



Supplement of

Long-term Trends in PM_{2.5} Chemical Composition and Its Impact on Aerosol Properties: Field Observations from 2007 to 2020 in Pearl River Delta, South China

Yunfeng He et al.

Correspondence to: Xiang Ding (xiangd@gig.ac.cn) and Michael Boy (michael.boy@helsinki.fi)

The copyright of individual parts of the supplement might differ from the article licence.

S1. Eliminating anthropogenic impact on Cl^- by XGBoost (eXtreme Gradient Boosting)

XGBoost, a promising tool in machine learning (ML), has recently been used in atmospheric research. The performance of XGBoost surpasses that of traditional analysis methods for both nonlinear and linear questions. (Hou et al., 2022). The introduction and Python package is available online (<https://github.com/dmlc/xgboost>). In this study, we used this method to decouple the impacts of anthropogenic sources on Cl^- in $\text{PM}_{2.5}$ as much as possible. A total of 6 independent variables, including levoglucosan (biomass burning marker), picene (coal combustion marker), meteorological parameters (wind speed, temperature, RH, and solar radiation, representing marine activity) and one dependent variable, Cl^- , were fed to the model. All these samples were randomly divided into two groups: a training set accounting for 70% and a testing set accounting for 30% ($R^2 = 0.41$). Then, we replaced daily concentrations of levoglucosan and picene with their average concentrations in 2007-2020 to eliminate changes in anthropogenic sources on Cl^- . As shown in Fig S10, it decreased slightly at a rate of $-2\% \text{ yr}^{-1}$, indicating that the influence of marine emissions on $\text{PM}_{2.5}$ had been almost unchanged.

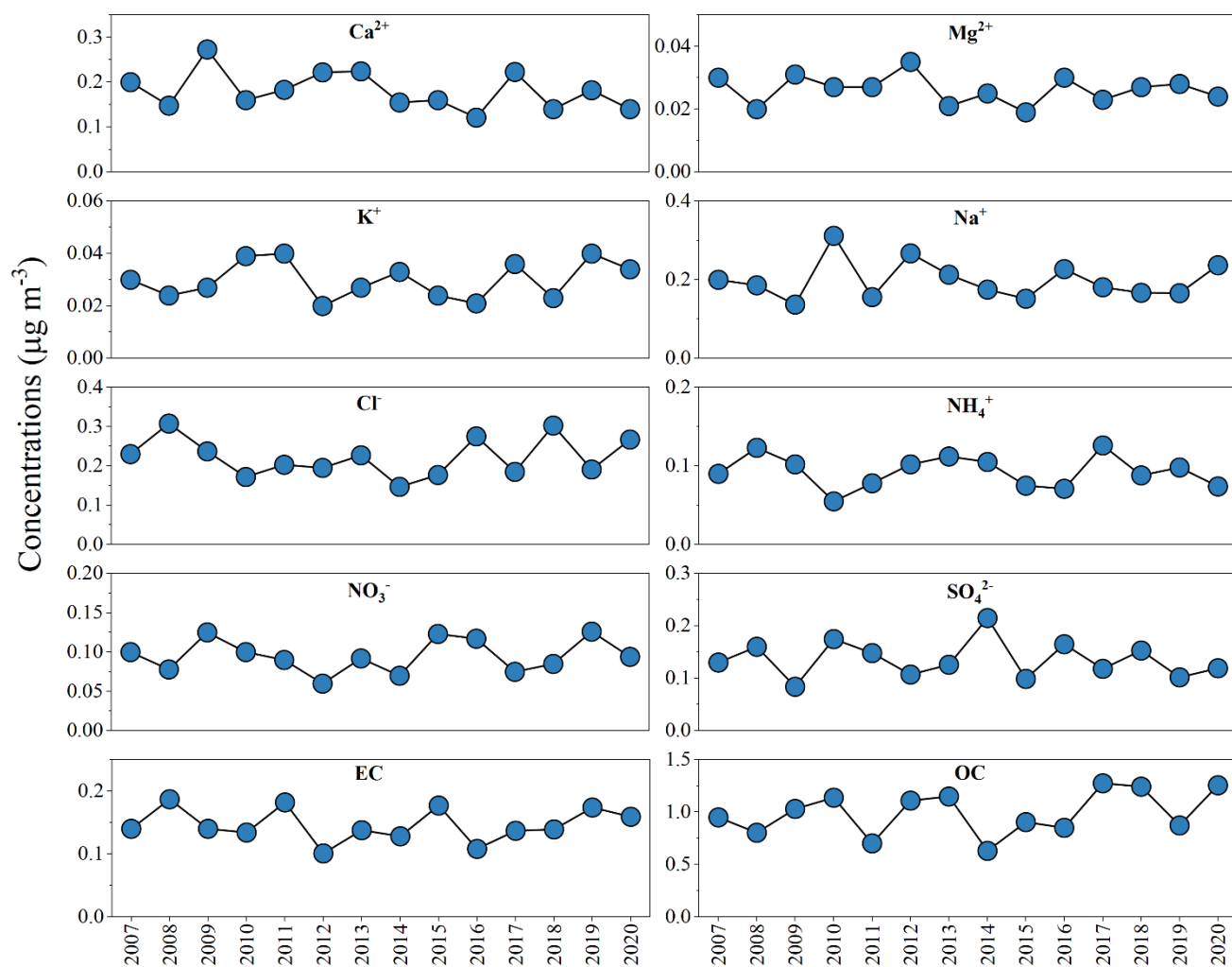
Table S1. The variations in PM_{2.5} main components, meteorological parameters, and other species from 2007 to 2020.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
I. Gaseous pollutants ($\mu\text{g m}^{-3}$)														
SO₂	71.3 ± 21.4	73.5 ± 20.1	59.8 ± 20.3	49.5 ± 20.3	35.6 ± 9.9			32.4 ± 11.0	23.8 ± 8.8	20.8 ± 7.1	26.4 ± 6.3	18.3 ± 3.9	14.0 ± 3.2	12.2 ± 3.5
NO₂	58.3 ± 10.3	72.1 ± 28.0	71.0 ± 16.5	49.6 ± 16.3	59.0 ± 14.8			54.2 ± 17.8	49.1 ± 17.3	40.6 ± 19.9	54.8 ± 19.8	53.4 ± 16.3	47.6 ± 19.7	44.7 ± 13.5
O₃	75.3 ± 31.4	63.2 ± 17.3	51.6 ± 25.0	69.1 ± 27.0	54.3 ± 22.8			53.0 ± 18.1	78.9 ± 37.0	59.9 ± 31.2	72.0 ± 38.7	60.1 ± 27.7	105.1 ± 21.9	74.5 ± 37.5
II. PM_{2.5} main components ($\mu\text{g m}^{-3}$)														
PM_{2.5}	87.1 ± 15.5	81.2 ± 18.0	72.4 ± 21.3	73.7 ± 37.5	64.2 ± 13.6	47.1 ± 14.0	65.4 ± 24.2	53.2 ± 14.6	43.0 ± 17.9	33.8 ± 12.8	49.1 ± 18.2	36.0 ± 9.6	40.2 ± 11.2	34.0 ± 11.3
OM	30.9 ± 7.6	36.3 ± 12.6	27.5 ± 11.2	26.9 ± 12.7	24.3 ± 8.7	15.6 ± 6.9	30.4 ± 14.1	16.7 ± 6.1	10.6 ± 5.1	9.5 ± 4.1	14.7 ± 4.9	13.2 ± 4.8	13.7 ± 3.6	14.6 ± 4.3
POC	10.4 ± 3.1	13.5 ± 3.4	10.9 ± 4.5	11.0 ± 4.9	9.4 ± 3.1	4.5 ± 1.4	11.0 ± 3.6	8.2 ± 4.6	4.4 ± 1.9	4.4 ± 1.7	4.3 ± 2.1	5.0 ± 2.1	4.3 ± 1.8	4.4 ± 1.1
SOC	7.7 ± 2.2	7.2 ± 2.4	5.4 ± 1.9	4.4 ± 2.0	4.7 ± 1.5	3.9 ± 1.6	6.0 ± 4.3	1.6 ± 0.5	1.6 ± 0.7	1.1 ± 0.4	4.7 ± 2.0	2.5 ± 0.9	4.0 ± 1.2	3.9 ± 1.7
EC	3.6 ± 1.1	4.2 ± 1.1	5.5 ± 2.3	3.1 ± 1.4	3.1 ± 1.0	1.9 ± 0.6	2.5 ± 0.8	3.0 ± 1.7	1.4 ± 0.6	2.3 ± 0.9	1.9 ± 0.9	1.7 ± 0.7	1.4 ± 0.6	1.6 ± 0.4
SO₄²⁻	22.2 ± 6.0	17.1 ± 5.5	17.0 ± 5.9	16.3 ± 6.5	14.2 ± 4.7	10.5 ± 4.4	13.1 ± 9.1	9.7 ± 3.2	10.1 ± 4.6	8.0 ± 2.7	8.1 ± 3.4	6.1 ± 2.0	6.4 ± 1.9	6.6 ± 2.9
NO₃⁻	6.7 ± 3.1	9.2 ± 4.2	11.5 ± 4.6	8.4 ± 4.8	9.6 ± 4.0	5.8 ± 5.4	9.6 ± 8.4	3.6 ± 2.3	5.6 ± 5.2	4.6 ± 3.8	5.5 ± 3.8	5.3 ± 2.9	3.1 ± 2.2	6.3 ± 3.4
NH₄⁺	6.6 ± 1.7	4.9 ± 2.4	7.1 ± 2.3	6.9 ± 3.8	6.6 ± 2.3	4.8 ± 2.1	6.6 ± 4.4	3.7 ± 1.4	4.5 ± 2.4	3.4 ± 1.7	4.0 ± 1.7	3.1 ± 1.3	3.0 ± 0.9	3.4 ± 1.4
Cl⁻	1.01 ± 0.54	1.61 ± 1.27	1.80 ± 1.03	1.49 ± 1.16	1.46 ± 0.93	1.21 ± 0.74	1.44 ± 1.12	0.38 ± 0.33	0.55 ± 0.37	0.52 ± 0.34	0.75 ± 0.66	0.75 ± 0.72	0.36 ± 0.36	0.39 ± 0.15
Na⁺	0.97 ± 0.66	0.93 ± 0.61	0.89 ± 0.18	0.60 ± 0.29	0.56 ± 0.16	0.40 ± 0.14	0.56 ± 0.26	0.36 ± 0.10	0.27 ± 0.08	0.25 ± 0.14	0.34 ± 0.14	0.52 ± 0.34	0.27 ± 0.14	0.48 ± 0.20
K⁺	1.49 ± 0.57	2.23 ± 0.01	0.97 ± 0.41	1.20 ± 0.61	1.14 ± 0.46	0.69 ± 0.32	1.13 ± 0.81	0.60 ± 0.21	0.49 ± 0.25	0.30 ± 0.13	0.59 ± 0.24	0.35 ± 0.15	0.46 ± 0.15	0.48 ± 0.22
Mg²⁺	0.15 ± 0.05	0.11 ± 0.05	0.23 ± 0.09	0.04 ± 0.04	0.08 ± 0.02	0.07 ± 0.03	0.10 ± 0.05	0.07 ± 0.02	0.05 ± 0.01	0.04 ± 0.01	0.04 ± 0.02	0.05 ± 0.02	0.05 ± 0.02	0.21 ± 0.12
Ca²⁺	1.30 ± 0.57	1.46 ± 0.72	0.27 ± 0.17	0.33 ± 0.14	0.53 ± 0.12	0.47 ± 0.35	0.85 ± 0.49	0.64 ± 0.16	0.27 ± 0.21	0.26 ± 0.13	0.52 ± 0.22	0.45 ± 0.15	0.27 ± 0.14	0.10 ± 0.03
III. Meteorological parameters														
Temperature (°C)	22.2 ± 2.1	17.2 ± 2.9	17.0 ± 3.1	19.7 ± 3.2	19.9 ± 3.8	22.1 ± 1.4	20.9 ± 1.1	20.2 ± 4.4	25.1 ± 2.4	23.8 ± 3.8	21.3 ± 3.2	22.4 ± 2.9	21.7 ± 2.7	20.2 ± 5.1
RH (%)	57 ± 11	47 ± 12	67 ± 13	64 ± 11	57 ± 11	61 ± 7	50 ± 13	57 ± 14	63 ± 8	67 ± 7	56 ± 13	63 ± 12	48 ± 11	55 ± 13
SSR (W m⁻²)	161.3 ± 41.3	156.5 ± 28.3	135.2 ± 36.3	141.8 ± 51.4	134.7 ± 36.9	101.2 ± 40.4	125.2 ± 46.2	95.3 ± 49.1	149.0 ± 36.2	122.4 ± 44.0	128.9 ± 53.1	122.2 ± 39.9	165.0 ± 29.4	145.5 ± 31.9
Wind speed (m s⁻¹)	1.3 ± 0.2	1.3 ± 0.4	1.7 ± 0.5	1.7 ± 0.7	1.8 ± 0.9	1.2 ± 0.6	1.3 ± 0.3	1.9 ± 0.6	1.5 ± 0.3	1.6 ± 0.4	1.7 ± 0.4	1.6 ± 0.5	1.6 ± 0.3	1.7 ± 0.4
IV. Other species														
ALWC ($\mu\text{g m}^{-3}$)	20.6 ± 10.0	11.3 ± 7.9	28.8 ± 11.4	22.0 ± 10.8	19.3 ± 9.9	13.5 ± 5.6	12.0 ± 7.5	10.8 ± 5.4	12.5 ± 6.2	12.0 ± 7.3	9.9 ± 6.5	11.0 ± 6.7	5.1 ± 2.5	9.5 ± 3.9
pH	1.51 ± 1.07	2.60 ± 0.71	1.94 ± 0.29	1.97 ± 1.00	2.54 ± 0.37	2.55 ± 0.43	2.69 ± 0.42	2.29 ± 0.33	2.13 ± 0.33	2.05 ± 0.46	2.60 ± 0.45	2.66 ± 0.37	2.31 ± 0.63	2.86 ± 0.49

Table S2. Sensitivity analysis of long-term trends with and without the year 2020.

	Slope (with 2020)	Slope (without 2020)	Uncertainty
PM _{2.5}	-4.0 **	-4.2 **	4%
OM	-1.70 **	-1.91 **	12%
EC	-0.23 **	-0.25 **	9%
SO ₄ ²⁻	-1.13 **	-1.21 **	7%
NO ₃ ⁻	-0.40 **	-0.47 **	18%
NH ₄ ⁺	-0.31 **	-0.33 **	6%
Cl ⁻	-0.10 **	-0.10 **	1%
Na ⁺	-0.05 **	-0.06 **	20%
K ⁺	-0.10 **	-0.12 **	16%
Mg ²⁺		-0.01 *	
Ca ²⁺	-0.06 *	-0.06 *	3%

15 One asterisk, two asterisks denote p value < 0.05, 0.01, respectively. Blank cells denote p value > 0.05. The uncertainty was calculated as:
 $Uncertainty = \frac{|Slope_{with} - Slope_{without}|}{|Slope_{with}|}$.



20 Figure S1. Annual variations of measured components in blank filter samples.

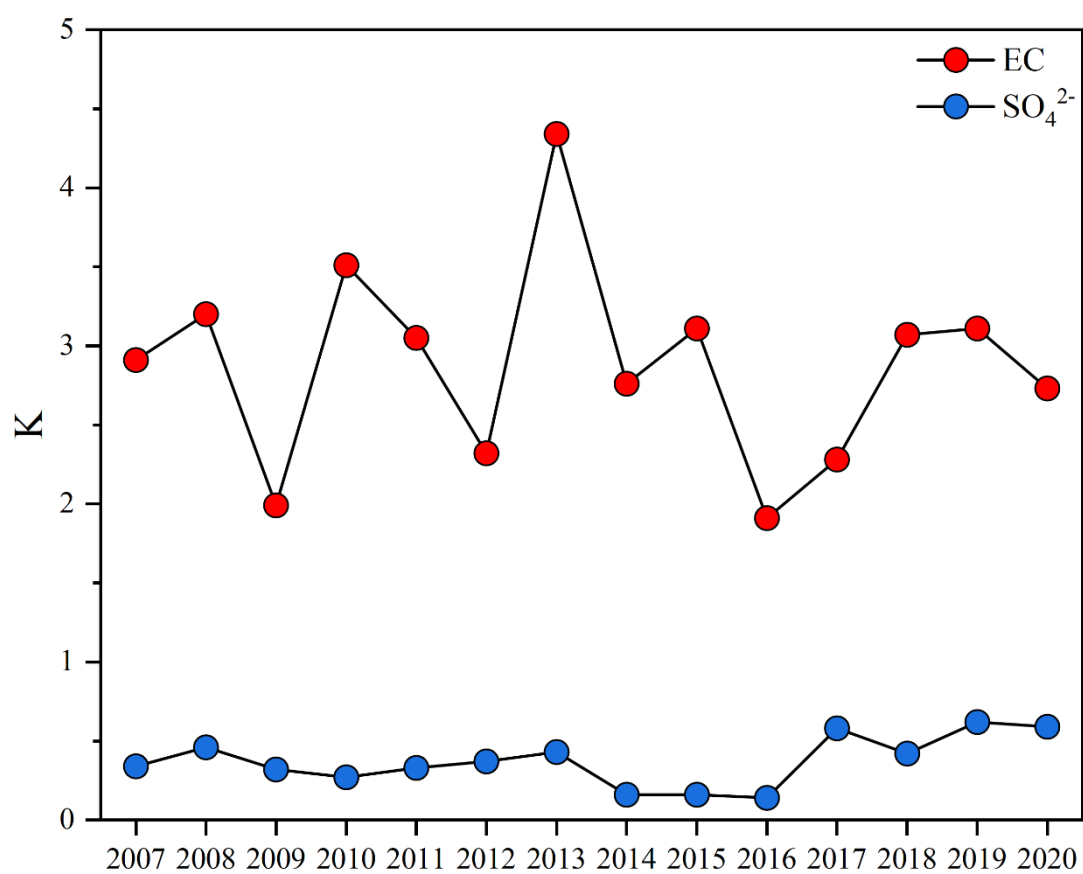
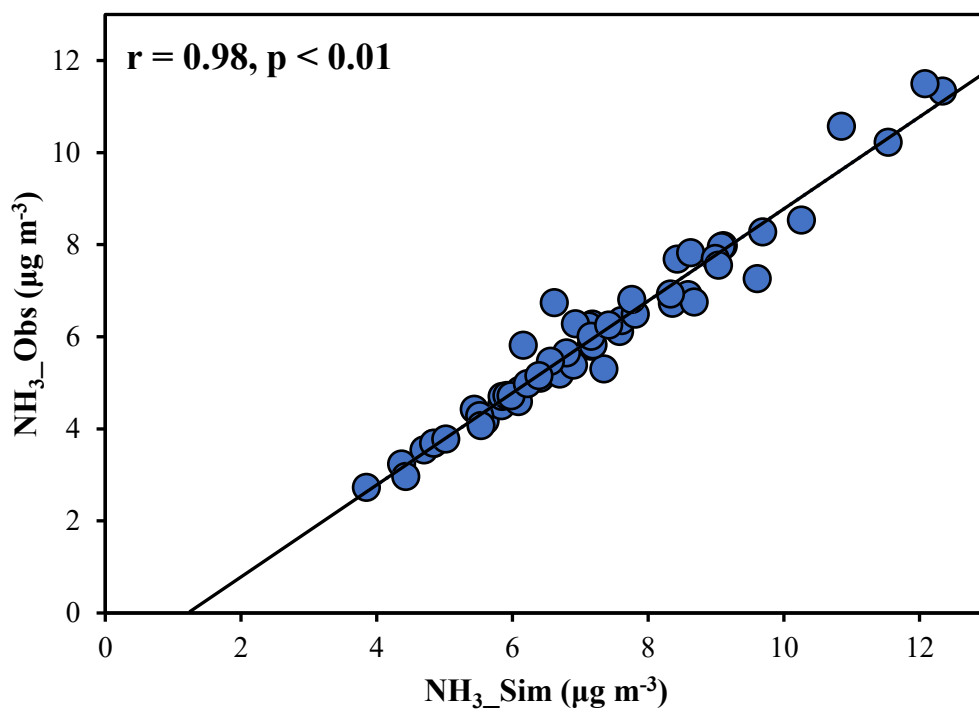


Figure S2. Variations of K values for EC and SO_4^{2-} , calculated by Bayesian Inference approach.



25

Figure S3. Observations and ISORROPIA-II simulation of gas-phase NH_3 in GIG.

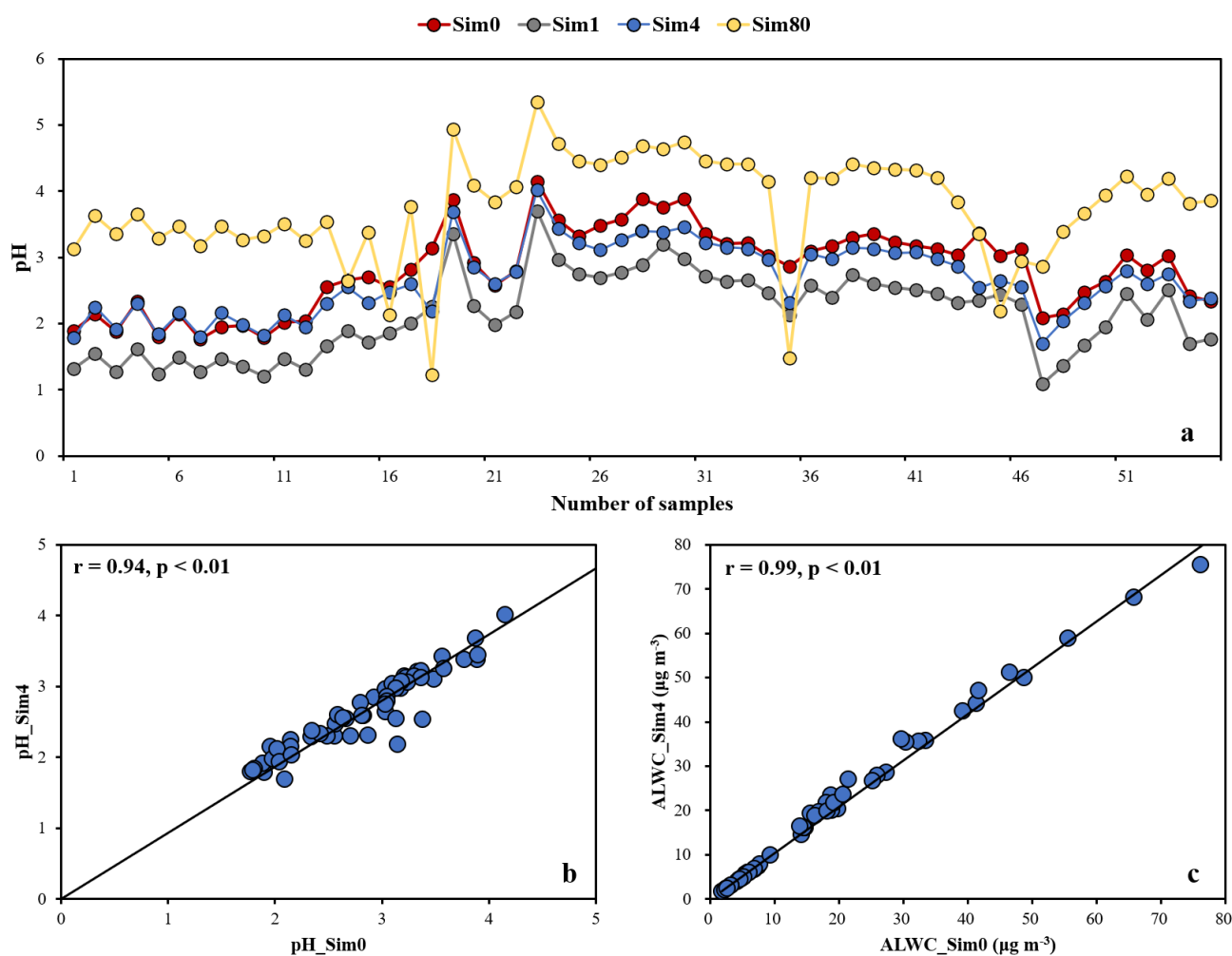
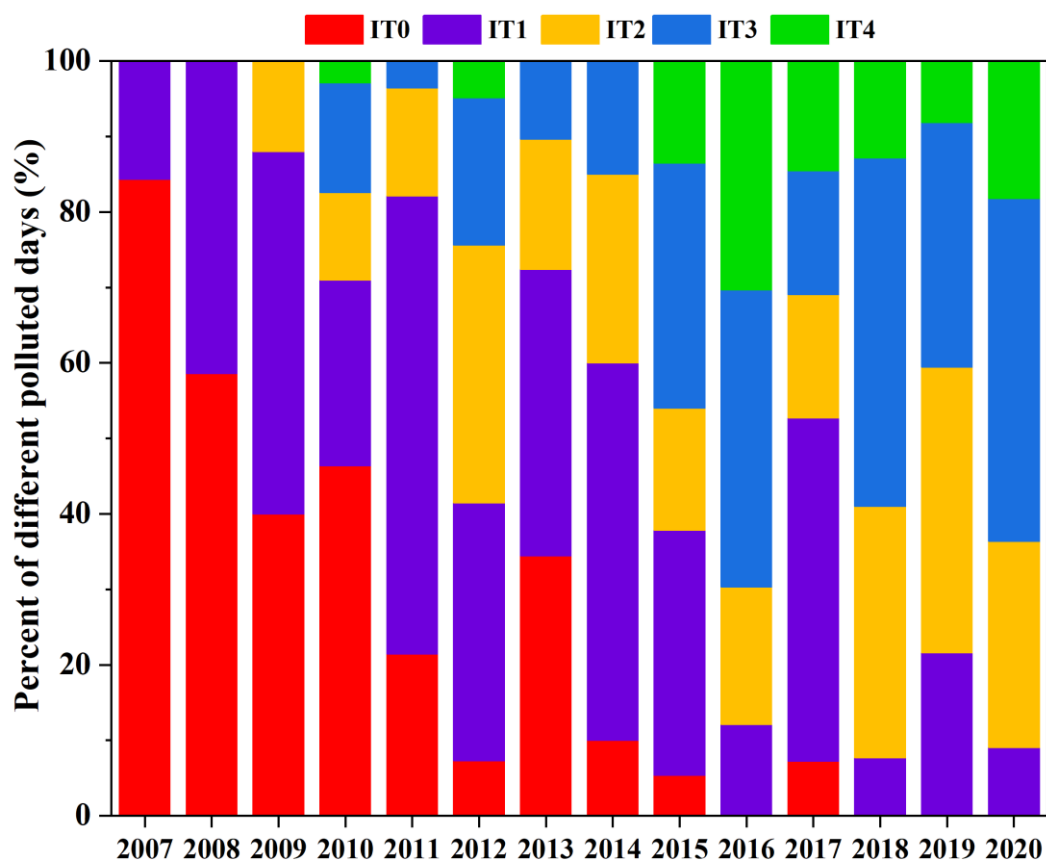


Figure S4. Comparison of predicted pH, ALWC with and without gas-phase input in GIG. (a) Sim0 represents results with gas-phase input, Sim1, Sim4, Sim80 represent results without gas-phase input and performed one, four, and eighty iterations respectively. It showed predicted aerosol pH rose as number of iterations increased. Four iterations without gas-phase input would generate the optimal result. (b) and (c) showed pH and ALWC calculated by this way had the greatest agreement with the one which has gas-phase input.



35 Figure S5. Percent of different polluted days in WQS from 2007 to 2020.

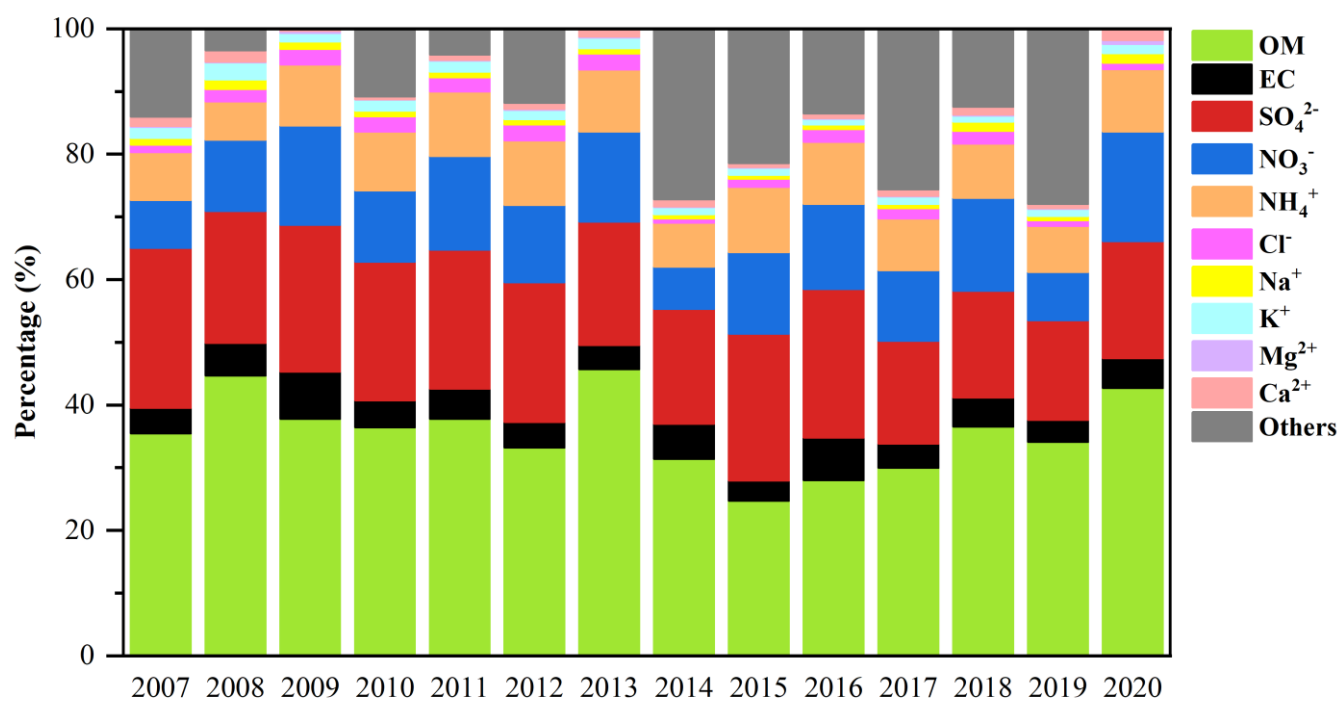


Figure S6. The variations in mass fraction of PM_{2.5} chemical composition.

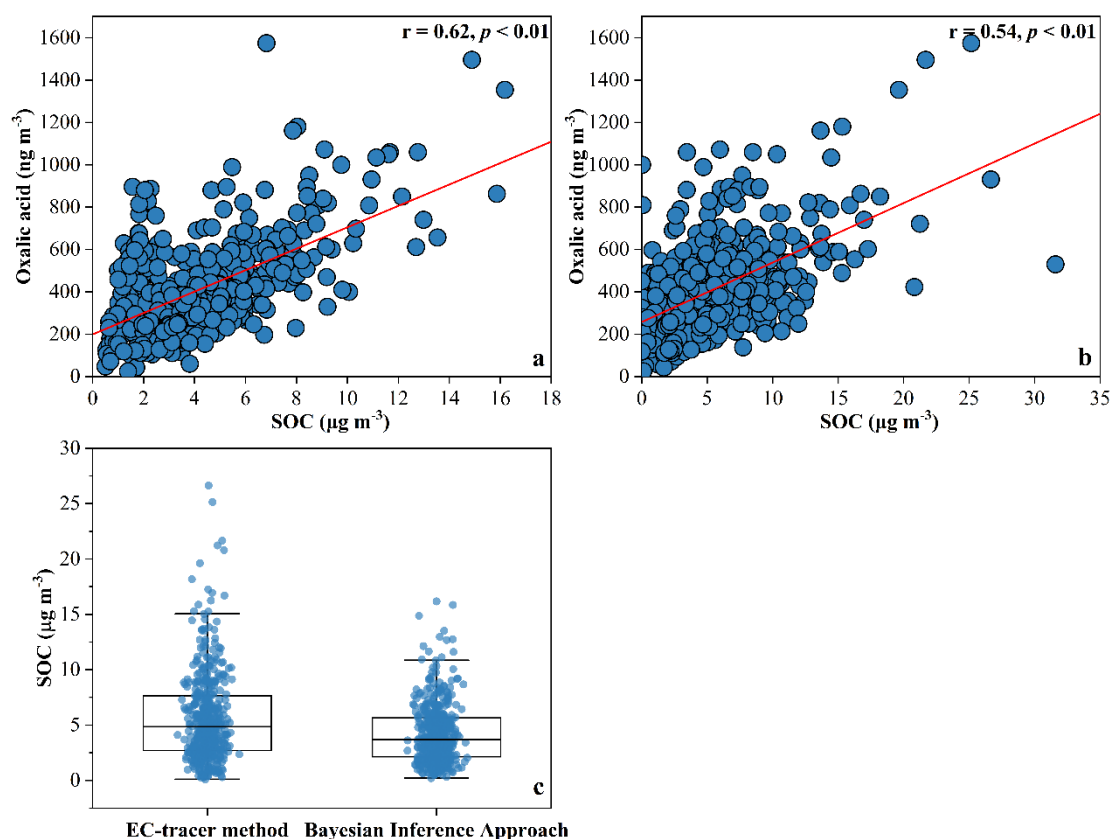


Figure S7. Oxalic acid, a typical secondary organic molecular tracer, exhibited a significant correlation ($p < 0.01$) with SOC estimated using both Bayesian Inference approach (a) and EC-tracer method (b). The Pearson's correlation coefficient for the Bayesian Inference approach (0.62) was higher than that for the EC-tracer method (0.54), suggesting that SOC derived from the Bayesian Inference approach was more reliable. In addition, the average SOC concentration estimated by the Bayesian Inference approach ($4.1 \pm 2.6 \mu\text{g m}^{-3}$) was lower than that by the EC-tracer method ($5.8 \pm 4.2 \mu\text{g m}^{-3}$) (c).

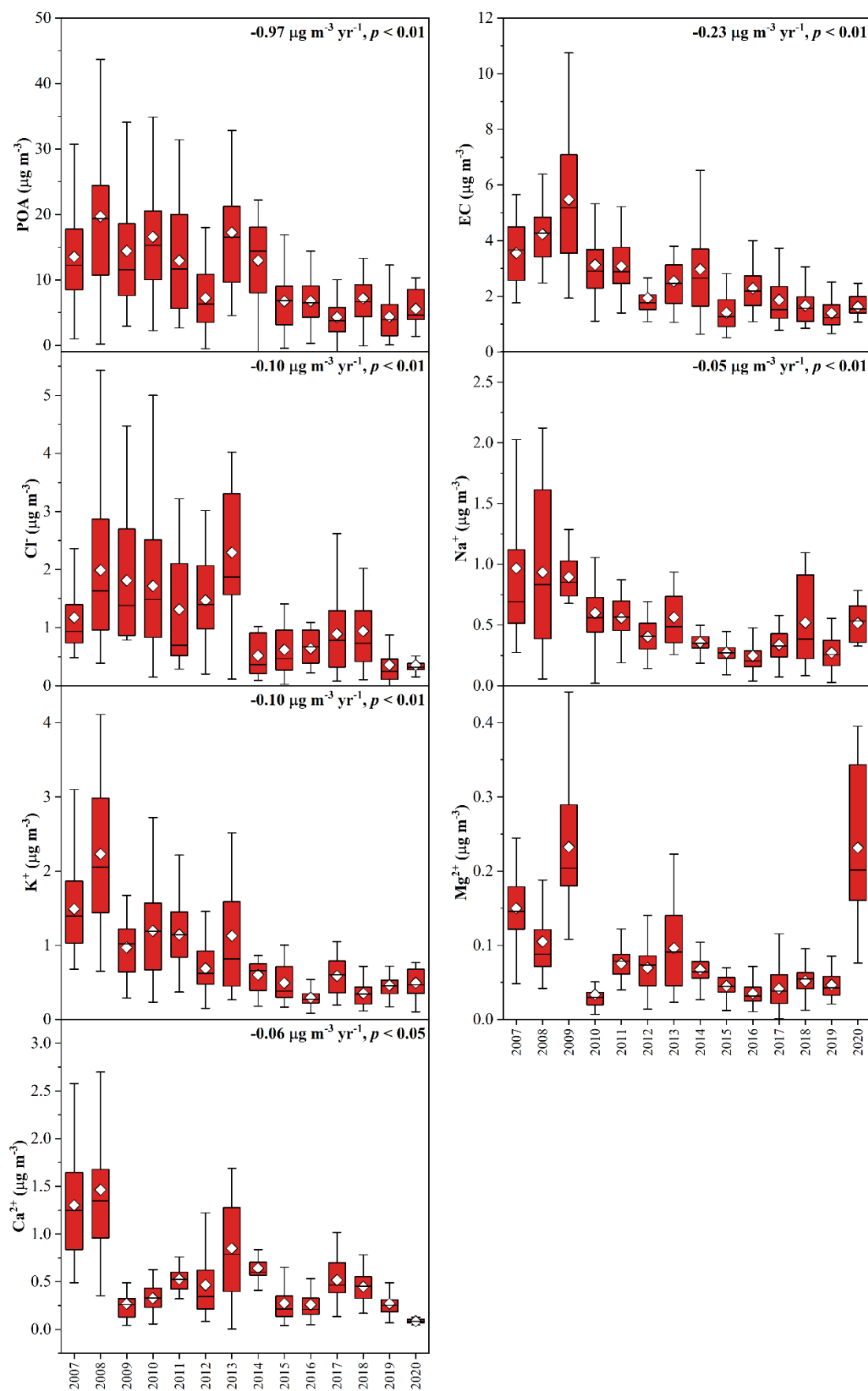
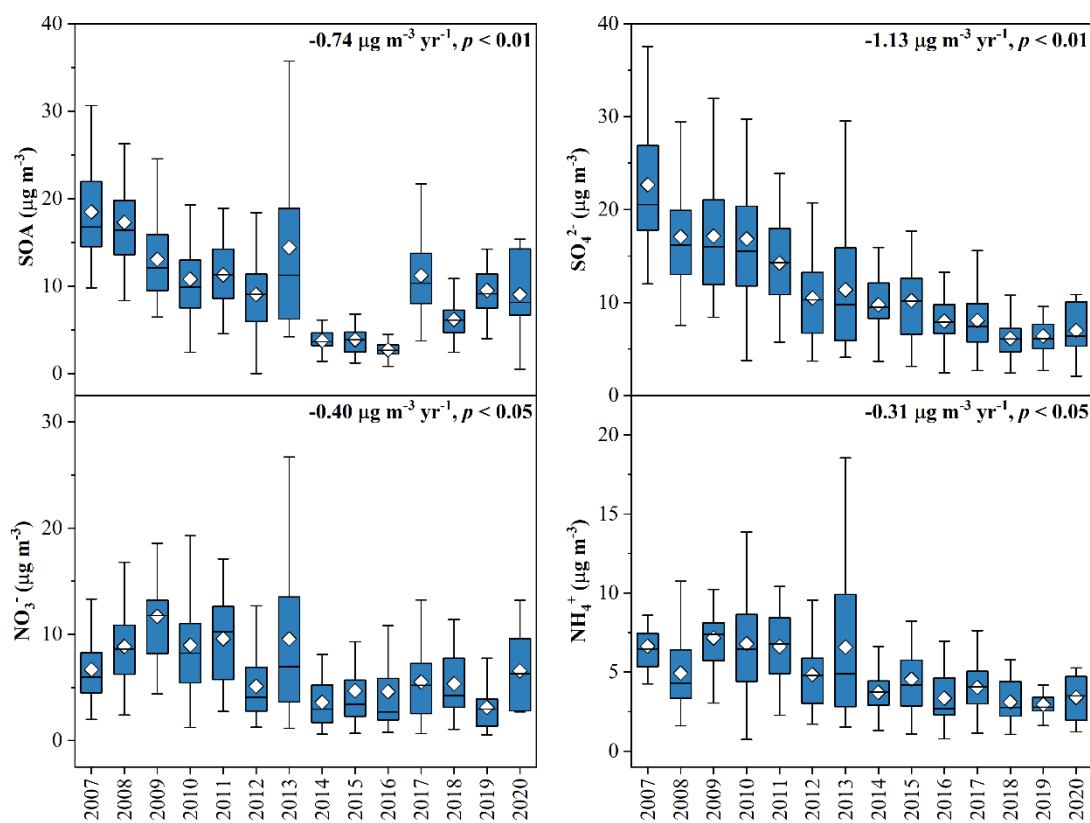


Figure S8. Concentrations of PM_{2.5} primary species during 2007-2020.



50 **Figure S9. Concentrations of PM_{2.5} secondary species during 2007-2020.**

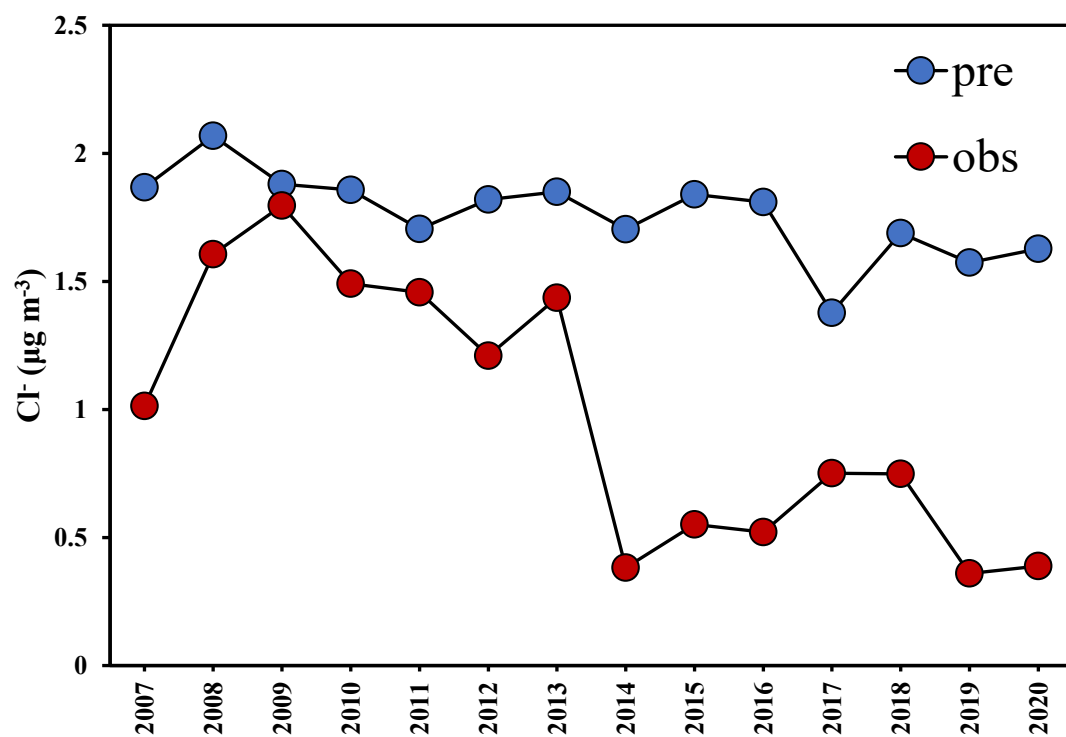
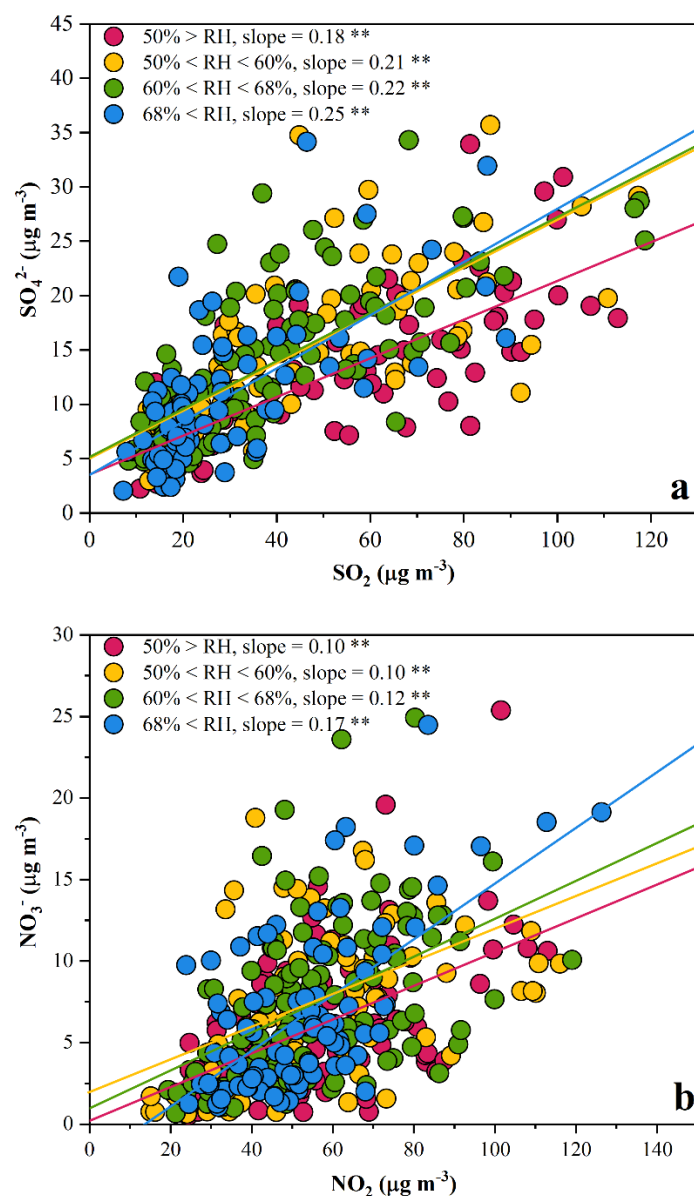
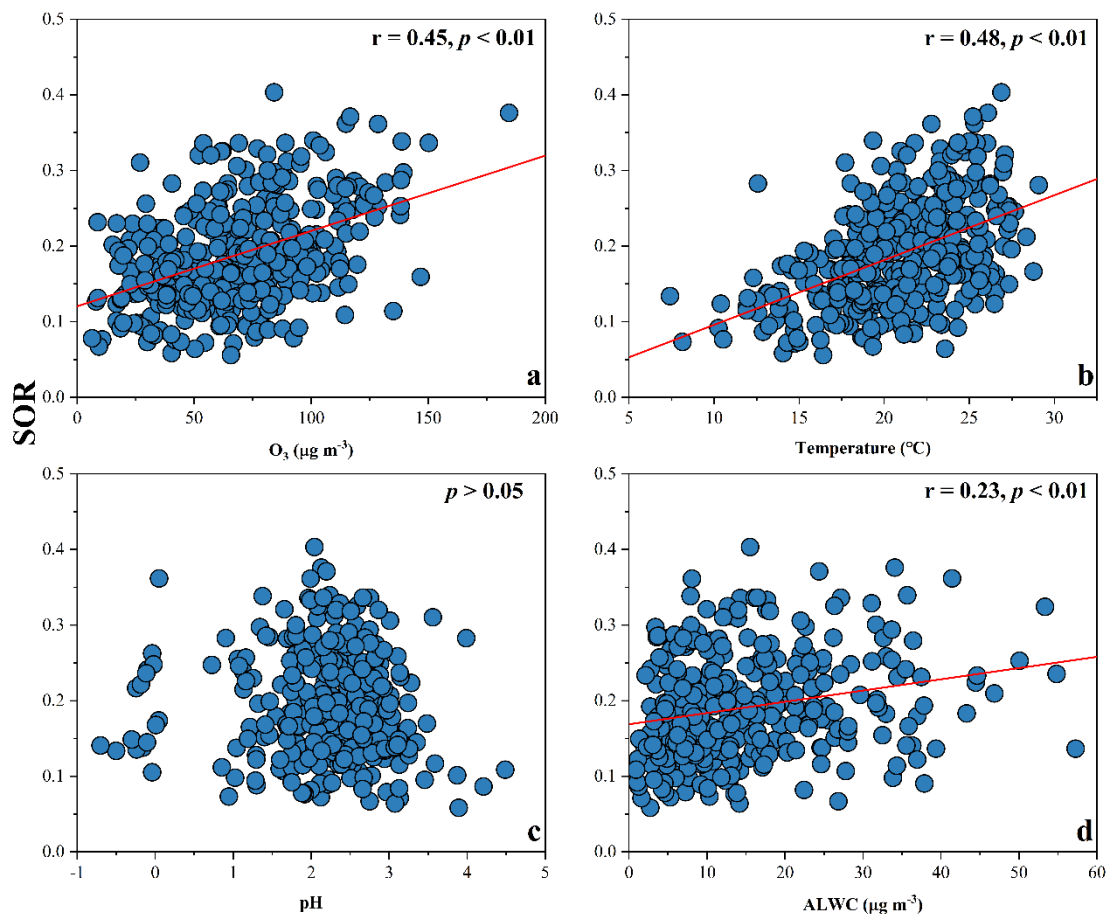


Figure S10. The observation (red) and prediction (blue) concentration of Cl^- . After eliminating variations in anthropogenic sources on Cl^- , it decreased slightly at a rate of $-2\% \text{ yr}^{-1}$ during 2007-2020 (blue).



55 **Figure S11. Correlations between $\text{SO}_4^{2-}/\text{SO}_2$ (a), as well as $\text{NO}_3^-/\text{NO}_2$ (b). Two asterisks denote p value less than 0.01. All samples were categorized into four groups according to the quartile ranges of RH. The slope became greater with rising RH, indicating conversion of primary pollutants to secondary species was more efficient.**



60 Figure S12. Correlations between sulfur oxidation rate (SOR) and O_3 (a), temperature (b), pH (c), as well as ALWC (d).

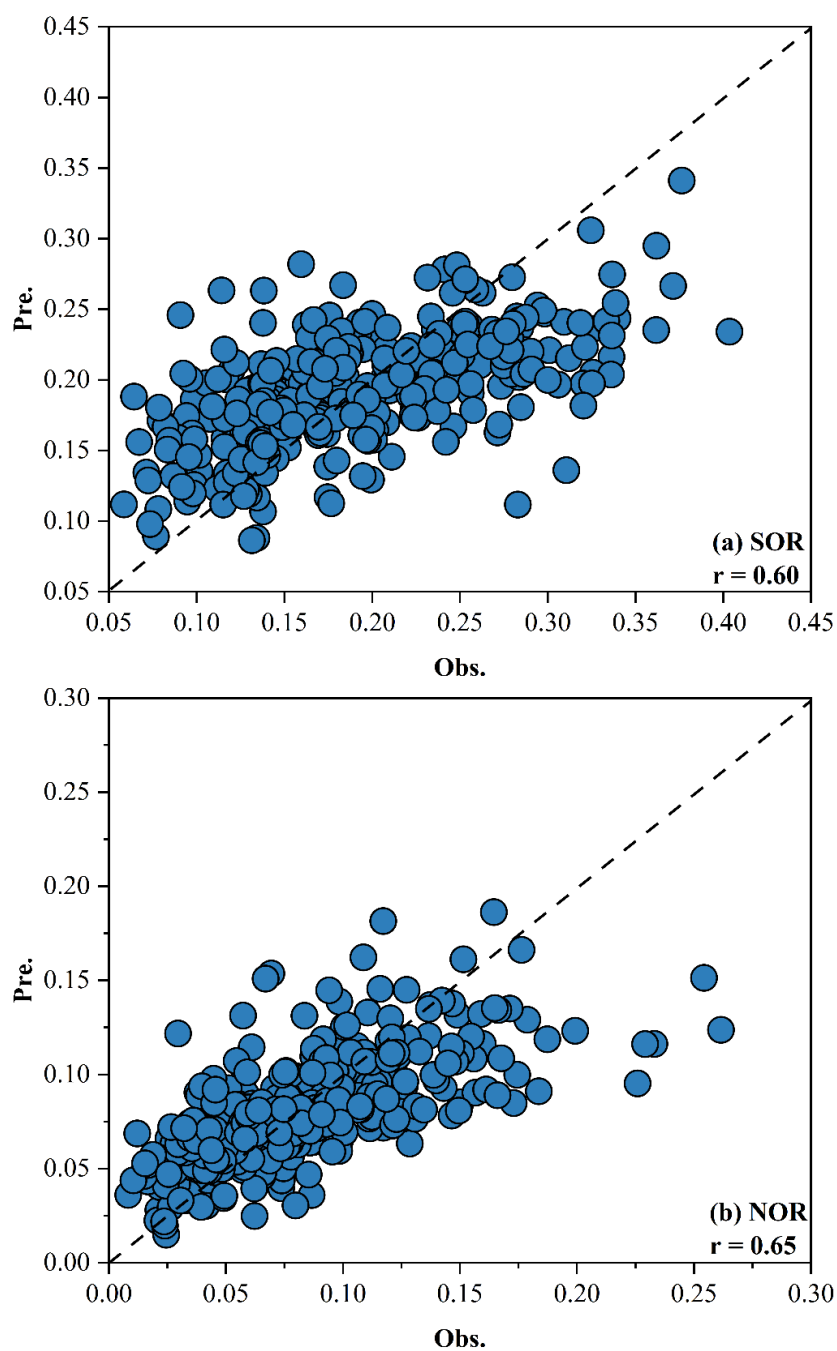


Figure S13. Correlations between predictions and observations of SOR (a) and NOR (b). Solid line is 1:1.

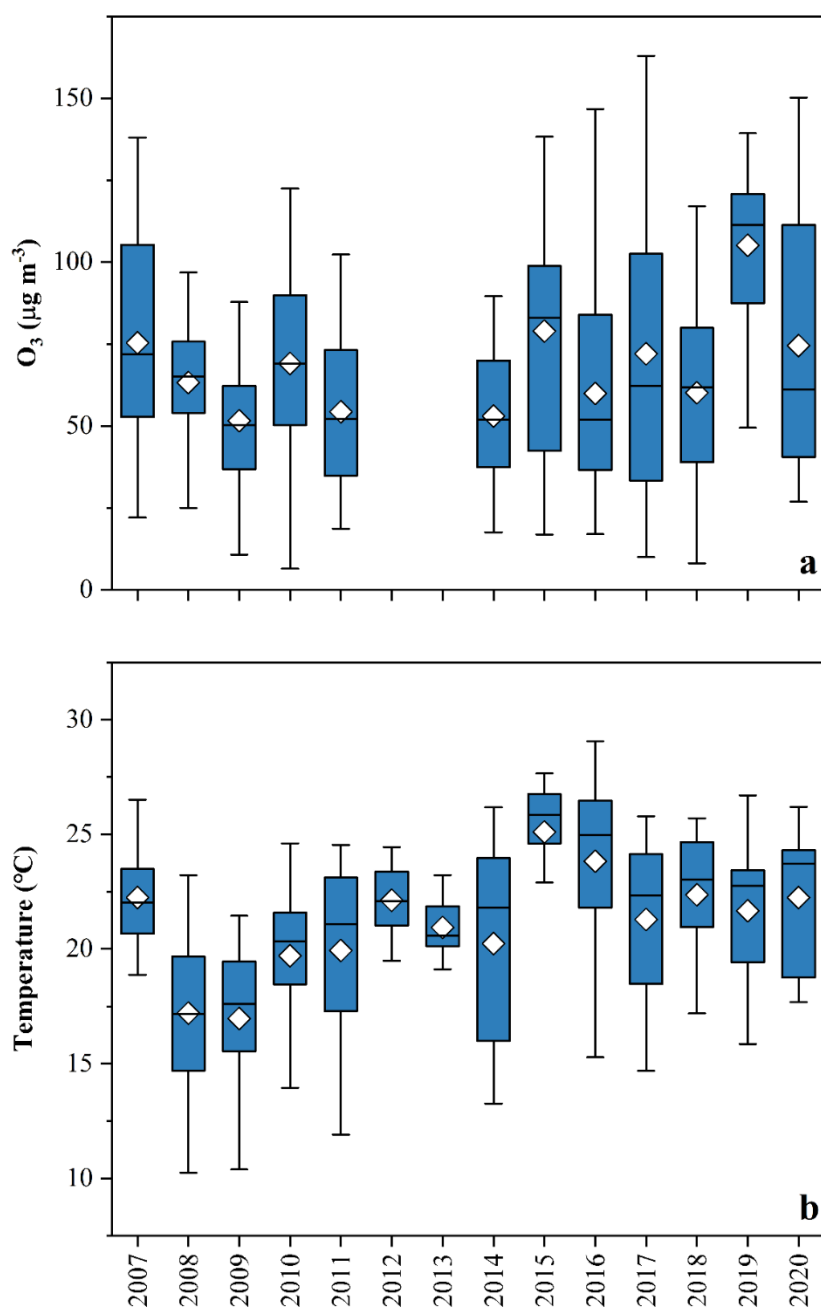


Figure S14. Variations of O₃ (a) and temperature (b) during 2007-2020.

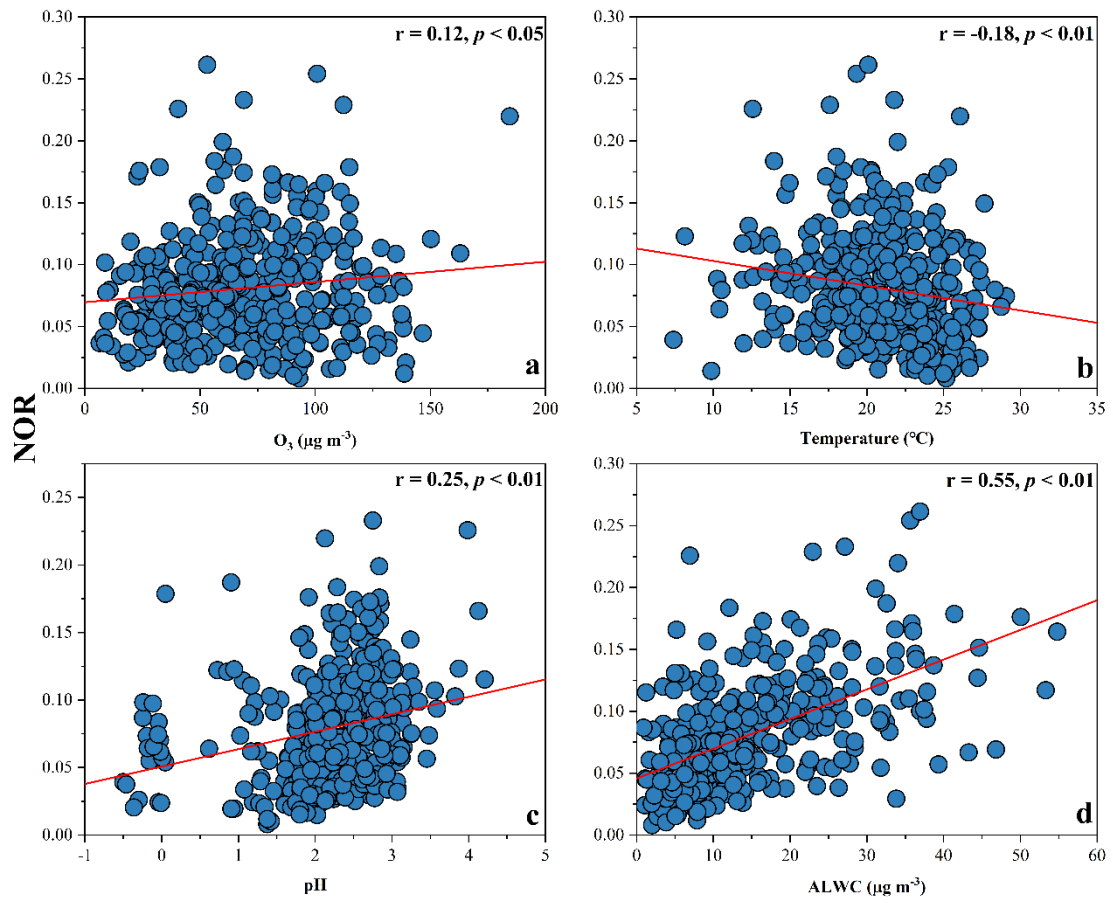


Figure S15. Correlations between nitrogen oxidation rate (NOR) and O_3 (a), temperature (b), pH (c), as well as ALWC (d).

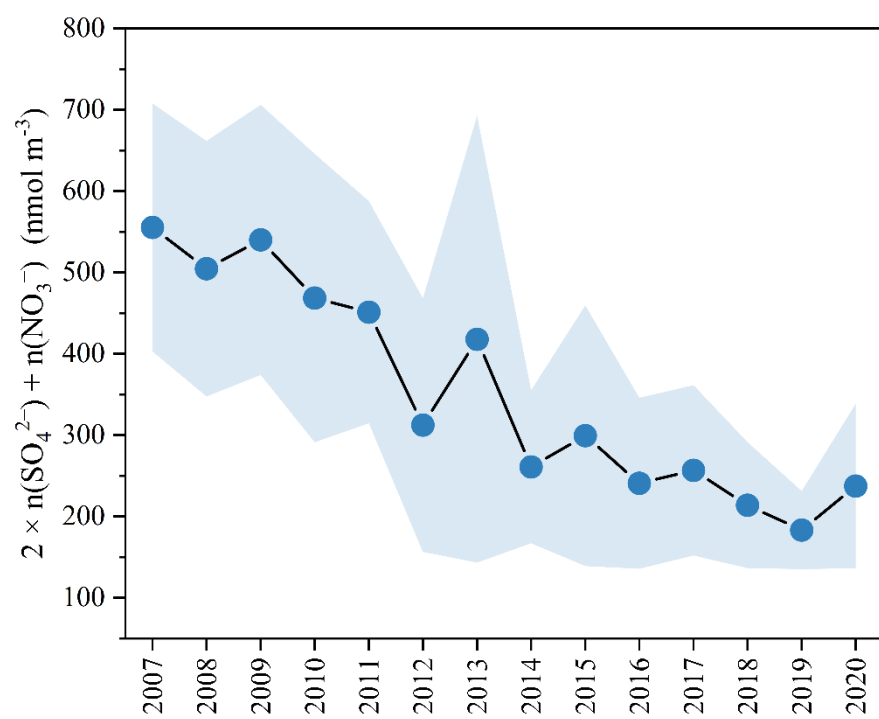


Figure S16. Variations in $2 \times n(\text{SO}_4^{2-}) + n(\text{NO}_3^-)$.

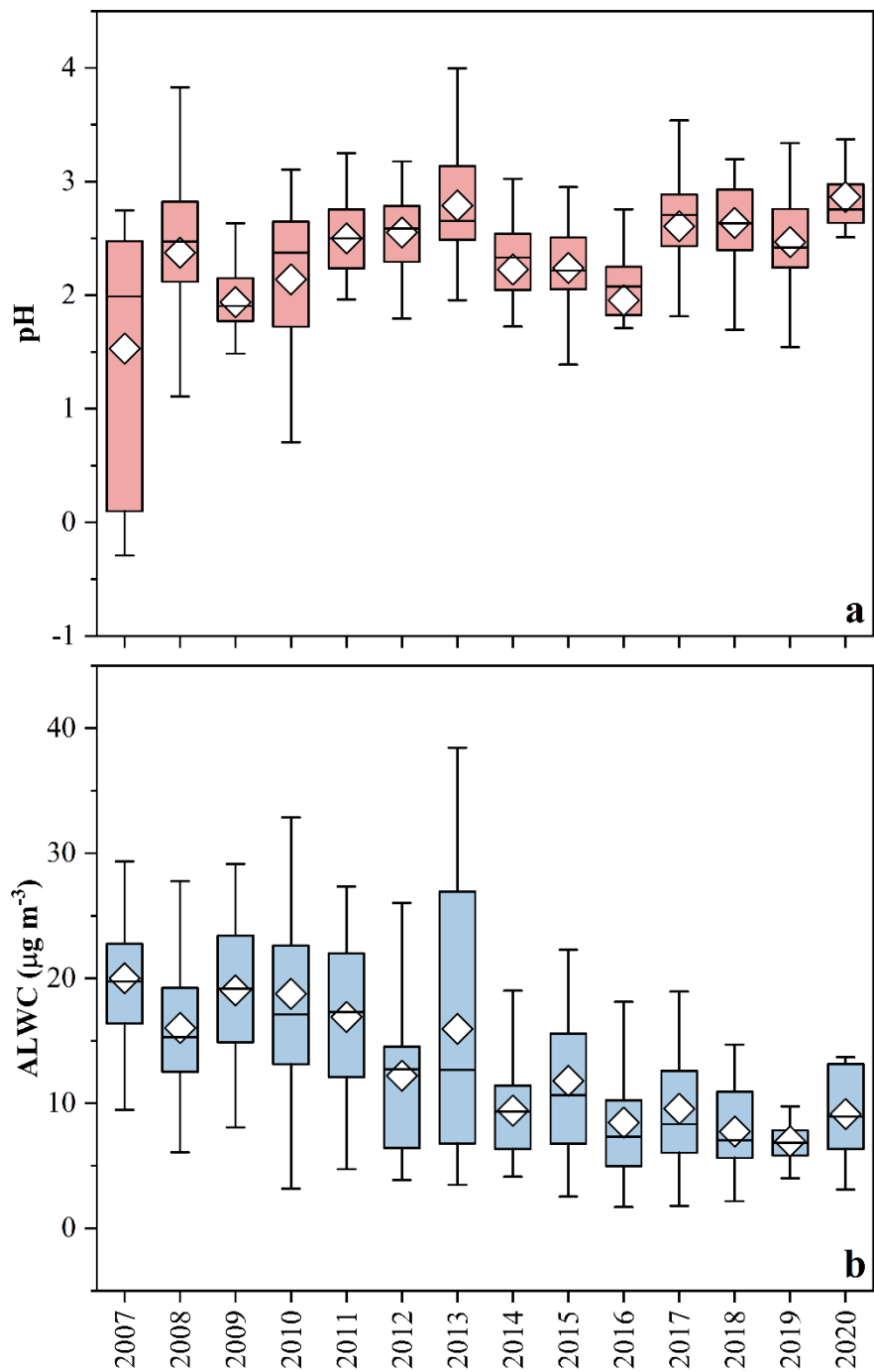
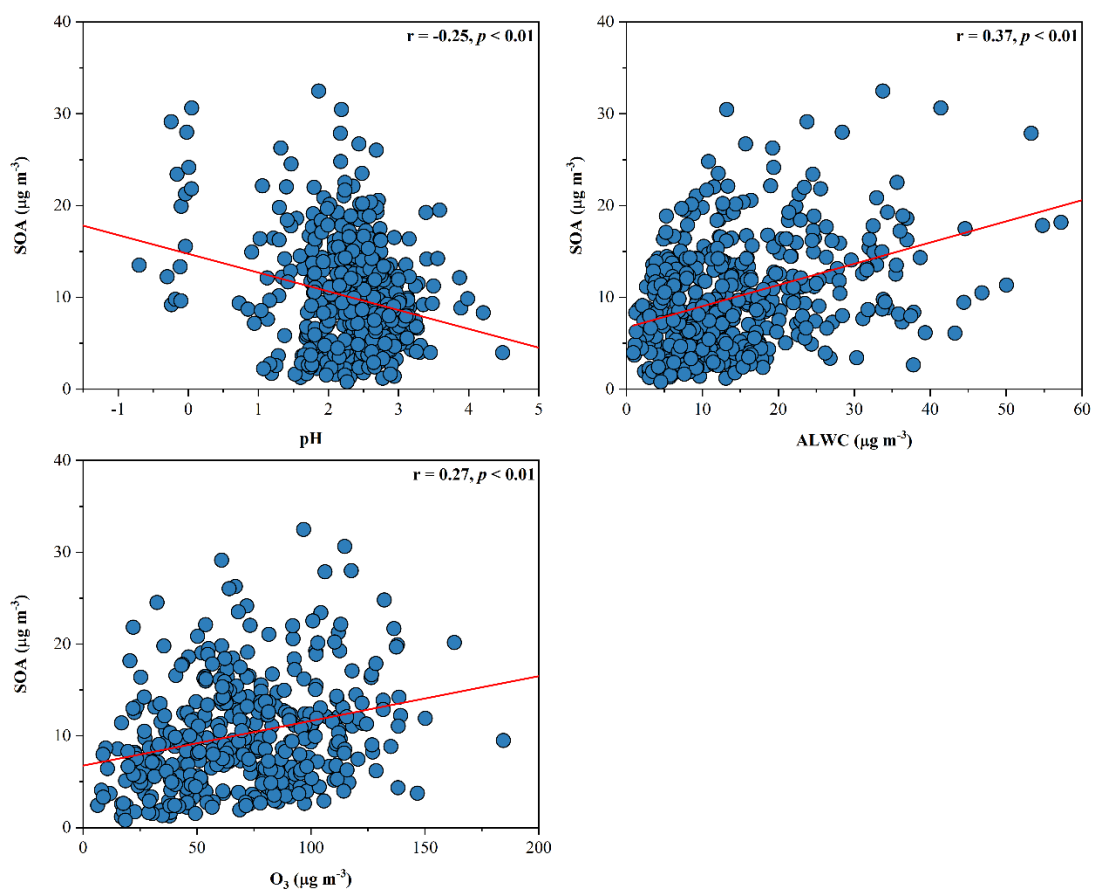


Figure S17. Recalculation of pH (a) and ALWC(b) using average temperature and RH as model input. The upward trend in pH and downward trend in ALWC still exist.



75

Figure S18. Correlations between SOA and pH, ALWC, as well as O_3 .

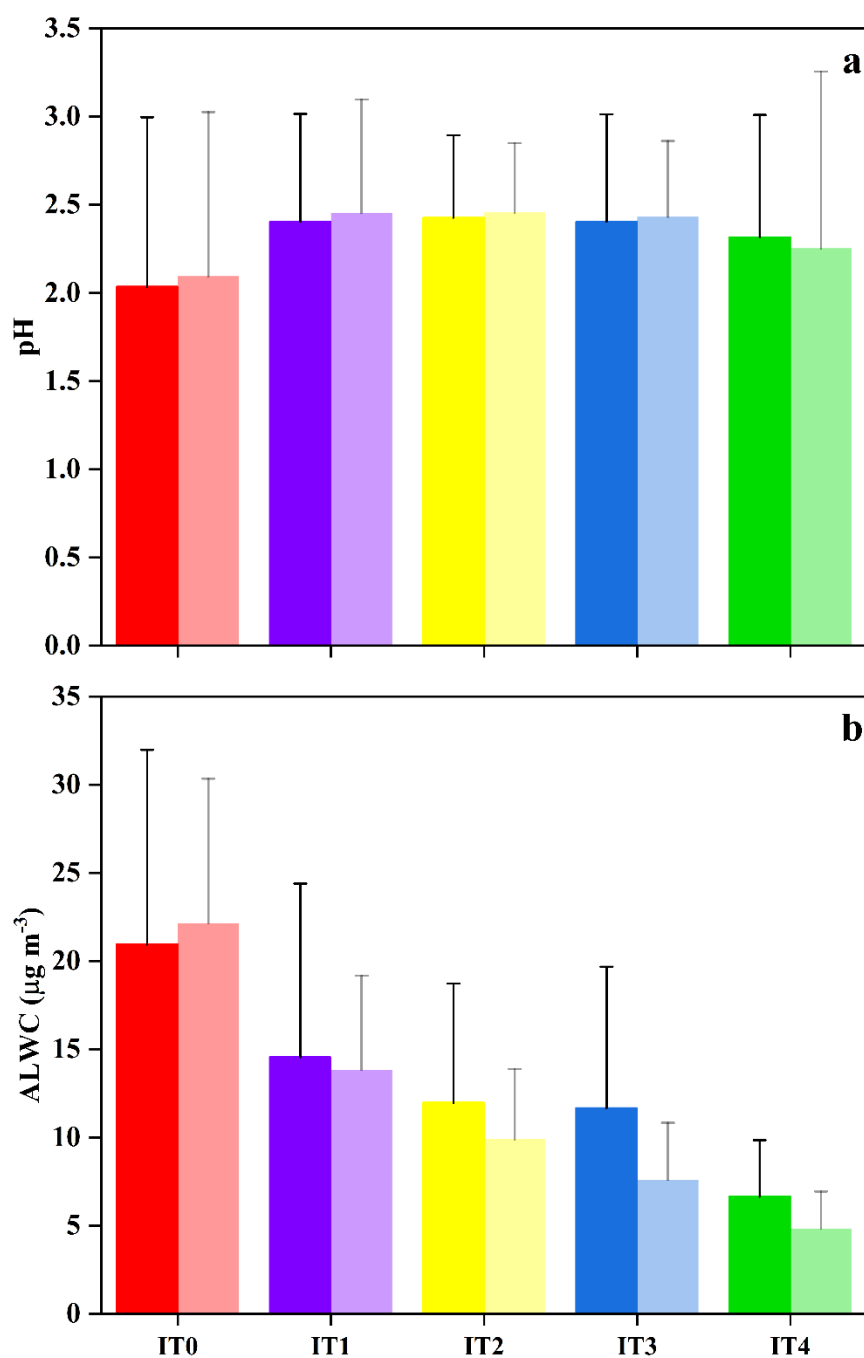


Figure S19. Variations in pH (a) and ALWC (b) under different pollutant levels. The dark color bars represented original model prediction, while the light color bars represented recalculation by average temperature and RH. A low pH occurred under high pollutant level (IT0). With decrease of pollutant levels (IT0-IT4), ALWC exhibited a downward trend.

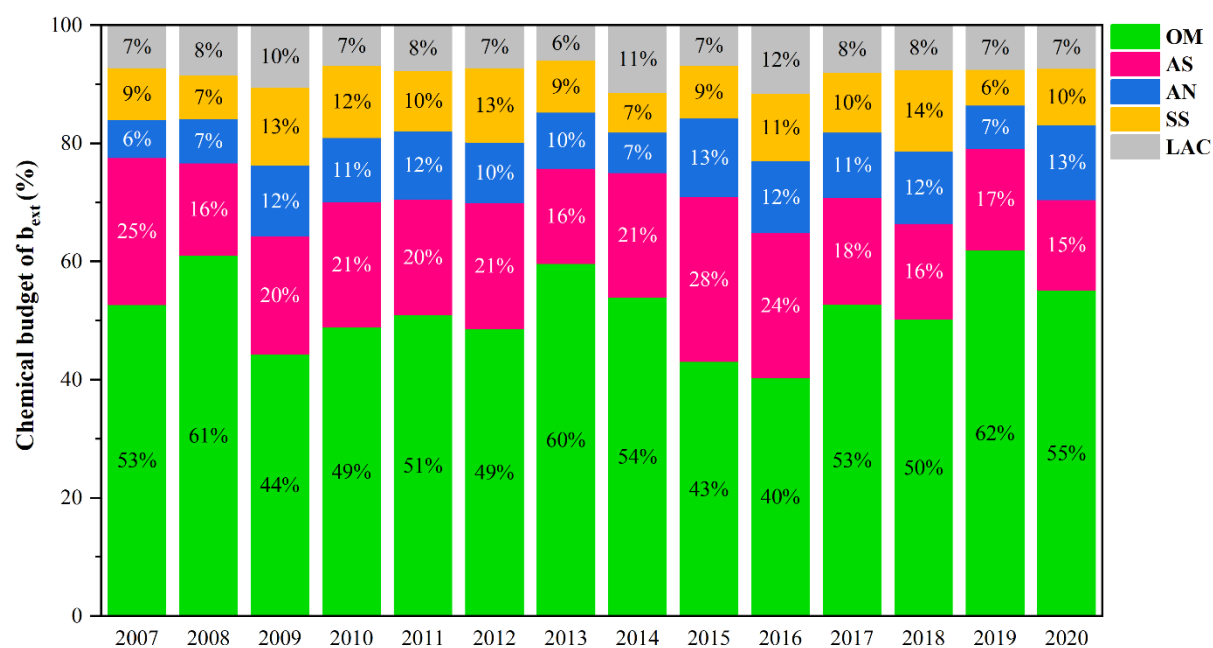


Figure S20. Chemical budget of b_{ext} from different components in $PM_{2.5}$.

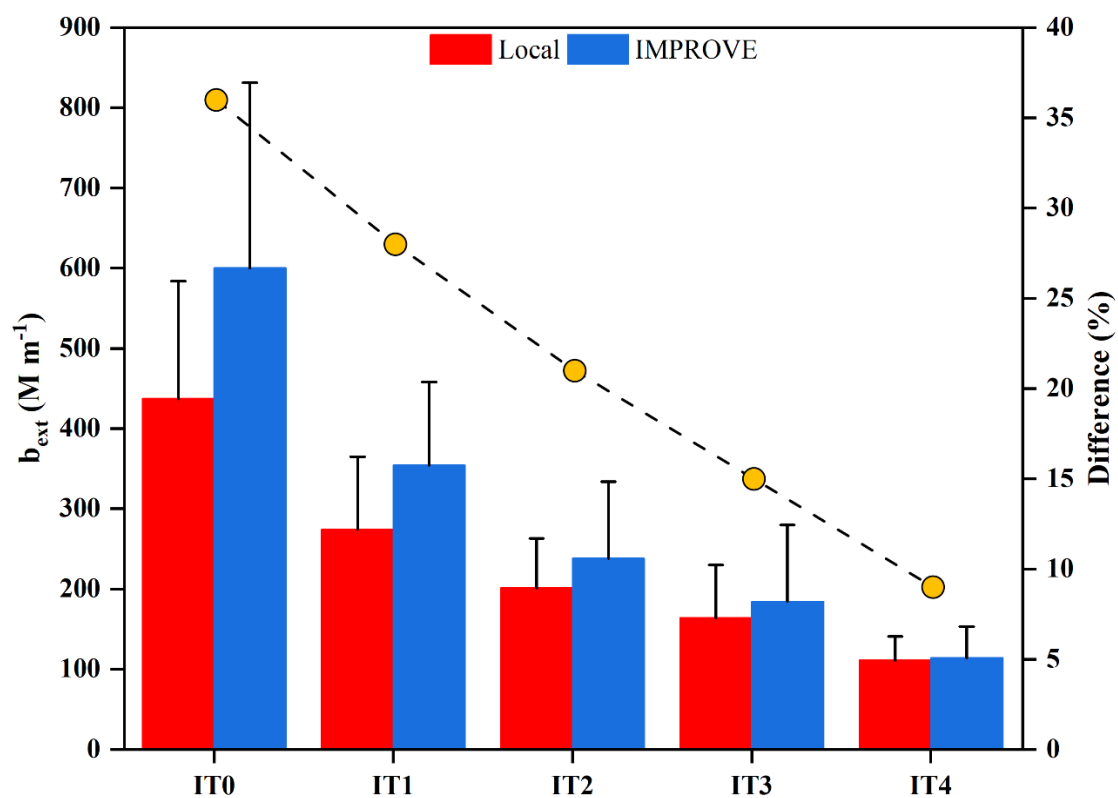


Figure S21. The difference of b_{ext} under different pollutant levels. Bars represent b_{ext} and circles represent differences between Local scheme and the revised IMPROVE scheme. When pollutant levels were high (IT0), the difference was above 30%, while the difference was under 10% when pollutant levels were low (IT4).

- 90 Bougiatioti, A., Nikolaou, P., Stavroulas, I., Kouvarakis, G., Weber, R., Nenes, A., Kanakidou, M., and Mihalopoulos, N.: Particle water and pH in the eastern Mediterranean: source variability and implications for nutrient availability, *Atmos. Chem. Phys.*, 16, 4579-4591, <https://doi.org/10.5194/acp-16-4579-2016>, 2016.
- Fang, Z. Y., Dong, S. W., Huang, C. P., Jia, S. G., Wang, F., Liu, H. M., Meng, H., Luo, L., Chen, Y. Z., Zhang, H. H., Li, R., Zhu, Y. J., and Tang, M. J.: On using an aerosol thermodynamic model to calculate aerosol acidity of coarse particles, *J. Environ. Sci.*, 148, 46-56, <https://doi.org/10.1016/j.jes.2023.07.0011>, 2025.
- 95 Fountoukis, C. and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+ - Ca^{2+} - Mg^{2+} - NH_4^+ - Na^+ - SO_4^{2-} - NO_3^- - Cl^- - H_2O aerosols, *Atmos. Chem. Phys.*, 7, 4639-4659, <https://doi.org/10.5194/acp-7-4639-2007>, 2007.
- Guo, H., Xu, L., Bougiatioti, A., Cerully, K. M., Capps, S. L., Hite Jr, J. R., Carlton, A. G., Lee, S. H., Bergin, M. H., Ng, N. L., Nenes, A., and Weber, R. J.: Fine-particle water and pH in the southeastern United States, *Atmos. Chem. Phys.*, 15, 5211-5228, <https://doi.org/10.5194/acp-15-5211-2015>, 2015.
- 100 Hou, L. L., Dai, Q. L., Song, C. B., Liu, B. W., Guo, F. Z., Dai, T. J., Li, L. X., Liu, B. S., Bi, X. H., Zhang, Y. F., and Feng, Y. C.: Revealing Drivers of Haze Pollution by Explainable Machine Learning, *Environ. Sci. Technol. Lett.*, 9, 112-119, <https://doi.org/10.1021/acs.estlett.1c00865>, 2022.
- 105 Nenes, A., Pandis, S. N., and Pilinis, C.: ISORROPIA: A New Thermodynamic Equilibrium Model for Multiphase Multicomponent Inorganic Aerosols, *Aquat. Geochem.*, 4, 123-152, <https://doi.org/10.1023/A:1009604003981>, 1998.
- Wen, L., Xue, L. K., Wang, X. F., Xu, C. H., Chen, T. S., Yang, L. X., Wang, T., Zhang, Q. Z., and Wang, W. X.: Summertime fine particulate nitrate pollution in the North China Plain: increasing trends, formation mechanisms and implications for control policy, *Atmos. Chem. Phys.*, 18, 11261-11275, <https://doi.org/10.5194/acp-18-11261-2018>, 2018.
- 110 Zhou, M., Zheng, G. J., Wang, H. L., Qiao, L. P., Zhu, S. H., Huang, D. D., An, J. Y., Lou, S. R., Tao, S. K., Wang, Q., Yan, R. S., Ma, Y. G., Chen, C. H., Cheng, Y. F., Su, H., and Huang, C.: Long-term trends and drivers of aerosol pH in eastern China, *Atmos. Chem. Phys.*, 22, 13833-13844, <https://doi.org/10.5194/acp-22-13833-2022>, 2022.