# ENVIRONMENTAL RESEARCH
## LETTERS

CrossMark

# Long term variation of microphysical properties of black carbon in Beijing derived from observation and machine learning

Kang Hu[1,2], Dantong Liu[2,*] ![ORCID], Siyuan Li[2], Yangzhou Wu[2,3], Baiwan Pan[2], Shitong Zhao[2], Xiaotong Jiang[2,4], Shuo Ding[2,5], Ping Tian[6,7], Dawei Hu[8], Chenjie Yu[9], Ye Wang[10], Fei Wang[6,7], Delong Zhao[6,7], Yunfei Wu[11], Deping Ding[6,7] and Hong liao[1]

[1] Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Nanjing University of Information Science & Technology, Nanjing 210044, People's Republic of China
[2] Department of Atmospheric Sciences, School of Earth Sciences, Zhejiang University, Hangzhou 310058, People's Republic of China
[3] Guangxi Key Laboratory of Environmental Pollution Control Theory and Technology, Guilin University of Technology, Guilin 541004, People's Republic of China
[4] College of Biological and Environmental Engineering, Bin Zhou University, Binzhou 256600, Shandong, People's Republic of China
[5] Department of Environmental Engineering, College of Quality and Safety Engineering, China Jiliang University, Hangzhou 310018, People's Republic of China
[6] Beijing Key Laboratory of Cloud, Precipitation and Atmospheric Water Resources, Beijing 100089, People's Republic of China
[7] Field Experiment Base of Cloud and Precipitation Research in North China, China Meteorological Administration, Beijing 100089, People's Republic of China
[8] Centre for Atmospheric Sciences, School of Earth and Environmental Sciences, University of Manchester, Manchester M13 9PL, United Kingdom
[9] Université de Paris Cité and Univ Paris Est Creteil, CNRS, LISA, Paris, France
[10] Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Joint International Research Laboratory of Climate and Environment Change (ILCEC)/ Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, People's Republic of China
[11] Key Laboratory of Regional Climate-Environment for Temperate East Asia (RCETEA), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, People's Republic of China
* Author to whom any correspondence should be addressed.

E-mail: dantongliu@zju.edu.cn

## Abstract

The microphysical attributes of black carbon (BC) can determine its absorption and hygroscopic properties. However, long-term information is difficult to obtain from the field. In this study, the BC properties including mass concentration, the coating volume ratio (VR) relative to the refractory BC (rBC), the rBC diameter and the fraction of cloud condensation nuclei (CCN), are derived from a number of field experiments using a random forest model. This model effectively derives the long-term BC microphysical properties in the Beijing region from 2013 to 2020 using continuous measurements of particulate matter, gas, BC mass concentration and meteorological parameters. The results reveal notably higher BC coatings (mean VR = 7.2) and a greater fraction of CCN-like BC (51%) in the winter compared to other seasons. Following the implementation of national air pollution control measures in 2017, BC mass exhibited a substantial reduction of 60% (29%) in the winter (summer), and VR decreased by 45% (24%). Apart from the influence of meteorological variations, these can be attributed to the declined primary emissions and the gas precursors which are associated with secondary formation of BC coatings. The reduction of both BC mass loading and coatings leads to its solar absorption decreasing by 50%, and the fraction of CCN-like BC (likely in clouds) decreasing by 23%. Environmental regulation will therefore continue to reduce both direct and indirect radiative impacts of BC in this region.

# 1. Introduction

The mixing state of atmospheric black carbon (BC) influences its light-absorbing (Liu *et al* 2017) and hygroscopic properties (Tritscher *et al* 2011, Liu *et al* 2013), which determine its direct and indirect radiative impacts. BC can acquire non-BC components soon after emission when aging in the atmosphere (Liu *et al* 2020). More coatings associated with BC can alter its initial fractal structure to be more spherical (Hu *et al* 2021b). A more spherical BC-containing particle (BCc) may be more appropriate when applied as a core–shell like structure (Hu *et al* 2022), and its light-absorbing properties can be considerably enhanced through the lensing effect (Jacobson 2001), often by a factor of 1.8–2 (Schnaiter *et al* 2005, Khalizov *et al* 2009, Cappa *et al* 2012). The coatings on the BC were observed to introduce additional heating at a rate of up to $0.1 \mathrm{~K~h}^{-1}$ over the North China Plain region (Zhao *et al* 2020), and results in a positive radiative forcing of $+0.1$ to $+4.2 \mathrm{~Wm}^{-2}$. The coatings of BC were found to be highly linked to high pollution events, and a reduction in emissions may lead to co-benefits of reducing both BC concentration and coatings (Zhang *et al* 2018).

The coatings on BC can increase its hygroscopicity because other non-BC substances are more hygroscopic than BC (Liu *et al* 2013). The addition of coatings can also enlarge the particle size, thus enhancing the ability of cloud condensation nuclei (CCN) for BCc. Observations in the Beijing region showed that under polluted conditions half the number of highly coated BC could be CCN-activated under a supersaturation (SS) of about 0.1% (Ding *et al* 2019), and up to 47% of BC mass could be activated under SS = 0.1% in central China (Hu *et al* 2021a). By measuring the coating size distribution of BCc, the BCc with larger than 200 nm was considered to be hydrophilic and was used to constrain the soluble fraction of BC in the Copernicus Atmosphere Monitoring Service global model (Ding and Liu 2022).

Measurements of BC coatings can be achieved by using the single particle soot photometer, which were conducted intensively over the Beijing region in recent years (Liu *et al* 2019), including some studies with a mixing state measurement covering a few years (Wu *et al* 2021). However, due to the non-routine measurements on the BC mixing state, the long-term and continuous information about the BC mixing state and the associated optical and hygroscopic properties are still not available. The PM$_{2.5}$ concentration in China has dramatically dropped since 2017 due to the implementation of the five-year Clean Air Action plan and the 'Three-Year Action Plan to Fight Air Pollution'. It is imperative to investigate the long-term variation of BC mixing state which may have been modified according to the change of pollution level. This will in turn access the effects of environmental policy in regulating BC to mitigate its direct and indirect radiative impacts.

In this study, we introduce a random forest model to link the measured microphysical properties of BC, including the coating volume ratio (VR) relative to refractory BC (rBC), the rBC diameter and the fraction of CCN, with the particulate matter (PM), gas, BC mass concentration and meteorological parameters between 2013 and 2020. Leveraging this established model, a continuous dataset of BC microphysical properties from 2013–2020 is derived. The radiative impact of BC before and after the implementation of environment regulation are evaluated based on this dataset.

# 2. Materials and methods

## 2.1. Site and instrumentation

The microphysical properties of BC were measured using a single-particle soot photometer (SP2, DMT Inc.). The incandescence signal was calibrated using the Aquadag standard (Acheson Inc., USA) and corrected with ambient rBC using a factor of 0.75 (Laborde *et al* 2012). The scattering signal was calibrated using the polystyrene latex sphere standard (PSL, Duke Scientific). Seven field experiments employing the SP2 instrument were conducted from 2013 to 2021 (summarized in table S1). These experiments included urban and semi-urban environments in the Beijing region.

The mass concentration of rBC can be measured through the incandescent signal of SP2. Assuming an rBC density of $1.8 \mathrm{~g~cm}^{-3}$ enables the determination of the core diameter ($D_c$) of rBC. The mass median diameter of rBC (MMD) is determined by identifying the BC core size at which the BC mass is evenly distributed both below and above this point. The scattering signal was reconstructed using the leading edge only method to account for the coating evaporation (Gao *et al* 2007, Liu *et al* 2014). The core size of rBC corresponding to the scattering signal was then employed in the Mie lookup table to determine the corresponding coated particle diameter ($D_p$) (Liu *et al* 2014, Taylor *et al* 2015). The VR is determined by summing up the volume of coated BC and rBC for all measured BC particles:

$$\mathrm{VR} = \frac{\sum_i D_{ve,i}^3}{\sum_i D_{c,i}^3} - 1, \qquad (1)$$

where $D_{ve,i}$ and $D_{c,i}$ represents the $i$th coated particle diameter and uncoated rBC diameter, respectively.

The proportion of BC particles which are likely to serve as CCN are calculated assuming that all BC with a coated particle size greater than 200 nm can be activated (Ding and Liu 2022). Considering that particle size is the predominant parameter influencing the activation of particles, this has also been validated by
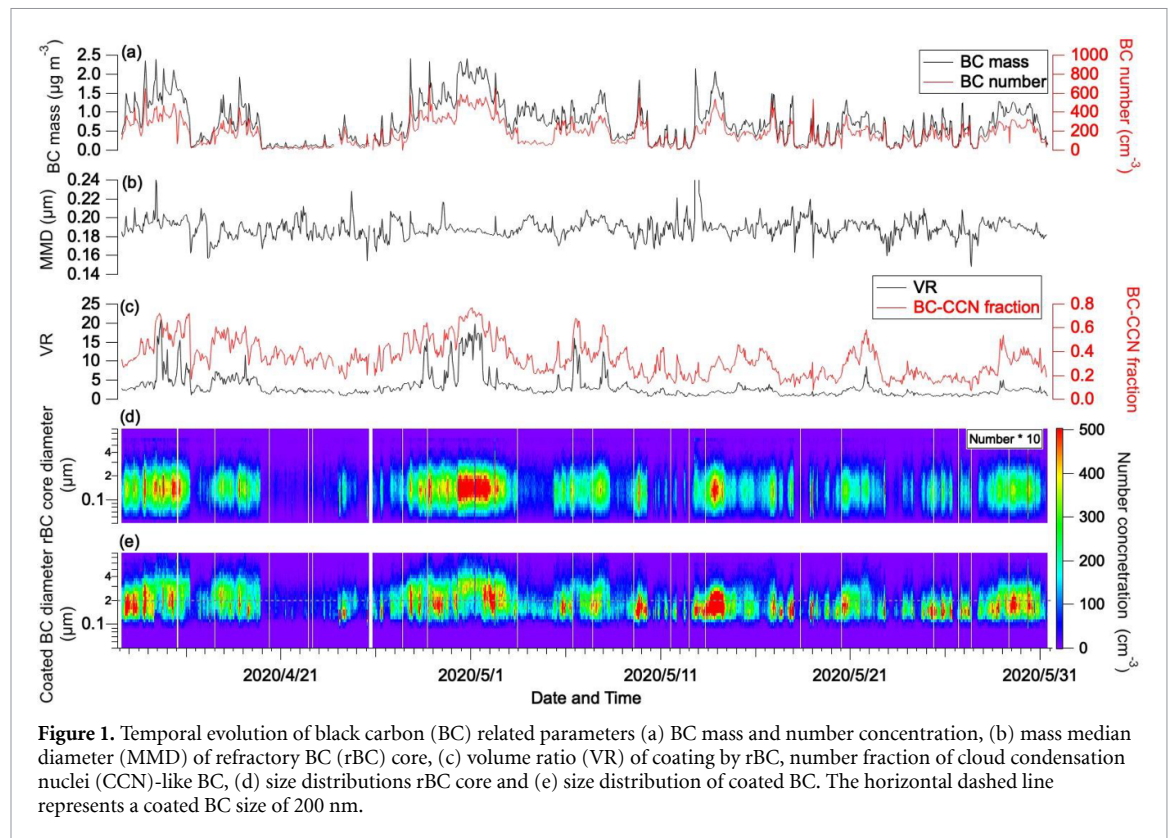
**Figure 1.** Temporal evolution of black carbon (BC) related parameters (a) BC mass and number concentration, (b) mass median diameter (MMD) of refractory BC (rBC) core, (c) volume ratio (VR) of coating by rBC, number fraction of cloud condensation nuclei (CCN)-like BC, (d) size distributions rBC core and (e) size distribution of coated BC. The horizontal dashed line represents a coated BC size of 200 nm.

previous researchers (Dusek *et al* 2006, Ervens *et al* 2007). A temporal evolution example of microphysical properties of BC measured by SP2 is shown in figure 1.

The Aethalometer AE33 (Magee Scientific Inc., USA) was employed for continuous monitoring of eBC mass (Petzold *et al* 2013) at a semi-urban site in Beijing (117.12° E, 40.14° N, 50 m a.s.l.). The absorption coefficient at 880 nm (Abs$_{880}$) were used to derive eBC mass due to the less interference by brown carbon at this wavelength. Dual spot measurements at various flow rates were conducted to automatically correct for the filter loading effect (Drinovec *et al* 2015). The multiscattering factor (C-value) was determined using a photoacoustic soot spectrometer (PASS-3, DMT) at an overlapping wavelength, resulting in a factor of 2.37 at 880 nm. The concentration of eBC mass was derived from the light attenuation and a fixed mass absorption cross-section (MAC) of 16.6 m$^2$ g$^{-1}$ (Aruna *et al* 2013).

**2.2. Calculation of absorbing properties of BC**
The MAC of BC under the partial core–shell model is computed based on particle size, refractive index, and Mie theory, as well as whether the BC particle contains absorption enhancement (equation (3)) (Hu *et al* 2022). Specifically, the calculation considers an external mixing state without a lensing effect on absorption, with the fraction of BC mass calculated as $F_{ns} = -0.27\log(\text{VR}) + 0.64$. Conversely, the remaining BC particles are assumed to have a core–shell

structure. Within this study, the MAC$_{880}$ is calculated using a refractive index (RI) of $1.95 + 0.79i$ (Bond and Bergstrom 2006) for rBC and $1.48 + 0i$ (Liu *et al* 2015) for coatings. The RI of rBC within the visible wavelength range is assumed to remain constant (Jacobson 2001, Bond and Bergstrom 2006). The mass-weighted MAC for bulk BCc particles is expressed as:

$$\text{MAC} = \frac{\sum_i \text{MAC}_i \times m_i}{\sum_i m_i} \quad (2)$$

where MAC$_i$ and $m_i$ represent the MAC and mass of BC within the $i^{\text{th}}$ bin, respectively.

The calculated MAC of BCc under the partial core–shell model is:

$$\text{MAC}_{\text{Partial Core-shell}} = F_{ns} \times \text{MAC}_{\text{uncoat}} + (1 - F_{ns}) \times \text{MAC}_{\text{coated}} \quad (3)$$

where $F_{ns}$ denotes the mass fraction of BC particles without absorption enhancement, MAC$_{\text{coated}}$ and MAC$_{\text{uncoat}}$ are determined using equation (2) with and without coatings, respectively.

The solar absorption of BC at 370–880 nm is determined using the approximated results of the Beer–Lambert law (Kirchstetter and Thatcher 2012):

$$\int I_0(\lambda) \left[ \frac{I_0 - I}{I_0}(\lambda, \text{BC}) \right] d\lambda$$
$$= \int I_0(\lambda)[1 - e^{-(\text{MAC}_{\text{BC}} \times [\frac{\lambda_0}{\lambda}]^{\text{AAE}_{\text{BC}}} \times C_{\text{BC}} \times h_{\text{ABL}})}] \quad (4)$$

where $MAC_{BC}$ denote the MAC-sections of BC at the reference wavelength ($\lambda_0$), which is set at 370 nm. The $AAE_{BC}$ is calculated using a fixed value of 1 (Clarke *et al* 2007, Olson *et al* 2015, Jiang *et al* 2022). $C_{BC}$ represents the mass concentration of BC. $h_{ABL}$ refers to the planet boundary layer height, obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data (ERA5). $I_0(\lambda)$ is the solar emission flux obtained from the Discrete Ordinate Radiative Transfer solvers module within the corresponding time period at 117.12 °E, 40.14 °N.

## 2.3. Deriving the long-term microphysical properties of BC

A random forest method (Breiman 2001) is employed to derive the microphysical properties of BC. The random forest method was chosen for the robustness in managing small sample sizes and high-dimensional feature spaces, thereby avoiding overfitting (Breiman 2001, Biau and Scornet 2016). The data from 2013.01–2020.08 is used for model training, and the remaining data in 2021.03–05 for testing. The output parameters derived from the training model include MMD, VR, and the fraction of CCN-like BC from SP2 measurements obtained during experiments (table S1, figure S1). The input parameters included the actual observed BC mass, $SO_2$, $O_3$, $NO_2$, CO, $PM_{2.5}$, and temperature, relative humility (RH), surface air pressure (SP) anomaly, planet boundary layer height (PBLH), U wind, V wind from ECMWF (table S2, figures S1–S3). The selected meteorological parameters are chosen based on their impact on aerosol transport, aging processes, and concentration variations. $PM_{2.5}$, $SO_2$, $O_3$, $NO_2$, and CO from monitoring stations (figure S3) are chosen for specific reasons. $PM_{2.5}$ is selected because figure S4 indicates a strong correlation between high $PM_{2.5}$ and increased BC mass and VR. $SO_2$ is considered due to the significant amount generated from coal combustion in China, acting as a precursor for sulfate formation in BC coatings. $O_3$ is included as it reflects atmospheric oxidation, influencing the secondary generation of BC coatings. Finally, $NO_2$ and CO are both products associated with BC sources. Coefficient of determination ($R^2$) is used to measure the performance of the random forest model.

Due to the rarity of high pollution conditions during the entire observation period, the model significantly underestimates the simulation of BC microphysical properties under such conditions (Wei *et al* 2020). This study initially defined a high-pollution indicator, referring to the sum of observed results for each parameter exceeding the monthly mean by twice the standard deviation. The raw high pollution data accounted for approximately 3.2%. In addition, this study introduced the synthetic minority oversampling technique (SMOTE) method (Torgo 2010, Ghorbani and Ghousi 2020), generating new synthetic samples along the line between high concentration data points and their nearest neighbors to oversample the high-pollution data (Chawla *et al* 2002, Chawla *et al* 2003). Following SMOTE resampling, the proportion of high pollution data increased to 25%.
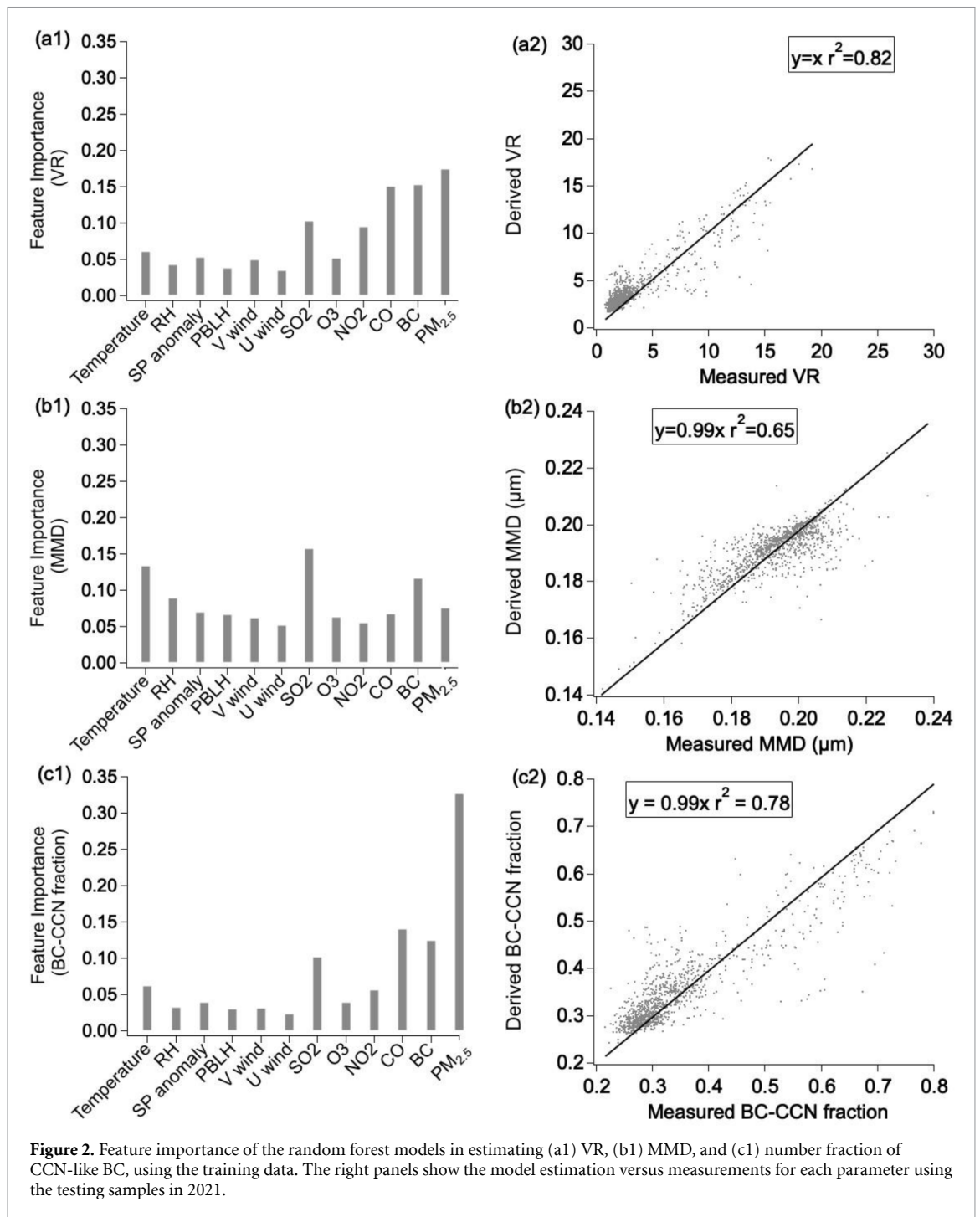
The first stage of training establishes the relationship between the high-pollution indicator and all predictors. The hyperparameters, which consist of n_estimators (the number of decision trees in the forest), max_depth (the maximum depth of each decision tree), max_features (the maximum number of features considered for splitting a node), min_samples_split (the minimum number of samples required to split a node), and bootstrap (an optional setting for sampling data with or without replacement), are adjusted to achieve the optimal random forest model. The optimal hyperparameters are determined as follows: n_estimators = 110, max_depth = 17, max_features = 'sqrt', min_samples_split = 4, and bootstrap = 'True'. The predicted high-pollution indicator is then utilized as a predictor in the second stage model. In the subsequent stage, the model uses the residual between actual measurements and simulated results as the dependent variable for training. This approach enhances the responsiveness of prediction factors to result variations, consequently improving prediction accuracy.

After obtaining the relation between the input parameters and BC microphysical properties, we are able to derive the MMD, VR and the fraction of CCN-like BC once all long-term input parameters (figure 3). The AE33 measured eBC mass was firstly converted to BC mass based on the existing relation between eBC and rBC. The other long-term input parameters include gas and $PM_{2.5}$ concentrations and the meteorological data (figure S5).

## 3. Results and discussion

### 3.1. Factors in determining BC microphysical properties

Figure S4 shows both BC and $PM_{2.5}$ mass concentration are correlated VR. For instance, high VR was frequently associated with elevated $PM_{2.5}$ and BC mass. The variation of data points from all experiments can be explained by mapping the $PM_{2.5}$ and BC mass. This is consistent with the random forest model results that BC and $PM_{2.5}$ mass concentrations had the most feature importance. This results from the fact that the most coated BC occurred during the polluted period when both $PM_{2.5}$ and rBC mass were high. The intensive coagulation and condensation processes during pollution causes substantial formation

**Figure 2.** Feature importance of the random forest models in estimating (a1) VR, (b1) MMD, and (c1) number fraction of CCN-like BC, using the training data. The right panels show the model estimation versus measurements for each parameter using the testing samples in 2021.

of non-BC substances in BC (Akagi *et al* 2012, Moffet *et al* 2016). Overall, VR is related to pollutant emissions and is less affected by changes in meteorological conditions.

Among the factors influencing the MMD results, $SO_2$ is the most significant (figure 2(b1)). MMD in Beijing displays evident seasonal variations (Liu *et al* 2019, Hu *et al* 2020), attributed to the influence of residential emissions, which notably affect MMD variations during winter compared to relatively uniform emission sectors in warmer seasons (Zhang *et al* 2009). Residential emissions in China generally lead to higher $SO_2$, which corresponds to

larger MMD, while BC emissions from traffic sources correspond to smaller MMD measurements. Thus, the influence of temperature primarily focuses on changes in residential emissions attributed to heating activities. The derived information from the hourly dataset may have included some diurnal variations, but the monthly average to some extent minimizes the effect in hourly time scale. BC mass represents intrinsic microphysical properties of BC, and it serves as a dependable parameter for inverting other microphysical properties, particularly for BC particle size.

Figure 2(c1) shows the most sensitive parameters for CCN-like fraction are slightly different from VR,
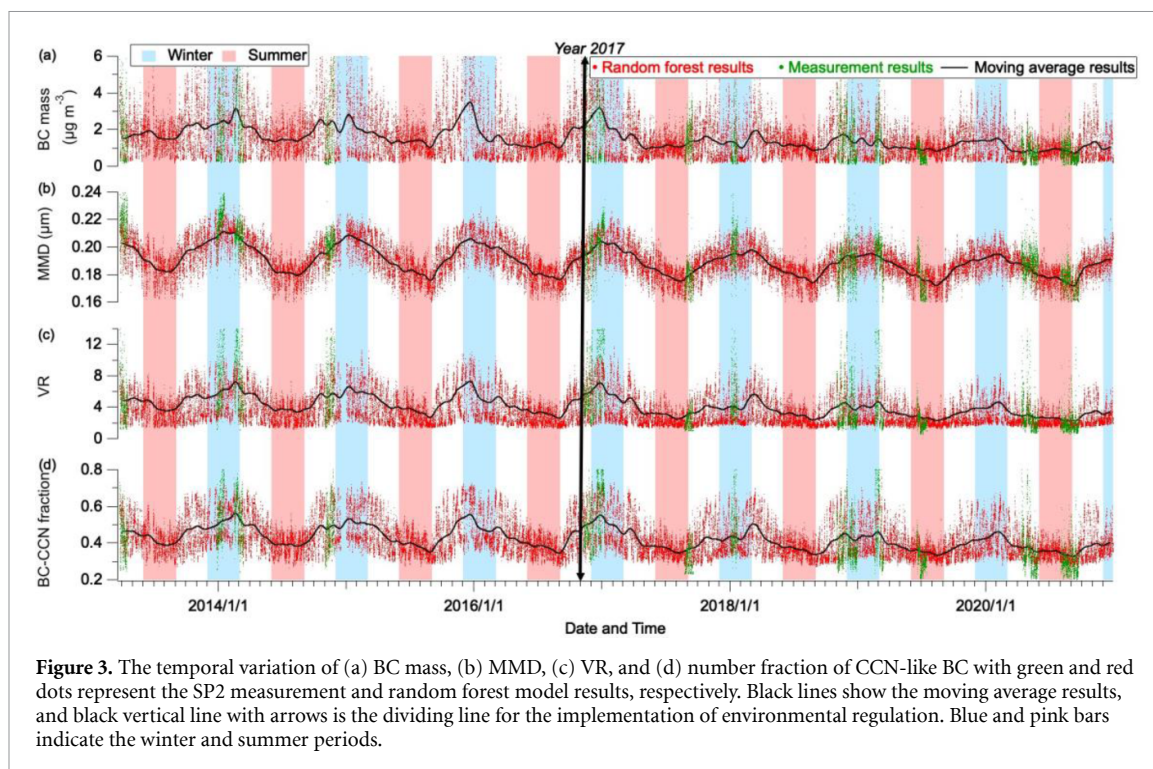
**Figure 3.** The temporal variation of (a) BC mass, (b) MMD, (c) VR, and (d) number fraction of CCN-like BC with green and red dots represent the SP2 measurement and random forest model results, respectively. Black lines show the moving average results, and black vertical line with arrows is the dividing line for the implementation of environmental regulation. Blue and pink bars indicate the winter and summer periods.

because the particle size of BC depends on both the core size (MMD) and VR. $PM_{2.5}$, CO and rBC mass are the three parameters with the highest importance, which is the same as VR, but $PM_{2.5}$ clearly had a more significant influence on the CCN-like fraction. The additional introduction of $SO_2$ and temperature may be due to the seasonal influence by which additional sources with seasonal patterns may have influenced the overall size of BC.
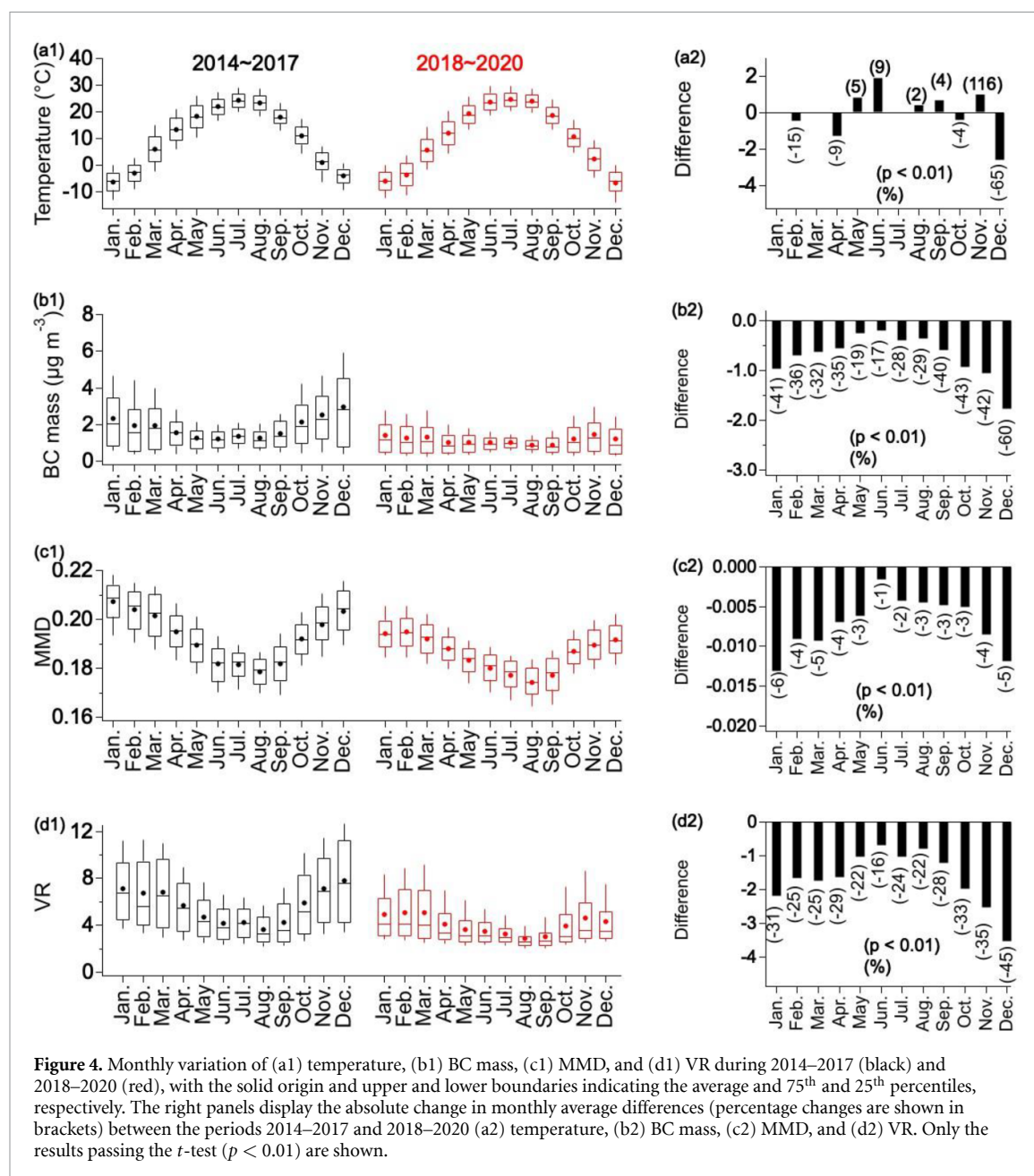
The analysis above reveals a satisfactory agreement between the simulation and observation, with three parameters (VR, MMD and the fraction of CCN-like BC) exhibiting explained variation of $R^2$ = 0.82, 0.65, 0.78, respectively (figures 2(a2)–(c2)).

### 3.2. Impacts of environmental regulations
In 2017, China concluded the first five-year Clean Air Action Plan leading to a notable improvement in pollution control. Consequently, there has been an obvious decline in BC mass loading in Beijing after 2017 (figure 3). Consequently, an obvious decline in BC mass loading in Beijing after 2017 can be generally seen (figure 3). The following analysis compares the characteristics of BCc before and after the 2017 environmental action plan. Figure 4(b1) showed the seasonal BC mass loading peaked in wintertime, which was due to the residential heating activities in cold season (Zhang *et al* 2009). As figure 4(b2) shows, the most noticeable reduction of BC mass loading was for the highly polluted winter months by 36%–60%. The BC mass loading in the summer had not shown

significant drop, this may be due to the low residential heating sources in the summer and the reduction of BC sources from non-residential sources were not significantly reduced in the warm season. This is consistent with the more apparent reduction of MMD in winter (figure 4(c)), because BC from coal burning tends to have a larger core size than other sources, and the more apparent reduction of residential coal burning in winter will have a more dramatic reduction of MMD in this season. Both emission and meteorology may influence the characteristics of pollutants, and the influence of emission and meteorology can be separated via modeling method (Cheng *et al* 2019, Kanaya *et al* 2020). The exact separation between emission and meteorology is not the scope of this study, but we use several years average before and after year 2017 (when the environment action was implemented) to normalize the yearly variations of meteorological conditions. The difference between the two datasets before and after year 2017 therefore very likely results from the change of emission, but the influence of meteorology is unable to be completely excluded.

The coatings on BC (as indicated by VR) showed the highest in the winter, which was about 165% higher than in summer (figure 4(d1)). This is consistent with previous observations (Liu *et al* 2019). This results from the high co-emissions of both BC and other pollutants in the winter, and the other non-BC substances facilitate the formation of coatings in BC (Zhang *et al* 2018). In addition, the colder temperate may also drive more gas phases to aerosols that produce the coatings. The reduction of primary emission

**Figure 4.** Monthly variation of (a1) temperature, (b1) BC mass, (c1) MMD, and (d1) VR during 2014–2017 (black) and 2018–2020 (red), with the solid origin and upper and lower boundaries indicating the average and 75th and 25th percentiles, respectively. The right panels display the absolute change in monthly average differences (percentage changes are shown in brackets) between the periods 2014–2017 and 2018–2020 (a2) temperature, (b2) BC mass, (c2) MMD, and (d2) VR. Only the results passing the $t$-test ($p < 0.01$) are shown.

has also reduced the precursors for the formation of secondary pollutants. BC coatings are mainly composed of secondary substances which condense or coagulate with primary particles. A reduction of secondary substances will lead to the reduction of coatings on BC. This can be clearly reflected by the significant reduction of VR in the winter by 45% after the environmental action (figure 4(d)), and this reduction was less in summer.

The $MAC_{880}$ calculated by different optical models is shown in figure 5. The partial core–shell model are proven to effectively reduce the overestimation of BCc absorption if considering all BC particles are in core–shell structure (Hu *et al* 2022). The modeled absorption coefficient were also compared with the observation by the aethalometer. The partial core–shell model was found to be closest to observation (figure S7). Consistent with the decreased VR, $MAC_{880}$ showed a mean 0.6 and 0.3 m$^2$ g$^{-1}$ decrease in winter and summer respectively (figure 5(a2)). The solar absorption absorbed by BC is shown in figure 5(b), which exhibits the largest in spring but lower in winter and summer. This is because though winter had the highest BC mass and MAC, the solar radiation was the lowest, while springtime solar radiation started to enhance when the BC mass was still high. This is consistent with the largest decrease in solar absorption in spring (5–8 W m$^{-2}$, figure 5(b2)) but lower for other seasons before and after the environmental action plan.
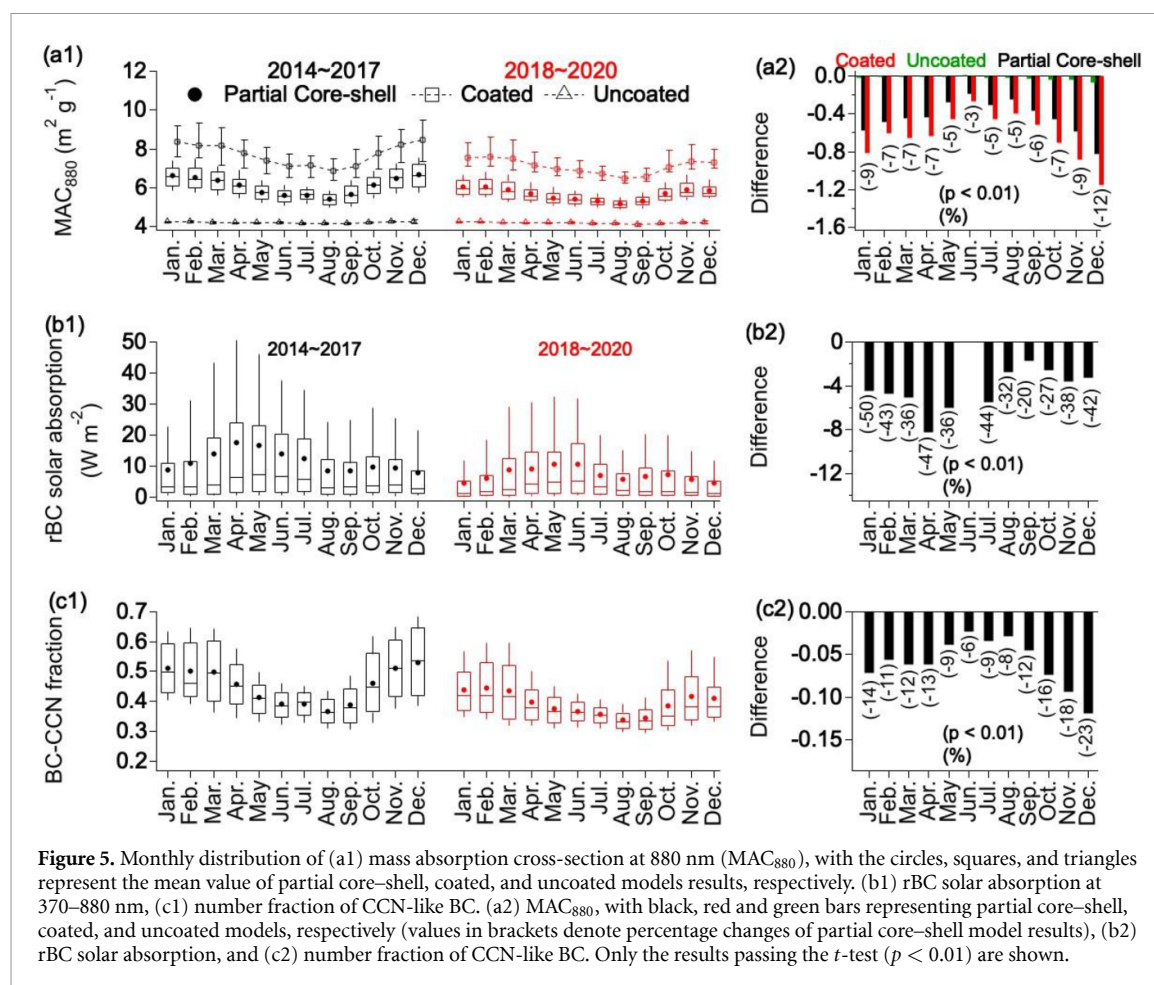
**Figure 5.** Monthly distribution of (a1) mass absorption cross-section at 880 nm ($MAC_{880}$), with the circles, squares, and triangles represent the mean value of partial core–shell, coated, and uncoated models results, respectively. (b1) rBC solar absorption at 370–880 nm, (c1) number fraction of CCN-like BC. (a2) $MAC_{880}$, with black, red and green bars representing partial core–shell, coated, and uncoated models, respectively (values in brackets denote percentage changes of partial core–shell model results), (b2) rBC solar absorption, and (c2) number fraction of CCN-like BC. Only the results passing the *t*-test ($p < 0.01$) are shown.

Figure 5(c1) showed the fraction of CCN-like BC exhibited apparent seasonal variations both before and after 2017, which was the highest in winter and lowest in summer. This is highly consistent with the variation of VR in that a higher VR means a larger coated size of BCc which is more likely CCN. Because of the decrease of both BC core size and coatings, the overall coated BC size had dropped significantly, leading to a reduction of 0.06–0.12 for the fraction of CCN-like BC in winter and 0.02–0.03 in summer (figure 5). This will mean BC is less likely to be incorporated into clouds, exerting less indirect impacts compared to that before 2018.

## 4. Conclusion

To investigate the long-term variation of the microphysical properties of BC, this study introduces a random forest model to link the BC microphysical properties with BC mass, PM, gas concentrations and meteorological parameters from several field experiments. Based on this, the long-term BC MMD, coating VR relative to rBC and CCN-like BC from year 2013–2020 are derived. The results show that the implementation of national air pollution regulations and changes in meteorological conditions led to a

significant reduction in BC mass loading by 17%–60%, especially by reducing the emissions in winter from residential heating activities. The clear reduction of BC coatings and CCN-like fraction suggests the reduction of secondary substances besides the primary emission. We find that most solar absorption occurs at springtime when the solar radiation starts to increase while BC emission and coatings are still high at the end of the heating period. In particular, under such conditions near the top of the boundary layer, it will further inhibit the development of the boundary layer, thereby exacerbating high pollution concentrations. Thus, the pollution regulation therefore needs more attention during this season. The results imply that the air pollution control has an effective impact on both the direct and indirect radiative effects of BC. However, the reduction of coatings on BC may decrease its hygroscopicity and reduce the efficiency of wet deposition, which may increase its lifetime in the atmosphere and needs further evaluation.

## Data availability statement

## Acknowledgments

## ORCID iD

Dantong Liu ⬤ https://orcid.org/0000-0003-3768-1770

## References

Akagi S K *et al* 2012 Evolution of trace gases and particles emitted by a chaparral fire in California *Atmos. Chem. Phys.* **12** 1397–421

Aruna K, Kumar T V L, Rao D N, Murthy B V K, Babu S S and Moorthy K K 2013 Black carbon aerosols in a tropical semi-urban coastal environment: effects of boundary layer dynamics and long range transport *J. Atmos. Sol. Terr. Phys.* **104** 116–25

Biau G and Scornet E 2016 A random forest guided tour *Test* **25** 197–227

Bond T C and Bergstrom R 2006 Light absorption by carbonaceous particles: an investigative review *Aerosol Sci. Technol.* **40** 27–67

Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

Cappa C D *et al* 2012 Radiative absorption enhancements due to the mixing state of atmospheric black carbon *Science* **337** 1078–81

Chawla N V *et al* 2003 SMOTEBoost: improving prediction of the minority class in boosting *Knowledge Discovery in Databases: PKDD 2003: 7th European Conf. on Principles and Practice of Knowledge Discovery in Databases, (Cavtat-Dubrovnik, Croatia, 22–26, September 2003) Proceedings 7* (Springer) pp 107–19

Chawla N V, Bowyer K W, Hall L O and Kegelmeyer W P 2002 SMOTE: synthetic minority over-sampling technique *J. Artif. Intell. Res.* **16** 321–57

Cheng J *et al* 2019 Dominant role of emission reduction in PM 2.5 air quality improvement in Beijing during 2013–2017: a model-based decomposition analysis *Atmos. Chem. Phys.* **19** 6125–46

Clarke A *et al* 2007 Biomass burning and pollution aerosol over North America: organic components and their influence on spectral optical properties and humidification response *J. Geophys. Res. Atmos.* **112** 1–13

Ding S *et al* 2019 Size-related physical properties of black carbon in the lower atmosphere over Beijing and europe *Environ. Sci. Technol.* **53** 11112–21

Ding S and Liu D 2022 Evaluation of the CAMS reanalysis for atmospheric black carbon and carbon monoxide over the north China plain *Environ. Pollut.* **314** 120286

Drinovec L *et al* 2015 The "dual-spot" Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation *Atmos. Meas. Tech.* **8** 1965–79

Dusek U *et al* 2006 Size matters more than chemistry for cloud-nucleating ability of aerosol particles *Science* **312** 1375–8

Ervens B, Cubison M, Andrews E, Feingold G, Ogren J A, Jimenez J L, DeCarlo P and Nenes A 2007 Prediction of cloud condensation nucleus number concentration using measurements of aerosol size distributions and composition and light scattering enhancement due to humidity *J. Geophys. Res. Atmos.* **112** 1–15

Gao R S *et al* 2007 A novel method for estimating light-scattering properties of soot aerosols using a modified single-particle soot photometer *Aerosol Sci. Technol.* **41** 125–35

Ghorbani R and Ghousi R 2020 Comparing different resampling methods in predicting students' performance using machine learning techniques *IEEE Access* **8** 67899–911

Hu D *et al* 2021a Direct quantification of droplet activation of ambient black carbon under water supersaturation *J. Geophys. Res. Atmos.* **126** e2021JD034649

Hu K *et al* 2021b Measurements of the diversity of shape and mixing state for ambient black carbon particles *Geophys. Res. Lett.* **48** 1–12

Hu K *et al* 2022 Identifying the fraction of core–shell black carbon particles in a complex mixture to constrain the absorption enhancement by coatings *Environ. Sci. Technol. Lett.* **9** 272–9

Hu K, Zhao D, Liu D, Ding S, Tian P, Yu C, Zhou W, Huang M and Ding D 2020 Estimating radiative impacts of black carbon associated with mixing state in the lower atmosphere over the northern North China Plain *Chemosphere* **252** 126455

Jacobson M Z 2001 Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols *Nature* **409** 695–7

Jiang X *et al* 2022 Connecting the light absorption of atmospheric organic aerosols with oxidation state and polarity *Environ. Sci. Technol.* **56** 12873–85

Kanaya Y, Yamaji K, Miyakawa T, Taketani F, Zhu C, Choi Y, Komazaki Y, Ikeda K, Kondo Y and Klimont Z 2020 Rapid reduction in black carbon emissions from China: evidence from 2009–2019 observations on Fukue Island, Japan *Atmos. Chem. Phys.* **20** 6339–56

Khalizov A F, Xue H, Wang L, Zheng J and Zhang R 2009 Enhanced light absorption and scattering by carbon soot aerosol internally mixed with sulfuric acid *J. Phys. Chem.* A **113** 1066–74

Kirchstetter T and Thatcher T 2012 Contribution of organic carbon to wood smoke particulate matter absorption of solar radiation *Atmos. Chem. Phys.* **12** 6067–72

Laborde M *et al* 2012 Single particle soot photometer intercomparison at the AIDA chamber *Atmos. Meas. Tech.* **5** 3077–97

Liu D *et al* 2014 Size distribution, mixing state and source apportionment of black carbon aerosol in London during wintertime *Atmos. Chem. Phys.* **14** 10061–84

Liu D *et al* 2017 Black-carbon absorption enhancement in the atmosphere determined by particle mixing state *Nat. Geosci.* **10** 184–8

Liu D *et al* 2019 Contrasting physical properties of black carbon in urban Beijing between winter and summer *Atmos. Chem. Phys.* **19** 6749–69

Liu D, Allan J, Whitehead J, Young D, Flynn M, Coe H, McFiggans G, Fleming Z L and Bandy B 2013 Ambient black carbon particle hygroscopic properties controlled by mixing state and composition *Atmos. Chem. Phys.* **13** 2015–29

Liu D, He C, Schwarz J P and Wang X 2020 Lifecycle of light-absorbing carbonaceous aerosols in the atmosphere *npj Clim. Atmos. Sci.* **3** 1–18

Liu D, Taylor J W, Young D E, Flynn M J, Coe H and Allan J D 2015 The effect of complex black carbon microphysics on the determination of the optical properties of brown carbon *Geophys. Res. Lett.* **42** 613–9

Moffet R C, O'Brien R E, Alpert P A, Kelly S T, Pham D Q, Gilles M K, Knopf D A and Laskin A 2016 Morphology and mixing of black carbon particles collected in central California during the CARES field study *Atmos. Chem. Phys.* **16** 14515–25

Olson M R, Victoria Garcia M, Robinson M A, Van Rooy P, Dietenberger M A, Bergin M and Schauer J J 2015 Investigation of black and brown carbon multiple-wavelength-dependent light absorption from biomass and fossil fuel combustion source emissions *J. Geophys. Res. Atmos.* **120** 6682–97

Petzold A *et al* 2013 Recommendations for reporting" black carbon" measurements *Atmos. Chem. Phys.* **13** 8365–79

Schnaiter M, Linke C, Möhler O, Naumann K-H, Saathoff H, Wagner R, Schurath U and Wehner B 2005 Absorption amplification of black carbon internally mixed with secondary organic aerosol *J. Geophys. Res. Atmos.* **110** 1–11

Taylor J, Allan J D, Liu D, Flynn M, Weber R, Zhang X, Lefer B L, Grossberg N, Flynn J and Coe H 2015 Assessment of the sensitivity of core/shell parameters derived using the single-particle soot photometer to density and refractive index *Atmos. Meas. Tech.* **8** 1701–18

Torgo L 2010 *Data Mining with R: Learning with Case Studies* (Chapman and Hall/CRC) (https://doi.org/10.1201/9780429292859)

Tritscher T *et al* 2011 Changes of hygroscopicity and morphology during ageing of diesel soot *Environ. Res. Lett.* **6** 034026

Wei J *et al* 2020 Improved 1-km-resolution $PM_{2.5}$ estimates across China using the space-time extremely randomized trees *Atmos. Chem. Phys.* **20** 3273–89

Wu Y *et al* 2021 Effect of source variation on the size and mixing state of black carbon aerosol in urban Beijing from 2013 to 2019: implication on light absorption *Environ. Pollut.* **270** 116089

Zhang Q *et al* 2009 Asian emissions in 2006 for the NASA INTEX-B mission *Atmos. Chem. Phys.* **9** 5131–53

Zhang Y, Zhang Q, Cheng Y, Su H, Li H, Li M, Zhang X, Ding A and He K 2018 Amplification of light absorption of black carbon associated with air pollution *Atmos. Chem. Phys.* **18** 9879–96

Zhao D *et al* 2020 Vertical evolution of black carbon characteristics and heating rate during a haze event in Beijing winter *Sci. Total Environ.* **709** 136251