First retrieval of 24-hourly 1-km-resolution gapless surface ozone (O 3) from space in China using artificial intelligence: diurnal variations and implications for air quality and phytotoxicity

Fan Cheng[1], Zhanqing Li[2], Zeyu Yang[1], Ruohan Li[3], Dongdong Wang[3], Aolin Jia[4], Ke Li[5], Bin Zhao[6,7], Shuxiao Wang[6,7], Dejia Yin[6,7], Shengyue Li[6,7], Wenhao Xue[8], Maureen Cribb[2], and Jing Wei[2]

[1]College of Global Change and Earth System Science, Faculty of Geographical Science, Beijing Normal University
[2]Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland
[3]Department of Geographical Sciences, University of Maryland
[4]Department of Environment Research and Innovation, Institute of Science and Technology (LIST)
[5]Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, School of Environmental Science and Engineering, Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Nanjing University of Information Science and Technology
[6]School of Environment, State Key Joint Laboratory of Environmental Simulation and Pollution Control, Tsinghua University
[7]State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex
[8]School of Economics, Qingdao University

February 25, 2024

First retrieval of 24-hourly 1-km-resolution gapless surface ozone (O_3) from space in China using artificial intelligence: diurnal variations and implications for air quality and phytotoxicity

Fan Cheng[1], Zhanqing Li[2*], Zeyu Yang[1], Ruohan Li[3], Dongdong Wang[3], Aolin Jia[4], Ke Li[5], Bin Zhao[6,7], Shuxiao Wang[6,7], Dejia Yin[6,7], Shengyue Li[6,7], Wenhao Xue[8], Maureen Cribb[2], Jing Wei[2*]

1. College of Global Change and Earth System Science, Faculty of Geographical Science, Beijing Normal University, Beijing, China
2. Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
3. Department of Geographical Sciences, University of Maryland, College Park, MD, USA
4. Department of Environment Research and Innovation, Luxembourg Institute of Science and Technology (LIST), Belvaux, Luxembourg
5. 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
6. State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China
7. State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, China
8. School of Economics, Qingdao University, Qingdao, China

* Correspondence:

Zhanqing Li (zhanqing@umd.edu), and Jing Wei (weijing@umd.edu)

Abstract

Surface ozone (O_3) is a crucial ambient pollutant gas that poses substantial risks to both human health and ecosystems. Nonetheless, there is a scarcity of high-spatial-resolution hourly surface O_3 data, particularly during the day when this information is needed due to the strong diurnal variations of O_3. We thus determined a best-performing artificial intelligence model to derive 24-hourly 1-km-resolution surface O_3 concentrations in China from a large array of satellite and surface data, which can portray well the diurnal variations of O_3 concentration. The overall sample-based cross-validated coefficients of determination (root-mean-square error) are 0.89 (15.74 $\mu g/m^3$), 0.91 (14.91 $\mu g/m^3$), and 0.85 (16.31 $\mu g/m^3$) during the full day (00:00–23:00 local time, or LT), daytime (08:00–17:00 LT), and nighttime (18:00–07:00 LT), respectively. The surface O_3 level generally rises from sunrise, around 07:00 LT, reaching a peak at ~15:00 LT, then continuously declining overnight. The magnitude of the diurnal variation amounts to 180% relative to its diurnal mean level. During daytime, solar radiation in the ultraviolet and shortwave spectral bands, along with temperature, explain more than half (32% and 24%) of the diurnal variations using the interpretable SHapley Additive exPlanations (SHAP) method, while nighttime O_3 levels are dominated by temperature (31%) and relative humidity (16%). In 2018, approximately 59%, 93%, and 100% of populated areas were susceptible to O_3 exposure risk for at least one day, with the maximum daily average 8-h O_3 levels surpassing the World Health Organization's recommended daily air quality standards of 160 $\mu g/m^3$, 120 $\mu g/m^3$, and 100 $\mu g/m^3$, respectively. Approximately 65%, 70%, and 99% of vegetated areas in China exceed the minimum critical levels for O_3 mixing ratios, as determined by the sum of all hourly values ≥ 0.06 $\mu mol\ mol^{-1}$ (SUM06), the sigmoidally weighted sum of all hourly values (W126), and accumulates over the threshold of 40 $nmol\ mol^{-1}$ (AOT40), respectively. Notably, gross primary productivity stands out as the most responsive indicator to surface O_3 pollution across various vegetated types in China, especially concerning the Hourly O_3 Accumulates without Threshold (AOT0, $R = -0.37$–-0.53, $p < 0.001$).

Keywords

Diurnal O_3 variations; (Explainable) Artificial Intelligence; SHAP; Air quality; Vegetation phytotoxicity

1. Introduction

Ozone (O_3) serves as a crucial trace gas in the atmosphere, primarily distributed in the stratosphere. It efficiently absorbs ultraviolet radiation, shielding virtually all Earth's living organisms and ecosystem from harmful effects. However, O_3 in the lower troposphere, particularly around the ground, injures human health and suppresses plant growth. As a greenhouse gas, it exerts radiative effects that leads to lower evaporation rates and relative humidity, altered precipitation patterns, and changes in atmospheric circulation (Allen et al., 2012; Fu and Tian, 2019; Lu et al., 2019; Stevenson et al., 2013). Being a primary air pollutant, its damage to human health is linked to various respiratory and cardiovascular diseases, such as kidney disease, circulatory disease, respiratory disease, and stroke (Brauer et al., 2016; Lin et al., 2018; Niu et al., 2022; Cai et al., 2023; C. Chen et al., 2023). Its harmful impacts on vegetation lead to reductions in carbon assimilation by most plants (Fares et al., 2013), gross primary productivity (GPP) (Yue and Unger, 2014), crop yield (Lin et al., 2018), and thus food supply (Wilkinson et al., 2012). When this highly reactive oxidant infiltrates leaves through stomata, the generation of additional reactive oxygen will trigger oxidative stress. Consequently, this hampers photosynthesis, impedes plant growth, therefore reducing yields (Ainsworth et al., 2012).

China, as one of the most populous countries with rapid development in the world, has suffered from significant air quality problems during the last four decades. In recent years, especially since 2013, China has enforced various strict air pollution control policies to significantly reduce anthropogenic pollutant emissions, leading to a notable air quality improvement, with a large reduction (39% during 2013–2020) in $PM_{2.5}$ concentrations (Wei et al., 2021a). On the contrary, surface O_3 pollution has worsened seriously during the same period (Huang et al., 2019; Y. Wang et al., 2020) at an average increasing rate of 2.49 $\mu g/m^3/yr$ ($p < 0.001$). The area surpassing the daily standard [i.e., maximum daily average 8-h (MDA8) O_3 = 160 $\mu g/m^3$] has also expanded considerably (Wei et al., 2022). In particular, the severity of surface O_3 pollution has now exceeded that of $PM_{2.5}$, becoming the primary pollutant affecting urban air quality (Wei et al., 2022; Liu et al., 2023; H. Wang et al., 2023). More effective surface O_3 control measures in the future are thus urgently needed.

To address the escalating problem of surface O_3 pollution, continuously monitoring and ascertaining its mass concentration is imperative. Ground-based observations have high precision and reliability, enabling the real-time monitoring of surface O_3 concentrations at specific sites. However, due to the uneven distribution of sites, achieving full coverage of O_3 monitoring remains a significant challenge. Chemical transport models such as the CMAQ, GEOS-Chem, and WRF-Chem models can simulate surface O_3 at a high temporal resolution (every one or several hours), but their accuracies are highly uncertain, and spatial resolutions are typically coarse, often at the degrees level. As the resolution increases, computational costs increase drastically. In particular, the surface O_3 from chemical reanalysis products (e.g., MERRA-2 and ERA5) have very large uncertainties in China compared with ground measurements (e.g., coefficient of determination (R^2) < 0.1 and root-mean-square error (RMSE) > 47 $\mu g/m^3$) (N. Wang et al., 2015; J. Hu et al., 2016; Qiao et al., 2019; Hou et al., 2022; Wei et al., 2022). Satellite remote sensing can provide O_3 retrievals of total column amount and vertical profiles from a series of instruments, such as the Tropospheric Monitoring Instrument (TROPOMI) and Ozone Monitoring Instrument (OMI), enabling us to monitor spatially continuous O_3 from space, together with other sources of data pertaining to surface O_3 (Zhu et al., 2022; J. Chen et al., 2022; Kang et al., 2021).

Trace amounts of O_3 are affected by numerous other factors through complex relationships, making highly accurate retrievals using conventional statistical approaches challenging. In recent years, considerable efforts have thus been undertaken to obtain surface O_3 concentrations using machine-learning (ML) approaches (Capilla, 2016; Ma et al., 2021; Song et al., 2022; Wang et al., 2022). We, for example, have used advanced ML to develop a long-term surface O_3 dataset with high accuracy in China called ChinaHighO_3 (Wei et al., 2022) that has gained widespread adoption for tracking air pollution (Y. Chen et al., 2022; Xia et al., 2022) and many public health studies (Zhang et al., 2022; Cai et al., 2023). However, most prior studies, including ours, have mainly concentrated on the daily (MDA8) scale, with only a handful delving into the diurnal hourly scale. Y. Zhang et al. (2023) have applied a bagged-tree model to generate hourly (09:00–16:00 LT) ground-level O_3 concentrations at a 5-km resolution over China by integrating the hourly Himawari-8 shortwave radiation product. B. Chen et al. (2023) built a deep-learning (DL) model to acquire hourly (10:00–15:00 LT) 5-km surface O_3 concentrations from Himawari-8 top-of-the-atmosphere radiation. Wang et al. (2022) explored a self-adaptive geospatially local approach for estimating hourly (09:00–18:00 LT) 2-km surface O_3 concentrations across China using Himawari-8 AHI brightness temperatures at multiple thermal infrared bands. However, these studies have only estimated hourly surface O_3 during the daytime (usually less than 10 hours), failing to provide comprehensive 24-hour coverage. 24-hour data are of utmost importance for the calculation of not only air quality metrics like MDA8 but also O_3-exposure phytotoxicity indices such as 12-hour average surface O_3 concentrations (M12), Accumulation of surface O_3 concentrations without Threshold (AOT0), and SUM of surface O_3 concentrations ≥ 0.06 μmol mol^{-1} (SUM06). As far as air quality is concerned, O_3 as a pollutant is as important during daytime as at nighttime. Retrievals from most previous studies also have large data gaps due to the presence of clouds that handicaps optical satellite remote sensing, seriously limiting their applications.

For the first time, we attempt to derive *24-hourly 1-km-resolution gapless* surface O_3 concentrations across China using a best-performing model by making use of ample satellite, ground, and model datasets pertinent to O_3, such as surface shortwave radiation from the Himawari-8 geostationary satellite and temperature retrievals, and many other factors influencing surface O_3 concentrations. The best-performing model was selected from 15 different ML and DL models, considering both model accuracy and efficiency. After being cross-validated independently against ground measurements, the O_3 products undergo a comprehensive analysis of spatial and temporal variations throughout both the daytime and nighttime, with their driving factors identified and quantified by leveraging the Explainable Artificial Intelligence (XAI) - SHAP (SHapley Additive exPlanations) method. Additionally, using the 24-hour data, we compute both the MDA8 O_3 and various O_3-exposure phytotoxicity indices and assess the short-term health risks of exposure to surface O_3 pollution, as well as the adverse impacts of O_3 pollution on vegetation.

2. Materials and methods

2.1 Data sources

2.1.1 Surface O_3 observations

This study employs ground-level hourly O_3 observations (μg/m³) from the Ministry of Ecology and Environment in China from a total of 1558 monitoring stations in 2018 (c.f. Figure 1 in Wei et al., 2020). Flagged invalid data are first excluded. Note that the observation status transitioned from standard (273 K, 1013 kPa) to room (298 K, 1013 kPa) conditions after 31 August 2018 (Wei et al., 2022). Data after this date were thus adjusted by multiplying by 1.09375 to maintain data consistency (MEE, 2018).

2.1.2 Ancillary data for surface O_3 retrievals

Surface O_3, an important secondary pollutant in the atmosphere, is influenced by various factors during its formation and dissipation. Downward shortwave radiation (DSR) and land surface temperature (LST) are the two most important factors influencing diurnal surface O_3 concentrations (Wei et al., 2022). Here, hourly DSR and LST data from geostationary satellites are adopted by virtue of their high spatial and temporal variations. The 1-km DSR hourly data are obtained from the Geostationary-NASA Earth Exchange Level 2 product. It was generated using a physical-based look-up-table approach from data collected from the new-generation geostationary Advanced Baseline Imager and Advanced Himawari Imager data (R. Li et al., 2023). Hourly LST data are derived from the Global Hourly All-sky-LST (GHA-LST) product with a downscaled 1-km-resolution, generated by combining a constellation of geostationary Earth orbit LST retrievals from the CGLS and MODIS MxD21 LST products. All-sky hourly LSTs are obtained using a spatiotemporal assimilation to address satellite gaps (Jia et al., 2023). In addition to shortwave radiation, hourly ultraviolet (UV) radiation, sourced from the ERA5 reanalysis with complete spatial coverage, has also been incorporated into our modeling, which plays a crucial role in impacting surface O_3 concentrations by catalyzing O_3 cycle initiating and controlling O_3 generating rate (Barnard et al., 2003; Seinfeld and Pandis, 2016).

Anthropogenic emissions of gaseous precursors are key ingredients of the photochemically generated O_3, including NO_X, VOCs, and CO, obtained from the 1-km-resolution daily Air Benefit and Cost and Attainment Assessment System-Emission Inventory version 2.0 (ABaCAS-EI v2.0) dataset covering China (S. Li et al., 2023). Population distribution is also employed to represent the anthropogenic emissions of precursors collected from the 1-km annual LandScan™ product. Meteorological variables have significant and diverse impacts on air pollutants. Employed in our model are the following most influential ones from hourly ERA5 global reanalysis data: boundary layer height (BLH), relative humidity (RH), total precipitation (TP), surface pressure (SP), wind speed (WS), and wind direction (WD) (calculated from the u- and v-components of winds). The following variables attributed to surface conditions are also included: the Shuttle Radar Topography Mission 90-m DEM and MODIS 1-km normalized difference vegetation index (NDVI) products. Altogether, we have gathered and employed a total of 21 variables for daytime and 19 for nighttime, with details provided in Table 1. All ancillary data are resampled and reaggregated to match DSR (0.01° × 0.01°) using the bidirectional interpolation method, following our previous study (Wei et al., 2023).

Table 1. An overview of data sources employed in this study, where TR and SR stand for temporal resolution and spatial resolution, respectively.

Data	Full name of the variable	Abbreviation	Unit	TR	SR	Source
Ground-level O_3	Ground-level O_3 measurements	O_3 sites	μg/m³	1 hour	Site	CNEMC
Solar Radiation	Downward shortwave radiation	DSR	W/m^2	1 hour	1 km	GeoNEX-L2
Temperature	Land surface temperature	LST	K	1 hour	1 km	GHA-LST
Meteorological Factors	Boundary layer height	BLH	m	1 hour	0.25°	ERA5
	Ultraviolet radiation	UV	W/m^2	1 hour	0.25°	
	Relative humidity	RH	%	1 hour	0.1°	
	Total precipitation	TP	m	1 hour	0.1°	
	Surface pressure	SP	hPa	1 hour	0.1°	
	Wind direction	WD	m/s	1 hour	0.1°	
	Wind speed	WS	m/s	1 hour	0.1°	
O_3 Precursor	Nitrogen oxides	NOx	mol/cm^2	1 day	1 km	ABaCAS-EI
	Volatile organic compounds	VOC	mol/cm^2	1 day	1 km	
	Carbon monoxide	CO	mol/cm^2	1 day	1 km	
Other Factors	Digital elevation model	DEM	m	-	90 m	SRTM
	Population	POP	people	1 year	1 km	LandScan™
	Normalized difference vegetation index	NDVI	/	16 day	1 km	MODIS

2.2 AI Model establishing and selection

Tropospheric O_3 accounts for 8–15% of total O_3, and surface O_3 constitutes 34–83% of the troposphere (David and Nair, 2011). As such, the satellite-retrieved total column O_3 amount only plays a minor role in dictating the variations in surface O_3 levels, rendering its remote sensing from satellite highly challenging. Besides exploiting various pertinent data sources, it is equally imperative to find the best model that can most effectively and efficiently extract any useful information for which AI has been proven to be most competent. To find the best-performing one, we applied 15 models, including eight ML and seven DL models. For the ML ones, we choose the original Decision Tree (DT) and six DT-derived ensemble-learning models consisting of multiple base models, falling into two categories of bagging and boosting. Bagging models combine multiple independent base models through averaging or voting, including random forests (RF; Breiman, 2001) and extremely randomized trees (ERT; Geurts et al., 2006). Boosting models entail iteratively constructing base models, with each model refining its performance based on the feedback from the preceding model, including Adaptive Boosting (AdaBoost), Gradient Boosting Decision Tree, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). AdaBoost is one of the earliest techniques within the realm of boosting and assigns a higher weight to misclassified samples from the previous base model in each iteration (Freund and Schapire, 1997). GDBT constructs base models by progressively improving the loss function (Friedman, 2001), and both XGBoost and LightGBM are optimizations of the GDBT framework. XGBoost introduces training loss (second-order Taylor expansion) and regularization, while LightGBM applies a histogram optimization and gradient-based one-side sampling method (Chen and Guestrin, 2016; Ke et al., 2017). CatBoost is specially tailored for handling categorical features (Sagi and Rokach, 2018).

For DL, we select among the Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), Deep Belief Networks (DBN), Deep Residual Network (ResNet), Residual Next (ResNeXt), and Deep Forest (DF) models. The MLP model serves as the fundamental neural network model capable of approximating complex nonlinear functions (Du et al., 2022). CNN is employed for grid-pattern data and relies on the core of the convolutional layer that involves a series of operations like convolution (Yamashita et al., 2018). LSTM is a special recurrent neural network that effectively handles dependencies over long periods by using gate functions in its cell structure (Yu et al., 2019). DBN is a multi-layered neural network containing multiple restricted Boltzmann machines (Hinton et al., 2006). ResNet is designed to address network degradation issues in deeper neural networks by using shortcut connections to learn the residual between desired and current outputs of a specific layer, alleviating problems like gradient disappearance and network degradation (He et al., 2015). ResNeXt is an upgraded version of ResNet that introduces a novel building block called the "cardinality bottleneck" (Xie et al., 2017). DF is a hybrid model combining various tree-based models, rather than neurons, in each middle layer to handle non-linear relationships, allowing for the capture of complex data structures (Zhou and Feng, 2019). The above-mentioned total of 15 AI models, each run separately, are adopted here to identify a best-performing model for retrieving hourly surface O_3 by comparing their accuracies and efficiencies using the same hourly training and validation datasets (i.e., 17:00 LT, N = 476,840).

Previous studies have indicated that incorporating spatiotemporal factors can enhance the accuracy of the model in predicting air pollutants, considering their significant spatiotemporal continuity (T. Li et al., 2017; Wei et al., 2021b). Consequently, here, a novel technique that assigns weighted effects based on polar coordinates with multidimensions is employed to compute the spatiotemporal factors (Wei et al., 2023; Sun et al., 2022), leading to the new extended 4-Dimensional Space-Time AI (4D-STAI) model. Spatial information is described within Euclidean space utilizing spherical coordinates (Equations 1–3), and temporal information is represented using three helix-shaped trigonometric vectors (Equations 4–6), encompassing both diurnal variations and seasonal cycles of air pollution:

$$S_1 = sin(2\pi \frac{Lon}{360}), \tag{1}$$

$$S_2 = \cos(2\pi \frac{Lon}{360})\sin(2\pi \frac{Lat}{180}) \ , \tag{2}$$

$$S_3 = \cos(2\pi \frac{Lon}{360})cos(2\pi \frac{Lat}{180}), \tag{3}$$

$$T_1 = \frac{DOY}{N}, \tag{4}$$

$$T_2 = \cos(2\pi \frac{DOY}{N}), \tag{5}$$

$$T_3 = sin(2\pi \frac{DOY}{N}), \tag{6}$$

where the *Lon* signifies the longitude of each grid and the *Lat* signifies the latitude; *N* indicates a year's number of days in total and it is 365 for the year 2018; and *DOY* refers to the day of the year.

2.3 Validation and analysis methods

Similar to many previous studies (Di et al., 2017; Zhan et al., 2018; Kang et al., 2021; Y. Wang et al., 2021), the 10-fold cross-validation (10-CV) method is utilized for assessing and comparing the performance of model, performed at the sample-based (out-of-sample) and station-based (out-of-station) levels. Sample-CV is segregated according to all training data samples to evaluate the model's overall accuracy, while spatial-CV is divided based on ground-based monitors to measure the spatial prediction accuracy (T. Li et al., 2017; Wei et al., 2022). These two methods involve randomly dividing the entire dataset into 10 subsets. In each iteration, the model is trained on nine data subsets, with the rest for testing. This ensures that the independence of training and test data. This process runs in turn for 10 iterations, ensuring that all data participate in the model validation process (Rodriguez et al., 2010).

Surface O_3 concentrations are affected by diverse factors, as stated before, all of which exhibit variations across time and space. To comprehend the factors driving diurnal fluctuations in surface O_3 levels, we employed the XAI methodology, and the game-theoretic SHAP approach is applied to explain the model output. Specifically, SHAP quantifies the significance of a feature by contrasting the predictions of the model when including and excluding that particular feature (Lundberg et al., 2020). We thus assess the importance of all variables for each hour using SHAP's TreeExplainer. Figure 1 shows the flowchart of retrieving 24-hourly gapless surface O_3 concentrations in our study.

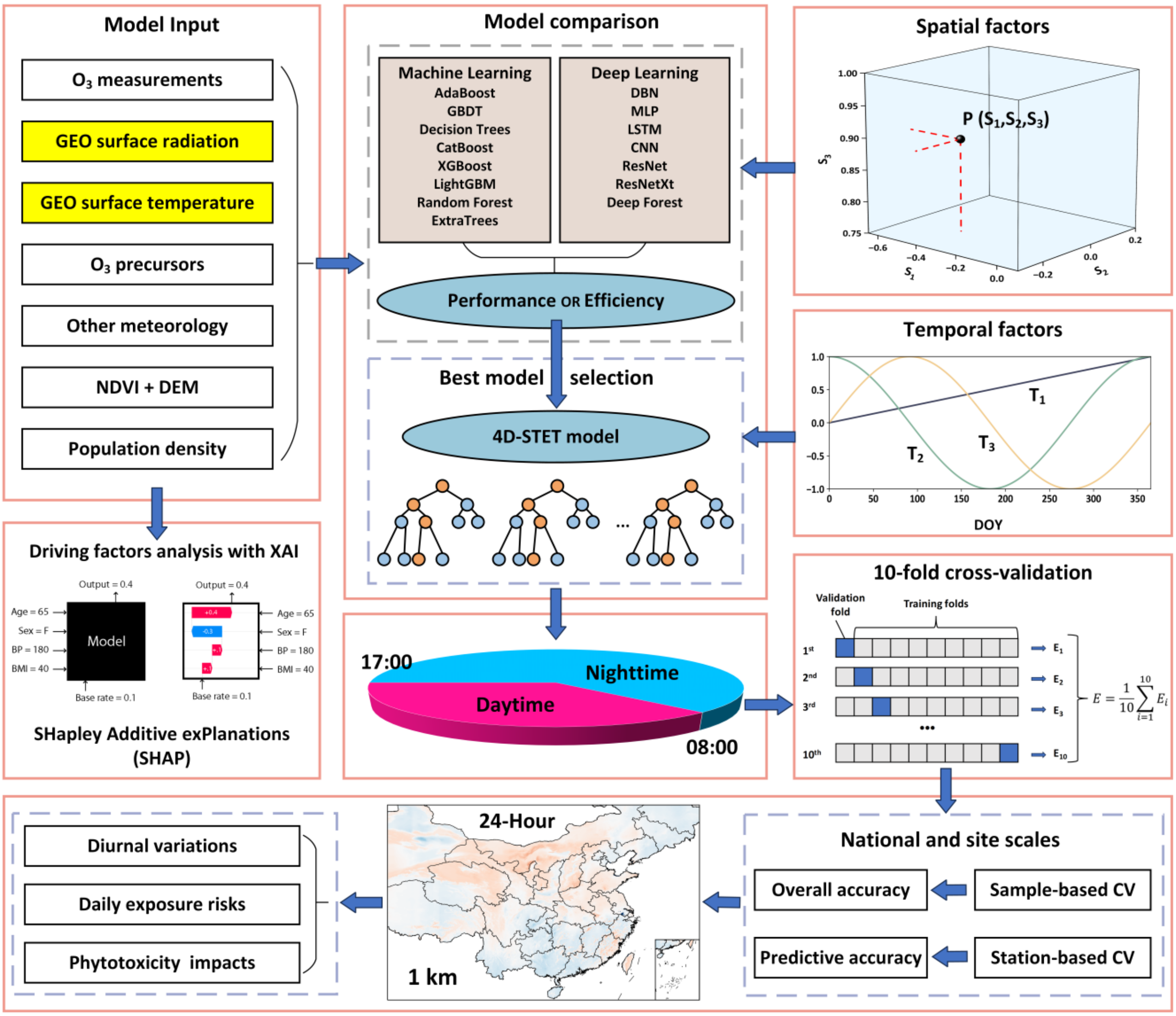

Figure 1. Flowchart of how satellite-derived 24-hourly gapless 1-km-resolution surface O_3 levels are retrieved across China in this study using AI.

2.4 O_3 phytotoxicity indices

Many studies have shown surface O_3 serves as an notable stressor in natural ecosystems, mainly affecting soil, biota and ecological processes (Kangasjarvi et al., 2005; Ainsworth et al., 2012; Super et al., 2015). We have chosen various O_3 phytotoxicity indices to investigate how vegetation responds to damage caused by surface O_3 exposure. MX refers to the hourly average value of O_3 within the specified period X, where M7 refers to the mean 7-hour O_3 concentration between 09:00–16:00 LT, and M12 refers to the mean 12-hour O_3 concentration between 08:00–20:00 LT, which mainly reflects the effects of O_3 levels on vegetation growth (Tong et al., 2009). This is expressed as

$$M7 = \frac{\sum_{i=1}^{n}[O_3]}{n}\ ppb\ (9 \ll i \leq 15), \quad (7)$$

$$M12 = \frac{\sum_{i=1}^{n}[O_3]}{n}\ ppb\ (8 \ll i \leq 19), \quad (8)$$

where the *i* represents the hours in UTC time, ranging from 0 to 23.

Fuhrer et al. (1997) and Grünhage et al. (1999) proposed the AOTX index representing O_3 the sum of hourly O_3 mixing ratios exceeding a threshold value (X nmol mol^{-1}) between 06:00–21:00 LT. AOT0 refers to AOTX when no threshold value is set (Equation 9), and AOT40 represents AOTX when the threshold value is set to 40 nmol mol^{-1} (Equation 10). AOT0 and AOT40 are usually used to measure the severity of vegetation damage caused by surface O_3 exposure. Generally, AOT40 is effective for assessing O_3 damage in highly polluted areas but may be less useful in regions with lower pollution levels. By contrast, AOT0 is more effectively across a broader range of pollution levels due to the retention of lower O_3 levels. To assess the extent of O_3 phytotoxicity, we use the maximum values of AOT0 and AOT40 over three consecutive months from April to September as annual results (Hayes and Bangor, 2017). Additionally, AOT40 causes damage to vegetation when it exceeds the thresholds of 3 ppm for agricultural crops and semi-natural vegetation, 5 ppm for forest trees, and 6 ppm for horticultural crops (Hayes and Bangor, 2017).

$$AOT0 = \sum_{i=1}^{n}([O_3])_i\ \ ppm\ (6 \ll i \leq 20) \quad (9)$$

$$AOT40 = \sum_{i=1}^{n}([O_3 - 40])_i\ \ ppm\ for\ [O_3]\ > 40\ ppb\ (6 \leq i \leq 20) \quad (10)$$

Heck and Cowling (1997) and Kohut (2007) introduced a SUM06 index representing the maximum cumulative value of hourly O_3 mixing ratios above 60 nmol mol^{-1} during 8:00–20:00 LT over three consecutive months from April to October (Equation 11). SUM06 is detrimental to vegetation when it exceeds the thresholds of 8–12 ppm for natural ecosystems, 10–16 ppm for tree seedings, and 15–20 ppm for crops (Heck and Cowling, 1997).

W126 index is a sigmoidally weighted hourly concentration (Lefohn and Runeckles 1987), calculated from the maximum of weighted cumulative values of hourly O_3 mixing ratios during

8:00–20:00 LT over three consecutive months from April to October (Equation 12). It exhibits a stronger response to elevated O_3 concentrations (Lefohn and Runeckles, 1987). W126 denotes damage to vegetation when it exceeds the specific thresholds of 5.9 ppm, 23.8 ppm, and 66.6 ppm for highly sensitive, moderately sensitive, and less sensitive species, respectively (Hayes and Bangor, 2017).

$$SUM06 = \sum_{i=1}^{n}([O_3])_i \ ppm \ for \ [O_3] \ > 60 \text{ ppb } (8 \ll i \leq 20) \quad (11)$$

$$W126 = \sum_{i=1}^{n}(w \times [O_3])_i \ ppm \ (8 \ll i \leq 20) \ w = \frac{1}{1+4403e^{(-0.126[O_3]_i)}} \quad (12)$$

To analyze the O_3 phytotoxicity to different vegetation types, the MODIS Land Use and Cover data were used to divide the land surface into three primary categories: forest, grassland, and cropland. We also assess and quantify the impacts of O_3 phytotoxicity on vegetation photosynthetic rate, growth situation, and yield by comparing six O_3 phytotoxicity indices (i.e., M7, M12, AOT0, AOT40, SUM06, and W126) with four vegetation abundance indices, i.e, NDVI, leaf area index (LAI), fraction of photosynthetically active radiation (FPAR), and GPP, collected from MODIS MOD13A2 16-day (1 km), MOD15A2H 8-day (500 m), MOD15A2H 8-day (500 m), and MOD17A2H 8-day (500 m) products, respectively.

3. Results and discussion

3.1 Model comparison and validation

3.1.1 Optimal model

Table 2 compares the model performance and efficiency among the 15 AI models in estimating hourly surface O_3 concentrations in China, utilizing the same hourly data samples (17:00 LT, N = 476,840). All eight ML models have fast training speeds and consume relatively small amounts of memory, of which the AdaBoost model shows the poorest performance. The two original GDBT and DT models have similar proficiencies in predicting hourly surface O_3, while the accuracies for their derived ensemble-learning models is improved, e.g., Catboost, XGBoost, RF, and LightGBM (e.g., CV-R^2 = 0.768, 0.805, 0.882, and 0.901, respectively). The ERT model operates swiftly (58 s) and performs the best (e.g., R^2 = 0.908) but uses the highest memory (10 GB). Among the seven DL models, the DBN model performs the worst and consumes a large amount of memory despite its fast training speed. The MLP model works better with improved accuracy and minimal memory but takes a significant amount of time (7300 s). The LSTM model exhibits enhanced performance when incorporating time series information, delivering results efficiently. With continuous optimization in both model architectures and loss functions, the accuracy of surface O_3 estimates consistently increases, e.g., CV-R^2 = 0.709, 0.749, and 0.755 for the CNN, ResNeXt, and ResNeXt models, respectively. However, their training speeds and memory requirements continue to increase. DF outperforms other DL models in terms of performance (e.g., R^2 = 0.904), yet it also has the most running time (15885 s) and memory (17 GB). Interestingly, most DL models perform not as accurately as and less efficiently than ML models in addressing regression problems (Grinsztajn et al., 2022) because DL is predominantly designed for handling more intricate computer vision tasks (e.g., object recognition and detection), requiring a vast amount of data samples (Shinde and Shah, 2018). When compared to the two best-performing ML and DL models, despite very similar accuracies (CV-R^2 = 0.908 versus 0.904, Slope = 0.877 versus 1.009), the ERT model saves approximately 270 and 2 times less time and memory than the DF model. The ERT model is thus chosen to estimate surface O_3 at the hourly scale in our study.

Table 2. Performance and efficiency comparison of different 4-dimensional (4D) space-time (ST) ML and DL models for estimating surface O_3 concentrations at 17:00 LT in China, where numbers in bold indicate the best evaluation indices.

Model Type	Model	R^2	Slope	RMSE	MAE	TSpeed (s)	TMemory (GB)
4D-STML	AdaBoost	0.461	0.390	41.42	33.32	312.22	0.0020
	GDBT	0.730	0.702	26.56	19.86	690.68	0.0007
	DT	0.756	0.866	26.09	17.55	**9.17**	0.0500
	CatBoost	0.768	0.741	24.63	18.37	34.036	0.0430
	XGBoost	0.805	0.790	22.53	16.49	250.08	0.8056
	RF	0.882	0.840	17.67	12.55	237.58	5.9900
	LightGBM	0.901	0.877	16.05	11.61	58.13	0.1210
	ERT	**0.908**	0.877	**15.55**	**10.93**	57.06	9.8000
4D-STDL	DBN	0.456	0.455	37.65	29.15	665.72	3.5620
	MLP	0.631	0.616	31.01	23.33	7317.21	**0.0005**
	LSTM	0.699	0.682	28.03	20.98	411.38	0.3700
	CNN	0.709	0.718	27.53	20.71	3518.96	1.7700
	ResNet	0.749	0.761	25.57	19.01	3970.83	1.7600
	ResNeXt	0.755	0.761	25.24	18.85	7918.61	1.8700
	DF	0.904	**1.009**	15.81	11.14	15885.32	17.1000

AdaBoost: Adaptive Boosting; GDBT: Gradient Boosting Decision Tree; DT: Decision Trees; CatBoost: Categorical Boosting; XGBoost: eXtreme Gradient Boosting; RF: Random Forest; LightGBM: Light Gradient Boosting Machine;

ERT: ExtraTrees; DBN: Deep Belief Network; MLP: Multilayer Perceptron; LSTM: Long Short Term Memory; CNN: Convolutional Neural Network; ResNet: Deep Residual Network; ResNeXt: ResNet Next; DF: Deep Forest.

3.1.2 Model performance

Figure 2 shows the sample-based overall accuracy of surface O_3 estimates at each hour from 0:00 to 23:00 LT using the 4-dimensional space-time extra-trees (4D-STET) model. The model accuracy varies for different hours. At 00:00 LT, O_3 estimates generally align closely with ground measurements, with a CV-R^2 of 0.78 and an RMSE of 17.08 μg/m³. The model performance slightly improves with comparable CV-R^2 (0.77–0.78) and lower RMSE values (14–17 μg/m³) until 07:00 LT. During the daytime, the model shows significant improvements, with increasing CV-R^2 (slopes closer to 1) and decreasing RMSEs. At 17:00 LT, the performance reaches its peak with a CV-R^2 of 0.91, a slope of 0.88, and a RMSE of 15.76 μg/m³. Subsequently, the performance of model deteriorates gradually. Overall, the model performs well across all hours, with CV-R^2 values above 0.75 and RMSE (MAE) values below 18 (13) μg/m³. Similar trends can be found in the station-based CV results. The model accuracy exhibits a gradual upward trend from midnight (0:00 LT) to the afternoon with increasing CV-R^2 and reduced RMSE values, reaching a peak at 17:00 LT (e.g., CV-R^2 = 0.89 and RMSE = 17.17 μg/m³), followed by a gradual decrease thereafter (Figure 3).

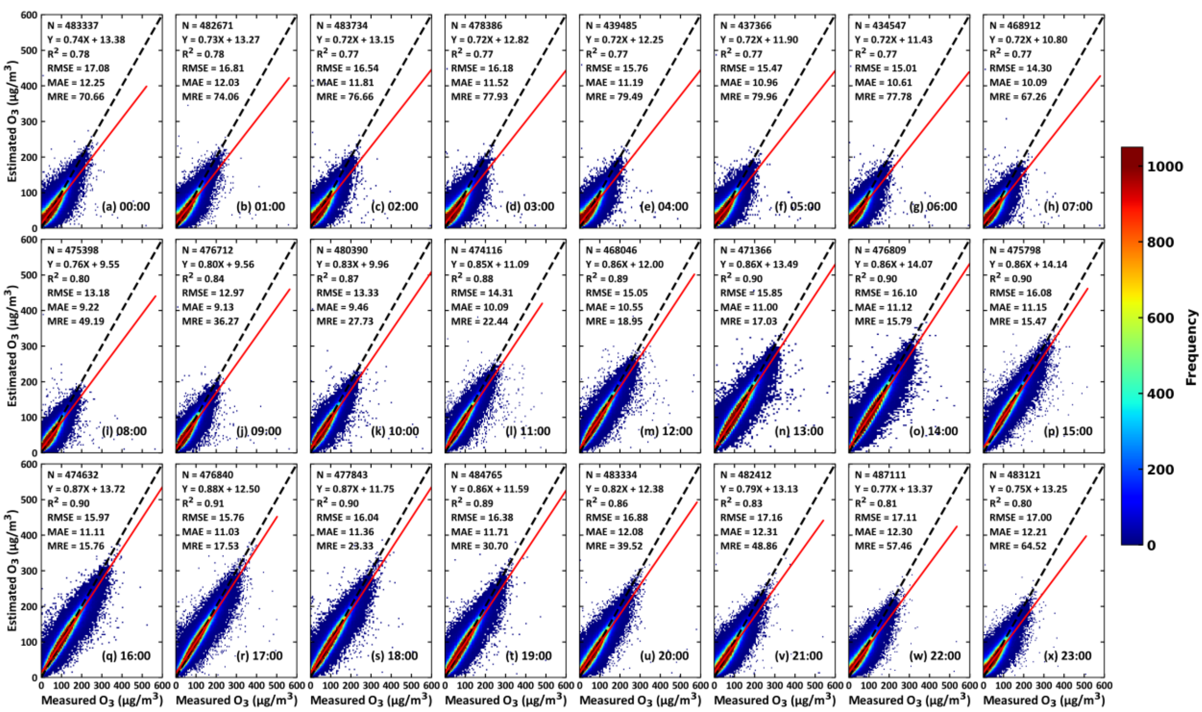

Figure 2. Out-of-sample cross-validation results of hourly O_3 estimates (μg/m³) from 00:00 to 23:00 LT for 2018 in China using the 4D-STET model. Black dashed lines denote 1:1 lines, and red solid lines denote best-fit lines from linear regression. The sample size (N), coefficient of determination (R^2), root-mean-square error (RMSE, μg/m³), mean absolute error (MAE, μg/m³), and mean relative error (MRE, %) are also given.

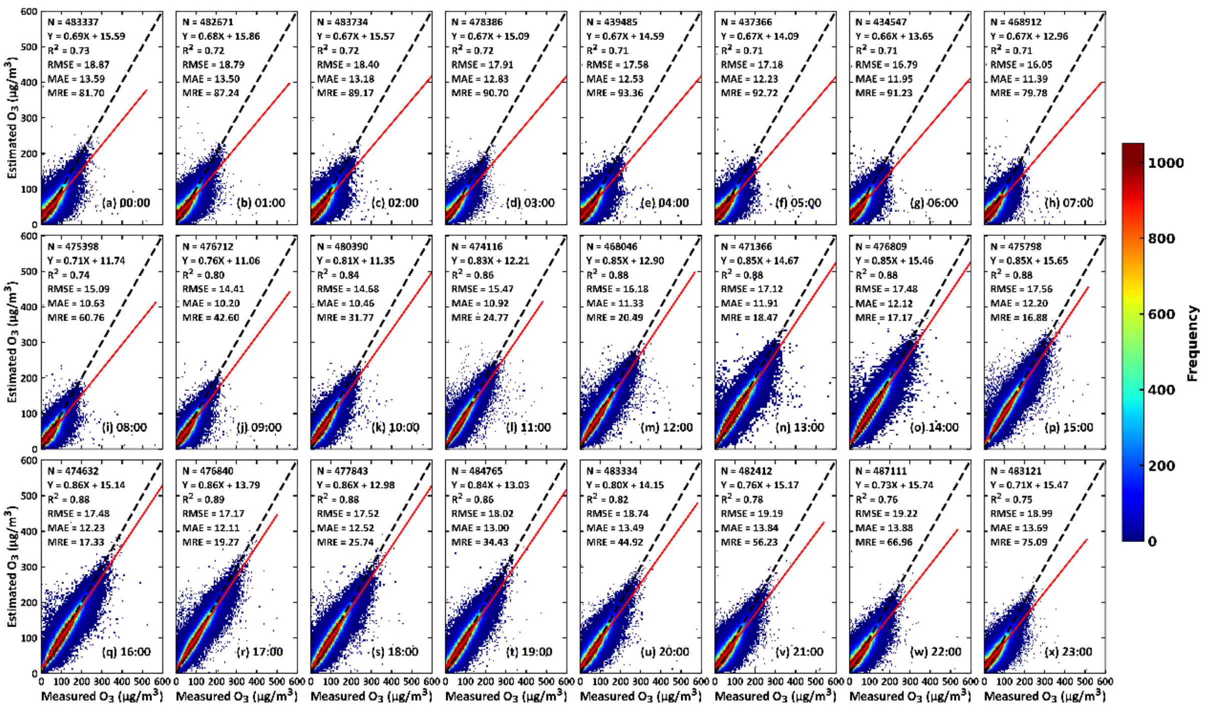

Figure 3. Same as Figure 2 but for out-of-station cross-validation results.

Overall, on the national scale, the model achieves a high overall (predictive) accuracy in retrieving hourly surface O_3 concentrations throughout the day, with average sample (spatial) CV-R^2 and RMSE values of 0.89 (0.86) and 15.74 (17.39) μg/m³. The superior performance of the model is also maintained during both daytime (e.g., CV-R^2 = 0.91 and 0.90, and RMSE = 14.91 and 16.31 μg/m³) and nighttime (e.g., CV-R^2 = 0.85 and 0.81, and RMSE = 16.31 and 18.14 μg/m³) (Table 3). The model also performs well in estimating and predicting all day, daytime, and nighttime hourly surface O_3 concentrations at the regional scale, especially in the Beijing-Tianjin-Hebei (BTH) region [e.g., sample (spatial) CV-R^2 = 0.90–0.94 (0.88–0.93)].

Table 3. Cross-validation (CV) statistics of hourly O_3 estimates (μg/m³) for all day, daytime, and nighttime periods in China and each typical region, using the 4D-STET model. All day represents 00:00–23:00 LT, daytime represents 08:00–17:00 LT, and nighttime represents the other hours.

Region	Period	Sample-CV			Spatial-CV		
		R^2	RMSE	MAE	R^2	RMSE	MAE
China	All day	0.89	15.74	11.10	0.86	17.39	12.33
	Daytime	0.91	14.91	10.38	0.90	16.31	11.41
	Nighttime	0.85	16.31	11.62	0.81	18.14	12.99
BTH	All day	0.93	15.53	10.62	0.91	17.10	11.38
	Daytime	0.94	14.66	9.78	0.93	16.67	10.52
	Nighttime	0.90	16.14	11.23	0.88	17.41	12.01
YRD	All day	0.88	16.80	11.40	0.87	17.51	11.98
	Daytime	0.90	17.37	11.67	0.89	17.99	12.18
	Nighttime	0.83	16.38	11.21	0.82	17.16	11.83
PRD	All day	0.88	17.21	11.85	0.86	18.46	12.79
	Daytime	0.89	18.38	12.56	0.88	19.29	13.24
	Nighttime	0.83	16.30	11.33	0.80	17.84	12.46

Figure 4 shows the model's accuracy of all hourly retrievals in 2018 across China at individual sites. Overall, our model demonstrates strong performance and adaptability in estimating surface hourly O_3 levels at most sites, without weak spatial patterns. At ~86% of the sites, sample-based CV-R^2 values exceed 0.80, and 80% (68%) of the sites have sample-based RMSE (MAE) values below 18 (13) μg/m³, particularly in locations within eastern and central China (CV-R^2 > 0.9) where the ground observation network is denser (Figure 4a-c). Spatial patterns for the spatial-CV results are similar, but the model exhibits an overall reduced accuracy in its predictive capability, with decreasing CV-R^2 values and increasing uncertainties for most sites across China (Figure 4d-f). Nevertheless, more than 81% and 82% (75%) of the sites still maintain reliability, with high spatial CV-R^2 > 0.80 and low RMSE (MAE) < 20 (14) μg/m³. Poor performance is primarily located at a few sites in western and northwestern China. This variance in the model's predictive ability is mainly caused by large differences in meteorological conditions and pollutant types and the small number of sites in western China. In general, surface O_3 retrievals are highly consistent across national, regional, and site scales, reaffirming the model's robust performance.

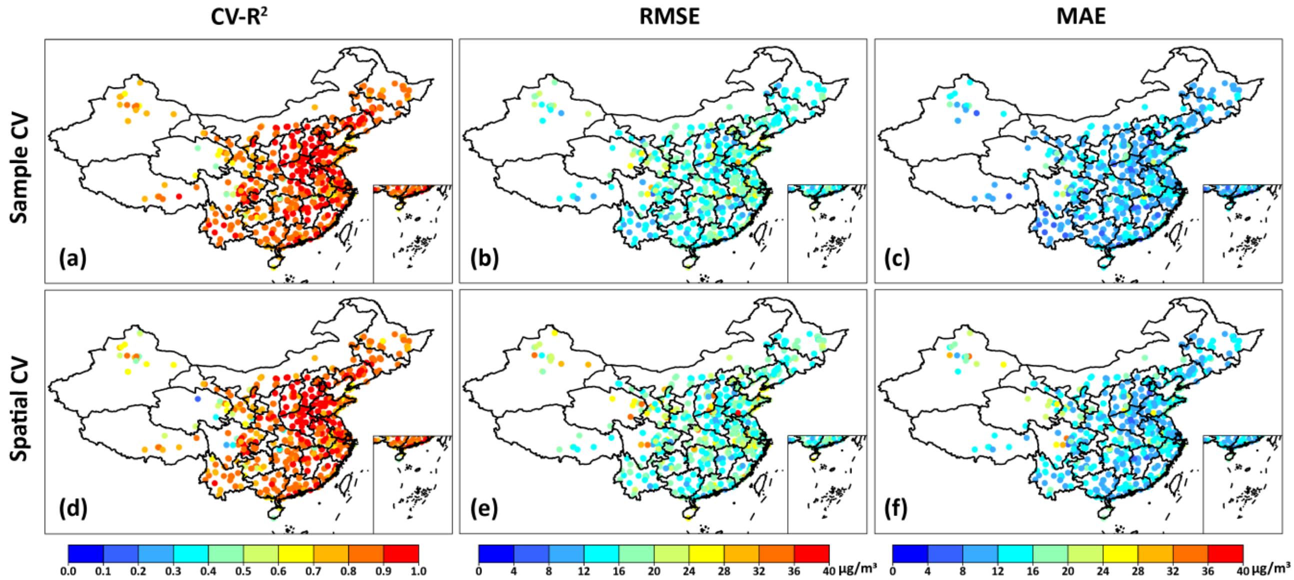

Figure 4. Individual-site-scale (a-c) out-of-sample (top row) and (d-f) out-of-station (bottom row) cross-validation results (including CV-R^2, RMSE, and MAE) for surface O_3 retrievals (μg/m³) collected from all hours in 2018 in China using the 4D-STET model.

3.2 Diurnal variations in surface O_3 and driving factors

Figure 5 shows satellite-derived gapless surface O_3 concentrations at a 1-km resolution for each hour throughout the day in China during the year 2018. As expected, surface O_3 has strong diurnal variations. At 08:00 LT, it is at its lowest level (average = 49.71 ± 9.37 μg/m³), gradually increasing as the Sun continues to rise. The increasing rate of surface O_3 concentrations is faster in northern China than in southern China, followed by a widespread growing trend in central and eastern China from 10:00 to 12:00 LT. It continues to rise notably over most regions in the domain, with a majority of values surpassing 100 μg/m³, reaching a peak at 15:00 LT (average = 96.64 ± 9.53 μg/m³). After that, areas with high O_3 pollution shrink rapidly, with average values dropping from 94.83 ± 9.52 μg/m³ at 17:00 LT to 61.29 ± 9.51 μg/m³ at 23:00 LT. The decreasing rate in southern China outpaces that in northern China, with the fastest decline observed in southeast China. During the daytime, surface O_3 concentrations in most areas exceed 80 μg/m³, with particularly high levels observed in the North China Plain and northwest China. By contrast, during the nighttime, surface O_3 levels consistently fall below 50 μg/m³, except in a few western and central regions. In general, surface O_3 concentrations during the daytime (08:00–17:00 LT) (average = 79.54 ± 7.98 μg/m³) are notably higher, ~1.3 times, than that at the nighttime (18:00–07:00 LT) (average = 62.72 ± 8.83

μg/m³). This difference is primarily ascribed to the complex interplay of various atmospheric processes, emissions, and photochemical reactions, e.g., higher oxidant OX (O_3 and NO_2) levels during the daytime (S. Han et al., 2011) and lower nighttime boundary layer height facilitating NO titration reactions that deplete nighttime O_3 (Liao et al., 2023). Similar diurnal variations in surface O_3 are observed at regional scales, with the highest values typically occurring around 15:00 LT (Figure 7). However, the BTH region seems to show more significant changes with substantial fluctuations in hourly surface O_3 variations.

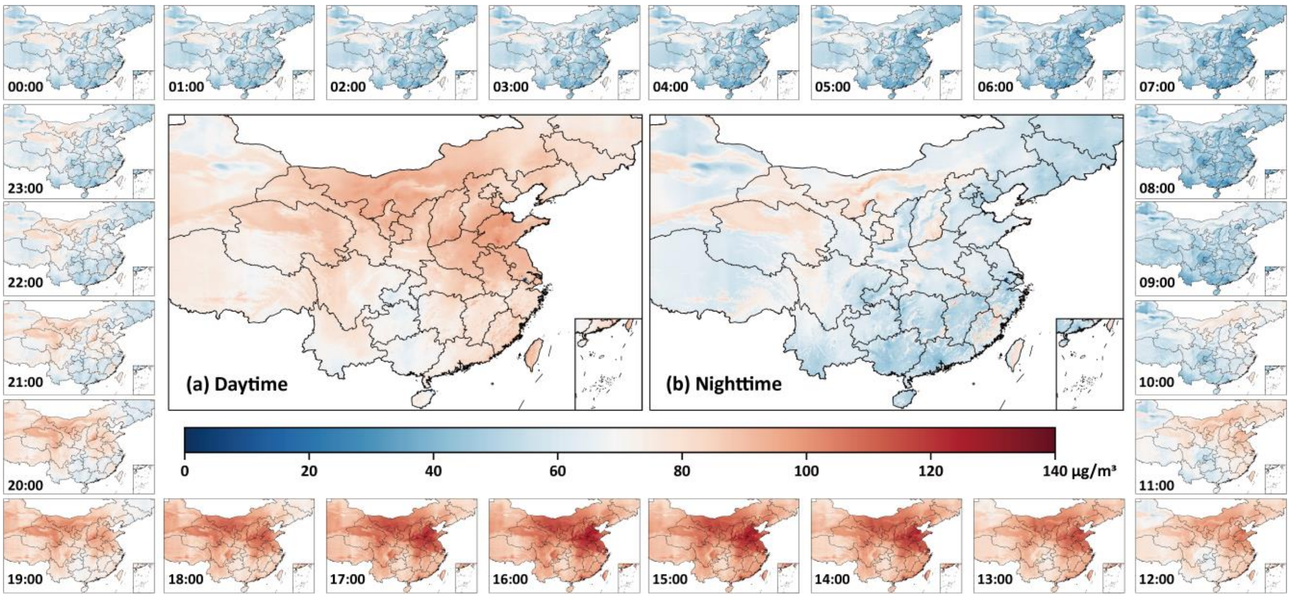

Figure 5. Satellite-derived 1-km-resolution surface O_3 concentrations for each hour throughout the day (00:00–23:00 LT, surrounding subplots), along with average maps during the (a) daytime (08:00–17:00 LT) and (b) nighttime (18:00–07:00 LT) in 2018 across China.

To gain a deeper insight into the driving factors affecting diurnal variations in surface O_3, we utilized XAI technology to compute the SHAP value for each of the variables and investigate their contributions at different hours throughout the day (Figure 6). In the morning hours (08:00 and 09:00 LT), the influencing factors are more intricate, with boundary layer height, wind, surface radiation (UV + shortwave), and temperature emerging as more significant contributors (SHAP = 10–18%). This can be explained by sunlight elevating solar radiant energy and near-surface temperatures, facilitating the photochemical reaction process (David and Nair, 2011; S. Han et al., 2011). Additionally, the intermittent vertical turbulent motion associated with boundary layer height and wind transport contributes to residual layer O_3 moving to near-surface, consequently elevating surface O_3 concentrations (Hu et al., 2012; Morris et al., 2010; Xu et al., 2020). During 10:00–17:00 LT, radiation (27–39%), temperature (13–32%), and RH (10–14%) consistently stand out as the three most influential factors. As the day progresses, however, the contribution of radiation gradually weakens, the role of temperature undergoes a significant upswing, and RH remains a relatively stable impact factor. The primary cause lies in increased human activities (R. Zhang et al., 2004), with heightened radiation and elevated temperatures substantially stimulating the production of atomic oxygen and oxidants and increasing photochemical reactions (Bloomer et al., 2009; Zhao et al., 2016; Wei et al., 2022; Zhang et al., 2023). RH can also impact the reaction process by interacting with water vapor, atomic oxygen, and cloud cover, and also affecting the dissipation of surface O_3 through dry deposition (Otero et al., 2018; Vautard et al., 2012). After solar radiation disappears towards evening, temperature becomes the most critical variable, but its contribution gradually decreases from ~44% (18:00 LT) to ~19% (07:00 LT). Other meteorological (e.g., RH = 16%, WS = 10%) and surface-related (e.g., DEM =11%) factors become increasingly more important in influencing surface O_3 variations. This may be caused by the dissipation of surface O_3 dominated by higher RH and lower WS at night (Tu et al., 2007; Gagliardi and Andenna, 2020).

In general, during the daytime, over half (56%) of the diurnal variations in surface O_3 can be attributed to surface radiation (32%) and surface temperature (24%). Other meteorological factors contribute ~35%, with RH (~11%), BLH (~9%), and WS (~5%) having a relatively larger influence. However, during the nighttime, LST contributes the most (~31%), accounting for nearly one-third, 7% higher than the daytime (24%). Other meteorological factors comprise approximately half (49%) of the influence, with the same three primary variables, i.e., RH, WS, and BLH, contributing at 16%, 9%, and 8%, respectively. Note that surface-related factors become more important during the nighttime compared to the daytime (13% versus 5%). Nevertheless, differences exist at the regional scale (Figure 7). In the BTH and YRD regions, LST contributions (36% and 27%) surpass radiation (21% and 22%) during the daytime, while WD, TP, and SP are more important meteorological factors during the nighttime. By contrast, in the PRD region, BLH and RH contribute the most (30% and 20%, respectively) during the daytime and nighttime, and meteorological factors contribute more significantly to surface O_3 variations compared to other regions. This region is closer to the sea, with southwest and southeast monsoons prevailing in the summer, and is affected by more weather systems (e.g., southwesterly wind, typhoons, and weak cyclones) (Jiang et al., 2015; H. Han et al., 2020).

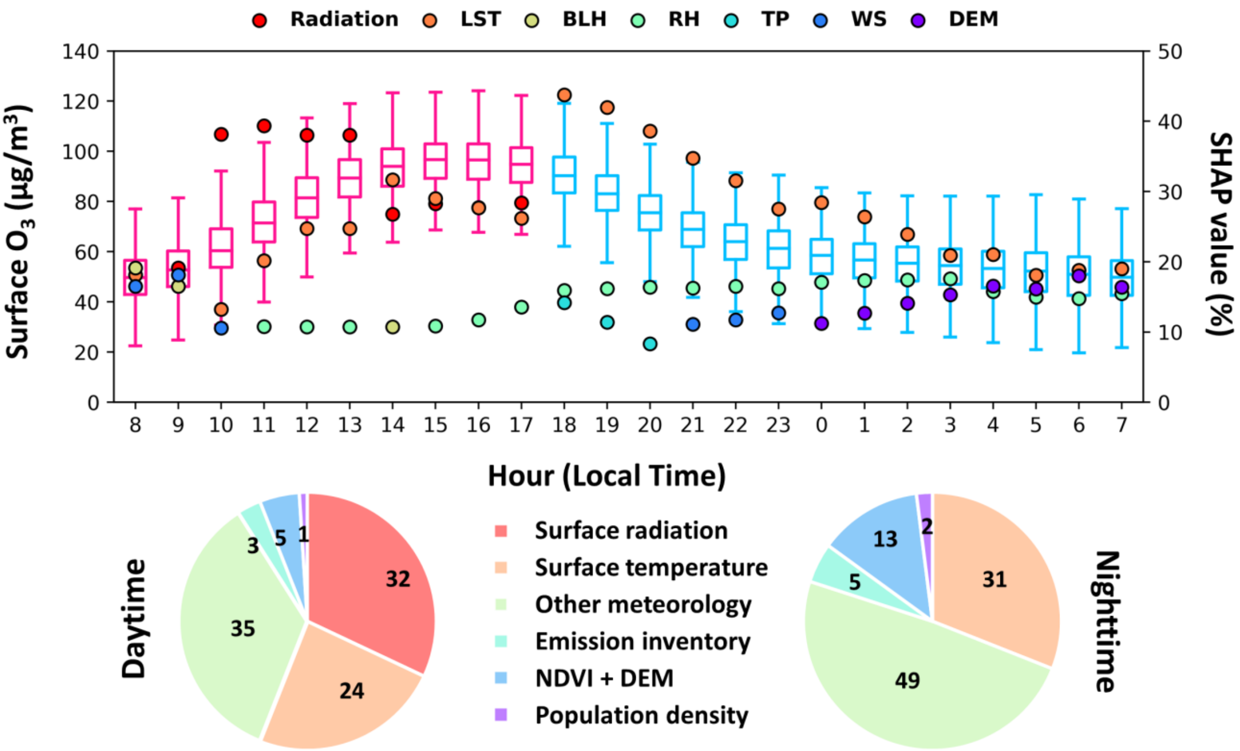

Figure 6. Time series of hourly surface O_3 variations (boxplots) and top-three driving factors (colored dots) throughout the day in 2018 in China. The two pie charts illustrate the contributions of driving-factor categories during the daytime (08:00–17:00 LT) and nighttime (18:00–07:00 LT), respectively. Surface radiation includes DSR and UV. Other meteorology includes BLH, RH, TP, SP, WD, and WS. Emission inventory includes NOx, VOC, and CO.

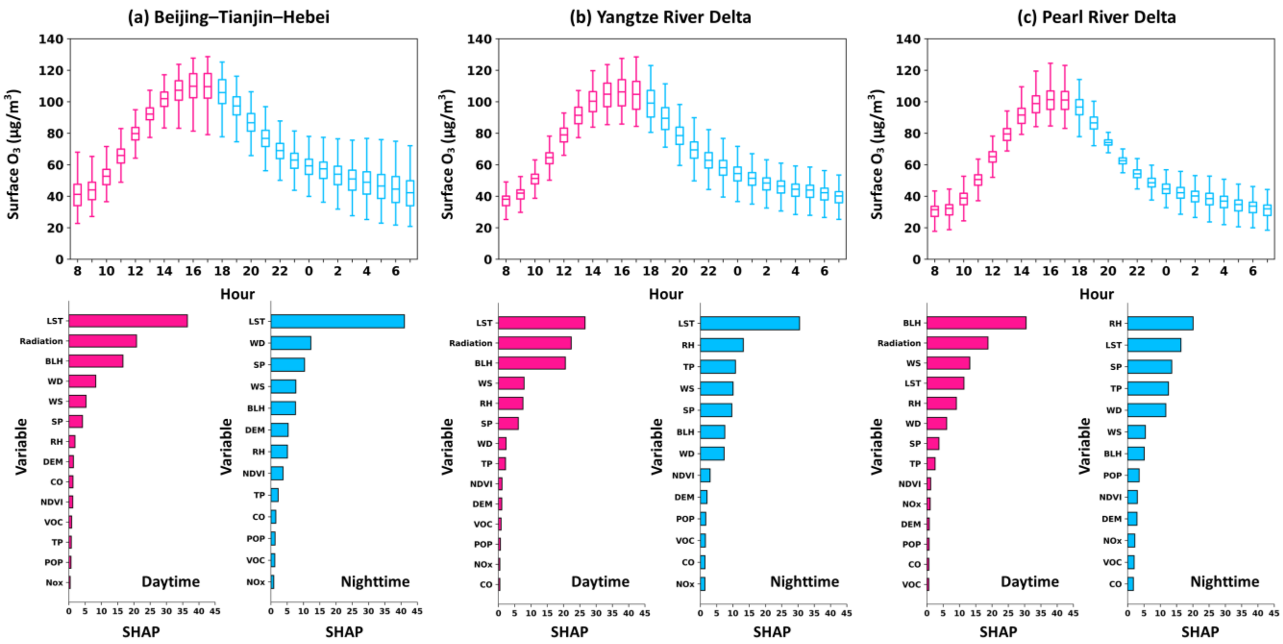

Figure 7. Box plots of diurnal surface O_3 concentrations (top row) and the sorted SHAP importance of each variable during the daytime (08:00–17:00 LT) and nighttime (18:00–07:00 LT) in 2018 for (a) the Beijing–Tianjin–Hebei region, (b) the Yangtze River Delta region, and (c) the Pearl River Delta region. In each box, the middle, lower, and upper horizontal black lines represent the mean bias, 25th percentile, and 75th percentile, respectively.

3.3 MDA8 O_3 levels and exposure risk

Using 24-hour data, we first calculate MDA8 O_3 concentrations across China and evaluate the population risk exposure to short-term O_3 pollution using World Health Organization (WHO) air quality standards updated in 2021 (WHO, 2021) (Figure 8). MDA8 O_3 concentrations mostly fall within the range of 79 to 109 μg/m³ (95th percentile), with a population-weighted (PWO_3) average of 96.8 μg/m³ in 2018 (Figure 8a). Serious pollution situations are mainly distributed in the North China Plain (especially in major parts of Shandong, Hebei, and Henan provinces: $PWO_3 > 120$ μg/m³) and north-central regions. By contrast, the remaining areas generally experience low levels, especially in northeast and southwest China ($PWO_3 < 80$ μg/m³). For the daily population risk of O_3 exposure, we found that ~44% (59%) of all (populated) areas in China encounter severe O_3 pollution, with at least one day surpassing the WHO's recommended short-term interim target 1 (i.e., IT1: MDA8 = 160 μg/m^3). However, the exposure risk is usually low (less than 10% of days) in most regions (Figure 8b). Regarding the short-term interim target 2 (i.e., IT2: MDA8 = 120 μg/m^3), areas exposed to a one-day risk expand significantly, reaching 85% in all areas and 93% in populated areas. The frequency also increases rapidly, with some eastern areas experiencing pollution for up to half of the year (Figure 8c). Most notably, when looking at the expected short-term air quality guidance (AQG) level (MDA8 = 100 μg/m^3), 100% of areas and the entire population are exposed to unhealthy air for at least one day, with a substantial risk intensity ranging from 20% to 70% across the domain (Figure 8d). These findings signify a serious risk of short-term O_3 exposure, underscoring the urgent requirement for environmental protection measures to control surface O_3 pollution, improve air quality, and promote future health benefits, especially in densely populated regions.

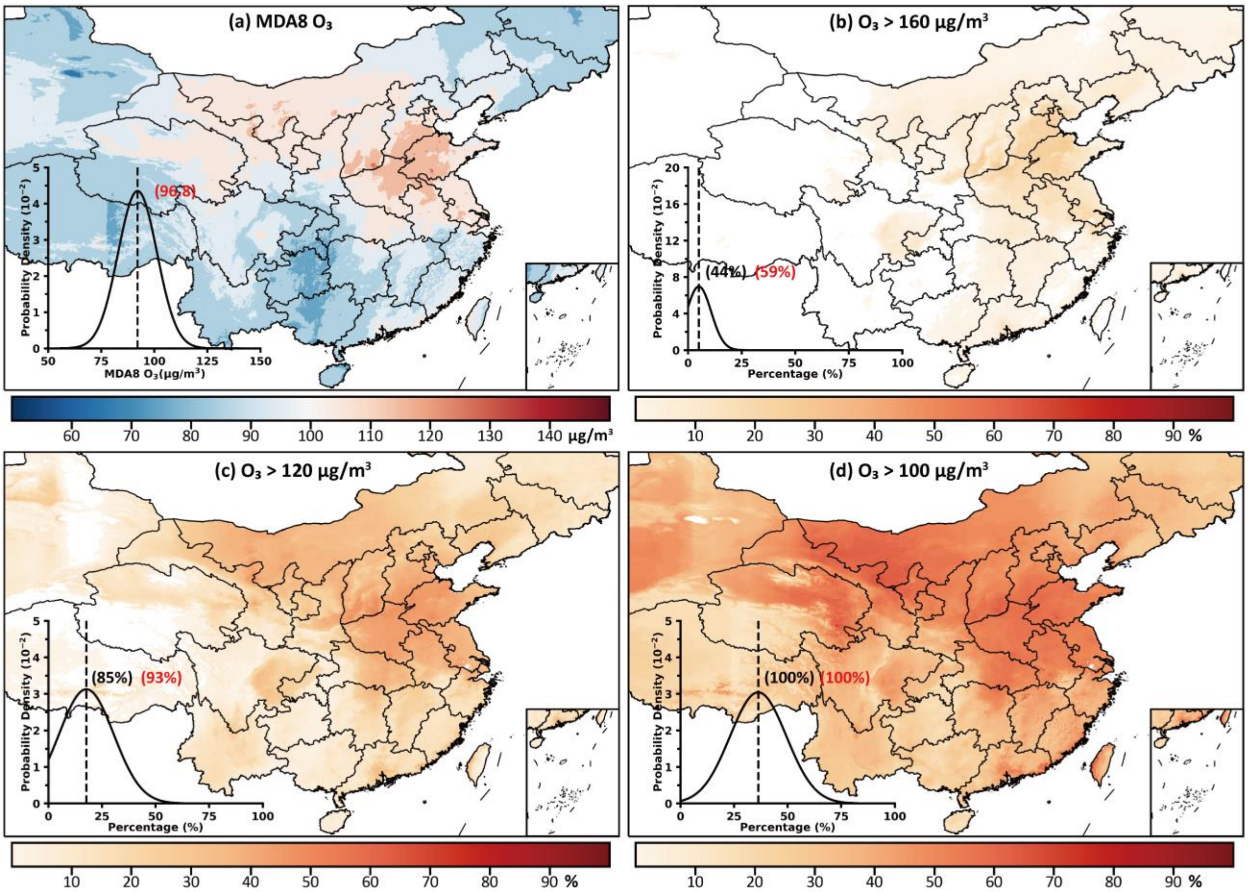

Figure 8. Spatial distribution of (a) MDA8 O_3 concentrations (μg/m³) and the percentage (%) of days exceeding the WHO-recommended short-term (b) interim target 1 (IT1: daily MDA8 O_3 = 160 μg/m^3), (c) interim target 2 (IT2: daily MDA8 O_3 = 120 μg/m^3), and (d) air quality guideline level (AQG: daily MDA8 O_3 = 100 μg/m^3) for 2018 in China. The inserted lower-left plots show probability density curves. The red number in (a) is the annual population-weighted average of MDA8 O_3 in China, and the black and red numbers in (b-d) indicate the percentages of pollution days for all and populated (population density > 0) regions, respectively.

3.4 Surface O_3 phytotoxicity indices and impacts

Figure 9 illustrates the spatial distribution of six main surface O_3 phytotoxicity indices calculated from 24-hour data in China for the year 2018. Specifically, M7 (mean 7-h O_3 concentration) and M12 (mean 12-h O_3 concentration) have similar spatial patterns, ranging between 30 to 43 ppb and 32 to 44 ppb (95th percentile), with an average of 36.8 and 37.3 ppb, respectively. Elevated values are predominantly concentrated in the western regions of Shandong province and scattered areas in northern China (Figure 9a-b). Conversely, most other areas maintain low levels, especially in western, southwest, and northeastern China (M7 and M12 < 35 ppb). There is a substantial disparity between AOT0 and AOT40, where the former ranges from 48 to 78 ppm (95th percentile), with an average of 62.6 ppm, and the latter is mostly within 27 ppm (average = 12.7 ppm). This distinction is attributed to AOT40 incorporating a threshold for hourly O_3 accumulation, while there is no criterion in AOT0. Nevertheless, extremely high AOT0 values are present in the North China Plain (particularly in Shandong and Tianjin), as well as in the western and central regions of Inner Mongolia and certain areas in northeastern China like Liaoning (Figure 9c). Similar spatial patterns are observed for AOT40, albeit with significantly lower levels. Note that ~85%, 91%, and 99% of vegetated areas in China exceed the defined critical levels of AOT40 at 6 ppm, 5 ppm, and 3 ppm, respectively. The spatial patterns of SUM06 (ranging from 0.12 to 45, average = 16.6 ppm) and W126 (ranging from 3 to 35, average = 12.9 ppm) are generally in close alignment with that of AOT40, but with higher levels in the North China Plain (Figure 9e-f). However, about 46%, 58%, and 65% of vegetated areas in China surpass the SUM06 critical levels of 15 ppm, 10 ppm, and 8 ppm, respectively. Furthermore, ~18% and 70% of vegetated areas in China are above the W126 critical levels of 23.8 ppm and 5.9 ppm, respectively. In general, the majority of vegetated areas in China experienced surface O_3 phytotoxicity, with the North China Plain being the most severely impacted region.

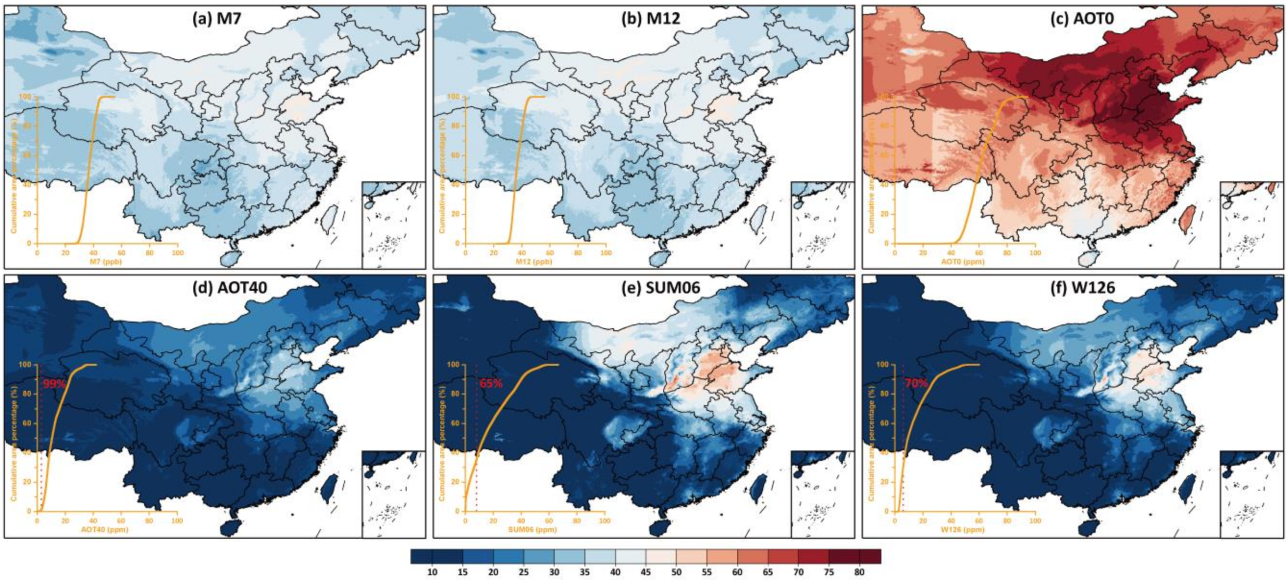

Figure 9. Spatial distributions of estimated (a) M7 (ppb), (b) M12 (ppb), (c) AOT0 (ppm), (d) AOT40 (ppm), (e) SUM06 (ppm), and (f) W126 (ppm) across China in 2018. The inserted lower-left plots show cumulative area percentages for vegetated areas in China. The red dotted lines in (d-f) show the lowest critical levels (i.e., 3, 8, and 5.9 ppm). The red numbers in (d-f) indicate cumulative percentages exceeding the specific critical level.

Additionally, we quantitatively investigated the influence of surface O_3 pollution on various types of vegetation (Figure 10), observing predominantly negative correlations between six O_3 phytotoxicity indices and four vegetation abundance indices. It is clear that vegetation growth and development are susceptible to exposure to surface O_3 pollution through phytotoxicity, with the extent of damage depending on the plant species, as implied by the varying strengths of the correlations. For croplands, AOT0, AOT40, and W126 are more associated with various vegetation abundance indices, particularly GPP, with Rs of -0.37 ($p < 0.001$), -0.31 ($p < 0.001$), and -0.26 ($p < 0.001$), respectively. Forests have stronger responses to O_3 pollution compared to croplands, showing heightened sensitivities, especially with AOT0 (R = -0.35–-0.53, $p < 0.001$), AOT40 (R = -0.27–-0.41, $p < 0.001$), and W126 (R = -0.25–-0.36, $p < 0.001$). For grasslands, the correlations between phytotoxicity indices and abundance indices are continuously strengthened, particularly with AOT0 (R = -0.47–-0.53, $p < 0.001$), M12 (R = -0.36–-0.42, $p < 0.001$), and AO40 (R = -0.31–-0.36, $p < 0.001$). Among all types of vegetation, AOT0 exhibits the most pronounced response to variations in vegetation growth, displaying the highest correlations with various abundance indices (R = -0.21–-0.53, $p < 0.001$), followed by M12 (R = -0.10–-0.42, $p < 0.001$), and AOT40 (R = -0.18–-0.41, $p < 0.001$). By contrast, M7 has a much weaker impact on vegetation, with the lowest correlations in nearly all cases, even exhibiting a positive correlation in croplands. In general, GPP has the strongest sensitivity to surface O_3 exposure, particularly in conjunction with AOT0, with the strongest correlation across all vegetated areas (R = -0.49, $p < 0.001$), as well as in croplands (R = -0.37, $p < 0.001$) and forests (R = -0.53, $p < 0.001$). This is further supported by its consistently strongest correlations with other O_3-exposure phytotoxicity indices. This can be attributed to ambient O_3's ability to enter leaves through stomata, causing damage to biological macromolecules and cell death (Kangasjarvi et al., 2005). This, in turn, reduces leaf stomatal conductance and photosynthetic rates (Ainsworth et al., 2012), ultimately leading to a decline in primary metabolism, leaf area, biomass, and a further reduction in GPP (Proietti et al., 2016; Jin et al., 2023).

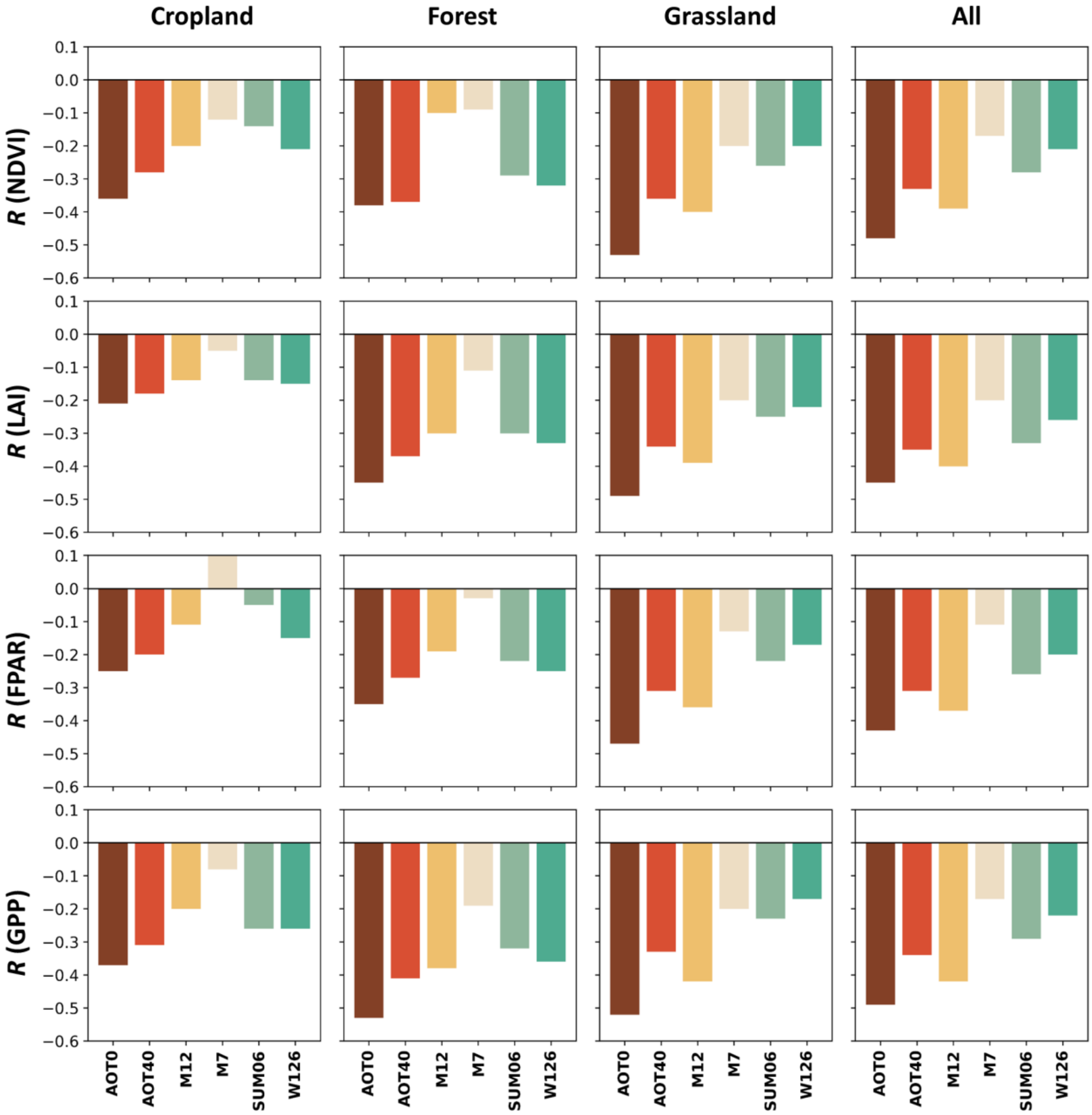

Figure 10. Correlation analysis between vegetation abundance indices (i.e., NDVI, LAI, FPAR, and GPP) and O_3-exposure phytotoxicity indices (i.e., AOT0, AOT40, M12, M7, SUM06, and W126) for various vegetated types, i.e., all vegetated areas (sample size, N = 9,461,375), forest (N = 1,827,010), grassland (N = 5,414,700), and cropland (N = 2,219,665) in China in 2018. All correlations are statistically significant at the 99.9% ($p < 0.001$) confidence level.

3.5 Comparison with related studies

Last, we compared our results with related research focusing on surface O_3 retrievals across China (Table 4). Most previous studies were concerned with the daily level, using MDA8 O_3 measurements calculated from hourly data as the baseline for model training (Liu et al., 2020; Xue et al., 2020; Song et al., 2022; Wei et al., 2022). In fact, MDA8 is not a straightforward multi-hour average but an iterative daily maximum 8-hour average, which can lead to substantial deviations in modeling interpolated results, particularly in remote areas lacking measurements. Accurate MDA8 calculations require 24 hours of retrievals, but few studies have addressed this issue (Y. Wang et al., 2022; Xue et al., 2022; B. Chen et al., 2023; Zhang et al., 2023). All of these studies have also exclusively focused on retrieving daytime surface O_3 concentrations, with durations (< 10 hours) falling significantly short of the requirements to calculate air quality and O_3-exposure phytotoxicity indices. Nevertheless, our model exhibits comparable or superior overall accuracy compared to the performance of AI models in previous studies conducted during the same daytime hours. In addition, their retrieved hourly surface O_3 concentrations often have sparse spatial resolutions (2–9 km), with severe spatial discontinuities due to large gaps of missing values in critical input satellite optical variables (e.g., Himawari-8 top-of-the-atmosphere radiation and brightness temperatures) caused by cloud contamination (Y. Wang et al.; 2022, B. Chen et al., 2023). These limitations significantly constrain their applicability in small-scale areas such as urban environments. Our study presents substantial improvements on all the abovementioned key aspects by first offering a spatially (100% coverage) continuous dataset of 24-hour surface O_3 concentrations across China, encompassing the full temporal range (0:00–23:00 LT) at a high resolution of 1 km.

Table 4. Model performance comparison in estimating hourly surface O_3 concentrations in China from previous studies.

Model	Duration	Spatial resolution	Overall accuracy			Missing values	Literature
			CV-R^2	RMSE	MAE		
BT	Daytime (09:00-16:00 LT)	5 km	0.87	18.30	13.30	No	Zhang et al., 2023
DF	Daytime (10:00-15:00 LT)	5 km	0.91	12.74	8.25	Yes	B. Chen et al. 2023
SGLboost	Daytime (09:00-18:00 LT)	2 km	0.85	19.04	-	Yes	Y. Wang et al., 2022
4D-STET	Daytime (09:00-16:00 LT)	1 km	0.91	15.00	10.45	No	This study
	Daytime (10:00-15:00 LT)		0.90	15.15	10.56		
	Daytime (09:00-18:00 LT)		0.91	15.19	10.60		
	Daytime (08:00-17:00 LT)		0.91	14.91	10.38		
	Nighttime (18:00-07:00 LT)		0.85	16.31	11.62		
	All day (00:00-23:00 LT)		0.89	15.74	11.10		

BT: bagged-tree; DF: deep forest; SGLboost: self-adaptive geospatially local categorical boosting; 4D-STET: 4-dimensional space-time extra-trees.

4. Summary and conclusions

Surface O_3 is a critical atmospheric pollutant gas influencing air quality, posing a major human health risk, as well as plant well-being risk. To overcome limitations encountered in previous studies, e.g., low temporal resolution (mostly daily, with only a few hourly observations during the daytime), sparse spatial resolution, and substantial spatial gaps in data retrievals, we refined a total of 15 AI models by introducing multidimensional spatiotemporal information to enhance their capabilities. The best-performing model (i.e., the 4D-STET model) was then selected to derive for the first time gapless surface O_3 concentrations in China at 24-hour temporal and 1-km spatial resolutions from the GEO DSR and LST products and many other ancillary data. Cross-validations demonstrate the robustness of our model in capturing the diurnal variations of surface O_3 concentrations, with an overall sample-based (station-based) CV-R^2 of 0.89 (0.86), 0.91 (0.90), and 0.85 (0.81), and RMSE values of 15.74 (17.39) μg/m^3, 14.91 (16.31) μg/m^3, and 16.31 (18.14) μg/m^3 during all times (00:00–23:00 LT), daytime (08:00–17:00 LT), and nighttime (18:00–07:00 LT), respectively. The availability of temporally continuous surface O_3 data facilitates our capacity to analyze diurnal variations, daily exposure risks, and phytotoxicity impacts at different spatiotemporal scales throughout China.

Surface O_3 levels showed strong diurnal variations, steadily rising from sunrise, peaking around 15:00 LT, and continuously decreasing thereafter. The XAI-SHAP analysis results revealed that shortwave and UV radiation, along with LST, explain about 56% of the surface O_3 variations during the daytime, while LST plays the most significant role during the nighttime, contributing approximately 31%. In 2018, approximately 59% (44%), 93% (85%), and 100% (100%) of populated areas (entire areas) faced short-term surface O_3 exposure risk for at least one day, with MDA8 O_3 surpassing the WHO's air quality standards of 160 μg/m^3, 120 μg/m^3, and 100 μg/m^3, respectively. Furthermore, ~99%, 91%, and 85% of vegetated areas in China exceeded the critical levels of AOT40 at 3 ppm, 5 ppm, and 6 ppm, respectively. For SUM06, ~65%, 58%, and 46% of vegetated areas surpassed the critical levels of 8 ppm, 10 ppm, and 15 ppm, respectively. As for W126, ~70% and 18% of vegetated areas exceeded the critical levels of W126 at 5.9 ppm and 23.8 ppm, respectively. These findings highlight the urgent need for environmental protection measures to mitigate surface O_3 pollution and promote the health of both the public and vegetation in the future. Furthermore, despite the consistent negative correlations, GPP demonstrates the strongest response to surface O_3 pollution among all vegetation (ozone-exposure) phytotoxicity indices, encompassing various vegetated types, especially when combined with AOT0 (R = -0.21–-0.53, $p < 0.001$). In a future study, we intend to apply our methodology to generate a long-term hourly surface O_3 dataset and provide more detailed insights into air quality and phytotoxic damage caused by surface O_3 pollution.

Data availability

CNEMC O_3 measurements are available at http://www.cnemc.cn. The Himawari-8 DSR product is available at https://zenodo.org/record/7023863. The GHA Land Surface Temperature product is available at http://glass.umd.edu/allsky_LST/GHA-LST/2018/. The ERA5 reanalysis is available at https://cds.climate.copernicus.eu/. The ABaCAS-EI v2.0 O_3 Precursor data is available at https://doi.org/10.6084/m9.figshare.21777005.v1. The SRTM DEM is available at https://www2.jpl.nasa.gov/srtm/. LandScan^{TM} population information is available at https://landscan.ornl.gov/. NDVI data is available at https://lpdaac.usgs.gov/products/mod13a2v061/. GPP data is available at https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD17A2H. LAI data is available at https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD15A2H. The MODIS Land Cover Type product is available at https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD12Q1.

Data Sharing

The generated 24-hour 1-km surface O_3 datasets and codes can be found at https://doi.org/10.5281/zenodo.10035857. They will be made publicly available once the paper is accepted.

Author contributions

ZL and JW designed the study. FC performed the research and wrote the initial draft of this paper. ZL, JW, and KL reviewed and edited the paper. ZY, KL, and WX assisted in processing the surface

O_3 ground measurements and relevant data. RL, DW, AJ, BZ, SW, DY, and SL provided and processed the GEO satellite products and emission inventory data. MC copyedited the article. All authors made substantial contributions to this work.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (42030606 and 42207541), as well as the Samsung Advanced Institute of Technology.

References

Ainsworth, E. A., Yendrek, C. R., Sitch, S., Collins, W. J., Emberson, L. D (2012). The effects of tropospheric ozone on net primary productivity and implications for climate change. Annu. Rev. Plant Biol. 63, 637–661. https://doi.org/10.1146/annurev-arplant-042110-103829

Allen, R.J., Sherwood, S.C., Norris, J.R., Zender, C.S., 2012. Recent northern hemisphere tropical expansion primarily driven by black carbon and tropospheric ozone. Nature 485, 350–354. https://doi.org/10.1038/nature11097

Barnard, W. F., Saxena, V. K., Wenny, B. N., DeLuisi, J. J., 2003. Daily surface UV exposure and its relationship to surface pollutant measurements. J. Air Waste Manag. Assoc. 53, 237–245. https://doi.org/10.1080/10473289.2003.10466134

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. https://doi.org/10.1029/2009GL037308

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R. V., Dentener, F., Dingenen, R. van, Estep, K., Amini, H., Apte, J. S., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P. K., Knibbs, L. D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J. L. T., Shaddick, G., Anderson, H. R., Vos, T., Forouzanfar, M. H., Burnett, R. T., Cohen, A., 2016. Ambient air pollution exposure estimation for the Global Burden of Disease 2013. Environ. Sci. Technol. 50, 79–88. https://doi.org/10.1021/acs.est.5b03709

Breiman, L., 2001. Random forests. Mach. Learn 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cai, M., Wei, J., Zhang, S., Liu, W., Wang, L., Qian, Z., Lin, H., Liu, E., McMillin, S. E., Cao, Y., Yin, P., 2023. Short-term air pollution exposure associated with death from kidney diseases: a nationwide time-stratified case-crossover study in China from 2015 to 2019. BMC Med. 21, 32. https://doi.org/10.1186/s12916-023-02734-9

Capilla, C., 2016. Prediction of hourly ozone concentrations with multiple regression and multilayer perceptron models. Int. J. SDP 11, 558–565. https://doi.org/10.2495/SDP-V11-N4-558-565

Chen, B., Wang, Y., Huang, J., Zhao, L., Chen, R., Song, Z., Hu, J., 2023. Estimation of near-surface ozone concentration and analysis of main weather situation in China based on machine learning model and Himawari-8 TOAR data. Sci. Total Environ. 864, 160928. https://doi.org/10.1016/j.scitotenv.2022.160928

Chen, C., Li, T., Sun, Q., Shi, W., He, M. Z., Wang, J., Liu, J., Zhang, M., Jiang, Q., Wang, M., Shi, X., 2023. Short-term exposure to ozone and cause-specific mortality risks and thresholds in China: Evidence from nationally representative data, 2013-2018. Environ. Int. 171, 107666. https://doi.org/10.1016/j.envint.2022.107666

Chen, J., Shen, H., Li, X., Li, T., Wei, Y., 2022. Ground-level ozone estimation based on geo-intelligent machine learning by fusing in-situ observations, remote sensing data, and model simulation data. Int. J. Appl. Earth Obs. Geoinf. 112, 102955. https://doi.org/10.1016/j.jag.2022.102955

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, arXiv:1603.02754 [cs], 785–794. https://doi.org/10.1145/2939672.2939785

Chen, Y., Zhou, Y., NixiaCiren, Zhang, H., Wang, C., GesangDeji, and Wang, X., 2022. Spatiotemporal variations of surface ozone and its influencing factors across Tibet: a

Geodetector-based study. Sci. Total Environ. 813, 152651. https://doi.org/10.1016/j.scitotenv.2021.152651
David, L. M., Nair, P. R., 2011. Diurnal and seasonal variability of surface ozone and NO_x at a tropical coastal site: association with mesoscale and synoptic meteorological conditions. J. Geophys. Res. 116, D10303. https://doi.org/10.1029/2010JD015076
Di, Q., Rowland, S., Koutrakis, P., Schwartz, J., 2017. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States (vol 67, pg 39, 2016). J. Air Waste Manage. Assoc. 67, 389–390.
Du, K., Leung, C., Mow, W., Swamy, M., 2022. Perceptron: learning, generalization, model selection, fault tolerance, and role in the deep learning era. MATHEMATICS 10. https://doi.org/10.3390/math10244730
Fares, S., Vargas, R., Detto, M., Goldstein, A. H., Karlik, J., Paoletti, E., Vitale, M., 2013. Tropospheric ozone reduces carbon assimilation in trees: estimates from analysis of continuous flux measurements. Global Change Biol. 19, 2427–2443. https://doi.org/10.1111/gcb.12222
Freund, Y., Schapire, R. E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139. https://doi.org/10.1006/jcss.1997.1504
Friedman, J. H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Statist. 29. https://doi.org/10.1214/aos/1013203451
Fu, T.-M., Tian, H., 2019. Climate change penalty to Ozone air quality: Review of current understandings and knowledge gaps. Curr Pollut Rep 5, 159–171. https://doi.org/10.1007/s40726-019-00115-6
Fuhrer, J., Skärby, L., Ashmore, M. R., 1997. Critical levels for ozone effects on vegetation in Europe. Environ. Pollut. 97, 91–106. https://doi.org/10.1016/S0269-7491(97)00067-5
Gagliardi, R. V., Andenna, C., 2020. A machine learning approach to investigate the surface ozone behavior. Atmosphere-basel 11, 1173. https://doi.org/10.3390/atmos11111173
Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63, 3–42. https://doi.org/10.1007/s10994-006-6226-1
Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on typical tabular data? Adv. Neural Inf. Process. Syst. 35, 507–520.
Grünhage, L., Jäger, H.-J., Haenel, H.-D., Löpmeier, F.-J., Hanewald, K., 1999. The European critical levels for ozone: improving their usage. Environ. Pollut. 105, 163–173. https://doi.org/10.1016/S0269-7491(99)00029-9
Han, H., Liu, J., Shu, L., Wang, T., Yuan, H., 2020. Local and synoptic meteorological influences on daily variability in summertime surface ozone in eastern China. Atmos. Chem. Phys. 20, 203–222. https://doi.org/10.5194/acp-20-203-2020
Han, S., Bian, H., Feng, Y., Liu, A., Li, X., Zeng, F., Zhang, X., 2011. Analysis of the relationship between O_3, NO, and NO_2 in Tianjin, China. Aerosol Air Qual. Res. 11, 128–139. https://doi.org/10.4209/aaqr.2010.07.0055
Hayes, F., Bangor, C., 2017. Chapter III. Mapping critical levels for vegetation, Modelling and mapping manual of the LRTAP Convention. https://unece.org/fileadmin/DAM/env/documents/2017/AIR/EMEP/Final__new_Chapter_3_v2__August_2017_.pdf
Hayes, F., Bangor, C., 2018. Supplement of Chapter III, Modelling and mapping manual of the

LRTAP Convention. https://icpvegetation.ceh.ac.uk/sites/default/files/ScientificBackgroundDocumentAOct2018.pdf

K. He, X. Zhang, S. Ren, J. Sun, 2016. Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. https://doi.org/10.1109/CVPR.2016.90

Heck, W. W., Cowling, E. B., 1997. The need for a long term cumulative secondary ozone standard - an ecological perspective. EM, January, 23–33.

Hinton, G. E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comp. 18, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Hou, T., Yu, S., Jiang, Y., Chen, X., Zhang, Y., Li, M., Li, Z., Song, Z., Li, P., Chen, J., Zhang, X., 2022. Impacts of chemical initial conditions in the WRF-CMAQ model on the ozone forecasts in eastern China. Aerosol Air Qual. Res. 22, 210402. https://doi.org/10.4209/aaqr.210402

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, 10,333–10,350. https://doi.org/10.5194/acp-16-10333-2016

Hu, X.-M., Doughty, D. C., Sanchez, K. J., Joseph, E., Fuentes, J. D., 2012. Ozone variability in the atmospheric boundary layer in Maryland and its implications for vertical transport model. Atmos. Environ. 46, 354–364. https://doi.org/10.1016/j.atmosenv.2011.09.054

Huang, X.-G., Zhao, J.-B., Cao, J.-J., Song, Y.-Y., 2019. Spatial-temporal variation of ozone concentration and its driving factors in China. Huan Jing Ke Xue 40, 1120–1131. https://doi.org/10.13227/j.hjkx.201807038

Jia, A., Liang, S., Wang, D., Ma, L., Wang, Z., Xu, S., 2023. Global hourly, 5 km, all-sky land surface temperature data from 2011 to 2021 based on integrating geostationary and polar-orbiting satellite data. Earth Syst. Sci. Data 15, 869–895. https://doi.org/10.5194/essd-15-869-2023

Jiang, Y. C., Zhao, T. L., Liu, J., Xu, X. D., Tan, C. H., Cheng, X. H., Bi, X. Y., Gan, J. B., You, J. F., Zhao, S. Z., 2015. Why does surface ozone peak before a typhoon landing in southeast China? Atmos. Chem. Phys. 15, 13,331–13,338. https://doi.org/10.5194/acp-15-13331-2015

Jin, Z., Yan, D., Zhang, Z., Li, M., Wang, T., Huang, X., Xie, M., Li, S., Zhuang, B., 2023. Effects of elevated ozone exposure on regional meteorology and air quality in China through ozone□vegetation coupling. J. Geophys. Res. Atmos. 128, e2022JD038119. https://doi.org/10.1029/2022JD038119

Kang, Y., Choi, H., Im, J., Park, S., Shin, M., Song, C.-K., Kim, S., 2021. Estimation of surface-level NO_2 and O_3 concentrations using TROPOMI data and machine learning over East Asia. Environ. Pollut. 288, 117711. https://doi.org/10.1016/j.envpol.2021.117711

Kangasjarvi, J., Jaspers, P., Kollist, H., 2005. Signalling and cell death in ozone-exposed plants. Plant Cell Environ. 28, 1021–1036. https://doi.org/10.1111/j.1365-3040.2005.01325.x

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: a highly efficient gradient boosting decision tree, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 3149–3157.

Kohut, R., 2007. Assessing the risk of foliar injury from ozone on vegetation in parks in the U.S. National Park Service’s Vital Signs Network. Environ. Poll. 149, 348–357. https://doi.org/10.1016/j.envpol.2007.04.022

Lefohn, A. S., Runeckles, V. C., 1987. Establishing standards to protect vegetation—ozone exposure/dose considerations. Atmos. Environ. (1967) 21, 561–568. https://doi.org/10.1016/0004-6981(87)90038-2

Li, R., Wang, D., Wang, W., Nemani, R., 2023. A GeoNEX-based high-spatiotemporal-resolution product of land surface downward shortwave radiation and photosynthetically active radiation. Earth Syst. Sci. Data 15, 1419–1436. https://doi.org/10.5194/essd-15-1419-2023

Li, S., Wang, S., Wu, Q., Zhang, Y., Ouyang, D., Zheng, H., Han, L., Qiu, X., Wen, Y., Liu, M., Jiang, Y., Yin, D., Liu, K., Zhao, B., Zhang, S., Wu, Y., Hao, J., 2023. Emission trends of air pollutants and CO□ in China from 2005 to 2021. Earth Syst. Sci. Data Disc. 2023, 1–22. https://doi.org/10.5194/essd-2022-464

Li, T., Shen, H., Yuan, Q., Zhang, X., 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, 11,985-11,993. https://doi.org/10.1002/2017GL075710

Liao, Z., Pan, Y., Ma, P., Jia, X., Cheng, Z., Wang, Q., Dou, Y., Zhao, X., Zhang, J., Quan, J., 2023. Meteorological and chemical controls on surface ozone diurnal variability in Beijing: a clustering-based perspective. Atmos. Environ. 295, 119566. https://doi.org/10.1016/j.atmosenv.2022.119566

Lin, Y., Jiang, F., Zhao, J., Zhu, G., He, X., Ma, X., Li, S., Sabel, C. E., Wang, H., 2018. Impacts of O_3 on premature mortality and crop yield loss across China. Atmos. Environ. 194, 41–47. https://doi.org/10.1016/j.atmosenv.2018.09.024

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, X., Yi, G., Zhou, X., Zhang, T., Bie, X., Li, J., Tan, H., 2023. Spatio-temporal variations of $PM_{2.5}$ and O_3 in China during 2013–2021: impact factor analysis. Environ. Pollut. 334, 122189. https://doi.org/10.1016/j.envpol.2023.122189

Lu, X., Zhang, L., Zhao, Y., Jacob, D.J., Hu, Y., Hu, L., Gao, M., Liu, X., Petropavlovskikh, I., McClure-Begley, A., Querel, R., 2019. Surface and tropospheric ozone trends in the Southern Hemisphere since 1990: possible linkages to poleward expansion of the Hadley circulation. Science Bulletin 64, 400–409. https://doi.org/10.1016/j.scib.2018.12.021

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9

Ma, R., Ban, J., Wang, Q., Zhang, Y., Yang, Y., He, M. Z., Li, S., Shi, W., Li, T., 2021. Random forest model based fine scale spatiotemporal O_3 trends in the Beijing-Tianjin-Hebei region in China, 2010 to 2017. Environ. Pollut. 276. https://doi.org/10.1016/j.envpol.2021.116635

Ministry of Ecology and Environment (MEE), 2018. Revision of the Ambien air quality standards (GB 3095–2012) (in Chinese): https://www.mee.gov.cn/gkml/sthjbgw/sthjbgg/201808/t20180815_451398.htm, last access: 22 October 2023.

Morris, G. A., Ford, B., Rappenglück, B., Thompson, A. M., Mefferd, A., Ngan, F., Lefer, B., 2010. An evaluation of the interaction of morning residual layer and afternoon mixed layer ozone in Houston using ozonesonde data. Atmos. Environ. 44, 4024–4034.

https://doi.org/10.1016/j.atmosenv.2009.06.057
Niu, Y., Zhou, Y., Chen, R., Yin, P., Meng, X., Wang, W., Liu, C., Ji, J. S., Qiu, Y., Kan, H., Zhou, M., 2022. Long-term exposure to ozone and cardiovascular mortality in China: a nationwide cohort study. Lancet Planet. Health 6, e496–e503. https://doi.org/10.1016/S2542-5196(22)00093-6
Otero, N., Sillmann, J., Mar, K. A., Rust, H. W., Solberg, S., Andersson, C., Engardt, M., Bergström, R., Bessagnet, B., Colette, A., Couvidat, F., Cuvelier, C., Tsyro, S., Fagerli, H., Schaap, M., Manders, A., Mircea, M., Briganti, G., Cappelletti, A., Adani, M., D'Isidoro, M., Pay, M.-T., Theobald, M., Vivanco, M. G., Wind, P., Ojha, N., Raffort, V., Butler, T., 2018. A multi-model comparison of meteorological drivers of surface ozone over Europe. Atmos. Chem. Phys. 18, 12,269–12,288. https://doi.org/10.5194/acp-18-12269-2018
Proietti, C., Anav, A., De Marco, A., Sicard, P., Vitale, M., 2016. A multi-sites analysis on the ozone effects on gross primary production of European forests. Sci. Total Environ. 556, 1–11. https://doi.org/10.1016/j.scitotenv.2016.02.187
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
Rodriguez, J. D., Perez, A., Lozano, J. A., 2010. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32, 569–575. https://doi.org/10.1109/TPAMI.2009.187
Sagi, O., Rokach, L., 2018. Ensemble learning: a survey. WIREs Data Mining Knowl. Discov. 8. https://doi.org/10.1002/widm.1249
Seinfeld, J. H., Pandis, S. N., 2016. Atmospheric chemistry and physics: from air pollution to climate change, John Wiley & Sons.
Shinde, P. P., Shah, S., 2018. A review of machine learning and deep learning applications, in: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 1–6. https://doi.org/10.1109/ICCUBEA.2018.8697857
Song, G., Li, S., Xing, J., Yang, J., Dong, L., Lin, H., Teng, M., Hu, S., Qin, Y., Zeng, X., 2022. Surface UV-assisted retrieval of spatially continuous surface ozone with high spatial transferability. Remote Sens. Environ. 274, 112996. https://doi.org/10.1016/j.rse.2022.112996
Stevenson, D.S., Young, P.J., Naik, V., Lamarque, J.-F., Shindell, D.T., Voulgarakis, A., Skeie, R.B., Dalsoren, S.B., Myhre, G., Berntsen, T.K., Folberth, G.A., Rumbold, S.T., Collins, W.J., MacKenzie, I.A., Doherty, R.M., Zeng, G., van Noije, T.P.C., Strunk, A., Bergmann, D., Cameron-Smith, P., Plummer, D.A., Strode, S.A., Horowitz, L., Lee, Y.H., Szopa, S., Sudo, K., Nagashima, T., Josse, B., Cionni, I., Righi, M., Eyring, V., Conley, A., Bowman, K.W., Wild, O., Archibald, A., 2013. Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). Atmos Chem Phys 13, 3063–3085. https://doi.org/10.5194/acp-13-3063-2013
Super, I., Vilà-Guerau de Arellano, J., Krol, M. C., 2015. Cumulative ozone effect on canopy stomatal resistance and the impact on boundary layer dynamics and CO_2 assimilation at the diurnal scale: a case study for grassland in the Netherlands. J. Geophys. Res. Biogeo. 120, 1348–1365. https://doi.org/10.1002/2015JG002996

Tong, D. Q., Mathur, R., Kang, D., Yu, S., Schere, K. L., Pouliot, G., 2009. Vegetation exposure to ozone over the continental United States: assessment of exposure indices by the Eta-CMAQ air quality forecast model. Atmos. Environ. 43, 724–733. https://doi.org/10.1016/j.atmosenv.2008.09.084

Tu, J., Xia, Z.-G., Wang, H., Li, W., 2007. Temporal variations in surface ozone and its precursors and meteorological effects at an urban site in China. Atmos. Res. 85, 310–337. https://doi.org/10.1016/j.atmosres.2007.02.003

Vautard, R., Moran, M. D., Solazzo, E., Gilliam, R. C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A. B., Jericevic, A., Prank, M., Segers, A., Silver, J. D., Werhahn, J., Wolke, R., Rao, S. T., Galmarini, S., 2012. Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations. Atmos. Environ. 53, 15–37. https://doi.org/10.1016/j.atmosenv.2011.10.065

Wang, H., Ke, Y., Tan, Y., Zhu, B., Zhao, T., Yin, Y., 2023. Observational evidence for the dual roles of BC in the megacity of eastern China: enhanced O_3 and decreased $PM_{2.5}$ pollution. Chemosphere 327, 138548. https://doi.org/10.1016/j.chemosphere.2023.138548

Wang, N., Guo, H., Jiang, F., Ling, Z. H., 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. https://doi.org/10.1016/j.scitotenv.2014.10.070

Wang, Y., Gao, W., Wang, S., Song, T., Gong, Z., Ji, D., Wang, L., Liu, Z., Tang, G., Huo, Y., Tian, S., Li, J., Li, M., Yang, Y., Chu, B., Petäjä, T., Kerminen, V.-M., He, H., Hao, J., Kulmala, M., Wang, Y., Zhang, Y., 2020. Contrasting trends of $PM_{2.5}$ and surface-ozone concentrations in China from 2013 to 2017. Natl. Sci. Rev. 7, 1331–1339. https://doi.org/10.1093/nsr/nwaa032

Wang, Y., Yuan, Q., Li, T., Zhu, L., Zhang, L., 2021. Estimating daily full-coverage near surface O_3, CO, and NO_2 concentrations at a high spatial resolution over China based on S5P-TROPOMI and GEOS-FP. ISPRS-J. Photogramm. Remote Sens. 175, 311–325. https://doi.org/10.1016/j.isprsjprs.2021.03.018

Wang, Y., Yuan, Q., Zhu, L., Zhang, L., 2022. Spatiotemporal estimation of hourly 2-km ground-level ozone over China based on Himawari-8 using a self-adaptive geospatially local model. Geosci. Front. 13, 101286. https://doi.org/10.1016/j.gsf.2021.101286

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., Cribb, M., 2019. Estimating 1-km-resolution $PM_{2.5}$ concentrations across China using the space-time random forest approach. Remote Sens. Environ. 231, 111221. https://doi.org/10.1016/j.rse.2019.111221

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, 3273–3289. https://doi.org/10.5194/acp-20-3273-2020

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M., 2021. 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., Li, K., Dickerson, R. R., Pinker, R. T., Wang, J., Liu, X., Sun, L., Xue, W., Cribb, M., 2022. Full-coverage mapping and spatiotemporal variations of ground-level ozone (O_3) pollution from 2013 to 2020 across China. Remote Sens. Environ. 270, 112775. https://doi.org/10.1016/j.rse.2021.112775

Wei, J., Li, Z., Chen, X., Li, C., Sun, Y., Wang, J., Lyapustin, A., Brasseur, G. P., Jiang, M., Sun, L., Wang, T., Jung, C. H., Qiu, B., Fang, C., Liu, X., Hao, J., Wang, Y., Zhan, M., Song, X., Liu, Y., 2023. Separating daily 1-km $PM_{2.5}$ inorganic chemical composition in China since 2000 via deep learning integrating ground, satellite, and model data. Environ. Sci. Technol. acs.est.3c00272, https://doi.org/10.1021/acs.est.3c00272

WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO

Wilkinson, S., Mills, G., Illidge, R., Davies, W. J., 2012. How is ozone pollution reducing our food supply? J. Experiment. Bot. 63, 527–536. https://doi.org/10.1093/jxb/err317

Xia, Y., Hu, Y., Huang, Y., Bian, J., Zhao, C., Wei, J., Yan, Y., Xie, F., Lin, J., 2022. Concurrent hot extremes and high ultraviolet radiation in summer over the Yangtze Plain and their possible impact on surface ozone. Environ. Res. Lett. 17, 064001. https://doi.org/10.1088/1748-9326/ac6c3c

Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K., 2017. Aggregated residual transformations for deep neural networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5987–5995. https://doi.org/10.1109/CVPR.2017.634

Xu, J., Xu, X., Lin, W., Ma, Z., Ma, J., Wang, R., Wang, Y., Zhang, G., Xu, W., 2020. Understanding the formation of high-ozone episodes at Raoyang, a rural site in the North China Plain. Atmos. Environ. 240, 117797. https://doi.org/10.1016/j.atmosenv.2020.117797

Xue, T., Zheng, Y., Geng, G., Xiao, Q., Meng, X., Wang, M., Li, X., Wu, N., Zhang, Q., Zhu, T., 2020. Estimating spatiotemporal variation in ambient ozone exposure during 2013–2017 using a data-fusion model. Environ. Sci. Technol. 54, 14,877–14,888. https://doi.org/10.1021/acs.est.0c03098

Xue, W., Zhang, J., Hu, X., Yang, Z., Wei, J., 2022. Hourly seamless surface O_3 estimates by integrating the chemical transport and machine learning models in the Beijing-Tianjin-Hebei region. Int. J. Environ. Res. Public Health 19, 8511. https://doi.org/10.3390/ijerph19148511

Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K., 2018. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629. https://doi.org/10.1007/s13244-018-0639-9

Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation 31, 1235–1270. https://doi.org/10.1162/neco_a_01199

Yue, X. Unger, N., 2014. Ozone vegetation damage effects on gross primary productivity in the United States. Atmos. Chem. Phys. 14, 9137–9153. https://doi.org/10.5194/acp-14-9137-2014

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, T., 2022. Risk of illness-related school absenteeism for elementary students with exposure to $PM_{2.5}$ and O_3. Sci. Total Environ. 842, 156824. .

Zhang, R., Lei, W., Tie, X., and Hess, P., 2004. Industrial emissions cause extreme urban ozone diurnal variability. Proc. Natl. Acad. Sci. U.S.A. 101, 6346–6350. https://doi.org/10.1073/pnas.0401484101

Zhang, Y., Wang, W., He, J., Jin, Z., Wang, N., 2023. Spatially continuous mapping of hourly ground ozone levels assisted by Himawari-8 shortwave radiation products. GIScience Remote Sens. 60.

Zhao, W., Fan, S., Guo, H., Gao, B., Sun, J., Chen, L., 2016. Assessing the impact of local meteorological variables on surface ozone in Hong Kong during 2000–2015 using quantile and multiple line regression models. Atmos. Environ. 144, 182–193. https://doi.org/10.1016/j.atmosenv.2016.08.077

Zhou, Z.-H., Feng, J., 2019. Deep forest. Natl. Sci. Rev. 6, 74–86. https://doi.org/10.1093/nsr/nwy108

Zhu, J., Tai, A. P. K., Hung Lam Yim, S., 2022. Effects of ozone–vegetation interactions on meteorology and air quality in China using a two-way coupled land–atmosphere model. Atmos. Chem. Phys. 22, 765–782 https://doi.org/10.5194/acp-22-765-2022