



Full length article

# Ozone response modeling to NO<sub>x</sub> and VOC emissions: Examining machine learning models

Cheng-Pin Kuo<sup>a</sup>, Joshua S. Fu<sup>a,b,\*</sup><sup>a</sup> Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA<sup>b</sup> Department of Atmospheric Sciences, National Central University, Taoyuan, Taiwan

## ARTICLE INFO

## Keywords:

Ozone  
Emission control  
Forecasting  
Machine learning  
Measurement-model fusion

## ABSTRACT

Current machine learning (ML) applications in atmospheric science focus on forecasting and bias correction for numerical modeling estimations, but few studies examined the nonlinear response of their predictions to precursor emissions. This study uses ground-level maximum daily 8-hour ozone average (MDA8 O<sub>3</sub>) as an example to examine O<sub>3</sub> responses to local anthropogenic NO<sub>x</sub> and VOC emissions in Taiwan by Response Surface Modeling (RSM). Three different datasets for RSM were examined, including the Community Multiscale Air Quality (CMAQ) model data, ML-measurement-model fusion (ML-MMF) data, and ML data, which respectively represent direct numerical model predictions, numerical predictions adjusted by observations and other auxiliary data, and ML predictions based on observations and other auxiliary data.

The results show that both ML-MMF ( $r = 0.93\text{--}0.94$ ) and ML predictions ( $r = 0.89\text{--}0.94$ ) present significantly improved performance in the benchmark case compared with CMAQ predictions ( $r = 0.41\text{--}0.80$ ). While ML-MMF isopleths exhibit O<sub>3</sub> nonlinearity close to actual responses due to their numerical base and observation-based correction, ML isopleths present biased predictions concerning their different controlled ranges of O<sub>3</sub> and distorted O<sub>3</sub> responses to NO<sub>x</sub> and VOC emission ratios compared with ML-MMF isopleths, which implies that using data without support from CMAQ modeling to predict the air quality could mislead the controlled targets and future trends. Meanwhile, the observation-corrected ML-MMF isopleths also emphasize the impact of transboundary pollution from mainland China on the regional O<sub>3</sub> sensitivity to local NO<sub>x</sub> and VOC emissions, which transboundary NO<sub>x</sub> would make all air quality regions in April more sensitive to local VOC emissions and limit the potential effort by reducing local emissions.

Future ML applications in atmospheric science like forecasting or bias correction should provide interpretability and explainability, except for meeting statistical performance and providing variable importance. Assessment with interpretable physical and chemical mechanisms and constructing a statistically robust ML model should be equally important.

## 1. Introduction

Air pollution has gained great attention owing to its adverse effects on human health (Apte et al., 2018; Kuo et al., 2021), climate (Dong et al., 2019), agriculture (Tai and Martin, 2017), ecosystems (Zarnetske et al., 2021), and visibility (Huang et al., 2014). In regional air quality management, controlling local anthropogenic emissions is a common way to improve regional air quality. Predicting air quality under designed emission control scenarios by using chemical transport models (CTMs) like the Community Multiscale Air Quality (CMAQ) model has been much studied (Arnold and Dennis, 2006; Che et al., 2011).

Moreover, to meet the prompt and various needs of policymakers, response surface modeling (RSM) was developed to assess the improved or changed air quality based on designed emission control strategies without extra CTM simulations. That is, RSM can retrieve the nonlinear equation between ambient air pollutant concentrations (e.g. O<sub>3</sub>) and multiple precursors emissions (e.g. NO<sub>x</sub> and VOCs) from multiple emission sources based on an ensemble of CTM simulations with changing O<sub>3</sub> precursors emissions (Wang et al., 2011; Xing et al., 2018; Zhu et al., 2015), and the user can apply the retrieved nonlinear equation to timely estimate the changed air pollutant concentrations based on input emission ratios and support emission control strategies. For

\* Corresponding author at: Department of Civil and Environmental Engineering, The University of Tennessee, 851 Neyland Drive, Knoxville, TN 37996, USA.  
E-mail address: [jfsu@utk.edu](mailto:jfsu@utk.edu) (J.S. Fu).

<https://doi.org/10.1016/j.envint.2023.107969>

Received 3 January 2023; Received in revised form 5 April 2023; Accepted 8 May 2023

Available online 12 May 2023

0160-4120/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

example, RSM has been intensively applied to assess the NO<sub>x</sub> and VOC emission control strategies of O<sub>3</sub> pollution (Li et al., 2022; Xing et al., 2018; Zhao et al., 2015). RSM can identify ambient O<sub>3</sub> sensitivity to NO<sub>x</sub> and VOC emission by O<sub>3</sub> isopleth and further classify the isopleth regimes into two chemical regimes, NO<sub>x</sub>-limited and VOC-limited. In the NO<sub>x</sub>-limited regimes, O<sub>3</sub> increases with increased NO<sub>x</sub> emissions and exhibits limited response to VOC emissions, and vice versa. Classification of the O<sub>3</sub> formation regime in the isopleth can assist policymakers to determine whether NO<sub>x</sub> or VOC emissions should be controlled preferentially in emission control strategies (Gipson et al., 1980). However, although RSM was employed to improve regional air quality in several previous studies, these studies were still based on simulated results and neglected the bias between the modeled estimations and observations in the benchmark case (Kelly et al., 2021; Li et al., 2022; Xing et al., 2022), which could largely affect the nonlinearity between pollutants and precursor emission changes.

To forecast air quality and support air quality policies, machine learning (ML) or machine intelligence has been rapidly developed and intensively implemented in environmental science and air quality management (Kang et al., 2018; Zheng et al., 2021). ML in atmospheric science applications for predicting air quality is mostly driven by historical air quality data, real-time monitoring data, measurements, satellite images, land-use information, or emission activity-related indices. ML is relatively easy to execute and can provide more accurate predictions compared with CTMs, which still need complicated data-preparing processes and computationally-intensive time and resources, and have a larger modeling bias. However, ML is also debated and remains low persuasiveness due to its black-box modeling process and failure to provide interpretability and explainability concerning physical/chemical mechanisms (Fu et al., 2022; Zheng et al., 2021).

In previous applications, ML can serve as a bias corrector to adjust modeling results. Several measurement-model fusion (MMF) (Fu et al., 2022) techniques in post-analysis have been developed in recent years to adjust CTM results based on observations (Geng et al., 2020; Lu et al., 2020; Requia et al., 2020; Sayeed et al., 2021). Air pollutant estimations without correction by observations could underestimate or overestimate improved air quality and derived environmental benefits (Fu et al., 2022). In addition, ML can also forecast air quality based on historical observations and other auxiliary data (e.g. meteorological and land-use data) without involving CTM results and still have good performance (Ausati and Amanollahi, 2016; Song et al., 2015; Zhou et al., 2019). However, whether ML either serves as a bias corrector or a forecaster, few ML studies examined pollutants' sensitivity to their precursor

emissions based on observation-corrected results.

In this study, we used the maximum daily 8-hour O<sub>3</sub> average (MDA8 O<sub>3</sub>) as the target index and selected Taiwan as the study region due to its island geography, high-density air quality monitoring stations ( $n = 73$ ), and three-year-updated emission inventory. The goal of this study is to (1) verify the capability of ML to correct CMAQ modeling results (denoted as ML-MMF data) and predict MDA O<sub>3</sub> without CMAQ-modeled results (denoted as ML data) in the benchmark-case modeling performance and (2) examine O<sub>3</sub> nonlinear responses to all anthropogenic NO<sub>x</sub> and VOC emission ratios by O<sub>3</sub> isopleths by using CMAQ, ML-MMF, and ML modeling results.

## 2. Methodology

The technical work is illustrated in Fig. 1. Three types of data (CMAQ, ML-MMF, and ML) were prepared for RSM; CMAQ data were direct outputs from CMAQ-modeled results; ML-MMF data are the corrected estimates by observation and CMAQ inputs (emissions, boundary conditions, meteorology, and land-use), and the constructed model was employed to predict O<sub>3</sub> based on the changed CMAQ outputs, the other CMAQ inputs and changed emissions (kg/day) under different emission scenarios; ML outputs are predictions constructed by using O<sub>3</sub> observations and CMAQ inputs, and the constructed model was utilized to predict O<sub>3</sub> based on the changed emissions (kg/day) and the other CMAQ inputs under different emission scenarios. Second, RSM was executed for each dataset to predict O<sub>3</sub> under different NO<sub>x</sub> and VOC emission ratios in proportion to the baseline emissions (emission ratio = 1) in the benchmark case. O<sub>3</sub> isopleths were finally constructed for each dataset and were validated by out-of-samples and observations.

### 2.1. Data preparation

Air quality regions in this study were categorized into six regions: Northern (NT), Chu-Miao (CM), Central (CT), Yun-Chia-Nan (YCN), Kao-Ping (KP), and Eastern (ET). A total of 73 air quality monitoring stations with hourly O<sub>3</sub> measurements were included (Fig. 2). The modeling period included January, April, July, and October 2016 to represent different seasons. Meteorological fields were firstly simulated by WRF (version 3.8), and hourly O<sub>3</sub> concentrations were simulated by the CMAQ model (version 5.2) with the gas-phase chemistry module, Carbon Bond 6, (Sarwar et al., 2008) and aerosol module, AERO6 (Appel et al., 2013) mechanisms. The configurations of the simulation domain nested four layers from East Asia (81 km × 81 km) to Taiwan island (3

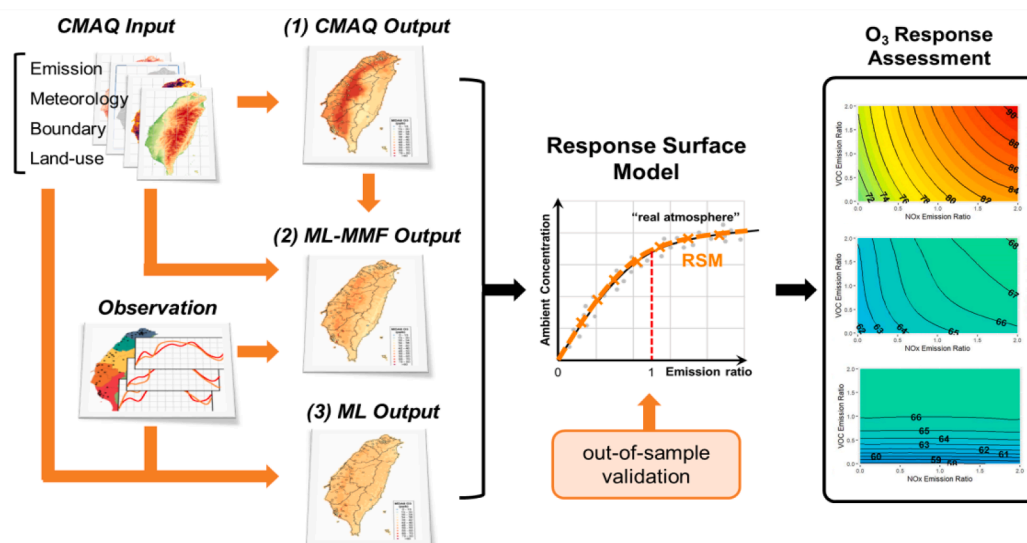


Fig. 1. Technical flowchart of this study.

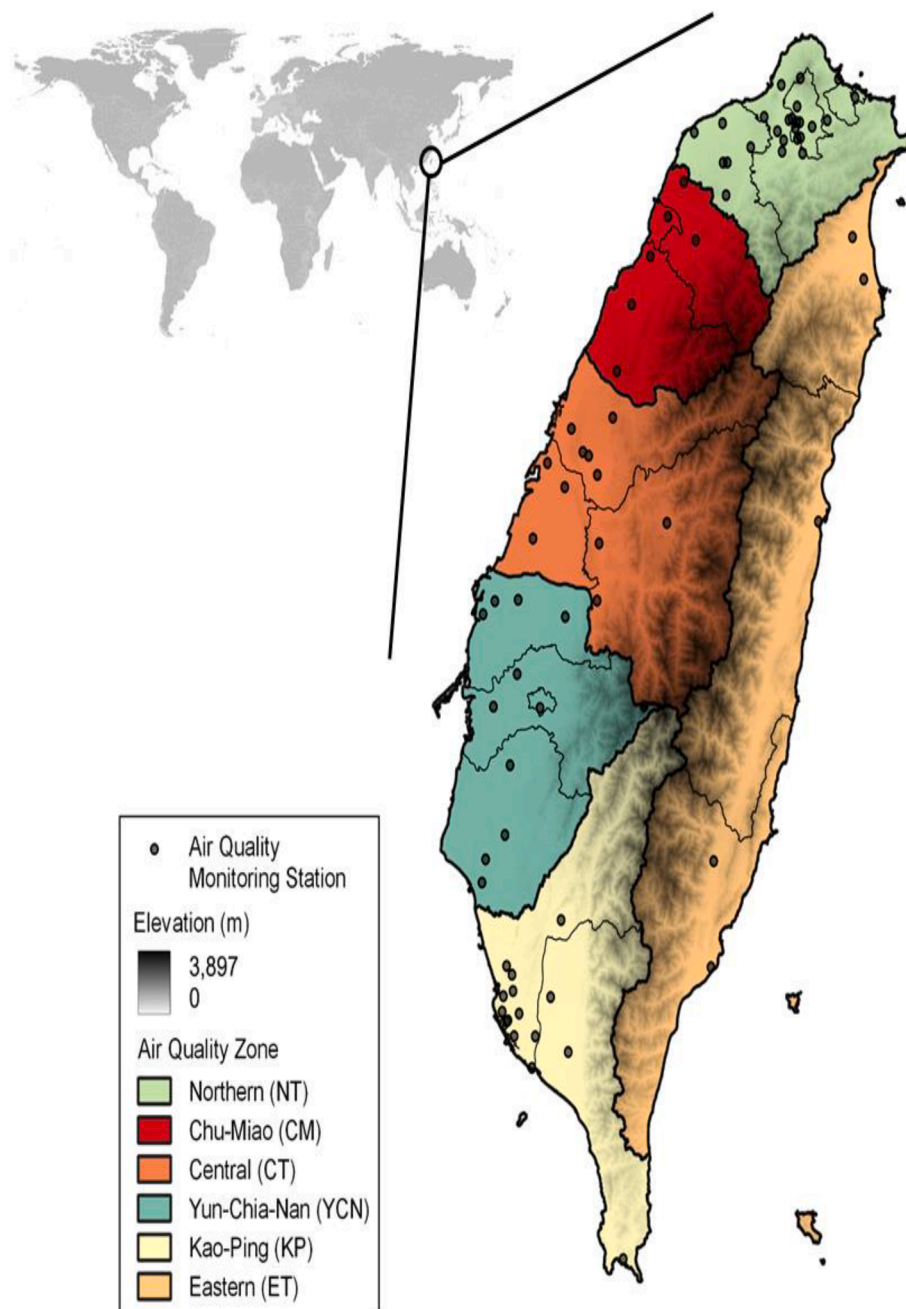


Fig. 2. Air quality monitoring stations ( $n = 73$ ) and six air quality regions in Taiwan.

$\text{km} \times 3 \text{ km}$ ) which covers 90 (row)  $\times$  135 (column) horizontal grid cells (Lai and Lin, 2020). Daily emission data from Taiwan Emission Data System (TEDS) version 10.0 with  $3 \text{ km} \times 3 \text{ km}$  resolution developed by Taiwan EPA including industrial, mobile, fugitive, and biogenic sources were used. Before CMAQ modeling, emission distribution and speciation were processed by the USEPA SMOKE program (U.S. EPA, 2018), which can apply temporal and spatial allocation and chemical speciation for industrial, mobile, and fugitive sources from TEDS. The modeling performance assessment of meteorology factors and CAMQ-modeled  $\text{O}_3$  are shown in Tables A1 and A2, respectively.

The details of the developed ML-MMF and ML framework were illustrated in Appendix I, and the major difference between ML-MMF and ML framework is predicting  $\text{O}_3$  with and without CMAQ direct outputs for prediction, as shown in the following conceptual equations:

$$\text{MLMMF } \text{O}_3 \text{ predictions} = f(\text{CMAQ}, \text{Emis}, \text{BC}, \text{Met}, \text{Land})$$

$$\text{ML } \text{O}_3 \text{ predictions} = f(\text{Emis}, \text{BC}, \text{Met}, \text{Land})$$

where CMAQ is CAMQ output; Emis are emission data (kg/day); BC are boundary conditions; Met are meteorological variables; Land are land-use variables. A total of 4039 grid cells with local emission, meteorology, and land-use information was considered. Basically, Both ML-MMF and ML models employed five techniques including the k-nearest neighbors' regression (KNN) (Kramer, 2013), regression tree (RT) (Loh, 2008), random forest (RF) (Deng et al., 2001), gradient boosted tree models (GBM) (Natekin and Knoll, 2013), and convolutional neural network (CNN) (Albawi et al., 2017), which can construct the nonlinearity between input variables (emissions, boundary conditions, meteorology, and land-use data) and observed  $\text{O}_3$  concentrations in the benchmark case. The best model with higher accuracy, reasonable spatial predictions, and no overfitting tendency was selected to predict

O<sub>3</sub> concentrations (Appendix I); 60% and 40% of the data set were selected as the training dataset and the testing dataset, respectively; The 10-fold cross-validation was also conducted to optimize hyper-parameters of the training dataset based on the uncertainty of modeling performance. The selected model for ML-MMF and ML would be further utilized to predict O<sub>3</sub> based on CMAQ outputs (not for the ML model), the changed emissions (kg/day), and the other fixed auxiliary data (boundary conditions, meteorology, and land-use data) under different emission scenarios.

The variables selected for ML-MMF and ML were related to emissions of precursor species including NO<sub>x</sub> and VOCs, boundary conditions that affect the background level of O<sub>3</sub>, NO<sub>x</sub>, and VOCs, meteorological factors involved with photochemical reactions and transport fluxes of air, and time-independent land-use geographical information (Table 1). Boundary conditions are the averaged values of all the grid cells around the modeling domain, and the temporal variations of meteorological and air pollution information during modeling period were considered. Meteorological factors on 850 hPa and 690 hPa represent the weather conditions of the mixing layer and lower troposphere layer (Lu et al., 2021). MDA8 O<sub>3</sub> would be the dependent variable because O<sub>3</sub> concentrations usually are higher during the daytime due to the existence of ultraviolet energy (Ware et al., 2016). For each month, the modeling performance was evaluated by correlation coefficient (r), mean absolute error (MAE), and root mean square error (RMSE) as shown in the following equation.

$$r = \frac{\sqrt{\sum_{i=1}^n \sum_{j=1}^m (M_{ij} - \bar{M})(O_{ij} - \bar{O})}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^m (M_{ij} - \bar{M})^2 \sum_{i=1}^n \sum_{j=1}^m (O_{ij} - \bar{O})^2}}$$

$$MAE = \frac{\sum |M_{ij} - O_{ij}|}{N \times M}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (M_{ij} - O_{ij})^2}{N \times M}}$$

where  $M_{ij}$  is the modeled O<sub>3</sub> for  $i$  th station in  $j$  th day;  $O_{ij}$  the observed O<sub>3</sub> for  $i$  th station in  $j$  th day;  $\bar{M}$  is the average of modeled O<sub>3</sub> for all

**Table 1**  
Selected variables for ML-MMF and ML model.

Dataset	Source	Variables
Emission	TEDS 10	NO <sub>x</sub> and VOCs emissions from point, mobile, fugitive, and biogenic sources (g/s).
Boundary conditions	WRF/ GEOS- Chem	air temperature (850 and 690 hPa) (K), relative humidity (850 and 690 hPa)(%), NO <sub>3</sub> , HNO <sub>3</sub> , N <sub>2</sub> O <sub>5</sub> , NO, NO <sub>2</sub> , VOCs and O <sub>3</sub> at surface level (ppmV)
Meteorological variables	WRF	Surface Surface pressure (Pa), PBL height (m), temperature at 2 m (K), mixing ratio at 2 m (unitless), wind speed (m/s), U wind component and V wind component (m/s), solar radiation reaching ground (W/m <sup>2</sup> ), precipitation (cm), total cloud fraction (unitless), average liquid water content of cloud (g/m <sup>3</sup> ) Pressure level (850 and 690 hPa) Air temperature (K), potential vorticity (m <sup>2</sup> K/kg/s), vertical velocity (m/s), U wind component and V wind component (m/s)
Land-use	CMAQ	Elevation (m), urban percent (%)

stations;  $\bar{M}$  is the average of observed O<sub>3</sub> for all stations;  $N$  is the number of stations;  $M$  is the number of days.

## 2.2. Response surface modeling (RSM)

RSM can retrieve non-linear O<sub>3</sub> responses to anthropogenic NO<sub>x</sub> and VOC emission ratios, which are changed ratios of emission compared with baseline emission in the benchmark case (emission ratio = 1). First, to generate the control matrix of anthropogenic NO<sub>x</sub> and VOC emission ratios for each air quality region, the Latin hypercube sampling (LHS) method was utilized to estimate enough sample size of the number of CMAQ simulations for RSM while providing enough statistical power and saving computing resources. The LHS design can provide flexibility, which selects a number of runs based on limited computing resources (USEPA, 2006), and also can capture the nonlinearity of ozone. The number of CMAQ runs for RSM was decided by the emission ratio range (0 %–200 %) and a number of control sectors such as power plants and mobile sources, the accuracy of O<sub>3</sub> response to NO<sub>x</sub> and VOC emissions, and the available computing resources. Typically, the number of CMAQ runs for RSM around 50 simulations is enough for constructing a statistically robust model (Li et al., 2022). Extra simulations for individual air quality regions were also conducted due to higher O<sub>3</sub> concentrations in spring, fall, and winter. Emission ratios of NO<sub>x</sub> and VOC are designed from 0% to 200%, and the number of CMAQ runs for six air quality regions is shown in Table A3.

In each air quality region, RSM involves multiple precursor emissions (NO<sub>x</sub> and VOC) and anthropogenic emission sectors (industrial, mobile, and fugitive sources) from multiple cities and counties, as shown in Fig. A1 and Table A4. Multiple city/county-level NO<sub>x</sub> and VOC emission sectors (Table A4) were used for each air quality region. Self-adaptive RSM (SA-RSM) based on regression method and stepwise selection was employed to predict O<sub>3</sub> based on designed emission ratios of NO<sub>x</sub> and VOC, and the followed multidimensional kriging method was used to illustrate O<sub>3</sub> isopleths to show O<sub>3</sub> nonlinear response to NO<sub>x</sub> and VOC emission ratios (Li et al., 2022; USEPA, 2006; Xing et al., 2019, 2018). In each air quality region, the averages of grid cells with daily MDA8 O<sub>3</sub> higher than 60 ppb were used for RSM based on WHO and Taiwan EPA standards (World Health Organization, 2005). Also, higher O<sub>3</sub> concentrations were identified to be more sensitive to anthropogenic emissions of NO<sub>x</sub> and VOC (Xing et al., 2011). In each region, the following equation linking the O<sub>3</sub> concentration ( $\Delta\text{Conc}$ ) in each grid to city/county-level NO<sub>x</sub> or VOC emission ratios of the emission sectors and stepwise selection that can automatically select polynomial variables for all grid cells with 0.15 of the entering level and the leaving level were utilized (Li et al., 2022):

$$\Delta\text{Conc}(m, n) = \sum_{i=1}^k X_i \bullet (\Delta E_i)^{a_i} + \sum_{i=1}^k \sum_{j=1}^{k-1} X_{ij} \bullet (\Delta E_i)^{b_i} (\Delta E_j)^{c_j}$$

where  $\Delta\text{Conc}$  is the changed daily MDA8 O<sub>3</sub> concentration (response) from the baseline scenario (the benchmark case, emission ratio = 1) in the grid ( $m, n$ );  $\Delta E_i$  is the changed emission ratios from 1 of  $k$  city/county-level NO<sub>x</sub> or VOC emission sectors;  $\Delta E_j$  is the other changed emission ratios from 1 of emission sector other than  $\Delta E_i$ ;  $a_i$ ,  $b_i$ , and  $c_j$  are the nonnegative integer powers of  $\Delta E_i$  and interaction terms of  $\Delta E_i$  and  $\Delta E_j$ . We set  $a_i$  from 1 to 3 and  $b_i$  and  $c_j$  from 1 to 2, respectively. A total of 2782 CMAQ simulations were used for RSM. To assure the performance of RSM, a total of 107 CMAQ simulations were used for out-of-sample validation (Table A3). The out-of-samples were randomly selected based on the specified emission ratios (0 %–200 %) by the LHS method as well. The RSM performance was evaluated by correlation coefficient (r), Mean normalized error (Mean NE), and Maximum NE (Max NE) (Xing et al., 2018) as shown in the following equations.



$$r = \sqrt{\frac{[\sum_{i=1}^N (M_i - \bar{M})(O_i - \bar{O})]}{\sum_{i=1}^N (M_i - \bar{M})^2 \sum_{i=1}^N (O_i - \bar{O})^2}}$$

$$\text{Mean NE} = \frac{1}{N} \sum_{i=1}^N \frac{|M_i - O_i|}{O_i}$$

$$\text{Max NE} = \max\left(\frac{|M_i - O_i|}{O_i}\right)$$

where  $M_i$  and  $O_i$  are the average of grid cells for RSM predictions and CAMQ, ML-MMF, or ML predictions in the  $i$ th simulation over different emission scenarios;  $N$  is the number of scenarios;  $\bar{M}$  and  $\bar{O}$  are the average of the RSM predictions and CAMQ, ML-MMF, or ML predictions in the  $i$ th simulation over different emission scenarios.

Finally, the response of  $O_3$  concentrations to changes in anthropogenic NOx and VOC emission ratios from all sources and all cities and counties in each air quality region was illustrated by  $O_3$  isopleths, which can be divided into NOx-limited (NOx-sensitive and VOC-rich) and VOC-limited (VOC-sensitive and NOx-rich) regimes and used to help determine whether NOx or VOC emissions should be controlled more aggressively in strategies to alleviate ground-level  $O_3$  concentrations (Gipson et al., 1980).

### 3. Results and discussion

#### 3.1. Benchmark case modeling performance

The performance of CMAQ, ML-MMF, and ML modeling compared with observations is shown in Table 2. Compared with ML-MMF and ML predictions, the CMAQ predictions have lower performance for all months ( $r = 0.41\text{--}0.80$ ,  $\text{RMSE} = 13.45\text{--}21.19$  ppb). After CMAQ data were adjusted by the auxiliary data (emission, meteorology, boundary condition, and land-use data) and corrected by observations, ML-MMF modeling presents significantly better performance for all months ( $r = 0.93\text{--}0.94$ ,  $\text{RMSE} = 4.49\text{--}7.43$  ppb). The improved performance highlights the benefits of adding auxiliary data in ML-MMF applications.

**Table 2**  
Modeling performance of CMAQ, ML-MMF, and ML of benchmark case in selected months in 2016 based on monitoring stations ( $n = 73$ ).

Data source	Index	Month				
		January	April	July	October	
CMAQ	r	0.41	0.53	0.79	0.80	
	RMSE (ppb)	13.45	17.87	21.19	19.09	
	MAE (ppb)	9.72	13.61	17.33	14.54	
ML-MMF	Train	r*	0.93 (0.15)	0.93 (0.19)	0.93 (0.21)	0.93 (0.17)
		RMSE (ppb)*	4.54 (0.36)	6.83 (0.63)	5.94 (0.63)	7.43 (0.74)
		MAE (ppb)*	3.36 (0.26)	5.14 (0.49)	4.26 (0.34)	5.34 (0.38)
	Test	r	0.93	0.93	0.93	0.94
		RMSE (ppb)	4.49	6.74	6.19	7.19
		MAE (ppb)	3.41	5.08	4.51	5.11
ML	Train	r*	0.92 (0.17)	0.92 (0.20)	0.90 (0.17)	0.90 (0.17)
		RMSE (ppb)*	4.66 (0.44)	7.18 (0.64)	7.04 (0.35)	8.94 (0.59)
		MAE (ppb)*	3.49 (0.27)	5.35 (0.42)	5.14 (0.33)	6.58 (0.47)
	Test	r	0.93	0.92	0.89	0.90
		RMSE (ppb)	4.62	7.10	7.35	8.94
		MAE (ppb)	3.49	5.41	5.41	6.40

\* The standard deviations of r, RMSE, and MAE for the trained models by 10-fold cross-validation are illustrated in the brackets.

Even though excluding CMAQ output and purely using auxiliary data for ML modeling, the ML model still maintains comparable modeling performance ( $r = 0.89\text{--}0.94$ ,  $\text{RMSE} = 4.62\text{--}8.94$  ppb). Both ML-MMF and ML have no overfitting tendency considering their similar performance in training and testing data. Averages of observed and modeled MDA8  $O_3$  for CMAQ, ML-MMF, and ML in selected months are presented in Fig. 3. Fig. 3(a) shows obvious overestimations of CMAQ compared with observations, especially in the regions along with the west side of the mountains and the central basin. Fig. 3(b)(c) shows the ML-MMF and ML estimations are much close to observations and have lower modeling bias (RMSE and MAE) compared with CMAQ.

Monthly concentrations of CMAQ, MMF, and ML are presented in Fig. A2. Among all selected months, Fig. A2(a) shows CMAQ remain overestimated in all months and has spatial specificity in each month, in which CT, YCN, and KP region are overestimated in January, NT, CM, and CT regions are overestimated in April, and whole west regions overestimated in July and October. On the other hand, Fig. A2(b)(c) shows the ML-MMF model and ML model present similar spatial distributions for all months except for the ET region in July and October. The different estimations in the ET region are due to fewer air quality monitoring stations in the region, so the ML model needs to learn the data from the monitoring sites in the western regions.

#### 3.2. Adjusted $O_3$ seasonal patterns

In Taiwan, frequent long-range transboundary pollution events from mainland China with higher air pollutant concentrations occur in spring, fall, and winter when northeast monsoon prevails (Dong et al., 2019; Huang et al., 2019). Since there is no standard method to identify long-range transboundary pollution in Taiwan, our method to classify event and non-event days is illustrated in Appendix II. In the modeling period, there were 14, 13, and 3 event days in January, April, and October, respectively, and no event days occurred in July. Fig. 4 presents the monthly MDA8  $O_3$  average of event and non-event days based on ML-MMF estimations. In January, transboundary plumes mostly contained higher PM, NOx, and CO, so there is no significant  $O_3$  concentration difference between the event and non-event days. In April, the MDA8  $O_3$  averages of 13 event days show a significant impact on the whole island. Compared with event days,  $O_3$  exceedances on non-event days mostly occur in central mountainous areas, which may result from higher biogenic VOC emissions in spring (Chang et al., 2005). In July, higher  $O_3$  level in northern Taiwan is due to the prevalence of southwest and south monsoon, and NOx, VOC, and formatted  $O_3$  transport from southern regions accumulated in northern regions. In October, both event days and non-event days show similar distribution, but event days present higher  $O_3$  concentration in western areas, especially in the NT region.

#### 3.3. RSM performance

Because higher  $O_3$  concentrations are more sensitive to anthropogenic NOx and VOC emissions (Xing et al., 2011), we performed RSM for  $O_3$  exceedance days (MDA8  $O_3 > 60$  ppb) for April, July, and October and excluded January due to no  $O_3$  exceedance days. RSM performances by using CMAQ, ML-MMF, and ML data for six air quality regions (NT, CM, CT, YCN, KP, and ET) of the event and non-event days in April, July, and October are illustrated in Table A5. Most of the RSM results meet the statistical requirement of mean NE  $< 3\%$  and max NE  $< 10\%$  considering  $r$  (0.855–1.000), mean NE (0.02%–2.62%), and max NE (0.08%–9.88%), except for the CM region's October (max NE = 11.48%) and the YCN region's April (max NE = 11.38%) and October (max NE = 12.04%) of CMAQ data and KP's April ( $r = 0.759$ ) of ML-MMF data. Good RSM performance represents that RSM can reproduce CMAQ, ML-MMF, or ML outputs by only using county/city-level NOx and VOC emission ratios without extra CMAQ simulations or ML-MMF/ML modeling. The higher max NE may be due to unstable RSM performance when NOx and VOC emission ratios are very low (0%) or very

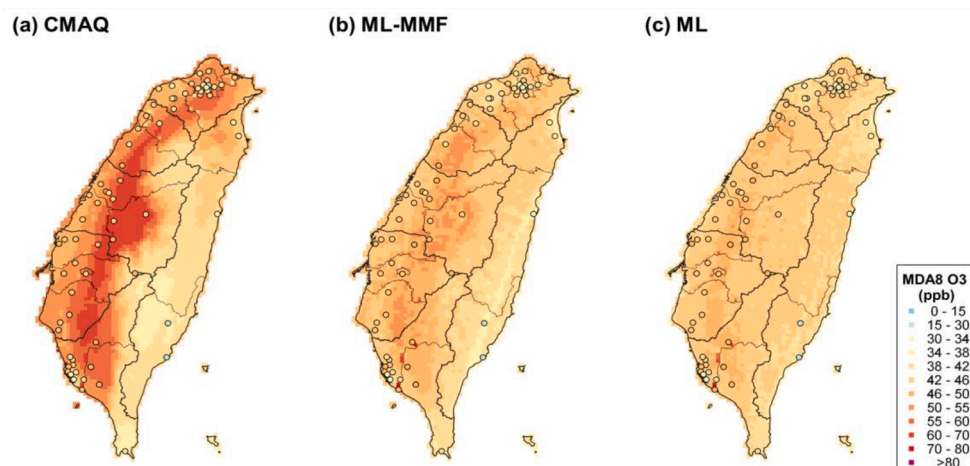


Fig. 3. Observed MDA8 O<sub>3</sub> (circles) and modeled estimations of (a) CMAQ, (b) ML-MMF, and (c) ML.

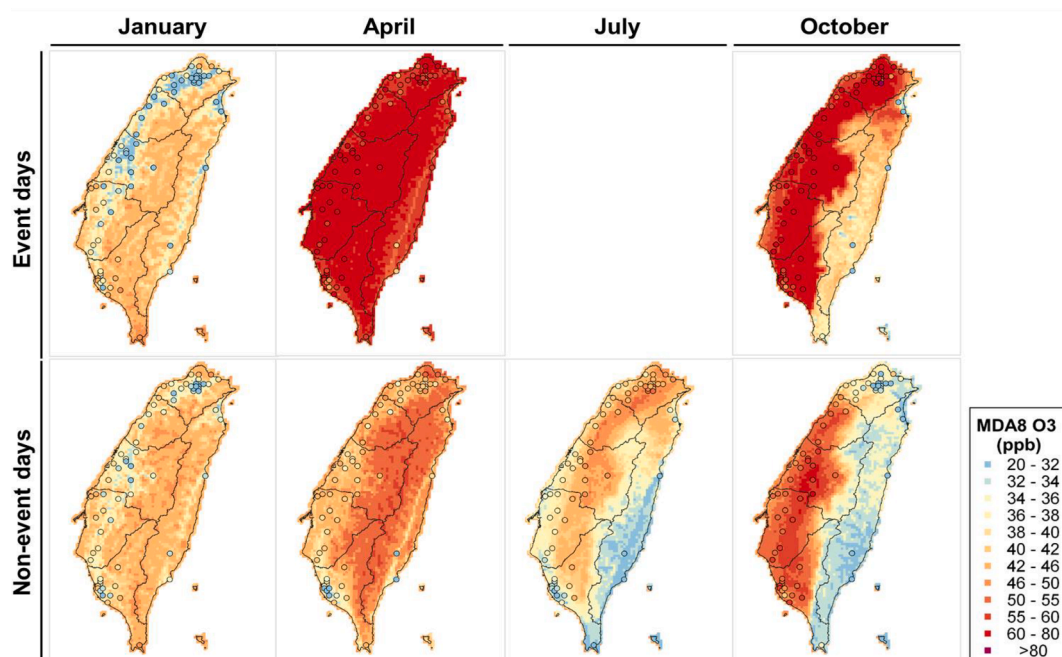


Fig. 4. Monthly MDA8 O<sub>3</sub> concentrations for event days (with long-range transboundary pollution) and non-event days in selected months by ML-MMF data. No event days occurred in July.

high (200 %) (Luo et al., 2021). Although the limited performance could be improved by adding more simulations under these extreme emission ratios, these scenarios are still impractical and hardly verified by observations. For example, it is almost impossible to measure pollutant concentrations without any anthropogenic emissions in urban areas. Furthermore, the validation of the baseline emission ratio (emission ratio = 1, Fig. A3) compared with observations shows a satisfactory performance (Mean NE = 2–6 %). Second, the higher bias may imply the high impact of non-local emissions, which means air pollutants or their precursors from upwind countries or other air quality regions dominate downwind local air quality, so local emissions only have a limited impact on local air quality and hardly explain the temporal variation of observed concentrations. As previously mentioned, in April and October, long-range transboundary air pollution transported from mainland China and plumes from upwind regions frequently accumulate and deteriorate air quality in southern air quality regions.

#### 3.4. O<sub>3</sub> sensitivity to local NO<sub>x</sub> and VOC emissions

Monthly combined O<sub>3</sub> isopleths of all air quality regions in April, July, and October for the event and non-event days by using CMAQ, ML-MMF, and ML data are presented in Fig. 5, and the O<sub>3</sub> isopleths for the individual zone are shown in Figs. A4–A8. All isopleths for each air quality region show the averages of grid cells having MDA8 O<sub>3</sub> concentrations higher than 60 ppb in the region, and the x-axis and y-axis represent the NO<sub>x</sub> and VOCs emission ratios for all sources in all cities and counties in the region.

First, ML-MMF isopleth is considered as the better tool to explain O<sub>3</sub> nonlinear responses to anthropogenic NO<sub>x</sub> and VOC emissions due to its observation-based adjustment and CTM basis, and ML-MMF isopleths show that the improved effort by reducing local emissions should be less than what CMAQ model simulated. For example, the CMAQ isopleth in July ranges from 72 ppb to 90 ppb concerning different emission ratios, but the ML-MMF isopleth presents a narrower range, which is from 62

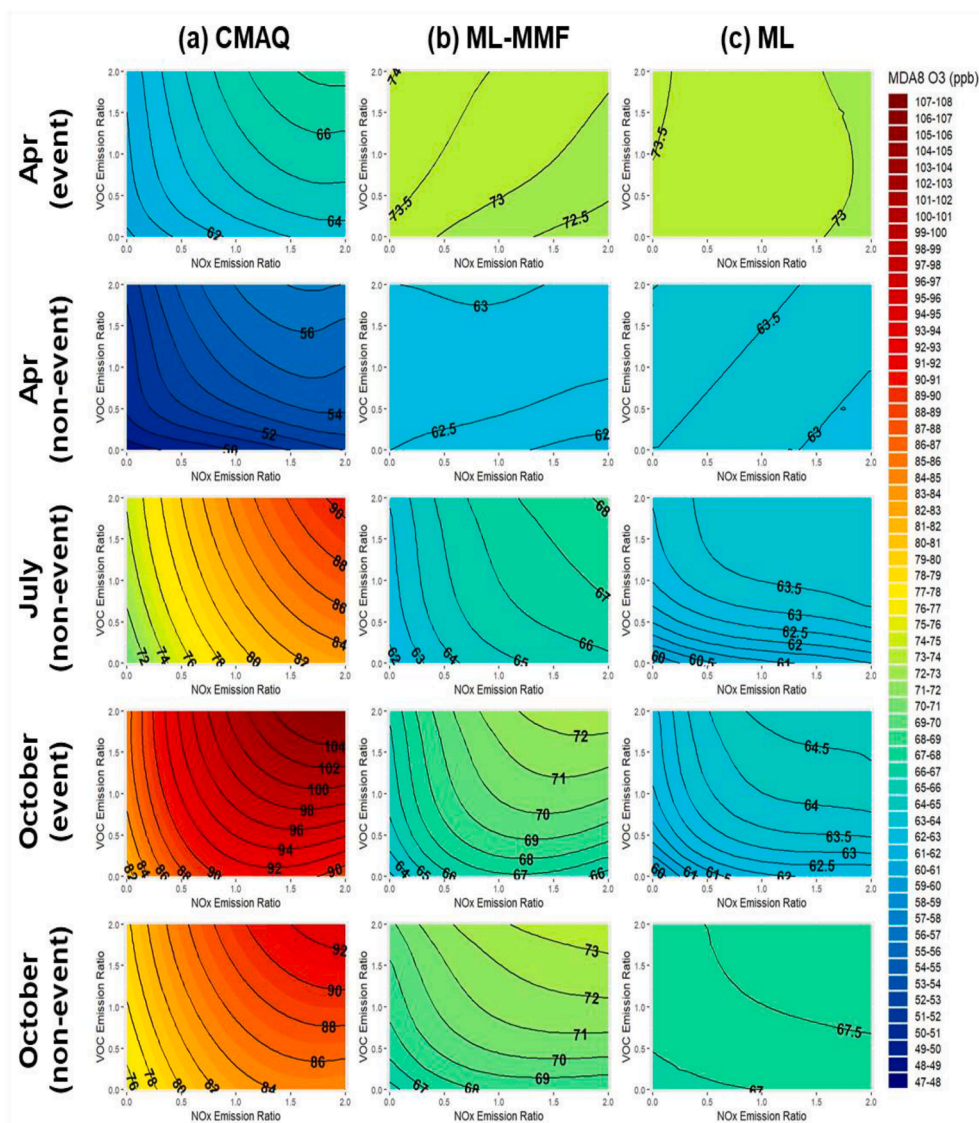


Fig. 5. Monthly combined O<sub>3</sub> isopleths from all air quality regions (NT, CM, CT, YCN, KP, and ET) for the event and non-event days in selected months by using CMAQ, ML-MMF, and ML data.

ppb to 68 ppb. The reduced range of isopleths emphasizes the importance of adjustment by observations, and CMAQ-based RSM results may bias the improved air quality by emission control strategies.

Second, the ML-MMF O<sub>3</sub> isopleths show changed O<sub>3</sub> sensitivity after fusing with observations. For example, CMAQ isopleths in April with a combined regime of NO<sub>x</sub>-limited and VOC-limited trends (Fig. 5(a)) are corrected to VOC-limited in ML-MMF isopleths (Fig. 5(b)) for the event and non-event days. If further comparing CMAQ modeled NO<sub>x</sub> and VOCs with observations (Fig. A9), CMAQ modeled estimations in the benchmark case were found to overestimate NO<sub>x</sub> in the NT region and underestimate VOCs compared with observations for all regions, which means the real O<sub>3</sub>-NO<sub>x</sub>-VOC system should be more VOC-limited and NO<sub>x</sub>-rich.

Third, ML O<sub>3</sub> isopleths present diverse O<sub>3</sub> ranges and distorted NO<sub>x</sub> and VOC regimes compared with ML-MMF O<sub>3</sub> isopleths, which implies ML without the support from CMAQ output could provide biased air quality predictions concerning its different O<sub>3</sub> response to emissions, although ML model performances well in the benchmark case. For example, for the non-event days of the NT region in April (Fig. A4), the ML isopleth identifies an obvious VOC-limited trend, but the ML-MMF isopleth shows a combined NO<sub>x</sub>-limited and VOC-limited trend. For the other months, the ML O<sub>3</sub> isopleths also present a lower O<sub>3</sub>

concentration level and narrower range, even though the trend may be similar to the ML-MMF O<sub>3</sub> isopleths.

The disparity between ML-MMF and ML O<sub>3</sub> isopleths implies that previous air quality forecasting studies (Ausati and Amanollahi, 2016; Song et al., 2015; Zhou et al., 2019) without involving CTM results and only using historical observations or other auxiliary data to predict future air quality may be potentially biased, even though the ML models met statistical requirements. Because the ML model could deviate the pollutant's responses to its precursor emissions from the real responses of air pollutant concentrations due to its lack of physical and chemical basis. One major explanation is the lower variable importance priority of emission variables in the ML model. The variable importance priority of emission variables is relatively lower than other variables like meteorological variables (Fig. A11), so the model missed the link between O<sub>3</sub> responses with NO<sub>x</sub> and VOC emissions. The lower variable importance priority is because most emission inventories are constructed by top-down methodology with routine monthly and weekly variation, the regular temporal variation of emission data has limited capability to explain dramatically varied pollutants. On the contrary, CMAQ-modeled O<sub>3</sub> has a higher importance priority in the ML-MMF model (Fig. A11), so its nonlinear information to NO<sub>x</sub> and VOC emissions still remains in ML-MMF outputs. Another potential reason is that the ML model only learns



grid-level information without considering the air pollutants transported and reacted between grids, so the ML predictions are unable to reflect the accumulated pollutants from the upwind grids or air quality regions, even adding zonal emissions as independent variables cannot improve the performance. Moreover, the limited sample size from observations could be not large enough for the ML model, because O<sub>3</sub> exceedance (>60 ppb) does not always occur around the monitoring stations within limited exceedance days (Fig. A10). Therefore, the ML model under certain seasons (e.g. non-event days in April) has limited data to construct a robust model. In summary, we suggest not using ML O<sub>3</sub> isopleths for emission control strategies.

Fourth, transboundary pollution from upwind may change O<sub>3</sub> sensitivity to local NO<sub>x</sub> and VOC emissions. In the NT region, which is the first region suffering from transboundary pollution from mainland China, the ML-MMF isopleths (Fig. A4) of non-event days in April remain a combined regime of NO<sub>x</sub>-limited and VOC-limited, but for

event days in April, the isopleth changes to serious VOC-limited. Because the northeast monsoon carries high NO<sub>x</sub> concentrations from the upwind, the NT region turns to a NO<sub>x</sub>-rich atmosphere and local O<sub>3</sub> response becomes more sensitive to VOC emissions. Even though there is no local NO<sub>x</sub> emission (emission ratio = 0), transboundary NO<sub>x</sub> still can support local O<sub>3</sub> formation. Moreover, if taking out boundary conditions and local meteorology during the fusion process (Fig. A12), boundary conditions and local meteorology play a significant role that changing O<sub>3</sub> sensitivity, which also proves the significant impact of transboundary pollution. The changed O<sub>3</sub> sensitivity affected by outside pollution from the upwind was also reported in California, where wildfires in the late summer would emit large amounts of VOCs that can be transported to urban areas and significantly change the urban O<sub>3</sub> sensitivity (Wu et al., 2022).

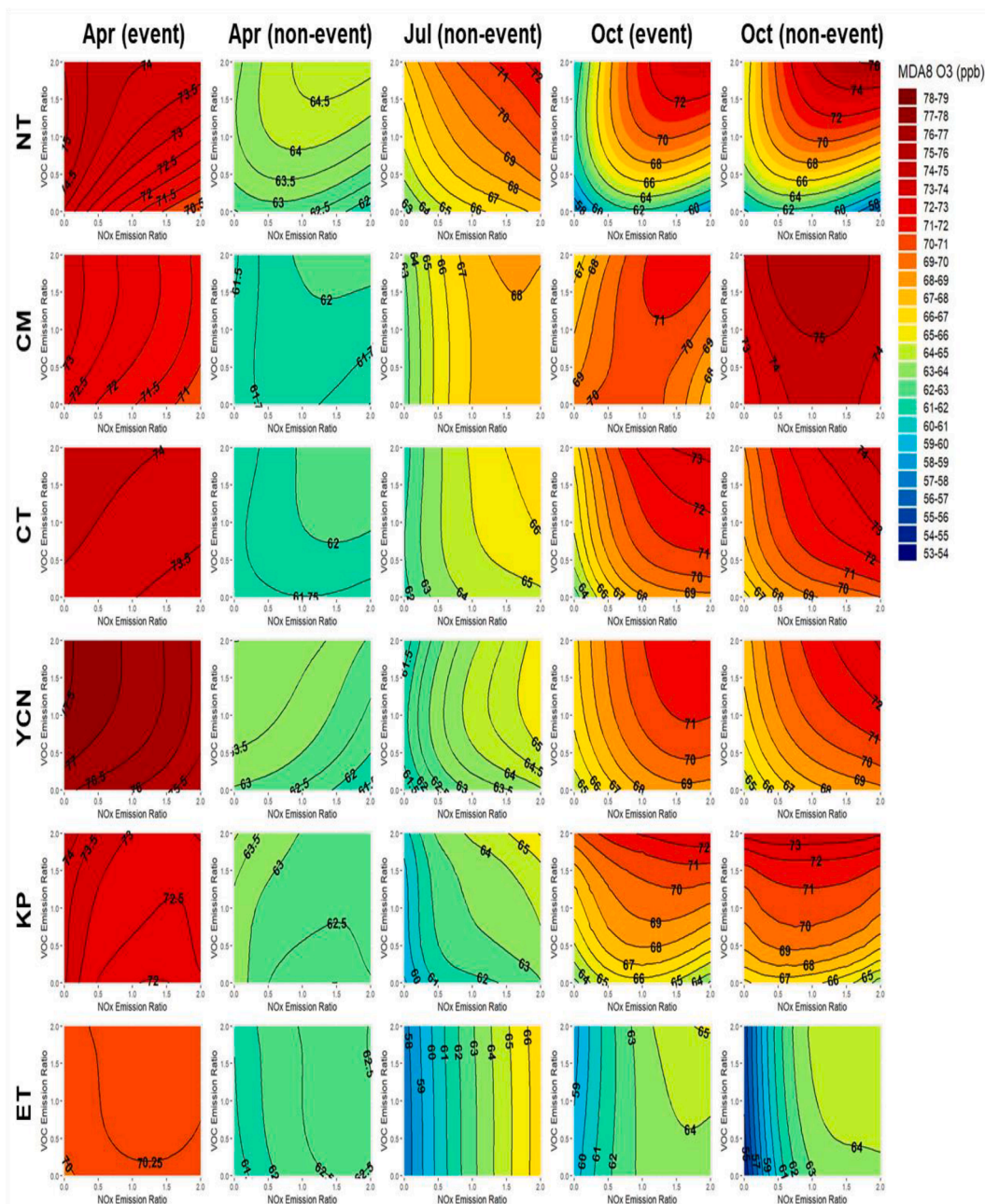


Fig. 6. Monthly O<sub>3</sub> isopleths for event and non-event days in selected months based on ML-MMF data.



### 3.5. Suggested emission control preference

Monthly O<sub>3</sub> isopleths for the event and non-event days based on ML-MMF data are shown in Fig. 6. The O<sub>3</sub> isopleths in April show VOC-limited trends for both the event and non-event days for all western regions, but the range reveals a limited improved effort by reducing VOC emission, which is around 0.25–4.5 ppb. Transboundary NO<sub>x</sub> could also accumulate for days and affect O<sub>3</sub> sensitivity to local emissions on non-event days. In July, controlling NO<sub>x</sub> emissions is a more effective way to lower O<sub>3</sub> concentrations in the CM, CT, YCN, and ET regions, and the NT and KP regions present combined NO<sub>x</sub>-limited and VOC-limited regimes. The more VOC-limited trend in the NT and KP regions may be due to more population and vehicle emissions in these regions.

In October, O<sub>3</sub> isopleths are similar between the event and non-event days, because transboundary pollution in October mostly carries PM and SO<sub>2</sub>, which are not related to the formation of O<sub>3</sub>. NT, CT, and YCN regions have a combined NO<sub>x</sub>-limited and VOC-limited trend, and the potential improvement can be significant (58–76 ppb) in the NT region. The more VOC-limited trend in the CM and KP regions could be due to the impact of local geography. Owing to the many plateaus and hills in the CM and KP regions, the wind would be slowed down by higher terrain roughness, causing the accumulation of pollutants like NO<sub>x</sub> from the upwind regions. In the ET region, the O<sub>3</sub> isopleths suggest an obvious NO<sub>x</sub>-limited trend for all months. In summary, most of the regions are VOC-limited in April and October but NO<sub>x</sub>-limited in July. The suggestion of controlling VOC emissions in fall and winter is also coherent with the other study in Taiwan (Chen et al., 2021).

### 3.6. Limitations and future works

This study still has some limitations. First, the ML-MMF model needs to rely on enough monitoring stations (sample size) and good-quality of emission inventory to build a robust model. Multiple monitoring stations can explain the variance of space-related variables like emissions, meteorology, and land uses between sites and provide enough statistical power, and the emission inventory should reflect the various emissions around the monitoring stations. In our case, the employed TEDS emission inventory is updated every three years with finer to 1 km spatial resolution, thus it can reflect emissions around stations within space and time. For those regions or countries without enough monitoring stations, using remote sensing data from satellites may be an alternative way to obtain observations, but satellite data could be still biased by cloud and vertical column densities (Wu et al., 2022). For regions or countries without proper emission inventory, using a global emission database like ECLIPSE (Stohl et al., 2015) could be a potential solution to provide emission data around stations, but the spatial distribution method from national emissions needs careful assessment.

Second, the ML-MMF RSM results or O<sub>3</sub> isopleths are hardly validated by observations. Even though the ML-MMF O<sub>3</sub> isopleths with baseline emission ratio (emission ratio = 1) were validated by observations and have a satisfactory performance (Fig. A3, Mean NE = 2–6 %), O<sub>3</sub> responses under extreme emission scenarios are hardly validated by observations. Because there are no observations to validate the estimations under extreme emission scenarios like zero (emission ratio = 0) or double (emission ratio = 2) anthropogenic emissions. Thus, the margins of O<sub>3</sub> isopleths still remained potential uncertainty. Further study to assess the uncertainty of O<sub>3</sub> isopleths should be investigated in the future.

## 4. Conclusion

ML models become mainstream in atmospheric science because of more available real-time monitoring data and measurements. But its black-box modeling process, failure to involve physical and chemical mechanisms, and weak cooperation with the existing numerical models lower its stringency and persuasion. Except for developing more

sophisticated model designs to increase predicting accuracy, the interpretability and explainability of predictions should be evaluated as well. In this study, the capability of the ML model to serve as a bias corrector (ML-MMF output) based on CMAQ modeled results and a forecaster (ML output) was examined and applied to predict O<sub>3</sub> nonlinear responses to anthropogenic NO<sub>x</sub> and VOC emissions. Three types of data were examined for constructing O<sub>3</sub> isopleths by RSM: CMAQ data, ML-MMF data (a bias corrector), and ML data (a forecaster).

Compared with CMAQ predictions ( $r = 0.41–0.80$ ), both ML-MMF ( $r = 0.93–0.94$ ) and ML predictions ( $r = 0.89–0.94$ ) showed significantly improved performance in the benchmark case. While ML-MMF isopleths exhibit O<sub>3</sub> nonlinearity close to actual responses due to their numerical base and observation-based correction, ML isopleths present different O<sub>3</sub> ranges and distorted NO<sub>x</sub> and VOC-limited regimes compared with ML-MMF O<sub>3</sub> isopleths even though the ML model meets the statistical requirement in the benchmark case. Without involving CMAQ results, the ML model presents biased predictions concerning its different O<sub>3</sub> responses to NO<sub>x</sub> and VOC emission ratios. It also implies that only using historical observations or other auxiliary data without support from CMAQ or other CTMs to forecast the air quality could mislead the future trend. Meanwhile, after being corrected by observations, ML-MMF data present changed O<sub>3</sub> sensitivity compared with CMAQ data. The corrected O<sub>3</sub> isopleths emphasize the impact of transboundary pollution from mainland China on the local O<sub>3</sub> sensitivity, which transboundary NO<sub>x</sub> in April would make all air quality regions in Taiwan more sensitive to local VOC emissions and limit the potential effort by reducing local VOC emissions.

It is advisable for future ML applications in atmospheric science like forecasting or bias correction to provide interpretability and explainability while requiring modeling performance. Failing to interpret the interaction between predicted air quality, emissions, and environmental factors may mislead controlled targets and air quality policies. Assessment with interpretable physical and chemical mechanisms and constructing a statistically robust ML model should be equally important.

### CRediT authorship contribution statement

**Cheng-Pin Kuo:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Joshua S. Fu:** Conceptualization, Investigation, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

Data will be made available on request.

### Acknowledgment

We greatly thank National Energy Research Scientific Computing Center (NERSC), National Center for Atmospheric Research (NCAR), and Oak Ridge Leadership Computing Facility (OLCF) at Oak Ridge National Laboratory (ORNL) for providing the computational resources used in this research. We would also like to acknowledge Professor Hsin-Chih Lai in the Department of Green Energy and Environmental Resources, Chang Jung Christian University for providing TEDS, WRF, and land-use data for CMAQ modeling. This study was supported by the grants from Taiwan EPA (110A252, 2021).

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2023.107969>.

## References

- Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). pp. 1–6. doi:10.1109/ICEngTechnol.2017.8308186.
- Appel, K.W., Pouliot, G.A., Simon, H., Sarwar, G., Pye, H.O.T., Napelenok, S.L., Akhtar, F., Roselle, S.J., 2013. Evaluation of dust and trace metal estimates from the Community Multiscale Air Quality (CMAQ) model version 5.0. *Geosci. Model Dev.* 6, 883–899. <https://doi.org/10.5194/gmd-6-883-2013>.
- Apte, J.S., Brauer, M., Cohen, A.J., Ezzati, M., Pope, C.A., 2018. Ambient PM<sub>2.5</sub> reduces global and regional life expectancy. *Environ. Sci. Technol. Lett.* 5, 546–551. <https://doi.org/10.1021/acs.estlett.8b00360>.
- Arnold, J.R., Dennis, R.L., 2006. Testing CMAQ chemistry sensitivities in base case and emissions control runs at SEARCH and SOS99 surface sites in the southeastern US. *Atmos. Environ.* 40, 5027–5040. <https://doi.org/10.1016/j.atmosenv.2005.05.055>.
- Ausati, S., Amanollahi, J., 2016. Assessing the accuracy of ANFIS, EEMD-GRNN, PCR, and MLR models in predicting PM<sub>2.5</sub>. *Atmos. Environ.* 142, 465–474. <https://doi.org/10.1016/j.atmosenv.2016.08.007>.
- Chang, K.H., Chen, T.F., Huang, H.C., 2005. Estimation of biogenic volatile organic compounds emissions in subtropical island - Taiwan. *Sci. Total Environ.* 346, 184–199. <https://doi.org/10.1016/j.scitotenv.2004.11.022>.
- Che, W., Zheng, J., Wang, S., Zhong, L., Lau, A., 2011. Assessment of motor vehicle emission control policies using Model-3/CMAQ model for the Pearl River Delta region, China. *Atmos. Environ.* 45, 1740–1751. <https://doi.org/10.1016/j.atmosenv.2010.12.050>.
- Chen, T.-F., Tsai, C.-Y., Chen, C.-H., Chang, K.-H., 2021. Effect of long-range transport from changing emission on ozone-NOx-VOC sensitivity: implication of control. *J. Innov. Technol.* 3, 39–49.
- Deng, W., Huang, Z., Zhang, J., Xu, J., 2001. Random forests. In: *The Elements of Statistical Learning*. Springer, New York, NY, pp. 587–604. <https://doi.org/10.1109/ICCECE51280.2021.9342376>.
- Dong, X., Fu, J.S., Huang, K., Zhu, Q., Tipton, M., 2019. Regional climate effects of biomass burning and dust in east asia: evidence from modeling and observation. *Geophys. Res. Lett.* 46, 11490–11499. <https://doi.org/10.1029/2019GL083894>.
- Fu, J.S., Carmichael, G.R., Dentener, F., Aas, W., Andersson, C., Barrie, L.A., Cole, A., Galy-Lacaux, C., Geddes, J., Itahashi, S., Kanakidou, M., Labrador, L., Paulot, F., Schwede, D., Tan, J., Vet, R., 2022. Improving estimates of sulfur, nitrogen, and ozone total deposition through multi-model and measurement-model fusion approaches. *Environ. Sci. Technol.* 56, 2134–2142. <https://doi.org/10.1021/acs.est.1c05929>.
- Geng, G., Meng, X., He, K., Liu, Y., 2020. Random forest models for PM<sub>2.5</sub> speciation concentrations using MISR fractional AODs. *Environ. Res. Lett.* 15 <https://doi.org/10.1088/1748-9326/ab76df>.
- Gipson, G.L., Freas, W.P., Kelly, R.F., Meyer, E.L., 1980. Guideline for use of city-specific EKMA in preparing ozone SIPs. Draft report. United States.
- Huang, K., Fu, J.S., Gao, Y., Dong, X., Zhuang, G., Lin, Y., 2014. Role of sectoral and multi-pollutant emission control strategies in improving atmospheric visibility in the Yangtze River Delta, China. *Environ. Pollut.* 184, 426–434. <https://doi.org/10.1016/j.envpol.2013.09.029>.
- Huang, K., Fu, J.S., Lin, N.H., Wang, S.H., Dong, X., Wang, G., 2019. Superposition of Gobi dust and southeast Asian biomass burning: the effect of multisource long-range transport on aerosol optical properties and regional meteorology modification. *J. Geophys. Res. Atmos.* 124, 9464–9483. <https://doi.org/10.1029/2018JD030241>.
- Kang, G.K., Gao, J.Z., Chiao, S., Lu, S., Xie, G., 2018. Air Quality Prediction: Big Data and Machine Learning Approaches. *Int. J. Environ. Sci. Dev.* 9, 8–16. doi:10.18178/ijesd.2018.9.1.1066.
- Kelly, J.T., Jang, C., Zhu, Y., Long, S., Xing, J., Wang, S., Murphy, B.N., Pye, H.O.T., 2021. Predicting the nonlinear response of PM<sub>2.5</sub> and ozone to precursor emission changes with a response surface model. *Atmosphere (Basel)* 12. <https://doi.org/10.3390/atmos12081044>.
- Kramer, O., 2013. Dimensionality Reduction with Unsupervised Nearest Neighbors. *Intell. Syst. Ref. Lib.* <https://doi.org/10.1007/978-3-642-38652-7>.
- Kuo, C.P., Fu, J.S., Wu, P.C., Cheng, T.J., Chiu, T.Y., Huang, C.S., Wu, C.F., Lai, L.W., Lai, H.C., Liang, C.K., 2021. Quantifying spatial heterogeneity of vulnerability to short-term PM<sub>2.5</sub> exposure with data fusion framework. *Environ. Pollut.* 285, 117266. doi:10.1016/j.envpol.2021.117266.
- Lai, H.C., Lin, M.C., 2020. Characteristics of the upstream flow patterns during PM<sub>2.5</sub> pollution events over a complex island topography. *Atmos. Environ.* 227, 117418 <https://doi.org/10.1016/j.atmosenv.2020.117418>.
- Li, J., Dai, Y., Zhu, Y., Tang, X., Wang, S., Xing, J., Zhao, B., Fan, S., Long, S., Fang, T., 2022. Improvements of response surface modeling with self-adaptive machine learning method for PM<sub>2.5</sub> and O<sub>3</sub> predictions. *J. Environ. Manage.* 303, 114210 <https://doi.org/10.1016/j.jenvman.2021.114210>.
- Loh, W., 2008. Classification and regression tree methods. *Encycl. Stat. Qual. Reliab.* 1, 315–323. <https://doi.org/10.5860/choice.45-6515>.
- Lu, H., Xie, M., Liu, X., Liu, B., Jiang, M., Gao, Y., Zhao, X., 2021. Adjusting prediction of ozone concentration based on CMAQ model and machine learning methods in Sichuan-Chongqing region, China. *Atmos. Pollut. Res.* 12, 101066 <https://doi.org/10.1016/j.apr.2021.101066>.
- Lu, X., Yuan, D., Chen, Y., Fung, J.C.H., Li, W., Lau, A.K.H., 2020. Estimations of long-term nss-SO<sub>4</sub>-2 and NO<sub>3</sub>-wet depositions over East Asia by use of ensemble machine-learning method. *Environ. Sci. Technol.* 54, 11118–11126. <https://doi.org/10.1021/acs.est.0c01068>.
- Luo, H., Zhao, K., Yuan, Z., Yang, L., Zheng, J., Huang, Z., Huang, X., 2021. Emission source-based ozone isopleth and isosurface diagrams and their significance in ozone pollution control strategies. *J. Environ. Sci. (China)* 105, 138–149. <https://doi.org/10.1016/j.jes.2020.12.033>.
- Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. *Front. Neurobot.* 7 <https://doi.org/10.3389/fnbot.2013.00021>.
- Requia, W.J., Di, Q., Silvern, R., Kelly, J.T., Koutrakis, P., Mickley, L.J., Sulprizio, M.P., Amini, H., Shi, L., Schwartz, J., 2020. An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States. *Environ. Sci. Technol.* 54, 11037–11047. <https://doi.org/10.1021/acs.est.0c01791>.
- Sarwar, G., Luecken, D., Yarwood, G., Whitten, G.Z., Carter, W.P.L., 2008. Impact of an updated carbon bond mechanism on predictions from the CMAQ modeling system: preliminary assessment. *J. Appl. Meteorol. Climatol.* 47, 3–14. <https://doi.org/10.1175/2007JAMC1393.1>.
- Sayed, A., Lops, Y., Choi, Y., Jung, J., Salman, A.K., 2021. Bias correcting and extending the PM forecast by CMAQ up to 7 days using deep convolutional neural networks. *Atmos. Environ.* 253, 118376 <https://doi.org/10.1016/j.atmosenv.2021.118376>.
- Song, Y., Qin, S., Qu, J., Liu, F., 2015. The forecasting research of early warning systems for atmospheric pollutants: a case in Yangtze River Delta region. *Atmos. Environ.* 118, 58–69. <https://doi.org/10.1016/j.atmosenv.2015.06.032>.
- Stohl, A., Aamaas, B., Amann, M., Baker, L.H., Bellouin, N., Bernsten, T.K., Boucher, O., Cherian, R., Collins, W., Daskalakis, N., Dusinska, M., Eckhardt, S., Fuglested, J.S., Harju, M., Heyes, C., Hodnebrog, H., Im, U., Kanakidou, M., Klimont, Z., Kupiainen, K., Law, K.S., Lund, M.T., Maas, R., MacIntosh, C.R., Myhre, G., Myriokefalitakis, S., Olivieri, D., Quaas, J., Quennehen, B., Raut, J.C., Rumbold, S.T., Samset, B.H., Schulz, M., Seland, Shine, K.P., Skeie, R.B., Wang, S., Yttri, K.E., Zhu, T., 2015. Evaluating the climate and air quality impacts of short-lived pollutants. *Atmos. Chem. Phys.* 15, 10529–10566. doi:10.5194/acp-15-10529-2015.
- Tai, A.P.K., Martin, M.V., 2017. Impacts of ozone air pollution and temperature extremes on crop yields: spatial variability, adaptation and implications for future food security. *Atmos. Environ.* 169, 11–21. <https://doi.org/10.1016/j.atmosenv.2017.09.002>.
- U.S. EPA, 2018. SMOKE v4. 6 User Manual.
- USEPA, 2006. Technical Support Document for the Proposed PM NAAQS Rule 48.
- Wang, S., Xing, J., Jang, C., Zhu, Y., Fu, J.S., Hao, J., 2011. Impact assessment of ammonia emissions on inorganic aerosols in East China using response surface modeling technique. *Environ. Sci. Technol.* 45, 9293–9300. <https://doi.org/10.1021/es2022347>.
- Ware, L.B., Zhao, Z., Koyama, T., May, A.K., Matthay, M.A., Lurmann, F.W., Balmes, J.R., Calfee, C.S., 2016. Long-term ozone exposure increases the risk of developing the acute respiratory distress syndrome. *Am. J. Respir. Crit. Care Med.* 193, 1143–1150. <https://doi.org/10.1164/rccm.201507-1418OC>.
- World Health Organization, 2005. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005: 1–21. doi:10.1016/0004-6981(88)90109-6.
- Wu, S., Lee, H.J., Anderson, A., Liu, S., Kuwayama, T., Seinfeld, J.H., Kleeman, M.J., 2022. Direct measurements of ozone response to emissions perturbations in California. *Atmos. Chem. Phys.* 22, 4929–4949. <https://doi.org/10.5194/acp-22-4929-2022>.
- Xing, J., Ding, D., Wang, S., Dong, Z., Kelly, J.T., Jang, C., Zhu, Y., Hao, J., 2019. Development and application of observable response indicators for design of an effective ozone and fine-particle pollution control strategy in China. *Atmos. Chem. Phys.* 19, 13627–13646. <https://doi.org/10.5194/acp-19-13627-2019>.
- Xing, J., Ding, D., Wang, S., Zhao, B., Jang, C., Wu, W., Zhang, F., Zhu, Y., Hao, J., 2018. Quantification of the enhanced effectiveness of NOx control from simultaneous reductions of VOC and NH<sub>3</sub> for reducing air pollution in the Beijing-Tianjin-Hebei region. *China. Atmos. Chem. Phys.* 18, 7799–7814. <https://doi.org/10.5194/acp-18-7799-2018>.
- 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. *Atmos. Chem. Phys.* 11, 5027–5044. <https://doi.org/10.5194/acp-11-5027-2011>.
- Xing, J., Zheng, S., Li, S., Huang, L., Wang, X., Kelly, J.T., Wang, S., Liu, C., Jang, C., Zhu, Y., Zhang, J., Bian, J., Liu, T.Y., Hao, J., 2022. Mimicking atmospheric photochemical modeling with a deep neural network. *Atmos. Res.* 265, 105919 <https://doi.org/10.1016/j.atmosres.2021.105919>.
- Zarnetske, P.L., Gurevitch, J., Franklin, J., Groffman, P.M., Harrison, C.S., Hellmann, J. J., Hoffman, F.M., Kothari, S., Robock, A., Tilmes, S., Visioni, D., Wu, J., Xia, L., Yang, C.E., 2021. Potential ecological impacts of climate intervention by reflecting sunlight to cool Earth. *Proc. Natl. Acad. Sci. U.S.A.* 118, 1–11. <https://doi.org/10.1073/pnas.1921854118>.
- 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. *Geosci. Model Dev.* 8, 115–128. <https://doi.org/10.5194/gmd-8-115-2015>.

Zheng, L., Lin, R., Wang, X., Chen, W., 2021. The development and application of machine learning in atmospheric environment studies. *Rem. Sens.* 13 <https://doi.org/10.3390/rs13234839>.

Zhou, Y., Chang, F.J., Chang, L.C., Kao, I.F., Wang, Y.S., 2019. Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *J. Clean. Prod.* 209, 134–145. <https://doi.org/10.1016/j.jclepro.2018.10.243>.

Zhu, Y., Lao, Y., Jang, C., Lin, C.-J., Xing, J., Wang, S., Fu, J.S., Deng, S., Xie, J., Long, S., 2015. Development and case study of a science-based software platform to support policy making on air quality. *J. Environ. Sci.* 27, 97–107.