Front. Environ. Sci. Eng. 2021, 15(2): 31
https://doi.org/10.1007/s11783-020-1323-0

RESEARCH ARTICLE

Enhancement of the polynomial functions response surface model for real-time analyzing ozone sensitivity

Jiangbo Jin[1], Yun Zhu (✉)[1,2], Jicheng Jang[1], Shuxiao Wang[3], Jia Xing[3], Pen-Chi Chiang[4,5], Shaojia Fan[2], Shicheng Long[1]

1 Guangdong Provincial Key Laboratory of Atmospheric Environment and Pollution Control, College of Environment and Energy, South China University of Technology, Guangzhou Higher Education Mega Center, Guangzhou 510006, China
2 Southern Marine Science and Engineering Guangdong Laboratory, Sun Yat-Sen University, Zhuhai 519000, China
3 State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
4 Graduate Institute of Environmental Engineering, Taiwan University, Taipei 10673, China
5 Carbon Cycle Research Center, Taiwan University, Taipei 10672, China

HIGHLIGHTS

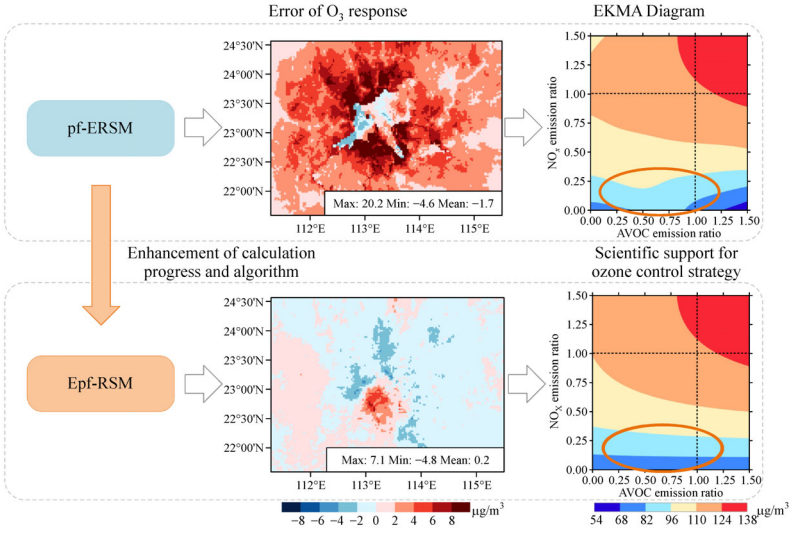
- The calculation process and algorithm of response surface model (RSM) were enhanced.
- The prediction errors of RSM in the margin and transition areas were greatly reduced.
- The enhanced RSM was able to analyze O_3-NO_x-VOC sensitivity in real-time.
- The O_3 formations were mainly sensitive to VOC, for the two case study regions.

ARTICLE INFO

Article history:
Received 14 February 2020
Revised 12 July 2020
Accepted 19 July 2020
Available online 10 September 2020

Keywords:
Response surface model
Hill-climbing algorithm
Ozone pollution
Precursor emissions
Control strategy

GRAPHIC ABSTRACT



ABSTRACT

Quantification of the nonlinearities between ambient ozone (O_3) and the emissions of nitrogen oxides (NO_x) and volatile organic compound (VOC) is a prerequisite for an effective O_3 control strategy. An Enhanced polynomial functions Response Surface Model (Epf-RSM) with the capability to analyze O_3-NO_x-VOC sensitivities in real time was developed by integrating the hill-climbing adaptive method into the optimized Extended Response Surface Model (ERSM) system. The Epf-RSM could single out the best suited polynomial function for each grid cell to quantify the responses of O_3 concentrations to precursor emission changes. Several comparisons between Epf-RSM and pf-ERSM (polynomial functions based ERSM) were performed using out-of-sample validation, together with comparisons of the spatial distribution and the Empirical Kinetic Modeling Approach diagrams. The comparison results showed that Epf-RSM effectively addressed the drawbacks of pf-ERSM with respect to over-fitting in the margin areas and high biases in the transition areas. The O_3 concentrations predicted by Epf-RSM agreed well with Community Multi-scale Air Quality simulation results. The case study results in the Pearl River Delta and the north-western area of the Shandong province indicated that the O_3 formations in the central areas of both the regions were more sensitive to anthropogenic VOC in January, April, and October, while more NO_x-sensitive in July.

✉ Corresponding author
E-mail: zhuyun@scut.edu.cn

1 Introduction

Ambient tropospheric ozone (O_3) has adverse effects on human health (Seltzer et al., 2018) and terrestrial vegetation (Yue et al., 2017). Recently, the treatment of soaring ground-level O_3 pollution has been receiving substantial attention in China after the concentration of $PM_{2.5}$ was effectively reduced (Sun et al., 2019). O_3 is a secondary pollutant that is formed in the troposphere by the reaction between nitrogen oxides (NO_x) and volatile organic compound (VOC) in the presence of sunlight. The complexity of the O_3 formation mechanism is shown through the sensitivity variations of O_3-NO_x-VOC in different spatial and temporal scales (Wang et al., 2017). Precisely identifying the photochemical regime of O_3 formation (i.e., NO_x-sensitive and VOC-sensitive) will help to determine which emissions should be focused on to efficiently abate O_3 pollution in a target region (Liao et al., 2007; Ye et al., 2016; Wang et al., 2019).

The isopleths of the O_3 concentrations versus NO_x and VOC emissions, known as the Empirical Kinetic Modeling Approach (EKMA) diagrams, are widely used to illustrate the O_3-NO_x-VOC sensitivity (Zhong et al., 2014; Sharma et al., 2016). Generally, the EKMA diagrams are obtained by interpolating dozens or hundreds of control scenarios with different combinations of precursor emission changes, which are simulated with observation-based models (OBMs) or emission-based models (EMBs) (Su et al., 2018; Chen et al., 2019; Zhang et al., 2020). OBMs are used to calculate the ozone production rate and relative incremental reactivity of precursors to generate O_3 isopleths plots (Tan et al., 2018a; Tan et al., 2018b). The OBM-based EKMA diagrams can identify the photochemical regime of O_3 formation, but they are inconvenient for determining the precursor reductions required to achieve a specific O_3 concentration (Cardelino and Chameides, 1995; 2000). Contrastingly, the EBM-based EKMA diagrams, which are drawn based on EBM-simulated nonlinear responses of O_3 concentration to precursor emission changes, allow qualitative and quantitative analysis of O_3 sensitivity (Ou et al., 2016; Collet et al., 2018). Nevertheless, the heavy computing burden involved in simulating control scenarios with EBMs presents a great challenge to provide real-time sensitivity analysis for O_3. The Response Surface Model (RSM) is a promising model to address this challenge. The RSM has been developed using statistical regression structures to approximate EBM functions (Xing et al., 2011). The development of RSM methodologies has allowed for the continuous improvement of the efficiency of the prediction systems owing to the reduced number of scenarios required to generate the RSM (see Table S1). Furthermore, the methodology of O_3 sensitivity analysis in the RSM has also been kept updated. The conventional RSM can efficiently provide a real-time response of O_3 concentration to precursor emission perturbations. However, the complex statistical technique involved in the conventional RSM creates difficulty in the investigation of the nonlinearity of the predicted system (Xing et al., 2018). Therefore, Xing et al. (2018) developed a series of polynomial functions to represent the response of O_3 concentrations to precursor emissions. The polynomial function based RSM (pf-RSM) can directly calculate the peak ratio (PR) to quickly identify the O_3 formation regime in the baseline condition. Furthermore, through combining multiple single-region pf-RSM systems and a total-region pf-RSM system mathematically, a polynomial function based Extended Response Surface Model (pf-ERSM) can be developed from the pf-RSM (Fang et al., 2020). The pf-ERSM exhibits good performance in quantifying the nonlinear relationship between O_3 concentrations and precursor emissions in multiple regions but there still exist two drawbacks: 1) over-fitting in the margin areas (Xing et al., 2011) where emissions are cut down more than 60%; and 2) existing relatively high biases in the transition areas which are the significant change areas for the change ratios of pollutant concentration between adjacent regions. These factors affect more or less the results of the sensitivity analysis for O_3 using the pf-ERSM.

Aiming to address the above two drawbacks, we enhanced the pf-ERSM by innovatively integrating the hill-climbing algorithm with it and optimizing the calculation process of ERSM. The enhanced pf-ERSM was referred to as Enhanced polynomial functions Response Surface Model (Epf-RSM) in the remainder of this paper. Two pilot case studies in China, one in the Pearl River Delta (PRD) and the other in the north-western area of the Shandong province (NWShD), were conducted to compare the performance of Epf-RSM and pf-ERSM for catching the nonlinearity of O_3 response to precursor emissions in real-time.

2 Materials and methods

The operation process for developing and applying Epf-RSM for O_3 (Epf-RSM_O_3) was shown in Fig. 1. First, the multi-dimensional experiment design for the control matrix was developed, with the control factors selected based on the emission control strategy targets (i.e., precursors/species, regions). Second, the Weather Research and Forecasting (WRF, version 3.9.1) model was used to simulate the meteorological condition in the base year, to drive simulations with the Community Multi-scale Air Quality (CMAQ, version 5.2) under various simulation scenarios from the control matrix. Third, the best suited polynomial functions were fitted based on the CMAQ simulation results, where a hill-climbing adaptive method was used. Thereafter, the best suited polynomial functions predicted the nonlinear response of O_3 concentration to multi-region precursor emission changes in the

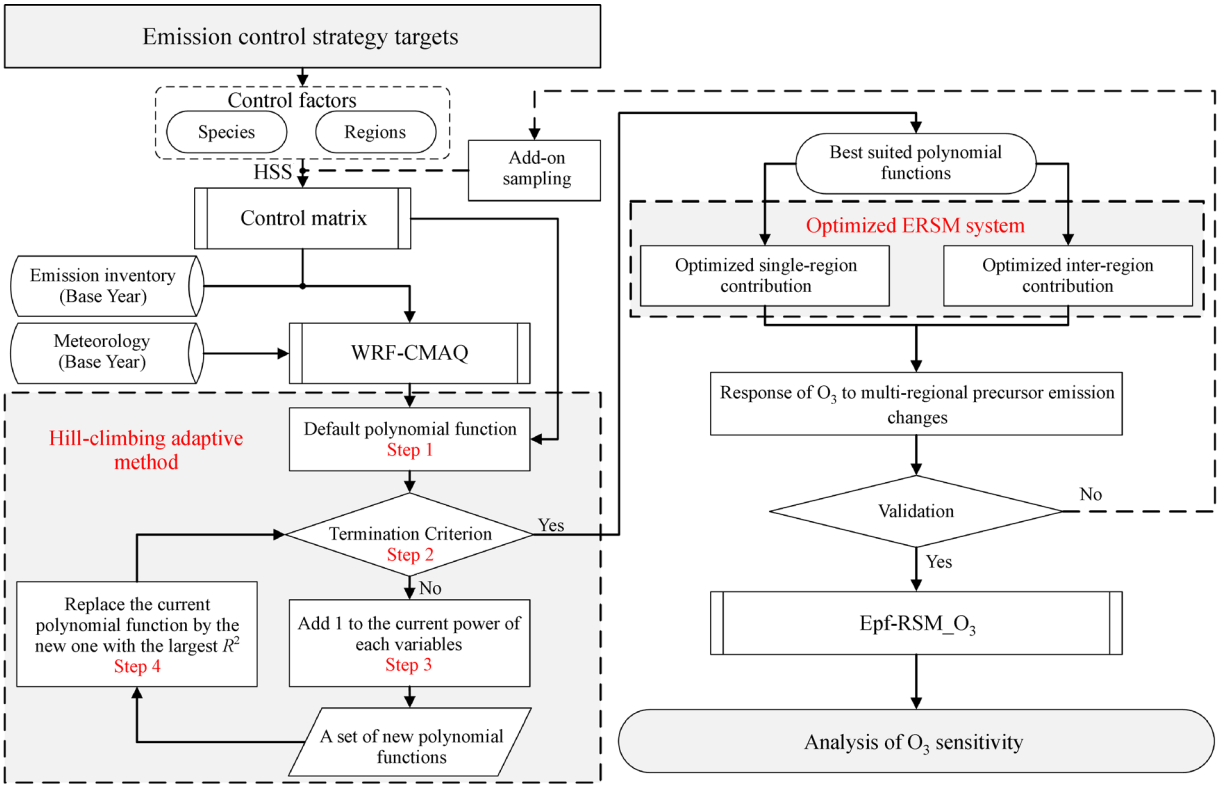


Fig. 1 Key operation process for the development and application of an Enhanced polynomial functions Response Surface Model (Epf-RSM) (Note: HSS: the Hammersley quasi-random Sequence Sample; WRF-CMAQ: the Weather Research and Forecasting coupled with Community Multi-scale Air Quality).

optimized ERSM system. Finally, after the robustness of the developed Epf-RSM_O_3 was validated, the system was ready to perform a series of real-time analyses.

2.1 WRF-CMAQ modeling domain and configuration

The WRF-CMAQ was performed using three nested modeling domains with grid resolutions of 27 km (outermost domain, d01), 9 km (middle domain, d02), and 3 km (innermost domain, d03). The d01 domain that covered approximately 4725 km × 3348 km on a Lambert conformal projection was centered at 31.1°N, 112.3°E for the PRD case study. The d02 domain that covered approximately 1197 km × 1197 km was centered at 22.8°N, 117.8°E, while the d03 that covered the entire PRD region (336 km × 444 km) with 112 × 148 grid cells was centered at 21.1°N, 117.0°E. The d03 was divided into seven regions, including Shunde (SD), Foshan excluding Shunde (FS), Guangzhou (GZ), Zhongshan (ZS), Jiangmen (JM), Dongguan & Shenzhen (DG&SZ), and all other regions in the PRD (OTH). The north-western area of the Shandong province (NWShD) case study was covered by a d01 domain of approximately 1701 km × 1970 km centered at 34.3°N, 115.4°E. The d02 domain that covered approximately 819 km × 819 km was centered at 36.9°N, 121.4°E, while the d03 that covered the NWShD (363 km × 327 km) with 121 × 109 grid cells was centered at 36.5°N, 119.9°E. The NWShD was also divided into seven regions, including Jinan (JN), Dezhou (DZ), Binzhou (BZ), Liaocheng (LC), Taian (TA), Zibo (ZB), and all other regions in the NWShD (OTH). Additionally, the average of data values from the national-controlled air-monitoring sites in the cities or the regions was chosen to represent the air quality of the individual city or region when conducting sensitivity analysis. The spatial distribution of air-monitoring sites was shown in Fig. S1.

The WRF model provided the meteorological input data files for CMAQ. The initial and boundary conditions for d01 were based on default profiles in CMAQ, and those for d02 and d03 were extracted from simulation results on the d01 and d02, respectively. The simulation periods were January, April, July, and October, which represented winter, spring, summer, and autumn respectively. These simulation periods were recorded in 2015 for the PRD and 2017 for the NWShD, respectively. The Tsinghua University provided the emission inventories for d01 and d02 for two case studies (Ma et al., 2017; Ding et al., 2019). Additionally, the emission inventory of the PRD (d03) was established by the collaborative research team from Tsinghua University and the South China University of Technology, and the emission inventory of the NWShD (d03) was developed by the Jinan Ecological Environment

Bureau. The performance of WRF-CMAQ modeling system was evaluated as shown in Text S1.

2.2 Hill-climbing adaptive method

The nonlinearity between O_3 concentrations and its precursor emissions is mainly caused by the nonlinear behavior of pollutants in the atmospheric processes (Xue et al., 2014; Pu et al., 2017). Heyes et al. (1996) used a general formula to develop a fitting-based model for simplifying the nonlinearity, and they found that a simplified formula could easily be applied to optimize control strategies (Heyes et al., 1997). In the pf-RSM, Xing et al. (2018) adopted a series of fifth-order polynomial functions (Eq. (1)) to characterize the O_3 responsiveness to precursor emissions.

$$\Delta\mathrm{Conc} = \sum_{i=1}^{n} \mathrm{X}_i \cdot (\mathrm{E_{NO_X}})^{\mathrm{a}_i} \cdot (\mathrm{E_{SO_2}})^{\mathrm{b}_i} \cdot (\mathrm{E_{NH_3}})^{\mathrm{c}_i} \cdot (\mathrm{E_{AVOC}})^{\mathrm{d}_i} \quad (1)$$

Where:ΔConc is the response of O_3 concentration at either an individual grid or aggregated grids in the target region; $\mathrm{E_{NO_X}}$, $\mathrm{E_{SO_2}}$, $\mathrm{E_{NH_3}}$, and $\mathrm{E_{AVOC}}$ represent the ratios of NO_x, SO_2, NH_3, and anthropogenic VOC (AVOC) emission changes, respectively; n is the number of terms in the function; X_i is the coefficient of the ith term; and a_i, b_i, c_i, and d_i are the exponents of $\mathrm{E_{NO_X}}$, $\mathrm{E_{SO_2}}$, $\mathrm{E_{NH_3}}$, and $\mathrm{E_{NH_3}}$ in the ith term, respectively.

However, it was found that not every receptor (single grid cell or region) was suitable to use the high-order (e.g., fifth-order) functions to capture the nonlinear chemistry. For instance, the absolute value of the coefficient of $(\mathrm{E_{NO_X}})^5$ (2.30) was much smaller than that of $(\mathrm{E_{NO_X}})^4$ (38.89) in the polynomial function of Tianjin in July (see Table S4 in Xing et al. (2018)). Due to the $\mathrm{E_{NO_X}}$ variation was usually in the range of [-1, 1] (i.e., from 100% cut down and 100% increase) in the experimental designs (Xing et al., 2011; Xing et al., 2017; Xing et al., 2018), the effect of $(\mathrm{E_{NO_X}})^5$ was so little as to be ignored. For a receptor like Tianjin, the use of a fifth-order polynomial function would result in a waste of computing resources. Moreover, using the high-order functions was more likely to cause the Runge phenomenon (Lin and Sun, 2015), which could result in an excessive distortion of the response surface, especially in the margin areas where emissions were cut down more than 60% (Fig. S4). Therefore, it was highly recommended that each receptor used the suitable polynomial function.

A hill-climbing adaptive method (HCAM) was developed that dynamically matched the nonlinearity in different receptors to find the best suited polynomial function for each receptor. Basically, the hill-climbing method is an iterative algorithm and is generally faster than the stochastic optimization (Guindon and Gascuel, 2003). This method usually begins with an arbitrary solution to the problem, then attempts to improve the solution by making an incremental change to the solution (Lozano et al., 2004). This study initialized the HCAM using a default polynomial function and proceeded to find a relatively better polynomial function by iteration. The maximum exponents of the four variables (i.e., $\mathrm{E_{NO_X}}$, $\mathrm{E_{SO_2}}$, $\mathrm{E_{NH_3}}$, and $\mathrm{E_{AVOC}}$) were 5, 1, 1, and 3, respectively, which were chosen by exponent examination in our previous study (Xing et al., 2018). All the possible polynomial functions were listed in Table S4.

Three indicators were used as the termination criterion for judging whether a polynomial function was the suitable one. First, the relative R^2 surplus ratio ($\partial\, R^2$, Eq. (2)) was developed to measure the surplus improvement potential for the fitting accuracy. Then, the change of Mean Bias (ΔMeanB, Eq. (3)) and change of Max Bias (ΔMaxB, Eq. (4)) were introduced to avoid the response surface from becoming excessively distorted. The thresholds of the three indicators, namely, the threshold of $\partial\, R^2$ (TRS), the threshold of mean bias (TMean), and the threshold of max bias (TMax), were determined as shown in Section 3.1. The iterative process is terminated when any of the following termination signals is received: (1) $\partial\, R^2$ is less than TRS, and ΔMeanB and ΔMaxB are larger than the corresponding thresholds; and (2) $\partial\, R^2$ is less than 0 or either ΔMeanB or ΔMaxB is greater than 0. When signal (1) occurs, it would be impossible to improve the fitting precision effectively; when signal (2) occurs, a continued increase in the exponents of the variables will create greater errors.

$$\delta \mathrm{R}_j^2 = \frac{\Delta \mathrm{R}_j^2}{1-\mathrm{R}_j^2} = \frac{\mathrm{R}_{j+1}^2 - \mathrm{R}_j^2}{1-\mathrm{R}_j^2} \quad (2)$$

$$\Delta\mathrm{MeanB}_j = \mathrm{MeanB}_{j+1} - \mathrm{MeanB}_j \quad (3)$$

$$\Delta\mathrm{MaxB}_j = \mathrm{MaxB}_{j+1} - \mathrm{MaxB}_j \quad (4)$$

Where:
j is the jth fitting polynomial function; and R^2, MeanB, and MaxB are described in Text S2.

The schematic diagram of HCAM was shown in Fig. 2, and the specific steps of HCAM were listed below:

Step 1: modeling the response of a receptor with the default polynomial function;

Step 2: determining whether the polynomial function met the termination criterion. If yes, outputting the result as a suitable polynomial function. Otherwise, using the current polynomial function for the next step by replacing the previous one;

Step 3: adding 1 to the exponent of each variable that did not reach the maximum exponent in the current polynomial function, thereby constructing a new set of polynomial functions;

Step 4: comparing the R^2 of these new polynomial

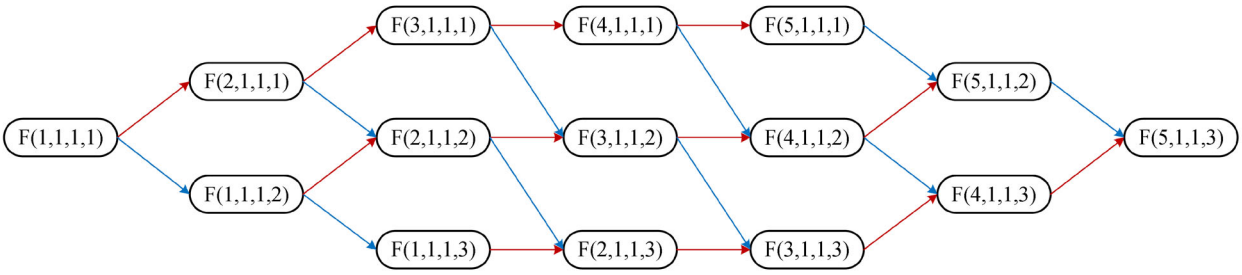


Fig. 2 The schematic diagrams of the hill-climbing adaptive method (HCAM) (Note: F(a, b, c, d) represents a polynomial function; a, b, c, and d represent the nonnegative integer exponents of E_{NO_X}, E_{SO_2}, E_{NH_3}, and E_{AVOC}, respectively; the red and blue lines represent either adding 1 to a or d, respectively).

functions to choose the one with the largest R^2, then returning to Step 2.

2.3 Optimization of ERSM calculation process

In the previous ERSM system (Xing et al., 2017), the multi-region contributions were divided into three components: (1) the impact of O_3 formed in the receptor by transported precursors from the source region (the contribution form CM); (2) the impact of direct regional transport of O_3 from the source region (the contribution form TP); and (3) inter-region contributions (IR) among multiple regions. Moreover, the IR was further subdivided into CM_IR (the inter-region contribution form CM) and TP_IR (the inter-region contribution form TP). The total contribution to O_3 concentration at the receptor grid t ($\Delta Conc_t$) was represented by Eq. S4 in Text S3, and the formulae for the four components were represented respectively by Eqs. S5−S8 in Text S3.

As shown in Eq. S5, the change ratios of precursor concentrations in the receptor associated with the emissions changes in the source region (denoted as "E'_{prec}") were calculated to estimate the contribution form CM. The E'_{prec} across grid cells (denoted as "$E'_{prec,t}$") showed a spatial difference (Fig. S5). However, the spatial difference was neglected because the E'_{prec} was calculated as the average of the aggregated grids in the target region (denoted as "$E'_{prec,T}$"). Therefore, using $E'_{prec,T}$ led to the final pf-ERSM-predicted $\Delta Conc_t$ deviating from the CMAQ-simulated $\Delta Conc_t$, especially in the significant change areas for $E'_{prec,t}$ (i.e., transition areas). To resolve this limitation, the contributions form CM and TP were integrated into the single-region contribution (denoted as "SR") to avoid the use of $E'_{prec,T}$. Correspondingly, the CM_IR and TP_IR were integrated into IR. The new formula for $\Delta Conc_t$ was provided in Eq. (5):

$$\Delta Conc_t = \sum_{S=1}^{N} SR_{S\to t} + IR_t \quad (5)$$

Where:
$\Delta Conc$ is the change of O_3 concentration in the receptor grid t of region T; N is the number of regions; $SR_{S\to t}$ is the contribution from CM and TP to $\Delta Conc$ associated with the emission changes in source region S; and IR_t is the inter-region contributions to $\Delta Conc$.

The O_3 concentrations in the receptor were affected by the emissions in the source region through two pathways, namely, CM and TP. Therefore, the single-region contribution from a source region was the sum of the contributions from CM and TP (see Eq. S6), and the formula of SR was expressed as follows:

$$SR_{S\to t} = RSM\_O3_{S\to t}(E_{prec,S}) \quad (6)$$

Where:
$E_{prec,S}$ is the change ratios of four precursors in source region S; and $RSM\_O3_{S\to t}(E_{prec,S})$ is the single-region RSM system that models the $\Delta Conc$ at receptor grid t to $E_{prec,S}$.

A total-region RSM (denoted as "RSM_{TT}") in which emissions in all regions changed simultaneously was used to explicitly represent the inter-region contributions (IR). The IR was calculated as the difference between the predictions in RSM_{TT} and the sum of SR, as shown below:

$$IR_t = RSMtt\_O3_{S\to t}(\overline{E}_{prec}) - \sum_{S=1}^{N} SR_{S\to t} \quad (7)$$

Where:
$\overline{E}_{prec}$ is the average of [$E_{prec,1}$, …, $E_{prec,T-1}$, $E_{prec,T+1}$,…, $E_{prec,N}$]; and $RSMtt\_O3_{S\to t}(\overline{E}_{prec})$ is the RSM_{TT} system that estimates the response of O_3 concentration at the receptor grid t to $\overline{E}_{prec}$.

2.4 Configuration of Response Surface Modeling systems

The experiment designs for generating Epf-RSM and pf-ERSM were similar to those described in our recent paper (Fang et al., 2020), which were summarized in Table 1. The O_3 response in each receptor to precursor emission changes in individual source region or in all regions (i.e., RSM/RSM_{TT}) was established by 42 training samples, which include one baseline scenario, one zero-out scenario, and 40 control scenarios. The 40 control

Table 1 Scenarios for Response Surface Modeling design

Short name	Objective	Control factor	Number of cases
Baseline	Baseline case	–	1
RSM	RSM method, to create single regional RSM in seven regions separately	Four precursors including NO_x, SO_2, NH_3, AVOC in each of the seven regions	For the PRD: 41 samples for each region in addition to the baseline case (total 288, $41 \times 7 + 1 = 288$), Hamersley quasi-random Sequence Sampling between 0.0 to 1.5 [a], the control matrix is shown in Table S5 For the NWShD: 41 samples for each region in addition to the baseline case (total 288, $41 \times 7 + 1 = 288$), Hamersley quasi-random Sequence Sampling between 0.0 to 2.0 [a], the control matrix is shown in Table S6
RSM_{TT}	Using RSM method, to create multiple regional RSM in seven regions	Four precursors including NO_x, SO_2, NH_3, AVOC in each of the seven regions	For the PRD: 41 samples for seven regions together in addition to the baseline case (total 42), Hamersley quasi-random Sequence Sampling between 0.0 to 1.5 [a], the control matrix is shown in Table S5 For NWShD: 41 samples for seven regions together in addition to the baseline case (total 42), Hamersley quasi-random Sequence Sampling between 0.0 to 2.0 [a], the control matrix is shown in Table S6
OOS	Out-of-sample validation	Four precursors including NO_x, SO_2, NH_3, AVOC in each of the seven regions	For the PRD: 10 samples for seven regions together, Hamersley quasi-random Sequence Sampling between 0.0 to 1.5 [a], the control matrix is shown in Table S7 For the NWShD: 15 samples for seven regions together, Hamersley quasi-random Sequence Sampling between 0.0 to 2.0 [a], the control matrix is shown in Table S8

Note: [a] the baseline = 1.

scenarios were randomly sampled by Hamersley quasi-random Sequence Sampling (HSS) (Hammersley, 1960) with a range of 0 to 1.5 (i.e., from 100% cut down and 50% increase) in the PRD, while with a range of 0 to 2 (i.e., from 100% cut down and 100% increase) in the NWShD, for further testing the applicability of the two ERSM systems over the different range of emission rates. Furthermore, two data sets were selected as out-of-sample scenarios to validate the prediction performance of the two ERSM systems. These included 10 samples for the PRD region (denoted as "OOS_PRD") and 15 samples for the NWShD region (denoted as "OOS_NWShD"). The two out-of-sample data sets were also sampled randomly by HSS.

3 Results and discussion

3.1 Termination criterion for the best suited polynomial functions

The grid cells containing the monitoring site were chosen to represent the air quality of the target regions, for investigating the optimal thresholds with three indicators (i.e., ∂R^2, ΔMeanB, and ΔMaxB). First, the HCAM was used to select the best suited polynomial function for each grid cell with artificially judging whether the iterative process should be terminated. Thereafter, the three indicators of these best suited polynomial functions were analyzed statistically. A strong correlation between R^2 and ∂R^2 of the best suited polynomial functions was shown in Fig. S6. Based on the Slogistic3 model in Origin 2017, a function of TRS (F_{TRS}) was designed to represent the correlation, as follows:

$$F_{TRS}(R^2) = a1 + \frac{a2}{1 + a3 \cdot \exp(-a4 \cdot R^2)} \tag{8}$$

This was followed by determining the four F_{TRS} coefficients (i.e., a1, a2, a3, and a4), TMean, and TMax by controlling variables. The indicator thresholds of termination criterion for SR (denoted as "TC_{SR}") were determined, as shown in Fig. S7. TC_{SR} was expressed as follows:

$$TC_{SR} = \begin{cases} \delta R_j^2 < 0.09 + \dfrac{0.645}{1 + 2.9 \times 10^{73} \times \exp(-168.6 \times R_j^2)} \\ \Delta MeanB j > -0.0004 \\ \Delta MaxB j > -0.001 \end{cases} \quad \text{or } \delta R_j^2 < 0 \text{ or } \Delta MeanB j > 0 \text{ or } \Delta MaxB j > 0 \tag{9}$$

The indicator thresholds of termination criterion for IR (denoted as "TC_{IR}") were determined, as shown in Fig. S8. TC_{IR} was expressed as follows:

$$TC_{SR}=\begin{cases}\delta R_j^2 < 0.09+\dfrac{0.645}{1+2.9\times10^{73}\times\exp(-168.6\times R_j^2)}\\ \Delta MeanBj > -0.00076\\ \\ \Delta MaxBj > -0.003\end{cases}\quad \text{or } \delta R_j^2<0 \text{ or } \Delta MeanBj > 0 \text{ or } \Delta MaxBj>0 \tag{10}$$

Finally, the HCAM with termination criterion was applied to select the best suited polynomial functions for each grid cell in the two case study domains. Fig. S9 depicted the spatial distribution of the E_{NO_X} and E_{AVOC} exponents in the best suited polynomial functions. These distributions showed that the E_{NO_X} and E_{AVOC} exponents in most grid cells were lower than those in previous polynomial functions, especially for July. The runtime of Epf-RSM was approximately 29.8%–51.7% lesser than that of pf-ERSM (Table S9), since the dimensions of the polynomial functions were lowered and the redundant computation process of pf-ERSM was optimized.

3.2 Performance comparison between Epf-RSM and pf-ERSM

3.2.1 Out-of-sample comparison

The out-of-sample comparison was used to compare the CMAQ-simulated and the Epf-RSM- or pf-ERSM-predicted O_3 concentrations for out-of-sample scenarios by the density scatterplots (Fig. 3). These included the density of points that were represented by the data point percent (defined in Text S4) and the best-fit lines (dashed lines) which were described by slope and intercept. Additionally, three statistical indexes were used to evaluate the prediction performance of the two ERSM systems (Fig. 3), namely, Correlation coefficient (R), mean normalized error (MeanNE), and the 95th maximal normalized error (95th MaxNE) (defined in Table S2).

As shown in Fig. 3, the point clustering in the Epf-RSM plots was denser near the one-to-one lines (red lines) than this in the pf-ERSM plots, suggesting that Epf-RSM had better prediction performance than pf-ERSM. Furthermore, the statistical indexes also demonstrated that Epf-RSM improved the prediction performance for O_3 concentrations, particularly for reducing 95th MaxNE. When comparing the pf-ERSM to Epf-RSM for all of the out-of-sample scenarios, the MeanNE and 95th MaxNE decreased, and R increased from 2.8, 10.2%, and 0.9914 to 0.6, 2.3%, and 0.9993, respectively.

3.2.2 Comparison of response spatial distribution

The prediction performance of the two ERSM-systems at different locations and times was compared using two out-of-sample scenarios that were selected from each pilot case study to represent two kinds of emission levels, namely, moderate and strict. The scenarios 5 and 3 were selected from OOS_PRD to represent the moderate and strict control scenarios of the PRD, respectively; the scenarios 8 and 3 were selected from OOS_NWShD to represent the moderate and strict control scenarios of the NWShD, respectively.

Figure 4 depicted the spatial distribution of CMAQ-simulated, Epf-RSM-predicted, and pf-ERSM-predicted responses for O_3, along with the errors of the two ERSMs (i.e., CMAQ-simulated responses minus ERSM-predicted responses) under two control scenarios for the PRD. Figure 5 showed similar results for the NWShD. In the two moderate control scenarios (Figs. 4(a) and 5(a)), the responses predicted when using Epf-RSM produced approximately the same results as the CMAQ simulation results; in contrast, those that were predicted using pf-ERSM deviated from those that were simulated using CMAQ in a few locations. Compared to moderate control scenarios, the errors of the Epf-RSM- and the pf-ERSM-predicted responses (denoted as "$errors_{epf}$" and "$errors_{pf}$," respectively) were higher under the strict control scenarios (see Figs. 4(b) and 5(b)). This was caused by the relatively poor prediction performance at the margin areas (Xing et al., 2011; Xing et al., 2018). Since the use of high-order polynomial functions at the margin areas was more likely to cause the Runge phenomenon, $errors_{pf}$ were prevalent in d03 domains. Epf-RSM introduced the HCAM to avoid the excessive distortion of the response surface, which improved the prediction performance at the margin areas. This resulted in $errors_{epf}$ being significantly less than $errors_{pf}$ under these strict control scenarios.

As shown in Figs. 4, 5, S10 and S11, the boundaries of some regions could be observed in the spatial distribution of $errors_{pf}$. It was because the transition areas that had relatively high $errors_{pf}$ were usually distributed on the border between adjacent regions. The precursor concentrations over the regional borders were mainly affected by the different emission changes in adjacent regions. This resulted in the difference of $E'_{prec,\ t}$ between regional borders and centers. Moreover, the spatial difference of $E'_{prec,\ t}$ was increasing with the gaps in control levels between adjacent regions growing, thus increasing $errors_{pf}$ in the transition areas. For example, since the NWShD had a wider range of emission changes, its gaps in randomly sampled control levels between adjacent regions were larger than those in the PRD, leading to greater $errors_{pf}$ in the NWShD. Epf-RSM as opposed to pf-ERSM, could reproduce the spatial variation of the nonlinear response

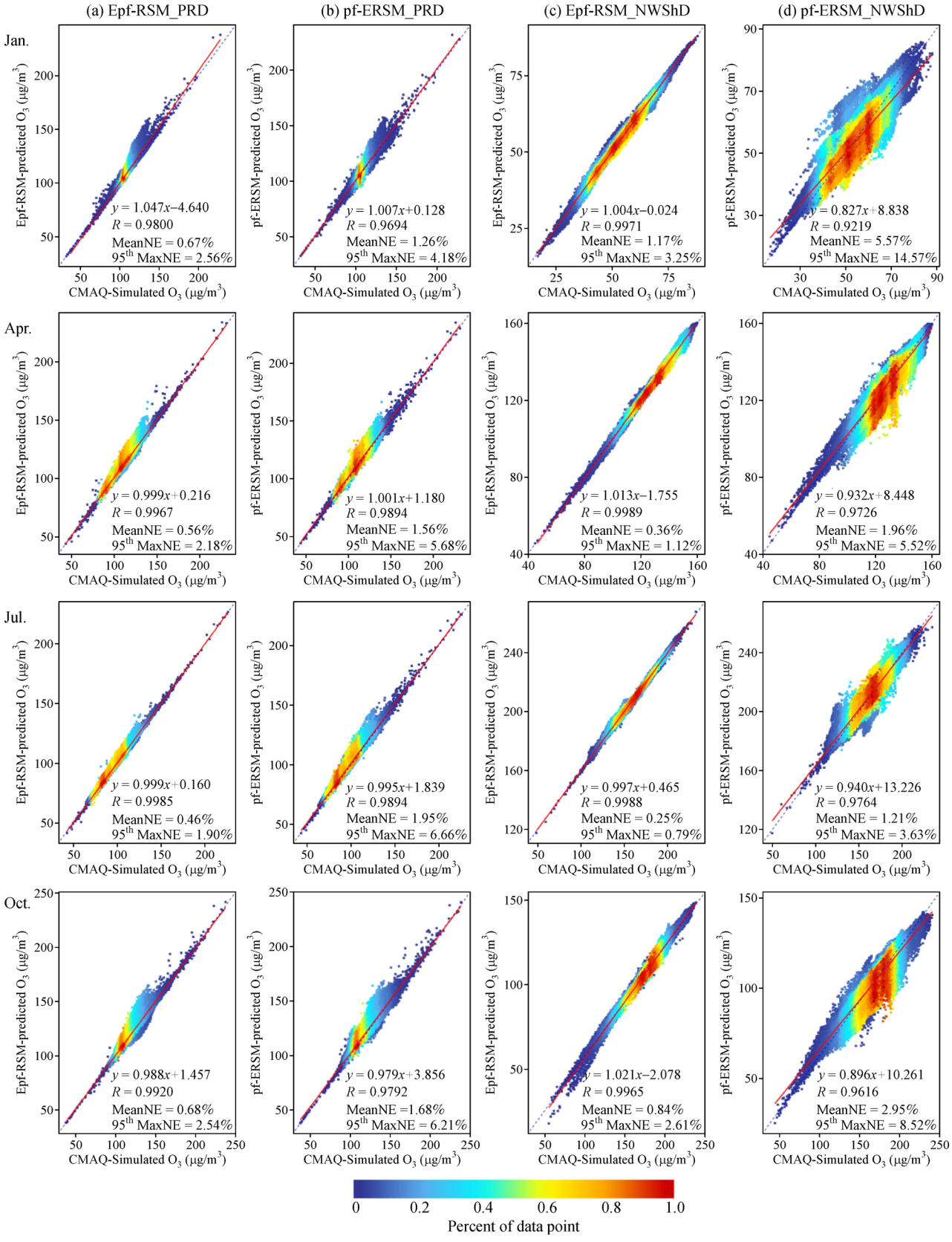


Fig. 3 Comparison of O_3 concentrations simulated by CMAQ with (a) Epf-RSM-predicted in the PRD, (b) pf-ERSM-predicted in the PRD, (c) Epf-RSM-predicted in the NWShD, (d) pf-ERSM-predicted in the NWShD, respectively (monthly averaged daily 1 h maxima O_3; unit: $\mu g/m^3$) (Note: the percent of data point represents the density of points falling at each grid cell on the plot with a resolution of 1 $\mu g/m^3 \times 1$ $\mu g/m^3$. The red lines are the one-to-one lines indicating perfect agreement, and the dashed lines are the best-fit lines described by slope and intercept).

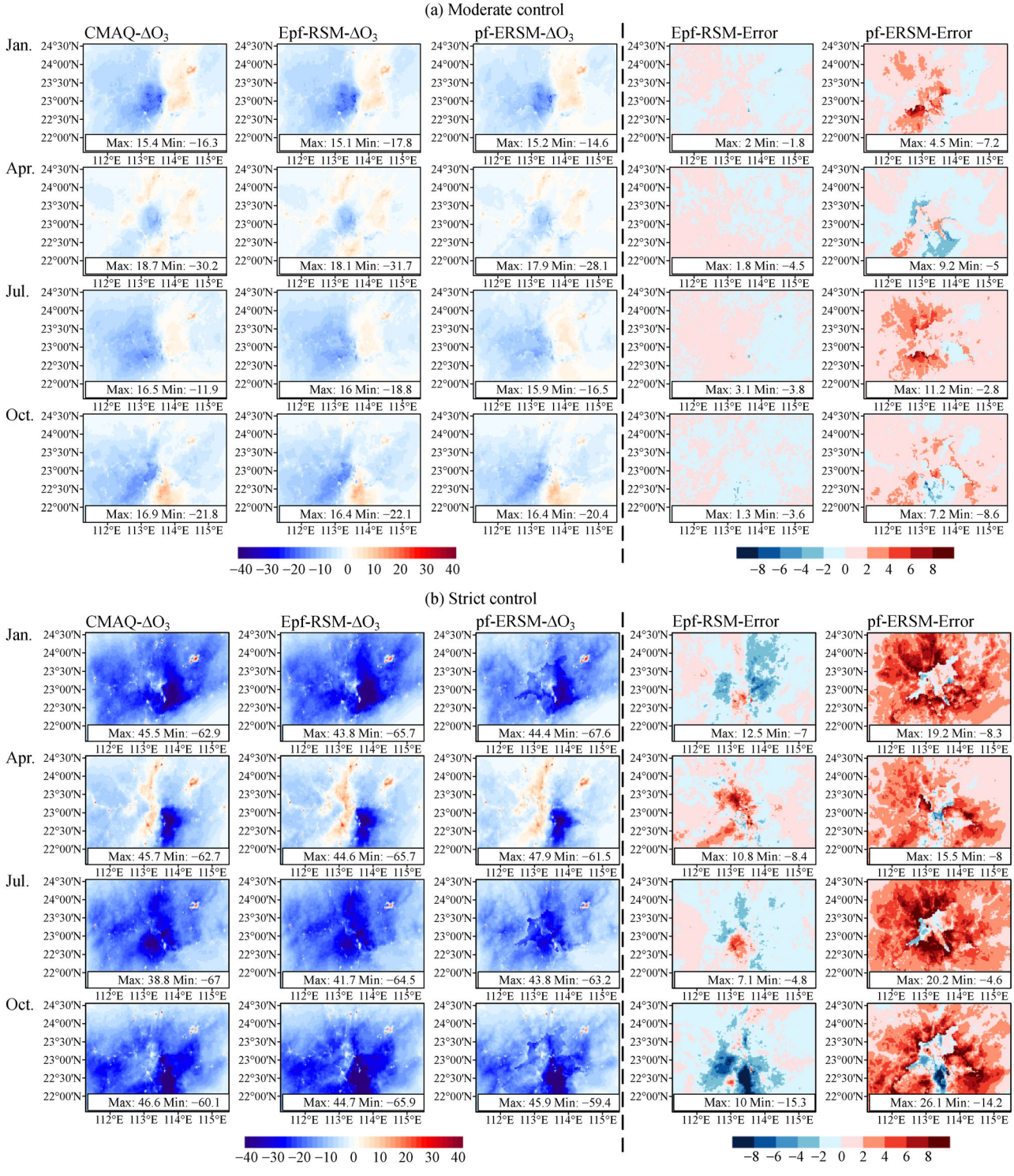


Fig. 4 Spatial distribution of CMAQ-simulated, Epf-RSM-predicted, and pf-ERSM-predicted O_3 responses, along with corresponding errors under (a) moderate control and (b) strict control scenarios of the PRD (monthly averaged daily 1 h maxima O_3 in 2015, unit: $\mu g/m^3$).

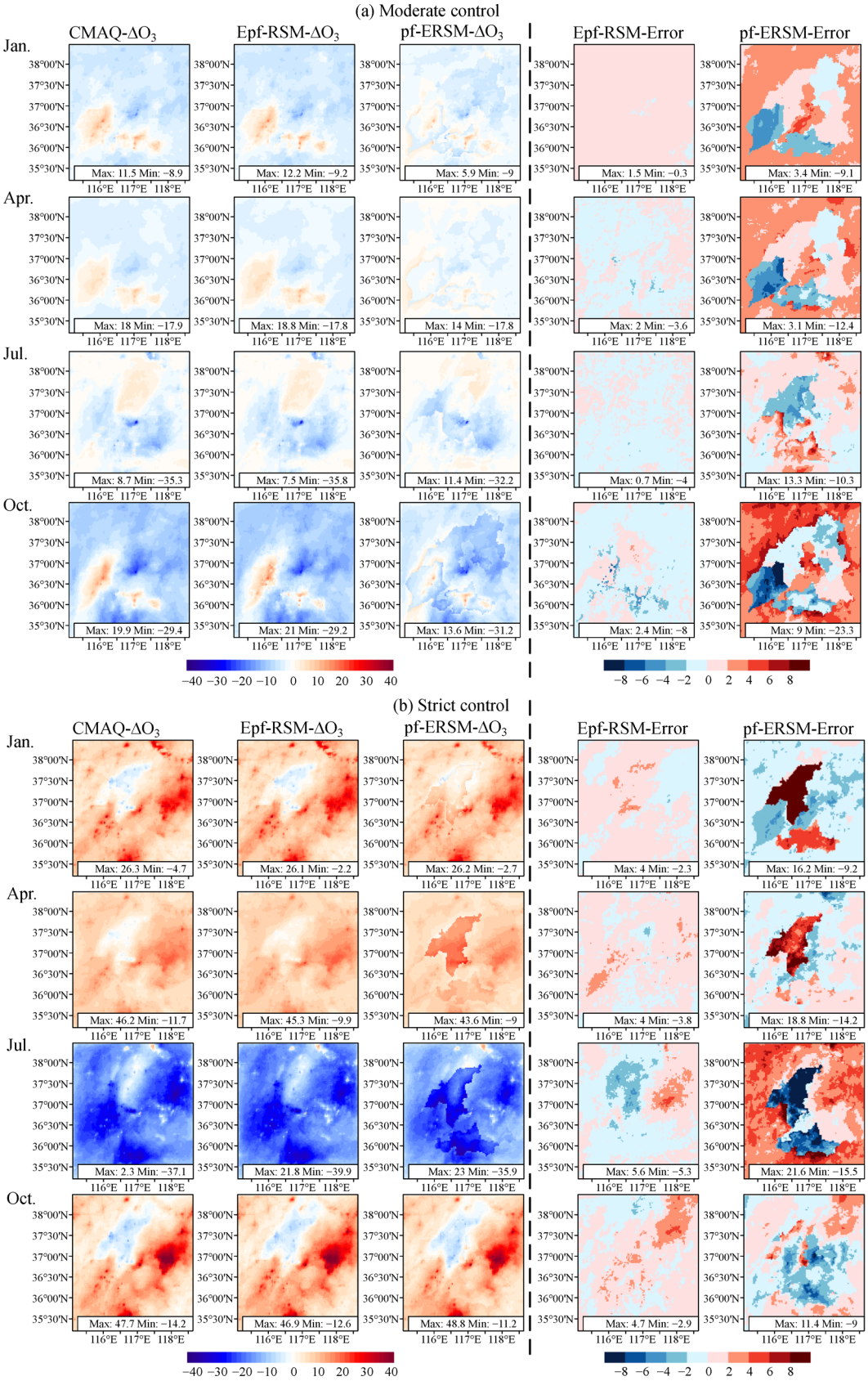


Fig. 5 Spatial distribution of CMAQ-simulated, Epf-RSM-predicted, and pf-ERSM-predicted O_3 responses, along with corresponding errors under (a) moderate control and (b) strict control scenarios of the NWShD (monthly averaged daily 1 h maxima O_3 in 2017, unit: μg/m^3).

over the transition areas due to directly fitting the polynomial function to predict the SR of each grid cell for avoiding the use of $E'_{\mathrm{prec},T}$.

3.2.3 Comparison of the EKMA diagrams

The EKMA diagrams were used to further evaluate the performance of the two ERSM systems for an O_3-NO_x-VOC sensitivity analysis. Figure 6 depicted the EKMA diagrams derived from (a) Epf-RSM for the central area of PRD (cPRD, i.e., the PRD without OTH), (b) pf-ERSM for the cPRD, (c) Epf-RSM for the central area of NWShD (cNWShD, i.e., the NWShD without OTH), and (d) pf-ERSM for the cNWShD. Figure 6 showed similar nonlinear responses for low to moderate emission cuts (approximately less than 60%) in the Epf-RSM-based and the pf-ERSM-based EKMA diagrams. However, obviously distorted isopleths were observed in the pf-ERSM-based EKMA diagrams for large emission cuts, especially when the emission cuts of NO_x exceeded 60%. These would affect the accuracy of pf-ERSM in quantify-

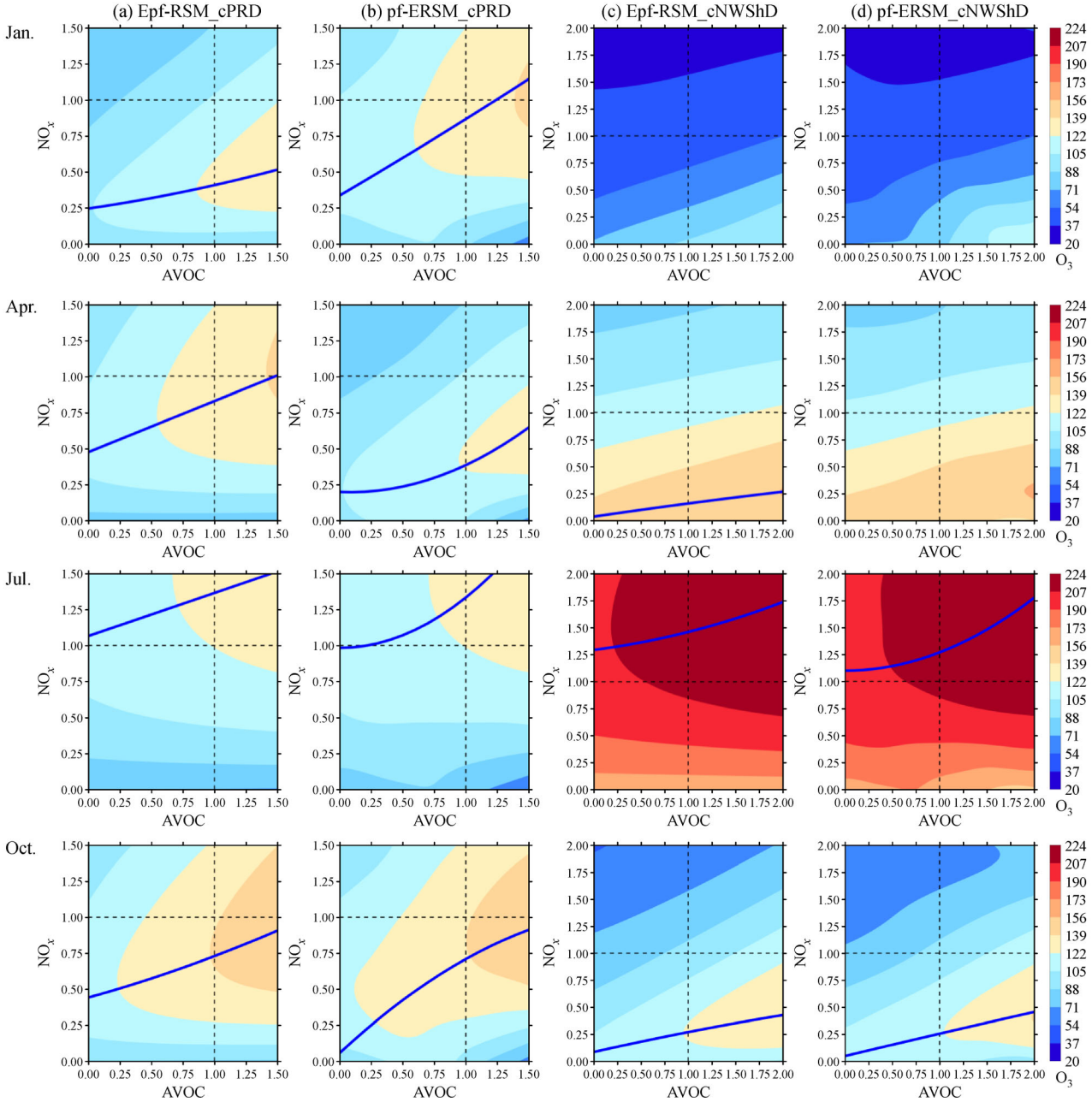


Fig. 6 Comparison of the EKMA diagrams as derived from the (a) Epf-RSM for the cPRD, (b) pf-ERSM for the cPRD, (c) Epf-RSM for the cNWShD, and (d) pf-ERSM for the cNWShD (monthly averaged daily 1 h maxima O_3; unit: μg/m^3) (Note: the *x*- and *y*-axis represent the ratios of current NO_x and AVOC emissions to base emissions respectively; the blue lines are ridgelines).

ing the effectiveness of large emission reductions. The results illustrated that pf-ERSM could only aid in developing precise O_3 control strategies for short-term O_3 mitigation (low to moderate emission reductions) but not for long-term attainment (large emission reductions). Contrastingly, Epf-RSM introduced the HCAM and optimized the ERSM calculation process, effectively increasing the accuracy of quantifying the effectiveness of both short-term and long-term O_3 reduction measures.

3.3 Data support for ozone control strategy

The Epf-RSM-based EKMA diagrams were useful for O_3-NO_x-VOC sensitivity analysis, which provided important information for policymakers to understand the ambient O_3 behavior. In an EKMA diagram, the NO_x-sensitive and VOC-sensitive regimes were separated by a ridgeline (blue line) corresponding to the 1 h maxima O_3 concentration for a given E_{AVOC} to produce (Ou et al., 2016). As shown in Fig. 6(a), the baseline scenarios of the cPRD in January, April, and October were above the ridgelines, indicating that O_3 formations were in VOC-sensitive regimes. The baseline scenario of the cPRD in July was below the ridgeline, showing that O_3 formation was in a NO_x-sensitive regime. For the cNWShD (Fig. 6(c)), the ridgelines were invisible or close to the NO_x zero-out line in January, April, and October, showing a strong VOC-sensitive condition, while the baseline scenario in July was NO_x-sensitive.

The domain-averaged PR value (defined in Text S5) of each region could be used for better understanding the nonlinearity between O_3 and its precursor emissions (Table 2). If PR < 1 (i.e., baseline), the baseline scenario was in a VOC-sensitive regime, otherwise, the baseline scenario was in a NO_x-sensitive regime (Xing et al., 2011; Xing et al., 2018). The domain-averaged PRs of the cPRD were 0.41, 0.83, 1.37, and 0.73, and those of the cNWShD were < 0.00, 0.16, 1.46, and 0.27 in January, April, July, and October, respectively; which showed the seasonal variation of O_3-NO_x-VOC sensitivity consistent with that seen in the EKMA diagdrams. Moreover, Table 2 showed that the domain-averaged PRs of Guangzhou were lower than those of other regions in the cPRD in four months; this was mainly attributed to the massive traffic and industrial emissions of NO_x in the Guangzhou (Yang et al., 2019). For the NWShD, the lower domain-averaged PR was observed in Zibo in July (Table 2); this might be because of the high NO_x emission from traffic and industrial emissions and coal-fired power plants in Zibo (Yao et al., 2019). These results were similar to the O_3 sensitivity analysis results for the same regions in the observation-based study that identified the O_3 sensitivity of China using an observable response indicator (Observable Peak Ratio) (Xing et al., 2019).

For preventing an increase in O_3 levels from the emission controls, the minimum ratio of AVOC to NO_x reduction (VNr, defined in Text S5) was calculated with Epf-RSM to help design an O_3 control strategy. The domain-averaged VNr values of the cPRD were estimated to be 1.27, 0.36, –0.81, and 0.37 in January, April, July, and October, respectively. The VNr in July was less than 0 due to the cPRD within the NO_x-sensitive regime, suggesting that O_3 concentration could be reduced by just controlling NO_x. Compared to the cPRD, the cNWShD was within stronger VOC-sensitive conditions in January, April, and October, which required more simultaneous AVOC control with NO_x. The domain-averaged VNr values of the cNWShD were estimated to be 4.17, 5.47, and 2.25 in January, April, and October, respectively. In July, the domain-averaged VNr value was

Table 2 The domain-averaged PR of each region

Region		Jan.	Apr.	Jul.	Oct.
PRD	Shunde	0.35	0.83	>1.50	0.69
	Foshan (excluding Shunde)	0.37	0.86	>1.50	0.72
	Guangzhou	0.31	0.74	1.22	0.56
	Zhongshan	0.40	0.81	>1.50	0.74
	Jiangmen	0.41	0.86	>1.50	0.73
	Dongguan & Shenzhen	0.49	0.74	1.37	0.87
	cPRD	0.41	0.83	1.37	0.73
NWShD	Jinan	< 0.00	0.19	1.65	0.28
	Dezhou	< 0.00	0.16	>2.00	0.27
	Binzhou	< 0.00	0.19	>2.00	0.32
	Liaocheng	< 0.00	0.15	1.75	0.23
	Taian	< 0.00	0.19	1.81	0.25
	Zibo	< 0.00	0.17	1.23	0.23
	cNWShD	< 0.00	0.16	1.46	0.27

estimated to be −1.66. For both regions, the simultaneous AVOC reduction could help to avoid the increase of O_3 along with NO_x controls under VOC-sensitive conditions in January, April, and October, while controlling NO_x only could reduce O_3 concentration under NO_x-sensitive condition in July.

4 Conclusions

The application of pf-ERSM for supporting an O_3 control strategy was hindered by the relatively poor performance in the margin and the transition areas. To break through this limitation, an Epf-RSM was developed. Epf-RSM innovatively integrated the HCAM into the optimized ERSM system to resolve the problem of relatively large errors in the margin and the transition areas, moreover, the calculation efficiency was further improved due to the algorithm dimension reduction and the optimization of the redundant computation process.

The prediction performance of Epf-RSM and pf-ERSM on O_3 concentrations in the two case study domains was compared. The results showed that Epf-RSM could reproduce the CMAQ simulations better than pf-ERSM with MeanNE and 95th MaxNE decreased from 2.8 and 10.2% to 0.6 and 2.3%, respectively. The spatial distribution maps and the EKMA diagrams demonstrated that Epf-RSM could effectively reduce the errors in the margin and the transition areas.

The O_3-NO_x-VOC sensitivity results analyzed by Epf-RSM indicated that the O_3 productions in both cPRD and cNWShD regions were limited by AVOC in January, April, and October, while they were NO_x-sensitive in July. The O_3 sensitivities of both regions varied for different seasons. Therefore, the seasonality of the chemical mechanism used to design an O_3 control strategy was considered. The minimum AVOC-to-NO_x reduction ratios were recommended to be 1.27, 0.36, and 0.37 for avoiding increasing O_3 concentrations in January, April, and October for the cPRD, respectively. Stronger VOC-sensitive regimes prevailed in cNWShD during the three months when compared to the cPRD. This required more simultaneous AVOC control with NO_x (i.e., at least 4.17, 5.47, and 2.25 times, respectively). However, more strict control of NO_x was recommended to lower the ambient O_3 for both regions in July.

The uncertainty of WRF-CMAQ simulation results will inevitably influence the results of the sensitivity analysis for O_3 because the Epf-RSM system is established based on the WRF-CMAQ simulation results. Hence, it is important to continually develop the emission inventory and improve the model performance of WRF-CMAQ for our future research in this domain.

Acknowledgements This work was supported by the Science and Technology Program of Guangzhou, China (No. 202002030188), the National Key Research and Development Program of China (No.2016YFC0207606), US EPA Emission, Air quality, and Meteorological Modeling Support (No. EP-D-12-044), the National Natural Science Foundation of China (Grant No. 21625701), the Fundamental Research Funds for the Central Universities (Nos. D2160320, D6180330, and D2170150) and the Natural Science Foundation of Guangdong Province, China (No. 2017A030310279).

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11783-020-1323-0 and is accessible for authorized users.

References

Cardelino C A, Chameides W L (1995). An observation-based model for analyzing ozone precursor relationships in the urban atmosphere. Journal of the Air & Waste Management Association, 45(3): 161–180

Cardelino C A, Chameides W L (2000). The application of data from photochemical assessment monitoring stations to the observation-based model. Atmospheric Environment, 34(12–14): 2325–2332

Chen X, Situ S P, Zhang Q, Wang X M, Sha C Y, Zhouc L Y, Wu L Q, Wu L L, Ye L M, Li C (2019). The synergetic control of NO_2 and O_3 concentrations in a manufacturing city of southern China. Atmospheric Environment, 201: 402–416

Collet S, Kidokoro T, Karamchandani P, Shah T (2018). Future-year ozone isopleths for South Coast, San Joaquin Valley, and Maryland. Atmosphere, 9(9): 354

Ding D, Xing J, Wang S X, Chang X, Hao J M (2019). Impacts of emissions and meteorological changes on China's ozone pollution in the warm seasons of 2013 and 2017. Frontiers of Environmental Science & Engineering, 13(5): 76

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

Guindon S, Gascuel O (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology, 52(5): 696–704

Hammersley J M (1960). Monte Carlo methods for solving multivariable problems. Annals of the New York Academy of Sciences, 86(3): 844–874

Heyes C, Schöpp W, Amann M, Bertok I, Cofala J, Gyarfas F, Klimont Z, Makowski M, Shibayev S (1997). A model for optimizing strategies for controlling ground-level ozone in Europe. Laxenburg: International Institute for Applied Systems Analysis

Heyes C, Schöpp W, Amann M, Unger S (1996). A reduced-form model to predict long-term ozone concentrations in Europe. Laxenburg: International Institute for Applied Systems Analysis

Liao K J, Tagaris E, Manomaiphiboon K, Napelenok S L, Woo J H, He S, Amar P, Russell A G (2007). Sensitivities of ozone and fine particulate matter formation to emissions under the impact of potential future climate change. Environmental Science & Technology, 41(24): 8355–8361

Lin H W, Sun L J (2015). Searching globally optimal parameter sequence for defeating Runge phenomenon by immunity genetic

algorithm. Applied Mathematics and Computation, 264: 85–98

Lozano M, Herrera F, Krasnogor N, Molina D (2004). Real-coded memetic algorithms with crossover hill-climbing. Evolutionary Computation, 12(3): 273–302

Ma Q A, Cai S Y, Wang S X, Zhao B, Martin R V, Brauer M, Cohen A, Jiang J K, Zhou W, Hao J M, Frostad J, Forouzanfar M H, Burnett R T (2017). Impacts of coal burning on ambient $PM_{2.5}$ pollution in China. Atmospheric Chemistry and Physics, 17(7): 4477–4491

Ou J M, Yuan Z B, Zheng J Y, Huang Z J, Shao M, Li Z K, Huang X B, Guo H, Louie P K K (2016). Ambient ozone control in a photochemically active region: Short-term despiking or long-term attainment? Environmental Science & Technology, 50(11): 5720–5728

Pu X, Wang T J, Huang X, Melas D, Zanis P, Papanastasiou D K, Poupkou A (2017). Enhanced surface ozone during the heat wave of 2013 in Yangtze River Delta region, China. Science of the Total Environment, 603: 807–816

Seltzer K M, Shindell D T, Malley C S (2018). Measurement-based assessment of health burdens from long-term ozone exposure in the United States, Europe, and China. Environmental Research Letters, 13(10): 104018

Sharma S, Chatani S, Mahtta R, Goel A, Kumar A (2016). Sensitivity analysis of ground level ozone in India using WRF-CMAQ models. Atmospheric Environment, 131: 29–40

Su R, Lu K D, Yu J Y, Tan Z F, Jiang M Q, Li J, Xie S D, Wu Y S, Zeng L M, Zhai C Z, Zhang Y H (2018). Exploration of the formation mechanism and source attribution of ambient ozone in Chongqing with an observation-based model. Science China. Earth Sciences, 61 (1): 23–32

Sun L, Xue L K, Wang Y H, Li L L, Lin J T, Ni R J, Yan Y Y, Chen L L, Li J, Zhang Q Z, Wang W X (2019). Impacts of meteorology and emissions on summertime surface ozone increases over central eastern China between 2003 and 2015. Atmospheric Chemistry and Physics, 19(3): 1455–1469

Tan Z F, Lu K D, Dong H B, Hu M, Li X, Liu Y H, Lu S H, Shao M, Su R, Wang H C, Wu Y S, Wahner A, Zhang Y H (2018a). Explicit diagnosis of the local ozone production rate and the ozone-NO_x-VOC sensitivities. Science Bulletin, 63(16): 1067–1076

Tan Z F, Lu K D, Jiang M Q, Su R, Dong H B, Zeng L M, Xie S D, Tan Q W, Zhang Y H (2018b). Exploring ozone pollution in Chengdu, southwestern China: A case study from radical chemistry to O_3-VOC-NO_x sensitivity. Science of the Total Environment, 636: 775–786

Wang N, Lyu X P, Deng X J, Huang X, Jiang F, Ding A J (2019). Aggravating O_3 pollution due to NO_x emission control in eastern China. Science of the Total Environment, 677: 732–744

Wang T, Xue L K, Brimblecombe P, Lam Y F, Li L, Zhang L (2017). Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects. Science of the Total Environment, 575: 1582–1596

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 NO_x control from simultaneous reductions of VOC and NH_3 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

Xue L K, Wang T, Gao J, Ding A J, Zhou X H, Blake D R, Wang X F, Saunders S M, Fan S J, Zuo H C, Zhang Q Z, Wang W X (2014). Ground-level ozone in four Chinese cities: precursors, regional transport and heterogeneous processes. Atmospheric Chemistry and Physics, 14(23): 13175–13188

Yang W, Chen H, Wang W, Wu J, Li J, Wang Z, Zheng J, Chen D (2019). Modeling study of ozone source apportionment over the Pearl River Delta in 2015. Environmental pollution (Barking, Essex: 1987), 253: 393–402

Yao Y R, He C, Li S Y, Ma W C, Li S, Yu Q, Mi N, Yu J, Wang W, Yin L, Zhang Y (2019). Properties of particulate matter and gaseous pollutants in Shandong, China: Daily fluctuation, influencing factors, and spatiotemporal distribution. Science of the Total Environment, 660: 384–394

Ye L M, Wang X M, Fan S F, Chen W H, Chang M, Zhou S Z, Wu Z Y, Fan Q (2016). Photochemical indicators of ozone sensitivity: Application in the Pearl River Delta, China. Frontiers of Environmental Science & Engineering, 10(6): 15

Yue X, Unger N, Harper K, Xia X, Liao H, Zhu T, Xiao J, Feng Z, Li J (2017). Ozone and haze pollution weakens net primary productivity in China. Atmospheric Chemistry and Physics, 17(9): 6073–6089

Zhang K, Xu J L, Huang Q, Zhou L, Fu Q Y, Duan Y S, Xiu G L (2020). Precursors and potential sources of ground-level ozone in suburban Shanghai. Frontiers of Environmental Science & Engineering, 14(6): 92

Zhong J, Cai X M, Bloss W J (2014). Modelling segregation effects of heterogeneous emissions on ozone levels in idealised urban street canyons: Using photochemical box models. Environmental Pollution, 188: 132–143