2019 (left) and air mass backward trajectory clusters (120 h) at the study site in the eastern IGP (right).

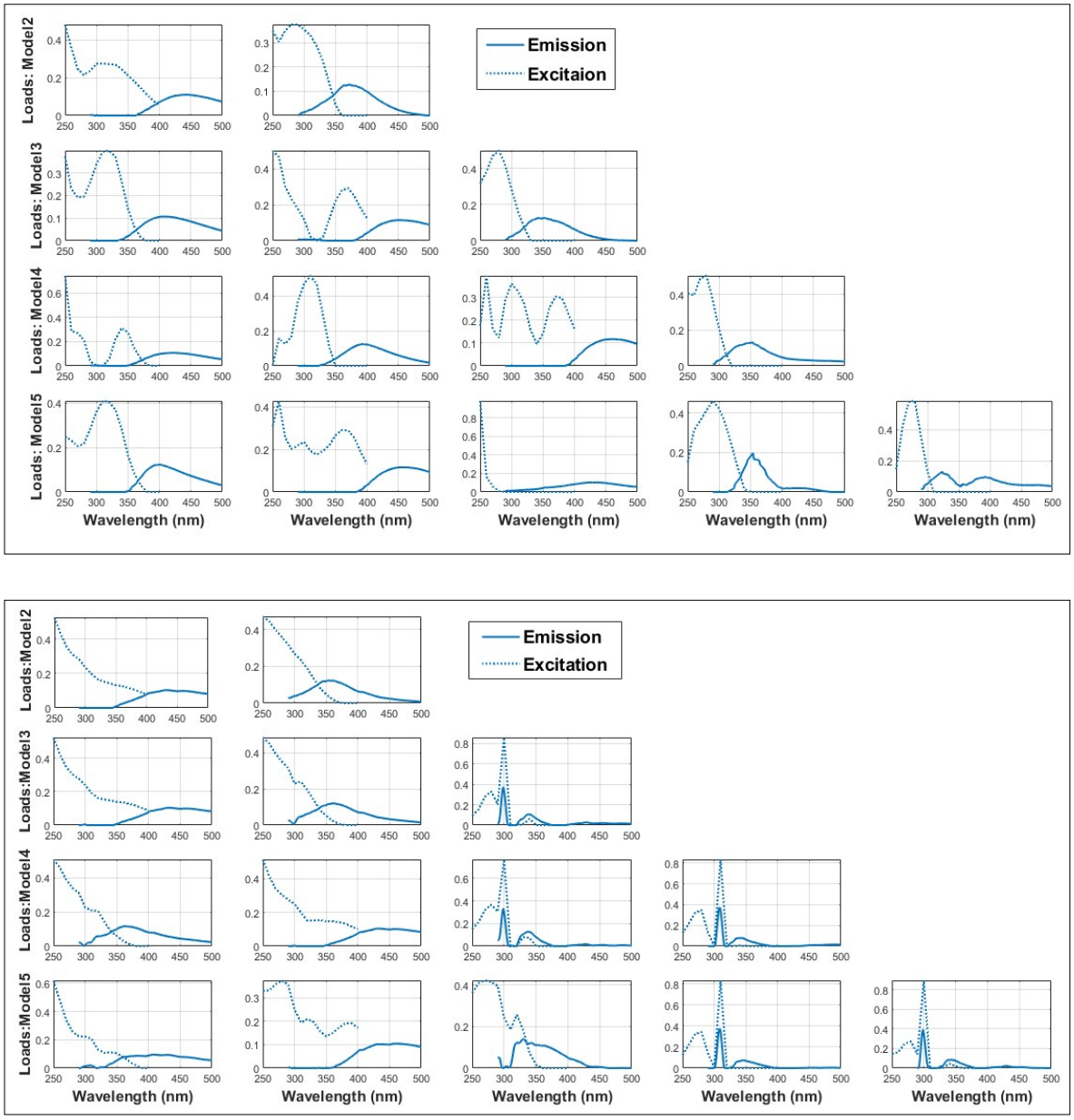

Fig. S2. Comparison of spectral loadings between 2-, 3-, 4-, and 5-component PARAFAC models of BrC_{aq} (top panel) and BrC_{me} (bottom panel).

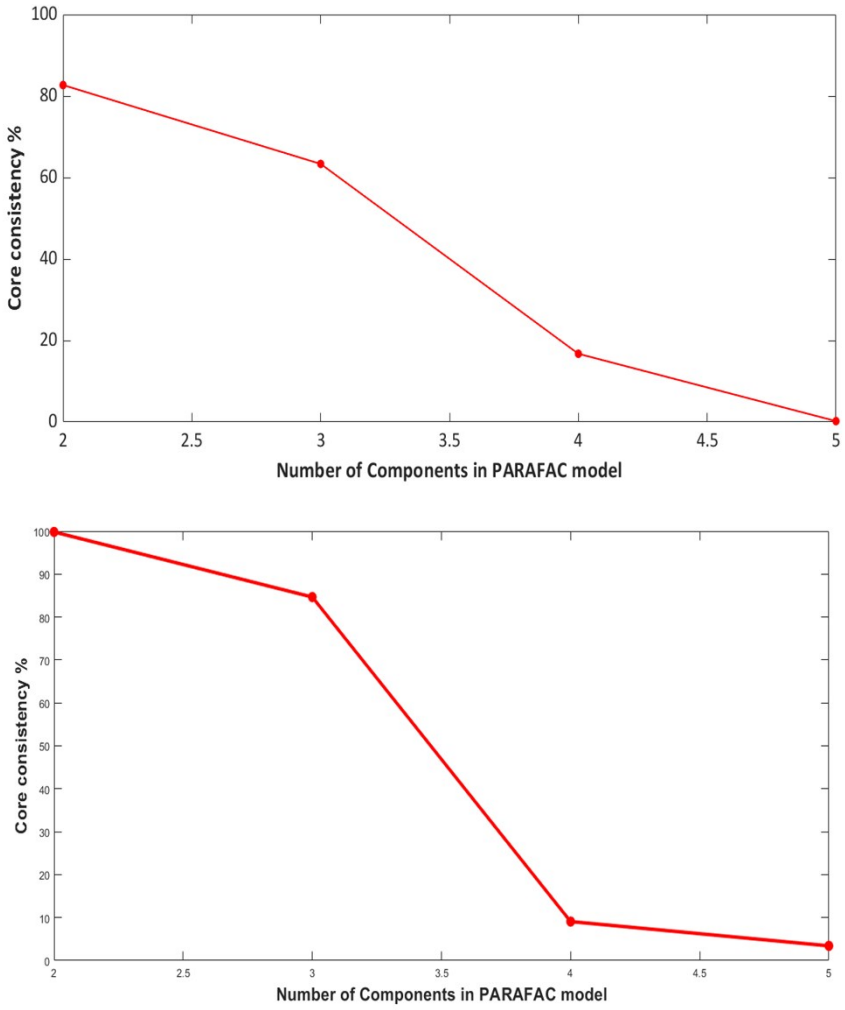

Fig. S3. Core consistency in 2-5 component PARAFAC models as a function of the number of components of BrC_{aq} (top) and BrC_{me} (bottom).

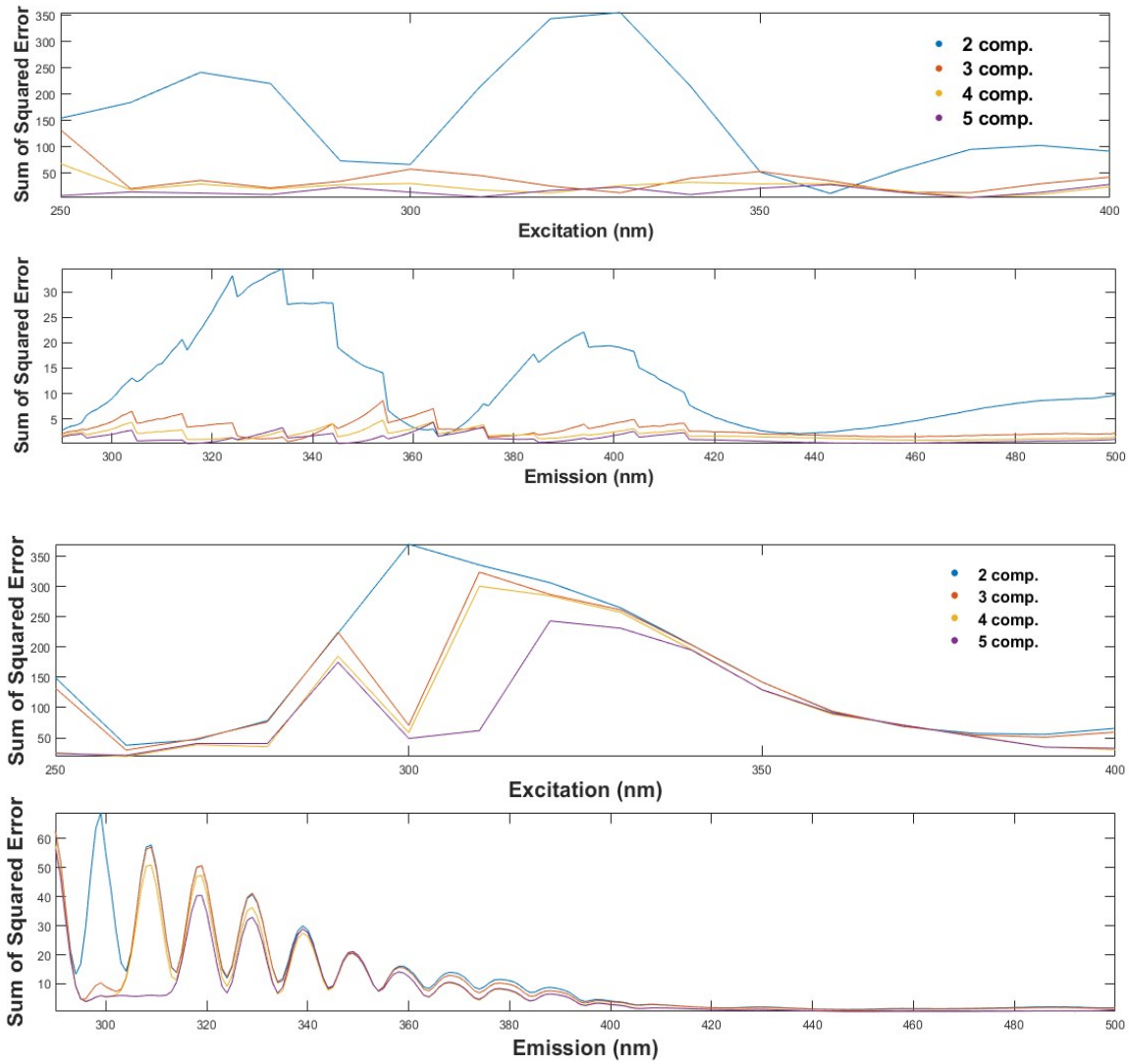

Fig. S4. Residual error analysis at individual excitation and emission wavelengths for the 2-5 component PARAFAC models for BrC_{aq} (top two plots) and BrC_{me} (bottom two plots).

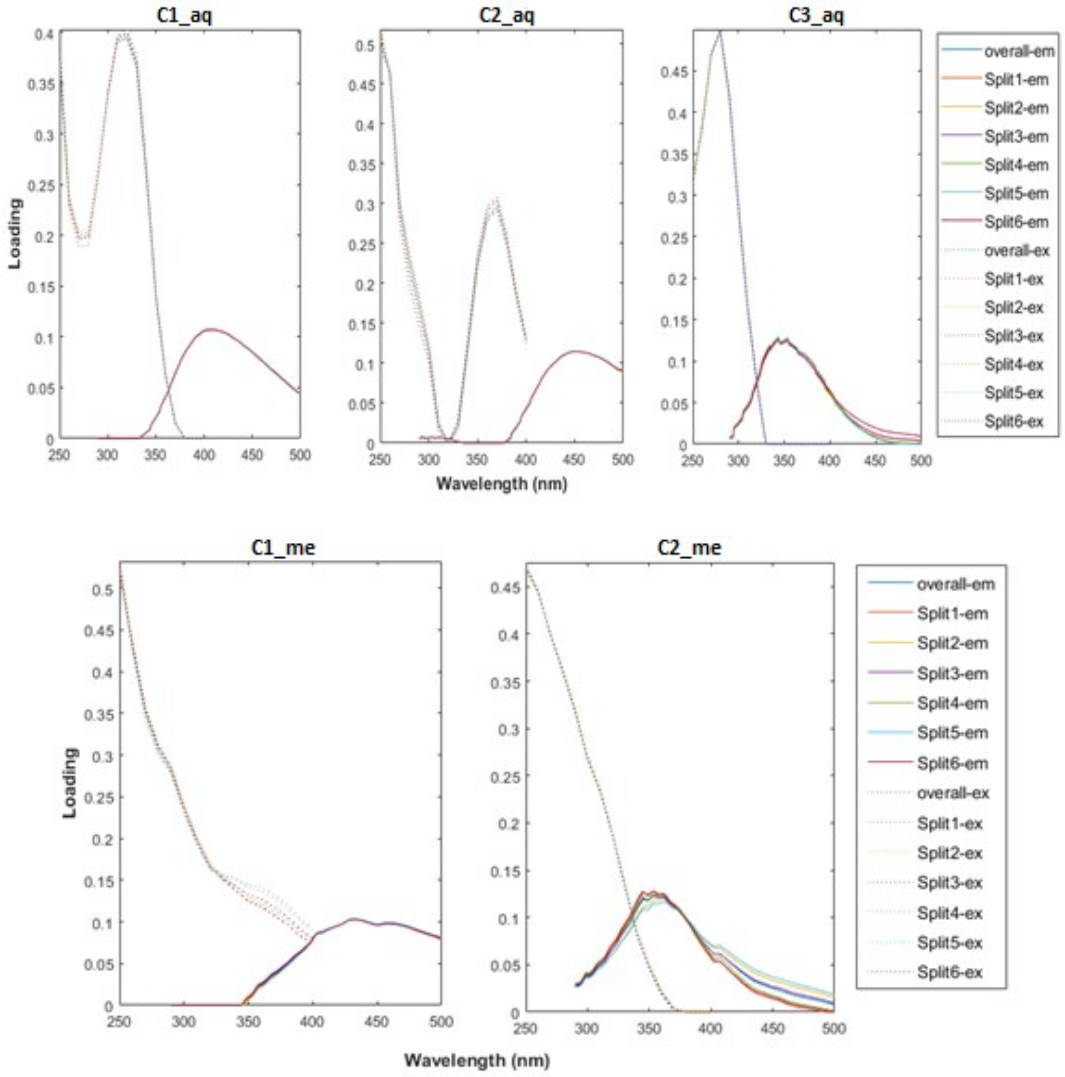

Fig. S5. Split-half analysis with split style $S_4C_6T_3$ of the 3-component PARAFAC model for BrC_{aq} (top panel) and the 2-component PARAFAC model for BrC_{me} (bottom panel).

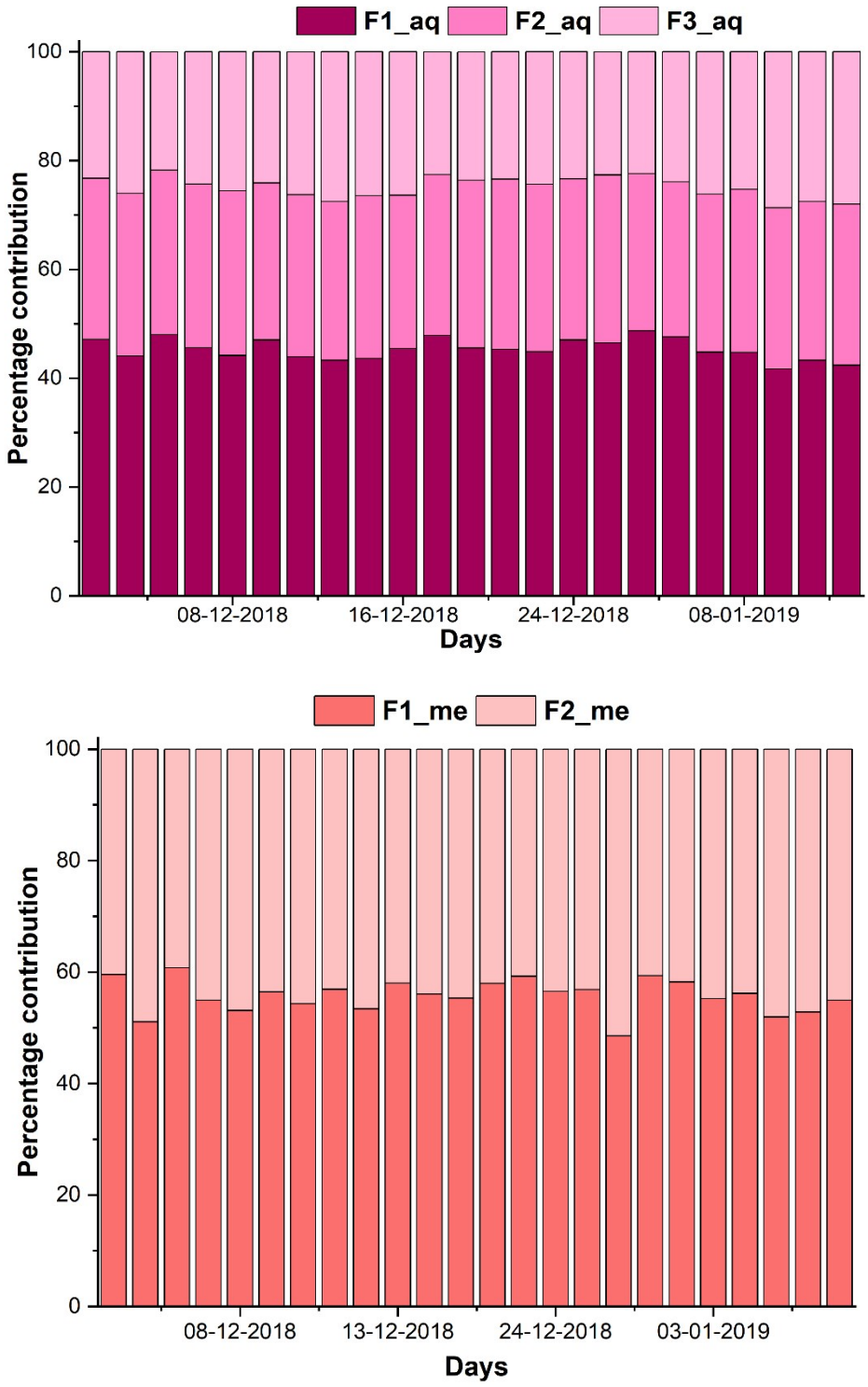

Fig. S6. Daily variation of relative contributions of the three PARAFAC-derived components for BrC_{aq} (top) and BrC_{me} (bottom) during the study period. Here, F1_aq, F2_aq and F3_aq are the fractional contributions of C1_aq, C2_aq and C3_aq, respectively, while F1_me and F2_me are the fractional contributions of C1_me and C2_me respectively.

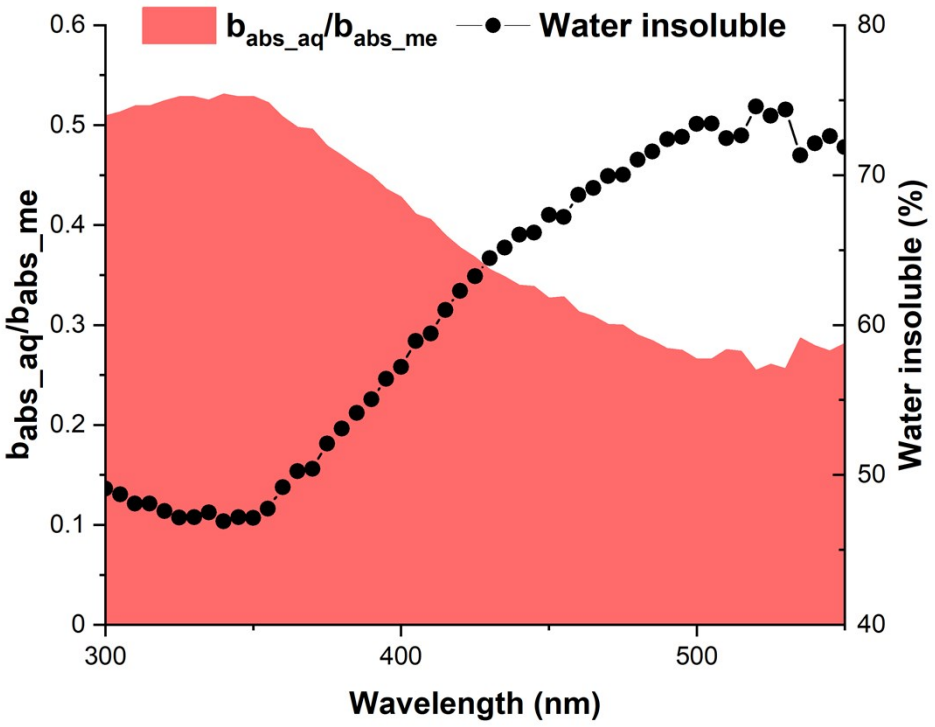

Fig. S7. Wavelength dependence (300-550 nm) of the aqueous and methanol-extractable components of b_{abs} and contribution of the water-insoluble fraction during the winter of 2018-2019.

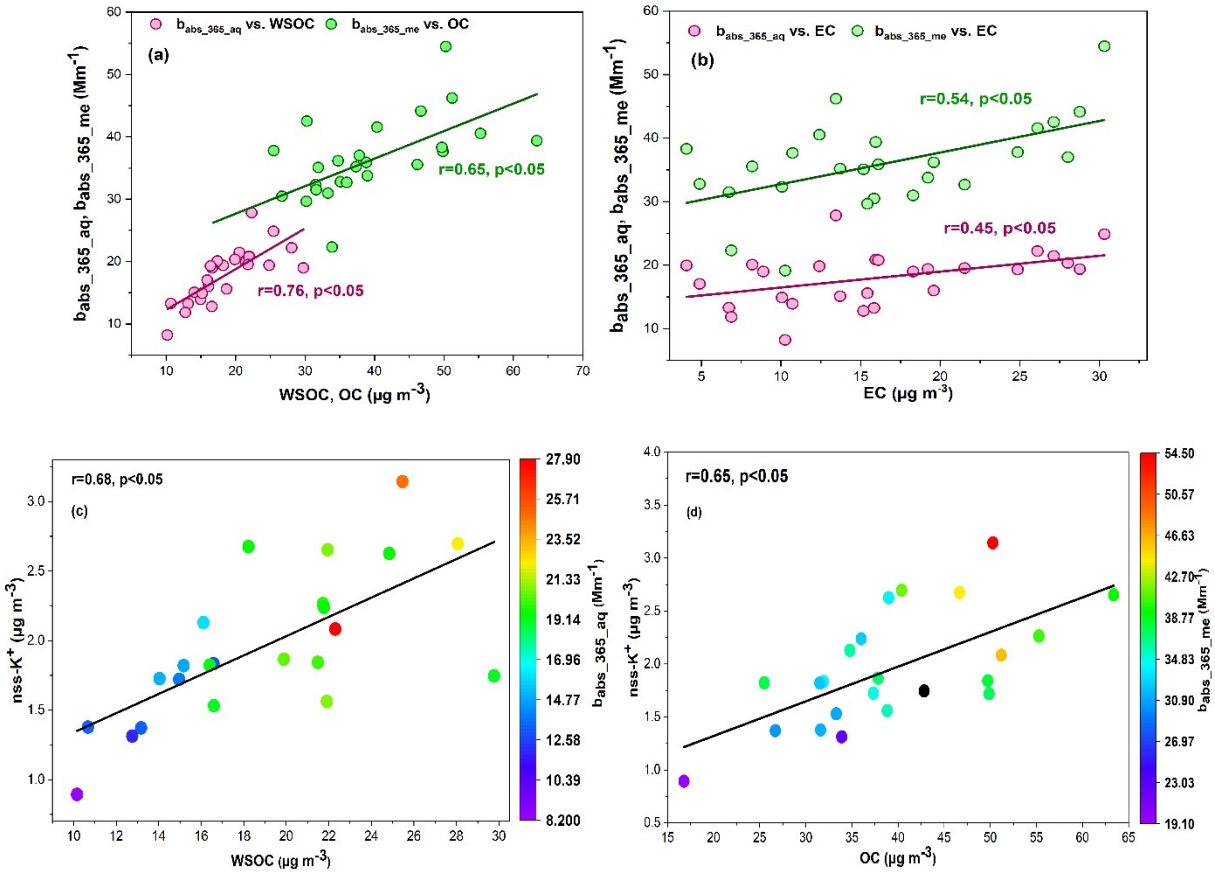

Fig. S8. Correlation of $b_{abs\_365\_aq}$ with WSOC, and $b_{abs\_365\_me}$ with OC (a); $b_{abs\_365\_aq}$ and $b_{abs\_365\_me}$ with EC (b); nss-K^+ with WSOC as a function of $b_{abs\_365\_aq}$ (c); and nss-K^+ with OC as a function of $b_{abs\_365\_me}$ (d).

References

1 K. R. Murphy, K. D. Butler, R. G. M. Spencer, C. A. Stedmon, J. R. Boehme and G. R. Aiken, Measurement of dissolved organic matter fluorescence in aquatic environments: An interlaboratory comparison, *Environ. Sci. Technol.*, 2010, **44**, 9405–9412.

2 D. N. Kothawala, K. R. Murphy, C. A. Stedmon, G. A. Weyhenmeyer and L. J. Tranvik, Inner filter correction of dissolved organic matter fluorescence, *Limnol. Oceanogr.: Methods*, 2013, **11**, 616–630.

3 E. M. Carstea, Fluorescence spectroscopy as a potential tool for in-situ monitoring of dissolved organic matter in surface water systems, *Water Pollut.*, 2012, InTech, DOI:10.5772/28979.

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

5 M. Bahram, R. Bro, C. Stedmon and A. Afkhami, Handling of Rayleigh and Raman scatter for PARAFAC modeling of fluorescence data using interpolation, *J. Chemom.*, 2006, **20**, 99-105.

6 R. Bro and H. A. L. Kiers, A new efficient method for determining the number of components in PARAFAC models, *J. Chemom.*, 2003, **17**, 274–286.

7 J. E. Birdwell and A. S. Engel, Characterization of dissolved organic matter in cave and spring waters using UV-Vis absorbance and fluorescence spectroscopy, *Org. Geochem.*, 2010, **41**, 270–280.

8 J. Qin, L. Zhang, X. Zhou, J. Duan, S. Mu, K. Xiao, J. Hu and J. Tan, Fluorescence fingerprinting properties for exploring water-soluble organic compounds in PM2.5 in an industrial city of northwest China, *Atmos. Environ.*, 2018, **184**, 203–211.

9 J. Liu, M. Bergin, H. Guo, L. King, N. Kotra, E. Edgerton and R. J. Weber, Size-resolved measurements

of brown carbon in water and methanol extracts and estimates of their contribution to ambient fine-particle light absorption, *Atmos. Chem. Phys.*, 2013, **13**, 12389–12404.

10 X. Zhang, Y. H. Lin, J. D. Surratt and R. J. Weber, Sources, composition and absorption Ångström exponent of light-absorbing organic components in aerosol extracts from the los angeles basin, *Environ. Sci. Technol.*, 2013, **47**, 3685–3693.

11 R. Satish and N. Rastogi, On the Use of Brown Carbon Spectra as a Tool to Understand Their Broader Composition and Characteristics: A Case Study from Crop-residue Burning Samples, *ACS Omega*, 2019, **4**, 1814–1853.

12 E.N. Kirillova, A. Andersson, S. Tiwari, A. K. Srivastava, D. S. Bisht, Ö. Gustafsson, Water-soluble organic carbon aerosols during a full New Delhi winter: Isotope-based source apportionment and optical properties, *J. Geophys. Res. Atmos.*, 2014, **119**, 3476-3485.

13 M. O. Andreae and A. Gelencsér, Black carbon or brown carbon? the nature of light-absorbing carbonaceous aerosols, *Atmos. Chem. Phys.*, 2006, **6**, 3131–3148.

14 Q. Chen, F. Ikemori and M. Mochida, Light absorption and excitation-emission fluorescence of urban organic aerosol components and their relationship to chemical structure, *Environ. Sci. Technol.*, 2016, **50**, 10859–10868.

15 R. Levinson, H. Akbari and P. Berdahl, Measuring solar reflectance-Part I: Defining a metric that accurately predicts solar heat gain, *Sol. Energy*, 2010, **84**, 1717–1744.