Supplement of

Inversion Algorithm of Black Carbon Mixing State Based on Machine Learning

Zeyuan Tian et al.

Correspondence to: Jiandong Wang (jiandong.wang@nuist.edu.cn)

1 SHAP interpretations

Figure S1 shows the scattering and incandescence signals of six typical particles used as input for the machine learning (ML) model. The Fig. S1 also indicates the specific distribution of seven important features, identified by the SHapley Additive exPlanation (SHAP) method, within the incandescence signals of different particles. These six typical particles have similar BC core diameters (D_c), but the coating thickness (CT) varies, decreasing sequentially from panel (a) to (f). Since the evaporation time of the coating of BC-containing particles is positively correlated with the CT, the time delay (Δt) decreases sequentially from panel (a) to (f). When the coating is thicker and Δt is longer, the seven important incandescence signal features are mostly distributed at the baseline position where the incandescence signal has not yet begun to rise, and their corresponding values are smaller. As the coating becomes thinner and Δt decreases, the position where the incandescence signal starts to rise from the baseline earlier. Consequently, the relative positions of the seven important incandescence signal features gradually approach the peak of the incandescence signal, and their corresponding values increase. Therefore, these seven important incandescence signal features exhibit the same characteristic: as their values increase, the corresponding SHAP values decrease, leading to a reduction in the entire particle diameter (D_p) of the BC-containing particles predicted by the model.

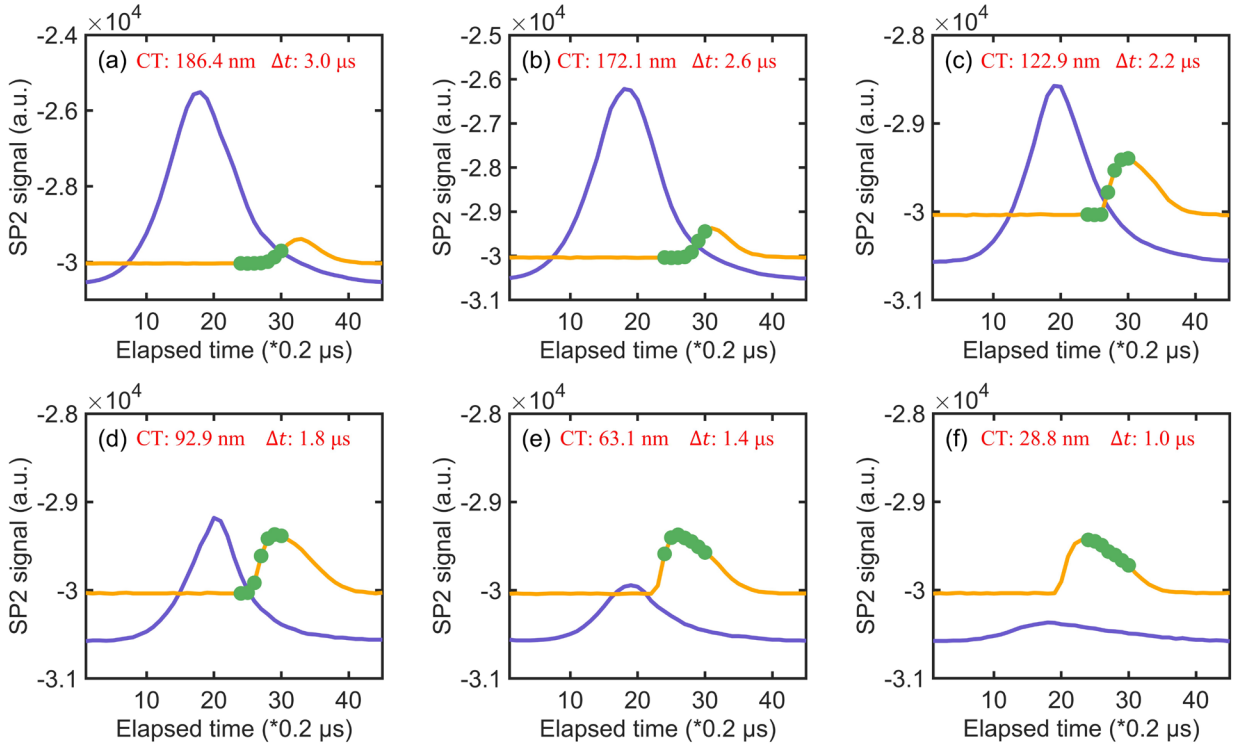

Figure S1. The scattering and incandescence signals of six typical particles used as input for the ML model. The green dots indicate the specific positions of seven important features, identified by the SHAP method, within the incandescence signals of different particles. From panel (a) to (f), the CT decreases sequentially, and correspondingly, Δt also gradually decreases.

2 Model application

Table S1. The statistical results of D_p and D_c of internally mixed BC obtained by applying the BC mixing state inversion model to the single-particle soot photometer (SP2) data in April 2022.

Statistical results	D_c	D_p
R^2	0.99	0.98
RMSE (nm)	0.21	8.28
MAE (nm)	0.14	4.19
Third Quartile (nm)	150.8	236.7
Median (nm)	130.8	200.0
First Quartile (nm)	114.0	177.5