Supplementary materials for

Characterising Mass-resolved Mixing State of Black Carbon in Beijing Using a Morphology-Independent Measurement Method

Chenjie Yu[1], Dantong Liu[1,5], Kurtis Broda[2], Rutambhara Joshi[1], Jason Olfert[2], Yele Sun[3], Pingqing Fu[3,6], Hugh Coe[1], James D. Allan[1,4]

[1]School of Earth and Environmental Science, University of Manchester, Manchester, M13 9PL, UK
[2]Department of Mechanical Engineering, University of Alberta, Alberta, T6G 2R3, Canada
[3]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[4]National Centre for Atmospheric Sciences, University of Manchester, Manchester, M13 9PL, UK
[5]Now at Department of Atmospheric Sciences, School of Earth Sciences, Zhejiang University, Zhejiang, 310027, China
[6]Now at Institute of Surface-Earth System Science, Tianjin University, Tianjin, 300072, China

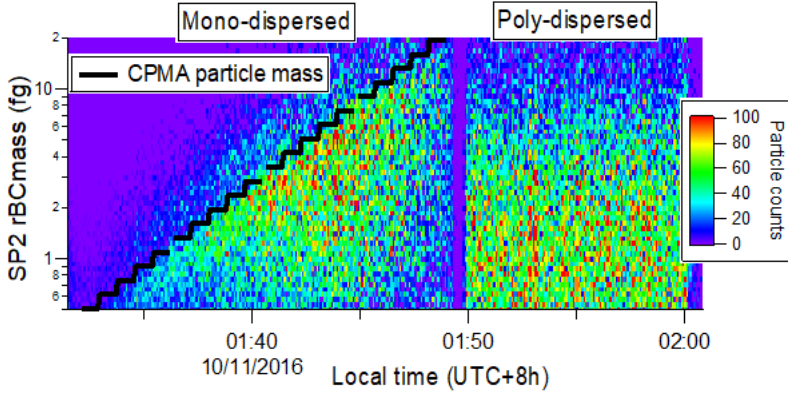*Correspondence to*: Dantong Liu (dantongliu@zju.edu.cn), James Allan (james.allan@manchester.ac.uk)

S1. CPMA-SP2 set up

Figure S1: Example for one CPMA-SP2 set up with mono-mass and poly-dispersed scan in sequence

S2. CPMA-SP2 inversion calculation

The two-variable distribution function ($\frac{\partial^2 N}{\partial logm_\mathrm{p} \partial logm_\mathrm{rBC}}$) is calculated by solving a Fredholm integral equation using a Twomey algorithm (Twomey, 1975). The number concentration distribution of rBC-containing particles can be related to the two-variable distribution for a single CPMA setpoint, i, using:

$$R_i = \frac{\partial^2 N}{\partial logm_\mathrm{p} \partial logm_\mathrm{rBC}}\bigg|_i \cdot \vec{\Gamma} \quad (1)$$

where R_i is the number distribution of SP2 response at the i th CPMA setpoint, and $\vec{\Gamma}$ is a kernel function applying a trapezoidal rule approximation for the response of the instrument. The kernel function (Collins et al., 2002) can be described as:

$$\vec{\Gamma} = \bar{\mathbb{E}}\left(m_\mathrm{p}\right) \sum_{\phi=1}^{\phi_{max}} \bar{f}\left(d_\mathrm{p}, \phi\right) \int_0^\infty \Omega\left(m_\mathrm{p}, Z\right) dlogm_\mathrm{p} \quad (2)$$

where $\bar{\mathbb{E}}\left(m_\mathrm{p}\right)$ is a correction factor accounting for particle losses in the CPMA, $\bar{f}(m_\mathrm{p}, \phi)$ is the charge fraction where ϕ represents the number of charges, and $\Omega\left(m_p, Z\right)$ is the CPMA transfer function where Z represents the electrical mobility. The

correction factor $\bar{\mathbb{E}}$ is determined by the ratio between total number concentration of the SP2 only measurement and total CPMA-SP2 scan number concentration over the same m_{rBC} range as the two-variable distribution. The correction factor $\bar{\mathbb{E}}$ can be estimated by comparing the total number concentration measured by the CPMA-SP2 to the SP2-only number concentration, where the total number concentration measured by the CPMA-SP2 is found by iterating the Twomey algorithm. To provide

5   initial guess to the inversion scheme, it was assumed that all the particles were singly charged ($\phi = 1$). Hence, the initial approximation for the two-variable distribution at scan i ($\left.\frac{\partial^2 N_{APPROX}}{\partial \log m_p \partial \log m_{rBC}}\right|_i$) is:

$$\left.\frac{\partial^2 N_{APPROX}}{\partial \log m_p \partial \log m_{rBC}}\right|_i = \frac{\left.\frac{dN}{d \log m_{rBC}}\right|_i}{\bar{f}(m_{p,i}, \phi=1)\hat{\beta}} \qquad (3)$$

10   Where $\left.\frac{dN}{d \log m_{rBC}}\right|_i$ is the number distribution of m_{rBC} of response at i th CPMA set points, and $\hat{\beta}$ the is the analytical solution to the integral of a triangular CPMA transfer function (Broda et al., 2018). After the calculation of the initial approximation, the two-variable distribution is corrected by applying the Twomey algorithm iteratively until the total number concentration of the final two-variable solution is close to the CPMA-SP2 original number concentration measurement. Finally, a weighted-average smoothing of the two variable distribution is performed to reduce noise.

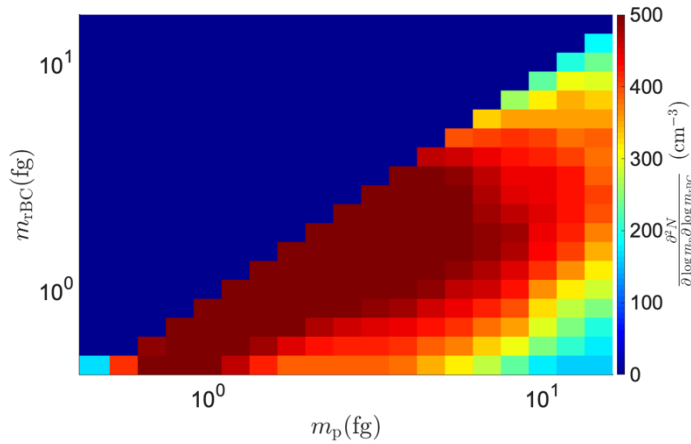15

S3. Original CPMA-SP2 two variable distribution example

Figure S2: Example for one CPMA-SP2 two variable distribution

20

S4. Mixing state of rBC-containing particles calculation

The detail of calculation is described in Riemer and West (2013). Consider there are N total rBC-containing particles in the atmosphere, each rBC-containing particle contains 2 types of material: rBC material and non-rBC material. The bins inside the two-variable distribution (see the example shown in Figure 4) represent different rBC-containing particles, and each bin contains a distinct m_p value and a distinct m_{rBC} value. The mass fractions of rBC core and non-rBC coating for rBC-containing particle i (p_i^{rBC} and $p_i^{non-rBC}$) are simply derived from the measured mass parameters:

$$p_i^{rBC} = \frac{m_{rBC}}{m_p} \qquad (4)$$

$$p_i^{non-rBC} = \frac{m_p - m_{rBC}}{m_p} \qquad (5)$$

For the mass fraction in the population, the mass fraction for total rBC (p^{rBC}) and total non-rBC ($p^{non-rBC}$) are described as:

$$p^{rBC} = \frac{M_{rBC}}{M_p} \qquad (6)$$

$$p^{non-rBC} = \frac{M_p - M_{rBC}}{M_p} \qquad (7)$$

Where M_p and M_{rBC} are the bulk rBC-containing particle mass and the bulk rBC mass which is derived from the CPMA-SP2 two-variable mass distribution results.

Compared with the previous study by Healy et al. (2014) which calculates the mass fraction parameters based on the morphology and density assumptions or the study by Bondy et al. (2018) which achieves the mass fraction through microscopy estimation, the CPMA-SP2 inversion metrics derives the mass parameters and mass fraction more directly.

The information Shannon mixing entropy is calculated via these mass fractions. The connection between the mass fraction of particle i and the Shannon mixing entropy within the rBC-containing particle i (H_i) is given:

$$H_i = -p_i^{rBC}\ln p_i^{rBC} - p_i^{non-rBC}\ln p_i^{non-rBC} \qquad (8)$$

The average particle mixing entropy can be expressed as:

$$H_\propto = \sum_{i=1}^{N} p_i H_i \qquad (9)$$

The Shannon entropy of the entire bulk population (H_γ) can be expressed as:

$$H_\gamma = -p^{\mathrm{rBC}}\ln p^{\mathrm{rBC}} - p^{\mathrm{non-rBC}}\ln p^{\mathrm{non-rBC}} \quad (10)$$

5   Later, the "diversity parameters" are used to describe the effective number of species in the single particle or the population of particles. The particle diversity for particle i is calculated from:

$$D_i = e^{H_i} \qquad (11)$$

10   The average single-particle diversity ($D_\propto$) and the bulk population diversity (D_γ) are inferred from:

$$D_\propto = e^{H_\propto} \qquad (12)$$
$$D_\gamma = e^{H_\gamma} \qquad (13)$$

15   Finally, the mixing state index (χ) term can be given from the diversity term above:

$$\chi = \frac{D_\propto - 1}{D_\gamma - 1} \qquad (14)$$

S5. Extrapolation methods

20   Firstly, the number distribution for total rBC-containing particle mass m_p has been fitted the log-normal distribution function:

$$\left.\frac{\mathrm{d}N}{\mathrm{dlog}m_\mathrm{p}}\right|_i = a_1 \cdot \exp\left(-\left(\frac{\log(m_{\mathrm{p},i})-b_1}{c_1}\right)\right) + a_2 \cdot \exp\left(-\left(\frac{\log(m_{\mathrm{p},i})-b_2}{c_2}\right)\right) \quad (15)$$

Where a_1, b_1, c_1, a_2, b_2, c_2 are all the constants for the distribution function.
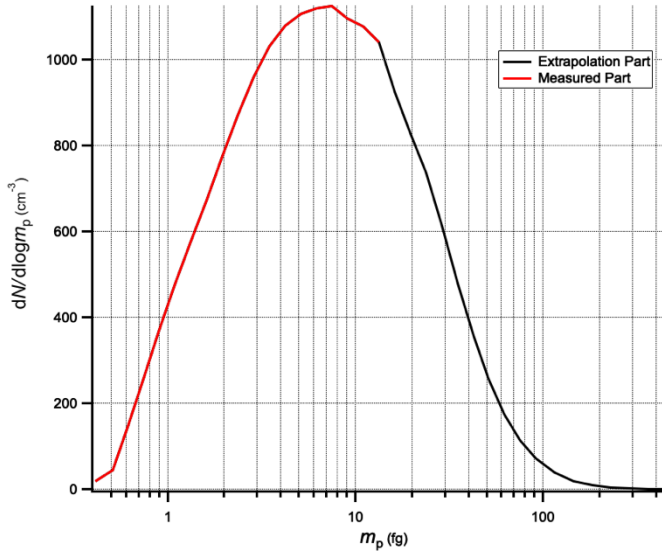
25

Figure S3: Example for the extrapolation for the size distribution of rBC-containing particle number concentration.

As the rBC-containing particle number concentration distribution is derived, the ratio (R_i) between the rBC-containing particle number concentration at the last measured bin and the number concentration at the extrapolated m_p bin i can be worked out:

$$R_i = \frac{dN/dlogm_p|_i}{dN/dlogm_p|_{m_p=15\ fg}} \quad (16)$$

Where $dN/dlogm_p|_i$ is the number concentration of rBC-containing particles at the extrapolated m_p bin i, and $dN/dlogm_p|_{m_p=15}$ is the number concentration of rBC-containing particles at the measured largest m_p bin (the bin with $m_p = 15$ fg). Then the two-dimensional function distribution $\frac{\partial^2 N}{\partial logm_p \partial logm_{rBC}}$ is extrapolated with two methods; One was based on fitting the two-variable distribution distribution as a function of m_{rBC} (Fit m_{rBC} Method), and other based on fitting the two-variable distribution distribution as a function of the mass ratio between the m_p and m_{rBC} (Fit Ratio Method):

$$\frac{\partial^2 N}{\partial logm_p \partial logm_{rBC}}\Bigg|_{Fit\ m_{rBC},(i,j)} = R_i \cdot \left(a_1 \cdot \exp\left(-\left(\frac{\log(m_{rBC,j})-b_1}{c_1}\right)\right) + a_2 \cdot \exp\left(-\left(\frac{\log(m_{rBC,j})-b_2}{c_2}\right)\right)\right) \quad (17)$$

$$\frac{\partial^2 N}{\partial logm_p \partial logm_{rBC}}\Bigg|_{Fit\ Ratio,(i,j)} = R_i \cdot \left(a_1 \cdot \exp\left(-\left(\frac{\log(m_{p,i}/m_{rBC,j})-b_1}{c_1}\right)\right) + a_2 \cdot \exp\left(-\left(\frac{\log(m_{p,i}/m_{rBC,j})-b_2}{c_2}\right)\right)\right) \quad (18)$$

6

While $\left.\dfrac{\partial^2 N}{\partial \log m_\mathrm{p} \partial \log m_\mathrm{rBC}}\right|_{(i,j)}$ represents the number concentration of two variable distribution at m_p bin i and m_rBC bin j; a_1,

b_1, c_1, a_2, b_2, c_2 are all the constants for the distribution function.

The examples for the extrapolation for the two periods are shown in Figure S4, the original measured part is on the left of the black line while the extrapolated part is on the right. Particles with larger rBC core are preferred for the Fit Ratio method.
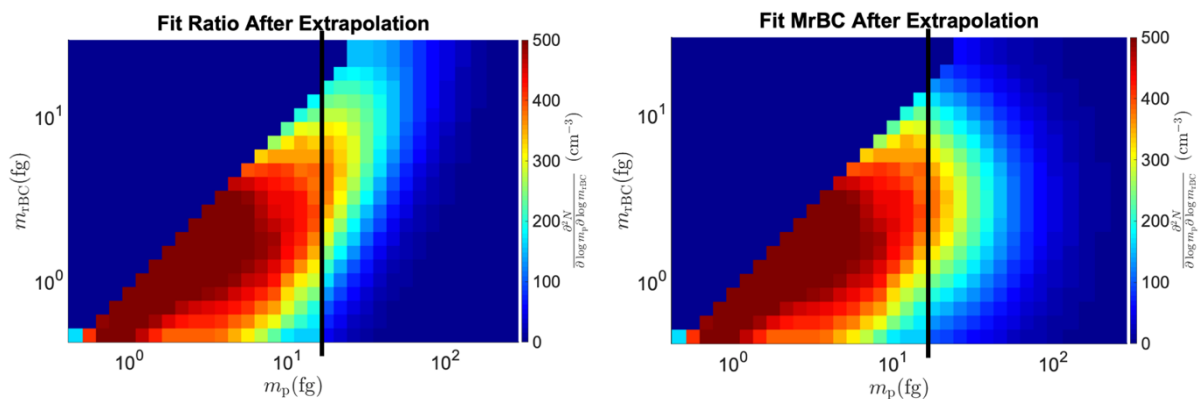

Figure S4: Examples for the two different extrapolation methods

S6. Bulk MR and NR-PM1/rBC mass concentration

Figure S5 illustrates the bulk MR value increased together with the NR-PM1 mass concentration and rBC mass concentration during the winter time. Compare to the winter campaign period, the average bulk MR varied less significantly in summer and the NR-PM1 and rBC concentration is lower.
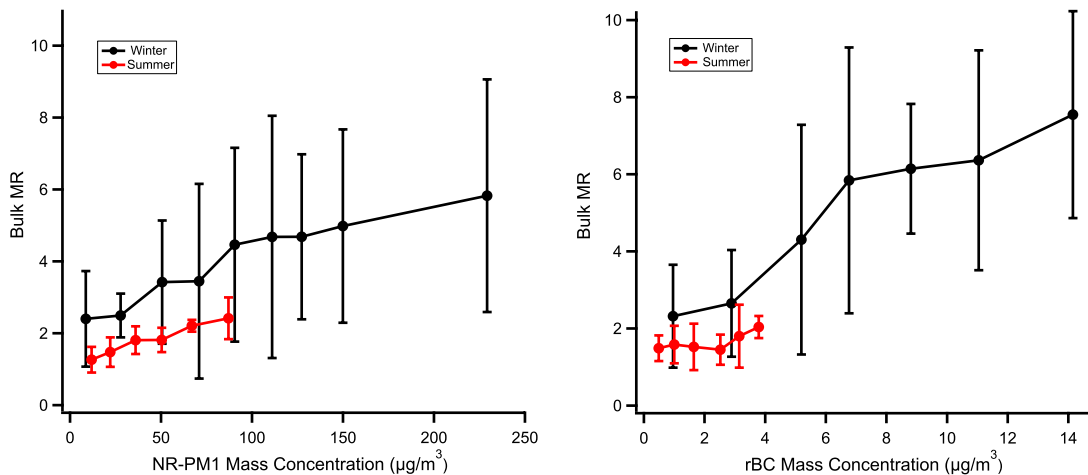
Figure S5: CPMA-SP2 bulk MR results as a function of NR-PM1 Mass Concentration and rBC Mass Concentration

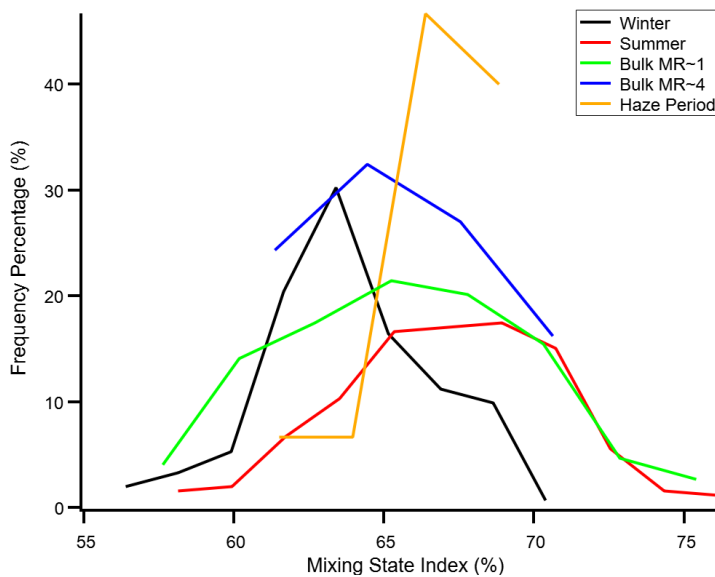## S7. Mixing state index frequency

5

Figure S6: Frequency of Mixing State Index.

Figure S6 is a histogram presents the percentage distribution of χ. The χ ranged from around 56% to around 71% for winter
10   and is ranged from around 58% to around 76% in summer. The χ value is mostly around 65% for the rBC-containing particles
with bulk MR~4 which indicates that thick coated rBC-containing particles are in the internal mixture state for the polluted

8

periods in winter. The investigation of haze period show that large fraction of rBC-containing particles is well internal mixed. For the rBC-containing particles with their bulk MR around 1, the χ varies from around 57% to around 75%. This widely variation range is the result of well internal mixture in summer and the poorly internal mixture in winter light polluted periods. Due to most of the rBC-containing particles with bulk MR~1 are found in the summer period, the χ distributed for MR~1 is

5 similar to the distributed for the summer period.

Reference

Bondy, A. L., Bonanno, D., Moffet, R. C., Wang, B., Laskin, A., and Ault, A. P.: The diverse chemical mixing state of aerosol particles in the southeastern United States, Atmospheric Chemistry and Physics, 18, 12595-12612, 10.5194/acp-18-12595-

10 2018, 2018.

Broda, K. N., Olfert, J. S., Irwin, M., Schill, G. P., McMeeking, G. R., Schnitzler, E. G., and Jäger, W.: A novel inversion method to determine the mass distribution of non-refractory coatings on refractory black carbon using a centrifugal particle mass analyzer and single particle soot photometer, Aerosol Science and Technology, 1-12, 10.1080/02786826.2018.1433812, 2018.

15 Collins, D. R., Flagan, R. C., and Seinfeld, J. H.: Improved Inversion of Scanning DMA Data, Aerosol Science and Technology, 36, 1-9, 10.1080/027868202753339032, 2002.

Healy, R. M., Riemer, N., Wenger, J. C., Murphy, M., West, M., Poulain, L., Wiedensohler, A., O'Connor, I. P., McGillicuddy, E., Sodeau, J. R., and Evans, G. J.: Single particle diversity and mixing state measurements, Atmospheric Chemistry and Physics, 14, 6289-6299, 10.5194/acp-14-6289-2014, 2014.

20 Riemer, N., and West, M.: Quantifying aerosol mixing state with entropy and diversity measures, Atmospheric Chemistry and Physics, 13, 11423-11439, 10.5194/acp-13-11423-2013, 2013.

Twomey, S.: Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimation of particle size distributions, Journal of Computational Physics, 18, 188-200, https://doi.org/10.1016/0021-9991(75)90028-5, 1975.

25