



## Full-coverage mapping and spatiotemporal variations of ground-level ozone (O<sub>3</sub>) pollution from 2013 to 2020 across China

Jing Wei<sup>a,b,\*</sup>, Zhanqing Li<sup>a,\*</sup>, Ke Li<sup>c</sup>, Russell R. Dickerson<sup>a</sup>, Rachel T. Pinker<sup>a</sup>, Jun Wang<sup>b</sup>, Xiong Liu<sup>d</sup>, Lin Sun<sup>e</sup>, Wenhao Xue<sup>f</sup>, Maureen Cribb<sup>a</sup>

<sup>a</sup> Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

<sup>b</sup> Department of Chemical and Biochemical Engineering, Iowa Technology Institute, and Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA, USA

<sup>c</sup> Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China

<sup>d</sup> Atomic and Molecular Physics Division, Harvard Smithsonian Center for Astrophysics, Cambridge, MA, USA

<sup>e</sup> College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China

<sup>f</sup> School of Economics, Qingdao University, Qingdao, China

### ARTICLE INFO

Editor: Dr. Menghua Wang

#### Keywords:

Ozone  
Air pollution  
Ensemble learning  
COVID-19  
China

### ABSTRACT

Ozone (O<sub>3</sub>) is an important trace and greenhouse gas in the atmosphere, posing a threat to the ecological environment and human health at the ground level. Large-scale and long-term studies of O<sub>3</sub> pollution in China are few due to highly limited direct ground and satellite measurements. This study offers a new perspective to estimate ground-level O<sub>3</sub> from solar radiation intensity and surface temperature by employing an extended ensemble learning of the space-time extremely randomized trees (STET) model, together with ground-based observations, remote sensing products, atmospheric reanalysis, and an emission inventory. A full-coverage (100%), high-resolution (10 km) and high-quality daily maximum 8-h average (MDA8) ground-level O<sub>3</sub> dataset covering China (called ChinaHighO<sub>3</sub>) from 2013 to 2020 was generated. Our MDA8 O<sub>3</sub> estimates (predictions) are reliable, with an average out-of-sample (out-of-station) coefficient of determination of 0.87 (0.80) and root-mean-square error of 17.10 (21.10) μg/m<sup>3</sup> in China. The unique advantage of the full coverage of our dataset allowed us to accurately capture a short-term severe O<sub>3</sub> pollution exposure event that took place from 23 April to 8 May in 2020. Also, a rapid increase and recovery of O<sub>3</sub> concentrations associated with variations in anthropogenic emissions were seen during and after the COVID-19 lockdown, respectively. Trends in O<sub>3</sub> concentration showed an average growth rate of 2.49 μg/m<sup>3</sup>/yr ( $p < 0.001$ ) from 2013 to 2020, along with the continuous expansion of polluted areas exceeding the daily O<sub>3</sub> standard (i.e., MDA8 O<sub>3</sub> = 160 μg/m<sup>3</sup>). Summertime O<sub>3</sub> concentrations and the probability of occurrence of daily O<sub>3</sub> pollution have significantly increased since 2015, especially in the North China Plain and the main air pollution transmission belt (i.e., the “2 + 26” cities). However, a decline in both was seen in 2020, mainly due to the coordinated control of air pollution and ongoing COVID-19 effects. This carefully vetted and smoothed dataset is valuable for studies on air pollution and environmental health in China.

### 1. Introduction

Ozone (O<sub>3</sub>) is an important atmospheric trace gas, where O<sub>3</sub> in the stratosphere plays a crucial role in absorbing ultraviolet rays, protecting the environment and humans. Tropospheric O<sub>3</sub> (< 12 km above the ground) is mainly produced by anthropogenic activities, affecting

radiative forcing at a global scale with implications on climate change (Checa-Garcia et al., 2018; Chen et al., 2007; Gaudel et al., 2018; Shindell et al., 2013; Sinha and Toumi, 1997). Exposure to high surface O<sub>3</sub> levels is highly related to increased human health risks, including cardiovascular and respiratory diseases (Bell et al., 2004; Lim et al., 2019; Turner et al., 2015). It also affects the ecosystem and agricultural

\* Corresponding authors at: Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA.

E-mail addresses: [weijing\\_rs@163.com](mailto:weijing_rs@163.com) (J. Wei), [zli@atmos.umd.edu](mailto:zli@atmos.umd.edu) (Z. Li).

<https://doi.org/10.1016/j.rse.2021.112775>

Received 23 May 2021; Received in revised form 18 September 2021; Accepted 25 October 2021

Available online 11 November 2021

0034-4257/© 2021 Elsevier Inc. All rights reserved.

production, e.g., inhibiting plant growth, promoting leaf senescence, and affecting crop yields (Ainsworth et al., 2012; Mills et al., 2018; Rai and Agrawal, 2012; Sitch et al., 2007).

Since the middle of the twentieth century, many countries around the world have observed tropospheric and ground-level O<sub>3</sub>. In 2013, the Chinese Ministry of Environment and Ecology (MEE) established a national air quality observation network to monitor real-time O<sub>3</sub>, particulate matter (PM), and other near-surface air pollutants (MEE, 2018). However, the construction and maintenance of ground networks require substantial manpower and material resources. As such, monitoring stations are sparsely distributed. Satellite remote sensing can make up for such a deficiency by providing spatially continuous atmospheric O<sub>3</sub> distributions. The Ozone Monitoring Instrument (OMI) on the Aura satellite, launched in 2004, provides a variety of widely used daily, global-coverage trace gas products, e.g., O<sub>3</sub>, nitrogen dioxide (NO<sub>2</sub>), and sulfur dioxide (SO<sub>2</sub>). Existing techniques from space mainly provide the total column O<sub>3</sub>, tropospheric O<sub>3</sub>, and ozone profiles at different vertical ranges (Liu et al., 2010). Near-surface O<sub>3</sub> typically accounts for only a few percent of total column O<sub>3</sub>, and the retrieval sensitivity to near-surface O<sub>3</sub> from ultraviolet measurements is limited. In some cases, tropospheric total column amounts can be helpful for understanding global- and regional-scale features, but values for O<sub>3</sub> in the planetary boundary layer are challenging to obtain and at exposure heights (~2 m), even more so. It is thus particularly difficult to extract near-surface O<sub>3</sub> concentrations from satellite measurements.

In recent years, much effort has been made to estimate near-surface O<sub>3</sub> concentrations using three main methodologies: chemical transport models, statistical models, and artificial intelligence. Chemical transport models mainly use mature models, e.g., WRF-Chem, CMAQ, and GEOS-Chem, to simulate O<sub>3</sub> at the ground level by considering chemical reactions and the transport of air pollutants (Di et al., 2017; Hu et al., 2016; Qiao et al., 2019; Wang et al., 2016a; Wang et al., 2015). Statistical models fit the relationships between the measured air pollution and their potential influential factors (e.g., satellite retrievals, precursors, and meteorology) by applying different regression methods, such as Land Use Regression (LUR; Beelen et al., 2009; Huang et al., 2017; Kerckhoffs et al., 2015; Son et al., 2018), Bayesian maximum entropy (BME; Adam-Poupard et al., 2014; Chen et al., 2020), the generalized additive model (GAM; Li et al., 2020b), and geographically weighted regression (GWR, Zhang et al., 2020). Artificial intelligence, i.e., machine and deep learning, allows for obtaining more accurate parameter estimates by mining valuable information from big data using different methods, e.g., neural network (Di et al., 2017), random forest (RF; Li et al., 2020b; Zhan et al., 2018), and eXtreme Gradient Boosting (XGBoost; Li et al., 2020b).

In general, chemical/numerical methods can provide high spatiotemporal coverage of near-surface O<sub>3</sub> simulations but are computationally intensive. Predictions with any chemical mechanism are sensitive in nonlinear ways to emissions and meteorology. Statistical models have been widely adopted because of their simplicity and rapidity, but they are sensitive to outliers and easily affected by collinear variables, leading to poor estimates. Artificial intelligence has become very popular recently due to its strong data-mining ability, but they are always directly applied and neglect the spatiotemporal heterogeneity of air pollution. Most past related studies are limited by input data sources, e.g., satellite total column gas products (e.g., OMI/Aura) with missing values, and meteorological products, e.g., National Centers for Environmental Prediction (NCEP), Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), and ERA-Interim, at low spatiotemporal resolutions (e.g., 3–6 h, 0.25°–0.625°).

Over the years, PM pollution has decreased significantly due to implemented environmental protection and control measures (Wang et al., 2020; Wei et al., 2021a–c; Xue et al., 2021; Zhang et al., 2019). By contrast, surface O<sub>3</sub> pollution has increased in China (Li et al., 2019; Lu et al., 2018; Wang et al., 2016b; Wang et al., 2020), creating a major public health concern (Shen et al., 2019). Compared with PM studies,

research on ground-level O<sub>3</sub> in China is more meager. Therefore, aimed at addressing the above problems, according to the idea of ensemble learning and considering the spatiotemporal variations in O<sub>3</sub> pollution, we extended a space-time extremely randomized trees (STET) model to derive daily ground-level O<sub>3</sub> concentrations with full spatial coverage at a resolution of 10 km from 2013 to 2020 across China. We also validated our O<sub>3</sub> estimates at different spatiotemporal scales and investigated the variations in daily and multi-year O<sub>3</sub> pollution via time series analyses across China.

## 2. Materials and methods

### 2.1. Data sources

#### 2.1.1. Surface O<sub>3</sub> observations

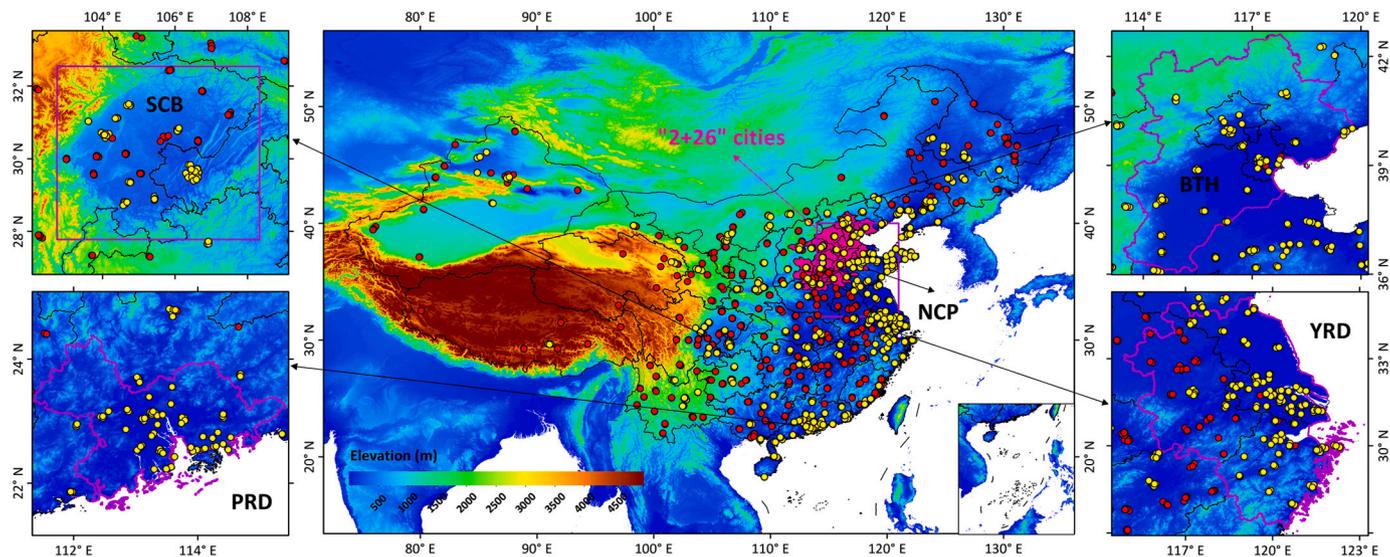
Used are hourly ground-based O<sub>3</sub> concentrations [in µg/m<sup>3</sup> at standard conditions (273 K, 1013 hPa)] collected by MEE across mainland China starting from an initial ~940 monitoring stations in 2013 and ending with ~1630 stations by 2020 (Fig. 1). We first removed invalid values and abnormal values due to instrument calibration issues. More importantly, since 31 August 2018, the reference state of gas observations was changed from the standard condition (i.e., 273 K and 1013 hPa) to room temperature and pressure (i.e., 298 K and 1013 hPa). The new measurements of O<sub>3</sub> concentrations (in µg/m<sup>3</sup>) were thus correspondingly rescaled by a factor of 1.09375 (MEE, 2018). For data presented here, 1 µg/m<sup>3</sup> is equivalent to 0.467 ppbv. Additionally, we averaged maximum O<sub>3</sub> concentrations over eight hours in a day to obtain MDA8 O<sub>3</sub> values at each station in China for each year from 1 January 2013 to 31 December 2020.

#### 2.1.2. Potential factors affecting surface O<sub>3</sub>

Surface O<sub>3</sub>, a secondary air pollutant, is the characteristic product of complex photochemical reactions affected by numerous natural and human factors. Most satellites (e.g., OMI) provide only total-column or tropospheric O<sub>3</sub> retrievals, rather than lower tropospheric O<sub>3</sub> retrievals where there are large differences in O<sub>3</sub> content. Long-term satellite O<sub>3</sub> products with high spatial resolutions are rarely available, and those existing satellite retrievals have numerous missing values. In our study, we provide a new approach for estimating high-resolution surface O<sub>3</sub> concentrations with full coverage using two crucial meteorological parameters, namely, solar radiation intensity and surface temperature (Bloomer et al., 2009; Lee et al., 2014; Li et al., 2020a). Thus, available surface solar radiation downwards (or downward shortwave radiation, DSR) and air temperature (TEM) measurements are used as main predictors for the ground-level O<sub>3</sub> estimation.

Other meteorological variables can also affect O<sub>3</sub>, e.g., an increase in relative humidity (RH) and surface pressure (SP) can have diverse effects on O<sub>3</sub> concentrations in the lower troposphere (He et al., 2017; Loughner et al., 2011; Taubman et al., 2006). A change in planetary boundary layer height (PBLH) also impacts O<sub>3</sub> pollution (Benish et al., 2020; Dickerson et al., 2007; Goldberg et al., 2014; Ma et al., 2011; Sánchez-Ccoylo et al., 2006). Winds, i.e., horizontal (WU) and vertical (WV) components, can affect the transport of O<sub>3</sub> and produce high O<sub>3</sub> levels in the downwind direction (Benish et al., 2020; Dickerson et al., 2007; Duan et al., 2008; Ma et al., 2011). Precipitation (PRE) and evaporation (ET) can also influence O<sub>3</sub> pollution by affecting mixing and photolysis rates (Dickerson et al., 1997; Meleux et al., 2007). The above nine daily meteorological variables were chosen from the latest released hourly ERA5 reanalysis dataset at a high spatial resolution reaching up to 0.1° × 0.1° (Hersbach et al., 2020). The spatiotemporal resolution of ERA5 reanalysis products is higher than from other atmospheric reanalysis products (e.g., NCEP and MERRA-2) used in previous studies (Di et al., 2017; Li et al., 2020b; Liu et al., 2020; Zhan et al., 2018).

Remote sensing measurements of OMI/Aura total-column O<sub>3</sub> products (Pawan, 2012) were also considered. NO<sub>2</sub> concentrations may have large impacts on O<sub>3</sub>, so OMI/Aura tropospheric NO<sub>2</sub> products were



**Fig. 1.** Spatial distributions of ground  $O_3$  monitoring stations (yellow dots) in 2013 and newly established stations (red dots) in 2020 across China, where the background is surface elevation (m). The black and magenta solid lines represent regional and provincial boundaries in China, respectively, and conglomerations of magenta dots show the locations of the “2 + 26” cities. Leftmost and rightmost smaller panels show four regions of interest in this study: the Sichuan Basin (SCB), the Pearl River Delta (PRD), the Beijing-Tianjin-Hebei (BTH) region, and the Yangtze River Delta (YRD). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

utilized (Krotkov et al., 2019). The LandScan™ product was also selected to provide population distribution (POP) information. Also Moderate Resolution Imaging Spectroradiometer land cover type (LUC) and NDVI products, and Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) data were employed to describe land-use and terrain changes across China. Anthropogenic emissions from fossil fuel combustion, industrial production, and vehicle exhaust are precursors affecting the formation of surface  $O_3$  (Li et al., 2020a). Direct emissions of three main  $O_3$  precursors, i.e., nitrogen oxides ( $NO_x$ ), volatile organic compounds (VOCs), and carbon monoxide (CO), whose concentrations were provided by the Multiresolution Emission Inventory for China (MEIC) (Li et al., 2017a; Zheng et al., 2018), were thus used.

Table S1 summarizes the ground-based, satellite remote sensing, atmospheric reanalysis, and model emission datasets used in this study. Except for meteorological conditions, the spatiotemporal resolutions of other auxiliary data are coarser than our targeted model resolution. The coarser-spatial-resolution variables (e.g., emissions and  $NO_2$ ) show smaller variations in space than meteorological variables. The coarser-temporal-resolution variables (i.e., DEM, LUC, and POP) change little over time. In addition, they are generally less important than the two main predictors (i.e., DSR and TEM) in estimating surface  $O_3$ . Therefore, similar to previous studies (Liu et al., 2020; Wei et al., 2021b; Zhan et al., 2018; Zhang et al., 2020), all finer and coarser-resolution auxiliary data were aggregated and resampled (regridded) to the same spatial resolution of  $0.1^\circ \times 0.1^\circ$  using the bilinear interpolation approach, and the same time interval.

## 2.2. STET modeling

In this study, a STET model was extended for estimating ground-level  $O_3$  concentrations (Wei et al., 2020, 2021a), based on the ensemble learning method called extremely randomized trees (extra-trees, or ERT) (Geurts et al., 2006).

### 2.2.1. Model training

First, all selected factors with potential effects on surface  $O_3$  were input to the ERT model for model training, followed by four main steps:

- 1) A training-and-validation dataset ( $N$ ) was generated by collocating surface  $O_3$  measurements, satellite data, meteorological variables, and model emissions at each surface monitoring station on each day of one year. The entire training dataset was then used to construct each decision tree.
- 2) For each binary tree, a random split ( $S, a$ ) was first generated according to surface  $O_3$  measurements by randomly selecting one arbitrary number ( $a_c$ ) between the maximum ( $a_{max}$ ) and minimum ( $a_{min}$ ) values. The training samples were next randomly assigned to two branches.
- 3) All auxiliary feature attributes ( $a_1, \dots, a_k$ ) in the node were traversed to get bifurcation values ( $s_1, \dots, s_k$ ) for all feature attributes, based on the Gini index (GI; Jiang et al., 2009). The best split ( $s^*$ ) was then determined when the scoring function was satisfied:  $score(s^*, S) = \max[Score(s_i, S)]$  (Geurts et al., 2006).
- 4) A decision tree was established using the classification-and-regression-trees (CART) algorithm (Breiman et al., 1984). Thousands of decision trees were constructed by repeating the above steps. Last, all weak classifiers were combined to form a strong classifier, i.e., ERT, allowing for parallel processing.

Fig. S1 shows a basic diagram of how trees begin their learning. Geurts et al. (2006) provide detailed information on how the ERT model works. The ERT model enables us to evaluate the importance of each independent variable for the surface  $O_3$  estimation, assigning an importance score to each variable, calculated according to the GI. It normalizes the cumulative changes in GIs before and after node branching for each feature during the model training (Jiang et al., 2009). The higher this score, the more important is this feature in the decision-tree construction. Variables with high scores make greater contributions to the model performance. Low-score variables have smaller effects on the model and may generate redundant information (Wei et al., 2020, 2021a). Variables with importance scores less than 1% are eliminated from the model to improve its efficiency and avoid overfitting caused by redundant input variables.

Per our analysis of each feature importance, DSR and TEM are the two most important variables for model construction, with high importance scores of 32% and 14%, respectively (Fig. S2). OMI  $NO_2$  and  $O_3$  products are also highly valuable, with importance scores of 6% and

5%. However, they can only provide trace gas information about the troposphere and the whole atmosphere. Other meteorological variables (especially ET and RH) and two land-related variables (i.e., DEM and NDVI) also have significant impacts on O<sub>3</sub> estimates, with importance scores ranging from 2% to 7%. The emissions of three main O<sub>3</sub> precursors (i.e., NO<sub>x</sub>, VOCs, and CO) influence the model to some degree, with importance scores of about 2%. In general, all 18 selected independent variables have an impact, with importance scores >1.5%. They are, therefore, all kept in the model.

2.2.2. Model extension

In the second stage, we extended a STET model for the surface O<sub>3</sub> estimation by considering the autocorrelation of O<sub>3</sub> pollution in space and its differences in time series using the original ERT model (Wei et al., 2021a). The position of one point in space is expressed by its longitude and latitude and the Haversine great-circle distances to the four corners and the center of the study region (i.e., 73.6°E-134.8°E, 15.8°N-53.7°N). The time is expressed by the day of the year (DOY), set to identify each raw data record on each day under different air pollution conditions. The above-mentioned independent variables, along with space and time

terms, are input into the model to build a robust ground-level O<sub>3</sub> estimation specific to China.

2.3. Validation method

In this study, the widely adopted out-of-sample (sample-based) 10-fold cross-validation (10-CV) method was selected to test the overall model performance in estimating near-surface O<sub>3</sub> concentrations (Di et al., 2017; Li et al., 2020b; Liu et al., 2020; Zhan et al., 2018). This method stipulates that all data samples are first randomly divided into 10 groups, of which 9 groups (i.e., 90% of the samples) are used for model training, and the rest (i.e., 10% of the samples) are used for model validation. This operation runs 10 times to ensure that all samples have been tested (Rodriguez et al., 2010).

Furthermore, an additional out-of-station (station-based) 10-CV method was employed to test the spatial predictive ability of the model in areas without ground-based measurements (Li et al., 2017b; Wei et al., 2020; Wu et al., 2021). It is performed using measurements from ground-based O<sub>3</sub> monitoring stations. The stations are randomly divided into 10 groups, of which data samples from 9 groups (i.e., 90%

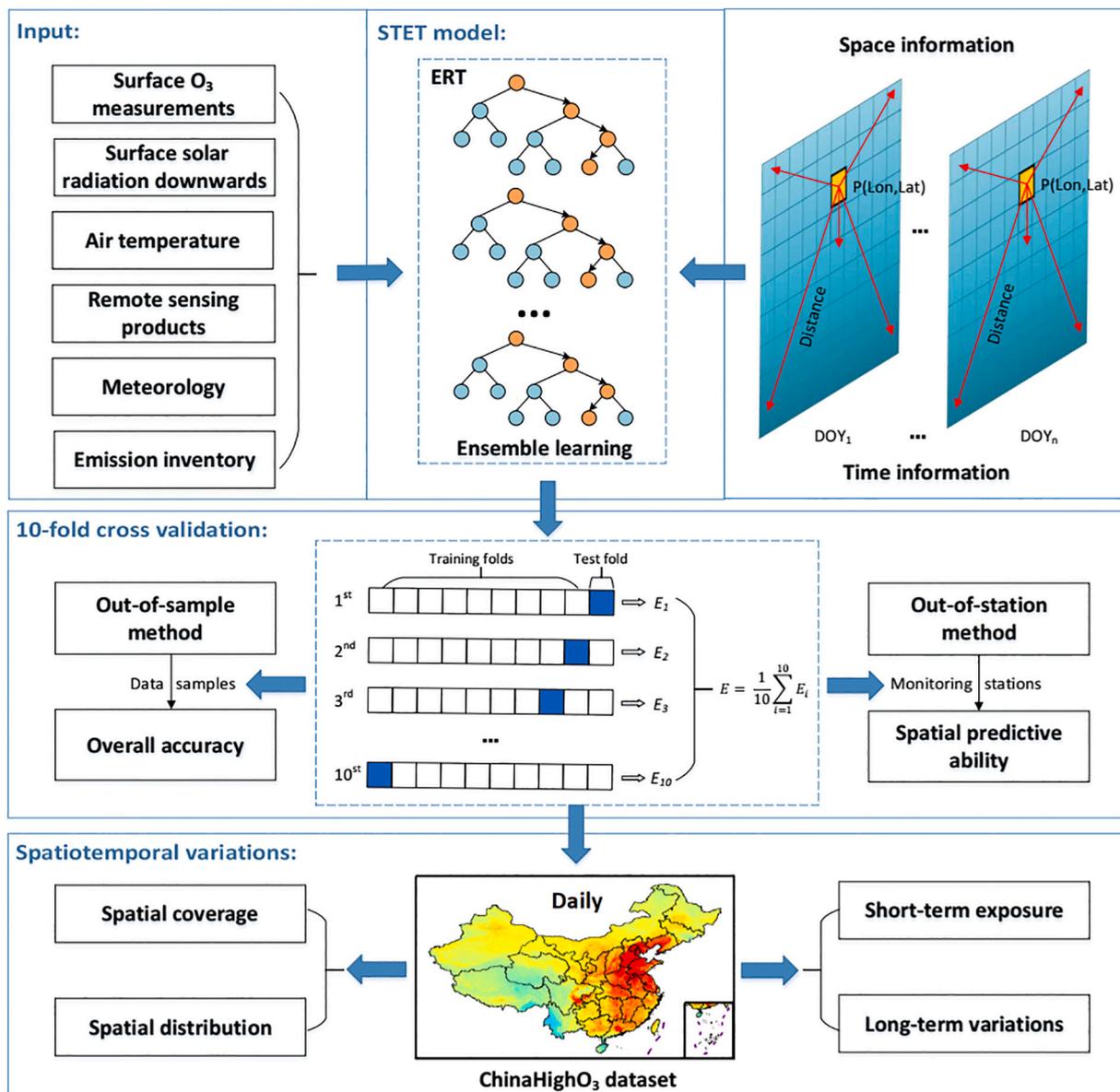


Fig. 2. Flowchart of the mapping process of the ChinaHighO<sub>3</sub> dataset in our study.

of the stations) and the remaining group (i.e., 10% of the stations) are employed for model training and validation, respectively. The training and validation samples thus represent data samples collected at different locations in space. This method enables us to evaluate the predictive accuracy of the model at locations where ground-based O<sub>3</sub> measurements are unavailable.

In addition, several main statistical metrics were used, including ordinary least squares (OLS; Zdaniuk, 2014) regression (e.g., slope and intercept), coefficient of determination (R<sup>2</sup>), root-mean-square error (RMSE), mean absolute error (MAE), and mean relative error (MRE). Deseasonalized O<sub>3</sub> monthly anomalies were adopted to calculate temporal trends (Wei et al., 2019) and used to analyze the long-term spatiotemporal variations in O<sub>3</sub> pollution across China. Fig. 2 shows the flowchart of the mapping process of the ChinaHighO<sub>3</sub> dataset in our study.

### 3. Results and discussion

#### 3.1. Accuracy assessment

##### 3.1.1. Overall model performance

First, we validated the overall performance of the developed model

using the out-of-sample approach at different spatial scales. Collocated are more than 3.5 million data samples ( $N = 3,567,344$ ) from 2013 to 2020 across China. The MDA8 O<sub>3</sub> estimates for the whole of China are highly consistent with surface measurements ( $CV-R^2 = 0.87$ ), with the slope and y-intercept equal to 0.87 and 11.8  $\mu\text{g}/\text{m}^3$ , respectively (Fig. 3a). The mean RMSE, MAE, and MRE values are 17.10  $\mu\text{g}/\text{m}^3$ , 11.29  $\mu\text{g}/\text{m}^3$ , and 18.38%, respectively, over the entire domain. Note that the overall accuracy of O<sub>3</sub> estimates has significantly improved compared to results derived from the original ERT model (i.e.,  $CV-R^2 = 0.78$ , slope = 0.81, RMSE = 22.39  $\mu\text{g}/\text{m}^3$ , and MAE = 14.88  $\mu\text{g}/\text{m}^3$ ) (Geurts et al., 2006). This confirms the necessity for including spatio-temporal information about O<sub>3</sub> pollution.

We further tested the model performance in typical regions in China. The model works best over the Beijing-Tianjin-Hebei (BTH) region (Fig. 3c) and the North China Plain (NCP, Fig. 3b), with out-of-sample CV-R<sup>2</sup> values of 0.91 and 0.89, respectively, and slopes from linear regression close to 1.0 (0.91 and 0.89, respectively). The model performance is slightly poorer (e.g.,  $CV-R^2 = 0.85$ – $0.86$ , and slope = 0.84– $0.86$ ) in the Yangtze River Delta (YRD, Fig. 3d), the Pearl River Delta (PRD, Fig. 3e), and the Sichuan Basin (SCB, Fig. 3f). Overall, the model uncertainty is generally small and stable with small differences (e.g., RMSE = 18.9–21.3  $\mu\text{g}/\text{m}^3$ , MAE = 12.1–13.6  $\mu\text{g}/\text{m}^3$ , and MRE =

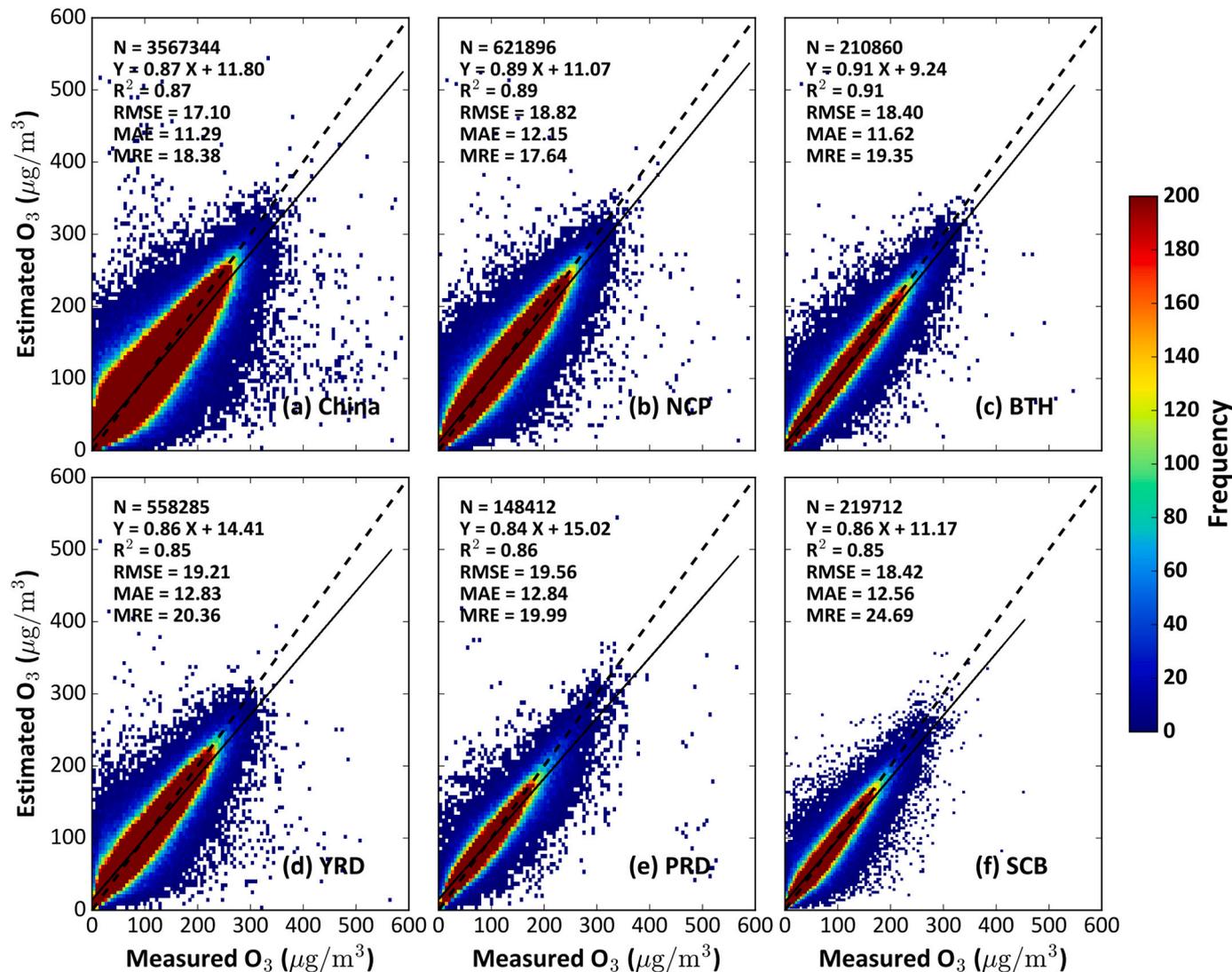


Fig. 3. Out-of-sample cross-validation results of MDA8 O<sub>3</sub> estimates ( $\mu\text{g}/\text{m}^3$ ) from 2013 to 2020 (a) in China, (b) the North China Plain (NCP), (c) the Beijing-Tianjin-Hebei (BTH) region, (d) the Yangtze River Delta (YRD), (e) the Pearl River Delta (PRD), and (f) the Sichuan Basin (SCB). Frequencies in the right-hand legend indicate the total number of data samples in each cell.

17.6–24.7%). These results suggest the varying robustness of our model at the regional scale in China, stemming chiefly from variable input parameters in terms of their density and accuracy.

The model performance for each separate year (Table 1) was also evaluated for the whole of China. The overall accuracy of the MDA8 O<sub>3</sub> estimates in the years since 2017 (i.e., out-of-sample CV-R<sup>2</sup> = 0.89–0.93, RMSE = 11.9–15.6 μg/m<sup>3</sup>, MAE = 7.9–10.8 μg/m<sup>3</sup>, and MRE = 10.3–15.0%) is generally better than that of the previous years (i.e., out-of-sample CV-R<sup>2</sup> = 0.79–0.82, RMSE = 19.1–22.4 μg/m<sup>3</sup>, MAE = 12.9–14.9 μg/m<sup>3</sup>, and MRE = 21.4–31.8%). The continuous increase in the density of monitoring stations, resulting in a sharp increase in the number of data samples, and instrument improvements and quality control upgrades explain this (Wei et al., 2021a). Overall, our model works well for the study period considered and for individual years in the study domain.

On the individual-station scale (Fig. 4), the sample size varies from site to site due to differences in the observational record and the number of useful data samples from 2013 to 2020. Except for a few stations established later in the study period, most stations have a sufficient number of data samples (Fig. S3a), with an average sample size (N) of 2230 and with more than 83% of the stations having at least 5 years of data samples (i.e., N > 1825). In terms of model accuracy, CV-R<sup>2</sup> values exceed 0.8 at ~83% of the stations, especially those located in central and eastern China (CV-R<sup>2</sup> > 0.9). In terms of model uncertainty, except for a few individual stations, ~83% of the stations have RMSE values <21 μg/m<sup>3</sup>, ~88% have MAE values <15 μg/m<sup>3</sup>, and ~85% have MRE values <25%. Overall, our model performs well at the station scale, with average CV-R<sup>2</sup>, RMSE, MAE, and MRE values of 0.86, 16.48 μg/m<sup>3</sup>, 11.23 μg/m<sup>3</sup>, and 18.36%, respectively.

### 3.1.2. Spatial predictive capability

Next, we focus on evaluating the ability of our model to predict surface O<sub>3</sub> spatially using the out-of-station approach at varying spatial scales. Over the entire domain, our surface O<sub>3</sub> predictions are well correlated to observations (e.g., CV-R<sup>2</sup> = 0.80, slope = 0.84), with mean RMSE, MAE, and MRE values of 21.10 μg/m<sup>3</sup>, 13.87 μg/m<sup>3</sup>, and 23.18% (Fig. S4a). These values are somewhat lower than the out-of-sample validation results (Fig. 3a), indicating a strong spatial predictive ability. The spatial predictive ability of the model also gradually increases over the years (Table 1), possibly due to the same reasons discussed above.

On the regional scale, the predictive ability of the model varies, i.e., better surface O<sub>3</sub> predictions were observed in the BTH (CV-R<sup>2</sup> = 0.87, RMSE = 21.51 μg/m<sup>3</sup>; Fig. S4c) and NCP (CV-R<sup>2</sup> = 0.84, RMSE = 21.99 μg/m<sup>3</sup>; Fig. S4b) regions, while relatively less accurate predictions were observed in the YRD, SCB, and PRD regions (CV-R<sup>2</sup> < 0.80, RMSE >22.6 μg/m<sup>3</sup>; Fig. S4d–f). Compared with the out-of-sample results (Fig. S4: b–f), the accuracy changed little. The evaluation metrics of the former two regions decreased slightly less than those of the other three regions. This is mainly due to the different densities of monitoring stations in each region.

Furthermore, the spatial predictive ability of the model shows spatial

differences on the individual-station scale (Fig. S5). The predictive accuracy of surface O<sub>3</sub> concentrations is poor in most stations located in western China, with large estimated uncertainties (e.g., CV-R<sup>2</sup> < 0.5, RMSE >24 μg/m<sup>3</sup>, MAE > 18 μg/m<sup>3</sup>, and MRE > 25%). By contrast, the model has a strong predictive ability in most stations in eastern China, with high CV-R<sup>2</sup> values >0.8 and small RMSE, MAE, and MRE values of <18, 12 μg/m<sup>3</sup>, and 15%, respectively. This difference in predictive ability between the eastern and western parts of China likely arises because aside from different environmental and population conditions, there are fewer monitoring stations in western China. At locations that have never had air pollution monitoring, such as remote desert and plateau areas, the uncertainty in the model predictions can be larger. This can only be truly quantified when new observations from these remote areas become available. In general, ~80%, 78%, 86%, and 70% of the stations have CV-R<sup>2</sup>, RMSE, MAE, and MRE values >0.7, < 24 μg/m<sup>3</sup>, 18 μg/m<sup>3</sup>, and 25%, with average values of 0.79, 20.08 μg/m<sup>3</sup>, 13.82 μg/m<sup>3</sup>, and 23.17%, respectively.

### 3.1.3. Temporal-scale validation

Fig. 5 shows the time series of the evaluation metrics used to evaluate how well the model estimated daily O<sub>3</sub> surface concentrations from 2013 to 2020. The daily sample size is large, ranging from 8003 to 10,060, with an average of 9766 and remaining unchanged over time (Fig. S3b). This is due to the unique advantage of the ChinaHighO<sub>3</sub> dataset, namely, its full coverage of China. While the performance varies somewhat seasonally, the magnitudes of the changes are moderate throughout the year, with CV-R<sup>2</sup>s ranging from 0.69 to 0.89 (average = 0.81), exceeding 0.75 on about 88% of the days in a year. The absolute uncertainties (i.e., RMSE and MAE) of the O<sub>3</sub> estimates vary seasonally, i.e., low in spring and winter but high in summer. By contrast, the relative uncertainty (MRE) shows an opposite seasonal variation. Surface MDA8 O<sub>3</sub> concentrations are relatively high in summer in most mid-latitude regions of China (Gong et al., 2018). Reasons for the larger errors in summer are the longer diurnal cycle and greater variation in summertime O<sub>3</sub>. The averaged variables used in the model may not reflect conditions associated with a high O<sub>3</sub> content in the afternoon, while observations are likely driven by the afternoon peaks. In general, average RMSE, MAE, and MRE values are 18.82 μg/m<sup>3</sup>, 11.27 μg/m<sup>3</sup>, and 18.42%, and < 20, 15 μg/m<sup>3</sup>, and 20% μg/m<sup>3</sup> on ~86%, 99%, and 75% of the days, respectively.

Monthly mean MDA8 O<sub>3</sub> estimates for each year were also evaluated (Fig. S6). High accuracy is seen, with strong slopes from linear regression of 0.83–0.97, high R<sup>2</sup> values of 0.86–0.97, and small uncertainties, with RMSE and MAE (MRE) values ranging from 5.5 and 4.0 μg/m<sup>3</sup> (4.4%) to 13.4 and 9.0 μg/m<sup>3</sup> (13.8%) among different years. In general, the data quality of the monthly O<sub>3</sub> estimates (N = 119,194) is reliable (e.g., R<sup>2</sup> = 0.93, RMSE = 9.42 μg/m<sup>3</sup>, MAE = 6.91 μg/m<sup>3</sup>, and MRE = 8.56%) during the entire study period of 2013–2020. This allows us to accurately analyze the spatiotemporal distributions of and variations in O<sub>3</sub> pollution in China.

**Table 1**

Statistics describing cross-validation results of MDA8 O<sub>3</sub> estimates (μg/m<sup>3</sup>) for each year from 2013 to 2020 in China.

Year	Sample size	Overall accuracy				Spatial predictive ability			
	N	R <sup>2</sup>	RMSE	MAE	MRE	R <sup>2</sup>	RMSE	MAE	MRE
2013	115,663	0.79	21.99	14.83	22.90	0.63	29.57	19.82	31.18
2014	325,152	0.80	22.39	14.81	31.80	0.65	30.02	20.52	45.79
2015	519,391	0.79	20.89	13.90	27.75	0.64	27.73	18.87	37.69
2016	516,746	0.82	19.19	12.99	21.49	0.72	23.71	16.38	27.45
2017	527,483	0.89	15.52	10.79	14.99	0.85	17.82	12.54	17.36
2018	520,002	0.91	14.10	9.64	13.34	0.88	15.66	10.82	15.00
2019	520,381	0.92	13.99	9.49	13.17	0.91	15.31	10.48	10.48
2020	522,526	0.93	11.96	7.97	10.27	0.92	12.96	8.71	11.33
All	3,567,344	0.87	17.10	11.29	18.38	0.80	21.10	13.87	23.18

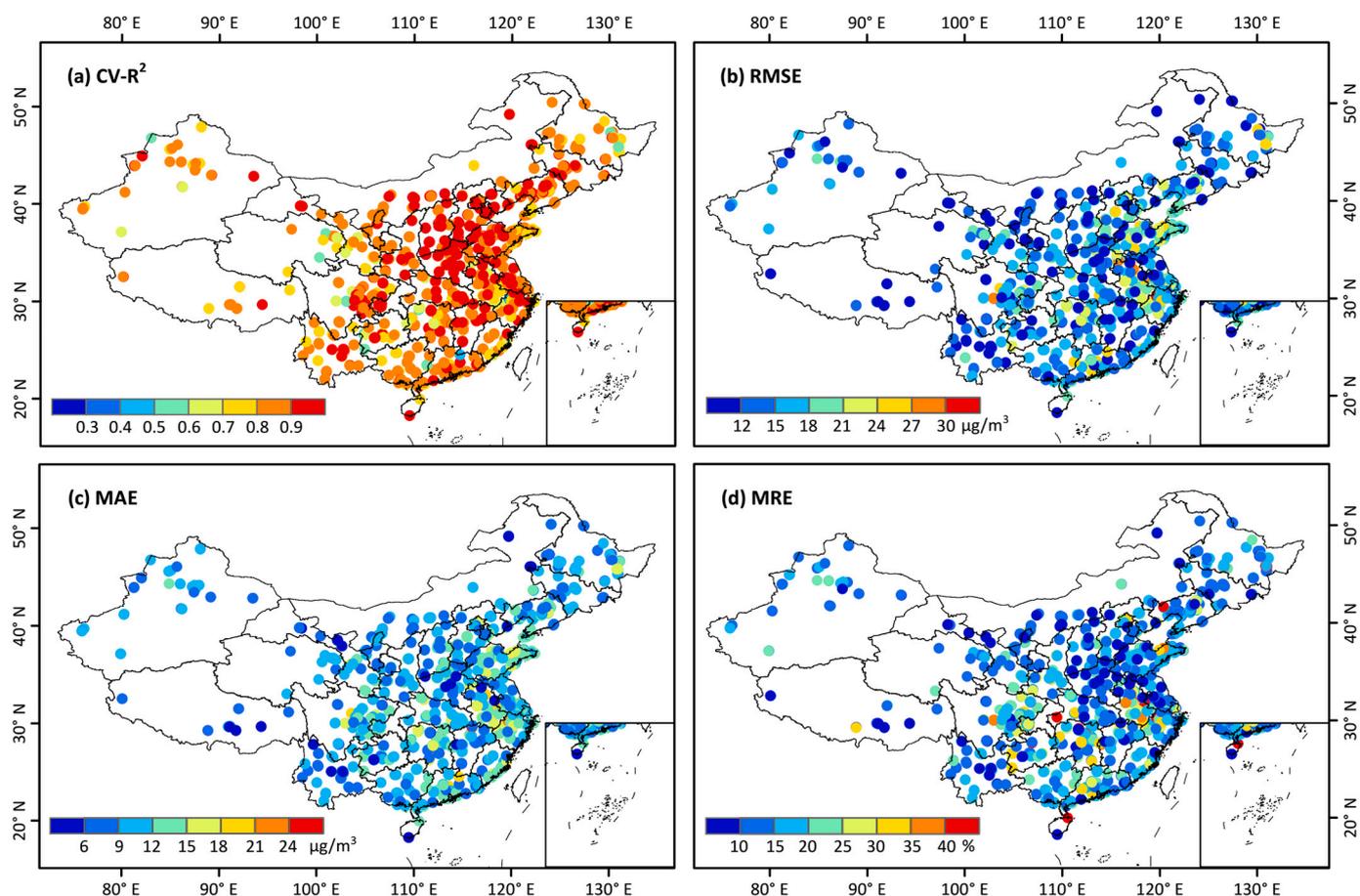


Fig. 4. Individual-site-scale out-of-sample cross-validation results of MDA8 O<sub>3</sub> estimates ( $\mu\text{g}/\text{m}^3$ ) from 2013 to 2020 in China.

### 3.2. Spatiotemporal surface O<sub>3</sub> variations

#### 3.2.1. The ChinaHighO<sub>3</sub> dataset

Combining ground-based observations, satellite remote sensing data, atmospheric reanalysis products, and emission inventory, we generated the ChinaHighO<sub>3</sub> dataset using the STET model, which belongs to one of the series of long-term, full-coverage, high-resolution and high-quality ground-level air pollutants for China (i.e., ChinaHighAirPollutants, CHAP). The ChinaHighO<sub>3</sub> dataset includes daily MDA8 O<sub>3</sub> maps from 1 January 2013 to 31 December 2020, which was released on 30 December 2020 and is constantly updated. It overcomes the problem of missing data in optical remote sensing products caused by cloud contaminations and can provide full-coverage ground-level O<sub>3</sub> distributions over China. Monthly, seasonal, and annual MDA8 O<sub>3</sub> maps from 2013 to 2020 are also available (Table S2).

Fig. 6 presents two typical examples of MDA8 O<sub>3</sub> maps, i.e., 18 June and 11 November 2019, and the annual map for 2019 in China. Our O<sub>3</sub> estimates can cover the whole of China on these two individual days. In general, the O<sub>3</sub> concentration is particularly high ( $> 150 \mu\text{g}/\text{m}^3$ ) in northern China and much lower ( $< 80 \mu\text{g}/\text{m}^3$ ) in southern China on 18 June 2019 (average =  $118.7 \pm 36.1 \mu\text{g}/\text{m}^3$ ). High emissions of three main O<sub>3</sub> precursors (i.e., NO<sub>x</sub>, VOCs, and CO) are mainly observed in eastern China, especially the NCP (Fig. S7). In general, differences in surface O<sub>3</sub> concentrations between northern and southern China were on the order of 2% on this day. A different situation was observed on 11 November 2019 (average =  $77.9 \pm 24.6 \mu\text{g}/\text{m}^3$ ). On an annual scale, differences in O<sub>3</sub> distribution between northern and southern China in 2019 decreased, with an average level of  $98.3 \pm 11.3 \mu\text{g}/\text{m}^3$ . The great differences between northern and southern China on different days are mainly due to differences in sunlight and ozone chemical formation in

different seasons. A comparison with ground-based observations shows highly consistent spatial patterns on both daily and annual scales across China. These results thus illustrate that spatially continuous O<sub>3</sub> data are important for air pollution studies at those places without monitoring stations.

Fig. 7 shows mean MDA8 O<sub>3</sub> maps for different seasons from 2013 to 2020 across China. O<sub>3</sub> concentrations changed significantly on a seasonal scale and were extremely high in the summer (average =  $103.6 \pm 18.0 \mu\text{g}/\text{m}^3$ ), especially in the NCP (average =  $138.8 \pm 13.5 \mu\text{g}/\text{m}^3$ ), followed by spring (average =  $99.4 \pm 9.2 \mu\text{g}/\text{m}^3$ ). By contrast, wintertime O<sub>3</sub> concentrations in China were much lower (average =  $69.9 \pm 7.7 \mu\text{g}/\text{m}^3$ ), especially in the BTH region (average =  $55.4 \pm 7.6 \mu\text{g}/\text{m}^3$ ). The spatial pattern of O<sub>3</sub> in autumn (average =  $80.9 \pm 10.1 \mu\text{g}/\text{m}^3$ ) was similar but generally higher than that in winter across China, especially in southeastern areas. Fig. S8 shows zoomed-in summer mean O<sub>3</sub> maps for four key regions in China. The ChinaHighO<sub>3</sub> dataset can reflect and describe well the distribution of and variation in ozone pollution at the local, even urban, scale due to its high spatial resolution of 10 km. All four typical regions experienced different degrees of O<sub>3</sub> pollution in summer, especially the BTH (average =  $142.9 \pm 14.5 \mu\text{g}/\text{m}^3$ ) and YRD (average =  $113.9 \pm 13.7 \mu\text{g}/\text{m}^3$ ) regions.

#### 3.2.2. A short-term severe O<sub>3</sub> pollution event

In addition to the atmospheric environment and air quality, short-term exposure to air pollution has a significant impact on human health (Almeida et al., 2011; Giani et al., 2020; Wong et al., 2007). It is difficult for traditional ground-based observations and previous satellite estimates to accurately capture air pollution on a wide scale due to the uneven distribution of surface monitoring stations. Our study helps make up for this deficiency by monitoring the pollution exposure at

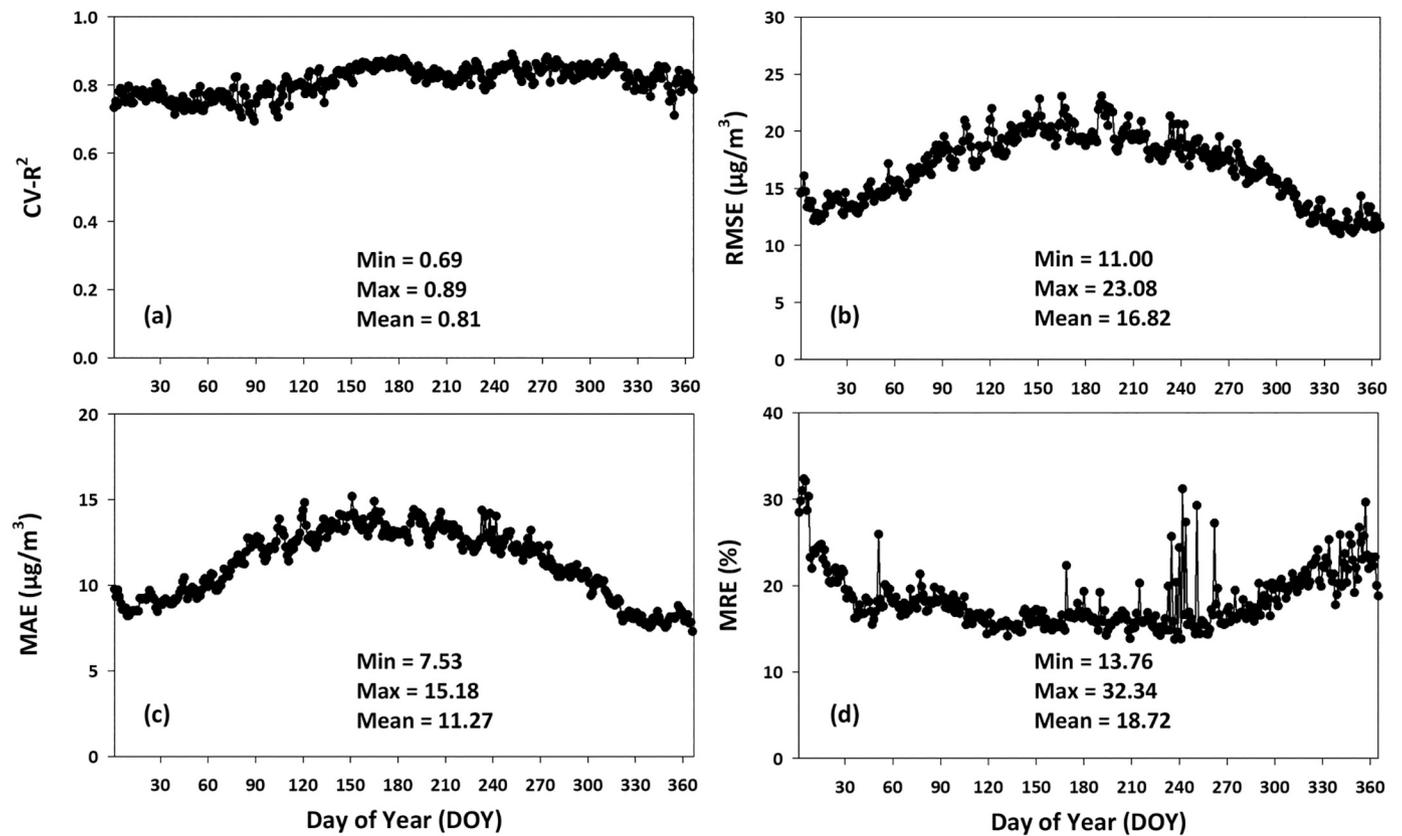


Fig. 5. Time series of daily variations of validation results of MDA8 O<sub>3</sub> estimates (µg/m<sup>3</sup>) from 2013 to 2020. Minimum, maximum, and mean values are given in each panel.

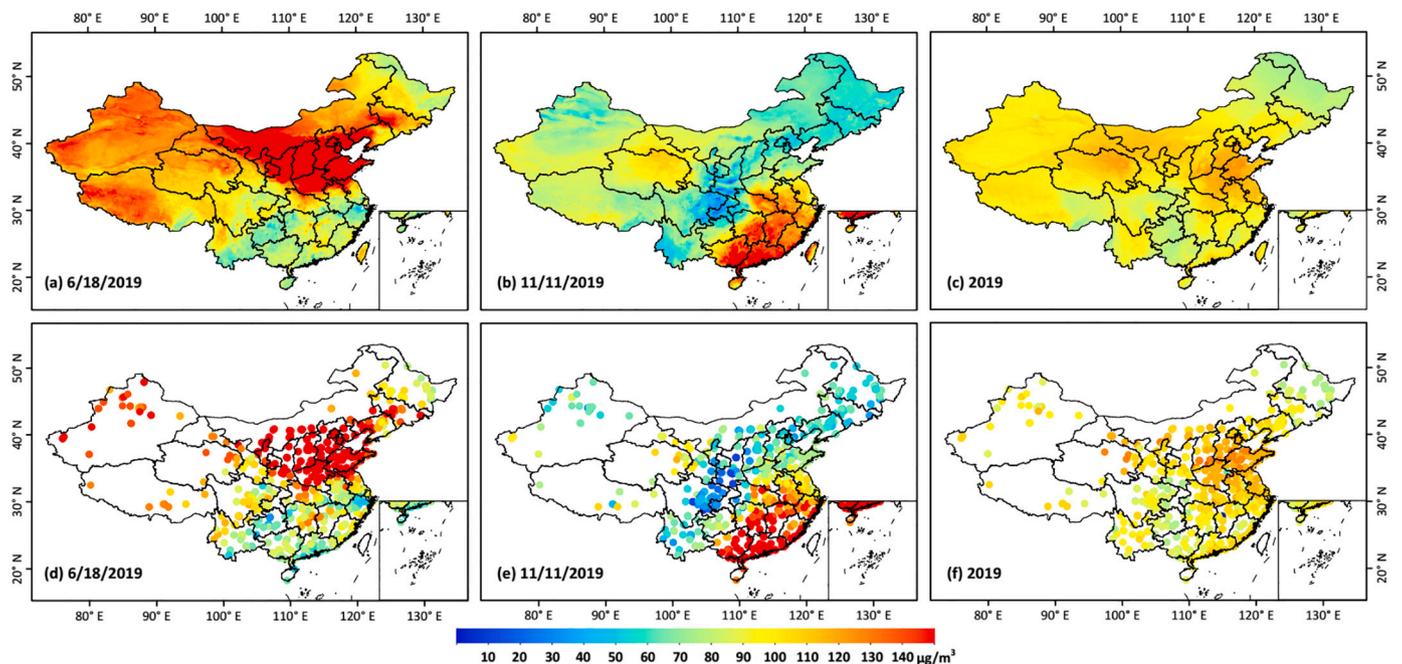


Fig. 6. (a-c) STET-model-derived and (d-f) ground-based MDA8 O<sub>3</sub> maps on 18 June 2019 (a & d), 11 November 2019 (b & e), and the annual mean map for 2019 (c & f) in China.

unprecedented spatiotemporal scales, i.e., generating spatially continuous and full-coverage daily surface O<sub>3</sub> maps. This allows users to quickly obtain more accurate estimates of the distribution of and variation in O<sub>3</sub> pollution at any location, especially in those areas with no or

minimal ground-based measurements.

We closely examined a severe surface O<sub>3</sub> pollution episode that occurred from 23 April to 8 May in 2020 in eastern China (Fig. 8). Before 25 April, O<sub>3</sub> was at a low level across the whole country, then gradually

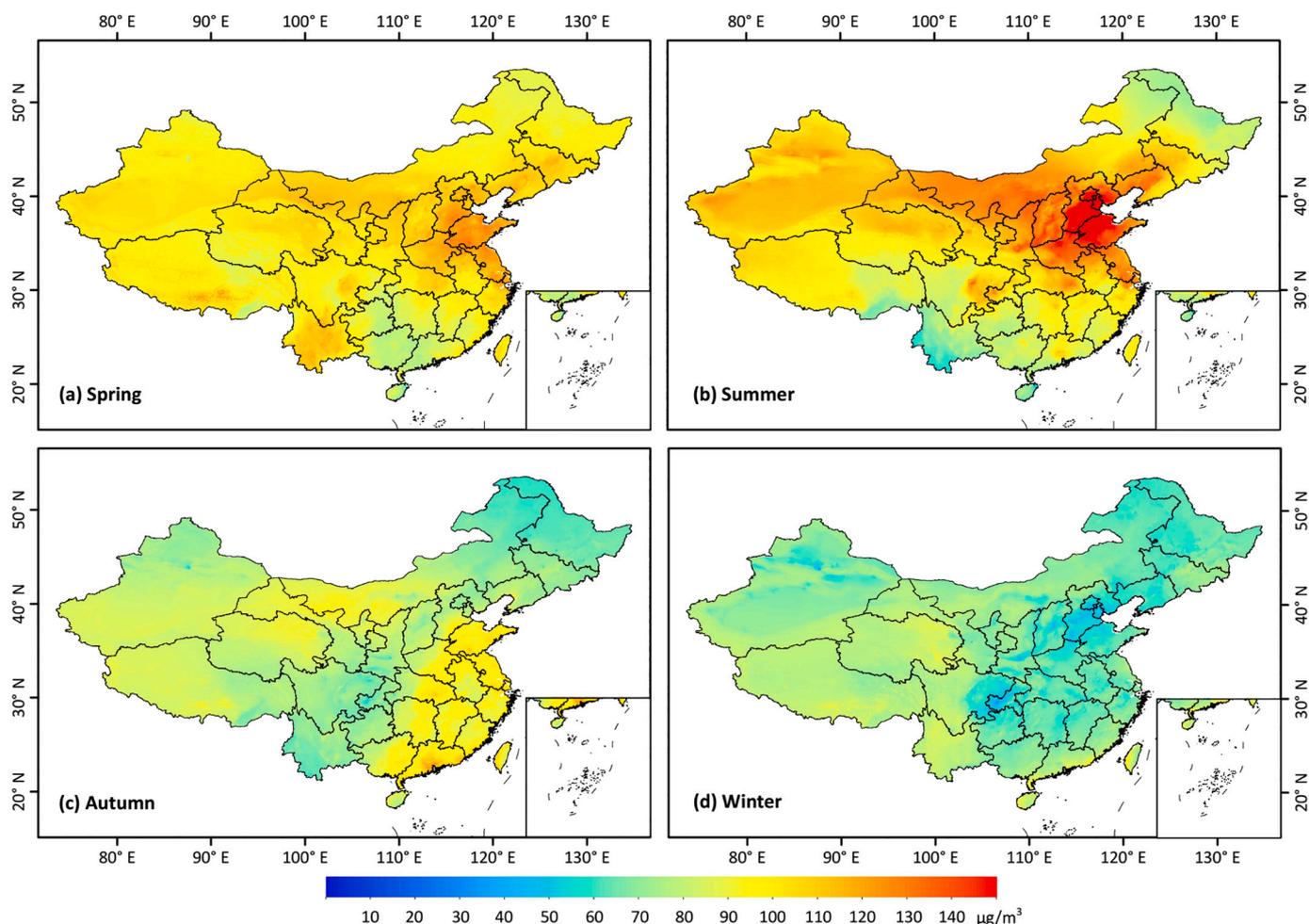


Fig. 7. Multi-year seasonal mean MDA8 O<sub>3</sub> maps (horizontal resolution = 10 km) averaged over the period 2013–2020 across China.

increased. On 28 April, the O<sub>3</sub> levels in all “2 + 26” cities (Fig. 1), the main air pollution transmission belt in the BTH and its surrounding area (“2” refers to Beijing and Tianjin, and “26” refers to 26 prefecture-level cities in Hebei Province), had exceeded the ambient air quality standard, i.e., MDA8 O<sub>3</sub> = 160 µg/m<sup>3</sup> (Fig. S9). More severe O<sub>3</sub> pollution occurred in most other areas on 29 April, with maximum values of  $124.0 \pm 30.2$  µg/m<sup>3</sup> and  $181.0 \pm 17.8$  µg/m<sup>3</sup> in China and YRD, respectively (Fig. S9). On 30 April, BTH experienced the maximum level of O<sub>3</sub> pollution (average =  $232.1 \pm 47.2$  µg/m<sup>3</sup>), remaining high until 2 May, when >50% of the cities in China exceeded the daily ozone standard. The air quality was significantly improved in northern BTH starting on 3 May, but central and southern China still suffered from light to moderate pollution, with some cities experiencing severe pollution. This national heavy pollution event lasted for nearly a week.

Surface O<sub>3</sub> concentrations were generally low in SCB before 25 April, gradually forming into regional pollution on 26 April, polluting the Chengdu Plain and southern and northeast Sichuan to varying degrees. By 28 April, most cities exceeded the ambient air quality standard. Pingyuan and southern Sichuan were heavily polluted, and O<sub>3</sub> concentrations remained high, reaching a maximum on 3 May, with an average value of  $184.3 \pm 29.8$  µg/m<sup>3</sup> in SCB (Fig. S9). On 6 May, the polluted air moved southward, gradually decreasing in pollution intensity. After 7 May, accompanied by cooling and precipitation, this episode of ozone pollution ended, and the air quality improved to good or excellent. This episode of severe regional pollution lasted for about 11 days, the first severe ozone pollution event with a long duration and wide coverage in Sichuan province since the start of 2020.

### 3.2.3. Changes during the COVID-19 pandemic

The coronavirus disease (COVID-19) broke out in Wuhan, Hubei Province at the end of 2019, quickly spreading to the whole country due mainly to the Spring Festival (WHO, 2020; Zu et al., 2020). To prevent the further spread of COVID-19, Hubei Province went into lockdown starting at 10 am on 23 January 2020, soon followed by almost all other major cities in China and lasting for about three weeks (Su et al., 2020; Tian et al., 2020). To gain further insight into ozone changes associated with COVID-19, O<sub>3</sub> changes in China are examined before (Period I: 1–25 January), during (Period II: 26 January to 17 February), and after (Period III: 18 February to 31 March) the COVID-19 outbreak. Considering the increase in O<sub>3</sub> in recent years, only compared are the relative differences in O<sub>3</sub> concentrations across eastern China between 2020 and 2019 during the three periods (Fig. 9).

Before the COVID-19 outbreak, O<sub>3</sub> concentrations remained near historical values, with relative changes within  $\pm 10\%$ . During the lockdown, significant increases in O<sub>3</sub> concentrations were seen in most parts of eastern China, especially in Hubei Province and its surrounding area, showing a relative change of >40%. Because O<sub>3</sub> formation rates over northern China are under a NO<sub>x</sub>-saturated regime, a reduction in NO<sub>x</sub> would enhance O<sub>3</sub> generation rates (Benish et al., 2020; Liu et al., 2021; Shi and Brasseur, 2020). By contrast, an opposite decline in O<sub>3</sub> concentrations was observed in the PRD, mainly caused by meteorological changes and reductions in both VOCs and NO<sub>x</sub> emissions during the lockdown (Wang et al., 2021). In addition, different from northern China, O<sub>3</sub> formation rates over the PRD are under a NO<sub>x</sub>-limited regime, so the same reduction in NO<sub>x</sub> would diminish O<sub>3</sub> generation rates (Liu et al., 2021; Wang et al., 2021). After the COVID-19 outbreak, O<sub>3</sub>

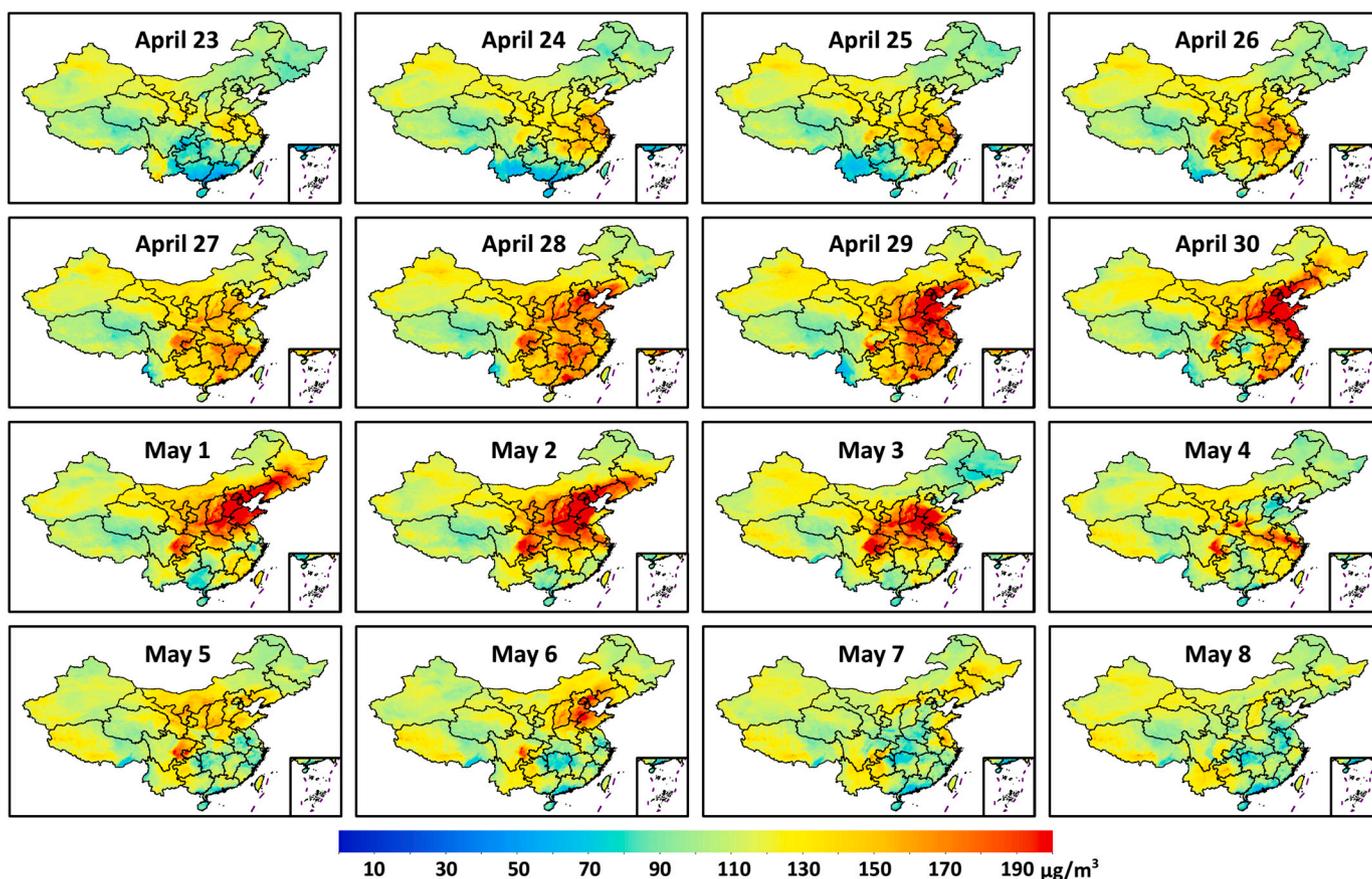


Fig. 8. A typical example of a severe O<sub>3</sub> pollution event that occurred from 23 April 2020 to 8 May 2020 in eastern China.

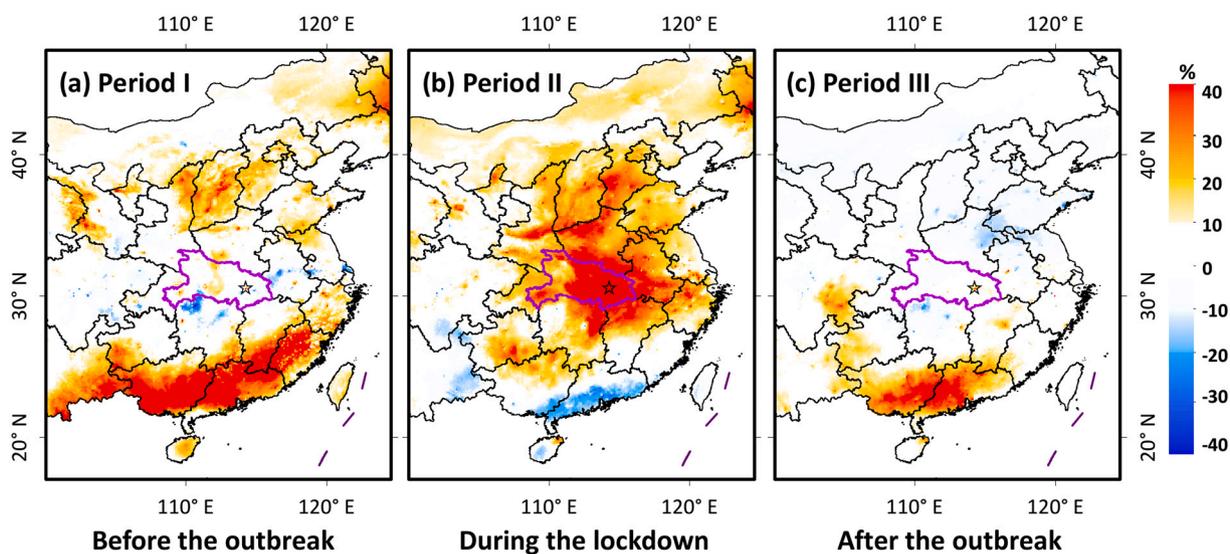


Fig. 9. Relative changes (%) in mean MDA8 O<sub>3</sub> concentration ( $\mu\text{g}/\text{m}^3$ ) in 2020 (during the COVID-19 epidemic) and 2019 during the same periods: (a) Period I (1–25 January) before the outbreak, (b) Period II (26 January to 17 February) during the lockdown, and (c) Period III (18 February to 31 March) after the outbreak in eastern China. The area outlined in magenta shows the boundary of Hubei Province, and the magenta star shows the location of Wuhan City. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

concentrations changed little (within  $\pm 10\%$ ) compared with concentrations in the previous year in most areas of eastern China, indicating that life had returned to normal. In southern China, there was a contrasting increase in O<sub>3</sub> concentrations, likely related to increases in NO<sub>x</sub> and temperature (Wang et al., 2021). The rate of ozone production varies nonlinearly with VOC and NO<sub>x</sub> emissions, and air quality can

initially worsen when NO<sub>x</sub> emissions are reduced, but the total amount of ozone produced ultimately increases with increasing NO<sub>x</sub> emissions (Lin et al., 1988). Strict NO<sub>x</sub> controls must eventually be implemented to protect the environment and human health.

### 3.2.4. Long-term variations in the recent decade

Fig. 10 shows the MDA8 O<sub>3</sub> trends (μg/m<sup>3</sup>/yr) during the study period (2013–2020), calculated from monthly anomalies across China. Surface O<sub>3</sub> concentrations varied from national to regional scales during the recent eight years. In general, most areas of the country show significant increasing O<sub>3</sub> pollution, with an average of 2.49 μg/m<sup>3</sup>/yr ( $p < 0.001$ ), especially in central China ( $> 5 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.05$ ) and the NCP ( $\sim 4.42 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ). The BTH and YRD regions had the stronger increasing trends, i.e., 3.84 and 3.43 μg/m<sup>3</sup>/yr ( $p < 0.001$ ), respectively. The other two typical regions, i.e., SCB ( $\sim 1.78 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ) and PRD ( $\sim 1.41 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.05$ ), show relatively low but clear increasing trends. The increase in O<sub>3</sub> over city clusters is closely associated with a decrease in NO<sub>x</sub> emissions and PM<sub>2.5</sub> concentrations (Li et al., 2019; Wang et al., 2020; Wei et al., 2021a; Zhang et al., 2019) and meteorological variations (Li et al., 2020a). By contrast, seen are opposite weakening trends in several coastal provinces in southern China (e.g., Guangxi and Zhejiang).

Next, we investigated surface O<sub>3</sub> variations under the background of different implemented environmental policies (Table 2). During the Clear Air Action Plan (2013–2017), China showed a significant increasing trend of 1.33 μg/m<sup>3</sup>/yr ( $p < 0.05$ ), especially in the NCP ( $\sim 4.58 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ) and BTH ( $\sim 4.38 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ) regions. Increasing trends were also found in the YRD and SCB regions. By contrast, O<sub>3</sub> pollution overall declined in the PRD region. During the Blue-Sky Defense Plan (2018–2020), O<sub>3</sub> concentrations continued to increase by 7.2% and 2.5–5.4% in China and typical urban agglomerations, respectively, in 2020 compared to 2017. Considering the entire study period, O<sub>3</sub> pollution increased the most in China ( $\sim 4.40 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ) and most typical regions during the period 2015–2019, especially the NCP ( $\sim 6.33 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ) and YRD ( $\sim 5.60 \mu\text{g}/\text{m}^3/\text{yr}$ ,  $p < 0.001$ ).

Taking into account the seasonal differences in O<sub>3</sub> discussed above, we next focus on the spatiotemporal variations in summertime mean MDA8 O<sub>3</sub> concentrations from 2013 to 2020 over eastern China (Fig. 11). Ozone levels remained at a high level in summer among different years in China, with an average value of  $> 90 \mu\text{g}/\text{m}^3$ . It was higher during the period 2017–2019 than in previous years, especially in the NCP ( $> 120 \mu\text{g}/\text{m}^3$ ). This was closely associated with rising temperatures and an increase in the number of hot days in the NCP (Li et al.,

2020a). Changes in O<sub>3</sub> have been diverse in the recent eight years, e.g., O<sub>3</sub> concentrations were higher in 2014 than in 2013 in most areas of China, yet generally decreased in 2015, especially in southern China. Ozone pollution had increased significantly since 2016, reaching a maximum in 2019 ( $\sim 117.4 \pm 23.6 \mu\text{g}/\text{m}^3$ ), especially in the NCP ( $\sim 159.7 \pm 14.1 \mu\text{g}/\text{m}^3$ ). This may be due to the decreasing PM<sub>2.5</sub> concentrations by  $\sim 15\%$  in the NCP (Li et al., 2020a; Wei et al., 2021a); NO<sub>x</sub> emission reduction is also suspected as an important driving factor of O<sub>3</sub> increase in recent years. However, the dominant reason remains controversial. By contrast, in 2020, overall O<sub>3</sub> pollution levels decreased in China and in most typical regions in China (Table S3). The coordinated control measures of fine PM and O<sub>3</sub> implemented by the Chinese government (Xiang et al., 2020) may explain this, as well as the ongoing effects of COVID-19 in China. These results are highly consistent with those previously reported, based on ground-based measurements made from 2013 to 2019 (Li et al., 2020a; Lu et al., 2020; Wang et al., 2020). Our predicted results also show similar patterns in spatial distribution compared to those derived from OMI/Aura satellite observations (Liu et al., 2020; Zhang et al., 2020) and air quality model simulations (Hu et al., 2016; Xue et al., 2020).

We also calculated the percentage of O<sub>3</sub>-polluted days (i.e., MDA8 O<sub>3</sub>  $> 160 \mu\text{g}/\text{m}^3$ ) for each grid in eastern China for each year from 2013 to 2020 (Fig. 12). In 2013 and 2014, O<sub>3</sub> pollution was mainly found in the eastern and southern provinces of China. Overall, the probability of occurrence was generally low ( $< 10\%$ ) in most areas. The area covered by O<sub>3</sub> pollution generally decreased from 2014 to 2015, especially in southern China. From then on, the area covered by O<sub>3</sub> pollution continuously expanded until 2020, covering most areas of eastern China. More importantly, the probability of occurrence of O<sub>3</sub> pollution increased significantly from 2017 to 2019, especially in the NCP. For example, 23% of the days in 2019 exceeded the accepted O<sub>3</sub> standard. At the regional scale, the proportion of days exceeding the daily O<sub>3</sub> standard also gradually increased in four typical regions, reaching 21%, 12%, 7%, and 3% in the BTH, YRD, PRD, and SCB regions in 2019, respectively (Fig. 13). By contrast, the probability of occurrence of O<sub>3</sub> pollution declined in most areas of northern China (e.g., NCP, BTH, and YRD) in 2020. Similar conclusions have been reported in previous studies (Liu et al., 2020; Xue et al., 2020; Zhan et al., 2018).

Fig. 14 shows the evolution of MDA8 O<sub>3</sub> concentrations for each year

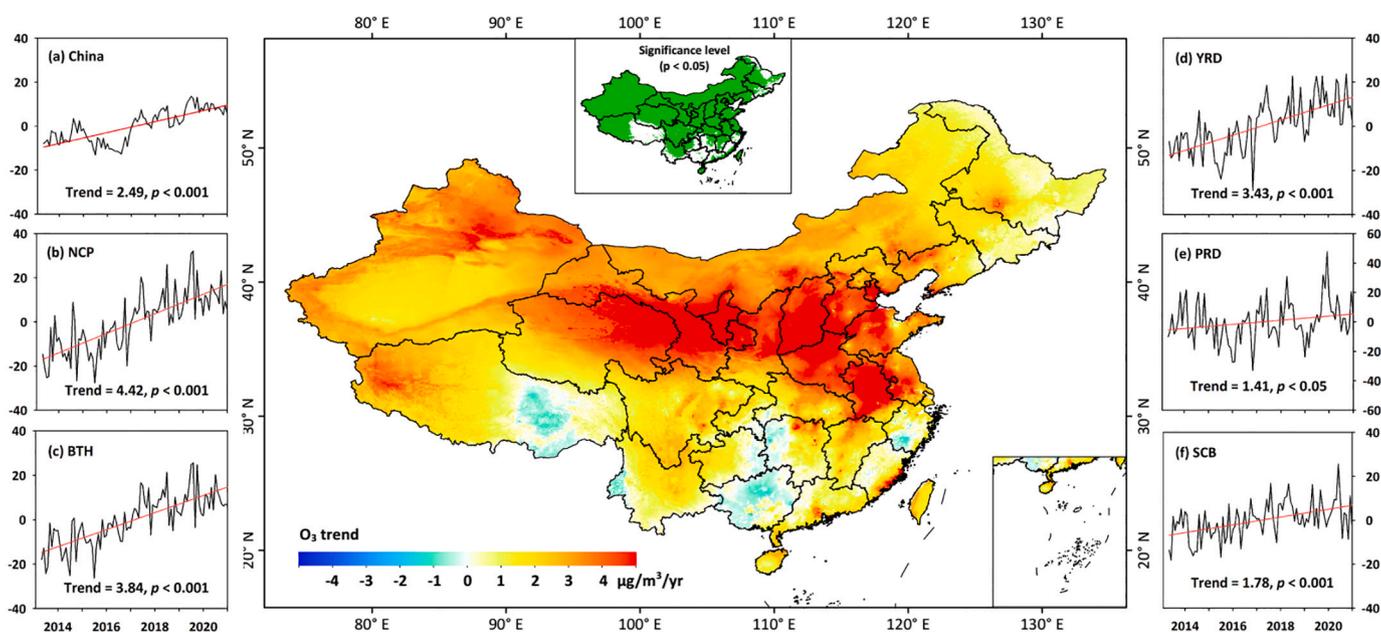


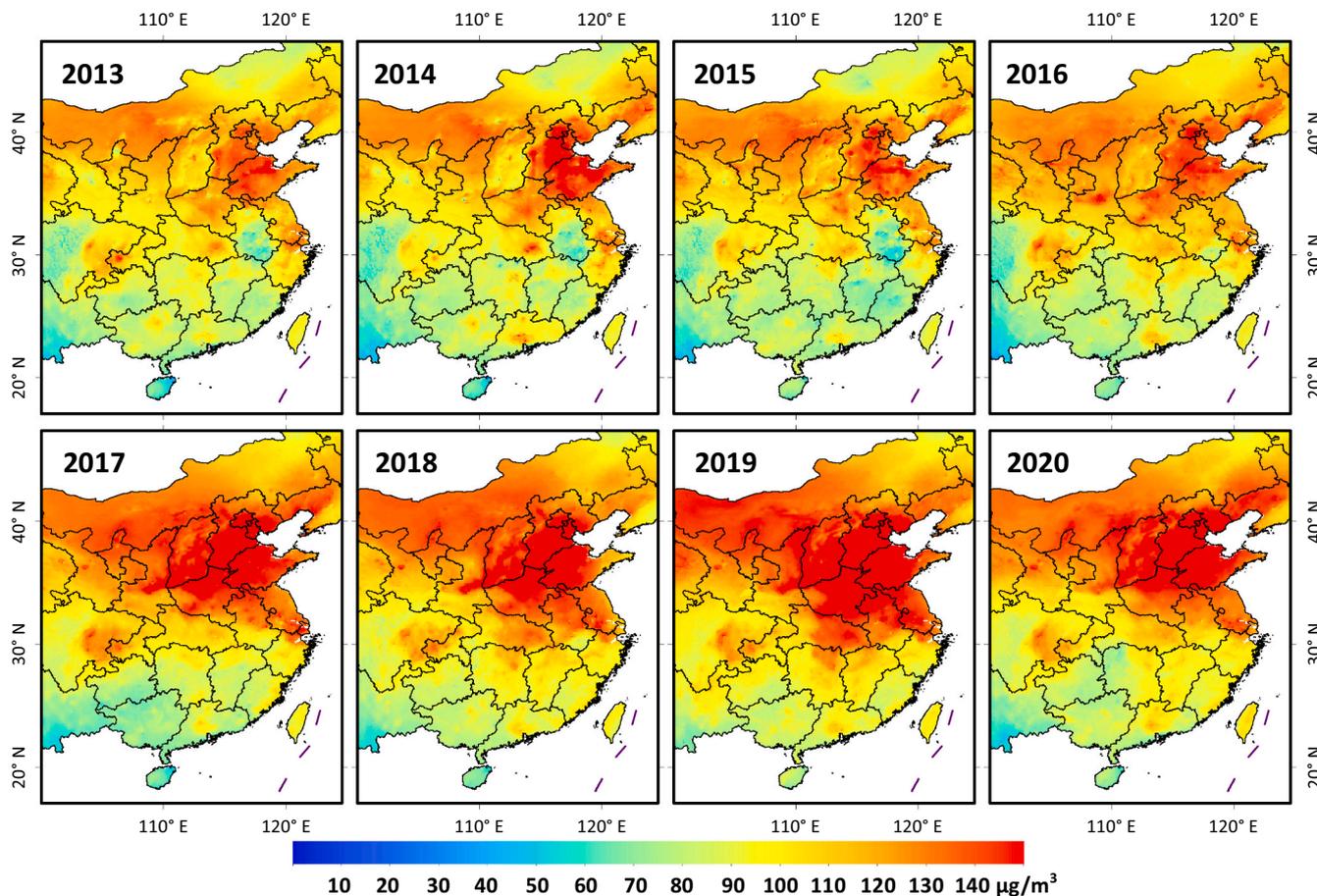
Fig. 10. Linear MDA8 O<sub>3</sub> trends (μg/m<sup>3</sup>/yr) calculated from de-seasonalized monthly MDA8 O<sub>3</sub> anomalies from 2013 to 2020 across China. The surrounding panels show the variations of monthly MDA8 O<sub>3</sub> anomalies in (a) China and (b–f) five typical regions.

**Table 2**

Statistics describing MDA8 O<sub>3</sub> trends (μg/m<sup>3</sup>/yr) and relative changes (%) in annual mean MDA8 O<sub>3</sub> concentrations (μg/m<sup>3</sup>) from 2013 to 2020 in China and in each typical region.

Region	2013–2020	2013–2017	2015–2019	2017	2020	2017–2020
	Trend (p)	Trend (p)	Trend (p)	Mean	Mean	Changed by
China	2.49 (<0.001)	1.33 (<0.01)	4.40 (<0.001)	91.8 ± 10.1	98.4 ± 10.8	7.2%
NCP	4.42 (<0.001)	4.58 (<0.001)	6.33 (<0.001)	108.8 ± 3.4	113.5 ± 4.1	4.3%
BTH	3.84 (<0.001)	4.78 (<0.001)	4.90 (<0.001)	104.8 ± 4.7	107.4 ± 7.2	2.5%
YRD	3.43 (<0.001)	2.94 (<0.01)	5.60 (<0.001)	102.8 ± 8.6	108.4 ± 8.2	5.4%
PRD	1.41 (<0.001)	-0.72 (0.56)	4.38 (<0.001)	89.8 ± 5.3	94.2 ± 6.0	4.9%
SCB	1.78 (<0.001)	2.37 (<0.001)	2.14 (<0.001)	82.9 ± 5.7	85.3 ± 5.8	2.8%

BTH: Beijing-Tianjin-Hebei; NCP: North China Plain; PRD: Pearl River Delta; SCB: Sichuan Basin; YRD: Yangtze River Delta.



**Fig. 11.** Spatial distributions of summertime mean MDA8 O<sub>3</sub> concentration (μg/m<sup>3</sup>) from 2013 to 2020 in eastern China.

at the “2 + 26” cities in northern China, where pollution is of particular concern to the public. Until 2015, O<sub>3</sub> concentrations were generally lower than 120 μg/m<sup>3</sup> in most cities, with much fewer days exceeding the air quality standard (i.e., MDA8 O<sub>3</sub> = 160 μg/m<sup>3</sup>) than those after 2016. Over time, the number of days with high O<sub>3</sub> concentrations gradually increased from year to year. In particular, a significant increase in O<sub>3</sub> concentration was captured from May to August in each year from 2017 to 2019. The MDA8 O<sub>3</sub> concentrations in almost all cities frequently exceeded 200 μg/m<sup>3</sup>, indicating a severe risk of ozone exposure.

### 3.3. Discussion

#### 3.3.1. Uncertainty and error analysis

We first investigated the effects of varying the number of training samples on modeled surface O<sub>3</sub> concentrations. For this purpose, we

gradually increased the proportion of training samples from 50% to 90% for model building, with the rest of the samples used for validation by applying different N-fold (i.e., 2, ..., 10) CV methods using 2020 data from China (Table S4). Overall, with an increase in training samples, the overall accuracy and spatial predictive ability of the STET model gradually improved, with increasing CV-R<sup>2</sup> values and decreasing estimation uncertainties. Small changes in each evaluation indicator were found, even when the training sample changed by as much as 40%, indicating that our model is stable and robust (e.g., CV-R<sup>2</sup> > 0.90 and RMSE < 14.1 μg/m<sup>3</sup>). This is mainly attributed to the unique advantage of the full-coverage mapping, which provides a large enough sample size to cover most surface O<sub>3</sub> conditions and variations across mainland China. It also benefits from the robustness of ensemble learning, which has a strong anti-noise ability (Breiman, 2001; Geurts et al., 2006).

We trained and built separate models for each characteristic region and compared their prominent features (Fig. S10) and model

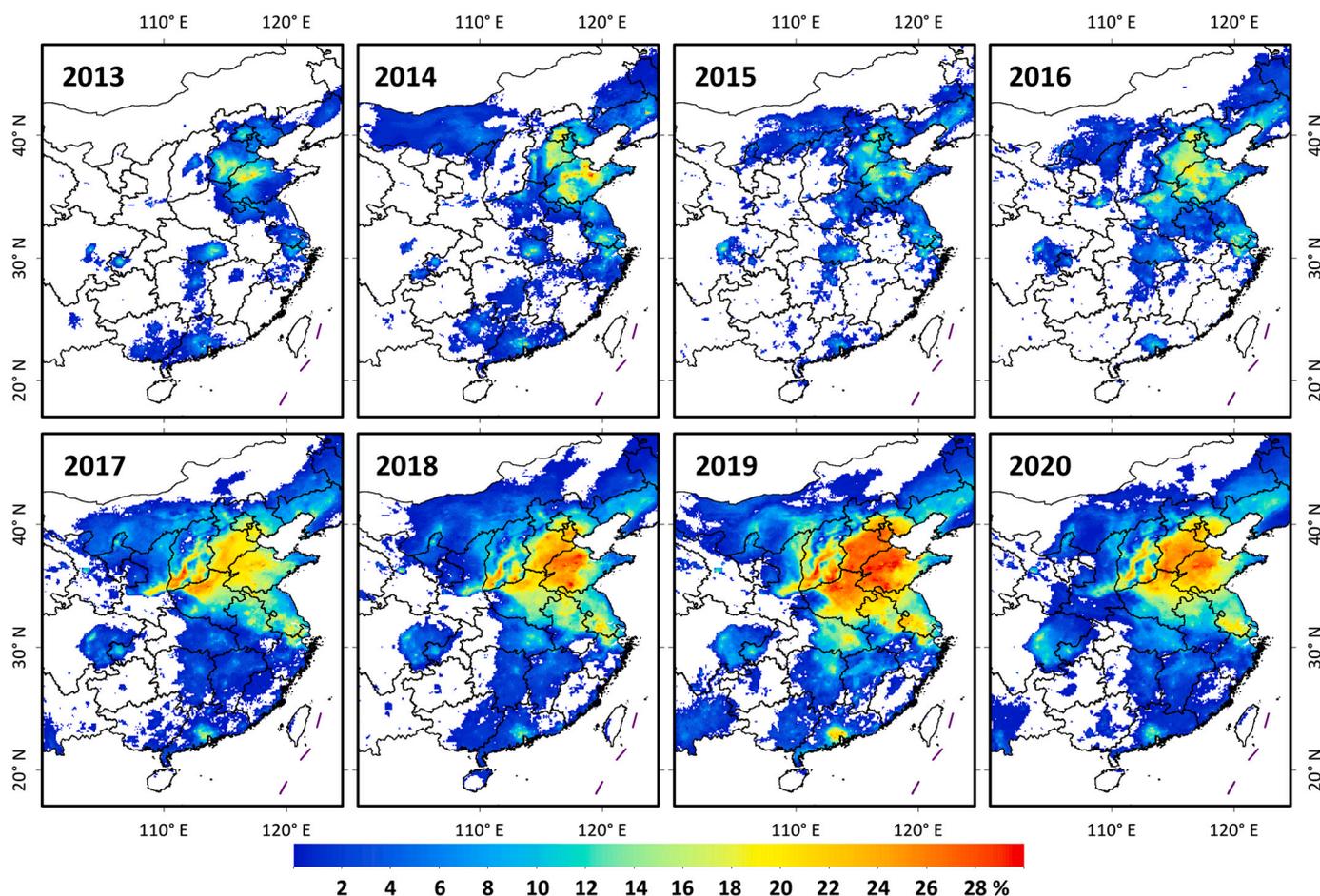


Fig. 12. Spatial distributions of the percentage of days exceeding the ambient O<sub>3</sub> standard (i.e., MDA8 O<sub>3</sub> concentrations >160 µg/m<sup>3</sup>) from 2013 to 2020 in China.

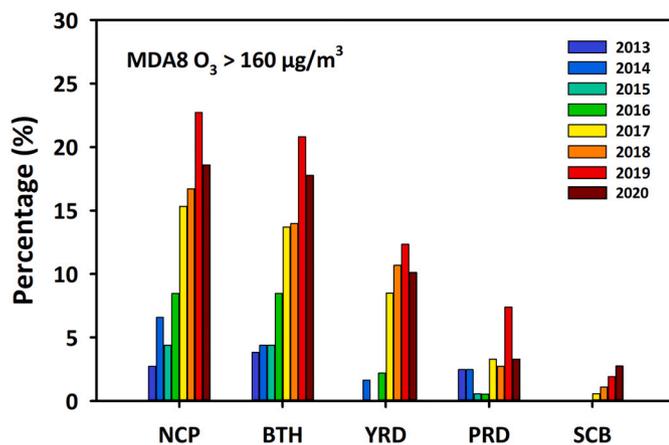


Fig. 13. Percentage of days exceeding the ambient air quality standard (i.e., MDA8 O<sub>3</sub> concentrations >160 µg/m<sup>3</sup>) from 2013 to 2020 in each typical region.

performances (Table S5) against the national model. The top-scoring features for the regional models are similar to those for the national model, e.g., ERA5 DSR, TEM, ET, RH, and OMI NO<sub>2</sub> and O<sub>3</sub> (Fig. S2). However, there were numerical differences in the importance scores for each variable. The model had different accuracies and spatial predictive abilities at the regional scale, with causes closely related to the density and spatial distribution of ground-level monitoring stations. Geographic, meteorological, and population conditions are also

different in each region. The performance of the national model was generally better with smaller estimation uncertainties than any one regional model, but differences in the statistical metrics were small. The national model involves a much bigger number of data samples that can cover more O<sub>3</sub> conditions. It can also consider the impact of adjacent regions, especially transition areas. Full-coverage mapping provides the richest dataset to train a robust model.

### 3.3.2. Comparison with related ozone datasets

We next compared our ChinaHighO<sub>3</sub> dataset with long-term atmospheric reanalysis products generated from chemical models, including MERRA-2 and ERA5, which have similar spatiotemporal coverages. For this purpose, 3-h MERRA-2 and 1-h ERA5 O<sub>3</sub> mixing ratio (unit: kg kg<sup>-1</sup>) simulations at horizontal resolutions of 0.625° × 0.5° and 0.25° × 0.25° were collected to calculate daily 14:00 local time MDA8 O<sub>3</sub> concentrations at the ground level (µg/m<sup>3</sup>) for the year 2020 in China, validated with corresponding ground-based measurements (Fig. S11). The ground-level O<sub>3</sub> simulations from the chemical reanalysis products are poor, with large uncertainties (e.g., R<sup>2</sup> < 0.1 and RMSE >47 µg/m<sup>3</sup>). The main reason is that the chemical reactions in the assimilation models are substantially simplified, mainly reflecting the impact of dynamic processes on stratospheric and tropospheric O<sub>3</sub> (Knowland et al., 2017). Our surface O<sub>3</sub> estimates are highly consistent with ground-based measurements (e.g., R<sup>2</sup> = 0.96 and RMSE = 8.6 µg/m<sup>3</sup>), a significant improvement over the chemical reanalysis products.

We also compared the O<sub>3</sub> estimates derived from our ensemble approach and from basic kriging techniques by simply interpolating data from all available O<sub>3</sub> stations in China. MDA8 maps obtained by two widely used kriging techniques, namely, ordinary kriging and universal

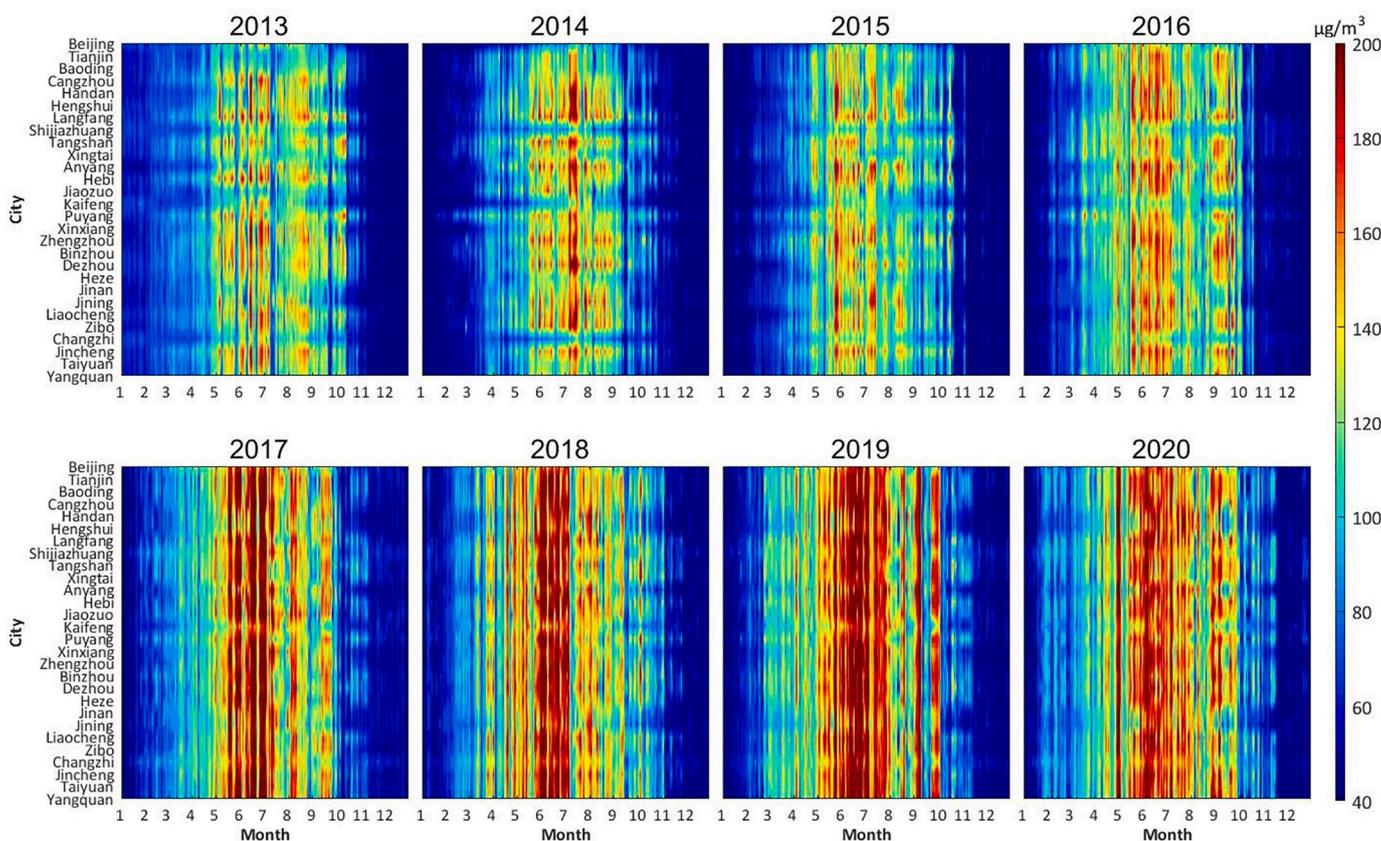


Fig. 14. Heat maps of MDA8 O<sub>3</sub> concentration (µg/m<sup>3</sup>) for each year from 2013 to 2020 at the “2 + 26” cities in China.

kriging, have similar spatial patterns and can basically capture information about surface O<sub>3</sub> (Fig. S12a–b). They are in good agreement with our results in southeastern China, with small differences within ±5%. However, significant differences >40% are seen in central, southwest, and northwest China (Fig. 12c–d). This is closely associated with the decreasing density of ground-based monitoring stations that the kriging methods are highly dependent on. In particular, severe block noise easily occurs in areas with few or no observation stations (areas outlined in red in Fig. S12). In addition, kriging results change smoothly in space in areas with complex terrains (e.g., southwest China). The main reason is that kriging only considers spatial correlations among observation stations, neglecting natural and human factors that impact air pollution (Fig. S2). Ensemble learning can make up for this deficit by making full use of data mining to build a robust conversion model from abundant, potentially influential factors on air pollution. Cross validation further illustrates that our model can obtain more reliable O<sub>3</sub> predictions by significantly decreasing the uncertainties by 16–28% compared to

kriging techniques.

### 3.3.3. Comparison with related studies

We compared results from our study with those from previous related studies, which also used the same out-of-sample 10-CV approach on MEE network O<sub>3</sub> observations for the same study period focusing on China (Table 3). Our algorithm yields a higher accuracy with smaller estimation uncertainties (CV-R<sup>2</sup> > 0.83, RMSE < 15 µg/m<sup>3</sup>) than the RF (CV-R<sup>2</sup> = 0.69, RMSE = 26.0 µg/m<sup>3</sup>; Zhan et al., 2018), XGBoost (CV-R<sup>2</sup> = 0.78, RMSE = 21.47 µg/m<sup>3</sup>; Liu et al., 2020), data fusion (CV-R<sup>2</sup> = 0.70, RMSE = 26.20 µg/m<sup>3</sup>; Xue et al., 2020), GWR (CV-R<sup>2</sup> = 0.77, MAE = 8.14 µg/m<sup>3</sup>; Zhang et al., 2020), and LUR/BME (CV-R<sup>2</sup> = 0.80, RMSE = 23.5 µg/m<sup>3</sup>; Chen et al., 2020) models at different temporal scales for the same study period.

In addition, different studies have relied on different main predictors, i.e., key variables input to the model for estimating surface O<sub>3</sub> concentrations, such as satellite-based total-column O<sub>3</sub>/NO<sub>2</sub> or CH<sub>2</sub>O,

Table 3  
Comparison of model performances based on previous O<sub>3</sub> studies focused on China as a whole.

Model	Temporal resolution	Validation			Study period	Main Predictors	Missing values	Literature
		R <sup>2</sup>	RMSE	MAE				
RF	Daily	0.69	26.00	–	2015	MERRA2	Yes	Zhan et al. (2018)
XGBoost	Daily	0.78	21.47	–	2013–2017	OMI O <sub>3</sub> , MERRA-2	Yes	Liu et al. (2020)
Data fusion	Daily	0.70	26.20	16.70	2013–2017	CTM simulations	Yes	Xue et al. (2020)
GWR	Monthly	0.77	–	8.14	2014	OMI NO <sub>2</sub> , CH <sub>2</sub> O	Yes	Zhang et al. (2020)
LUR/BME	Daily	0.80	23.50	–	2015–2017	In situ observations	Yes	Chen et al. (2020)
STET	Daily	0.78	21.16	14.09	2015	ERA5 DSR and TEM	No	This study
		0.81	20.27	13.38	2013–2017			
		0.83	18.88	12.72	2015–2017			
		0.90	12.43	8.82	2014			

LUR/BME: land-use regression/Bayesian maximum entropy; CTM: chemical transport model; GWR: geographically weighted regression; RF: random forest; STET: space-time extremely randomized trees; XGBoost: eXtreme Gradient Boosting.

MERRA-2 reanalysis data, model simulations, or in situ observations. These other O<sub>3</sub> datasets contain a large number of missing values at coarse or false (e.g., forced resampling) spatial resolutions (i.e., 0.25°–0.625°), limited by input data sources. Our study overcomes these issues and is an improvement on previous studies, providing a daily full-coverage (spatial coverage = 100%) and true-spatial-resolution (~0.1° × 0.1°) O<sub>3</sub> dataset for China generated from two main predictors, i.e., DSR and TEM. The dataset developed here constitutes a nearly continuous record of ground-level O<sub>3</sub> concentrations from 2013 to 2020 in China.

#### 4. Summary and conclusions

Ground-level O<sub>3</sub> is a major pollutant affecting human health. To compensate for the sparse and inhomogeneous coverage of ground-based O<sub>3</sub> networks and the low data quality, missing values, and low resolution of many existing satellite-based O<sub>3</sub> estimates, we applied a spatiotemporal extremely randomized trees machine-learning model to develop a long-term, near-surface ozone product that can overcome or lessen the above limitations. Besides O<sub>3</sub> training data, input variables include surface solar radiation downwards, air temperature, meteorological variables, land use information and topography, population distribution information, and a pollution emission inventory. The MDA8 O<sub>3</sub> product (ChinaHighO<sub>3</sub>) with full coverage across China at a spatial resolution of 10 km from 2013 to 2020 was generated.

The estimates were evaluated against surface observations at varying spatiotemporal scales and compared with previous related studies. CV results illustrate that our model has a high overall accuracy (spatial predictive ability), with average out-of-sample (out-of-station) CV-R<sup>2</sup>, RMSE, MAE, and MRE values of 0.87 (0.80), 17.10 (21.10) µg/m<sup>3</sup>, 11.29 (13.87) µg/m<sup>3</sup>, and 18.38 (23.18) %, respectively. Note that currently, we can only evaluate the surface O<sub>3</sub> predictions by removing parts of the base dataset using different 10-CV approaches. Also, note that assessing the accuracy of predictions in locations where O<sub>3</sub> measurements have never been made still remains a challenging task. Overall, the China-HighO<sub>3</sub> product is superior to existing ones in terms of model accuracy, spatial coverage and resolution, and data record length.

Benefiting from the unique advantages of the ChinaHighO<sub>3</sub> dataset, a recent (April to May 2020) short-term national and regional severe O<sub>3</sub> pollution event was well captured. Also observed was a rapid increase in O<sub>3</sub> pollution during the COVID-19 lockdown, especially in Hubei and surrounding provinces (e.g., an increase of >30%), followed by a return to normal levels after the lockdown ended in China. This was not a repudiation of NO<sub>x</sub> controls. A long-term analysis also showed that O<sub>3</sub> concentrations have significantly increased by 2.49 µg/m<sup>3</sup>/yr ( $p < 0.001$ ) in China from 2013 to 2020, especially in the North China Plain (~4.42 µg/m<sup>3</sup>/yr,  $p < 0.001$ ). In addition, summertime O<sub>3</sub> concentrations after 2017 were much higher than in previous years due to rising temperatures and an increase in the number of hot days. The number of days exceeding the ambient O<sub>3</sub> air quality standard (MDA8 O<sub>3</sub> = 160 µg/m<sup>3</sup>) and the areal extent of high O<sub>3</sub> levels were also shown to be gradually increasing across China, especially in the “2 + 26” cities in the North China Plain. Our ChinaHighO<sub>3</sub> dataset will thus be useful for related studies on air pollution in China, especially those studies focused on environmental health.

In this study, daily MDA8 O<sub>3</sub> concentrations were first estimated to examine daily and interannual changes in surface O<sub>3</sub> in China. Since hourly ground measurements and core input predictors (i.e., downwelling solar surface radiation and air temperature) are available, our model can be employed to reasonably predict hourly surface O<sub>3</sub> concentrations. This will help toward reproducing and investigating the diurnal patterns of and variations in surface O<sub>3</sub>. Considering similarly available ground measurements and after determining the appropriate core input predictors (e.g., aerosol optical depth and tropospheric gas column amounts), our model can also be applied and extended to estimate other species of ground-level air pollutants (e.g., PM, NO<sub>2</sub>, SO<sub>2</sub> and

CO). Our future studies will focus on this.

#### Data availability

The ChinaHighAirPollutants (CHAP) dataset is open access and freely available to all users from <https://weijing-rs.github.io/product.html>. The ChinaHighO<sub>3</sub> dataset is freely available at <https://doi.org/10.5281/zenodo.4400042>.

#### Declaration of Competing Interest

None.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China (42030606) and the National Key R&D Program of China (2017YFC1501702). J. Wang's participation was made possible via the in-kind support from the University of Iowa.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2021.112775>.

#### References

- Adam-Poupard, A., Brand, A., Fournier, M., Jerrett, M., Smargiassi, A., 2014. Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, land-use regression (LUR), and combined Bayesian maximum entropy-LUR approaches. *Environ. Health Perspect.* 122 (9), 970–976.
- Ainsworth, E., Yendrek, C., Sitch, S., Collins, W., Emberson, L., 2012. The effects of tropospheric ozone on net primary productivity and implications for climate change. *Annu. Rev. Plant Biol.* 63, 637–661.
- Almeida, S., Casimiro, E., Calheiros, J., 2011. Short-term association between exposure to ozone and mortality in Oporto, Portugal. *Environ. Res.* 111 (3), 406–410.
- Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Sci. Total Environ.* 407 (6), 1852–1867.
- Bell, M., McDermott, A., Zeger, S., Samet, J., Dominici, F., 2004. Ozone and short-term mortality in 95 US urban communities, 1987–2000. *JAMA* 292 (19), 2372–2378.
- Benish, S.E., He, H., Ren, X., Roberts, S.J., Salawitch, R.J., Li, Z., Wang, F., Wang, Y., Zhang, F., Shao, M., Lu, S., Dickerson, R.R., 2020. Measurement report: aircraft observations of ozone, nitrogen oxides, and volatile organic compounds over Hebei Province, China. *Atmos. Chem. Phys.* 20, 14523–14545.
- Bloomer, B.J., Stehr, J.W., Piety, C.A., Salawitch, R.J., Dickerson, R.R., 2009. Observed relationships of ozone air pollution with temperature and emissions. *Geophys. Res. Lett.* 36 (9) <https://doi.org/10.1029/2009GL037308>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA. ISBN 978-0-412-04841-8.
- Checa-Garcia, R., Hegglin, M.I., Kinnison, D., Plummer, D.A., Shine, K.P., 2018. Historical tropospheric and stratospheric ozone radiative forcing using the CMIP6 database. *Geophys. Res. Lett.* 45, 3264–3273.
- Chen, W., Liao, H., Seinfeld, J., 2007. Future climate impacts of direct radiative forcing of anthropogenic aerosols, tropospheric ozone, and long-lived greenhouse gases. *J. Geophys. Res. Atmos.* 112 (D14), D14209.
- Chen, L., et al., 2020. A hybrid approach to estimating long-term and short-term exposure levels of ozone at the national scale in China using land-use regression and Bayesian maximum entropy. *Sci. Total Environ.* 752, 141780. <https://doi.org/10.1016/j.scitotenv.2020.141780>.
- Di, Q., Rowland, S., Koutrakis, P., Schwartz, J., 2017. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. *J. Air Waste Manage. Assoc.* 67 (1), 39–52.
- Dickerson, R.R., Kondragunta, S., Stenchikov, G., et al., 1997. The impact of aerosols on solar ultraviolet radiation and photochemical smog. *Science* 278 (5339), 827–830.
- Dickerson, R., et al., 2007. Aircraft observations of dust and pollutants over Northeast China: insight into the meteorological mechanisms of transport. *J. Geophys. Res. Atmos.* 112 (D24) <https://doi.org/10.1029/2007JD008999>.
- Duan, J., Tan, J., Yang, L., Wu, S., Hao, J., 2008. Concentration, sources and ozone formation potential of volatile organic compounds (VOCs) during ozone episode in Beijing. *Atmos. Res.* 88, 25–35.
- Gaudel, A., Cooper, O.R., Ancellet, Gérard, Barret, B., Ziemke, J., 2018. Tropospheric ozone assessment report: present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation. *Elem. Sci. Anth.* 6 (1), 39.

- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. *Machine Learn.* 63 (1), 3–42.
- Giani, P., Castruccio, S., Anav, A., Howard, D., Hu, W., Crippa, P., 2020. Short-term and long-term health impacts of air pollution reductions from COVID-19 lockdowns in China and Europe: a modelling study. *Lancet Planet. Health* 4 (10), E474–E482.
- Goldberg, D.L., Loughner, C.P., Tzortziou, M., Stehr, J.W., Pickering, K.E., Marufu, L.T., Dickerson, R.R., 2014. Higher surface ozone concentrations over the Chesapeake Bay than over the adjacent land: observations and models from the DISCOVER-AQ and CBODAQ campaigns. *Atmos. Environ.* 84, 9–19.
- Gong, X., Hong, S., Jaffe, D.A., 2018. Ozone in China: spatial distribution and leading meteorological factors controlling O<sub>3</sub> in 16 Chinese cities. *Aerosol Air Qual. Res.* 18 (9), 2287–2300.
- He, J., Gong, S., Yu, Y., Yu, L., Wu, L., Mao, H., Song, C., Zhao, S., Liu, H., Li, X., Li, R., 2017. Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities. *Environ. Pollut.* 223, 484–496.
- Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. *Q. J. Roy. Meteor. Soc.* 146 (730), 1999–2049.
- Hu, J., Chen, J., Ying, Q., Zhang, H., 2016. One-year simulation of ozone and particulate matter in China using WRF/CMAQ modeling system. *Atmos. Chem. Phys.* 16, 10333–10350.
- Huang, L., Zhang, C., Bi, J., 2017. Development of land use regression models for PM<sub>2.5</sub>, SO<sub>2</sub>, NO<sub>2</sub> and O<sub>3</sub> in Nanjing, China. *Environ. Res.* 158, 542–552.
- Jiang, R., Tang, W., Wu, X., Fu, W., 2009. A random forest approach to the detection of epistatic interactions in case-control studies. *BMC Bioinform.* 10, 135.
- Kerckhoffs, J., Wang, M., Meliefste, K., Malmqvist, E., Fischer, P., Janssen, N.A.H., Beelen, R., Hoek, G., 2015. A national fine spatial scale land-use regression model for ozone. *Environ. Res.* 140, 440–448.
- Knowland, K., Ott, L., Duncan, B., Wargan, K., 2017. Stratospheric intrusion-influenced ozone air quality exceedances investigated in the NASA MERRA-2 reanalysis. *Geophys. Res. Lett.* 44 (20), 10,691–10,701.
- Krotkov, N., Lok, N., Sergey, V., Edward, A., Eric, J., William, H., Joanna, and the OMI core team, 2019. OMI/Aura NO<sub>2</sub> Cloud-Screened Total and Tropospheric Column L3 Global Gridded 0.25 degree x 0.25 degree V3, NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC). <https://doi.org/10.5067/Aura/OMI/DATA3007>.
- Lee, Y., Shindell, D., Faluvegi, G., Wenig, M., Lam, Y., Ning, Z., Hao, S., Lai, C., 2014. Increase of ozone concentrations, its temperature sensitivity and the precursor factor in South China. *Tellus Ser. B Chem. Phys. Meteorol.* 66, 23455.
- Li, M., Zhang, Q., Kurokawa, J., Woo, J., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D., Carmichael, G., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., Zheng, B., 2017a. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. *Atmos. Chem. Phys.* 17, 935–963.
- Li, T., Shen, H., Yuan, Q., Zhang, X., Zhang, L., 2017b. Estimating ground-level PM<sub>2.5</sub> by fusing satellite and station observations: a geo-intelligent deep learning approach. *Geophys. Res. Lett.* 44 (23), 11,985–11,993.
- Li, K., Jacob, D.J., Liao, H., Shen, L., Zhang, Q., Bates, K., 2019. Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China. *Proc. Natl. Acad. Sci. U. S. A.* 116 (2), 422–427.
- Li, K., Jacob, D.J., Shen, L., Lu, X., De Smedt, I., Liao, H., 2020a. Increases in surface ozone pollution in China from 2013 to 2019: anthropogenic and meteorological influences. *Atmos. Chem. Phys.* 20, 11,423–11,433.
- Li, R., Zhao, Y., Zhou, W., Meng, Y., Zhang, Z., Fu, H., 2020b. Developing a novel hybrid model for the estimation of surface 8h ozone (O<sub>3</sub>) across the remote Tibetan Plateau during 2005–2018. *Atmos. Chem. Phys.* 20, 6159–6175. <https://doi.org/10.5194/acp-20-6159-2020>.
- Lim, C., et al., 2019. Long-term exposure to ozone and cause-specific mortality risk in the U.S. *Am. J. Respir. Crit. Care Med.* 200 (8) <https://doi.org/10.1164/rccm.201806-1161OC>.
- Lin, X., Trainer, M., Liu, S.C., 1988. On the nonlinearity of the tropospheric ozone production. *J. Geophys. Res. Atmos.* 93 (D12), 15879–15888.
- Liu, X., Bhartia, P., Chance, K., Spurr, R., Kurosu, T., 2010. Ozone profile retrievals from the ozone monitoring instrument. *Atmos. Chem. Phys.* 10, 2521–2537. <https://doi.org/10.5194/acp-10-2521-2010>.
- Liu, R., Ma, Z., Liu, Y., Shao, Y., Zhao, W., Bi, J., 2020. Spatiotemporal distributions of surface ozone levels in China from 2005 to 2017: a machine learning approach. *Environ. Int.* 142, 105823. <https://doi.org/10.1016/j.envint.2020.105823>.
- Liu, Y., Wang, T., Stavrou, T., Elguindi, N., Doumbia, T., Granier, C., Bouarar, I., Gaubert, B., Brasseur, G., 2021. Diverse response of surface ozone to COVID-19 lockdown in China. *Sci. Total Environ.* 789, 147739 <https://doi.org/10.1016/j.scitotenv.2021.147739>.
- Loughner, C.P., Allen, D.J., Pickering, K.E., Zhang, D.L., Shou, Y.X., Dickerson, R.R., 2011. Impact of fair-weather cumulus clouds and the Chesapeake Bay breeze on pollutant transport and transformation. *Atmos. Environ.* 45 (24), 4060–4072.
- Lu, X., et al., 2018. Severe surface ozone pollution in China: a global perspective. *Environ. Sci. Technol. Lett.* 5 (8), 487–494.
- Lu, X., Zhang, L., Wang, X., Gao, M., Li, K., Zhang, Y., Yue, X., Zhang, Y., 2020. Rapid increases in warm-season surface ozone and resulting health impact in China since 2013. *Environ. Sci. Technol. Lett.* 7 (4), 240–247.
- Ma, Z., Zhang, X., Xu, J., Zhao, X., Meng, W., 2011. Characteristics of ozone vertical profile observed in the boundary layer around Beijing in autumn. *J. Environ. Sci.* 23, 1316–1324.
- Meleux, F., Solmon, F., Giorgi, F., 2007. Increase in summer European ozone amounts due to climate change. *Atmos. Environ.* 41, 7577–7587.
- Mills, G., et al., 2018. Tropospheric ozone assessment report: present-day tropospheric ozone distribution and trends relevant to vegetation. *Elem. Sci. Anth.* 6 (1), 47.
- Ministry of Ecology and Environment (MEE), 2018. Revision of the Ambient air quality standards (GB 3095–2012) (in Chinese). available at: [http://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/201808/t20180815\\_629602.html](http://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/201808/t20180815_629602.html).
- Pawan, K., 2012. OMI/Aura Ozone (O<sub>3</sub>) Total Column Daily L2 Global Gridded 0.25 degree x 0.25 degree V3, Goddard Earth Sciences Data and Information Services Center (GES DISC). <https://doi.org/10.5067/Aura/OMI/DATA2025>.
- Qiao, X., Guo, H., Wang, P., Tang, Y., Ying, Q., Zhao, X., Deng, W., Zhang, H., 2019. Fine particulate matter and ozone pollution in the 18 cities of the Sichuan Basin in southwestern China: model performance and characteristics. *Aerosol Air Qual. Res.* 19, 2308–2319. <https://doi.org/10.4209/aaqr.2019.05.0235>.
- Rai, R., Agrawal, M., 2012. Impact of tropospheric ozone on crop plants. *Proc. Natl. Acad. Sci. India Section B: Biol. Sci.* 82, 241–257.
- Rodriguez, J., Perez, A., Lozano, J., 2010. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (3), 569–575.
- Sánchez-Ccoyllo, O.R., Ynoue, R.Y., Martins, L.D., de Fátima Andrade, M., 2006. Impacts of ozone precursor limitation and meteorological variables on ozone concentration in São Paulo, Brazil. *Atmos. Environ.* 40, 552–562.
- Shen, L., Jacob, D.J., Liu, X., Huang, G., Li, K., Liao, H., Wang, T., 2019. An evaluation of the ability of the ozone monitoring instrument (OMI) to observe boundary layer ozone pollution across China: application to 2005–2017 ozone trends. *Atmos. Chem. Phys.* 19, 6551–6560.
- Shi, X., Brasseur, G.P., 2020. The response in air quality to the reduction of Chinese economic activities during the COVID outbreak. *Geophys. Res. Lett.* 47, 11. <https://doi.org/10.1029/2020GL088070>.
- Shindell, D., Faluvegi, G., Nazarenko, L., Bowman, K., Lamarque, J., Voulgarakis, A., Schmidt, G., Pechony, O., Ruedy, R., 2013. Attribution of historical ozone forcing to anthropogenic emissions. *Nat. Clim. Chang.* 3 (6), 567–570.
- Sinha, A., Toumi, R., 1997. Tropospheric ozone, lightning, and climate change. *J. Geophys. Res. Atmos.* 102 (D9), 10,667–10,672.
- Sitch, S., Cox, P., Collins, W., Huntingford, C., 2007. Indirect radiative forcing of climate change through ozone effects on the land-carbon sink. *Nature* 448 (7155), 791–794.
- Son, Y., et al., 2018. Land use regression models to assess air pollution exposure in Mexico City using finer spatial and temporal input parameters. *Sci. Total Environ.* 639, 40–48.
- Su, T., Li, Z., Zheng, Y., Luan, Q., Guo, J., 2020. Abnormally shallow boundary layer associated with severe air pollution during the COVID-19 lockdown in China. *Geophys. Res. Lett.* 47, e2020GL090041 <https://doi.org/10.1029/2020GL090041>.
- Taubman, B.F., Hains, J.C., Thompson, A.M., Marufu, L.T., Doddridge, B.G., Stehr, J.W., Piety, C.A., Dickerson, R.R., 2006. Aircraft vertical profiles of trace gas and aerosol pollution over the mid-Atlantic United States: statistics and meteorological cluster analysis. *J. Geophys. Res. Atmos.* 111 (D10).
- Tian, H., et al., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368 (6491), 638–642.
- Turner, M., et al., 2015. Long-term ozone exposure and mortality in a large prospective study. *Am. J. Respir. Crit. Care Med.* 193 (10), 1134–1142.
- Wang, N., Guo, H., Jiang, F., Ling, Z., Wang, T., 2015. Simulation of ozone formation at different elevations in mountainous area of Hong Kong using WRF-CMAQ model. *Sci. Total Environ.* 505, 939–951.
- Wang, M., et al., 2016a. Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and PM<sub>2.5</sub>. *Environ. Sci. Technol.* 50 (10), 5111–5118.
- Wang, T., Xue, L., Brimblecombe, P., Lam, Y.F., Li, L., Zhang, L., 2016b. Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects. *Sci. Total Environ.* 575, 1582–1596.
- Wang, Y., et al., 2020. Contrasting trends of PM<sub>2.5</sub> and surface ozone concentrations in China from 2013 to 2017. *Natl. Sci. Rev.* 7 (8), 1331–1339.
- Wang, S., Zhang, Y., Ma, J., Zhu, S., Shen, J., Wang, P., Zhang, H., 2021. Responses of decline in air pollution and recovery associated with COVID-19 lockdown in the Pearl River Delta. *Sci. Total Environ.* 756, 143868. <https://doi.org/10.1016/j.scitotenv.2020.143868>.
- Wei, J., Peng, Y., Mahmood, R., Sun, L., Guo, J., 2019. Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products. *Atmos. Chem. Phys.* 19, 7183–7207.
- Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., Song, Y., 2020. Improved 1-km resolution PM<sub>2.5</sub> estimates across China using enhanced space-time extremely randomized trees. *Atmos. Chem. Phys.* 20 (6), 3273–3289.
- Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M.C., 2021a. Reconstructing 1-km-resolution high-quality PM<sub>2.5</sub> data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. *Remote Sens. Environ.* 252, 112136. <https://doi.org/10.1016/j.rse.2020.112136>.
- Wei, J., Li, Z., Xue, W., Sun, L., Fan, T., Liu, L., Su, T., Cribb, M., 2021b. The ChinaHighPM<sub>10</sub> dataset: generation, validation, and spatiotemporal variations from 2015 to 2019 across China. *Environ. Int.* 146, 106290. <https://doi.org/10.1016/j.envint.2020.106290>.
- Wei, J., Li, Z., Pinker, R., Wang, J., Sun, L., Xue, W., Li, R., Cribb, M., 2021c. Himawari-8-derived diurnal variations of ground-level PM<sub>2.5</sub> pollution across China using the fast space-time Light Gradient Boosting Machine (LightGBM). *Atmos. Chem. Phys.* 21 (10), 7863–7880.
- WHO, 2020. The World Health Organization. Coronavirus Disease (COVID-19) Pandemic. Available online: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- Wong, C., Vichit-Vadakan, N., Kan, H., Qian, Z., 2007. Public health and air pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. *Environ. Health Perspect.* 116 (9), 1195–1202.

- Wu, S., Huang, B., Wang, J., He, L., Wang, Z., Yan, Z., Lao, X., Zhang, F., Liu, R., Du, Z., 2021. Spatiotemporal mapping and assessment of daily ground NO<sub>2</sub> concentrations in China using high-resolution TROPOMI retrievals. *Environ. Pollut.* 273, 116456.
- Xiang, S., et al., 2020. Control of both PM<sub>2.5</sub> and O<sub>3</sub> in Beijing-Tianjin-Hebei and the surrounding areas. *Atmos. Environ.* 224, 117259.
- Xue, T., et al., 2020. Estimating spatiotemporal variation in ambient ozone exposure during 2013–2017 using a data-fusion model. *Environ. Sci. Technol.* 54, 14,877–14,888.
- Xue, W., Zhang, J., Zhong, C., Li, X., Wei, J., 2021. Spatiotemporal PM<sub>2.5</sub> variations and its response to the industrial structure from 2000 to 2018 in the Beijing-Tianjin-Hebei region. *J. Clean Prod.* 279, 123742 <https://doi.org/10.1016/j.jclepro.2020.123742>.
- Zdaniuk, B., 2014. *Ordinary Least-Squares (OLS) Model*. Springer Netherlands.
- Zhan, Y., Luo, Y., Deng, X., Grieneisen, M.L., Zhang, M., Di, B., 2018. Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment. *Environ. Pollut.* 233, 464–473. <https://doi.org/10.1016/j.envpol.2017.10.029>.
- Zhang, Q., et al., 2019. Drivers of improved PM<sub>2.5</sub> air quality in China from 2013 to 2017. *Proc. Natl. Acad. Sci. U. S. A.* 116 (49), 24,463–24,469. <https://doi.org/10.1073/pnas.1907956116>.
- Zhang, X., Zhao, L., Cheng, M., Chen, D., 2020. Estimating ground-level ozone concentrations in eastern China using satellite-based precursors. *IEEE Trans. Geosci. Remote Sens.* 58 (7), 4754–4763.
- Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., Zhang, Q., 2018. Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmos. Chem. Phys.* 18, 14095–14111.
- Zu, Z.Y., Jiang, M.D., Xu, P.P., Chen, W., Ni, Q.Q., Lu, G.M., Zhang, L.J., 2020. Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology* 296 (2), 200490.