Full-coverage mapping and spatiotemporal variations of ground-level ozone (O3) pollution from 2013 to 2020 across China

Wei Jing[1], Li Zhanqing[2], Li Ke[3], Dickerson Russell R.[2], Pinker Rachel T.[2], Wang Jun[1], Liu Xiong[4], Sun Lin[5], Xue Wenhao[6], and Cribb Maureen[2]

[1]University of Iowa
[2]University of Maryland
[3]Harvard University
[4]Harvard Smithsonian Center for Astrophysics
[5]Shandong University of Science and Technology
[6]Beijing Normal University

November 15, 2022

Abstract

Ozone (O3) is an important trace and greenhouse gas in the atmosphere yet, and it threatens the ecological environment and human health at the ground level. Large-scale and long-term studies of O3 pollution in China are few due to highly limited direct measurements whose accuracy and density vary considerably. To overcome these limitations, we employed the ensemble learning method of the extremely randomized trees model by utilizing the spatiotemporal information of a large number of input variables from ground-based observations, remote sensing, atmospheric reanalysis, and model simulation products to estimate ground-level O3. This method yields uniform, long-term and continuous spatiotemporal information of daily maximum eight-hour average (MDA8) O3 over China (called ChinaHighO3) from 2013 to 2020 at a 10 km resolution without any missing values (spatial coverage = 100%). Evaluation against observations indicates that our O3 estimations and predictions are reliable with an average out-of-sample (out-of-station) coefficient of determination (CV-R2) of 0.87 (0.80) and root-mean-square error of 17.10 (21.10) μg/m3 [units here are at standard conditions (273K, 1013hPa)], and are also robust at varying spatial and temporal scales in China. This high-quality and full-coverage O3 dataset allows us to investigate the exposure and trends in O3 pollution at both long- and short-term scales. Trends in O3 concentrations varied substantially but showed an average growth rate of 2.49 μg/m3/yr ($p < 0.001$) from 2013 to 2020 in China. Most areas show an increasing trend since 2015, especially in summer ozone over the North China Plain. Our dataset accurately captured a recent national and regional O3 pollution event from 23 April to 8 May in 2020. Rapid increase and recovery of O3 concentrations associated with variations in anthropogenic emissions were seen during and after the COVID-19 lockdown, respectively. This carefully vetted and smoothed dataset is valuable for studies on air pollution and environmental health in China.

Full-coverage mapping and spatiotemporal variations of ground-level ozone (O_3) pollution from 2013 to 2020 across China

Jing Wei[1,2*], Zhanqing Li[2*], Ke Li[3], Russell R. Dickerson[2], Rachel T. Pinker[2], Jun Wang[1], Xiong Liu[4], Lin Sun[5], Wenhao Xue[6], Maureen Cribb[2]

1. Department of Chemical and Biochemical Engineering, Iowa Technology Institute, and Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA, USA
2. Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
3. Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
4. Atomic and Molecular Physics Division, Harvard Smithsonian Center for Astrophysics, Cambridge, MA, USA
5. College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, China
6. College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

* Corresponding authors: zli@atmos.umd.edu; weijing_rs@163.com

Abstract

Ozone (O_3) is an important trace and greenhouse gas in the atmosphere yet, and it threatens the ecological environment and human health at the ground level. Large-scale and long-term studies of O_3 pollution in China are few due to highly limited direct measurements whose accuracy and density vary considerably. To overcome these limitations, we employed the ensemble learning method of the extremely randomized trees model by utilizing the spatiotemporal information of a large number of input variables from ground-based observations, remote sensing, atmospheric reanalysis, and model simulation products to estimate ground-level O_3. This method yields uniform, long-term and continuous spatiotemporal information of daily maximum eight-hour average (MDA8) O_3 over China (called ChinaHighO_3) from 2013 to 2020 at a 10 km resolution without any missing values (spatial coverage = 100%). Evaluation against observations indicates that our O_3 estimations and predictions are reliable with an average out-of-sample (out-of-station) coefficient of determination (CV-R^2) of 0.87 (0.80) and root-mean-square error of 17.10 (21.10) $\mu g/m^3$ [units here are at standard conditions (273K, 1013hPa)], and are also robust at varying spatial and temporal scales in China. This high-quality and full-coverage O_3 dataset allows us to investigate the exposure and trends in O_3 pollution at both long- and short-term scales. Trends in O_3 concentrations varied substantially but showed an average growth rate of 2.49 $\mu g/m^3/yr$ ($p < 0.001$) from 2013 to 2020 in China. Most areas show an increasing trend since 2015, especially in summer ozone over the North China Plain. Our dataset accurately captured a recent national and regional O_3 pollution event from 23 April to 8 May in 2020. Rapid increase and recovery of O_3 concentrations associated with variations in anthropogenic emissions were seen during and after the COVID-19 lockdown, respectively. This carefully vetted and smoothed dataset is valuable for studies on air pollution and environmental health in China.

Keywords: Ozone, ensemble learning, COVID-19, air pollution, China

1. Introduction

Ozone (O_3) is an important atmospheric trace gas, where O_3 in the stratosphere plays a crucial role in absorbing ultraviolet rays, protecting the environment and humans. Tropospheric O_3 (< 12 km above the ground) is mainly produced by anthropogenic activity and affects radiative forcing at global scale with implications on climate change (Sinha et al., 1997; Chen et al., 2007; Shindell et al., 2013; Checa-Garcia et al., 2018; Gaudel et al., 2018). Exposure to high surface O_3 is highly related to increased human health risks, including cardiovascular and respiratory disease (Bell et al., 2004; Turner et al., 2015; Lim et al., 2019). It also affects the ecosystem and agricultural production, e.g., inhibiting plant growth, promoting leaf senescence, and affecting crop yields (Sitch et al., 2007; Ainsworth et al., 2012; Rai and Agrawal, 2012; Mills et al., 2018).

Since the middle of the 20th century, many countries around the world have observed tropospheric and ground-level O_3. In 2013, the Chinese Ministry of Environment and Ecology (MEE) established a national air quality observation network to monitor real-time O_3, particulate matters (PM), and other near-surface air pollutants (MEE, 2018). However, the construction and maintenance of ground networks require a lot of manpower and material resources, and as such, monitoring stations are sparsely distributed. Satellite remote sensing can make up for such deficiency by providing spatially continuous atmospheric O_3 distributions. The OMI/Aura satellite, launched in 2004, provides a variety of widely-used daily, global-coverage trace gas products (e.g., O_3, NO_2, and SO_2). Existing techniques from space mainly provide the total column O_3, tropospheric O_3, and ozone profiles at different vertical ranges (Liu et al., 2010; Huang et al., 2018; Ziemke et al., 2006). Near-surface O_3 typically accounts only for a few percent of total column O_3, and the retrieval sensitivity to near-surface O_3 from ultraviolet measurements is limited. In some cases, tropospheric totals can be helpful for understanding global and regional scale features, but values for ozone in the planetary boundary layer are challenging and at exposure heights (~2 m) even more so. Thus, it is particularly difficult to extract the near-surface O_3 from satellite measurements.

In recent years, great effort has been made to estimate near-surface O_3 concentrations using three main methodologies: chemical transport models, statistical models, and artificial intelligence. Chemical transport methods mainly use mature models, e.g., WRF-Chem, CMAQ, and GEOS-Chem, to simulate the O_3 at the ground level by considering chemical reactions and transport of air pollutants (Di et al., 2017; Hu et al., 2016; Qiao et al., 2019; Wang et al., 2015, 2016). Statistical models fit the relationships between the measured air pollution and their potential influence 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; Song et al., 2018), Bayesian maximum entropy (BME; Adam-Poupart et al., 2014; Chen et al., 2020), 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 to obtain more accurate parameter estimates by mining valuable information from big data using different methods, e.g., neural networks (Di et al., 2017), random forests (RF; Li et al., 2020b; Zhan et al., 2018), and XGBoost (Liu et al., 2020).

In general, chemical/numerical methods can provide high spatiotemporal coverage of near-surface O_3 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., NCEP, MERRA2, and ERA-Interim, at low spatiotemporal resolutions (e.g., 3–6 h, 0.25°–0.625°).

Over the years, due to implemented environmental protection and control measures, PM pollution has decreased significantly (Zhang et al., 2019; Wang et al., 2020; Wei et al., 2021a), but O_3 pollution increased in China (Wang, et al., 2017; Lu et al., 2018; Li et al., 2019; Wang, 2020), attracting the major public health concern (Shen et al., 2019). Compared with PM studies, research on ground-level O_3 is more meager for China. Therefore, aimed at addressing the above problems, according to the idea of ensemble learning and considering the spatiotemporal variations in O_3 pollution, we extended a space-time extremely randomized trees (STET) model to derive the daily ground-level O_3 with full spatial coverage at a resolution of 10 km from 2013 to 2020 in China. Subsequently, we have tested the reliability of our O_3 retrievals at different spatial and temporal scales and investigated the short-term (i.e., on a daily basis) and long-term (i.e., multi-year time series) exposure and trends in O_3 pollution across China.

2. Materials and methods

2.1 Data sources

2.1.1 MEE network O_3 observations

Used are hourly ground-based O_3 concentrations [in μg/m^3 at standard conditions (273K, 1013hPa)] collected by MEE across mainland China starting from an initial ~940 monitoring stations in 2013 and ending with ~1630 stations by 2020 (Figure S1). We first removed invalid values and abnormal values due to instrument calibration issues (Guo et al., 2009). 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_3 concentrations (in μg/m^3) are thus correspondingly rescaled by a factor of 1.09375 (MEE, 2018). For data presented here, 1 μg/m^3 is equivalent to 0.467 ppbv. Additionally, we averaged maximum O_3 concentrations over eight hours in a day to obtain MDA8 O_3 values at each station in China for each year from January 1 2013 to December 31 2020.

2.1.2 Potential factors affecting surface O_3

Surface O_3, 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 the total-column or tropospheric O_3 retrievals, rather than lower tropospheric O_3 where there are large differences in O_3 content. Long-term satellite O_3 products with high spatial resolutions are rarely available, and these satellite retrievals have numerous missing values. In our study, we provide a new approach for estimating high-resolution surface O_3 with full coverage by using two crucial meteorological parameters, namely, solar radiation intensity and surface temperature (Bloomer et al., 2009; Lee et al., 2014; Li et al., 2020).

Atmospheric reanalysis. Available surface downward solar radiation (DSR) and air temperature (TEM) are used for ground-level O_3 estimation. Other meteorological variables can also affect O_3, e.g., an increase in relative humidity (RH) and surface pressure (SP) can pose diverse effects on O_3

concentrations in the lower troposphere (Taubman et al., 2006; Loughner et al., 2011; He et al., 2017). A change in planetary boundary layer height (BLH) can have variable impacts on O_3 pollution (Sánchez-Ccoyllo et al., 2006; Dickerson et al., 2007; Ma et al., 2011; Goldberg et al., 2014; Benish et al., 2020). Winds, i.e., horizontal (WU) and vertical (WV) components, can affect the transport of O_3 and produce high O_3 levels in the downwind direction (Dickerson et al., 2007; Duan et al., 2008; Ma et al., 2011; Benish et al., 2020). Precipitation (PRE) and evaporation (ET) can also influence O_3 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 data at a high spatial resolution reaching up to 0.1° × 0.1° (C3S, 2017). The spatiotemporal resolution of ERA5 reanalysis is higher than from other atmospheric reanalysis products (e.g., NCEP and MERRA2) used in previous studies (Di et al., 2017; Li et al, 2020b; Liu et al., 2020; Zhan et al., 2018).

Satellite remote sensing products. Remote sensing measurements of OMI/Aura total-column O_3 products (Pawan, 2012) are also considered. NO_2 concentrations may have large impacts on O_3; thus OMI/Aura tropospheric NO_2 products (Nickolay et al., 2019) are utilized. LandScan™ product is also selected to provide the population distribution (POP), and MODIS land cover type (LUC) and NDVI products, and SRTM DEM data are employed to describe the land-use and terrain changes across China.

Model simulations. Anthropogenic emissions from fossil fuel combustion, industrial production, and vehicle exhaust are precursors affecting the formation of surface O_3 concentrations (Li et al., 2020). Therefore, the direct emissions of three main O_3 precursors, i.e., NO_x, VOCs, and CO, are provided by the Multiresolution Emission Inventory for China (MEIC) (Li et al., 2017; Zheng et al., 2018) are used.

Table 1 summarizes the ground-based, satellite remote sensing, atmospheric reanalysis, and model data 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) have smaller variations in space than meteorological variables, while they (e.g., DEM, LUC, and POP) change little over time. In addition, they are generally less important than two main predictors (i.e., DSR and TEM) in estimating surface O_3 (Figure 1).

Therefore, similar to previous studies (Zhan et al., 2018; Liu et al., 2020; Zhang et al., 2020; Wei et al., 2021b), all the finer and coarser-resolution auxiliary data are resampled (regrided) to the same spatial resolution of 0.1°×0.1° using the bilinear interpolation approach, and the same time interval.

[Please insert Table 1 here]

2.2 Space-Time Extra-Trees modeling

In this study, a space-time Extra-Trees (STET) model is extended for estimating the ground-level O_3 concentrations (Wei et al., 2021a). It is based on the ensemble learning named the extremely randomized trees (extra-trees, or ERT) (Geurts et al., 2006).

2.2.1 Model training

First, all the selected factors with potential effects on surface O_3 are input to the ERT model for model training. This, to quantitatively evaluate the contribution of each variable on O_3 to see if further model adjustments are needed by removing redundant variables. Four main steps are followed:

1) A training and validation dataset (*N*) is generated by collocating the surface O_3 measurements, satellite data, meteorological variables, and model emissions at each surface monitor for each day in one year. Then the entire training dataset is used to construct each decision tree.
2) For each binary tree, a random split (*S, a*) is first generated according to the surface O_3 measurements by randomly selecting one arbitrary number (a_c) between the maximum (a_{max}) and the minimum (a_{min}) value; next, the training samples are randomly assigned to two branches.
3) All the auxiliary feature attributes (a_1, …, a_k) in the node are traversed to get the bifurcation values (s_1, …, s_k) for all feature attributes based on the Gini index (Jiang et al., 2009). Then the best split (s*) is determined when satisfying the scoring function: score(s*, S) = max[Score(s_i, S)] (Geurts et al., 2006).
4) A decision tree is established using the CART algorithm (Breiman et al., 1984), and then thousands of decision trees are constructed by repeating the above steps. Last, all the weak

classifiers are combined to form a strong classifier, i.e., extremely randomized trees, allowing parallel processing.

The ERT model enables us to evaluate the importance of each independent variable for surface O_3 estimation, named the importance score, calculated according to the Gini index (GI). It normalized the cumulative changes of GIs before and after node branching for each feature during the model training (Jiang et al., 2009). The higher this score, the more important this feature in the decision-tree construction. The variables with high scores make great contributions to the model performance; by contrast, low-score variables may pose small effects on the model or even bring redundant information (Wei et al., 2020, 2021a). Variables with an importance score of 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 (Figure 1). Satellite OMI NO_2 and O_3 products are also highly valuable with importance scores of 6% and 5%; but they can only provide trace gas information of 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_3 estimates with importance scores from 2% to 7%. The emissions of three main O_3 precursors (i.e., NO_x, VOCs, and CO) have influences on the model, with importance scores of about 2%. In general, all 18 selected independent variables have an impact with importance scores > 1.5%, which cannot be neglected, and are kept in the model.

[Please insert Figure 1 here]

2.2.2 Model extension

In the second stage, we extended a STET model for surface O_3 estimation by considering the autocorrelation of O_3 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_3 estimation specific to China.

2.3 Validation method

In this study, the out-of-sample (sample-based) 10-fold cross-validation (10-CV) method, which has been widely adopted, is selected to test the overall model performance in near-surface O_3 estimates (Di et al., 2017; Li et al., 2020b; Liu et al., 2020; Zhan et al., 2018). It 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 samples have been all tested (Rodriguez et al., 2010). Furthermore, an additional out-of-station (station-based) 10-CV method is employed to test the spatial prediction ability of the model in areas without ground-based measurements (Li et al., 2017; Wei et al., 2020; Wu et al., 2021). It is performed using the ground-based O_3 monitoring stations; the monitors are randomly divided into 10 groups, of which the data samples from 9 groups (i.e., 90% of the monitors) and the rest one group (i.e., 10% of the monitors) are employed for model training and validation. Thus, the training and validation samples are composed of data samples collected at different locations in space. This method enables us to evaluate the prediction accuracy of the model at locations where ground-based O_3 measurements are unavailable.

In addition, several main statistical metrics are used, including the ordinary least squares (OLS, Zdaniuk, 2014) regression (e.g., slope and intercept), coefficient of determination (R^2), root-mean-square error (RMSE), mean absolute error (MAE), and mean relative error (MRE). Deseasonalized O_3 monthly anomalies are adopted to calculate the temporal trends (Wei et al., 2019b) and used to analyze the long-term spatiotemporal variations in O_3 pollution across China. Figure 2 illustrates the flowchart of the mapping process of the ChinaHighO_3 dataset in our study.

[Please insert Figure 2 here]

3. Results and discussion

In this study, using ground-based observations, satellite remote sensing data, atmospheric reanalysis products, and model simulations, we have generated a full-coverage and high-quality near-surface O_3 dataset in China (i.e., ChinaHighO_3) at a 10 km resolution using the STET model. This dataset was released on 30 December 2020 (DOI: 10.5281/zenodo.4400043) and is constantly updated. The ChinaHighO_3 dataset includes daily MDA8 O_3 maps from 1 January 2013 to 31 December 2020. It overcomes the problem of missing data in optical remote sensing products caused by cloud contamination and can provide full-coverage ground-level O_3 distributions over China (i.e., 16–54°N, 74–135°E). Monthly, seasonal, and annual MDA8 O_3 maps from 2013 to 2020 are also available (Table S1).

3.1 Accuracy assessment

3.1.1 Overall model performance

First, we validate 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 over China. The MDA8 O_3 estimates are highly consistent (CV-R^2 = 0.87) with the surface measurements, the slope, and y-intercept equal to 0.87 and 11.8 μg/m^3 (Figure 3). The mean RMSE, MAE, and MRE values are 17.10 μg/m^3, 11.29 μg/m^3, and 18.38%, respectively, over the entire domain. Note that the overall accuracy of O_3 estimates has been significantly improved compared to results derived from the original ERT model (i.e., CV-R^2 = 0.78, slope = 0.81, RMSE = 22.39 μg/m^3, and MAE = 14.88 μg/m^3) (Geurts et al., 2006). This confirms the necessity for spatiotemporal information on O_3 pollution.

[Please insert Figure 3 here]

The model performance for each separate year (Table 2) was also evaluated. The overall accuracy of the MDA8 O_3 estimates in the years since 2017 (i.e., out-of-sample CV-R^2 = 0.89–0.93, RMSE = 11.9–15.6 μg/m^3, MAE = 7.9–10.8 μg/m^3, and MRE = 10.3–15.0%) is generally better than that of the previous years (i.e., out-of-sample CV-R^2 = 0.79–0.82, RMSE = 19.1–22.4 μg/m^3, MAE = 12.9–14.9 μg/m^3, and MRE = 21.4–31.8%). The main reasons being the continuous increase in density of the monitoring stations resulted in a sharp increase in the number of data samples (Wei et al., 2021a) and the instrument improvements and quality control upgrades. As shown, our model works well over the study period and for individual years over the study domain.

[Please insert Table 2 here]

Further tested was the model performance in typical regions in China (Figure 3b-f). The model works best over the Beijing-Tianjin-Hebei (BTH) region and the North China Plain (NCP) with out-of-sample CV-R^2 values of 0.91 and 0.89, respectively, and slopes from linear regression closest 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), the Pearl River Delta (PRD), and the Sichuan Basin (SCB). Overall, the model uncertainty is generally small and stable with small differences (e.g., RMSE = 18.9–21.3 μg/m^3, MAE = 12.1–13.6 μg/m^3, and MRE = 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.

On the individual-station scale (Figure 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 during the study period, most stations have sufficient data samples (Figure S2a), 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^2 values exceed 0.8 at ~83% of the stations, especially those located in central and eastern China (CV-R^2 > 0.9). In terms of model uncertainty, except for a few individual stations, ~83% of the stations have RMSE values < 21 μg/m^3, ~88% have MAE values < 15 μg/m^3, and ~85% have MRE values < 25%. Overall, our model performs well at the station scale, with average CV-R^2, RMSE, MAE, and MRE values of 0.86, 16.48 μg/m^3, 11.23 μg/m^3, and 18.36%, respectively.

[Please insert Figure 4 here]

3.1.2 Spatial prediction capability

Next, we focus on evaluating the spatial ability of our model to predict surface O_3 using the out-of-station approach at varying spatial scales. On the entire domain scale, our surface O_3 predictions are well correlated to the observations (e.g., CV-R^2 = 0.80, slope = 0.84) with a mean RMSE, MAE, and MRE values of 21.10 μg/m^3, 13.87 μg/m^3, and 23.18% (Figure 5a), which are a somewhat lower than the out-of-sample validation results (Figure 3a), indicating a strong spatial prediction ability. In addition, the spatial prediction ability of the model gradually increases over the years (Table 2).

On the regional scale, the prediction ability of the model varies differently (Figure 5b-f), where better surface O_3 predictions were observed in the BTH (e.g., CV-R^2 = 0.87, RMSE = 21.51 μg/m^3) and NCP (e.g., CV-R^2 = 0.84, RMSE = 21.99 μg/m^3) regions, while the opposite relatively less accurate predictions were observed in the YRD, SCB, and PRD regions (e.g., CV-R^2 < 0.80, RMSE > 22.6 μg/m^3). In comparison with the out-of-sample results (Figure 3b-f), the accuracy has not changed too much; nevertheless, the evaluation metrics of the former two regions declined slightly less than the other three regions. This is mainly caused by the differences in the density of monitoring stations among the regions.

[Please insert Figure 5 here]

Furthermore, the spatial prediction ability of the model shows spatial differences on the individual-station scale (Figure 6). The prediction accuracy of surface O_3 concentrations are poor with large estimated uncertainties (e.g., CV-R^2 < 0.5, RMSE > 24 μg/m^3, MAE > 18 μg/m^3, and MRE > 25%) in most stations located in western China. By contrast, the model has a strong prediction ability in most stations in eastern China with high CV-R^2 values > 0.8, and small RMSE MAE, and MRE values < 18, 12 μg/m^3, and 15%, respectively. Because the number of monitors is smaller in western China; moreover, the natural and human conditions are largely different from eastern China. For locations that have never had air pollution monitoring, such as remote desert and plateau areas, the uncertainty in the model predictions can be larger. It can only be truly quantified when new observations become available. In general, ~80%, 78%, 86% and 70% of the stations have CV-R^2, RMSE, MAE, and MRE values > 0.7, < 24 μg/m^3, 18 μg/m^3, and 25%, showing an average value of 0.79, 20.08 μg/m^3, 13.82 μg/m^3, and 23.17%, respectively.

[Please insert Figure 6 here]

3.1.3 Temporal-scale validation

Subsequently, we plot the time series of model performance in daily O_3 estimates during 2013-2020 (Figure 7). The daily sample size is large, ranging from 8003 to 10,060, with an average of 9766 and remains unchanged for a long period in particular (Figure S2b). This is due to the unique advantage of full coverage of our ChinaHighO_3 dataset across China. While the performance varies somewhat with the season, the magnitudes of the changes are rather moderate throughout the year, with CV-R^2s ranging from 0.69 to 0.89 (average = 0.81), exceeding 0.75 at about 88% of the days in a year. The absolute uncertainties (i.e., RMSE and MAE) of the ozone estimation have apparent seasonal variations, i.e., low in spring and winter, but high in summer (Fig. 7c-d); by contrast, the relative uncertainty (MRE) shows an opposite seasonal variation (Fig.7d). Surface MDA8 O_3 concentrations are relatively high in summer in most mid-latitude regions of China (Gong et al., 2018). The reason for the larger errors in summer is the greater diurnal cycles and variations in summer ozone, where the averaged variables used in the model may not reflect the conditions associated with high O_3 content in the afternoon, while the observations are likely driven by the afternoon peaks. In general, average RMSE, MAE, and MRE values are 18.82 μg/m^3, 11.27 μg/m^3, and 18.42%, and < 20, 15 μg/m^3, 20% on ~86%, and 99%, and 75% of the days, respectively.

[Please insert Figure 7 here]

The monthly mean MDA8 O_3 estimates for each year are also evaluated (Figure 8). High accuracy is seen, with strong slopes from linear regression of 0.83~0.97, high R^2 values of 0.86~0.97, and small uncertainties with RMSE and MAE (MRE) values ranging from 5.5 and 4.0 μg/m^3 (4.4%) to 13.4 and 9.0 μg/m^3 (13.8%) among different years. In general, the data quality of the monthly O_3 estimates (N = 119,194) is reliable (e.g., R^2 = 0.93, RMSE = 9.42 μg/m^3, MAE = 6.91 μg/m^3, and

MRE = 8.56%) during the entire study period of 2013–2020. This allows us to accurately analyze the spatial and temporal distributions of and variations in O_3 pollution in China.

[Please insert Figure 8 here]

3.2 Spatiotemporal surface O_3 variations

3.2.1 Spatial coverage and distribution

Figure 9 presents two typical examples of MDA8 O_3 maps on 18 June and 11 November 2019 and the annual map for 2019 in China. This dataset can uniquely capture MDA8 O_3 concentrations anywhere in the country (i.e., spatial coverage = 100%) on any given day. In general, the O_3 concentration is particularly high (> 150 μg/m^3) in northern China and much lower (< 80 μg/m^3) in southern China on 18 June 2019 (average = 118.7 ± 36.1 μg/m^3). High emissions of three main O_3 precursors (i.e., NOx, VOCs, and CO) are mainly observed in eastern China, especially the NCP (Figure S3). In general, about 2% differences in surface O_3 concentrations between northern and southern China are derived by the emissions on this day. A completely different situation was observed on 11 November 2019 (average = 77.9 ± 24.6 μg/m^3). On an annual scale, differences in O_3 distribution between northern and southern China in 2019 decreased, with an average level of 98.3 ± 11.3 μg/m^3. The great differences in surface O_3 concentrations between Northern and Southern China on different days are mainly dominated by differences in sunlight and ozone chemical formation during different seasons. A comparison with ground-based observations shows highly consistent spatial patterns on both daily and annual scales across China. As such, these results illustrate that spatially continuous O_3 data, which is important for those places without monitoring stations, can be provided.

[Please insert Figure 9 here]

Figure 10 shows mean MDA8 O_3 maps for different seasons from 2013 to 2020 across China. As evident, O_3 concentrations change significantly on a seasonal scale; they are extremely high in summer (average = 103.6 ± 18.0 μg/m^3), especially in the NCP (average = 138.8 ± 13.5 μg/m^3). Followed by spring (average = 99.4 ± 9.2 μg/m^3). By contrast, winter yields much lower O_3

concentrations in China (average = 69.9 ± 7.7 μg/m^3), especially the BTH (average = 55.4 ± 7.6 μg/m^3). The spatial pattern of O_3 in autumn (average = 80.9 ± 10.1 μg/m^3) is similar but generally higher than that in winter across China, especially in southeastern areas (Table S2). Figure S4 shows zoomed-in summer mean O_3 maps for four key regions in China. The ChinaHighO_3 dataset can reflect and describe well the distribution and variation in ozone pollution at the local, even urban scales due to its high spatial resolution of 10 km. All four typical regions experience different degrees of O_3 pollution in summer, especially the BTH (average = 142.9 ± 14.5 μg/m^3) and YRD (average = 113.9 ± 13.7 μg/m^3) regions.

[Please insert Figure 10 here]

3.2.2 A short-term severe O_3 pollution event

We closely examined a severe surface O_3 pollution episode that occurred from 23 April to 8 May in 2020 in eastern China (Figure 11). Before 25 April, O_3 was at a low level across the whole country, then gradually increased. On 28 April, the O_3 levels at BTH and all surrounding "2+26" cities (Figure S1) had exceeded the ambient air quality standard, i.e., MDA8 O_3 = 160 μg/m^3 (Figure S5). More severe O_3 pollution occurred in most other areas on 29 April, with a maximum value of 124.0 ± 30.2 μg/m^3 and 181.0 ± 17.8 μg/m^3 in China and YRD (Figure S6). On 30 April, BTH experienced the maximum level of O_3 pollution (average = 232.1 ± 47.2 μg/m^3), and remained high till 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_3 concentrations were generally low in SCB before 25 April, and gradually formed into regional pollution on 26 April, when the Chengdu Plain and southern and northeast Sichuan were polluted to varying degrees. By 28 April, most cities exceeded the ambient air quality standard; Pingyuan and southern Sichuan were heavily polluted, and the O_3 concentrations remained high and reached the maximum on 3 May with an average value of 184.3 ± 29.8 μg/m^3 in SCB (Figure S6). On 6 May, the polluted air moved southward, gradually decreasing in the pollution intensity. After 7 May, accompanied by cooling and precipitation, this episode of ozone pollution ended, and the air quality changed 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.

[Please insert Figure 11 here]

3.2.3 Changes during the COVID-19 pandemic

Coronavirus (COVID-19) broke out in Wuhan, Hubei Province at the end of 2019, and quickly spread to the whole country due mainly to the Spring Festival (WHO, 2020; Zu et al., 2020). To prevent the further spread of COVID-19, the entire Hubei Province was on lockdown starting at 10am on 23 January 2020, soon followed by almost all other major cities in China, which lasted about three weeks (Su et al., 2020; Tian et al., 2020). To gain further insight of the ozone changes associated with the COVID-19, O_3 changes in China are examined before (Period 1: 1–25 January), during (Period 2: 26 January to 17 February), and after (Period 3: 18 February to 31 March) the COVID-19 outbreak. Considering the increase in O_3 in recent years, only compared are the relative difference in O_3 concentrations across eastern China between 2020 and 2019 during the three periods (Figure 12).

Before the COVID-19 outbreak, O_3 concentrations remained near the historical values with relative changes within ± 10%. During the lockdown, significant increases in O_3 concentrations were seen in most parts of eastern China, especially in Hubei Province and its surrounding areas, showing a relative change > 40%. In contrast, an opposite decline in O_3 concentrations was observed in the PRD, mainly caused by the sharp decline in NO_x emissions after the lockdown (Ding et al., 2020; Feng et al., 2020). Because O_3 formation rates over northern China are under the NO_x-saturated regime, a reduction in NO_x would enhance the O_3 generation rates (Liu and Wang, 2020; Shi and Brasseur, 2020; Benish et al., 2021). Ozone formation rates over the PRD are under the NO_x-limited regime, so the same reduction in NO_x would diminish the O_3 generation rates (Liu and Wang, 2020; Wang et al., 2021; Li et al., 2021). After the COVID-19 outbreak, O_3 concentrations changed little (within ± 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 is a contrasting increase in O_3

concentrations, likely related to increases in NO_x and temperature (Wang et al., 2021). Although sensitivities change, the total amount of ozone produced, and the size of plume scale with NO_x emissions, but the rate of ozone production is nonlinear, and air quality can worsen with initial emissions controls (Lin et al., 1988).

[Please insert Figure 12 here]

3.2.4 Long-term variations in the recent decade

Figure 13 shows the MDA8 O_3 trends (μg/m^3/yr) during the study period (2013–2020) calculated from monthly anomalies across China. Surface O_3 concentrations show diverse variations from the national to regional scales during the recent eight years. In general, most areas of the country show significant increasing O_3 pollution, with an average of 2.49 μg/m^3/yr ($p < 0.001$), especially in central China (> 5 μg/m^3/yr, $p < 0.05$) and NCP (~4.42 μg/m^3/yr, $p < 0.001$). The BTH and YRD regions had the stronger increasing trends by 3.84 and 3.43 μg/m^3/yr ($p < 0.001$), respectively. In addition, other two typical regions, i.e., SCB (~1.78 μg/m^3/yr, $p < 0.001$), and PRD (~1.41 μg/m^3/yr, $p < 0.05$) showed relatively low but obvious increasing trends. The increase in O_3 over city clusters are closely associated with a decrease in NO_x emissions and $PM_{2.5}$ concentrations (Li et al., 2019; Zhang et al., 2019; Wang et al., 2020; Wei et al., 2021a) and meteorological variations (Li et al., 2020). By contrast, seen are opposite weakening trends in several coastal provinces in southern China (e.g., Guangxi and Zhejiang).

[Please insert Figure 13 here]

Next investigated are the variations in surface O_3 pollution under the background of different implemented environmental policies (Table 3). During the Clear Air Action Plan (2013–2017), China showed a significant increasing trend of 1.33 μg/m^3/yr ($p < 0.05$), especially in the NCP (~4.58 μg/m^3/yr, $p < 0.001$) and BTH (~4.38 μg/m^3/yr, $p < 0.001$) regions. In addition, increasing trends were also found in the YRD and SCB regions. By contrast, O_3 pollution overall declined in the PRD region. During the Blue-Sky Defense Plan (2018–2020), O_3 concentrations continued to increase by 7.2% and 2.5–5.4% in China and typical urban agglomerations in 2020 than those in 2017, respectively. In particular, considering the entire study period, O_3 pollution increased the most in China (~4.40 μg/m^3/yr, p < 0.001) and most typical regions during the period 2015–2019, especially NCP (~6.33 μg/m^3/yr, p < 0.001) and YRD (~5.60 μg/m^3/yr, p < 0.001).

[Please insert Table 3 here]

Taking into account the seasonal differences in O_3 discussed above, we focus on the spatiotemporal variations in summer mean MDA8 O_3 from 2013 to 2020 over eastern China (Figure 14). Ozone levels always remained at a high level in summer among different years in China, with an average value > 90 μg/m^3. It was higher during the period 2017–2019 than in previous years, especially in the NCP (> 120 μg/m^3). This is closely associated with the rising temperatures and increased number of hot days in the NCP (Li et al., 2020). Changes in O_3 have been diverse in recent eight years, i.e., O_3 concentrations were higher in 2014 than in 2013 in most areas of China, yet generally decreased in 2015, especially in southern China. O_3 pollution has increased significantly since 2016, reaching a maximum in 2019 (~117.4 ± 23.6 μg/m^3), especially in the NCP (~159.7 ± 14.1 μg/m^3). This may be due to the decreasing $PM_{2.5}$ concentrations by ~15% in the NCP (Li et al., 2020; Wei et al., 2021a), yet, the dominant reason remains controversial. By contrast, overall O_3 pollution decreased in China and in most typical regions in China in 2020 (Table S1). The coordinated control measures of fine particulate matter and O_3 implemented by the Chinese government (Xiang et al., 2020) may explain this, as well as the ongoing effects of the 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., 2020; Lu et al., 2020; Wang et al., 2020). Our predicted results also show similar patterns in spatial distribution compared to those derived from satellite OMI/Aura observations (Liu et al., 2020; Zhang et al., 2020) and air quality model simulations (Hu et al., 2016; Xue et al., 2020) in previous studies.

We have also calculated the percentage of O_3 polluted days (i.e., MDA8 O3 > 160 μg/m^3) for each grid in eastern China for each year from 2013 to 2020 (Figure 15). In 2013 and 2014, O_3 pollution was mainly found along the east and south provinces of China, but the probability of occurrence was generally low in most areas, with a percentage < 10%. The area where O_3 pollution occurred overall decreased from 2014 to 2015, especially in southern China. From then on, the area covered by O_3 pollution continuously expanded until 2020, covering most areas of eastern China. More importantly, the probability of occurrence of O_3 pollution increased significantly from 2017 to 2019, especially in the NCP; for example, 23% of the days in 2019 exceed the accepted O_3 standard. At the regional scale, the proportion of days exceeding the daily O_3 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 (Figure S7). By contrast, the probability of occurrence of O_3 pollution overall declined in most areas of Northern China (e.g., NCP, BTH, and YRD) in 2020. Similar conclusions have been reported by previous studies (Liu et al., 2020; Xue et al., 2020; Zhan et al., 2018).

Figure 16 shows the evolution of MDA8 O_3 concentrations for each year at the "2+26" cities in Northern China, where pollution is of particular concern to the public (Figure S1). Until 2015, O_3 concentrations were generally lower than 120 $\mu g/m^3$ in most cities, with much fewer days exceeding the air quality standard (i.e., MDA8 O_3 = 160 $\mu g/m^3$) than those after 2016. With time, the number of days with high O_3 concentrations gradually increased from year to year. In particular, a significant increase in O_3 concentrations can also be captured, i.e., from May to August in each year from 2017 to 2019, the MDA8 O_3 concentrations in almost all cities frequently exceeded 200 $\mu g/m^3$, indicating a severe risk of ozone exposure.

Surface monitoring stations are distributed unevenly across China and vary greatly in density from region to region. Most are located in urban areas, making it difficult to accurately predict air pollution on a wider scale. Our study helps make up for this deficiency by generating spatially continuous and full-coverage daily surface O_3 maps, allowing users to obtain more accurate estimates of distributions and variations of O_3 pollution, especially for those areas with no or minimal ground-based measurements. These maps can also help evaluate numerical models as well as pollution control measures and estimates of pollution exposure.

3.3 Discussion

3.3.1 Uncertainty and error analysis

We have first investigated the effects of varying training samples on the model results in surface O_3 estimates. For this purpose, we gradually increase the proportion of training samples from 50% to 90% for model building, and the rest of the samples are used for validation by applying different N-fold (i.e., 2, …, 10) CV methods using 2020 data over China (Table S4). In general, with the increase of training samples, the overall accuracy and spatial prediction ability of the STET model are gradually improved with increasing CV-R^2 values and declined estimation uncertainties. Small changes in each evaluation indicator have been found, even when the training sample has changed by as much as 40%, indicating that our model is stable and robust (e.g., CV-R^2 > 0.90 and RMSE < 14.1 μg/m^3). 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_3 conditions and variations across mainland China; in addition, it benefits from the robustness of ensemble learning, which has a strong anti-noise ability (Breiman, 2001; Geurts et al., 2006).

We have trained and built the models separately for each characteristic region and compared the prominent features (Figure S8) and model performance (Table S5) with the national model. The top-scoring features for regional models are similar to those for the national model, e.g., ERA5 DSR, TEM, ET, RH, and OMI NO_2 and O_3 (Figure 1). However, there are numerical differences in the importance scores for each variable. The model shows different accuracy and spatial prediction ability at the regional scale, with causes closely related to the density and spatial distribution of ground-level monitoring stations. The geographic, meteorological, and population conditions are different in each region. The performance of the national model is generally better with smaller estimation uncertainties than anyone regional model, but the differences in the statistics metrics are small. The whole model involves a much bigger number of data samples that can cover more O_3 conditions; it can also consider the impact of adjacent regions, especially the transition areas. Full coverage mapping provides the richest data set to train a robust model.

3.3.2 Comparison with chemical reanalysis products

We have compared our ChinaHighO_3 dataset with long-term atmospheric reanalysis products generated from chemical models, including MERRA2 and ERA5, which have spatiotemporal coverages. For this purpose, 3-hour MERRA2 and 1-hour ERA5 Ozone Mixing Ratio (OMR, unit:

kg kg^{-1}) simulations at a horizontal resolution of 0.25°×0.25° are collected to calculate the 2:00 p.m. and MDA8 O_3 concentrations at the ground level (μg/m^3) in 2020 in China and validated with the corresponding ground-based measurements, respectively (Figure S9). The ground-level O_3 simulations from the chemical reanalysis products are very poor, showing great uncertainties (e.g., $R^2 < 0.1$, RMSE > 47 μg/m^3). The main reason is that the chemical reactions in the assimilation models are substantially simplified and mainly reflect the impact of dynamic processes on stratospheric and tropospheric ozone (Knowland, et al., 2017). By contrast, our surface O_3 estimates are highly consistent with the ground-based measurements (e.g., $R^2 = 0.96$, RMSE = 8.6 μg/m^3), which seem to be better than the chemical reanalysis products.

3.3.3 Comparison with related studies

We have compared our study with previous related studies, which used the same out-of-sample 10-CV approach with the MEE network O_3 observations, for the same study period focusing on China (Table 4). Our algorithm yields a higher accuracy with smaller estimation uncertainties (CV-R^2 > 0.83, RMSE < 15 μg/m^3) than the RF (CV-R^2 = 0.69, RMSE = 26.0 μg/m^3; Zhan et al., 2018), XGBoost (CV-R^2 = 0.78, RMSE = 21.47 μg/m^3; Liu et al., 2020), data fusion (CV-R^2 = 0.70, RMSE = 26.20 μg/m^3; Xue et al., 2020), GWR (CV-R^2 = 0.77, MAE = 8.14 μg/m^3; Zhang et al., 2020), and LUR/BME (CV-R^2 = 0.80, RMSE = 23.5 μg/m^3; Chen et al., 2020) models at different temporal scales for the same study period.

In addition, different studies relied on different main predictors, i.e., key variables input to the model in estimating surface O_3 concentrations. These O_3 datasets in previous studies are derived from main predictors, including the satellite total-column O_3/NO_2, or CH_2O, MERRA-2 reanalysis, model simulations, or in situ observations, showing a large number of missing values at coarse or false (e.g., forced resampling) spatial resolutions (i.e., 0.25°–0.625°) limited by the input data sources. By contrast, our study overcomes these issues and is a large improvement on previous studies, which provides a daily full-coverage (spatial coverage = 100%) and true-spatial-resolution (~0.1° × 0.1°) O_3 dataset generated from two main predictors (i.e., DSR and TEM) provided by the ERA5 reanalysis. In addition, the dataset provided here constitutes the nearly continuous record of ground-level O_3 concentrations from 2013 to 2020 in China.

[Please insert Table 4 here]

4. Summary and conclusions

Ground-level O_3 is a major pollutant affecting our health. To compensate for the sparse and inhomogeneous coverage of ground-based ozone networks and the low data quality, missing values and low resolution of many existing satellite-based ozone estimation, we applied a spatiotemporal extremely randomized trees (STET) machine-learning model to develop a long-term near-surface ozone product that can overcome or lessen the above limitations. Besides ozone training data, the input variables include surface downward solar radiation, air temperature, meteorological variables, land use and topography, population distribution, and pollution emission inventory. The daily maximum 8-hour average (MDA8) O_3 product (ChinaHighO_3) with full coverage across China at a spatial resolution of 10 km from 2013 to 2020 are generated.

The estimates are evaluated against surface observations at varying spatiotemporal scales and compared with previous related studies. The cross-validation (CV) results illustrate that our model yields a high overall accuracy (spatial prediction ability) with an average out-of-sample (out-of-station) CV-R^2, RMSE, MAE, and MRE values of 0.87 (0.80), 17.10 (21.10) $\mu g/m^3$, 11.29 (13.87) $\mu g/m^3$, and 18.38 (23.18) %, respectively. Nevertheless, in the current stage, we can only evaluate the surface O_3 predictions by removing parts of the base data set using different 10-CV approaches, but the accuracy of predictions where there have never been O_3 measurements still remains a challenge. In particular, the ChinaHighO_3 product is superior to existing ones in terms of model accuracy, spatial coverage and resolution, and data record length.

The spatial distributions and temporal variations of ground-level O_3 concentrations are investigated during the recent decade. A long-term analysis showed that O_3 concentrations have significantly increased by 2.49 $\mu g/m^3/yr$ ($p < 0.001$) in China from 2013 to 2020, especially in the North China Plain (~4.42 $\mu g/m^3/yr$, $p < 0.001$). In addition, summer ozone changed diversely, which was much higher since 2017 than in previous years due to the rising temperatures and increased hot days. The number of days exceeding the ambient O_3 air quality standard (MDA8 O_3 = 160 $\mu g/m^3$) and the areal extent of high O_3 levels were also shown to be gradually increasing across China, especially in the "2+26" cities in the Northern China Plain. Benefiting from the unique advantages of the ChinaHighO_3 dataset, a recent short-term national and regional severe O_3 pollution event with its formation and dissipation from the end of April to the beginning of May 2020 was well captured. Also observed was a rapid increase in O_3 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 is not a repudiation of NO_x controls. Therefore, our ChinaHighO_3 dataset will be of great significance for the related studies on air pollution in China, especially for those focusing on environmental health.

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 is made possible via the in-kind support from the University of Iowa.

References

Adam-Poupart, 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.

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., and 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., and 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., Friedman, J., Olshen, R., and Stone, C., 1984. Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software. ISBN 978-0-412-04841-8.

Breiman, L., 2001. Random forests, Mach. Learn., 45, 5-32.

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, L., and Coauthors, 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

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.

Copernicus Climate Change Service (C3S), 2017. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), date of access. https://cds.climate.copernicus.eu/cdsapp#!/home.

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., and Coauthors, 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

Ding, J., van der A, R. J., Eskes, H. J., Mijling, B., Stavrakou, T., van Geffen, J. H. G. M., Veefkind,

J. P., 2020. NO_x emissions reduction and rebound in China due to the COVID-19 crisis. Geophys. Res. Lett., 46, e2020GL089912. https://doi.org/10.1029/2020GL089912

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.

Feng, S., Jiang, F., Wang, H., Wang, H., Ju, W., Shen, Y., Zheng, Y., Wu, Z., Ding, A., 2020. NO_x emission changes over China during the COVID-19 epidemic inferred from surface NO_2 observations. Geophys. Res. Lett., 47, e2020GL090080. https://doi.org/10.1029/2020GL090080

Gaudel, A., Cooper, O. R., Gérard Ancellet, 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 Lean. 63 (1), 3–42.

Goldberg, D. L., Loughner, C. P., Tzortziou, M., Stehr, J. W., Pickering, K. E., Marufu, L. T., and 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.

Guo, J., Zhang, X., Che, H., Gong, S., An, X., Cao, C., et al., 2009. Correlation between PM concentrations and aerosol optical depth in eastern China. Atmos. Environ. 43 (37), 5876–5886.

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.

Hu, J., Chen, J., Ying, Q., and 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 $PM_{2.5}$, SO_2, NO_2 and O_3 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 Bioinformatics 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., and 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.

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 B: Chem. Phys. Meteorol., 66, 23455.

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., 2020. Increases in surface ozone pollution in China from 2013 to 2019: anthropogenic and meteorological influences. Atmos. Chem. Phys. 20, 11,423–11,433.

Li, K., Jacob, D. J., Liao, H., Qiu, Y., Shen, L., Zhai, S., Bates, K., Sulprizio, M., Song, S., Lu, X., Zhang, Q., Zheng, B., Zhang, Y., Zhang, J., Lee, H. C., Kuk, S. K., 2021. Ozone pollution in the North China Plain spreading into the late-winter haze season. Proc. Natl. Acad. Sci. U. S. A., 118(10), e2015797118.

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., 2017. 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, 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_3) across the remote Tibetan Plateau during 2005–2018. Atmos. Chem. Phys. 20, 6159–6175. https://doi.org/10.5194/acp-20-6159-2020

Li, T., Shen, H., Yuan, Q., Zhang, X., and Zhang, L., 2017. Estimating ground-level $PM_{2.5}$ by fusing satellite and station observations: a geo-intelligent deep learning approach. Geophys. Res. Lett., 44(23), 11,985–11,993.

Lim, C., and Coauthors, 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, 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. Intl. 142, 105823. https://doi.org/10.1016/j.envint.2020.105823

Liu, X., Bhartia, P., Chance, K., Spurr, R., and 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, Y., Wang, T., 2020. Diverse response of atmospheric ozone to COVID-19 lockdown in China. Atmos. Ocean. Phys., arXiv:2008.10851.

Loughner, C. P., Allen, D. J., Pickering, K. E., Zhang, D. L., Shou, Y. X., and 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., and Coauthors, 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.

Ministry of Ecology and Environment (MEE), 2018. Revision of the Ambien air quality standards (GB 3095-2012) (in Chinese), available at: http://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/201808/t20180815_629602.html.

Meleux, F., Solmon, F., Giorgi, F., 2007. Increase in summer European ozone amounts due to climate change. Atmos. Environ. 41, 7577–7587.

Mills, G., and Coauthors, 2018. Tropospheric ozone assessment report: present-day tropospheric ozone distribution and trends relevant to vegetation. Elem. Sci. Anth. 6(1), 47.

Nickolay, A., Lok, N., Sergey, V., Edward, A., Eric, J., William, H., Joanna, and the OMI core team, 2019. OMI/Aura NO2 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), DOI:10.5067/Aura/OMI/DATA3007

Pawan, K., 2012. OMI/Aura Ozone (O3) Total Column Daily L2 Global Gridded 0.25 degree x 0.25 degree V3, Goddard Earth Sciences Data and Information Services Center (GES DISC), DOI: 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.

Ryan, W. F., 1995. Forecasting severe ozone episodes in the Baltimore metropolitan area. Atmos. Environ. 29(17), 2387-2398.

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. Nature Clim. Change 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., and Coauthors, 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., and 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., and Coauthors, 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., and Coauthors, 2015. Long-term ozone exposure and mortality in a large prospective study. Am. J. Respir. Crit. Care Med. 193(10), 1134–1142.

Wang, M., and Coauthors, 2016. Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and $PM_{2.5}$. Environ. Sci. Technol. 50(10), 5111–5118.

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, 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

Wang, T., Xue, L., Brimblecombe, P., Lam, Y. F., Li, L., and Zhang, L., 2016. Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects. Sci. Total Environ. 575, 1582-1596.

Wang, Y., and Coauthors, 2020. Contrasting trends of $PM_{2.5}$ and surface ozone concentrations in China from 2013 to 2017. Natl. Sci. Rev. 7(8), 1331–1339.

Wei, J., Peng, Y., Mahmood, R., Sun, L., and 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 $PM_{2.5}$ 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 $PM_{2.5}$ 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., and Cribb, M., 2021b. The ChinaHighPM_{10} dataset: generation, validation, and spatiotemporal variations from 2015 to 2019 across China. Environ. Intl. 146, 106290. https://doi.org/10.1016/j.envint.2020.106290

WHO, 2020. The World Health Organization. Coronavirus Disease (COVID-19) Pandemic. Available online: https://www.who.int/ emergencies/diseases/novel-coronavirus-2019

Wu, S., Huang, B., Wang, J., He, L., Wang, Z., Yan, Z., Lao, X., Zhang, F., Liu, R., and Du, Z., 2021. Spatiotemporal mapping and assessment of daily ground NO_2 concentrations in China using high-resolution TROPOMI retrievals. Environ. Pollut. 273, 116456.

Xiang, S., and Coauthors, 2020. Control of both $PM_{2.5}$ and O_3 in Beijing-Tianjin-Hebei and the surrounding areas. Atmos. Environ. 224, 117259.

Xue, T., and Coauthors, 2020. Estimating spatiotemporal variation in ambient ozone exposure during 2013–2017 using a data-fusion model. Environ. Sci. Technol. 54, 14,877–14,888.

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., and Coauthors, 2019. Drivers of improved $PM_{2.5}$ 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, Q., Streets, D., He, K., Klimont, Z., 2007. Major components of China's anthropogenic primary particulate emissions. Environ. Res. Lett. 2, No. 045027.

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.

Tables

Table 1. Summary of the data sources used in this study.

Category	Variable	Description	Unit	Spatial Resolution	Temporal Resolution	Data Source
Ground measurements	O_3	Ozone	μg/m^3	-	Hourly	CNEMC
Atmospheric reanalysis	DSR	Downwelling surface radiation	W/m^2	0.1°×0.1°	Hourly	ERA5 reanalysis
	BLH	Boundary layer height	m		Hourly	
	ET	Evaporation	mm			
	PRE	Precipitation	mm		Hourly	
	RH	Relative humidity	%		Hourly	
	TEM	2-m air temperature	K		Hourly	
	SP	Surface pressure	hPa		Hourly	
	WU	10-m u-component	m/s		Hourly	
	WV	10-m v-component	m/s		Hourly	
Satellite remote sensing products	O_3	total-column O_3	DU	0.25°×0.25°	Daily	OMI/Aura products
	NO_2	tropospheric NO_2	molec/cm^2			
	NDVI	Normalized difference vegetation index	-	0.05°×0.05°	Monthly	MODIS products
	LUC	Land-use cover	-	0.05°×0.05°	Annual	
	DEM	Surface elevation	m	90 m × 90 m	-	SRTM
	POP	Ambient population	-	1 km × 1 km	Annual	LandScan™
Emission inventory	NO_x	Nitric oxide metabolite	Mg/grid	0.25°×0.25°	Monthly	MEIC
	VOCs	Volatile organic compounds				
	CO	Carbon monoxide				

882 **Table 2.** Statistics of cross-validation results of MDA8 O_3 estimates ($\mu g/m^3$) for each year from
883 2013 to 2020 in China.

Year	Sample size	Overall accuracy				Spatial prediction ability			
	N	R^2	RMSE	MAE	MRE	R^2	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

884

Table 3. Statistics of MDA8 O_3 trends (μg/m^3/yr) and relative change (%) in annual mean MDA8 O_3 concentrations (μg/m^3) from 2013 to 2020 in China and 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 %

Table 4. Comparison of model performances with previous O_3 studies focused on China as a whole.

Model	Temporal resolution	Validation			Study period	Main Predictors	Missing values	Literature
		R^2	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_3, 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_2, CH_2O	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**	**No**	This study*
		0.81	**20.27**	**13.38**	**2013-2017**	**TEM**		
		0.83	**18.88**	**12.72**	**2015-2017**			
	Monthly	**0.90**	**12.43**	**8.82**	**2014**			

BME: Bayesian maximum entropy; CTM: chemical transport model; XGBoost: eXtreme Gradient Boosting

Figures

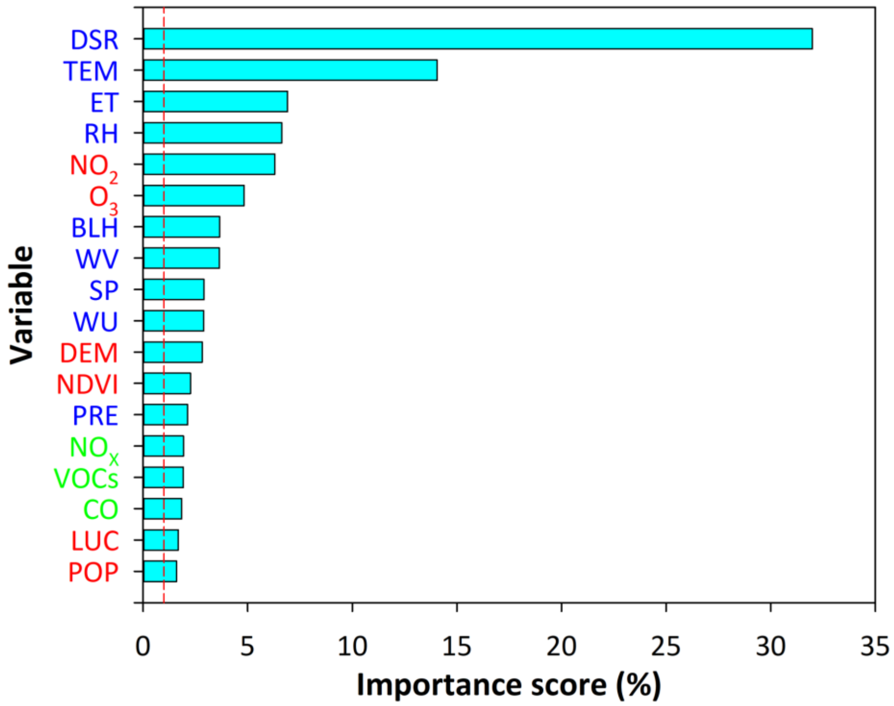
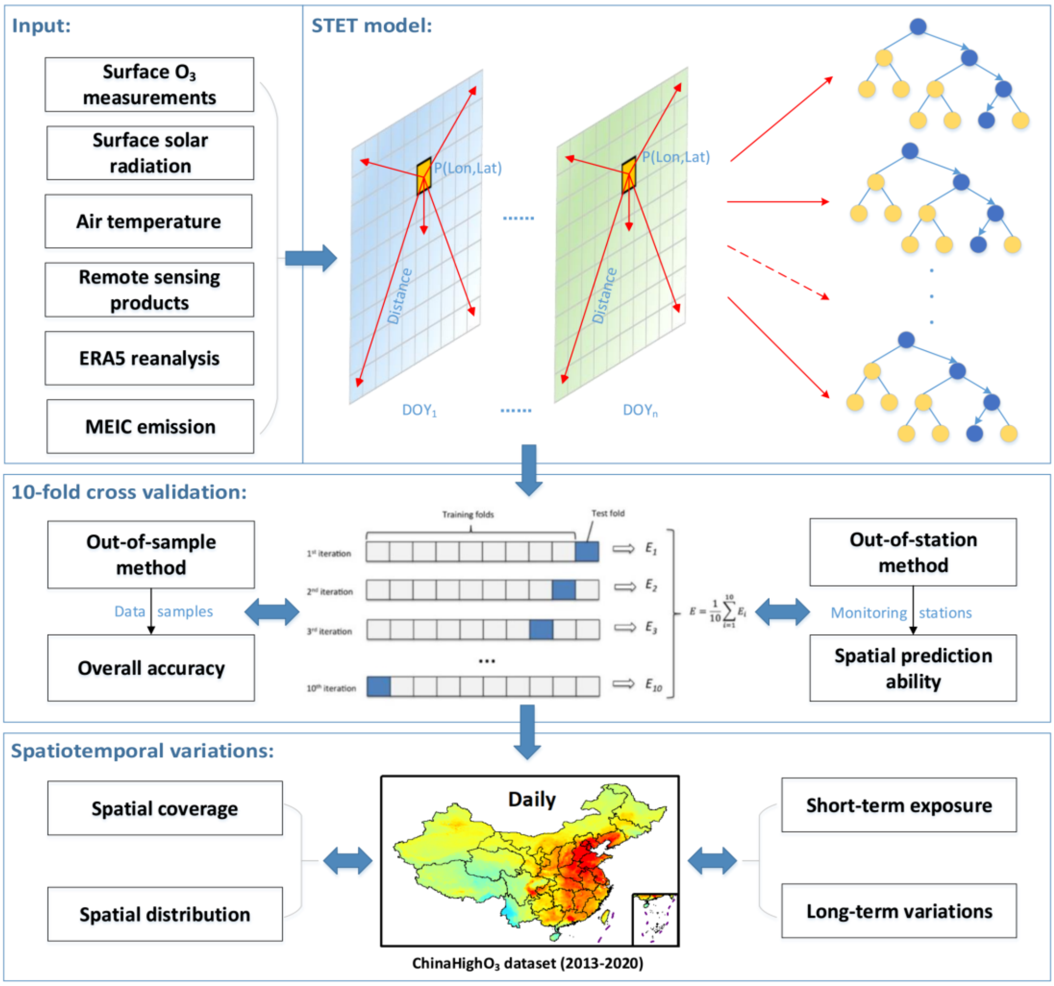


Figure 1. Sorted importance scores of variables used in estimating O_3 concentrations using the STET model, where red, blue, and green colors indicate variables from satellites, ERA5 reanalysis, and MEIC emission inventory, respectively. The vertical red dashed line shows the importance score of 1%.



896

897 Figure 2. Flowchart of the mapping process of the ChinaHighO_3 dataset for our study.

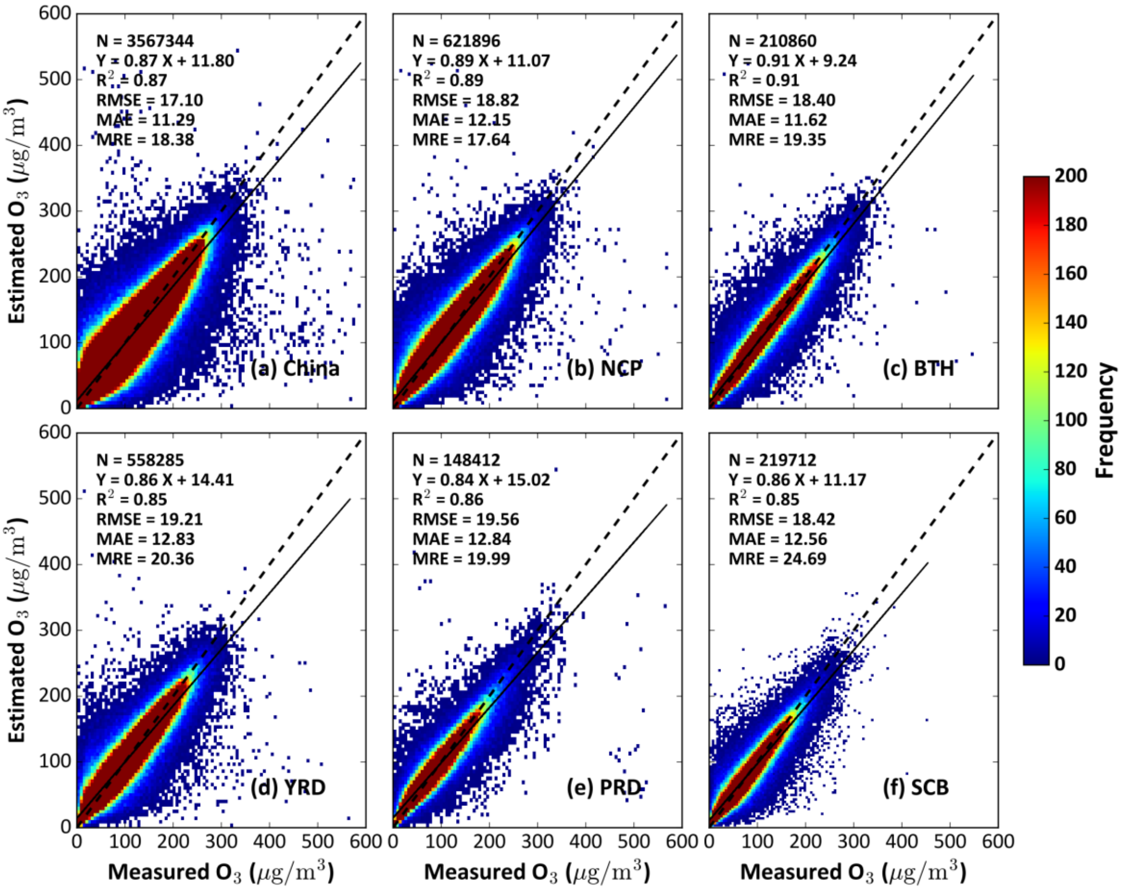


Figure 3. Out-of-sample cross-validation results of MDA8 O_3 estimates (μg/m^3) from 2013 to 2020 (a) in China, (b) North China Plain (NCP), (c) Beijing-Tianjin-Hebei (BTH) region, (d) Yangtze River Delta (YRD), (e) Pearl River Delta (PRD), and (f) Sichuan Basin (SCB). Frequency in the right legend indicates the total number of data samples in each cell.

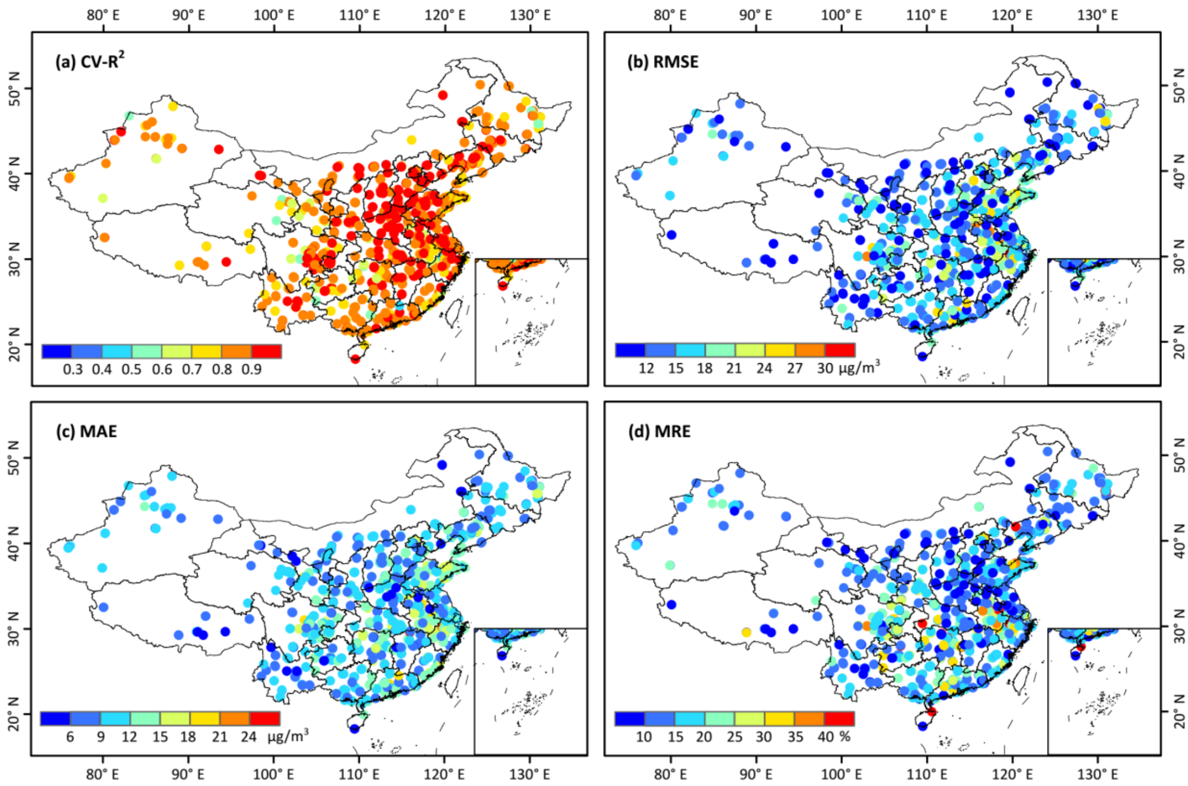


Figure 4. Individual-site-scale out-of-sample cross-validation results of MDA8 O_3 estimates (μg/m^3) from 2013 to 2020 in China.

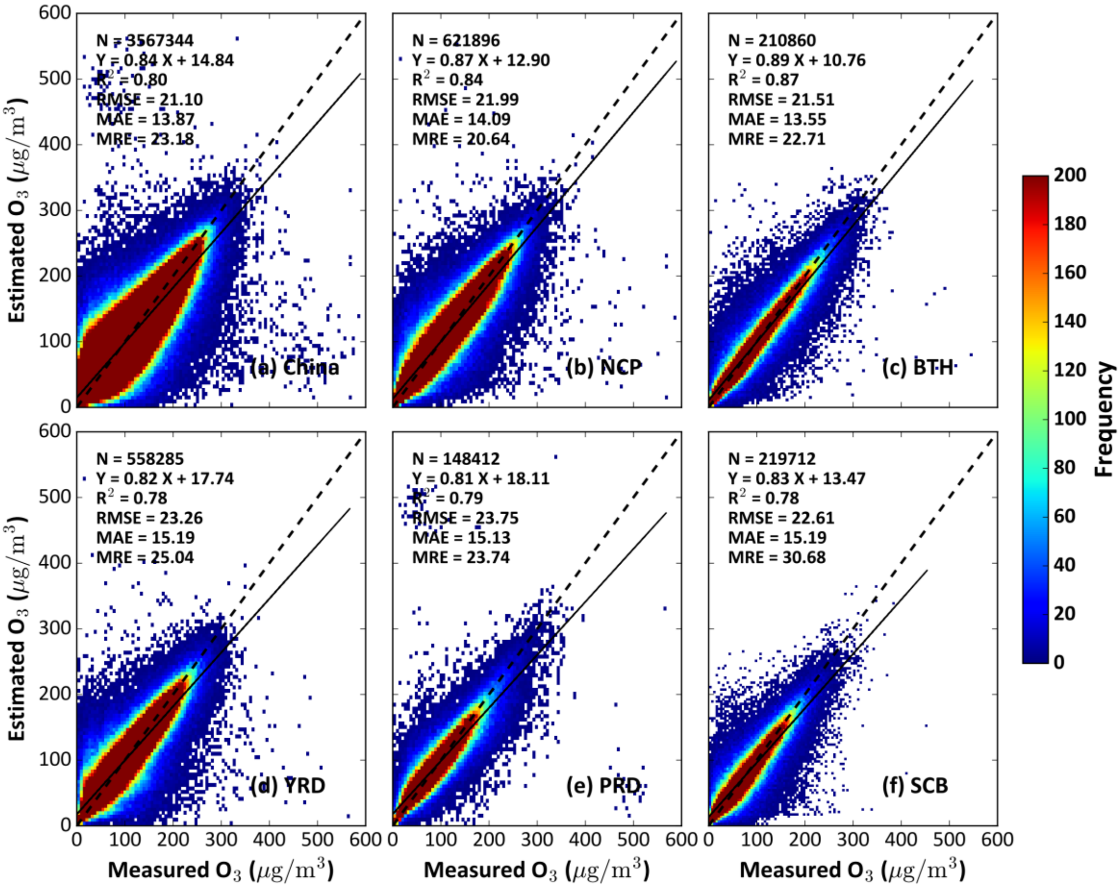


Figure 5. Same as Figure 4 but for out-of-station cross-validation results.

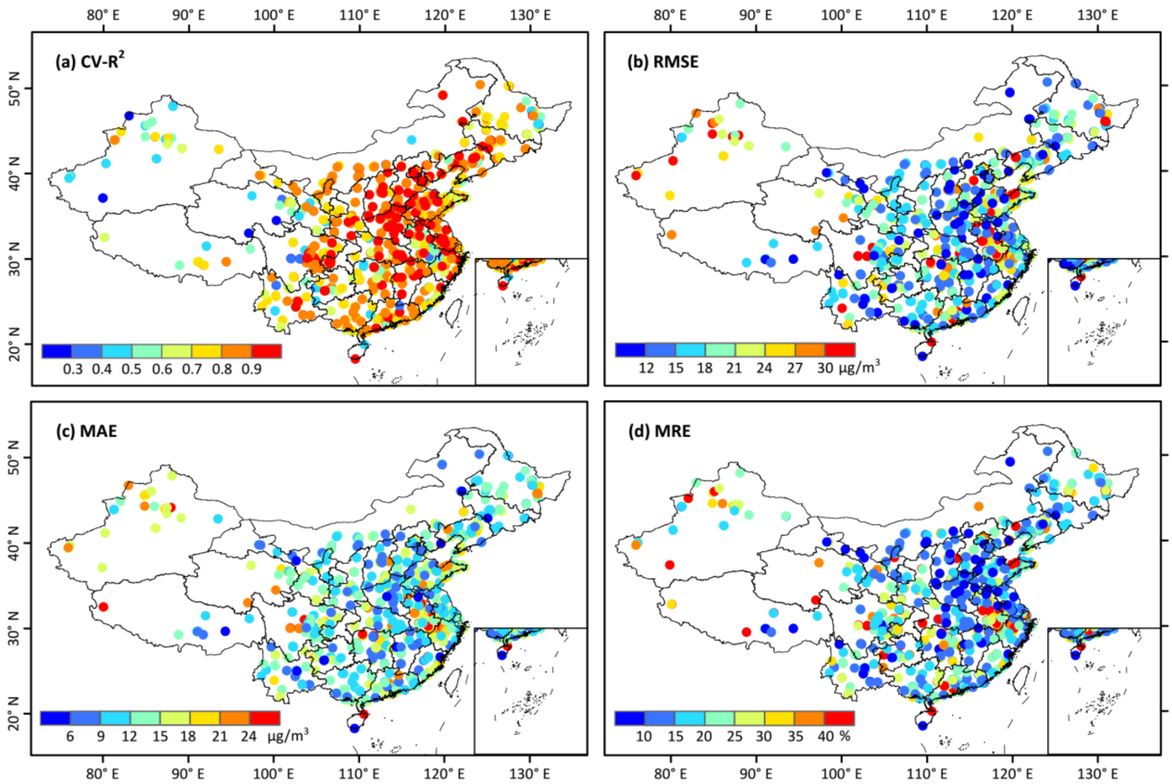


Figure 6. Same as Figure 5 but for out-of-station cross-validation results.

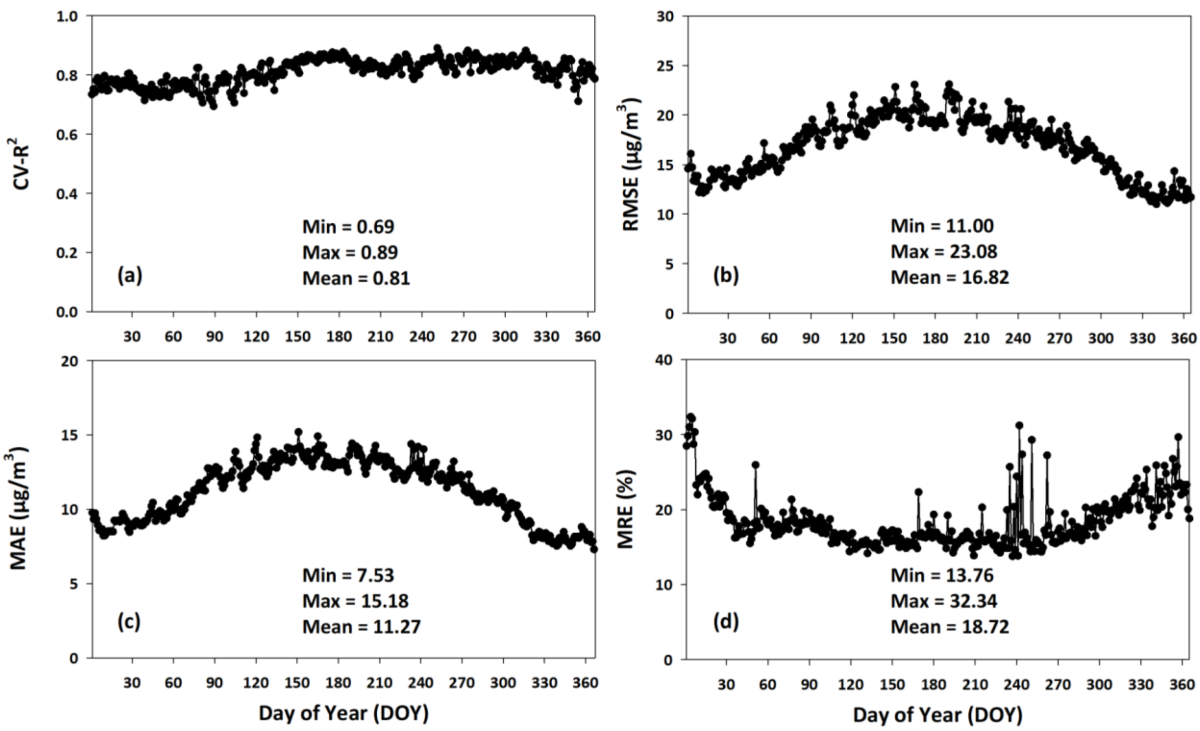


Figure 7. Time series of daily variations of the validation results of MDA8 O_3 estimates (μg/m^3) from 2013 to 2020. Minimum, maximum, and mean values are given in each panel.

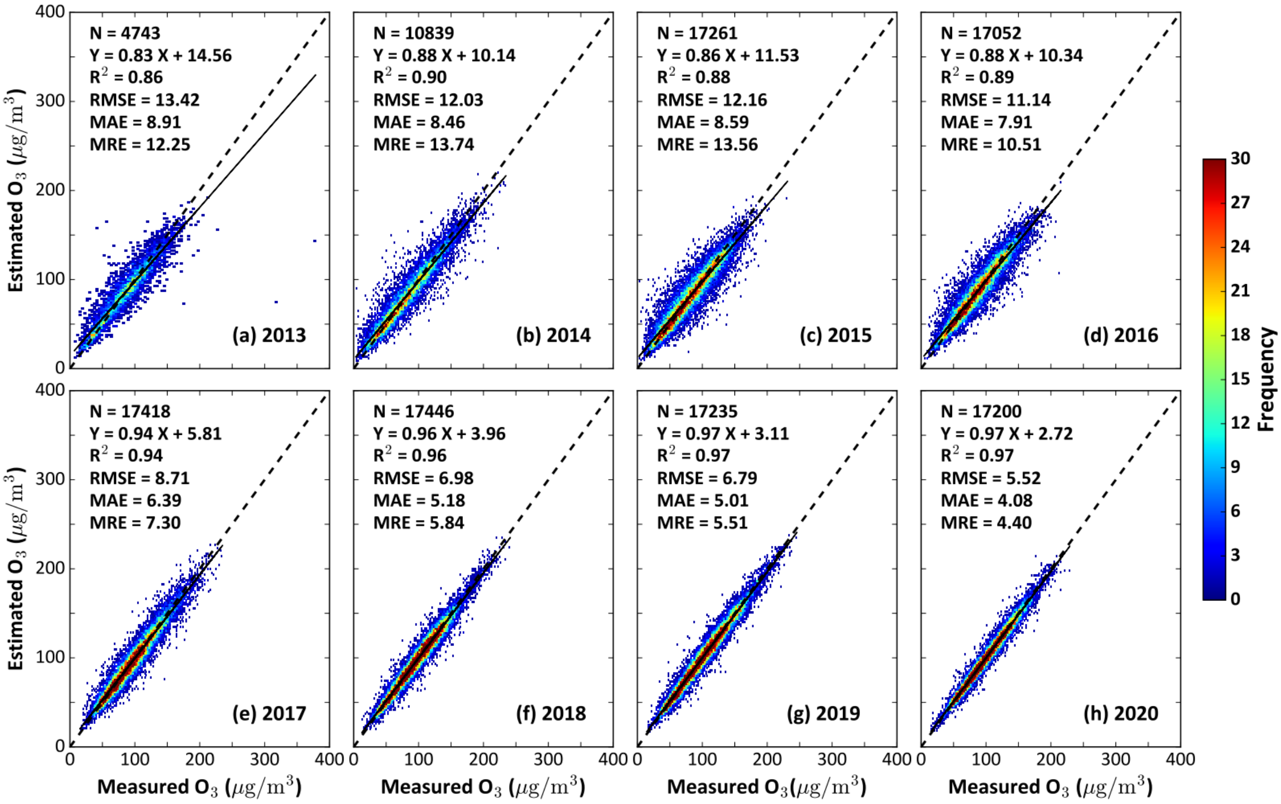


Figure 8. Validation of monthly composite MDA8 O_3 concentrations (μg/m^3) for (a-g) each year and (h) all years from 2013 to 2020 in China.

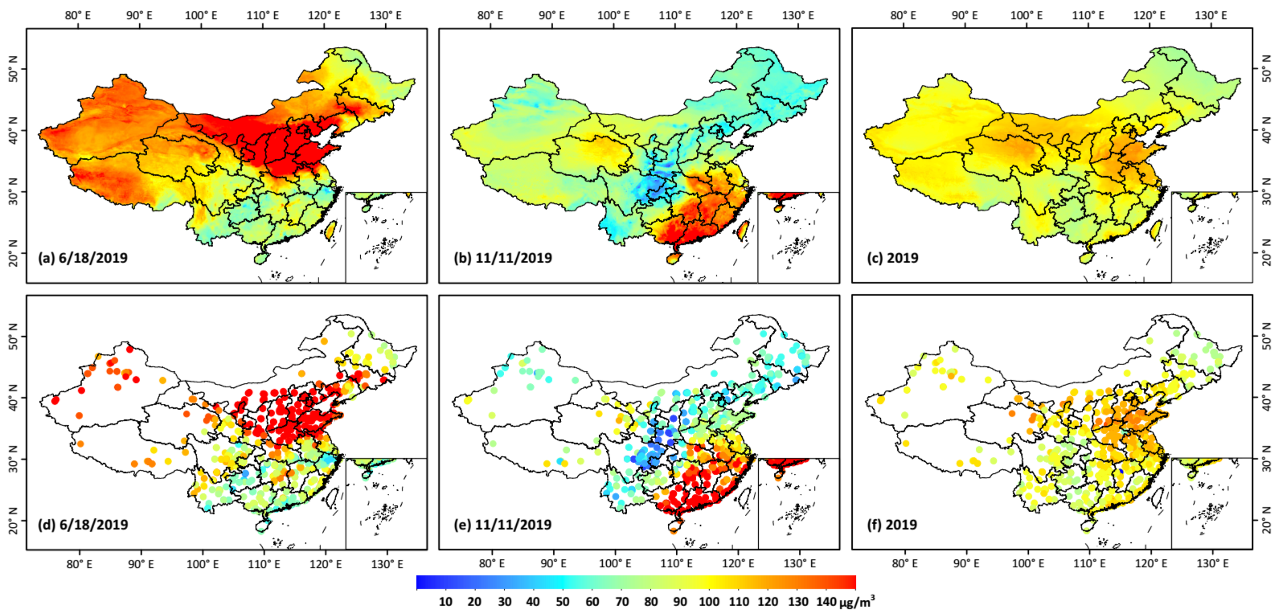


Figure 9. (a-c) STET-model-derived and (d-f) ground-based MDA8 O_3 maps on 18 June 2019 (a & d), 11 November 2019 (b & e), and the annual mean map for 2019 (c & f) covering China.

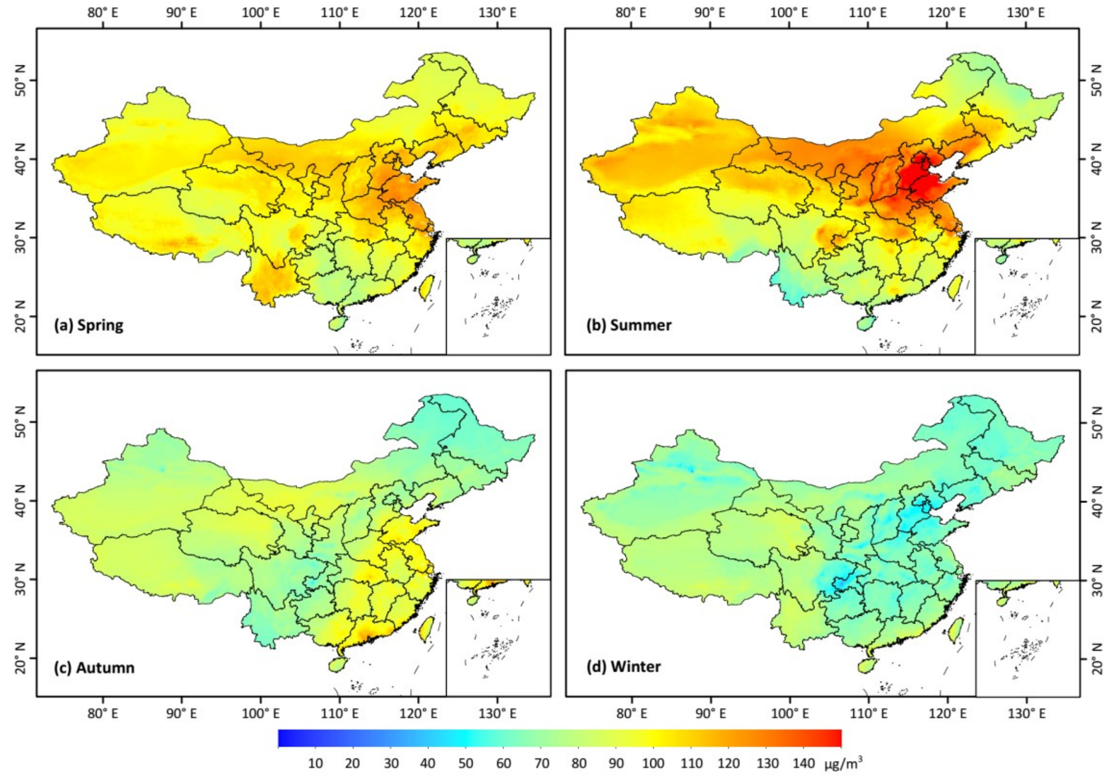


Figure 10. Multi-year seasonal mean MDA8 O_3 maps (10 km) averaged over the period 2013–2020 across China.

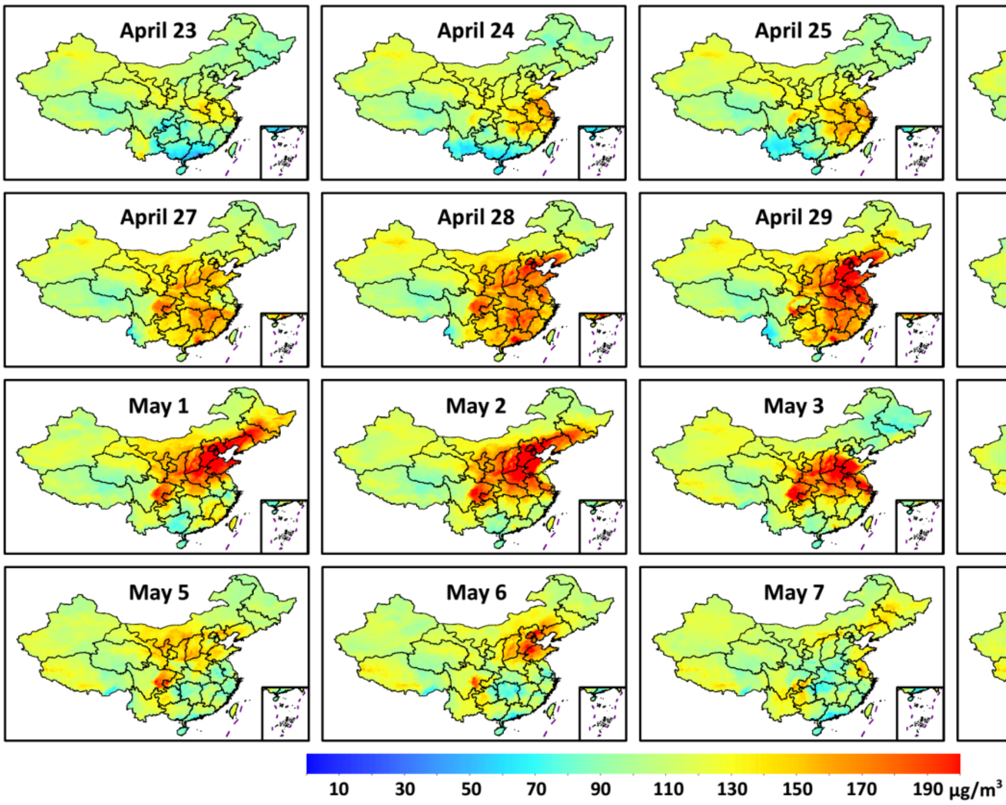


Figure 11. A typical example of a severe O_3 pollution event that occurred from 23 April 2020 to 8 May 2020 in eastern China.

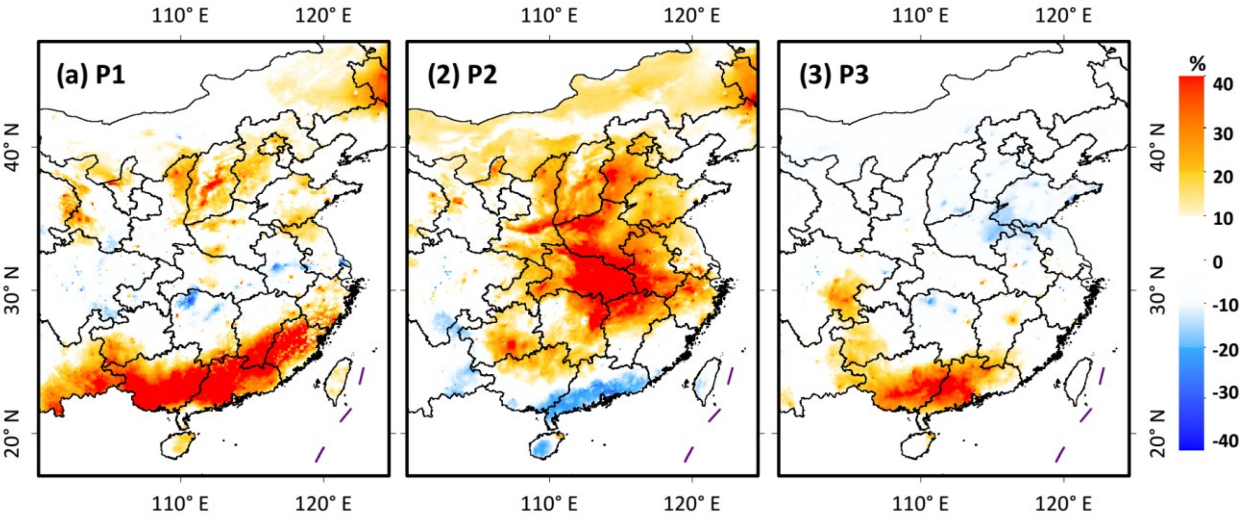


Figure 12. Relative changes (%) in mean MDA8 O_3 concentrations ($\mu g/m^3$) in 2020 (during the COVID-19 epidemic) and 2019 during the same periods: (a) Period 1 (P1, 1–25 January), (b) Period 2 (P2, 26 January to 17 February), and (c) Period 3 (P3, 18 February to 31 March) in eastern China.

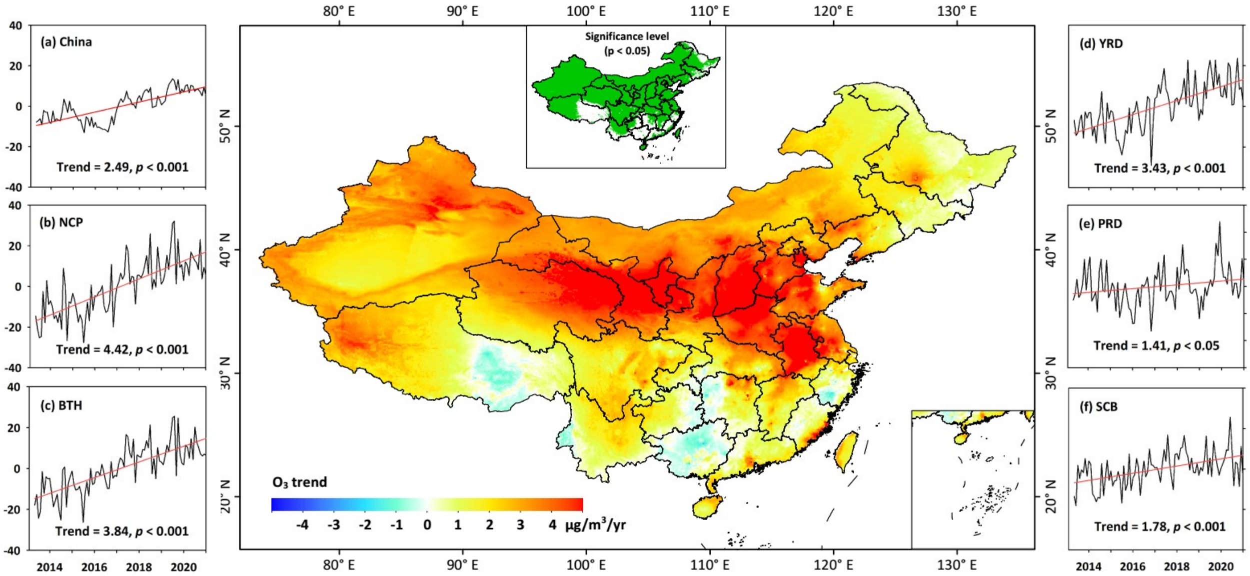


Figure 13. Linear MDA8 O_3 trends ($\mu g/m^3/yr$) calculated from de-seasonalized monthly MDA8 O_3 anomalies from 2013 to 2020 across China. The surrounding panels show the variations of monthly MDA8 O_3 anomalies in (a) China and (b-f) five typical regions.

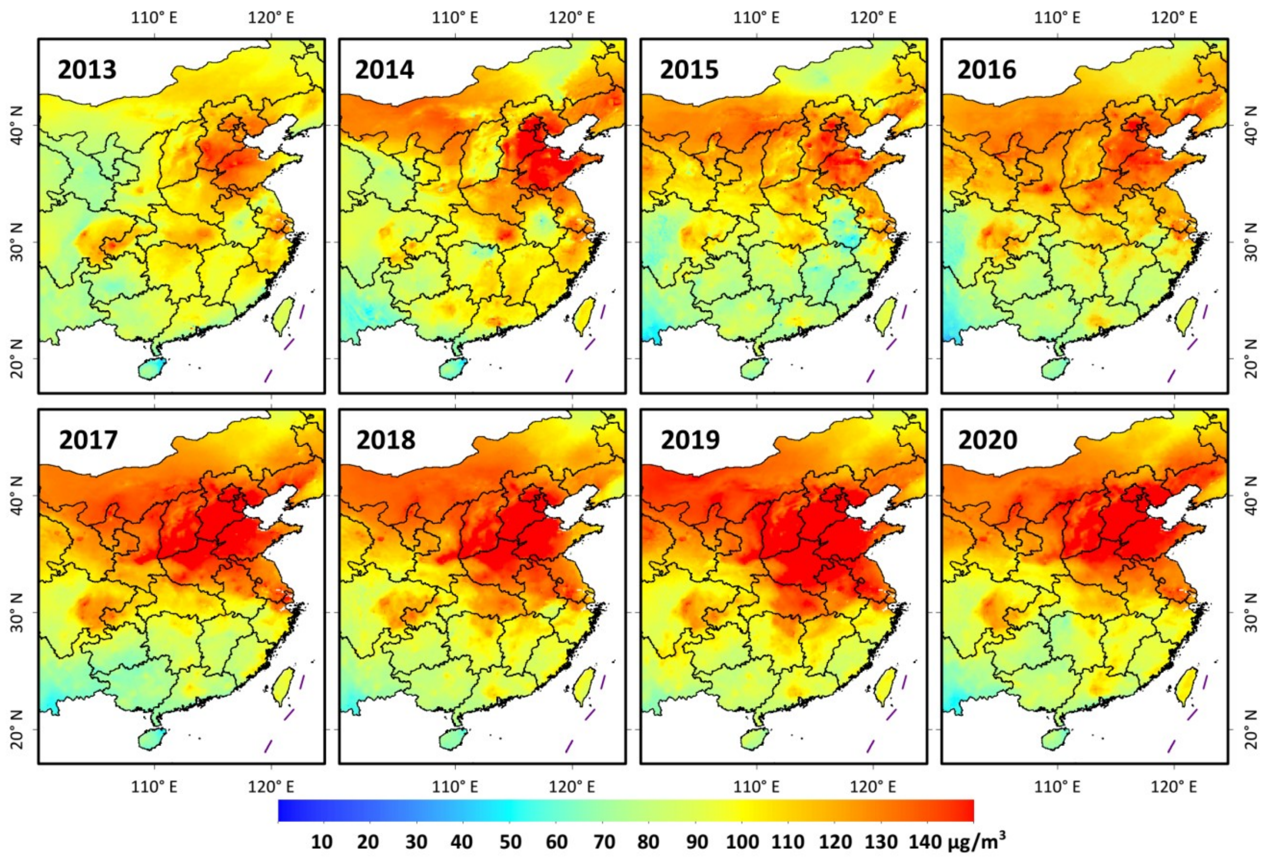
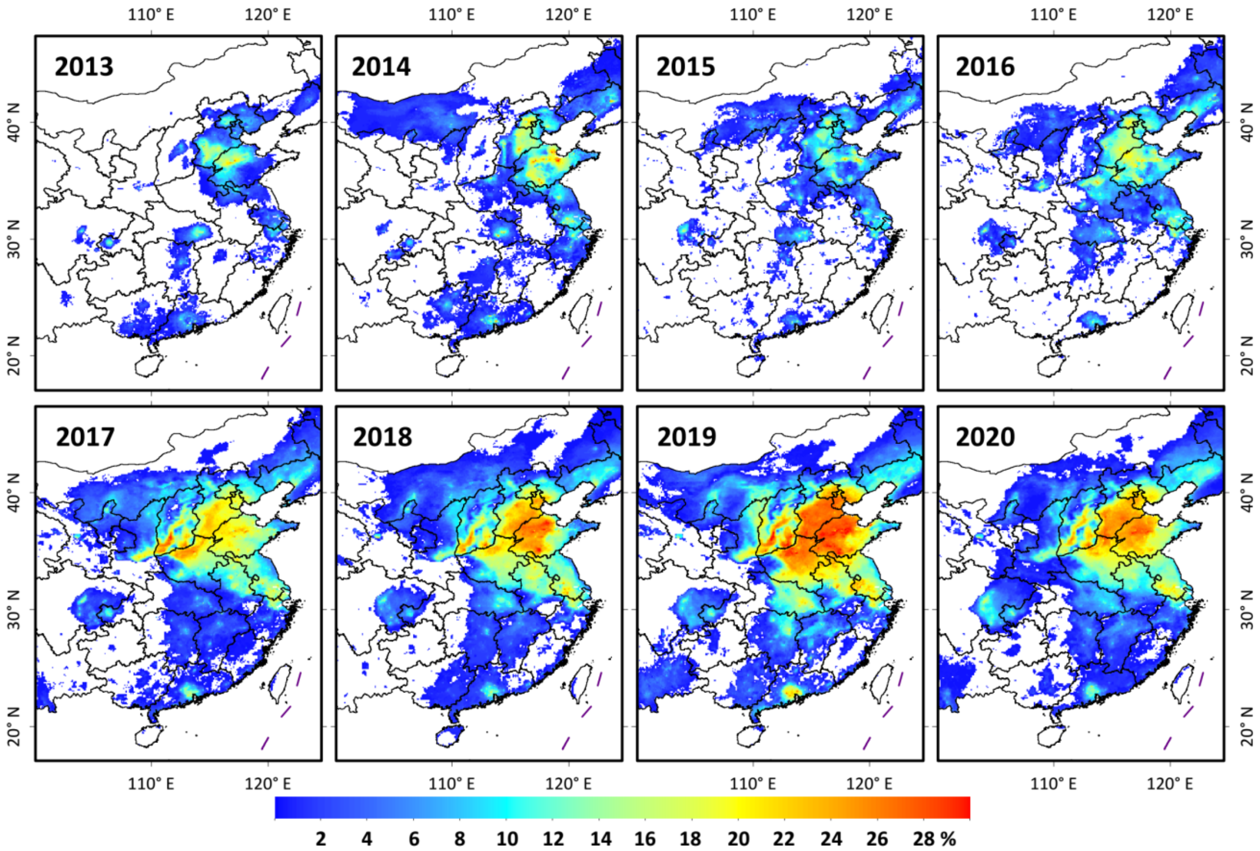


Figure 14. Spatial distributions of summer mean MDA8 O_3 concentrations (μg/m^3) from 2013 to 2020 in eastern China.



939
940 **Figure 15.** Spatial distributions of the percentage of days exceeding the ambient O_3 standard (i.e.,
941 MDA8 O_3 concentrations > 160 μg/m^3) from 2013 to 2020 in China.

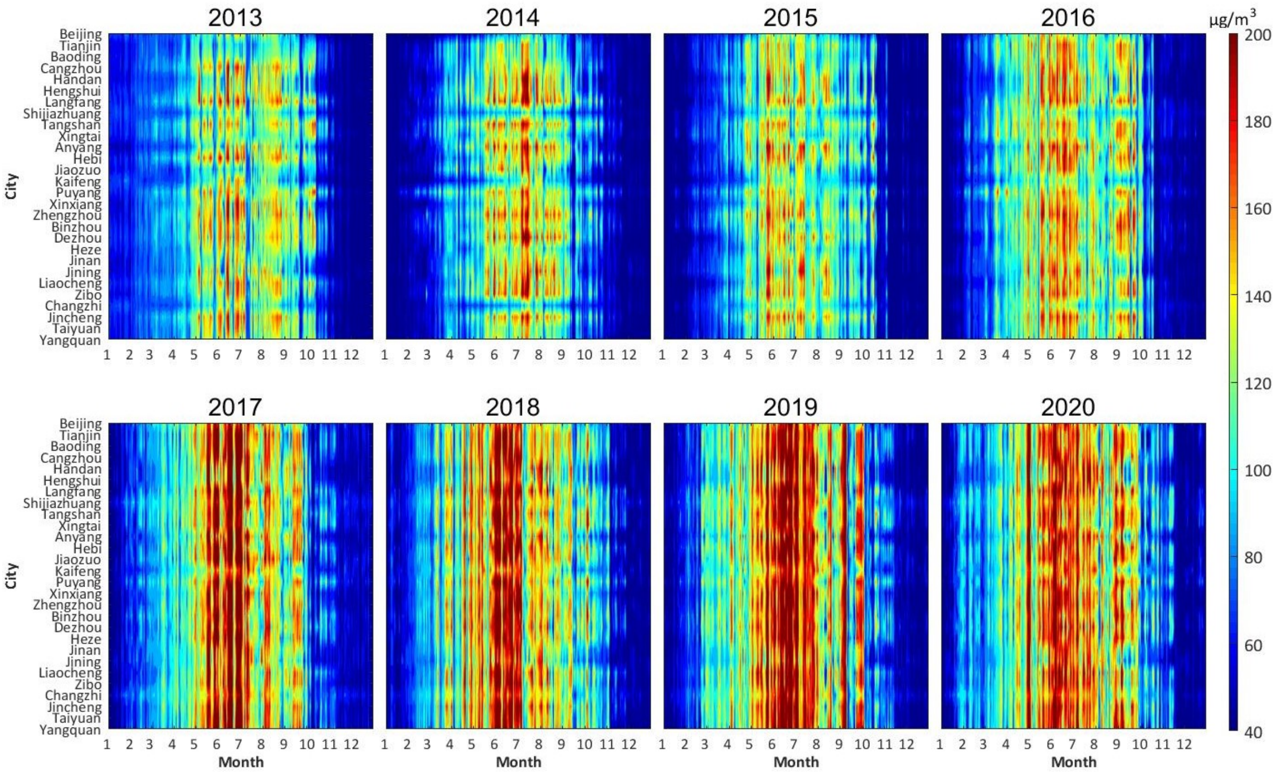


Figure 16. Heat maps of MDA8 O_3 concentrations ($\mu g/m^3$) for each year from 2013 to 2020 at the "2+26" cities in China.