



*Supplement of*

## **Understanding summertime peroxyacetyl nitrate (PAN) formation and its relation to aerosol pollution: insights from high-resolution measurements and modeling**

**Baoye Hu et al.**

*Correspondence to:* Jinsheng Chen ([jschen@iue.ac.cn](mailto:jschen@iue.ac.cn)) and Yuxiang Yang ([907460293@qq.com](mailto:907460293@qq.com))

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

The index of agreement (IOA) is calculated by the eq. S1 (Ghahremanloo et al., 2021):

$$IOA = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (S1)$$

Where,  $O_i$  denotes the observed values,  $\bar{O}$  is the average observed values,  $S_i$  represents the simulated value, and  $n$  is the number of samples.

To compare the PAN production rates ( $P(PAN)$ ) from observations and simulations, which were determined using Eq. S2 (Xu et al., 2021):

$$P(PAN) = \frac{PAN_2 - PAN_1}{t_2 - t_1} \quad (S2)$$

where  $t_1$  and  $t_2$  represent the start and end times, respectively, of the local photochemical PAN production identified for each day based on simulation results, and  $PAN_1$  and  $PAN_2$  are the corresponding PAN concentrations.

The net production of PAN ( $Net(PAN)$ ) involved the production pathway of  $PA+NO_2$ , and the loss of PAN was thermal decomposition and  $PAN+OH$  during the daytime (5:00-18:00 local time) (Liu et al., 2022; Zeng et al., 2019). The net production of PAN was calculated from eq. S3:

$$Net(PAN) = k_{PA+NO_2} [PA][NO_2] - k_{PAN} [PAN] - k_{PAN+OH} [PAN][OH] \quad (S3)$$

The relative incremental reactivity (RIR) was calculated based on modeling results to reflect the sensitivity of PAN formation toward its precursor levels. If the RIR value was positive, it meant that the increase of precursors enhances PAN formation, whereas negative RIR value indicated that the increase of precursors inhibited PAN production. Besides, the greater the absolute value of RIR, the more sensitive PAN formation is to this precursor. The RIR value was calculated from eq. S4:

$$RIR_x = \frac{Net(PAN_x) - Net(PAN_{x-\Delta x})}{Net(PAN_x)} \cdot \frac{\Delta x}{x} \quad (S4)$$

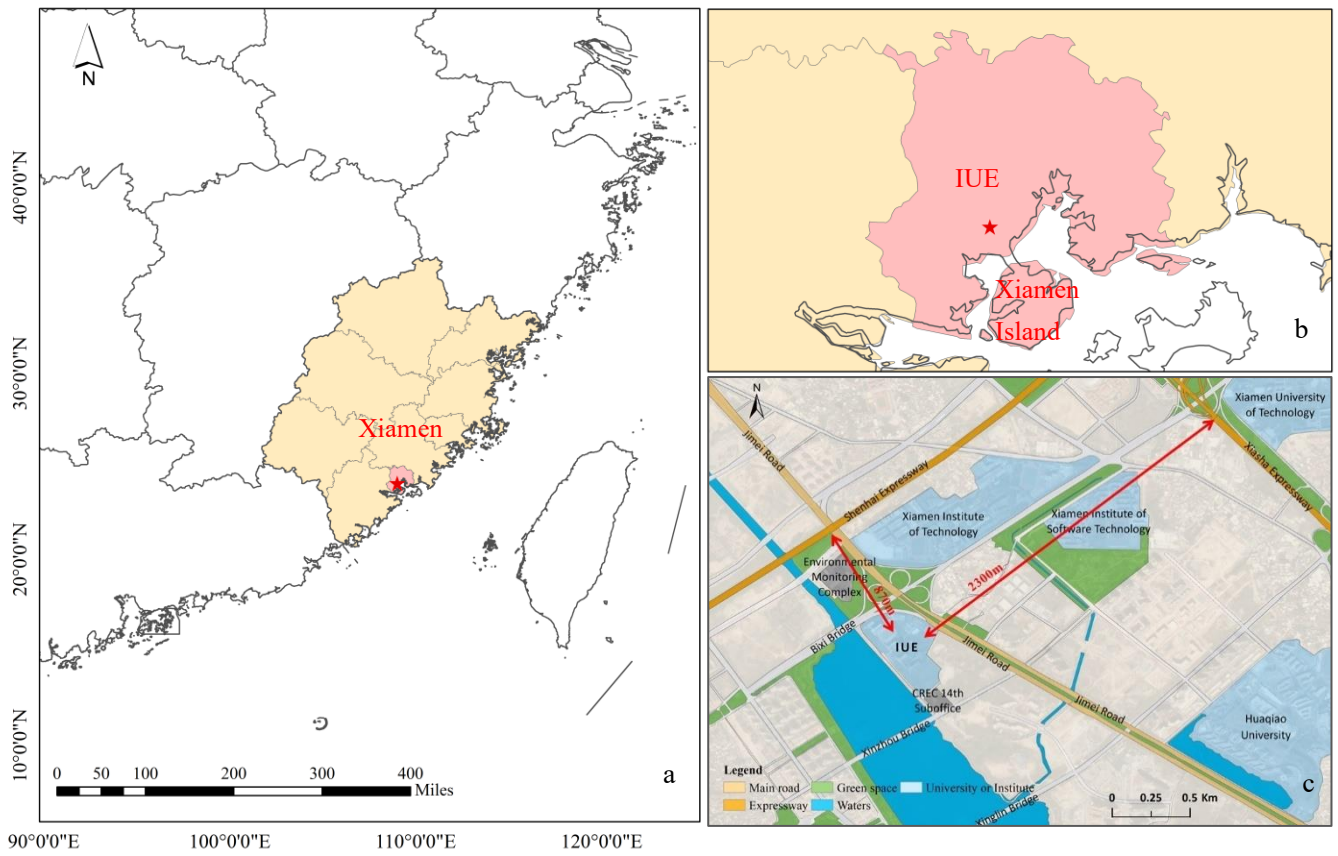
Where  $x$  represents a certain PAN precursor (e.g.,  $NO_x$ ,  $C_5H_8$ ,  $O_3$ , and HONO).  $\Delta x/x$  represents the hypothetical change of mixing ratio of  $x$  (20% in this study). During simulations of  $O_3$  and HONO, the model was not constrained by the OH modelling considering that  $O_3$  and HONO contribute to PAN production through formation of OH (Xue et al., 2014) (Figure S2).

the root-mean-squared error (RMSE) and mean absolute error (MAE) are calculated using eq. S5 and eq. S6, respectively (Hodson, 2022):

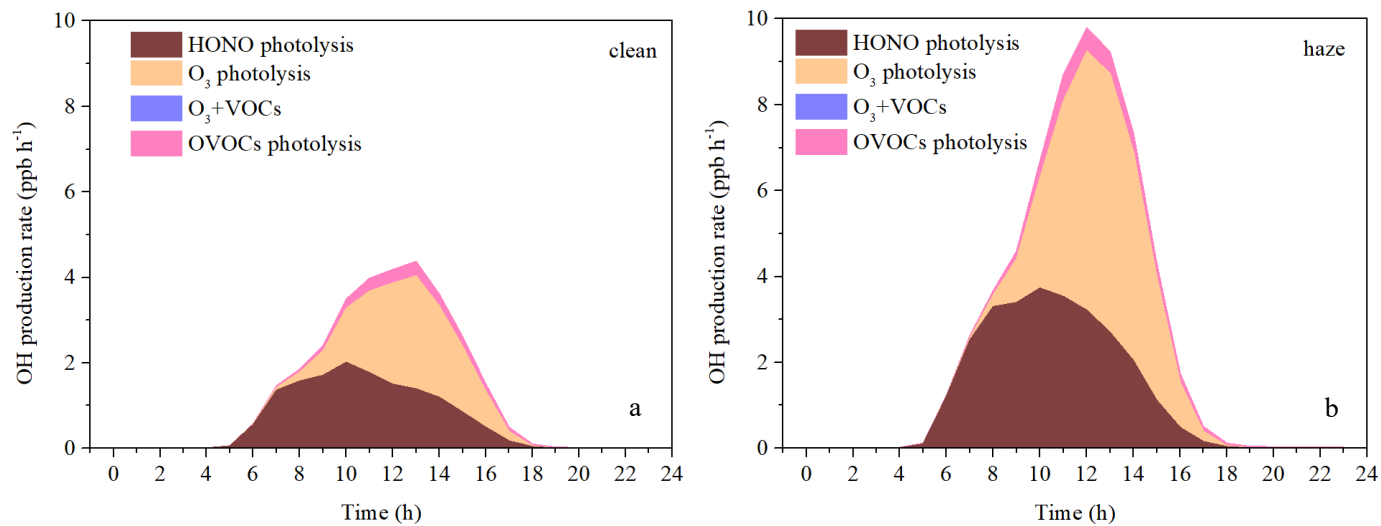
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (S5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (S6)$$

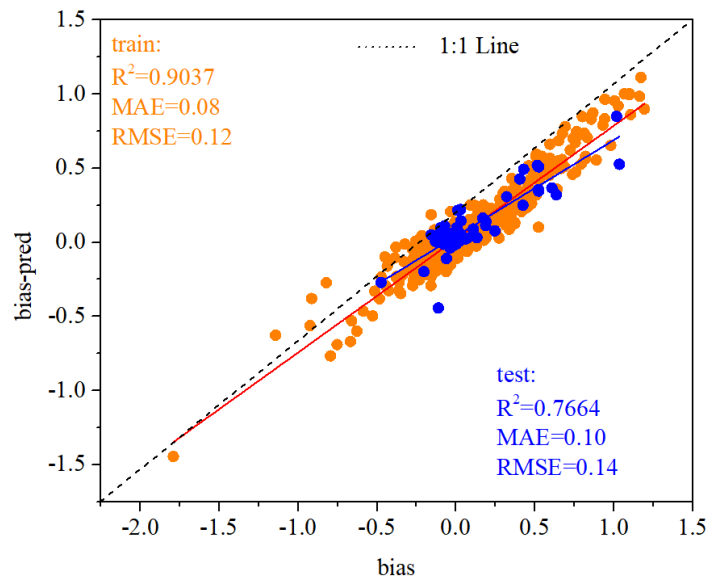
Where  $n$  is the number of observations,  $y_i$  is the observed value, and  $\hat{y}_i$  is the model's predicted value.



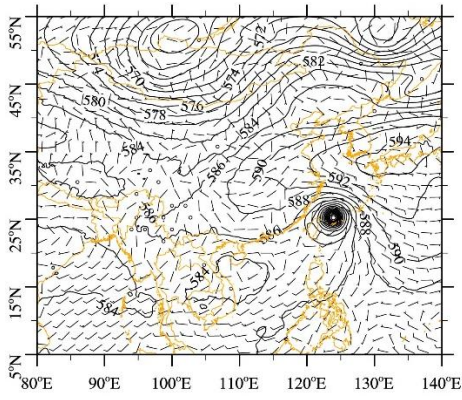
**Figure S1.** Location of Xiamen (a), position of IUE in Xiamen (b) and surrounding of IUE (c).



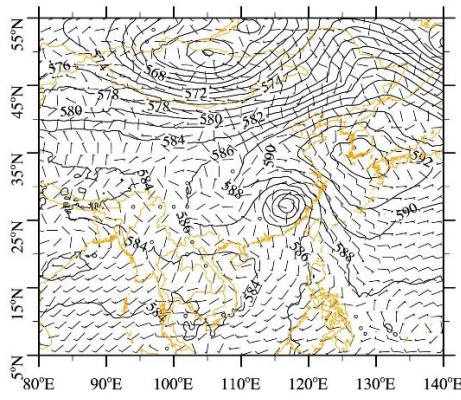
**Figure S2.** Model-simulated average primary production rates of OH during clean (a) and haze days (b).



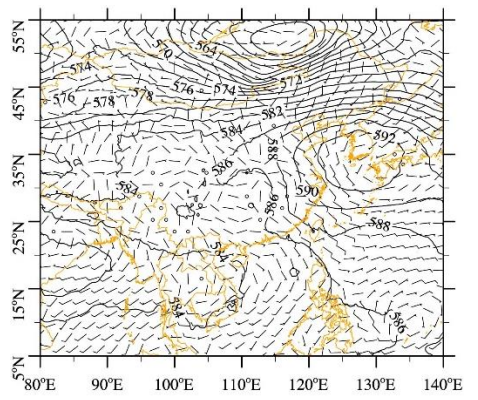
**Figure S3.** Model performance of XGBoost model. Orange dots represent the train set, blue dots represent the test set, and dotted black lines represent 1:1.



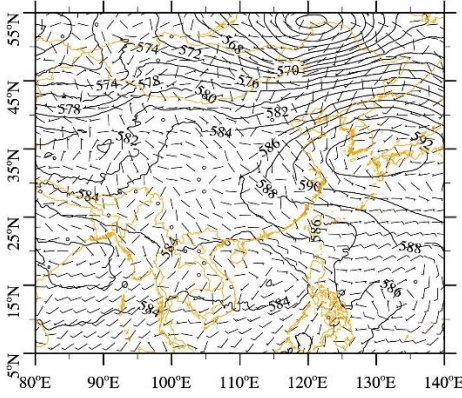
2018-7-10 12:00 UTC



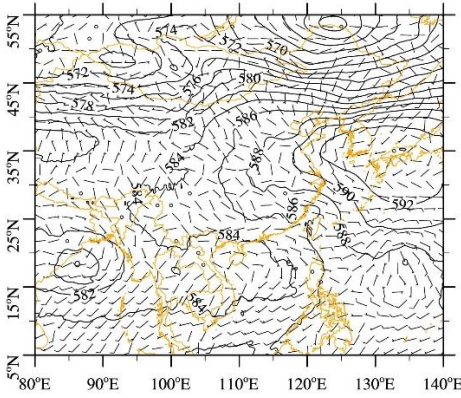
2018-7-11 12:00 UTC



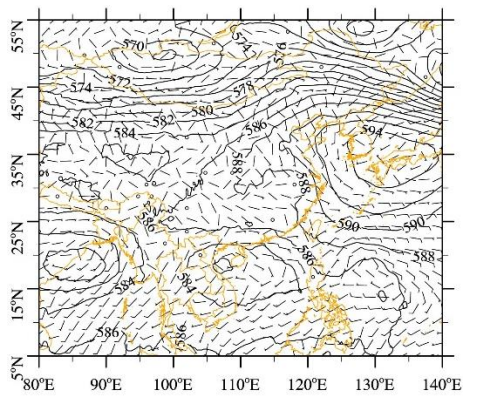
2018-7-12 12:00 UTC



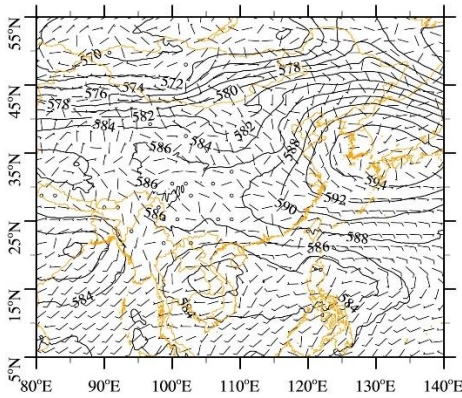
2018-7-13 12:00 UTC



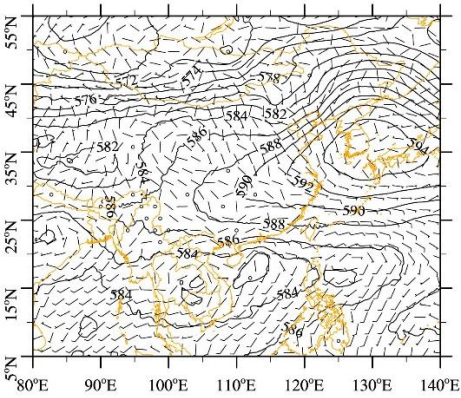
2018-7-14 12:00 UTC



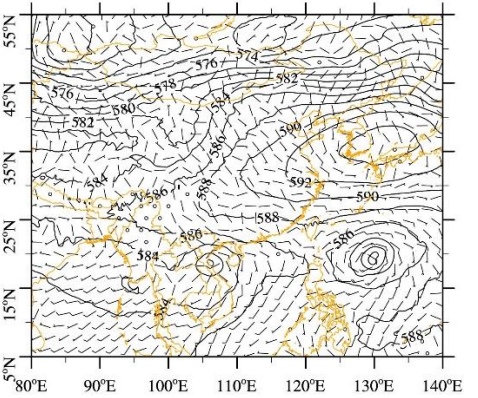
2018-7-15 12:00 UTC



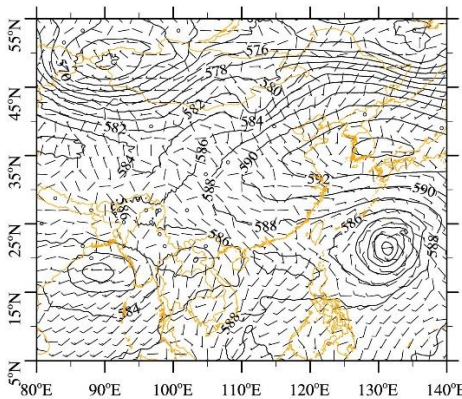
2018-7-16 12:00 UTC



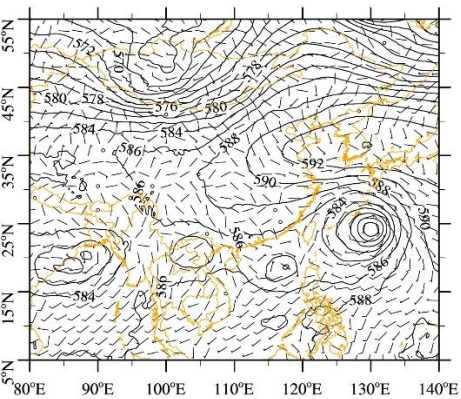
2018-7-17 12:00 UTC



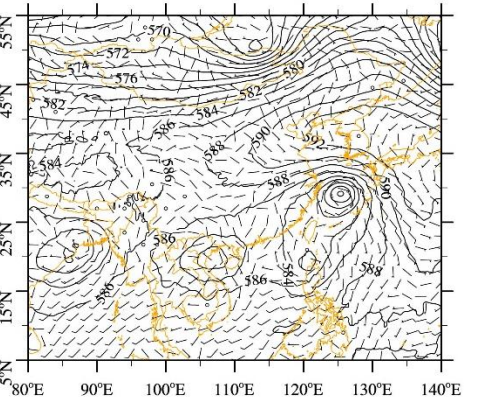
2018-7-18 12:00 UTC



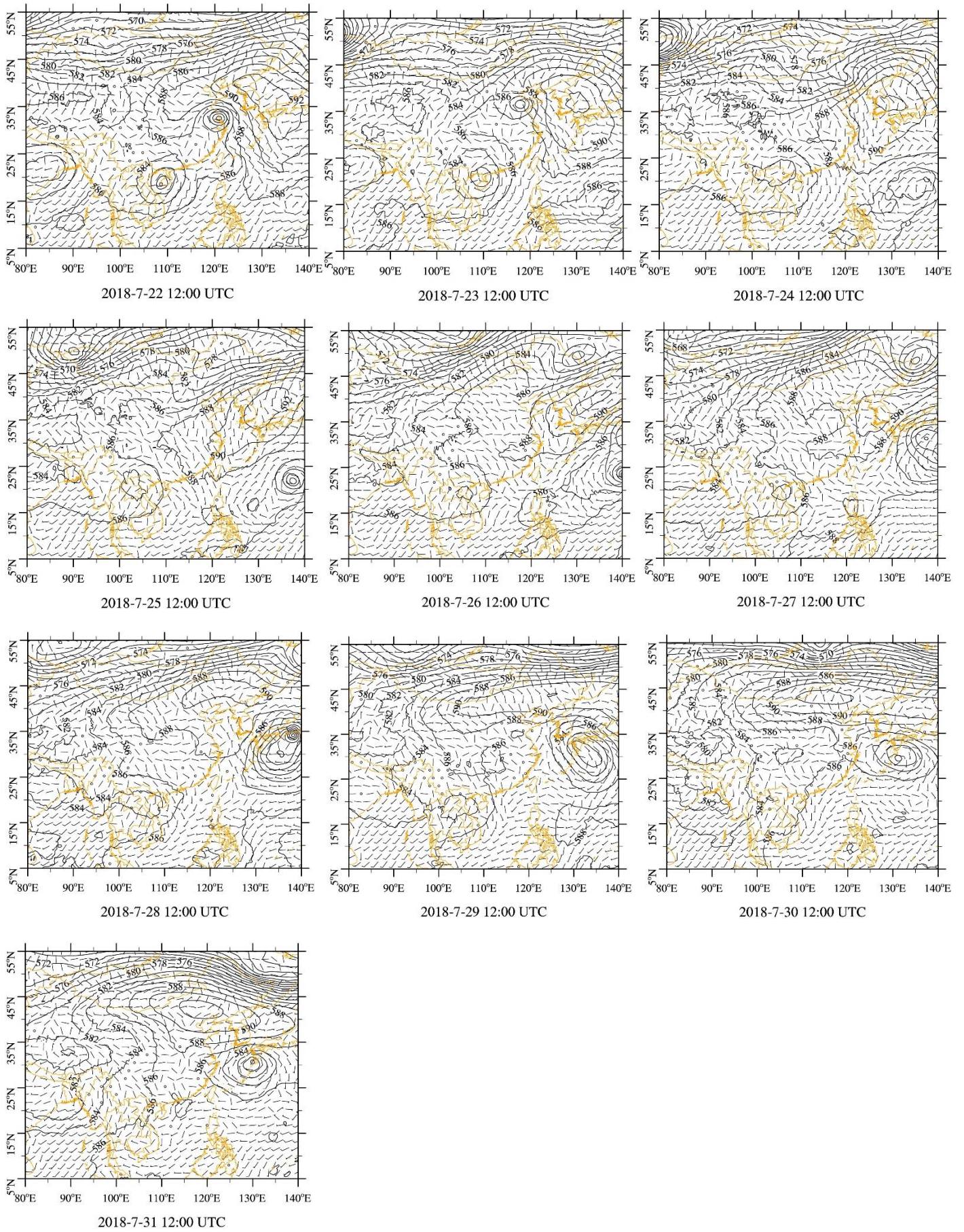
2018-7-19 12:00 UTC



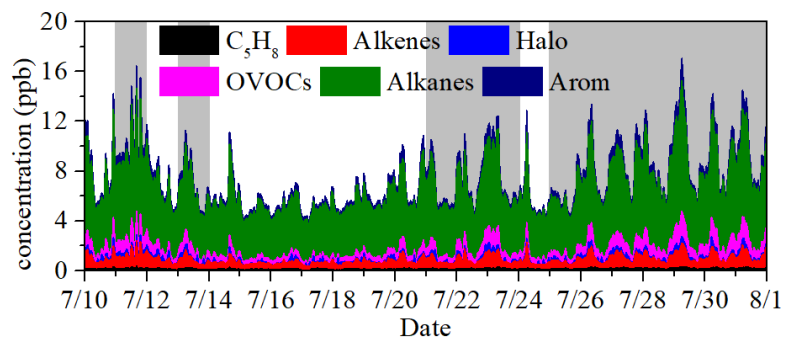
2018-7-20 12:00 UTC



2018-7-21 12:00 UTC

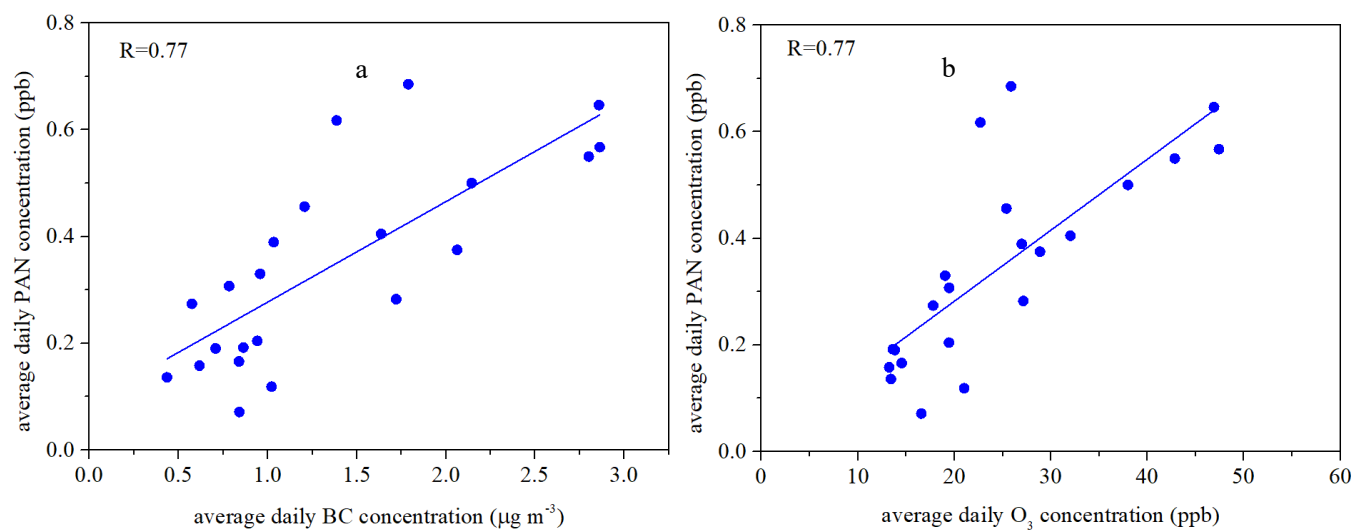


**Figure S4.** Synoptic situation at 500 hPa from 10 to 31 July.

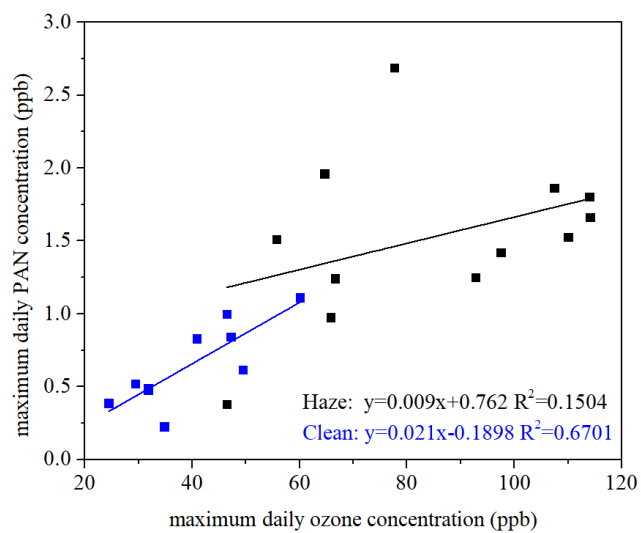


**Figure S5.** Time series of VOCs observed at IUE during 10-31 July 2018. The gray shading represents days when the PM<sub>2.5</sub> hourly daily maximum value exceeded 35 µg m<sup>-3</sup>.

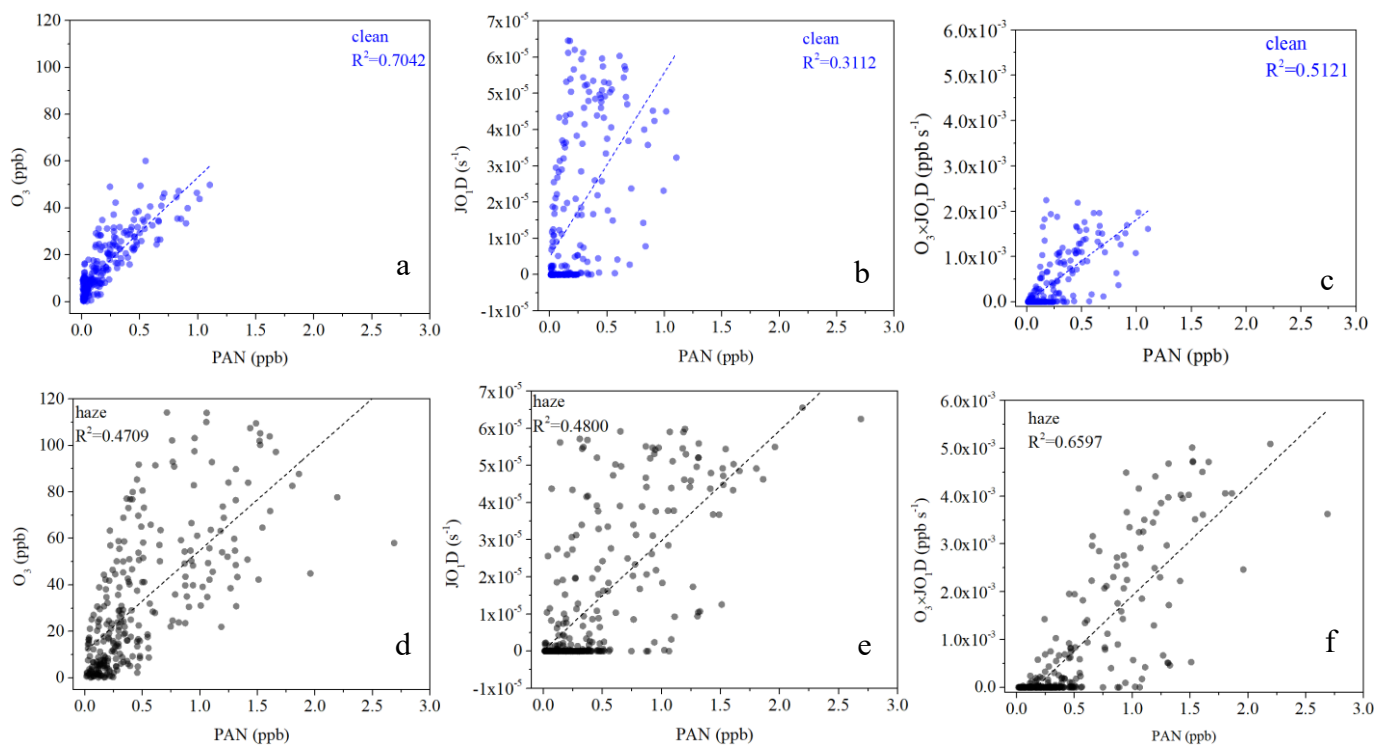




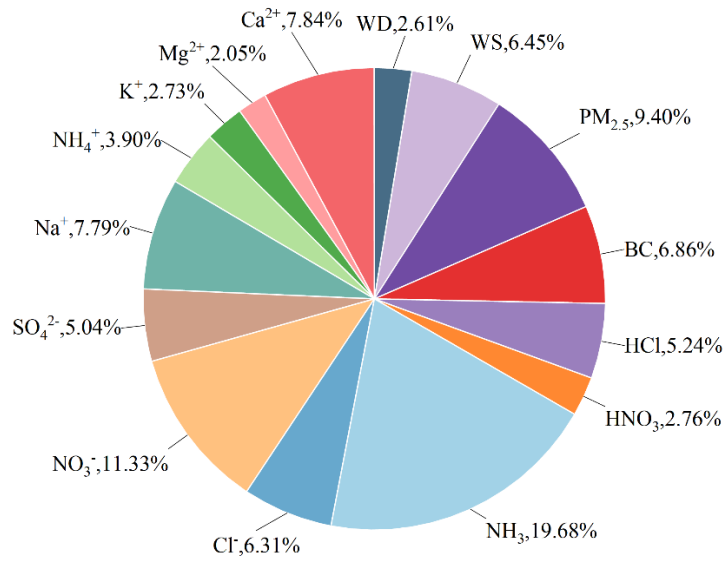
**Figure S6.** The correlation between the average daily values of PAN and BC (a), as well as the correlation between the maximum daily values of PAN and  $\text{O}_3$  (b).



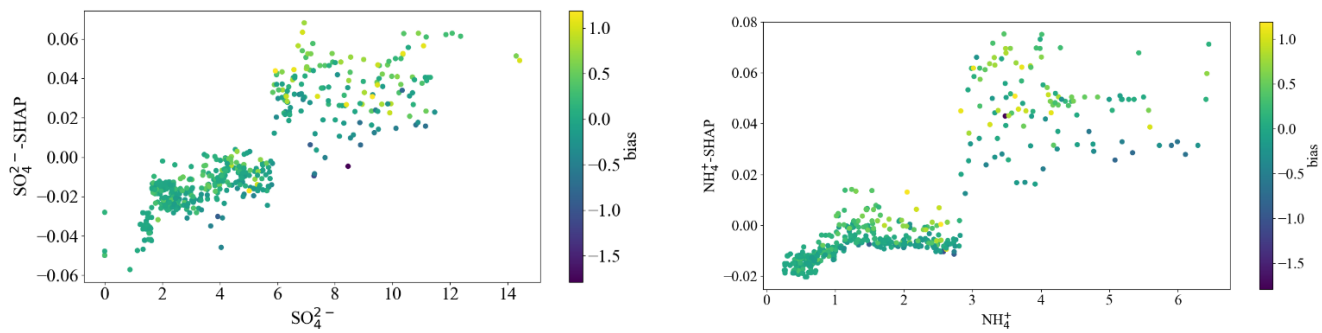
**Figure S7.** Correlation between PAN and O<sub>3</sub> daily maximum concentrations during haze and clean.



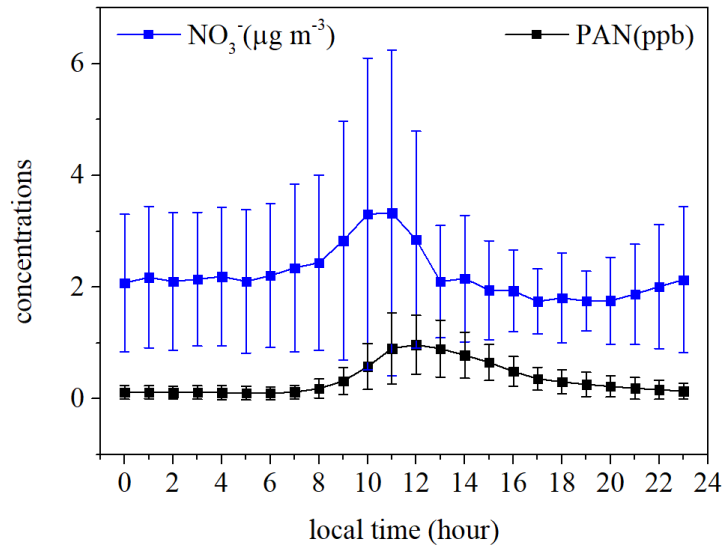
**Figure S8.** a, b, and c are scatter plots of PAN with O<sub>3</sub>, JO<sub>1</sub>D, and O<sub>3</sub> × JO<sub>1</sub>D during the cleaning period, while d, e, and f are scatter plots of PAN with O<sub>3</sub>, JO<sub>1</sub>D, and O<sub>3</sub> × JO<sub>1</sub>D during the haze period. The darker the color, the denser the data points.



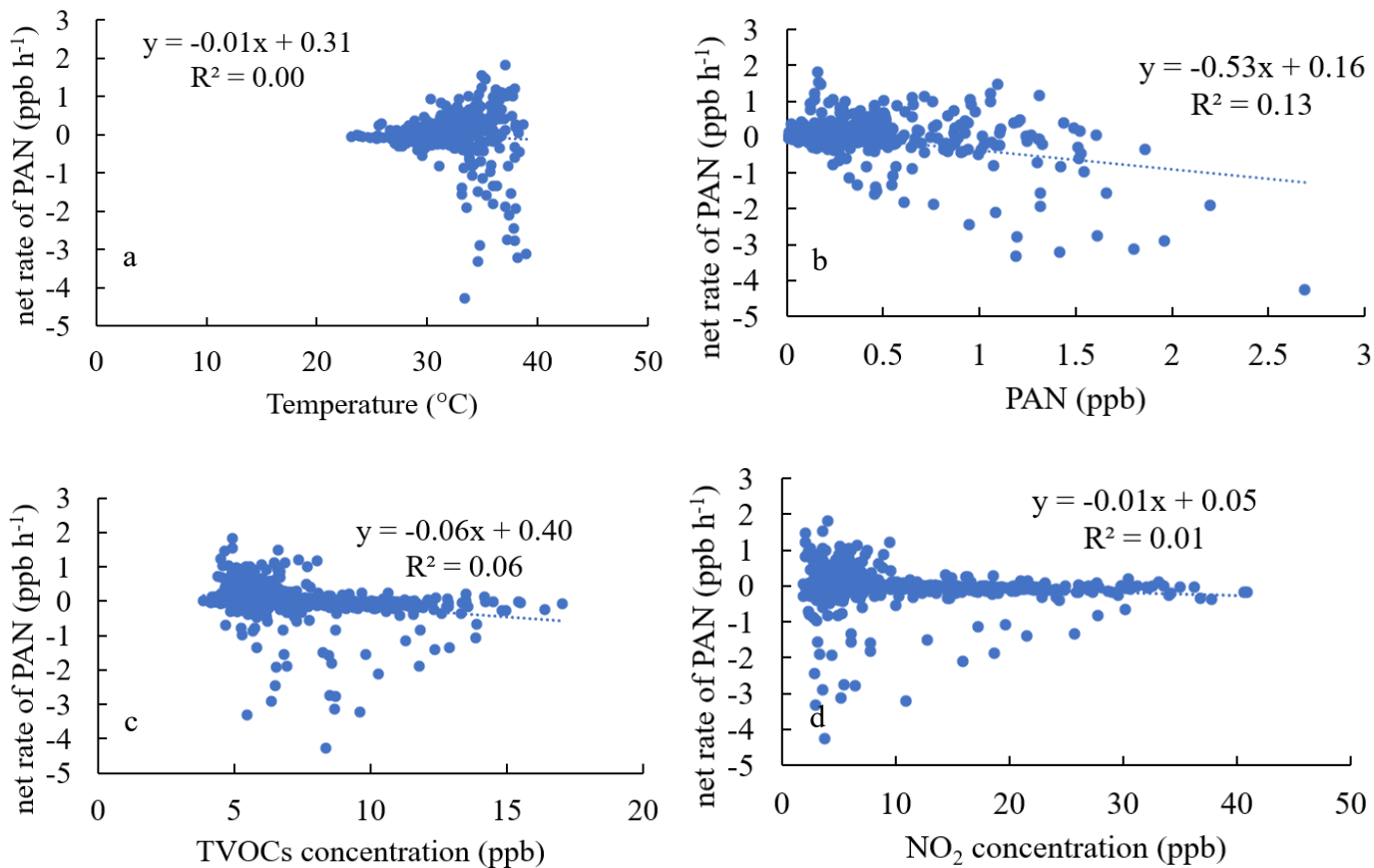
**Figure S9.** The average proportion of the absolute SHAP value for each feature during the whole observation period.



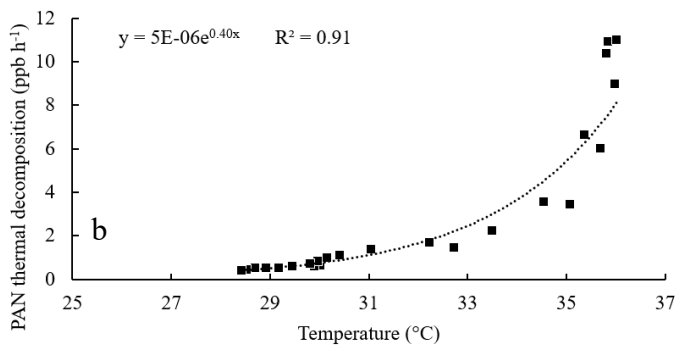
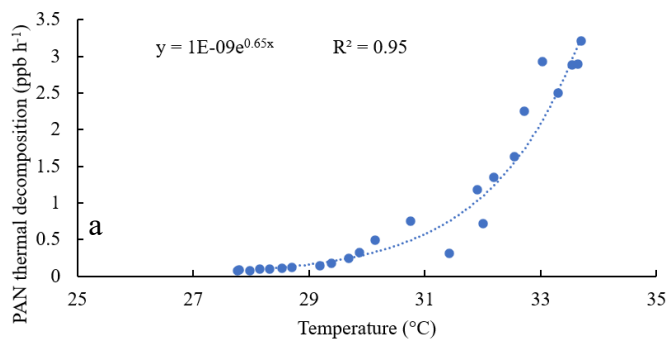
**Figure S10.** The scatter plot of  $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$  concentrations versus their SHAP values., and colored with the bias (the model simulation minus the observed value).



**Figure S11.** Diurnal variation of PAN and  $\text{NO}_3^-$ .

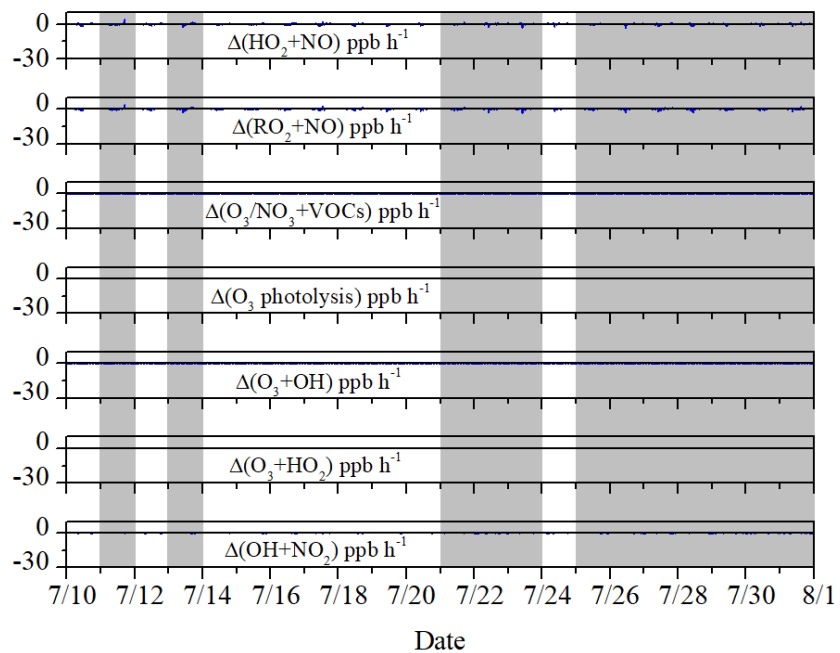


**Figure S12.** Correlation analysis of the net production rate of PAN with temperature (a), PAN (b), VOCs (c), and NO<sub>2</sub> (d) concentration, respectively.



**Figure S13.** Correlation between temperature and PAN thermal decomposition during clean (a) and haze (b) period.





**Figure S14.** Time series plot of the reaction of  $\Delta(\text{HO}_2+\text{NO})$ ,  $\Delta(\text{RO}_2+\text{NO})$ ,  $\Delta(\text{O}_3/\text{NO}_3+\text{VOCs})$ ,  $\Delta(\text{O}_3 \text{ photolysis})$ ,  $\Delta(\text{O}_3+\text{OH})$ ,  $\Delta(\text{O}_3+\text{HO}_2)$ , and  $\Delta(\text{OH}+\text{NO}_2)$ .

**Table S1.** Measured VOC concentrations during 10-31 July 2018 in Xiamen (units: ppt).

<b>Chemicals</b>	<b>Mean ± SD</b>	<b>Chemicals</b>	<b>Mean ± SD</b>
<b>Aromatics</b>	<b>549±295</b>	<b>Alkanes</b>	<b>5001±1378</b>
ethylbenzene	19±15	ethane	1315±180
o-xylene	21±16	propane	1059±490
m/p-xylene	51±39	isobutane	415±103
isopropylbenzene	4±0	n-butane	599±142
n-propylbenzene	6±1	isopentane	706±198
m-ethyltoluene	12±1	n-pentane	83±74
p-ethyltoluene	8±1	2,2-dimethylbutane	4±5
o-ethyltoluene	7±1	2,3-dimethylbutane	11±19
1,3,5-trimethylbenzene	6±1	2-methylpentane	12±16
1,2,4-trimethylbenzene	62±7	3-methylpentane	29±27
1,2,3-trimethylbenzene	6±1	n-hexane	213±110
benzene	120±59	2-methylhexane	62±12
toluene	183±168	cyclohexane	39±7
styrene	44±10	3-methylhexane	96±19
<b>Halocarbons</b>	<b>166±172</b>	n-heptane	64±14
1,3-dichloropropene	33±33	n-octane	23±4
trichloroethylene	2±6	n-nonane	13±2
trichloroethane	67±88	n-decane	13±2
tetrachloroethylene	4±6	n-undecane	25±5
tetrachloroethane	1±4	<b>Alkenes</b>	<b>747±337</b>
chloroethane	59±129	1-hexene	118±48
<b>OVOCs</b>	<b>699±356</b>	ethene	161±117
acetone	369±166	propene	135±34
butanone	266±158	1,3-butadiene	9±17
4-methyl-2-pentanone	4±2	1-pentene	1±1
methyl tert-butyl ether	60±38	trans-2-pentene	57±12
<b>isoprene</b>	<b>153±53</b>	butene	8±17

**Table S2.** The independent samples T-test between haze and clean period.

	Haze (mean±stdev)	Clean (mean±stdev)
$\Delta\text{HO}_2$ (ppb)	$8.64 \cdot 10^{-5} \pm 8.49 \cdot 10^{-4}$	$8.18 \cdot 10^{-5} \pm 5.76 \cdot 10^{-4}$
$\Delta\text{OH}$ (ppb)	$4.23 \cdot 10^{-7} \pm 1.37 \cdot 10^{-5}$	$4.94 \cdot 10^{-7} \pm 1.49 \cdot 10^{-5}$
$\Delta\text{RO}_2$ (ppb)	$-6.55 \cdot 10^{-4} \pm 2.28 \cdot 10^{-3}$	$-6.11 \cdot 10^{-4} \pm 1.43 \cdot 10^{-3}$
$\Delta\text{NO}_2$ (ppb)	$-0.22 \pm 0.48^{**}$	$-0.11 \pm 0.27$
$\Delta\text{NO}$ (ppb)	$-0.05 \pm 0.17^{**}$	$0.03 \pm 0.09$

Note: \*\* The significance level is 0.01 between haze and clean period.

## Reference

- Ghahremanloo, M., Lops, Y., Choi, Y., and Yeganeh, B.: Deep Learning Estimation of Daily Ground-Level NO<sub>2</sub> Concentrations From Remote Sensing Data, *J. Geophys. Res.: Atmos.*, 126, 10.1029/2021jd034925, 2021.
- Hodson, T. O.: Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not, *Geosci. Model Dev.*, 15, 5481-5487, 10.5194/gmd-15-5481-2022, 2022.
- Liu, T., Chen, G., Chen, J., Xu, L., Li, M., Hong, Y., Chen, Y., Ji, X., Yang, C., Chen, Y., Huang, W., Huang, Q., and Wang, H.: Seasonal characteristics of atmospheric peroxyacetyl nitrate (PAN) in a coastal city of Southeast China: Explanatory factors and photochemical effects, *Atmos. Chem. Phys.*, 22, 4339-4353, 10.5194/acp-22-4339-2022, 2022.
- Xu, W., Zhang, G., Wang, Y., Tong, S., Zhang, W., Ma, Z., Lin, W., Kuang, Y., Yin, L., and Xu, X.: Aerosol Promotes Peroxyacetyl Nitrate Formation During Winter in the North China Plain, *Environ. Sci. Technol.*, 55, 3568-3581, 10.1021/acs.est.0c08157, 2021.
- Xue, L., Wang, T., Wang, X., Blake, D. R., Gao, J., Nie, W., Gao, R., Gao, X., Xu, Z., Ding, A., Huang, Y., Lee, S., Chen, Y., Wang, S., Chai, F., Zhang, Q., and Wang, W.: On the use of an explicit chemical mechanism to dissect peroxy acetyl nitrate formation, *Environ. Pollut.*, 195, 39-47, 10.1016/j.envpol.2014.08.005, 2014.
- Zeng, L., Fan, G. J., Lyu, X., Guo, H., Wang, J. L., and Yao, D.: Atmospheric fate of peroxyacetyl nitrate in suburban Hong Kong and its impact on local ozone pollution, *Environ. Pollut.*, 252, 1910-1919, 10.1016/j.envpol.2019.06.004, 2019.