Appendix A. Supplementary Material

Text S1 Evaluation of WRF-CMAQ performance.

The meteorological observation data at Sugang (in Shunde, PRD) and Shengzhongzichangku (in Jinan, NWShD) were used to evaluate the performance of WRF model in the PRD and the NWShD. The definitions of the statistical indicator, i.e., the Normalized Mean Bias (NMB) and Correlation coefficients (R), were summarized in **Table S2**, and the performance of WRF model was summarized in **Table S3**. As shown in **Table S3**, the NMBs of temperature, wind speed, and relative humidity were varying from -95.88 to 1.22%, from -22.41 to 154.95%, and from -2.99 to 7.07%, respectively; and their R values were ranging from 0.7647 to 0.9257, from 0.4438 to 0.7623, and from 0.7436 to 0.9078, respectively. Furthermore, as shown in **Figs. S2-3**, the performance of CMAQ model was evaluated through comparisons between simulations and observations at three representative monitoring sits in each d03 domains, which represent the urban sites (GD Business College in the PRD and Shengzhongzichangku in the NWShD), industrial site (Sugang in the PRD and Diershuichang in the NWShD) and rural site (Dongchengzhushan in the PRD and Changqingdangxiao in the NWShD). The NMBs of CMAQ in predicting O_3 concentration over the PRD ranged from -15.91 to 19.52%, which were mostly within the recommended benchmark (NMB < ±15%) (Emery et al., 2017) except GD Business College in April (-15.91%) and Sugang in October (19.52%). The R values for O_3 predictions over the PRD ranged from 0.6752 to 0.8176, all meeting the benchmark (R > 0.5) (Emery et al., 2017). Compared to the PRD, the model performance of CMAQ for the NWShD was slightly worse. However, the most NMBs and R values for O_3 predictions over the NWShD (-12.34 ~ 13.12% and 0.5171 ~ 0.6913, respectively) met the recommended values, except that the NMBs in the three representative monitoring sites in January (-26.70, -17.43, and -25.64%, respectively) were lower, the NMB in Changqingdangxiao in July (21.61%) was higher, and the R values in Shengzhongzichangku in January (0.4884), April (0.4552), and October (0.4942) were slightly lower. Taken together, the magnitude and overall temporal trends of O_3 concentration in two case studies were reasonably predicted.

Text S2 Definitions of the indicators in the termination criterion.

Coefficient of Determination (R^2)

$$R^2 = \frac{\sum_{k=1}^{NS}\left(M_k - \bar{M}\right)^2}{\sum_{k=1}^{NS}\left(O_k - \bar{O}\right)^2} \quad \text{(S1)}$$

Mean Bias (MeanB)

$$\text{MeanB} = \frac{1}{NS}\sum_{k=1}^{NS}\frac{M_k - O_k}{\text{Base}} \quad \text{(S2)}$$

Max Bias (MaxB)

$$\text{MaxB} = \max\left(\frac{M_k - O_k}{\text{Base}}\right) \quad \text{(S3)}$$

Where:

M_k is ERSM-predicted the change of O_3 concentration (ΔConc) in control scenarios k; O_k is the CMAQ-predicted ΔConc in control scenario k; $\bar{M}$ is the average of ERSM-predicted ΔConc in all control scenarios; $\bar{O}$ is the average of CMAQ-predicted ΔConc in all control scenarios; NS is the number of control scenarios; and Base is the O_3 concentration in the base case.

Text S3 The calculation process of the ERSM (Xing et al., 2017).

$$\Delta \mathrm{Conc}_t = \sum_{S=1}^{N} \mathrm{CM}_{S \to t} + \sum_{S=1}^{N} \mathrm{TP}_{S \to t} + \mathrm{CM\_IR}_t + \mathrm{TP\_IR}_t \quad \text{(S4)}$$

$$\mathrm{CM}_{S \to t} = \mathrm{RSM\_O3}_{S \to t}\left(\mathrm{E'}_{\mathrm{prec},T}\right) = \begin{cases} \mathrm{RSM\_O3}_{S \to t}\left(\mathrm{RSM\_}\left(\mathrm{A,B,C,D}\right)_{S \to t}\left(\mathrm{E}_{\mathrm{prec},S}\right)\right) & \text{while } S \neq T \\ \mathrm{RSM\_O3}_{S \to t}\left(\mathrm{E}_{\mathrm{prec},S}\right) & \text{while } S = T \end{cases} \quad \text{(S5)}$$

$$\mathrm{TP}_{S \to t} = \begin{cases} \mathrm{RSM\_O3}_{S \to t}\left(\mathrm{E}_{\mathrm{prec},S}\right) - \mathrm{CM}_{S \to t} & \text{while } S \neq T \\ 0 & \text{while } S = T \end{cases} \quad \text{(S6)}$$

$$\mathrm{CM\_IR}_{S \to t} = \mathrm{RSM\_O3}_{T \to t}\left(\sum_{S=1}^{N} \mathrm{RSM\_(A,B,C,D)}_{S \to t}\left(\mathrm{E}_{\mathrm{prec},S}\right)\right) - \sum_{S=1}^{N} \mathrm{CM}_{S \to t} \quad \text{(S7)}$$

$$\mathrm{TP\_IR}_t = \mathrm{RSMtt\_O3}_{S \to t}\left(\overline{\mathrm{E}}_{\mathrm{prec}}\right) - \sum_{S=1}^{N} \mathrm{CM}_{S \to t} - \sum_{S=1}^{N} \mathrm{TP}_{S \to t} - \mathrm{CM\_IR}_t \quad \text{(S8)}$$

Where:

$\Delta \mathrm{Conc}_t$ is the change of O_3 concentration in receptor grid t in region T; N is the number of regions; $\mathrm{E}_{\mathrm{prec},S}$ is the change ratios of four precursors in source region S; $\mathrm{E'}_{\mathrm{prec},T}$ represents the change ratios of precursor concentrations in the aggregated grids of receptor region T associated with $\mathrm{E}_{\mathrm{prec},S}$; $\overline{\mathrm{E}}_{\mathrm{prec}}$ is the average of [$\mathrm{E}_{\mathrm{prec},1}$, …, $\mathrm{E}_{\mathrm{prec},T-1}$, $\mathrm{E}_{\mathrm{prec},T+1}$,…, $\mathrm{E}_{\mathrm{prec},N}$];

$\mathrm{RSM\_P}_{S \to t}\left(\mathrm{E'}_{\mathrm{prec},T}/\mathrm{E}_{\mathrm{prec},S}\right)$ is the single-regional RSM system that models the response of pollutant P (i.e. O_3 and its precursors) at receptor grid t to $\mathrm{E'}_{\mathrm{prec},T}$ or $\mathrm{E}_{\mathrm{prec},S}$; A, B, C, and D represent the precursors, i.e. NO_X, SO_2, NH_3, and AVOC, respectively; and

$\mathrm{RSMtt\_O3}_{S \to t}\left(\overline{\mathrm{E}}_{\mathrm{prec}}\right)$ represents the total-regional RSM system in which emissions in all regions changed simultaneously and is used to calculate the response of O_3 at receptor grid t to $\overline{\mathrm{E}}_{\mathrm{prec}}$.

Text S4 Definition of the percent of data point (P), as follows:

$$P_i = \frac{NumOP_i}{MaxOP} \tag{S9}$$

Where:

P_i is the percent of data point which represents the the density of points falling at *i*th grid; $NumOP_i$ is the number of overlapping points in *i*th grid; and MaxOP is the maximum number of overlapping points in grids.

Text S5 Definitions of chemical response indicators for ozone (Xing et al., 2011; Xing et al., 2018; Xing et al., 2019)

For O_3, the PR is the NO_X emissions that produce the maximum O_3 concentrations under baseline AVOC emissions, and the PR can be directly calculated from the polynomial function:

$$PR = 1 + E_{NOx}\Big|_{\frac{\partial \Delta Conc}{\partial E_{NOx}}=0} \quad \text{(S10)}$$

Where:

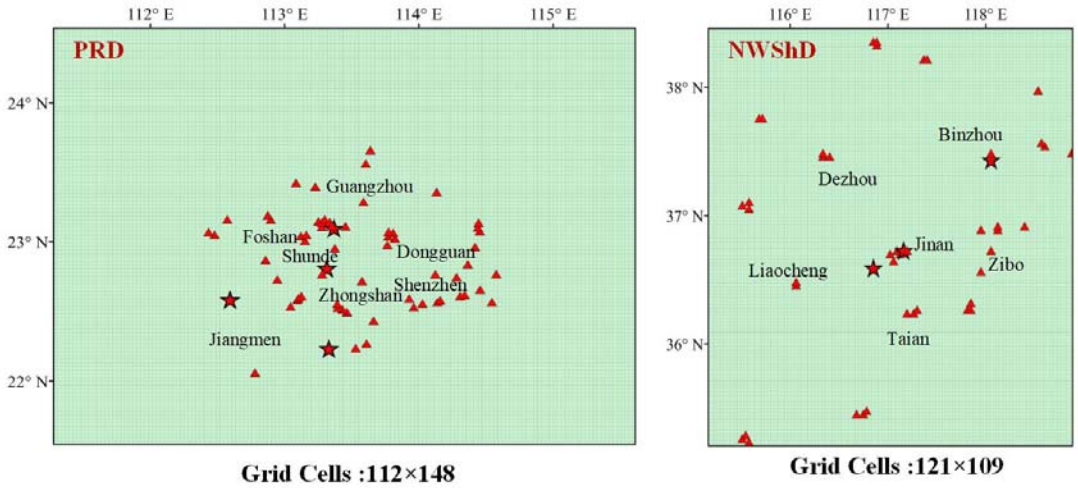
$\frac{\partial \Delta Conc}{\partial E_{NOx}}$ is the first derivation of the ΔConc to E_{NOx}.

The level of simultaneous control of AVOC to prevent an increase in O_3 levels from the NOx controls is defined by the ratio of AVOC to NOx (i.e., VNr) corresponding to the PR and is calculated as follows:

$$VNr = r\Big|_{\frac{\partial \Delta Conc}{\partial E_{NOx}}=0} \quad \text{when PR<1, } r = \frac{E_{AVOC}}{E_{NOx}} \quad \text{(S11)}$$

Where:

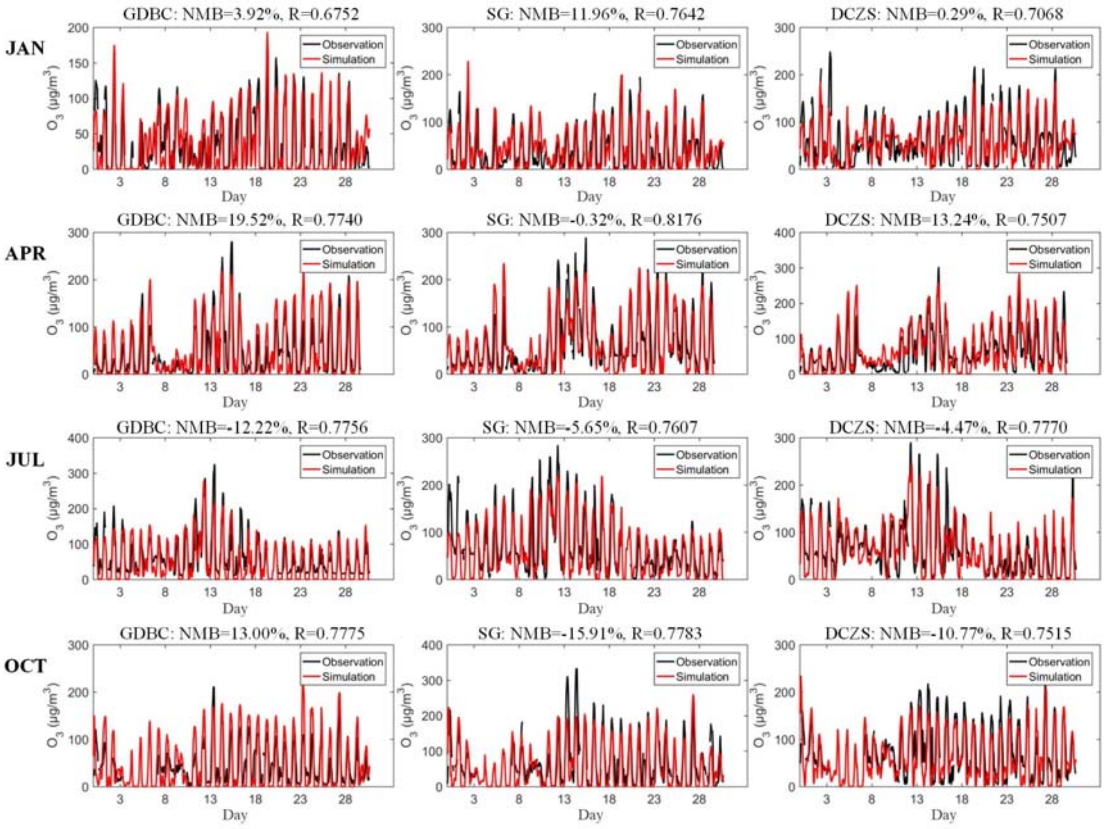
$\frac{\partial \Delta Conc}{\partial E_{NOx}}$ is the first derivative of the ΔConc to E_{NOx}, when $E_{AVOC} = r \times E_{NOx}$, and E_{SO_2} and E_{NH_3} are 0.



81

82 **Fig. S1 S**patial distribution of the national-controlled air-monitoring sites in the innermost
83 modeling domains for two case studies.

84 Note: the triangular points represent the monitoring sites, and the pentacle points represent the
85 selected monitoring sites for evaluation of the model performance.
86



87

88 **Fig. S2** The temporal variations of the observed (black line) and simulated (red line) hourly
89 average O_3 concentrations at three monitoring sites in the PRD, in 2015.

90 Note: GDBC-GD Business College; SG-Sugang; DCZS-Dongchengzhushan.

91

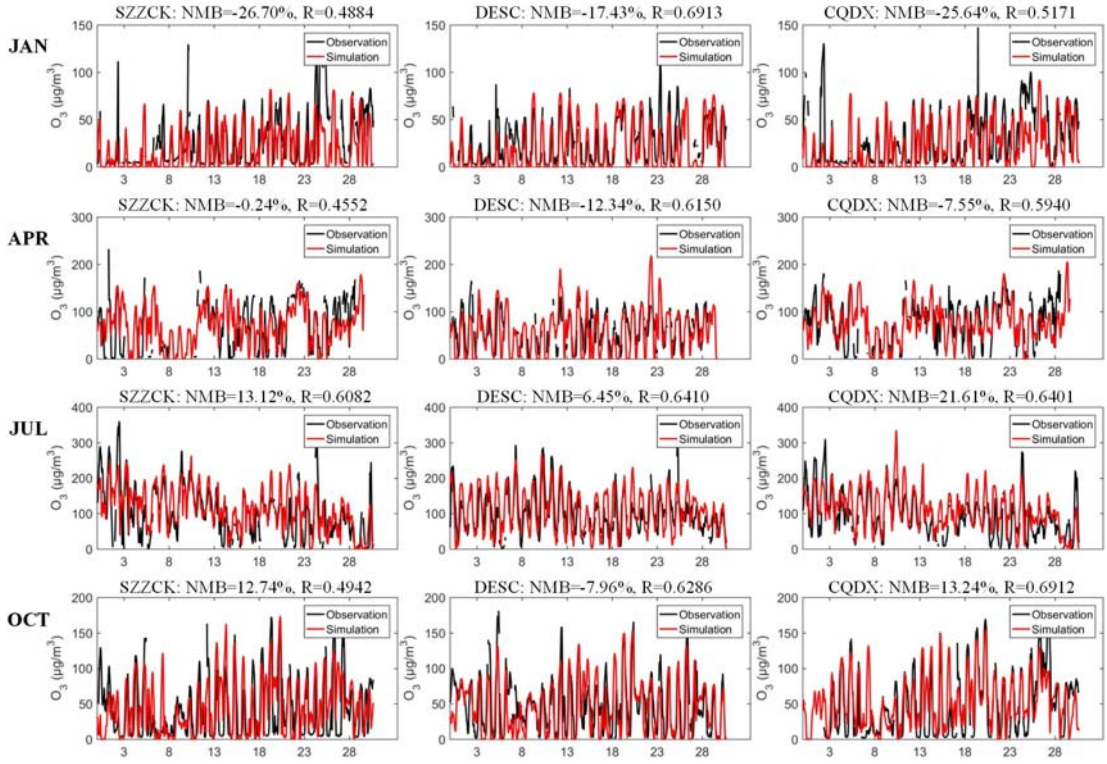


Fig. S3 The temporal variations of the observed (black line) and simulated (red line) hourly average O_3 concentrations at three monitoring sites in the NWShD, in 2017.

Note: SZZCK-Shenzhongzichangku; DESC-Diershuichang; CQDX-Changqingdangxiao.

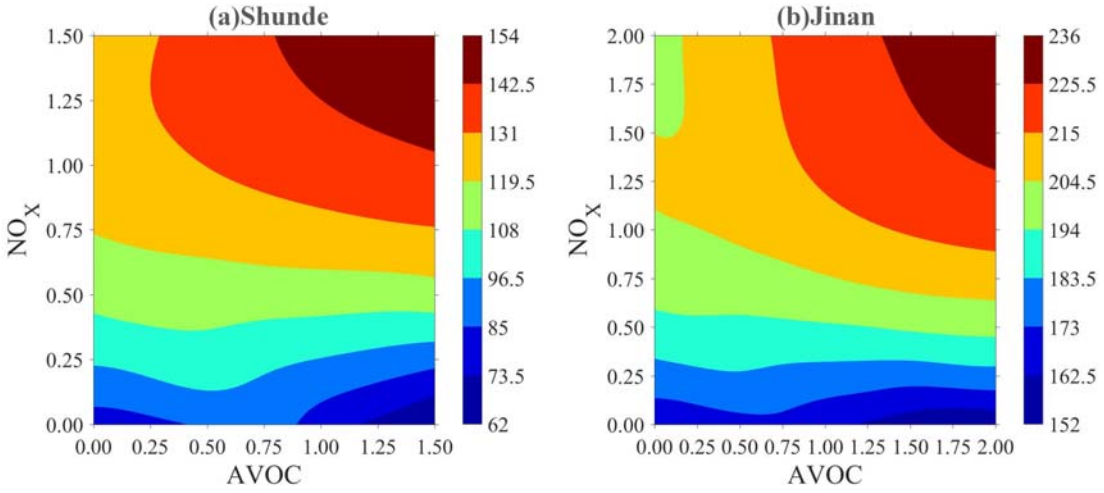


Fig. S4 The isopleths of O_3 in response to the simultaneous changes of precursor emissions in July, in (a) the Shunde, PRD and in (b) the Jinan, NWShD (monthly averaged daily 1 h maxima O_3; unit: $\mu g/m^3$).

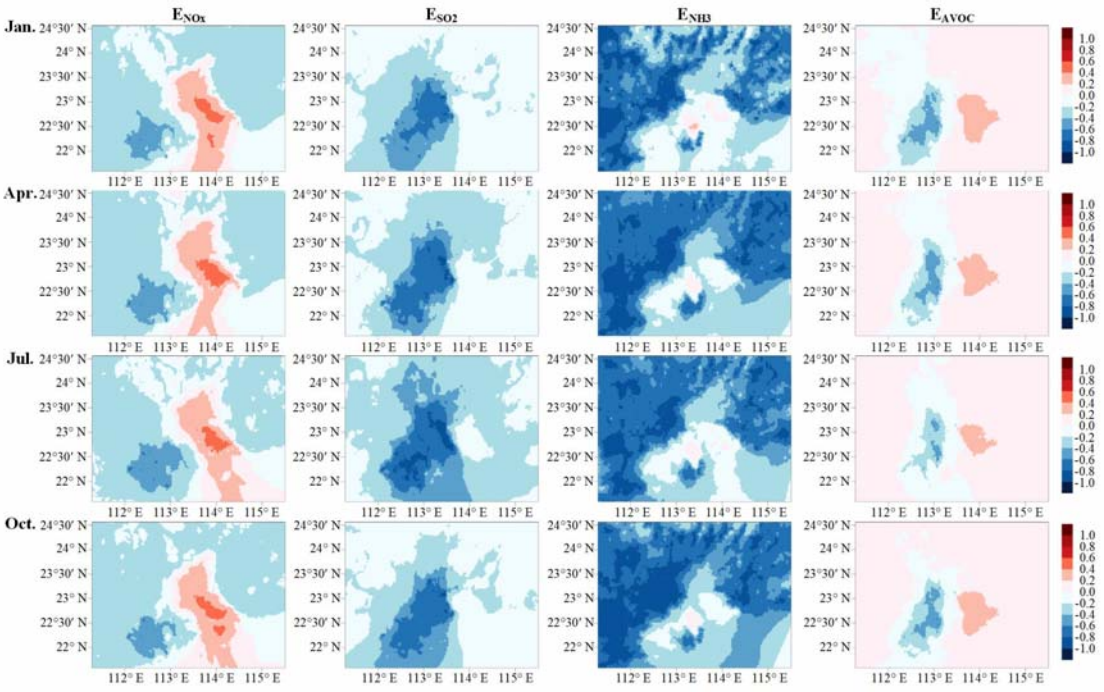
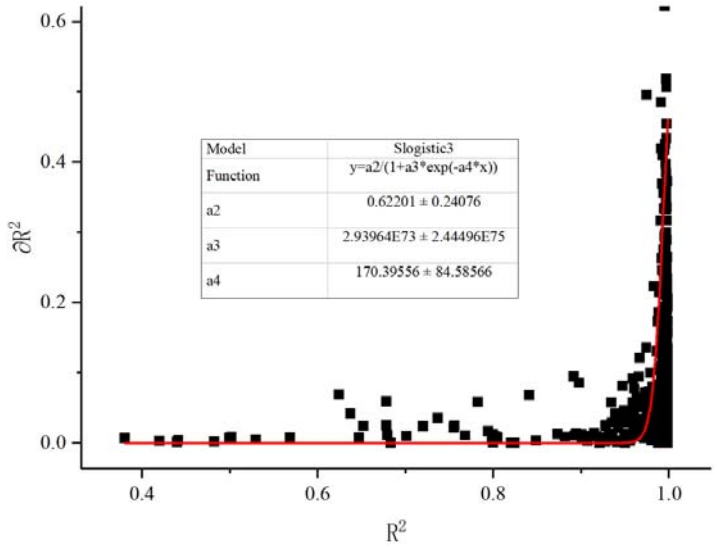


Fig. S5 Spatial distribution of pf-ERSM-predicted change ratios of precursors under one control scenario in the PRD.

Note: the control scenario is scenario 5 listed in **Table S5**.



106

107 **Fig. S6** The scatter plot of R^2 and ∂R^2 of the best suited polynomial functions (Black Square)
108 across grid cells containing monitoring sites.

109 Note: the red line is the fitted curve derived from the Slogistic3 model in Origin 2017.

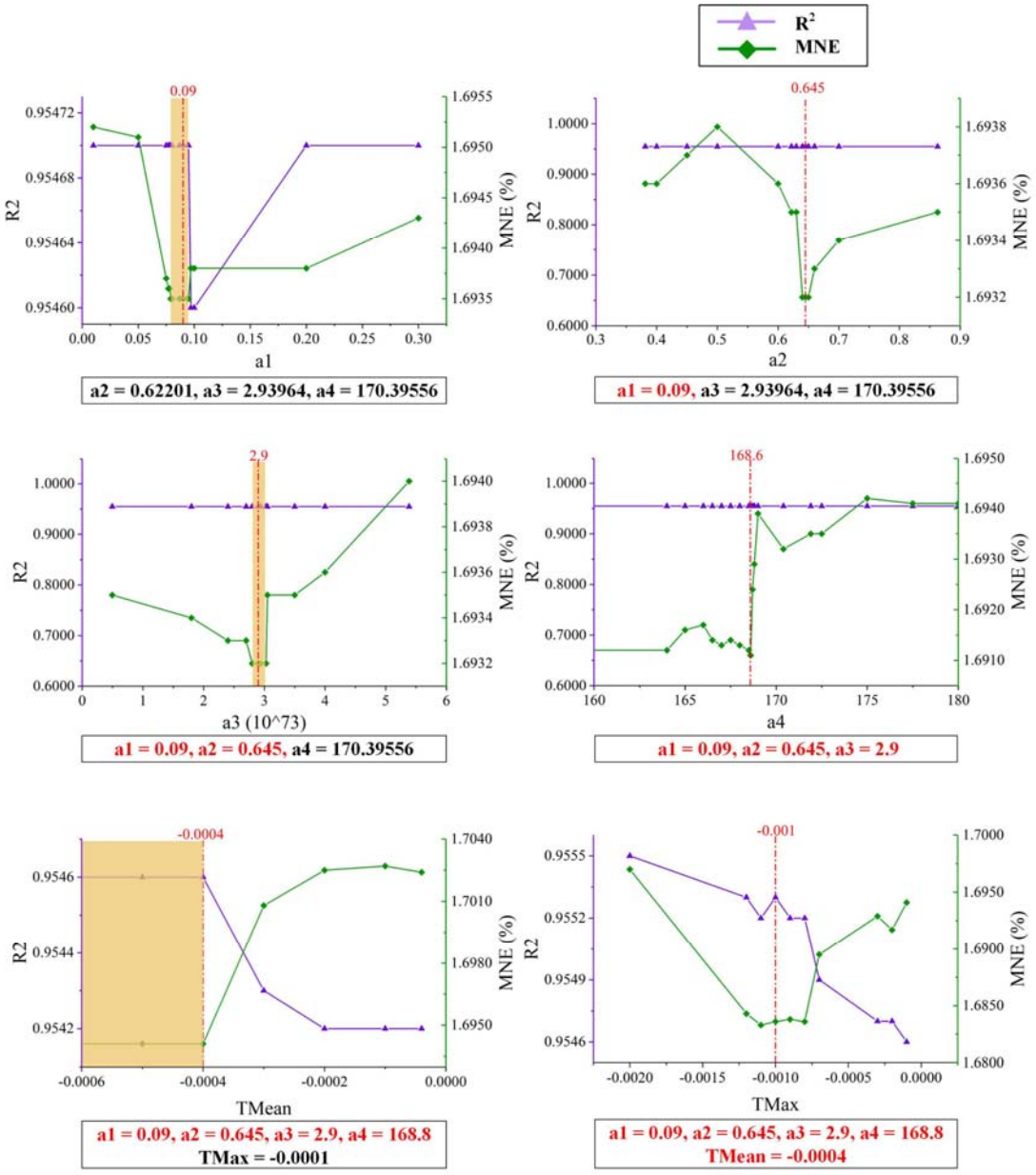


Fig. S7 The process of determining the thresholds of the termination criterion for single-region contribution.

Note: the orange shading represents the optimal range of the parameter value; the red dotted line represents the value determined for the current parameter; and the red fonts represent the parameter values that have been determined.

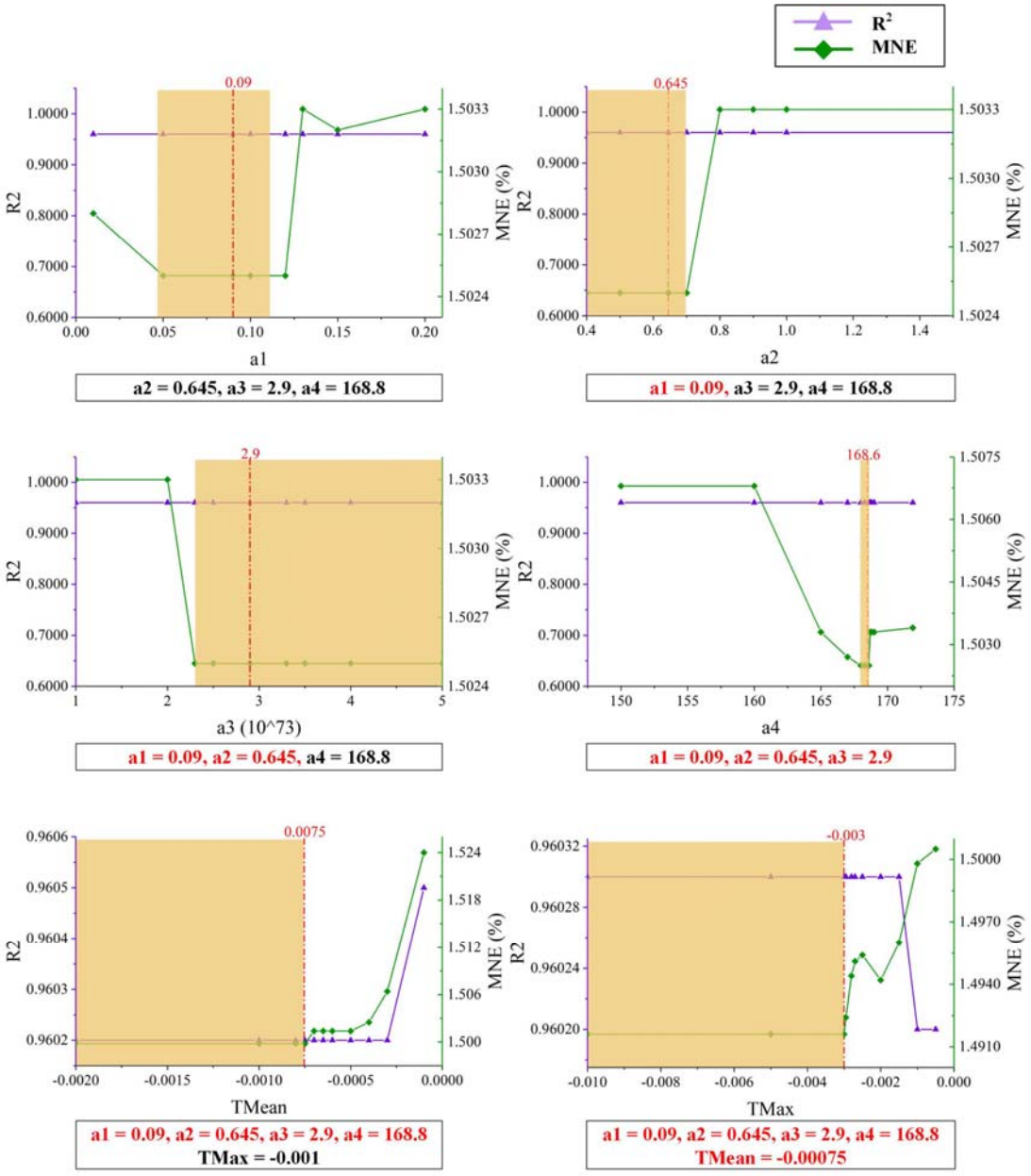
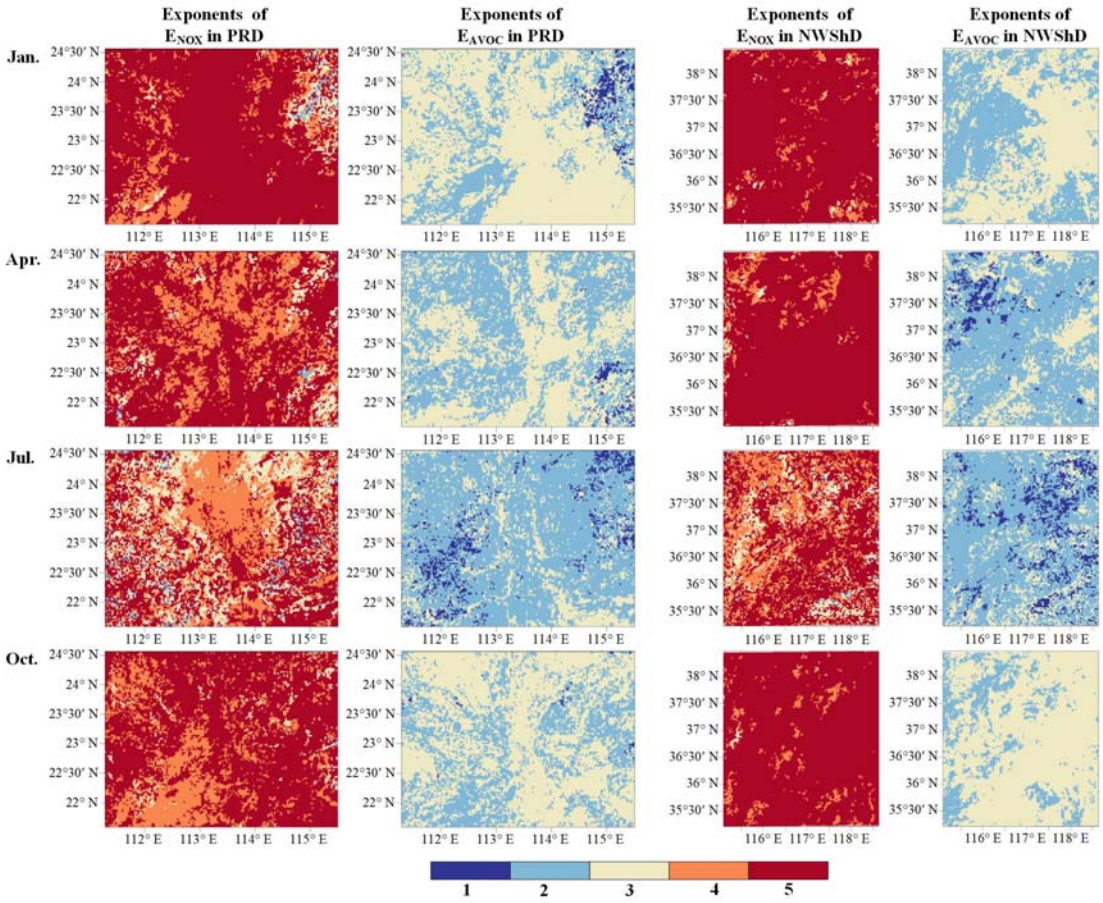


Fig. S8 The process of determining the thresholds of the termination criterion for inter-region contribution.

Note: the orange shading represents the optimal range of the parameter value; the red dotted line represents the value determined for the current parameter; and the red fonts represent the parameter values that have been determined.

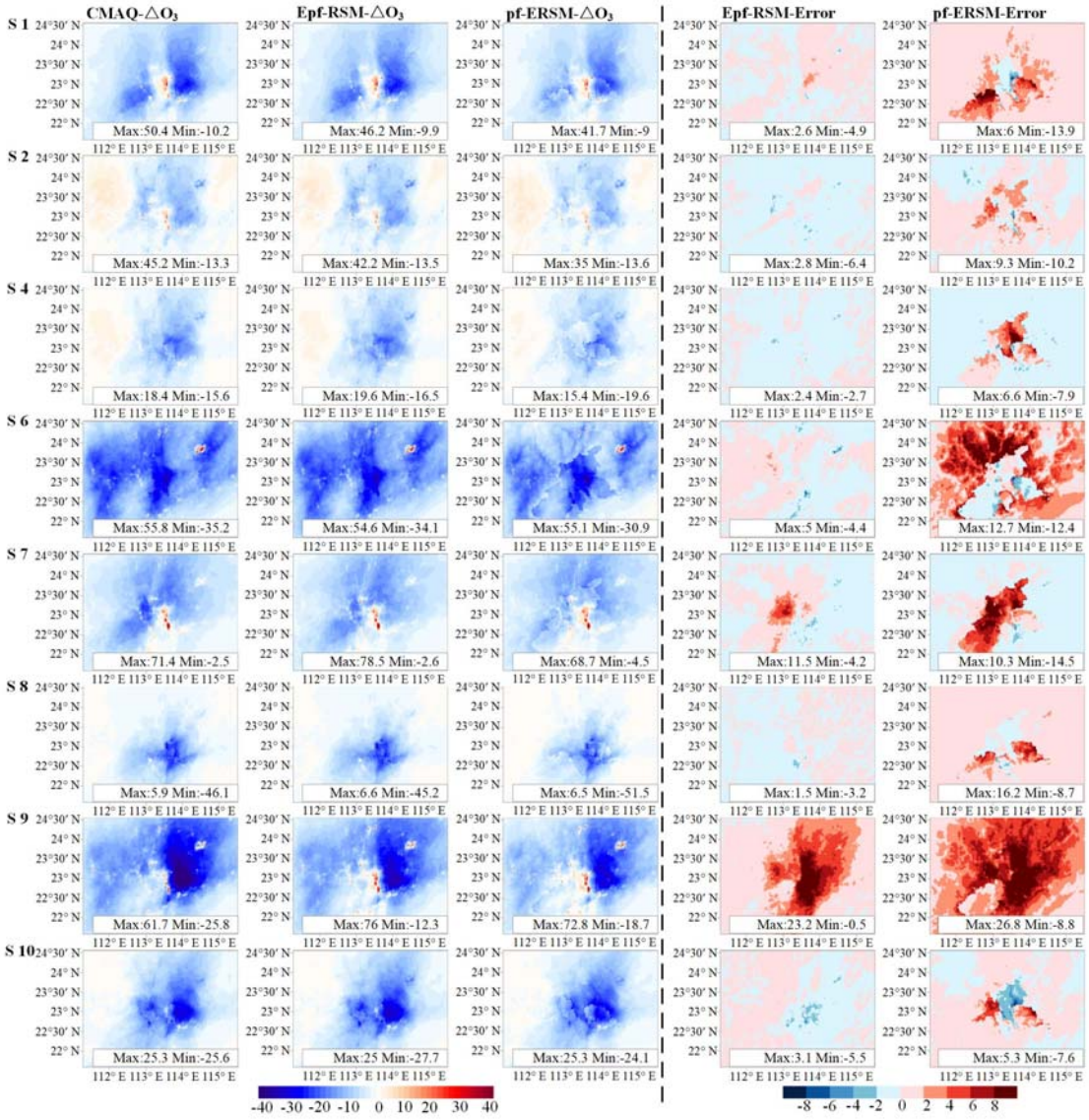


122

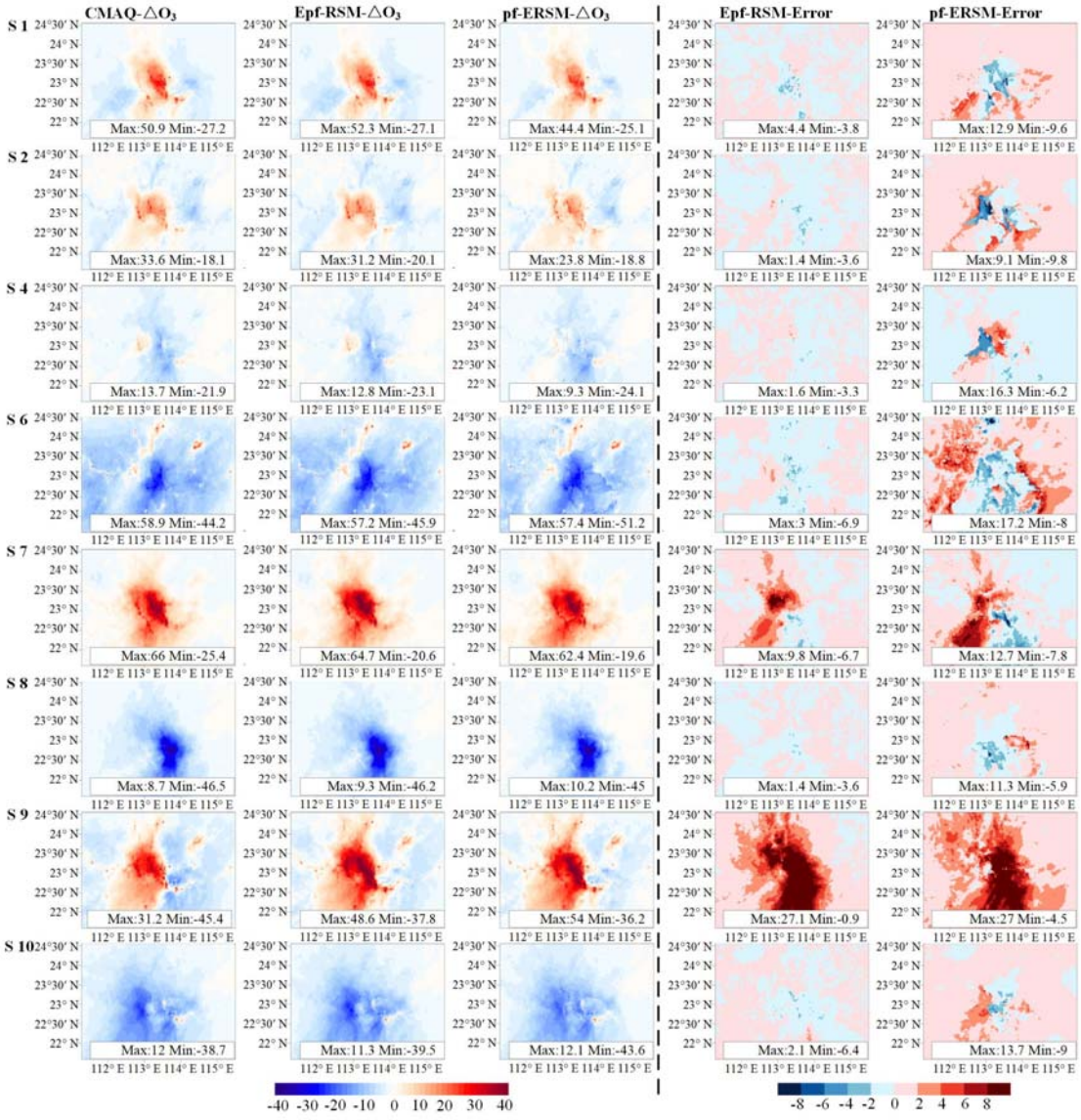
123 **Fig. S9** Spatial distribution of the exponents of E_{NOx} and E_{AVOC}.

124

(a) January


CMAQ-△O3
Epf-RSM-△O3
pf-ERSM-△O3
Epf-RSM-Error
pf-ERSM-Error
S 1
S 2
S 4
S 6
S 7
S 8
S 9
S 10
Max:50.4 Min:-10.2
Max:46.2 Min:-9.9
Max:41.7 Min:-9
Max:2.6 Min:-4.9
Max:6 Min:-13.9
Max:45.2 Min:-13.3
Max:42.2 Min:-13.5
Max:35 Min:-13.6
Max:2.8 Min:-6.4
Max:9.3 Min:-10.2
Max:18.4 Min:-15.6
Max:19.6 Min:-16.5
Max:15.4 Min:-19.6
Max:2.4 Min:-2.7
Max:6.6 Min:-7.9
Max:55.8 Min:-35.2
Max:54.6 Min:-34.1
Max:55.1 Min:-30.9
Max:5 Min:-4.4
Max:12.7 Min:-12.4
Max:71.4 Min:-2.5
Max:78.5 Min:-2.6
Max:68.7 Min:-4.5
Max:11.5 Min:-4.2
Max:10.3 Min:-14.5
Max:5.9 Min:-46.1
Max:6.6 Min:-45.2
Max:6.5 Min:-51.5
Max:1.5 Min:-3.2
Max:16.2 Min:-8.7
Max:61.7 Min:-25.8
Max:76 Min:-12.3
Max:72.8 Min:-18.7
Max:23.2 Min:-0.5
Max:26.8 Min:-8.8
Max:25.3 Min:-25.6
Max:25 Min:-27.7
Max:25.3 Min:-24.1
Max:3.1 Min:-5.5
Max:5.3 Min:-7.6
24°30′ N
24° N
23°30′ N
23° N
22°30′ N
22° N
112° E 113° E 114° E 115° E
-40 -30 -20 -10 0 10 20 30 40
-8 -6 -4 -2 0 2 4 6 8

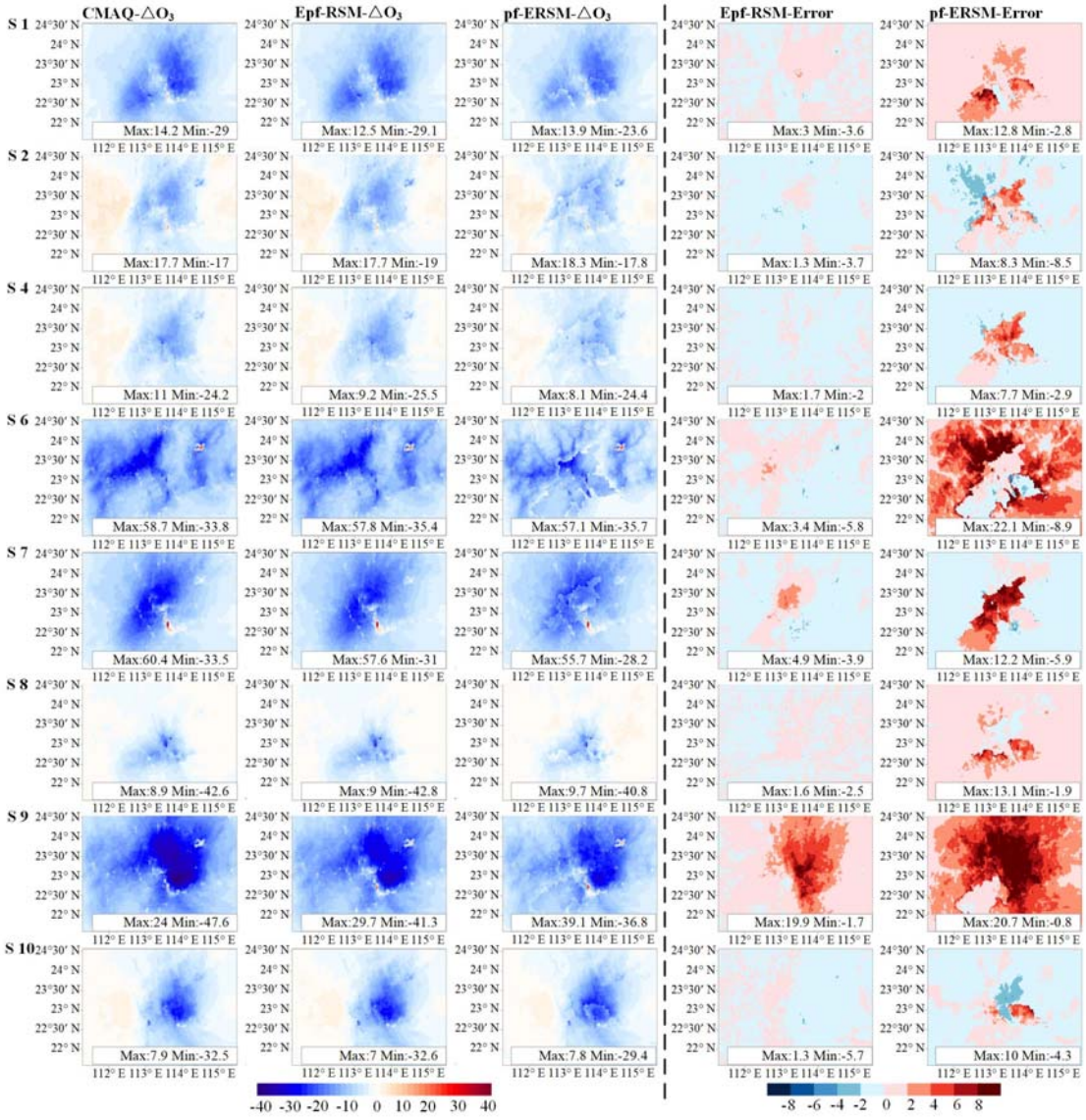
128 **(b) April**



129

130

(c) July


CMAQ-△O3
Epf-RSM-△O3
pf-ERSM-△O3
Epf-RSM-Error
pf-ERSM-Error
S 1
S 2
S 4
S 6
S 7
S 8
S 9
S 10
Max:14.2 Min:-29
Max:12.5 Min:-29.1
Max:13.9 Min:-23.6
Max:3 Min:-3.6
Max:12.8 Min:-2.8
Max:17.7 Min:-17
Max:17.7 Min:-19
Max:18.3 Min:-17.8
Max:1.3 Min:-3.7
Max:8.3 Min:-8.5
Max:11 Min:-24.2
Max:9.2 Min:-25.5
Max:8.1 Min:-24.4
Max:1.7 Min:-2
Max:7.7 Min:-2.9
Max:58.7 Min:-33.8
Max:57.8 Min:-35.4
Max:57.1 Min:-35.7
Max:3.4 Min:-5.8
Max:22.1 Min:-8.9
Max:60.4 Min:-33.5
Max:57.6 Min:-31
Max:55.7 Min:-28.2
Max:4.9 Min:-3.9
Max:12.2 Min:-5.9
Max:8.9 Min:-42.6
Max:9 Min:-42.8
Max:9.7 Min:-40.8
Max:1.6 Min:-2.5
Max:13.1 Min:-1.9
Max:24 Min:-47.6
Max:29.7 Min:-41.3
Max:39.1 Min:-36.8
Max:19.9 Min:-1.7
Max:20.7 Min:-0.8
Max:7.9 Min:-32.5
Max:7 Min:-32.6
Max:7.8 Min:-29.4
Max:1.3 Min:-5.7
Max:10 Min:-4.3
112° E 113° E 114° E 115° E
24°30′ N
24° N
23°30′ N
23° N
22°30′ N
22° N
-40 -30 -20 -10 0 10 20 30 40
-8 -6 -4 -2 0 2 4 6 8

(d) October

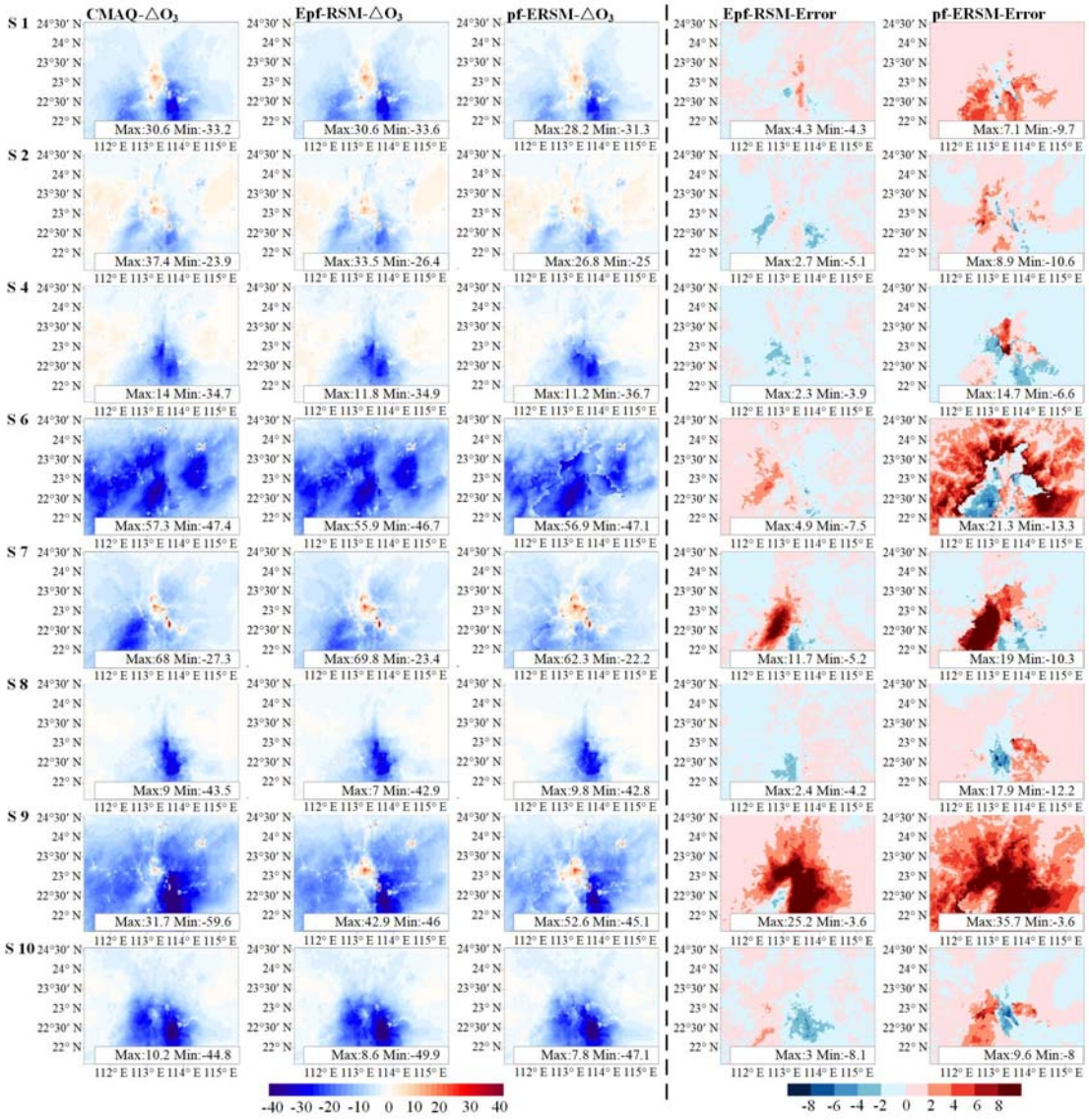
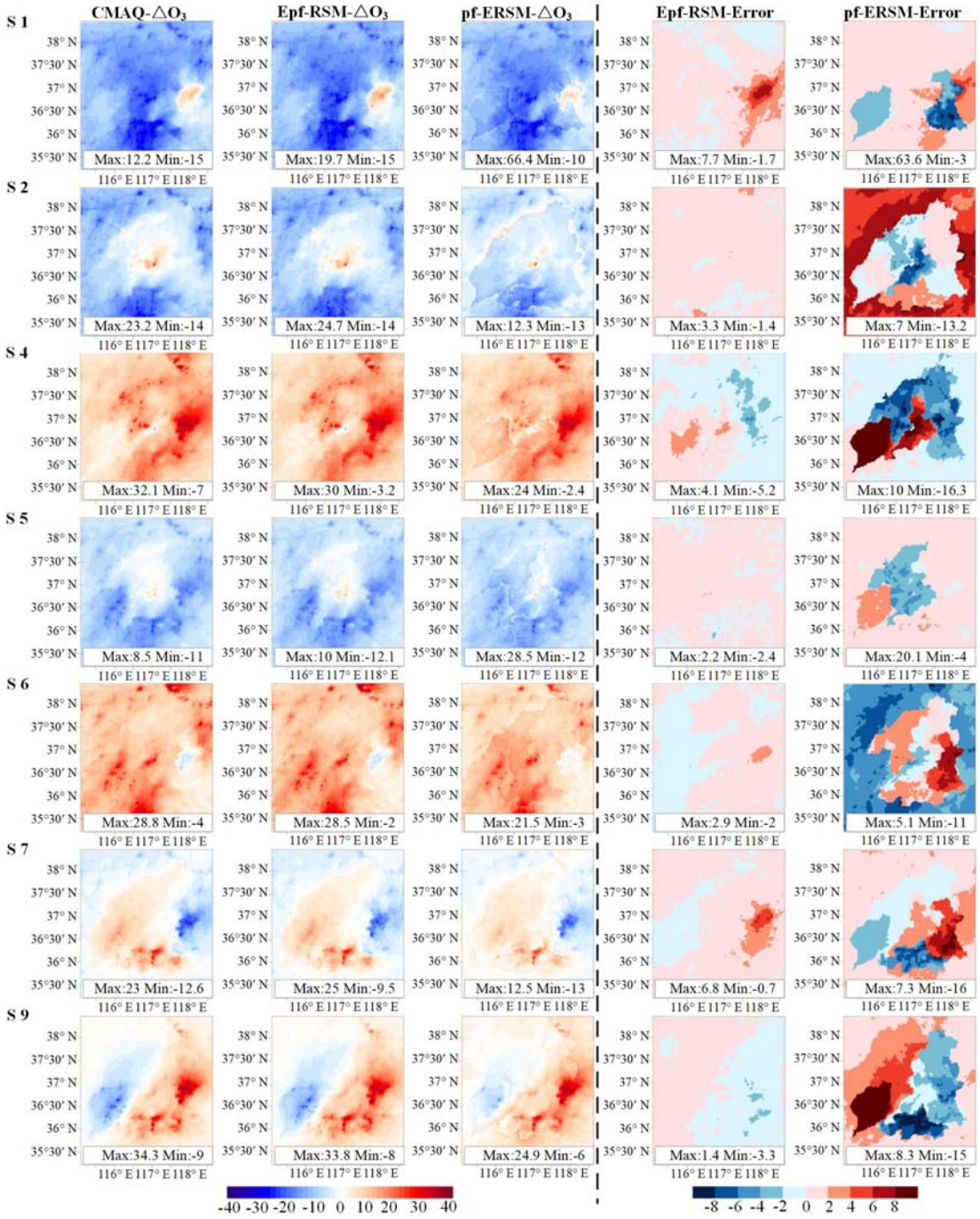


Fig. S10 Spatial distribution of CMAQ-simulated, Epf-RSM-predicted, pf-ERSM-predicted O_3 responses, and their errors under out-of-sample scenarios in the PRD (monthly averaged daily 1 h maxima O_3; unit: $\mu g/m^3$).

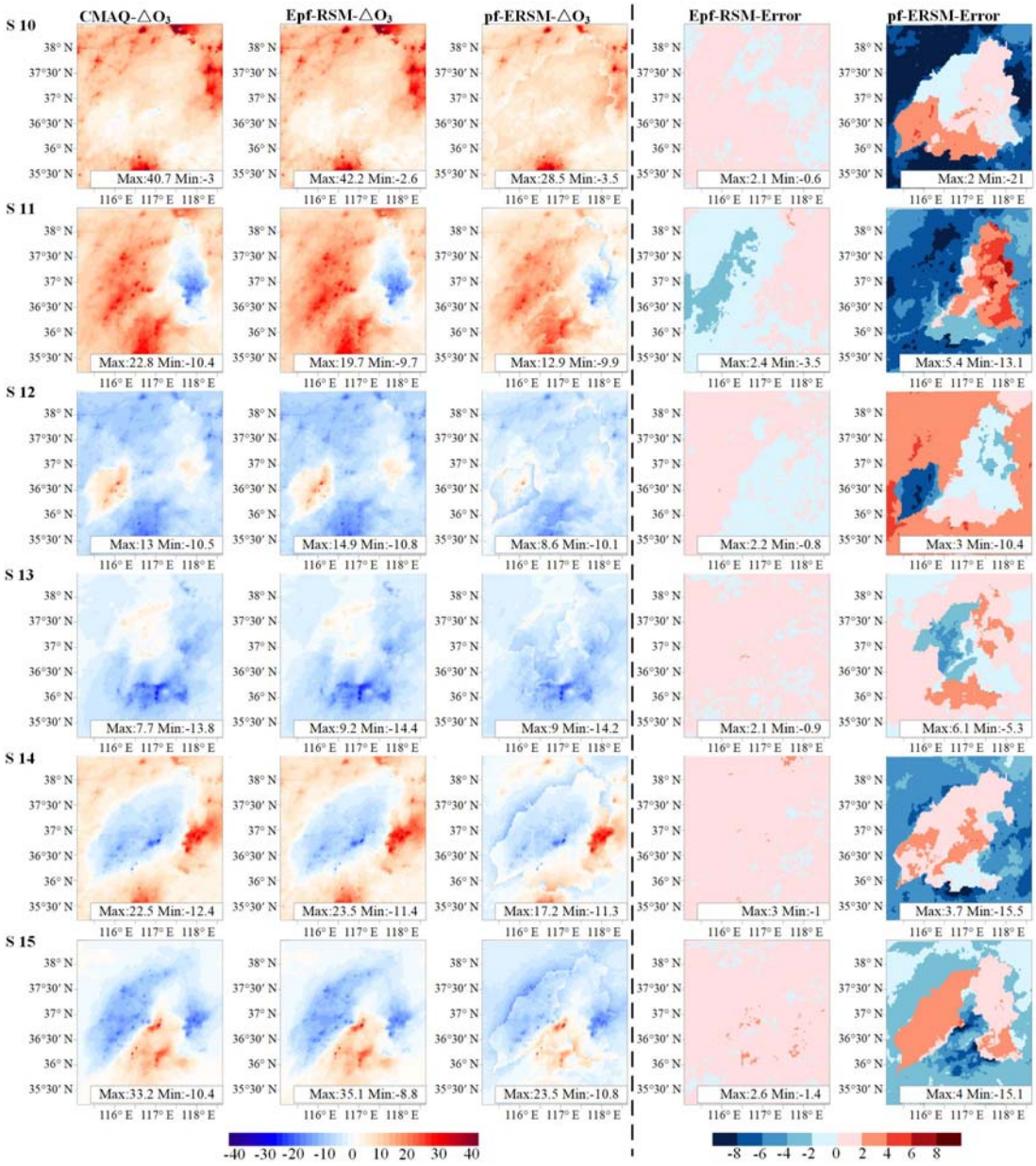
140

(a) January



141

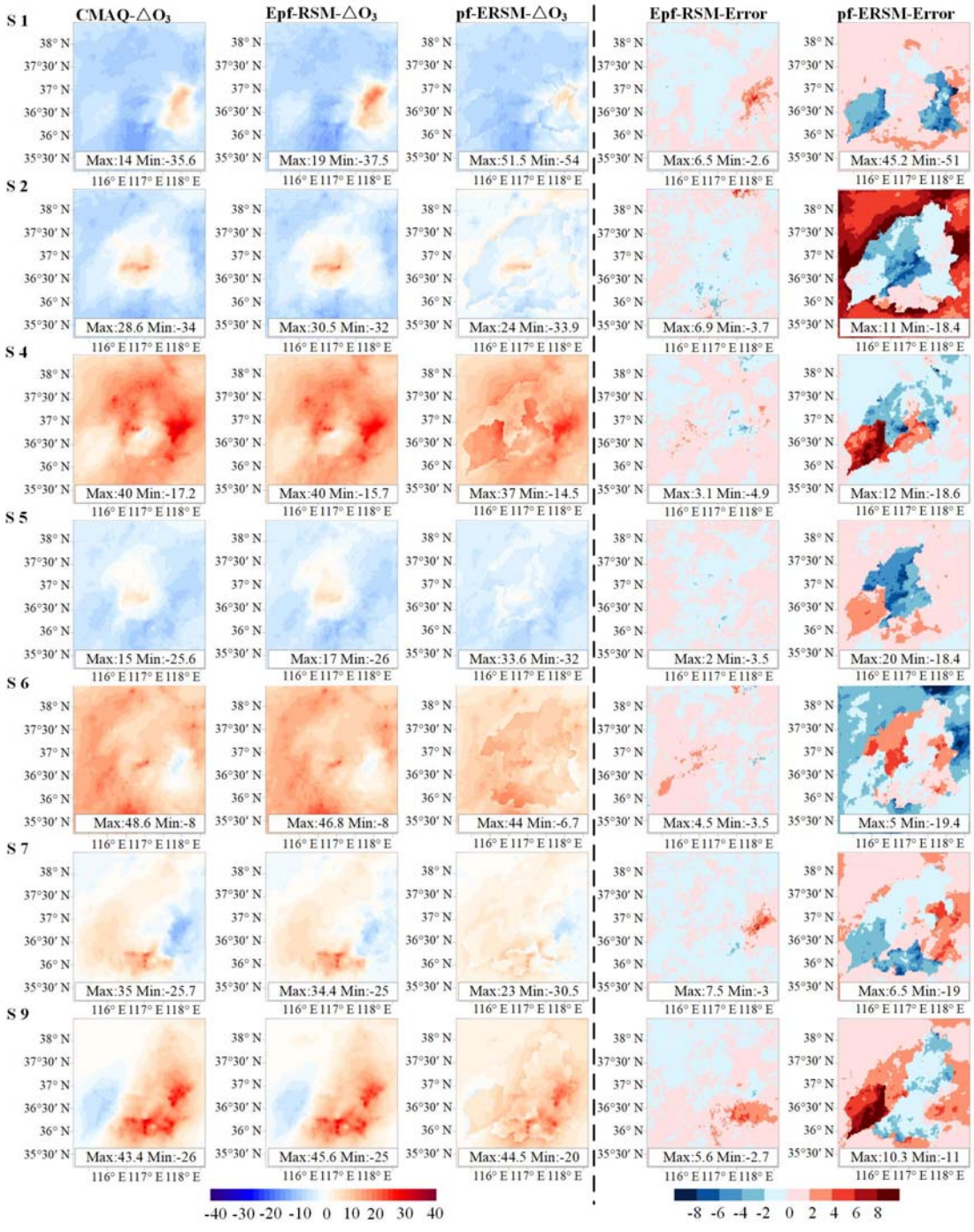
142 **(a) January (Cont.)**



143

144

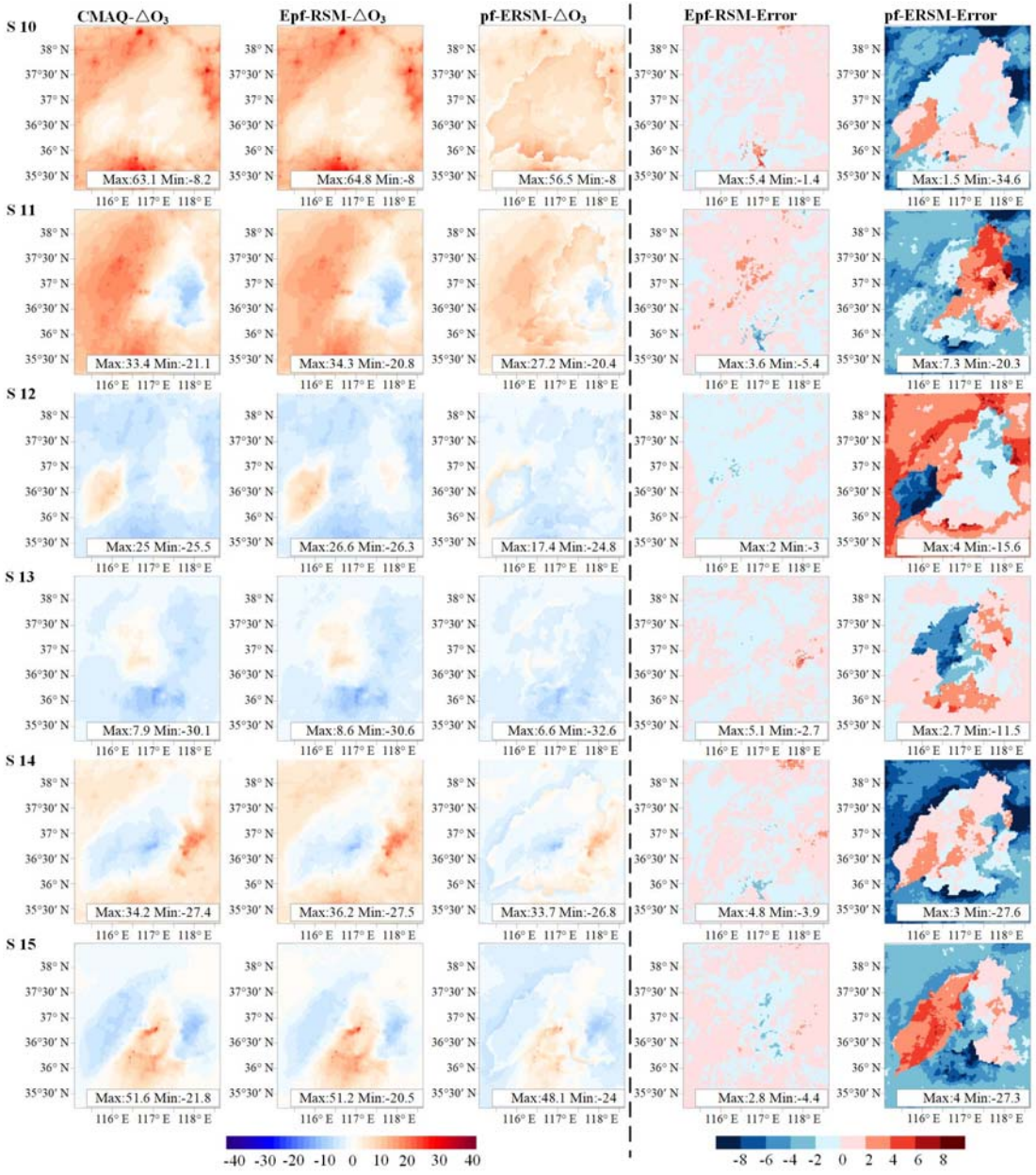
(b) April



145

146

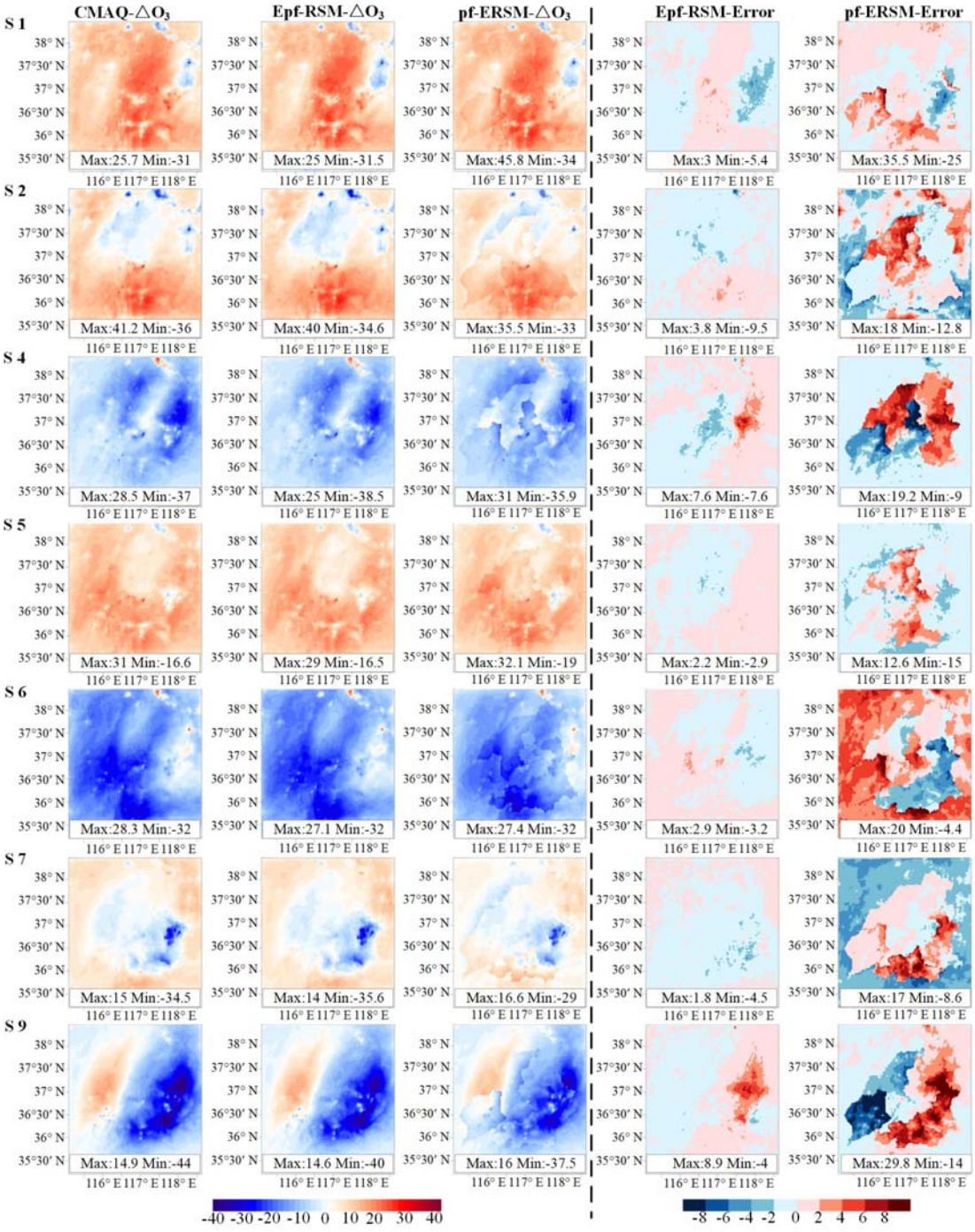
(b) April (Cont.)



147

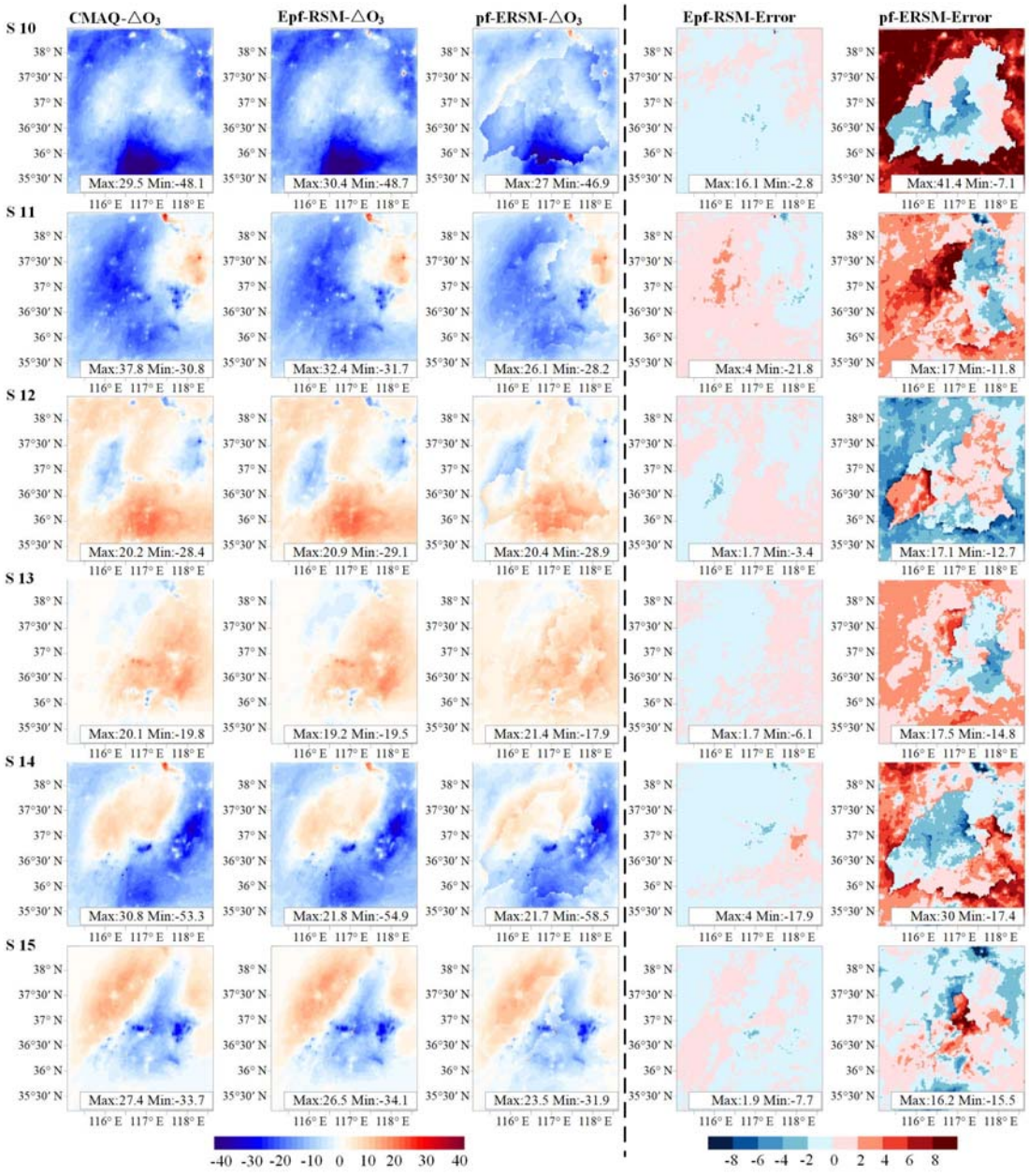
148

(c) July



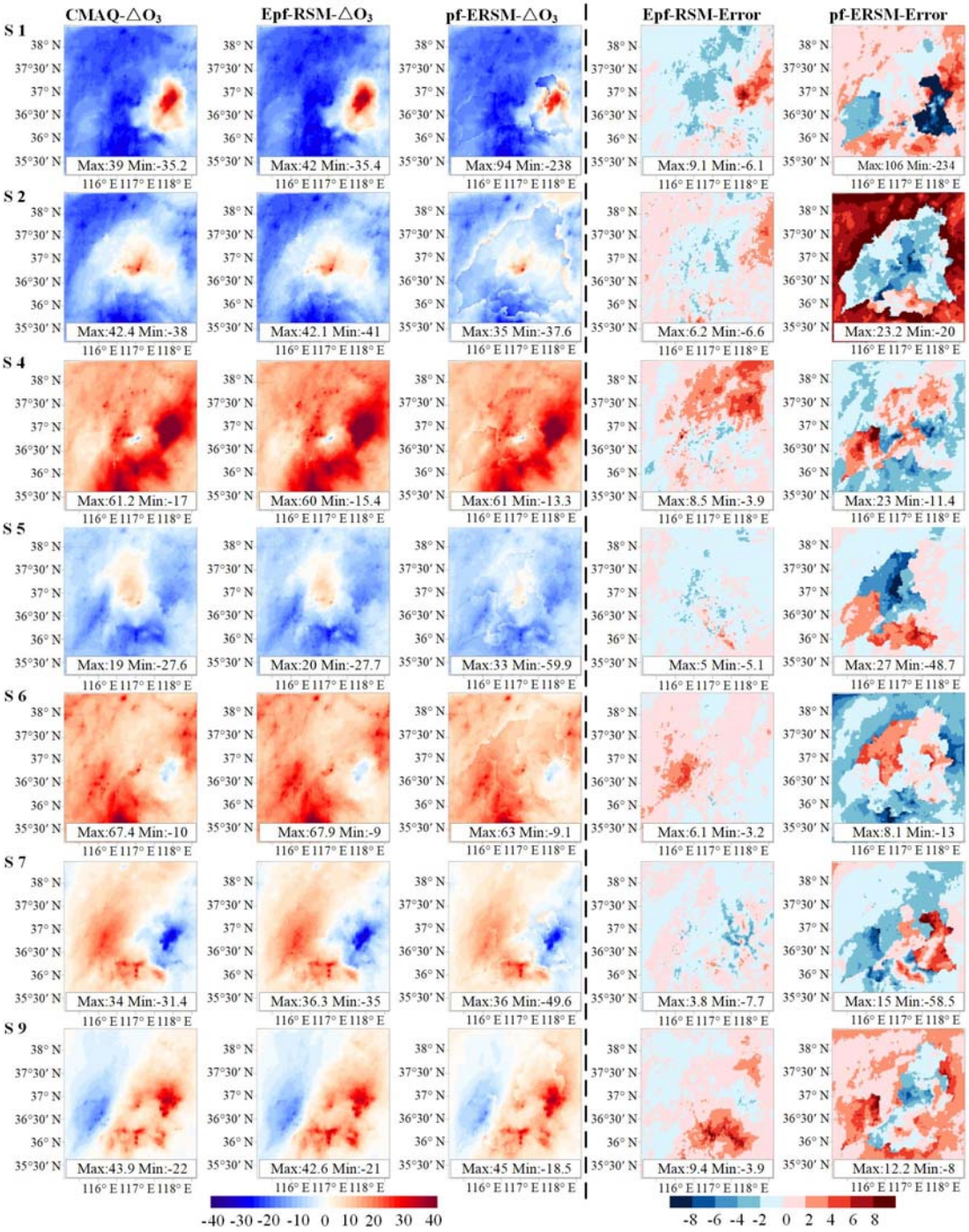
149

(c) July (Cont.)


CMAQ-△O3
Epf-RSM-△O3
pf-ERSM-△O3
Epf-RSM-Error
pf-ERSM-Error
S 10
S 11
S 12
S 13
S 14
S 15
Max:29.5 Min:-48.1
Max:30.4 Min:-48.7
Max:27 Min:-46.9
Max:16.1 Min:-2.8
Max:41.4 Min:-7.1
Max:37.8 Min:-30.8
Max:32.4 Min:-31.7
Max:26.1 Min:-28.2
Max:4 Min:-21.8
Max:17 Min:-11.8
Max:20.2 Min:-28.4
Max:20.9 Min:-29.1
Max:20.4 Min:-28.9
Max:1.7 Min:-3.4
Max:17.1 Min:-12.7
Max:20.1 Min:-19.8
Max:19.2 Min:-19.5
Max:21.4 Min:-17.9
Max:1.7 Min:-6.1
Max:17.5 Min:-14.8
Max:30.8 Min:-53.3
Max:21.8 Min:-54.9
Max:21.7 Min:-58.5
Max:4 Min:-17.9
Max:30 Min:-17.4
Max:27.4 Min:-33.7
Max:26.5 Min:-34.1
Max:23.5 Min:-31.9
Max:1.9 Min:-7.7
Max:16.2 Min:-15.5
38° N
37°30′ N
37° N
36°30′ N
36° N
35°30′ N
116° E 117° E 118° E
-40 -30 -20 -10 0 10 20 30 40
-8 -6 -4 -2 0 2 4 6 8

152

(d) October



153

(d) October (Cont.)

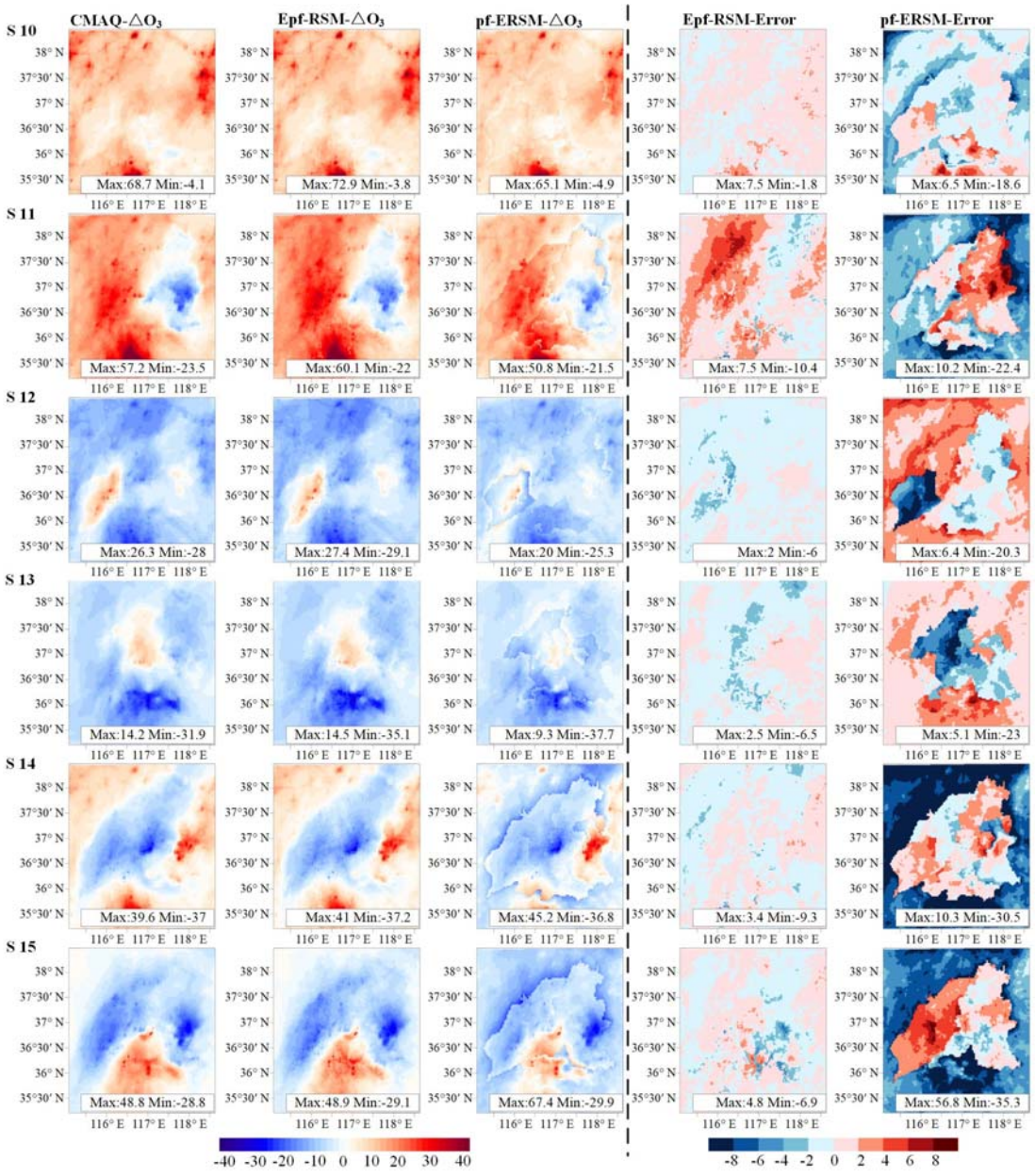


Fig. S11 Spatial distribution of CMAQ-simulated, Epf-RSM-predicted, pf-ERSM-predicted O_3 responses, and their errors under out-of-sample scenarios in the NWShD (monthly averaged daily 1 h maxima O_3; unit: μg/m^3).

159 **Table S1** The strengths and limitations of CMAQ and continually improved versions of RSM
160 systems for conducting sensitivity analysis.

Model	Statistical Techniques	Strength	Limitation
CMAQ [a] (Xing et al., 2011)	-	A powerful regulatory tool for comparing the efficacy of various emissions control strategies and policy decisions.	Conducting the sensitivity analysis requires a large number of control scenario runs which is inefficient using a brute force method.
RSM [b] (Xing et al., 2011)	Using Maximum Likelihood Estimation Experimental Best Linear Unbiased Predictors to approximate model functions of CMAQ.	The sensitivity analysis can be easily and quickly done using RSM and no extra CMAQ simulations are needed.	The required number of scenarios increases significantly with the increase in the number of control variables (increasing from tens of thousands of scenarios for 15 variables to hundreds of thousands of scenarios for 25 variables).
ERSM [c] (Zhao et al., 2015; Xing et al., 2017)	Combining multiple single-region RSM systems and a total-region RSM system mathematically.	Adding many regions to the ERSM system will not tremendously increase the number of scenarios.	Establishing an ERSM system still requires thousands of simulations with CMAQ.
Pf-ERSM [d] (Xing et al., 2018; Fang et al., 2020)	Using a series of polynomial functions to represent the response of O_3 concentrations to precursor emissions.	Compared to the traditional ERSM, the required number of scenarios to build pf-ERSM is reduced by nearly two thirds.	The high-order (fifth-order) polynomial functions and redundant computation process of the ERSM system increase the computational time.

161 Note: [a] CMAQ-Community Multi-scale Air Quality; [b] RSM-Response Surface Model; [c] ERSM-
162 Extended Response Surface Model; [d] pf-ERSM-polynomial function based Extended Response
163 Surface Model.

Table S2 Definitions of the statistical measures for model performance evaluation.

Static Indexes	Definition	Notes
Correlation coefficient (R)	$\frac{\sum_{t=1}^{NG}\left(M_t-\bar{M}\right)\left(O_t-\bar{O}\right)}{\sqrt{\sum_{t=1}^{NG}\left(M_t-\bar{M}\right)^2\sum_{t=1}^{NG}\left(O_t-\bar{O}\right)^2}}$	Unitless, $-1 \le R \le +1$ R = 1 is perfect correlation R = 0 is totally uncorrelated
Normalized Mean Bias (NMB)	$\frac{\sum_{t=1}^{NG}\left(M_t-O_t\right)}{\sum_{t=1}^{NG}O_t}\cdot 100\%$	$-\infty \le NMB \le +\infty$
Mean normalized error (MeanNE)	$\frac{1}{NG}\sum_{t=1}^{NG}\frac{\left\|M_t-O_t\right\|}{O_t}\cdot 100\%$	$0\% \le MeanNE \le +\infty$
Maximal normalized error (MaxNE)	$\max\left(\frac{\left\|M_t-O_t\right\|}{O_t}\right)\cdot 100\%$	$0\% \le MaxNE \le +\infty$

Note: M_t is ERSM-predicted change value of ozone concentration in grid t ($\Delta Conc_t$); O_t is the CMAQ-predicted $\Delta Conc_t$; $\bar{M}$ is the average of ERSM-predicted $\Delta Conc$ in the whole grids; $\bar{O}$ is the average of CMAQ-predicted $\Delta Conc$ in the whole grids; and *NG* is the number of grids.

Table S3 Validation of temperature, wind speed, and relative humidity at monitoring sites.

Site	Month	Temperature (℃)		Wind speed (m s^{-1})		Relative humidity (%)	
		NMB (%)	R	NMB (%)	R	NMB (%)	R
Sugang	Jan.	-3.34	0.9023	101.02	0.6813	7.07	0.8895
	Apr.	1.22	0.9257	-22.41	0.4843	6.20	0.9078
	Jul.	-0.13	0.7647	7.36	0.5116	2.99	0.7436
	Oct.	-3.69	0.8469	-7.61	0.7623	4.29	0.8633
Shengzhongzichangku	Jan.	-95.88	0.8250	120.82	0.6295	4.90	0.8022
	Apr.	-11.83	0.8907	154.95	0.5357	0.36	0.8621
	Jul.	-1.64	0.7989	100.34	0.4438	6.56	0.7731
	Oct.	-4.14	0.9120	96.14	0.5082	-2.99	0.8935

Table S4 Terms of functions in different polynomial functions.

	NO_X[c]	SO_2[c]	NH_3[c]	AVOC[c]	NO.	Terms in the Polynomial Functions
1[a]	1	1	1	1	4	
2	2	1	1	1	6	
3	1	1	1	2	6	
4	3	1	1	1	8	
5	1	1	1	3	8	
6	3	1	1	2	11	
7	2	1	1	3	11	
8	3	1	1	3	12	
9	4	1	1	1	10	
10	4	1	1	2	13	
11	4	1	1	3	15	
12	5	1	1	1	12	
13	5	1	1	2	16	
14[b]	5	1	1	3	19	

Note: [a] polynomial function 1 is the default polynomial function in HCAM; [b] polynomial function 14 is the final polynomial function in HCAM; [c] maximum exponents of precursors. The circles of different colors represent different variables (orange is E_{NOx}, red is E_{SO2}, green is E_{NH3}, and blue is E_{AVOC}); and the numbers in the circles are exponents of variables.

173 **Table S5** Control matrix of CMAQ scenarios for RSM design in the PRD region.

Cases	NO_X	SO_2	NH_3	AVOC
Baseline	1	1	1	1
Zero-out	0	0	0	0
Scenario 1	0.291	1.319	0.472	1.226
Scenario 2	1.463	1.245	1.293	1.158
Scenario 3	0.478	0.186	1.431	0.259
Scenario 4	0.004	0.158	1.393	0.062
Scenario 5	1.129	0.781	0.726	0.967
Scenario 6	1.420	1.365	1.140	0.160
Scenario 7	1.231	1.142	1.465	0.331
Scenario 8	0.809	0.931	0.185	1.409
Scenario 9	0.000	0.132	0.748	1.037
Scenario 10	1.096	0.033	1.103	0.002
Scenario 11	1.055	1.413	0.883	1.465
Scenario 12	1.011	0.911	1.027	0.068
Scenario 13	1.302	1.300	0.043	0.040
Scenario 14	1.389	1.145	1.500	0.229
Scenario 15	0.021	1.369	0.377	1.499
Scenario 16	1.340	0.421	0.904	1.403
Scenario 17	0.724	0.331	1.300	1.102
Scenario 18	0.099	1.484	1.258	0.006
Scenario 19	0.353	1.493	0.000	0.784
Scenario 20	0.217	0.640	0.109	0.316
Scenario 21	0.597	0.710	1.470	0.152
Scenario 22	0.277	0.494	0.158	0.668
Scenario 23	0.085	1.104	1.214	0.953
Scenario 24	1.500	1.461	0.427	1.329
Scenario 25	0.297	0.042	0.594	0.825
Scenario	0.644	0.030	0.572	0.104

Scenario 26				
Scenario 27	0.630	1.057	1.358	0.017
Scenario 28	0.037	0.668	0.310	0.732
Scenario 29	0.073	0.319	1.486	0.523
Scenario 30	0.444	0.269	0.254	0.608
Scenario 31	0.902	0.061	0.036	0.473
Scenario 32	1.490	0.001	0.693	1.476
Scenario 33	1.484	0.012	0.796	1.329
Scenario 34	0.136	0.466	0.008	1.192
Scenario 35	0.900	1.268	0.073	1.490
Scenario 36	0.793	0.604	1.040	0.903
Scenario 37	1.135	0.834	0.200	0.453
Scenario 38	1.461	1.497	0.363	1.275
Scenario 39	0.468	1.005	1.500	1.355
Scenario 40	1.307	0.278	0.022	0.621

174

Table S6 Control matrix of CMAQ scenarios for RSM design in the NWShD region.

Cases	NO_X	SO_2	NH_3	AVOC
Baseline	1	1	1	1
Zero-out	0	0	0	0
Scenario 1	1.466	1.995	0.324	1.205
Scenario 2	0.828	1.619	0.209	1.622
Scenario 3	0.615	1.399	1.157	0.398
Scenario 4	1.187	1.871	1.661	0.897
Scenario 5	0.450	1.461	1.884	0.410
Scenario 6	1.523	1.538	0.427	0.827
Scenario 7	1.336	1.678	0.962	1.031
Scenario 8	0.103	0.922	0.745	1.311
Scenario 9	0.155	0.547	1.045	0.082
Scenario 10	0.530	1.575	0.537	1.373
Scenario 11	0.099	1.918	0.174	0.664
Scenario 12	0.716	1.189	0.851	0.556
Scenario 13	1.658	0.208	0.264	0.634
Scenario 14	1.351	1.301	0.084	0.267
Scenario 15	1.097	0.808	1.935	1.951
Scenario 16	1.777	1.075	0.807	1.946
Scenario 17	0.271	1.836	1.954	1.871
Scenario 18	1.816	1.113	1.418	1.679
Scenario 19	0.364	0.988	0.037	0.335
Scenario 20	1.017	0.852	1.346	0.987
Scenario 21	0.464	0.319	1.358	0.462
Scenario 22	1.131	0.670	1.271	0.000
Scenario 23	0.021	0.191	1.053	0.791
Scenario 24	1.644	1.295	0.397	1.561
Scenario 25	1.747	0.637	1.471	1.490

Scenario 26	1.851	0.757	0.486	1.250
Scenario 27	1.986	1.204	0.659	1.060
Scenario 28	0.870	0.265	1.640	0.729
Scenario 29	0.652	1.414	1.578	1.114
Scenario 30	0.341	1.739	0.780	1.775
Scenario 31	0.940	0.571	0.559	1.810
Scenario 32	1.233	1.759	1.216	1.432
Scenario 33	1.553	0.130	1.840	1.746
Scenario 34	0.230	1.001	1.784	1.520
Scenario 35	1.436	0.096	1.143	0.154
Scenario 36	0.992	0.484	0.616	1.190
Scenario 37	0.580	0.024	1.521	0.928
Scenario 38	1.931	0.447	0.942	0.113
Scenario 39	1.293	0.364	1.743	0.248
Scenario 40	0.778	0.722	0.126	0.503

176 **Table S7** Control matrix of out-of-sample datasets for the PRD.

Region	Precursor	Out-of-sample									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Shunde	NO_X	1.355	0.140	0.980	1.291	0.612	1.182	0.396	0.767	1.428	0.021
	SO_2	0.875	0.482	1.421	0.262	0.906	0.301	0.709	0.110	0.018	1.113
	NH_3	0.505	0.110	0.768	1.371	0.314	1.020	0.901	0.284	1.196	1.425
	AVOC	1.417	0.285	0.447	1.127	0.938	0.709	1.351	0.884	0.681	0.097
Foshan(excluding Shunde)	NO_X	0.561	1.354	0.398	0.038	0.435	0.876	0.141	1.037	0.235	1.295
	SO_2	0.267	1.010	0.059	0.596	0.928	0.140	0.477	0.899	1.474	0.398
	NH_3	0.540	0.279	1.212	0.446	0.730	0.059	1.144	1.481	0.382	0.103
	AVOC	0.162	0.671	0.477	0.715	1.147	0.349	0.950	1.229	0.802	0.017
Guangzhou	NO_X	0.636	0.463	1.338	0.552	0.924	1.270	0.248	1.140	0.313	1.075
	SO_2	0.741	0.178	1.441	1.165	0.591	0.049	1.240	1.356	0.287	0.823
	NH_3	0.931	1.316	1.007	0.468	0.726	0.664	0.034	1.427	0.809	0.516
	AVOC	1.064	1.191	0.468	0.641	0.063	0.732	1.303	1.224	0.566	0.971
Zhongshan	NO_X	1.255	0.054	1.301	0.510	0.972	0.422	1.040	0.267	0.840	0.611
	SO_2	0.704	1.447	0.511	0.939	0.821	0.265	0.192	0.024	1.022	1.305
	NH_3	0.909	0.691	1.074	0.423	0.001	1.234	1.346	0.160	0.743	1.418
	AVOC	0.556	1.056	0.008	0.776	1.339	0.845	0.302	0.999	1.473	0.111
Jiangmen	NO_X	0.010	0.122	1.139	0.937	0.832	0.324	0.276	0.424	1.096	1.225
	SO_2	0.741	0.220	1.258	1.352	1.408	0.126	0.897	0.363	0.031	0.585
	NH_3	1.306	0.907	0.366	0.641	0.415	0.891	1.427	0.172	0.722	1.039
	AVOC	0.527	1.175	0.912	0.843	0.387	0.228	0.467	0.733	1.374	0.029
Dongguan & Shenzhen	NO_X	0.243	1.339	1.126	0.509	0.493	1.464	0.650	1.291	0.090	0.154
	SO_2	0.070	0.205	0.537	1.040	1.112	0.895	1.289	0.753	0.161	0.413
	NH_3	0.484	0.370	1.455	0.921	0.669	1.021	0.162	1.182	0.296	0.589
	AVOC	1.043	0.004	0.758	0.855	0.421	1.298	0.963	0.366	1.484	0.222
Other	NO_X	0.945	0.171	0.063	1.373	1.227	0.641	0.809	1.087	0.356	1.152
	SO_2	0.134	1.016	0.414	0.613	1.312	0.729	0.201	1.456	1.138	0.071
	NH_3	1.329	0.575	0.622	0.231	0.902	0.142	1.262	0.864	0.479	0.763
	AVOC	0.244	0.151	0.314	0.455	1.376	1.200	1.443	0.907	1.084	0.532

177

178 **Table S8** Control matrix of out-of-sample datasets for the NWShD.

Region	Precursor	Out-of-sample														
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Jinan	NO_X	0.389	0.410	0.596	1.787	1.726	0.251	1.149	1.492	0.767	1.349	1.238	0.883	1.033	1.966	0.031
	SO_2	0.709	1.517	1.070	0.527	1.900	0.615	1.699	0.950	1.759	0.872	1.455	1.258	0.375	0.196	0.086
	NH_3	0.436	1.265	1.439	1.558	1.095	1.776	0.260	0.560	1.910	0.941	1.721	0.679	0.065	0.851	0.358
	AVOC	1.620	1.447	0.984	1.190	1.284	0.534	0.864	0.135	1.570	1.947	0.440	0.391	1.767	0.744	0.025
Dezhou	NO_X	0.357	1.869	0.612	1.683	0.038	1.054	0.835	1.382	1.190	0.784	0.136	1.534	0.530	1.326	1.848
	SO_2	1.938	0.059	1.724	0.414	1.534	0.280	0.895	1.831	0.721	1.403	1.018	1.113	1.208	0.184	0.629
	NH_3	0.741	1.996	0.852	1.039	1.083	0.506	0.331	1.593	0.552	1.274	0.194	1.391	0.043	1.684	1.856
	AVOC	0.260	0.302	1.739	1.006	1.328	0.034	1.475	1.932	0.768	1.165	0.832	0.401	1.686	1.366	0.652
Binzhou	NO_X	1.143	0.169	1.587	1.654	0.099	0.607	0.848	0.782	0.340	0.980	1.890	0.492	1.813	1.349	1.240
	SO_2	0.299	1.028	0.799	0.234	0.630	1.550	1.275	1.159	1.828	1.434	0.808	1.619	0.080	0.469	1.925
	NH_3	1.297	1.468	0.418	1.668	0.909	0.262	0.622	0.034	1.114	0.377	1.873	1.394	1.833	0.723	0.945
	AVOC	1.936	0.608	1.059	1.214	1.558	0.699	1.849	1.153	1.681	0.815	0.397	0.421	1.428	0.100	0.202
Liaocheng	NO_X	1.040	0.104	1.788	0.836	1.336	0.226	0.671	0.547	1.975	1.333	0.438	0.284	1.184	1.595	1.654
	SO_2	0.464	1.332	0.141	1.386	1.775	1.617	1.985	0.619	0.101	0.992	1.518	0.893	0.694	1.094	0.296
	NH_3	1.859	1.239	1.712	0.353	0.401	0.118	0.945	1.402	0.686	0.825	1.175	1.967	1.593	0.185	0.563
	AVOC	1.596	0.668	1.022	0.599	0.333	0.910	1.650	1.923	1.367	1.180	1.286	1.847	0.153	0.488	0.097
Taian	NO_X	1.493	1.134	1.617	1.866	0.728	0.837	0.178	0.376	0.060	1.205	0.587	1.359	1.869	0.949	0.505
	SO_2	1.673	0.531	0.273	0.795	0.070	0.233	1.548	0.609	1.223	1.357	0.873	1.016	1.877	1.101	1.768
	NH_3	0.332	0.028	0.420	0.674	1.673	1.548	1.368	1.203	1.797	1.143	0.906	1.024	1.985	0.617	0.223
	AVOC	1.224	1.796	0.163	0.354	1.931	1.584	0.983	0.620	0.889	0.723	1.085	1.440	0.444	0.013	1.707
Zibo	NO_X	1.000	0.415	1.563	0.627	0.333	1.395	1.923	1.145	0.131	0.840	1.799	0.749	1.325	0.219	1.702
	SO_2	1.225	0.614	1.417	1.706	0.393	0.030	0.511	1.478	1.157	0.756	0.906	0.164	1.857	1.873	1.046
	NH_3	0.842	0.781	0.304	1.962	0.465	1.743	1.502	1.100	1.679	1.035	1.264	0.637	1.354	0.178	0.025
	AVOC	1.537	0.892	1.602	1.894	0.975	1.364	0.423	0.706	0.054	1.100	1.267	0.551	1.858	0.138	0.387
Other	NO_X	1.946	0.252	1.518	1.794	0.703	0.296	1.391	1.225	0.856	0.007	0.550	1.666	1.095	0.470	0.956
	SO_2	0.937	1.830	0.567	0.357	1.724	1.089	0.677	1.956	0.026	0.462	1.538	0.157	0.806	1.456	1.239
	NH_3	0.747	0.631	1.201	0.102	0.956	0.180	1.773	1.342	0.308	1.124	1.949	1.649	1.552	0.879	0.484
	AVOC	0.200	0.279	1.555	0.738	1.706	1.257	1.935	0.098	0.589	0.816	1.859	1.008	0.522	1.125	1.395

Table S9 The runtime of building Epf-RSM and pf-ERSM systems (unit: min).

Months	PRD		NWShD	
	Epf-RSM	pf-ERSM	Epf-RSM	pf-ERSM
Jan.	152	233	120	173
Apr.	143	238	99	172
Jul.	130	237	84	174
Oct.	150	238	125	178

Note: the systems were run on the workstation with Intel (R) Xeno (R) CPU 2.60 GHz, 32-core processor, and 124 GB RAM.

Reference

Emery C, Liu Z, Russell A G, Odman M T, Yarwood G, Kumar N (2017). Recommendations on statistics and benchmarks to assess photochemical model performance. Journal of the Air & Waste Management Association, 67(5): 582-598

Fang T, Zhu Y, Jang J, Wang S, Xing J, Chiang P-C, Fan S, You Z, Li J (2020). Real-time source contribution analysis of ambient ozone using an enhanced meta-modeling approach over the Pearl River Delta Region of China. Journal of Environmental Management, 268: 110650

Xing J, Ding D, Wang S X, Dong Z X, Kelly J T, Jang C, Zhu Y, Hao J M (2019). Development and application of observable response indicators for design of an effective ozone and fine-particle pollution control strategy in China. Atmospheric Chemistry and Physics, 19(21): 13627-13646

Xing J, Ding D, Wang S X, Zhao B, Jang C, Wu W J, Zhang F F, Zhu Y, Hao J M (2018). Quantification of the enhanced effectiveness of NOx control from simultaneous reductions of VOC and NH3 for reducing air pollution in the Beijing-Tianjin-Hebei region, China. Atmospheric Chemistry and Physics, 18(11): 7799-7814

Xing J, Wang S X, Jang C, Zhu Y, Hao J M (2011). Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology. Atmospheric Chemistry and Physics, 11(10): 5027-5044

Xing J, Wang S X, Zhao B, Wu W J, Ding D A, Jang C, Zhu Y, Chang X, Wang J D, Zhang F F, Hao J M (2017). Quantifying Nonlinear Multiregional Contributions to Ozone and Fine Particles Using an Updated Response Surface Modeling Technique. Environmental Science & Technology, 51(20): 11788-11798

Zhao B, Wang S X, Xing J, Fu K, Fu J S, Jang C, Zhu Y, Dong X Y, Gao Y, Wu W J, Wang J D, Hao J M (2015). Assessing the nonlinear response of fine particles to precursor emissions: development and application of an extended response surface modeling technique v1.0. Geoscientific Model Development, 8(1): 115-128