Supplementary Information for

Deciphering the impacts of meteorology on surface ozone variability in eastern China using explainable machine learning models

Xingpei Ye[1,2], Lin Zhang[1,3,4*], Xiaolin Wang[1], Ni Lu[1], Sebastian Hickman[2], Guo Luo[1], and Alex T. Archibald[2,5]

[1]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China

[2]Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

[3]Institute of Carbon Neutrality, Peking University, Beijing, China

[4]Center for Environment and Health, Peking University, Beijing, China

[5]National Centre for Atmospheric Science, University of Cambridge, Cambridge, United Kingdom

**Corresponding to*: Lin Zhang (zhanglg@pku.edu.cn)

**Corresponding to*: Lin Zhang (zhanglg@pku.edu.cn)

This file includes: 11 figures, 3 tables

Supplementary Figures S1-S11

Supplementary Tables S1-S3

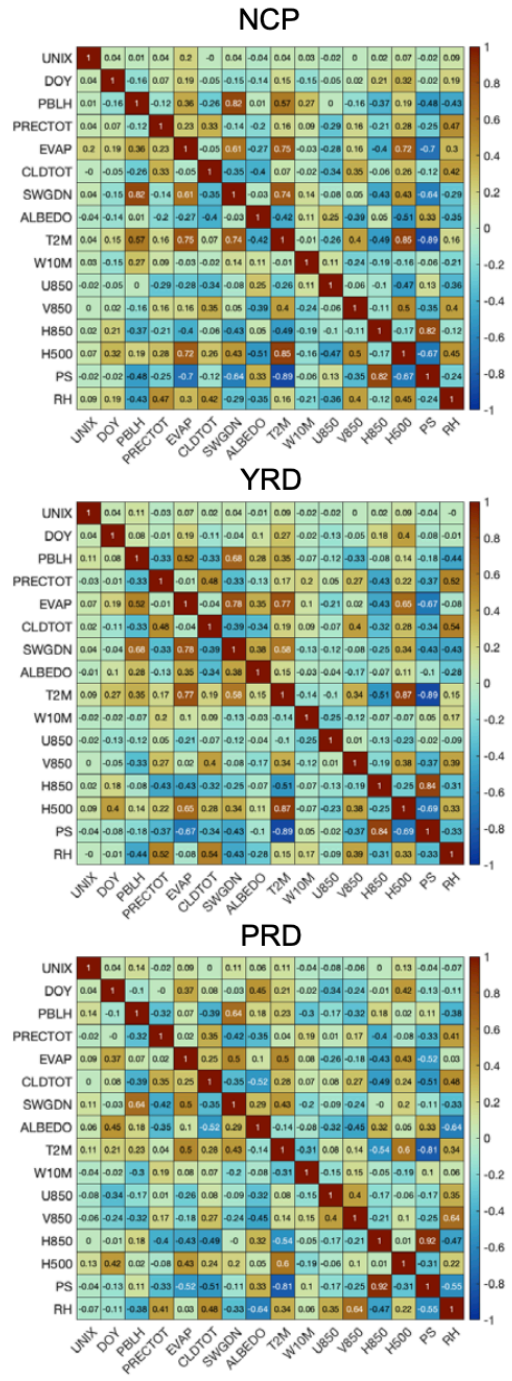


Figure S1. Pearson correlation heat map of all 16 input variables used in the machine learning models, revealing multicollinearity patterns.

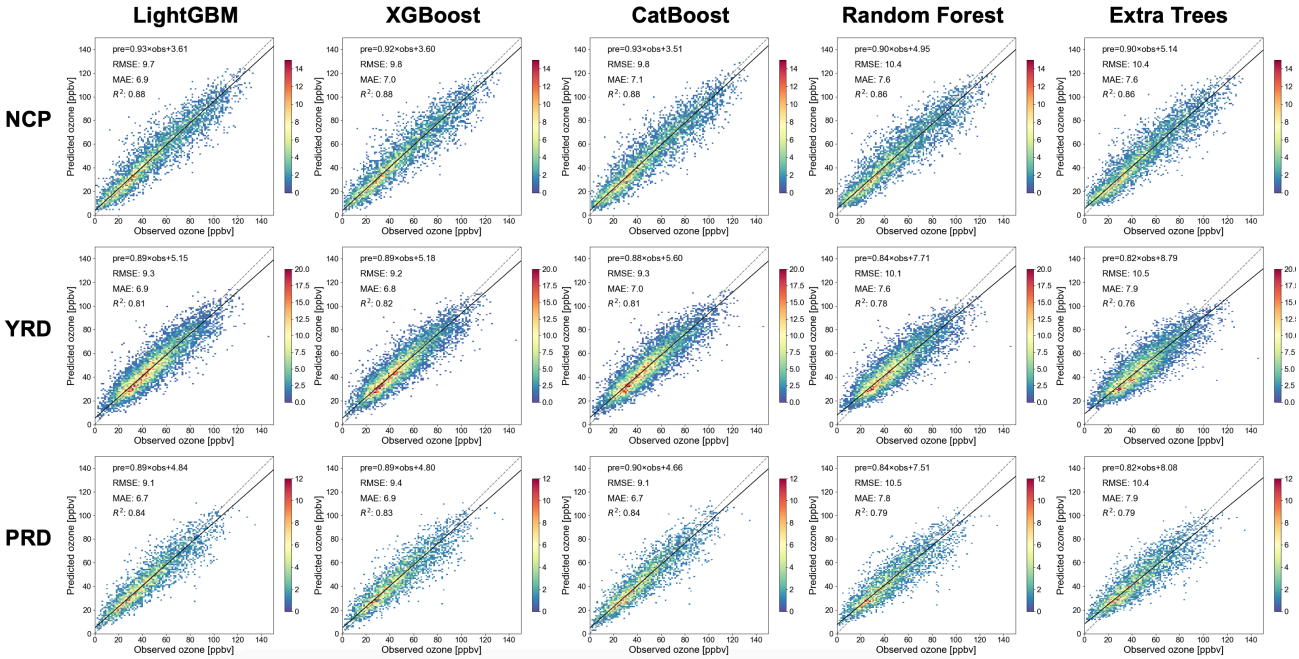


Figure S2. Density plots of observed and predicted MDA8 ozone concentrations across three regions for the testing dataset using five machine learning models. Colors indicate the number of data points in each ppbv bin.

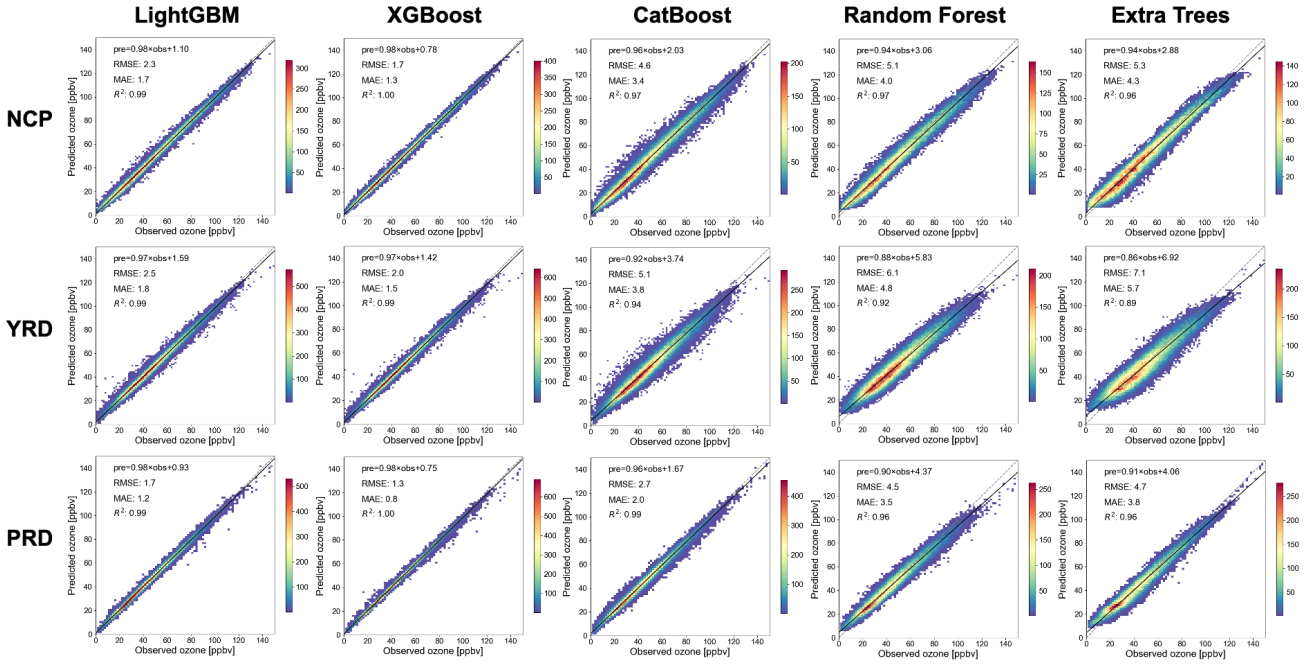


Figure S3. Density plots of observed and predicted MDA8 ozone concentrations across three regions for the training set using the five machine learning models. The color indicates the number of data pairs in each ppbv bin.

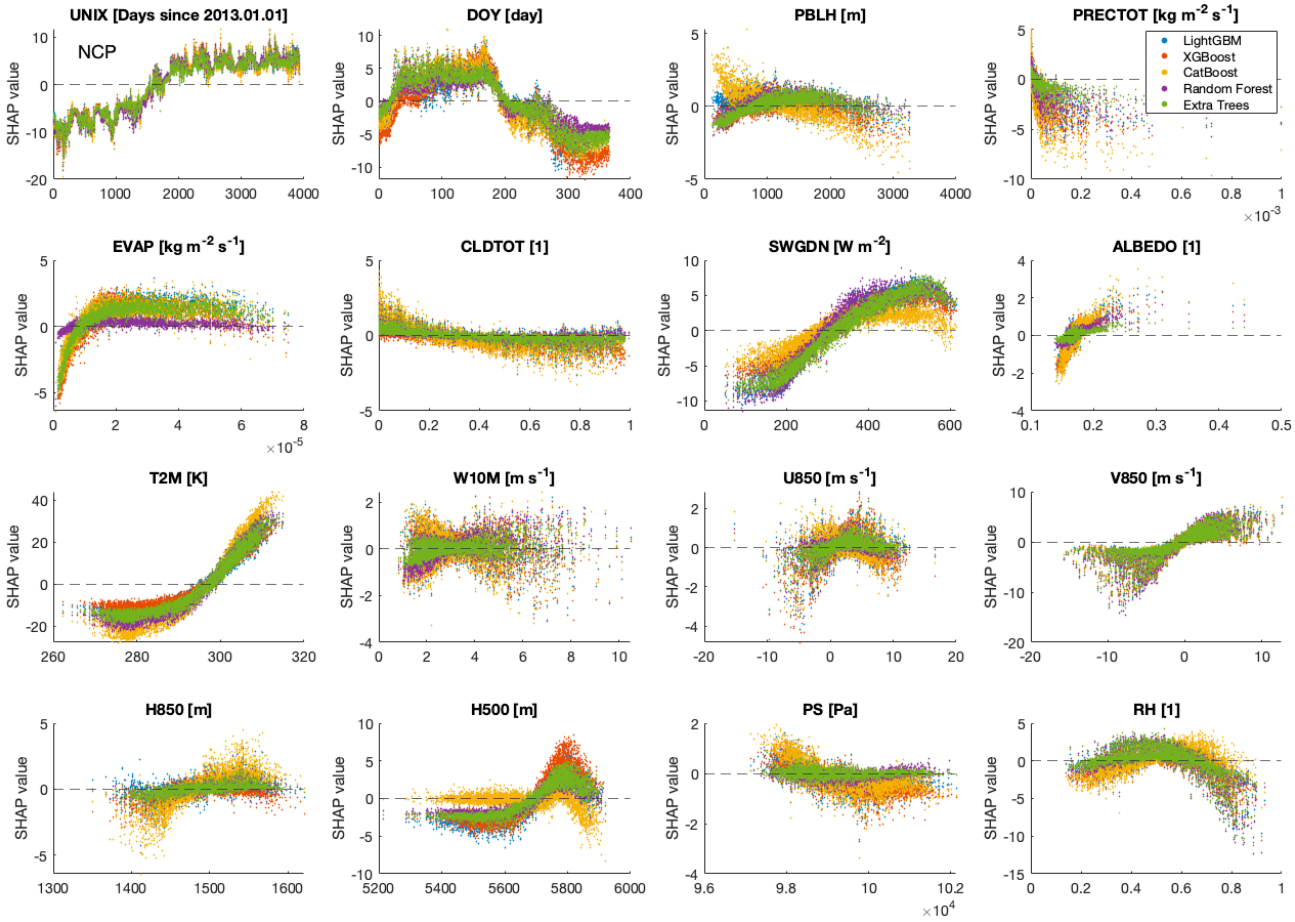


Figure S4. Dependence plots of meteorological variables and their SHAP values in NCP. Colors denote different machine learning models.

40

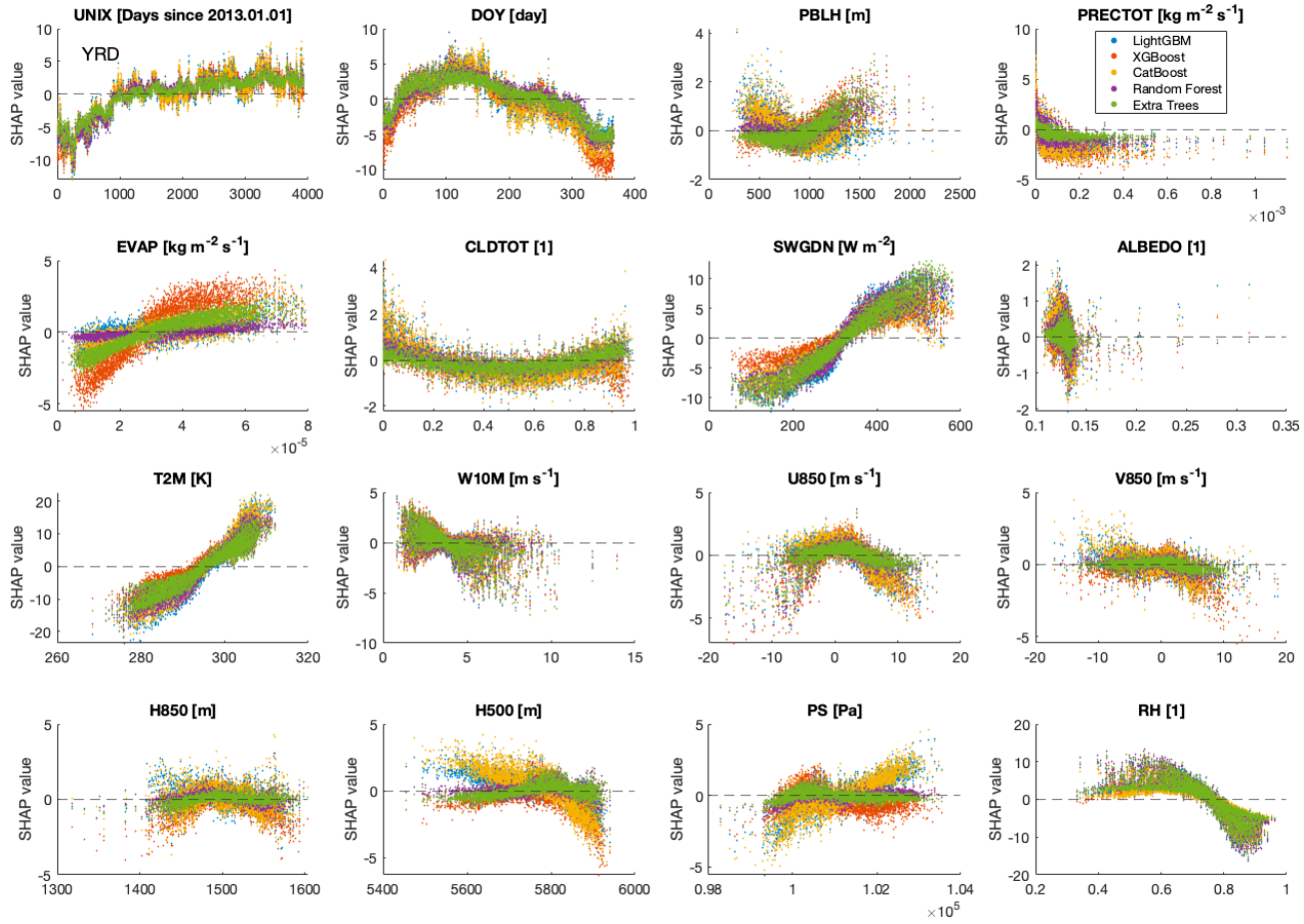
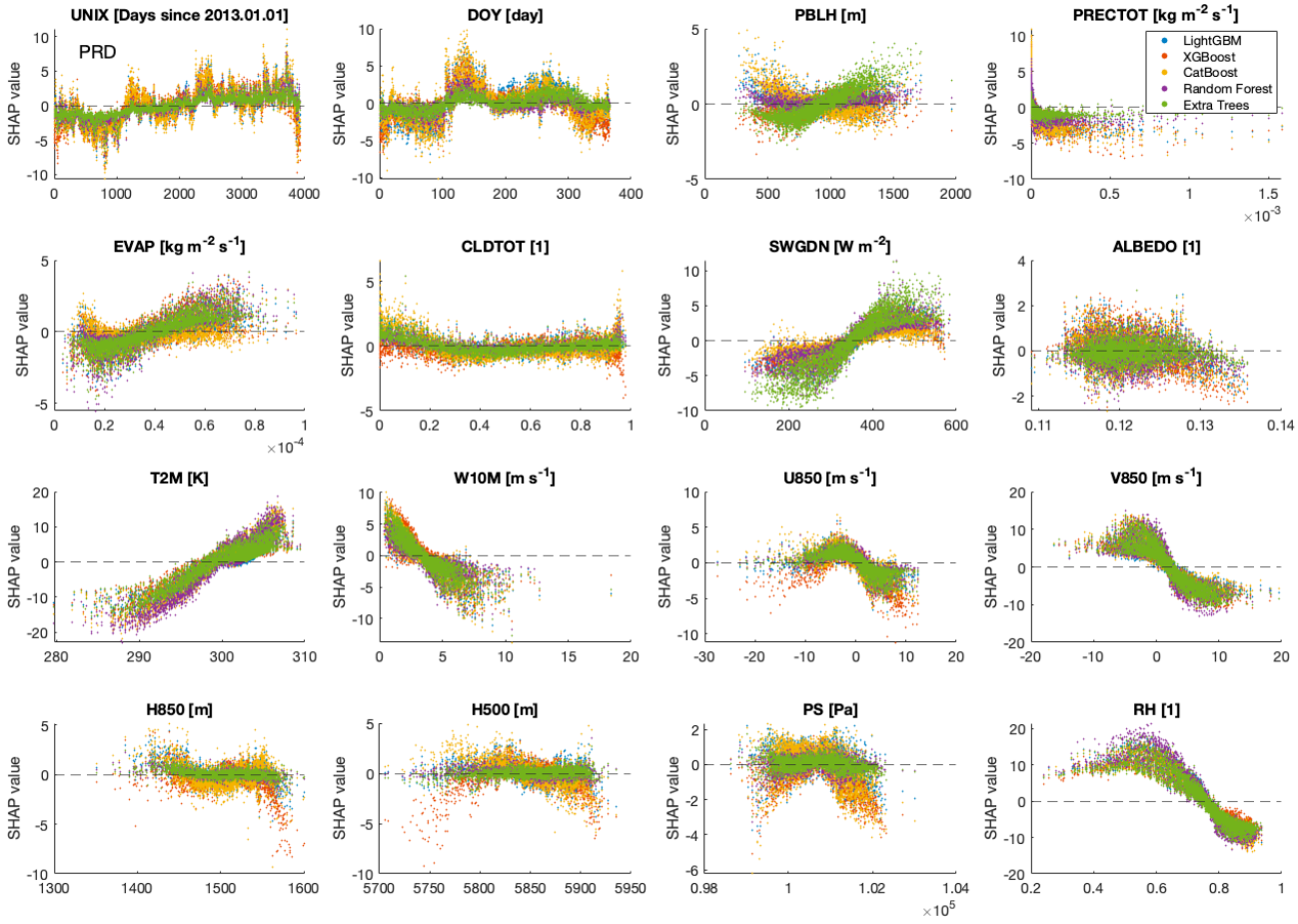


Figure S5. The same as Figure S4, but for the YRD region.



45

Figure S6. The same as Figure S4, but for the PRD region.

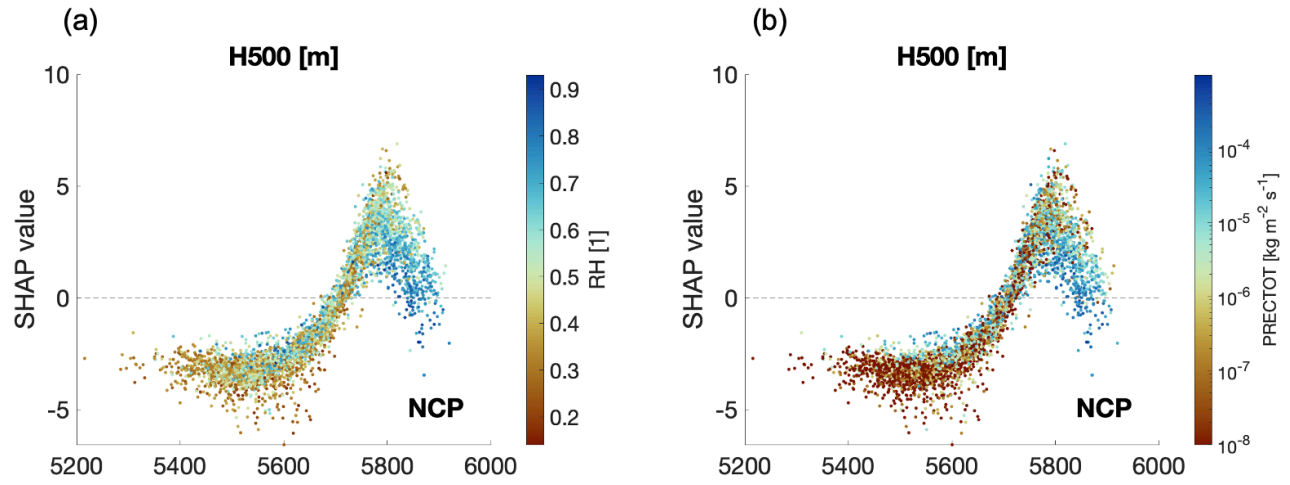


Figure S7. Dependence plots of H500 and its SHAP values (LightGBM model results) in NCP, color-coded by (a) relative humidity (RH) and (b) precipitation (PRECTOT).

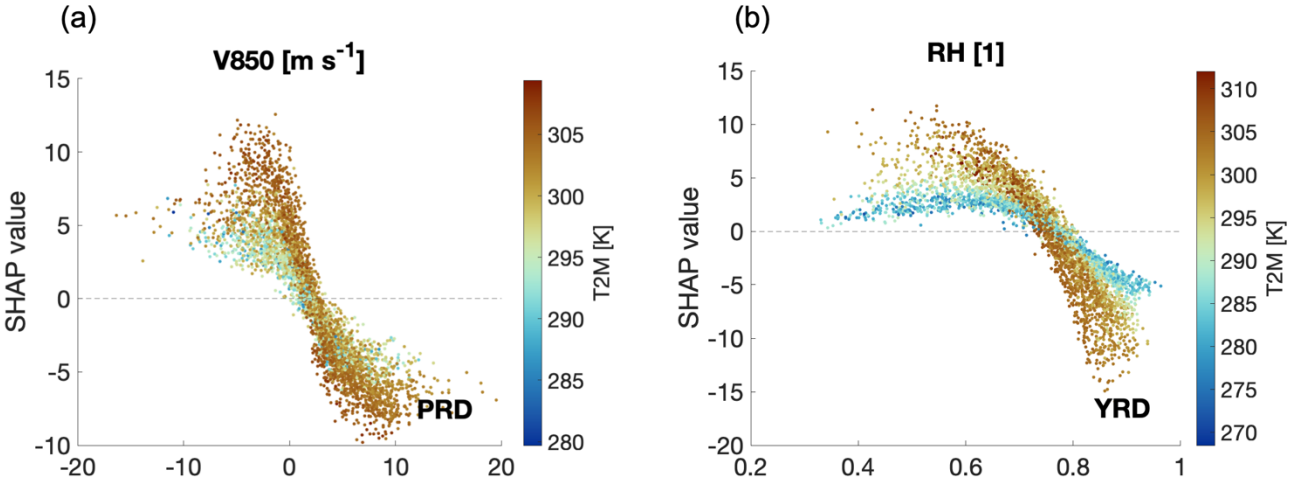


Figure S8. Dependence plots of SHAP values (LightGBM model results) for V850 in PRD and for RH in YRD, color-coded by the temperature level.

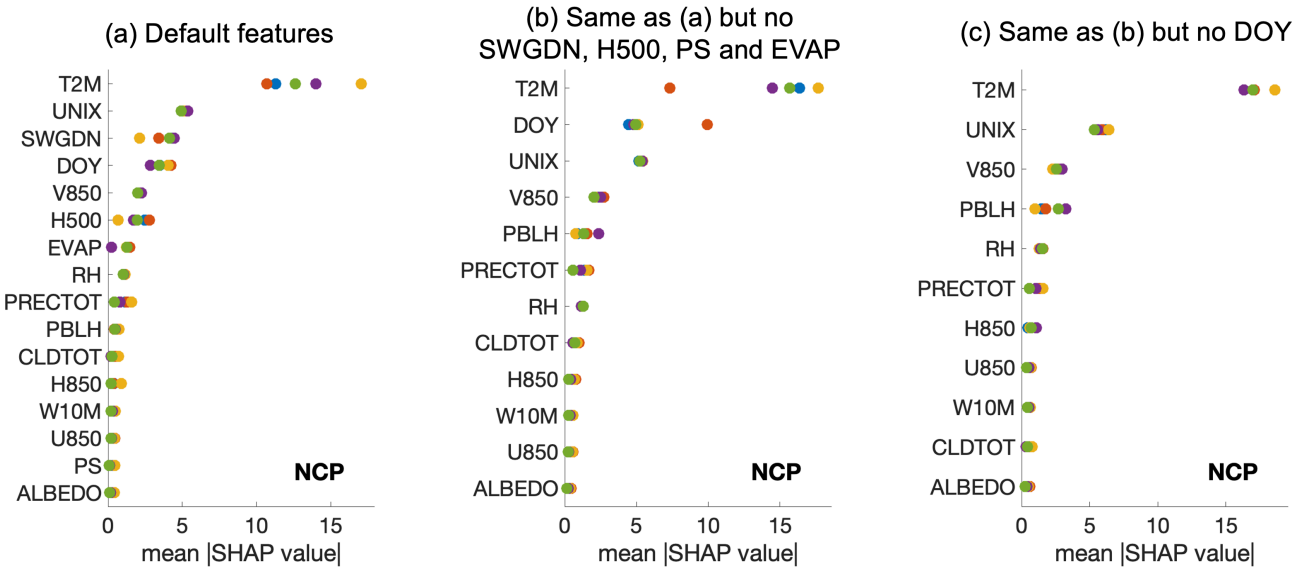


Figure S9. Global feature importance based on the mean absolute SHAP values in NCP. (a) Full model with all 16 features; (b) Same as (a) but excluding SWGDN, H500, PS, and EVAP; (c) Same as (b) but further excluding day of year (DOY). Colors denote different machine learning models.

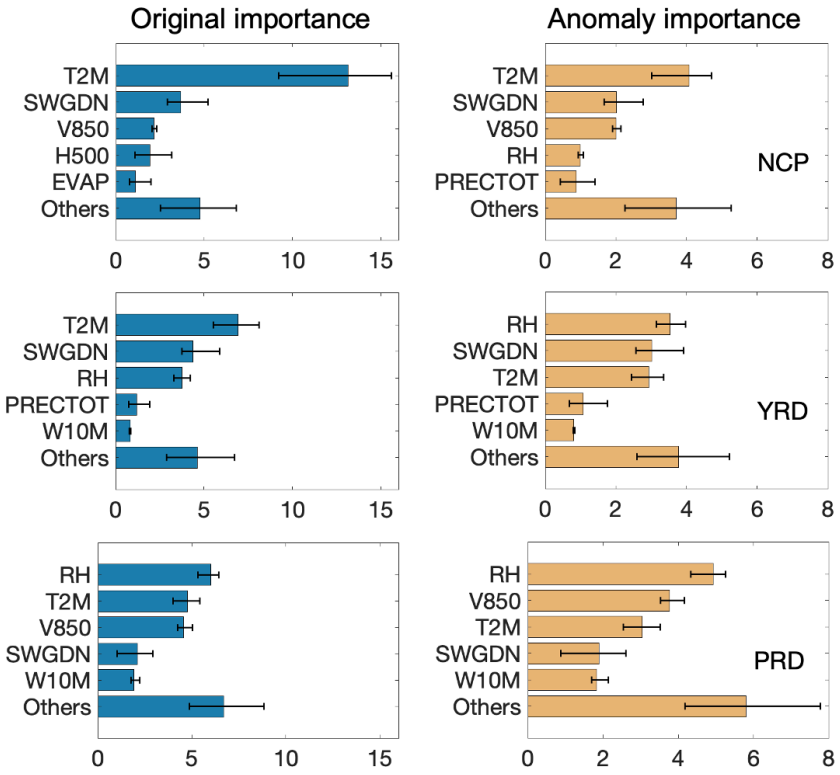


Figure S10. Comparison of global feature importance calculated from original SHAP values (G_i) and anomaly SHAP value (G'_i) after de-weathering. Error bars denote ranges across different machine learning models.

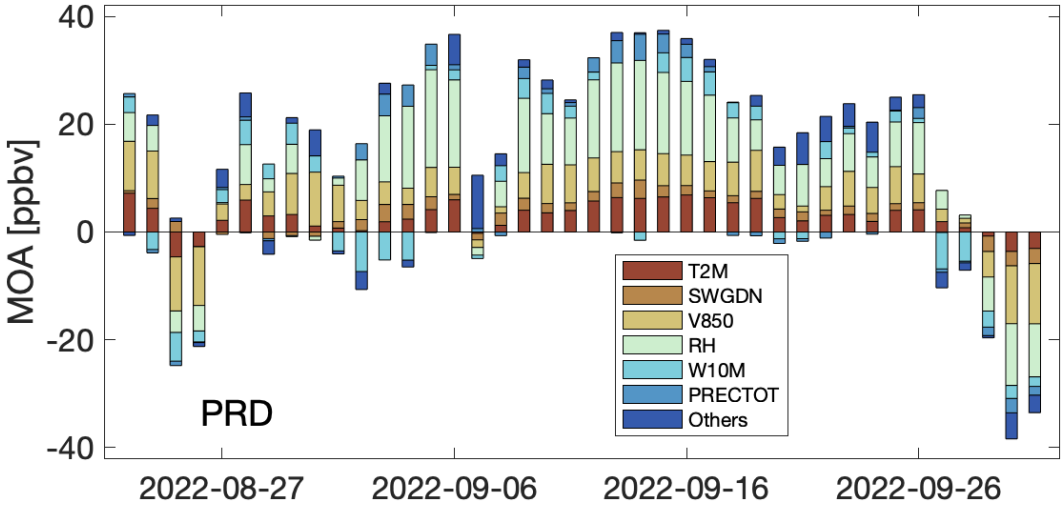


Figure S11. A persistent MOA event lasting over 30 days during autumn 2022 in the PRD region.

Table S1. Sensitivity of surface ozone to temperature (dO_3/dT) across three regions, as estimated by the dependence plots of SHAP values

Region	LightGBM	XGBoost	CatBoost	Random Forest	Extra Trees
NCP	1.74	1.77	2.70	2.15	1.98
YRD	1.25	0.85	1.27	0.91	0.81
PRD	1.11	1.13	1.43	1.29	0.80

Table S2. Statistics of meteorology-induced ozone anomalies (MOA) during warm seasons, averaged across five ML models. The first two columns report the mean and maximum values of positive MOA. The third column presents the average duration of consecutive positive MOA days. The right two columns show the percentage of positive MOA events lasting exactly 1 day and those exceeding 7 days. Ranges in parentheses indicate inter-model variation.

Region	Mean MOA [ppbv]	Max MOA [ppbv]	Duration [days]	Percentage of MOA events lasting 1 day [%]	Percentage of MOA events lasting >7 days [%]
NCP	11.7 (11.4–12.0)	43.3 (41.0–46.0)	3.1 (3.1–3.1)	28.1 (26.9–29.7)	6.9 (6.4–7.7)
YRD	11.5 (10.9–12.0)	37.5 (35.2–39.8)	3.4 (3.3–3.5)	25.1 (23.9–25.9)	8.4 (7.4–9.0)
PRD	15.9 (15.8–16.0)	52.8 (51.9–54.0)	4.3 (4.2–4.4)	26.2 (24.8–29.2)	10.8 (10.5–10.9)

Table S3. Rankings of global feature importance as determined by SHAP, Gain, and Permutation (Perm) methods.

	NCP			YRD			PRD		
rank	**SHAP**	**Gain**	**Perm**	**SHAP**	**Gain**	**Perm**	**SHAP**	**Gain**	**Perm**
1	T2M	T2M	T2M	T2M	SWGDN	T2M	RH	RH	RH
2	UNIX	SWGDN	UNIX	SWGDN	T2M	RH	T2M	V850	T2M
3	SWGDN	H500	DOY	RH	RH	DOY	V850	T2M	V850
4	DOY	DOY	SWGDN	DOY	UNIX	SWGDN	PRECTOT	SWGDN	ALBEDO
5	H500	UNIX	V850	UNIX	DOY	UNIX	W10M	DOY	W10M
6	V850	EVAP	H500	PRECTOT	ALBEDO	ALBEDO	SWGDN	W10M	DOY
7	EVAP	V850	RH	ALBEDO	W10M	PS	DOY	ALBEDO	PRECTOT
8	PRECTOT	RH	PRECTOT	PS	U850	W10M	ALBEDO	PRECTOT	UNIX
9	RH	PRECTOT	EVAP	W10M	PS	PRECTOT	UNIX	UNIX	U850
10	ALBEDO	PBLH	PS	U850	PRECTOT	U850	U850	U850	SWGDN
11	PS	U850	ALBEDO	EVAP	EVAP	EVAP	PS	EVAP	PS
12	W10M	PS	PBLH	H500	V850	V850	EVAP	H500	EVAP
13	U850	H850	U850	CLDTOT	H500	H500	H850	PS	H850
14	PBLH	ALBEDO	W10M	H850	H850	H850	H500	H850	H500
15	CLDTOT	W10M	H850	V850	CLDTOT	CLDTOT	PBLH	PBLH	PBLH
16	H850	CLDTOT	CLDTOT	PBLH	PBLH	PBLH	CLDTOT	CLDTOT	CLDTOT