Supporting Information for

Technical note: Reconstructing surface missing aerosol elemental carbon data in long-term series with ensemble learning

Qingxiao Meng[1], Yunjiang Zhang[1*], Sheng Zhong[2], Jie Fang[1], Lili Tang[3], Yongcai Rao[4], Minfeng Zhou[5], Jian Qiu[6], Xiaofeng Xu[7], Jean-Eudes Petit[8], Olivier Favez[9], Xinlei Ge[1]

[1]Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

[2]Jiangsu Environmental Monitoring Center, Nanjing, China

[3]Nanjing Rayleigh Scientific Instruments Co., Ltd., Nanjing, China

[4]Xuzhou Environmental Monitoring Center of Jiangsu, Xuzhou, China

[5]Suzhou Environmental Monitoring Center of Jiangsu, Suzhou, China

[6]Zhenjiang Environmental Monitoring Center of Jiangsu, Zhenjiang, China

[7]Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China

[8]Laboratoire des Sciences du Climat et de l'Environnement, CEA/Orme des Merisiers, Gif sur Yvette, France

[9]Institut National de l'Environnement Industriel et des Risques, Verneuil-en-Halatte, France

Correspondence to: Yunjiang Zhang (yjzhang@nuist.edu.cn)

Contents of this file

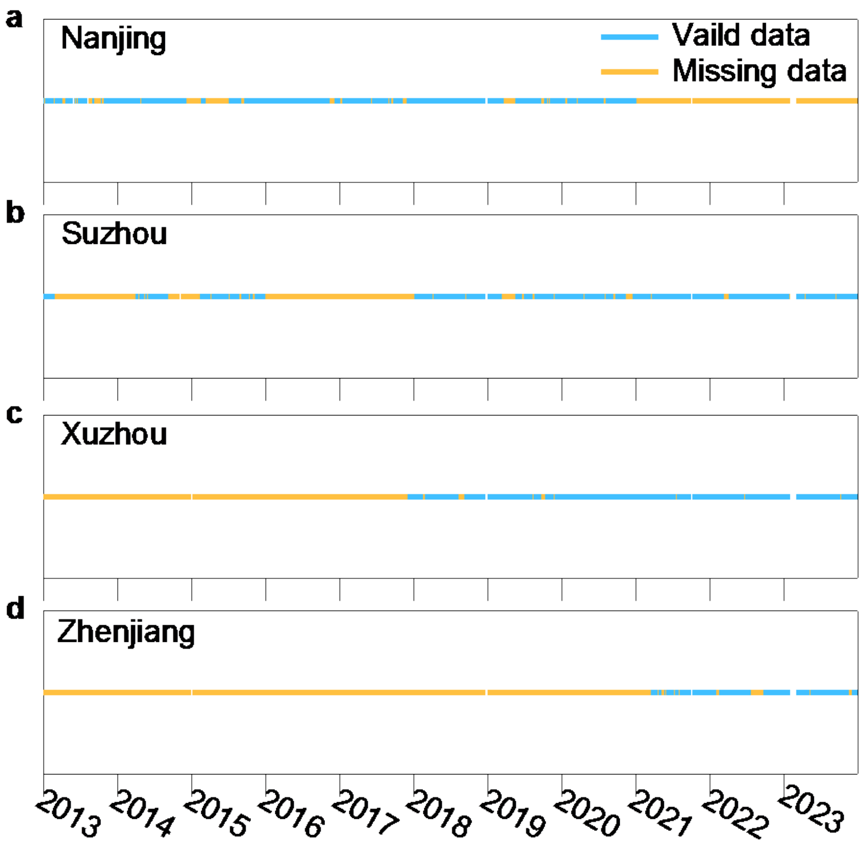

Figure S1. The amount of the vaild data and missing data from 2013 to 2023 in the four cities: a. Nanjing, b. Suzhou, c Xuzhou, and d. Zhenjiang. The blue line represents the amount of valid ground observation data, while the yellow line represents the amount of missing data reconstructed using the ensemble learning model.

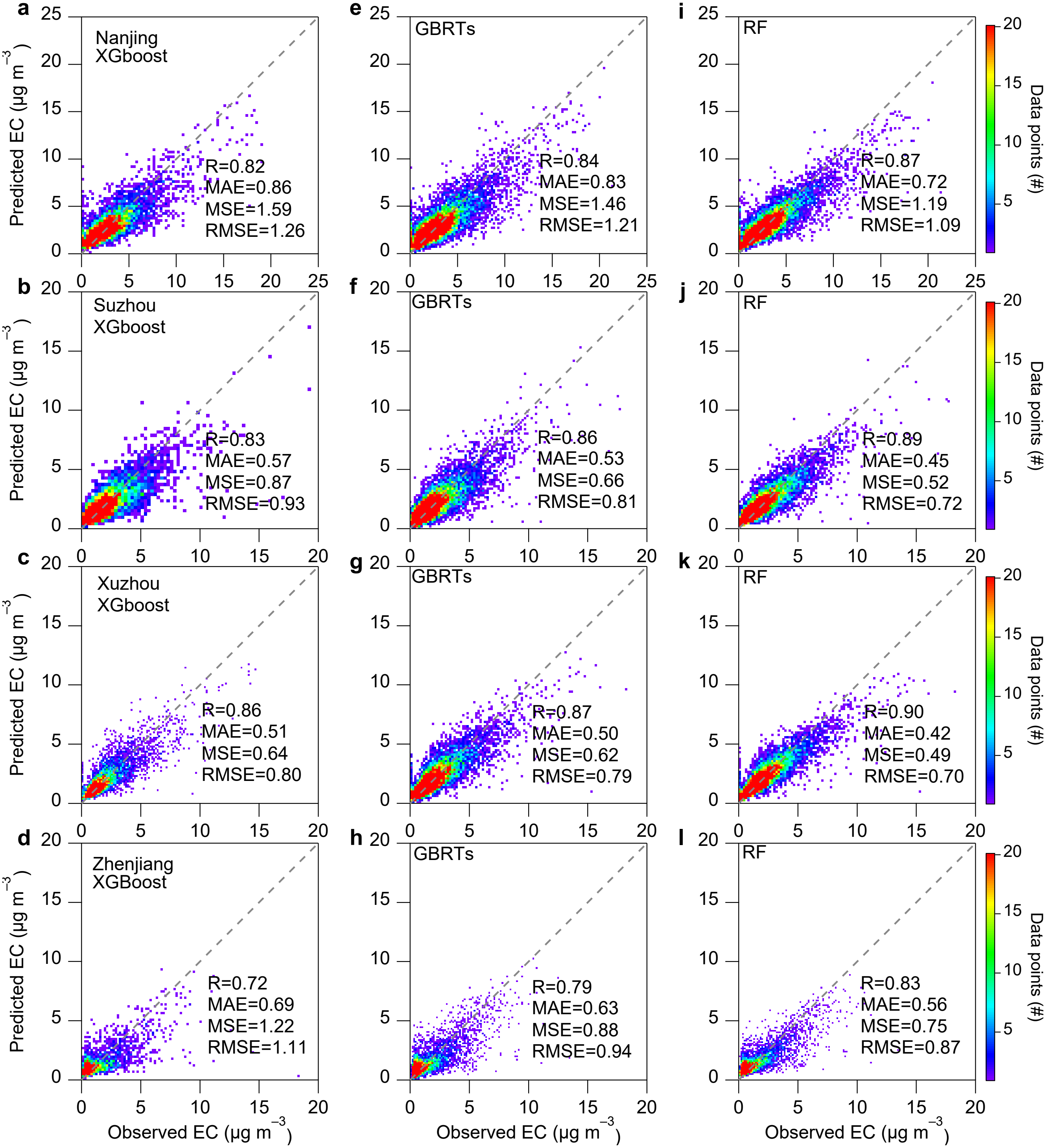

Figure S2. Performance evaluation of three machine learning models using the test set for reconstructing EC observation data (**a-d**, XGBoost; **e-h**, GBRTs; **i-l**, Random Forests).

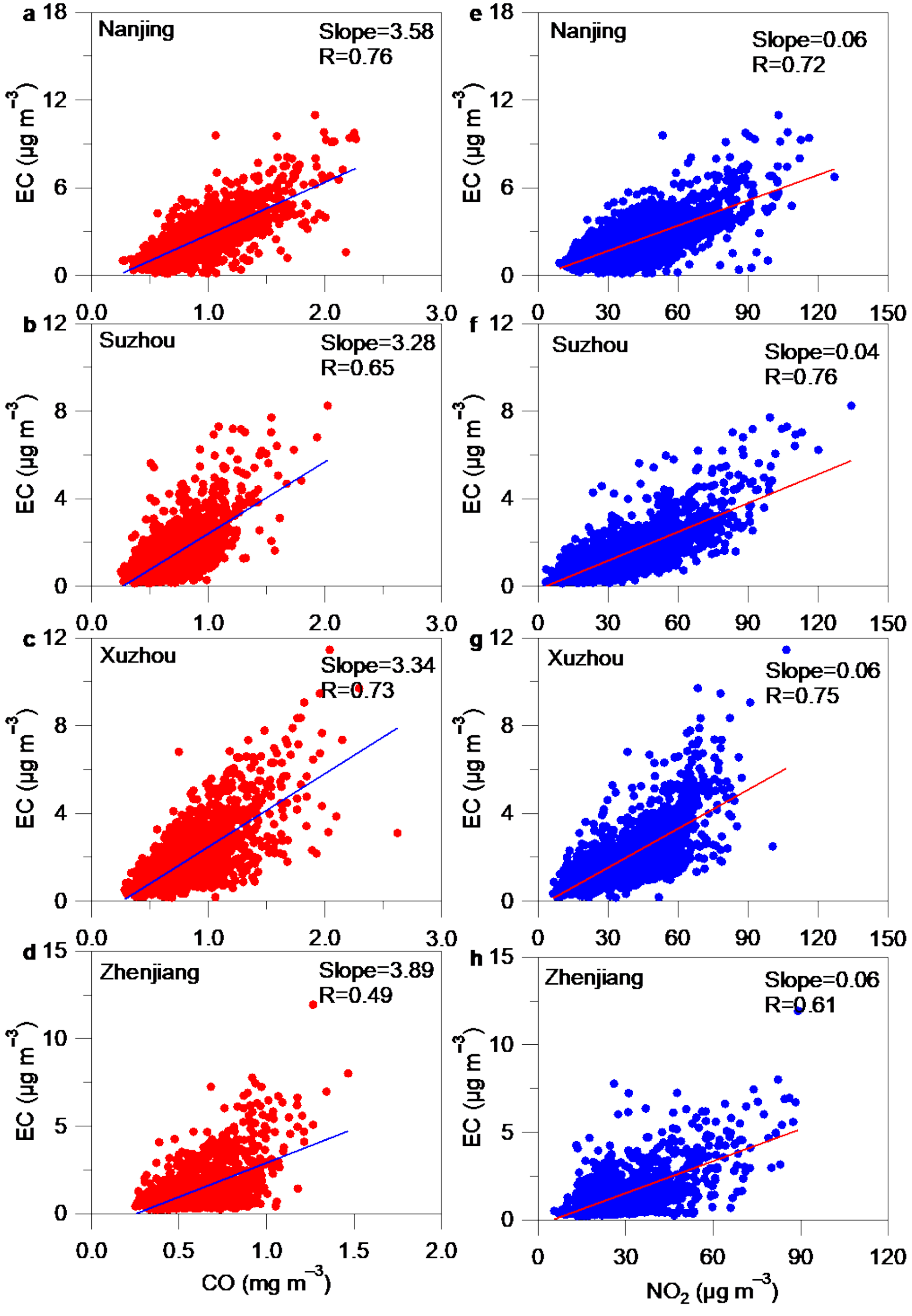

Figure S3. Correlation analysis between observed EC with CO and NO_2, respectively.

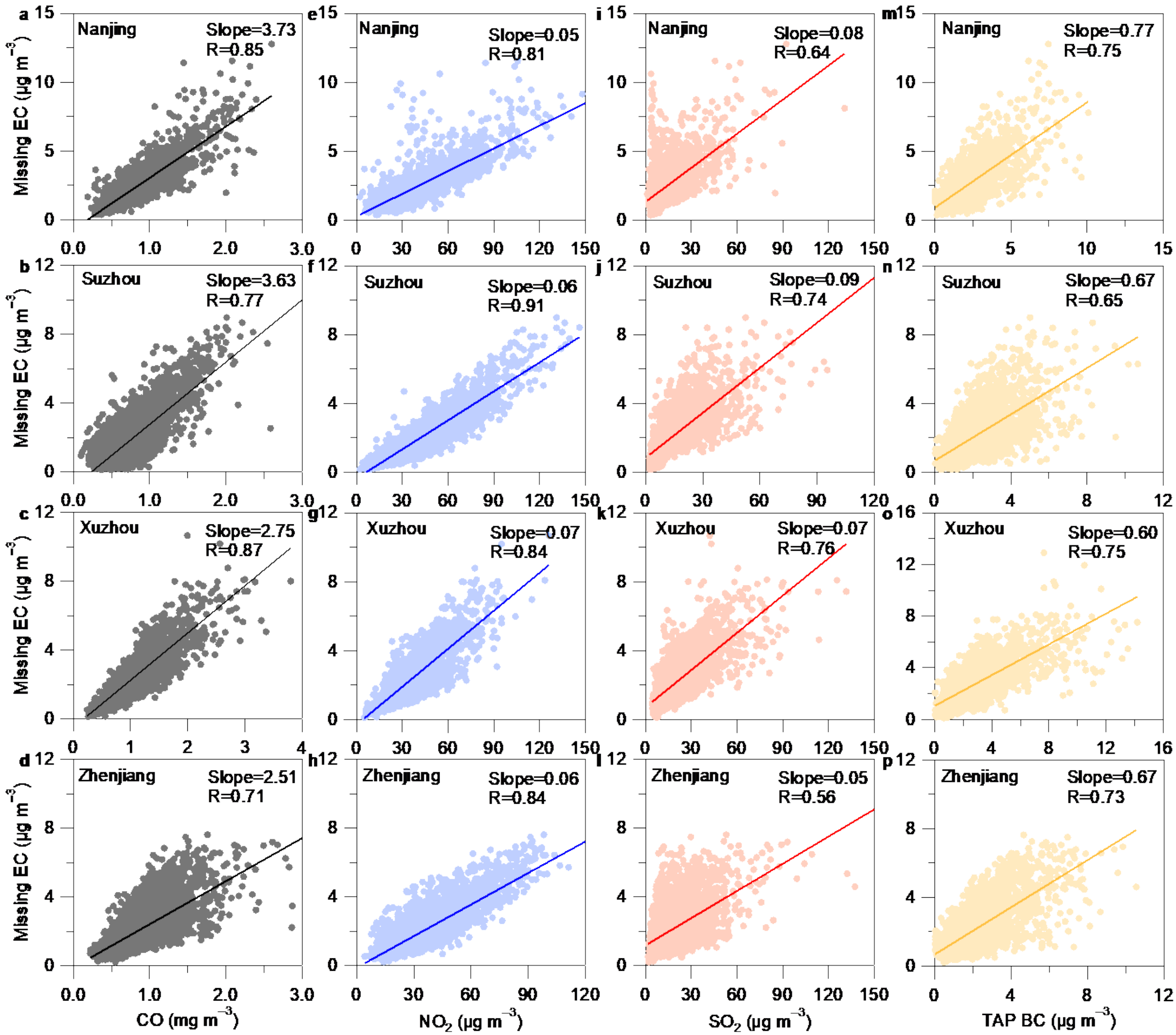

Figure S4. Correlation analysis between reconstructed EC and observed CO, SO_2, NO_2 and TAP BC.

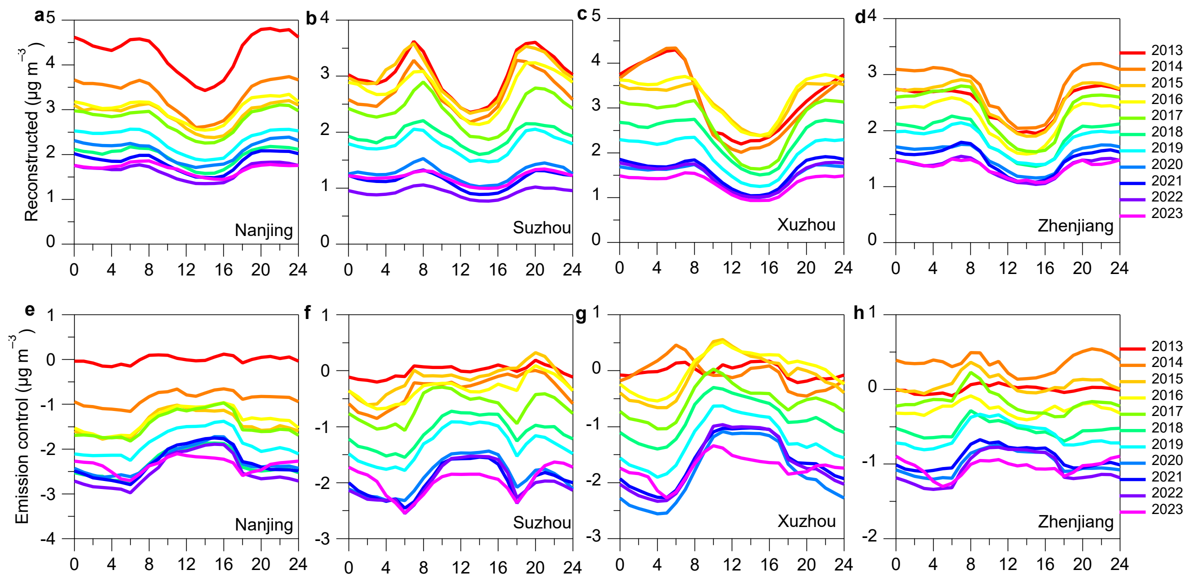

Figure S5. Annually diurnal variation of EC concentration. **a-d** Diurnal variation of EC in each year from 2013 to 2023. **e-h** Diurnal variation of emission-driven EC.

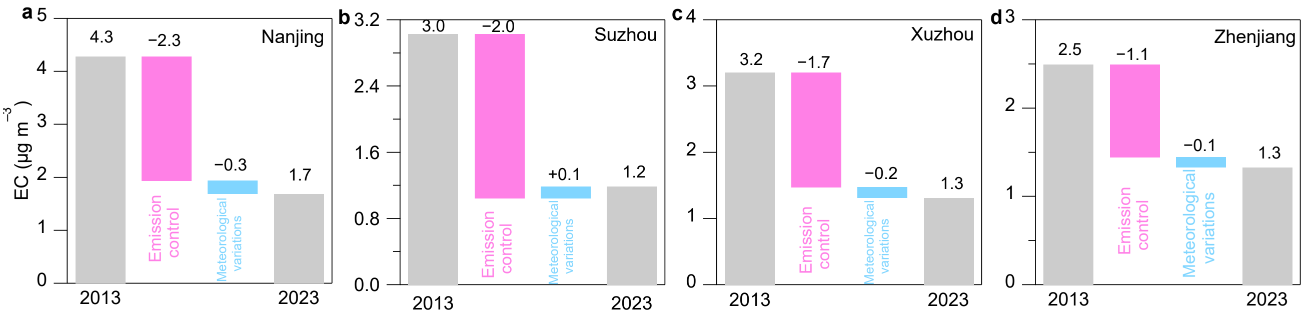

Figure S6. The contributions of EC concentration changes driven by anthropogenic emission control and meteorological conditions in four cities during 2013 and 2023.

Table S1. Summarization of field campaigns at four sites.

City	Longitude	Latitude
Nanjing	118.7538°	32.0551°
Suzhou	120.628°	31.2864°
Xuzhou	117.256°	34.2153°
Zhenjiang	119.6707	32.1875°

Table S2. Comparison Table of Meteorological Variables

Variable abbreviations	Meteorological variables	Unit
U10	10m u-component of wind	m s^{-1}
V10	10m v-component of wind	m s^{-1}
U850	850hPa u-component of wind	m s^{-1}
V850	850hPa v-component of wind	m s^{-1}
W850	850hPa w-component of wind	m s^{-1}
U650	650hPa u-component of wind	m s^{-1}
V650	650hPa v-component of wind	m s^{-1}
W650	650hPa w-component of wind	m s^{-1}
U500	500hPa u-component of wind	m s^{-1}
V500	500hPa v-component of wind	m s^{-1}
W500	500hPa w-component of wind	m s^{-1}
Tmx	Maximum 2m temperature	K
BLH	Boundary layer height	m
RH	Relative Humidity	Dimensionless
SR	Mean surface direct short-wave radiation flux	W m^{-2}
SP	Mean sea level pressure	Pa
TCC	Total cloud cover	Dimensionless
TP	Total precipitation	m

Table S3. Annual mean EC concentration data for the four representative cities in the Yangtze River Delta (Nanjing, Suzhou, Xuzhou, and Yangzhou) from 2013 to 2023.

Year	Nanjing ($\mu g\ m^{-3}$)	Suzhou ($\mu g\ m^{-3}$)	Xuzhou ($\mu g\ m^{-3}$)	Zhenjiang ($\mu g\ m^{-3}$)
2013	4.28	3.03	3.21	2.50
2014	3.33	2.72	3.11	2.78
2015	2.89	2.99	3.17	2.54
2016	3.01	2.76	3.26	2.21
2017	2.76	2.38	2.64	2.34
2018	1.97	1.93	2.30	1.89
2019	2.32	1.77	1.94	1.82
2020	2.09	1.27	1.47	1.54
2021	1.84	1.14	1.57	1.48
2022	1.63	0.92	1.49	1.34
2023	1.69	1.19	1.32	1.33

Table S4. Performance evaluation of test set for driving factor analysis model.

City	Method	R	MAE	MSE	RMSE
Nanjing	XGBoost	0.88	0.97	2.02	1.42
Nanjing	GBRTs	0.90	0.86	1.54	1.24
Nanjing	RF	0.92	0.76	1.35	1.16
Nanjing	EL	0.96	0.52	0.62	0.79
Suzhou	XGBoost	0.85	0.72	1.10	1.05
Suzhou	GBRTs	0.87	0.63	0.86	0.93
Suzhou	RF	0.89	0.58	0.81	0.90
Suzhou	EL	0.96	0.38	0.32	0.56
Xuzhou	XGBoost	0.82	0.71	1.17	1.08
Xuzhou	GBRTs	0.86	0.62	0.93	0.97
Xuzhou	RF	0.88	0.56	0.92	0.96
Xuzhou	EL	0.95	0.36	0.33	0.58
Zhenjiang	XGBoost	0.89	0.52	0.50	0.71
Zhenjiang	GBRTs	0.91	0.47	0.40	0.63
Zhenjiang	RF	0.93	0.42	0.34	0.59
Zhenjiang	EL	0.96	0.28	0.17	0.41

Table S5. Correlation analysis of data from the Reconstructed, MERRA-2, and Modeled methods with ground-based observation data.

City	Data	R	Slope
Nanjing	Reconstructed	0.97	0.86
Nanjing	MERRA-2	0.61	0.64
Nanjing	Modeled	0.68	0.54
Suzhou	Reconstructed	0.97	0.88
Suzhou	MERRA-2	0.59	0.87
Suzhou	Modeled	0.68	0.55
Xuzhou	Reconstructed	0.98	0.89
Xuzhou	MERRA-2	0.72	1.27
Xuzhou	Modeled	0.76	0.92
Zhenjiang	Reconstructed	0.97	0.87
Zhenjiang	MERRA-2	0.69	0.95
Zhenjiang	Modeled	0.65	0.38

Table S6. The results of the Mann-Kendall Test (MK) in this study are compared to those obtained using the De-weathered method.

City	Method	Slope	P_{value}
Nanjing	De-weathered	-0.208	<0.05
Nanjing	FEA	-0.214	<0.05
Suzhou	De-weathered	-0.221	<0.05
Suzhou	FEA	-0.229	<0.05
Xuzhou	De-weathered	-0.221	<0.05
Xuzhou	FEA	-0.204	<0.05
Zhenjiang	De-weathered	-0.142	<0.05
Zhenjiang	FEA	-0.138	<0.05

Table S7. Analysis of driving factors for CO, SO_2 and NO_2.

City	Year	Pollutant	Emission control	Meteorological variables
Nanjing	2013-2023	CO	-0.42 (mg m^{-3})	0.01 (mg m^{-3})
Nanjing	2013-2023	NO_2	-20.99(μg m^{-3})	-2.23(μg m^{-3})
Nanjing	2013-2023	SO_2	-16.60(μg m^{-3})	-2.65(μg m^{-3})
Suzhou	2013-2023	CO	-0.12 (mg m^{-3})	0.03 (mg m^{-3})
Suzhou	2013-2023	NO_2	-34.64(μg m^{-3})	2.17(μg m^{-3})
Suzhou	2013-2023	SO_2	-21.60(μg m^{-3})	1.00(μg m^{-3})
Xuzhou	2013-2023	CO	-0.83 (mg m^{-3})	-0.02 (mg m^{-3})
Xuzhou	2013-2023	NO_2	-18.52(μg m^{-3})	-0.65(μg m^{-3})
Xuzhou	2013-2023	SO_2	-42.74(μg m^{-3})	0.39(μg m^{-3})
Zhenjiang	2013-2023	CO	-0.57 (mg m^{-3})	0.03 (mg m^{-3})
Zhenjiang	2013-2023	NO_2	-3.18(μg m^{-3})	-0.37(μg m^{-3})
Zhenjiang	2013-2023	SO_2	-28.40(μg m^{-3})	1.25(μg m^{-3})