https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Open Access Earth System Science Data Discussions

Recurrent mapping of Hourly Surface Ozone Data (HrSOD) across China during 2005–2020 for ecosystem and human health risk assessment

Wenxiu Zhang[1,a], Di Liu[1,a], Hanqin Tian[2], Naiqin Pan[2,3], Ruqi Yang[4], Wenhan Tang[5], Jia Yang[6], Fei Lu[1], Buddhi Dayananda[7], Han Mei[8], Siyuan Wang[1], Hao Shi[1,*]

[1]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 10085, China

[2]Schiller Institute of Integrated Science and Society, Boston College, Chestnut Hill, MA 02467, U.S.

[3]College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL 36849, U.S.

[4]Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, U.S.

[5]Department of Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, IL 61801, U.S.

[6]Oklahoma State University, Natural Resource Ecology & Management, Stillwater, OK 74078, U.S.

[7]School of Agriculture and Food Sciences, The University of Queensland, Brisbane, QLD 4072, Australia

[8]Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Hong Kong 999077, China

[a]The authors contribute equally.

*Correspondence: Hao Shi (haoshi@recces.ac.cn)

Abstract

Surface ozone is an important air pollutant detrimental to human health and vegetation productivity. Regardless of its short atmospheric lifetime, surface ozone has significantly increased since the 1970s across the Northern Hemisphere, particularly in China. However, high temporal resolution surface ozone concentration data is still

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

lacking in China, largely hindering accurate assessment of associated environmental and human health impacts. Here, we collected hourly ground ozone observations (over 6 million records), meteorological data, remote sensing products, and social-economic information, and applied the Long Short-Term Memory (LSTM) recurrent neural networks to map hourly surface ozone data (HrSOD) at a 0.1° × 0.1° resolution across China during 2005–2020. Benefiting from its advantage in time-series prediction, the LSTM model well captured the spatiotemporal dynamics of observed ozone concentrations, with the sample-based, site-based, and by-year cross-validation coefficient of determination (R^2) values being 0.72, 0.65 and 0.71, and root mean square error (RMSE) values being 11.71 ppb (mean = 30.89 ppb), 12.81 ppb (mean = 30.96 ppb) and 11.14 ppb (mean = 31.26 ppb), respectively. Air temperature, atmospheric pressure, and relative humidity were found to be the primary influencing factors. Spatially, surface ozone concentrations were high in northwestern China and low in the Sichuan Basin and northeastern China. Among the four megacity clusters in China, namely the Beijing-Tianjin-Hebei region, the Pearl River Delta, the Yangtze River Delta, and the Sichuan Basin, surface ozone concentration kept decreasing before 2016. However, it tended to increase thereafter in the former three regions, though an abrupt decrease in surface ozone concentrations occurred in 2020. Overall, the HrSOD provides critical information for surface ozone pollution dynamics in China and can support fine-resolution environmental impact and human health risk assessment. The data set is available at https://doi.org/10.5281/zenodo.7415326 (Zhang et al., 2022).

1 Introduction

Ozone (O_3) is an important constituent of the atmosphere and is ubiquitously present in both the troposphere and the stratosphere. Stratospheric ozone protects life on Earth by absorbing harmful solar ultraviolet (UV) rays (Norval et al., 2007; Slaper et al., 1997; van der Leun et al., 2003). Tropospheric ozone is a major gaseous pollutant produced in a series of complex reactions between volatile organic compounds (VOCs) and

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

nitrogen oxides (NOx) in the presence of sunlight (Wang et al., 2017a). Exposure to high-concentration surface ozone can cause severe impacts on human health, inducing high morbidity in respiratory, cardiopulmonary, and cardiovascular diseases (Berman et al., 2012; Li et al., 2018; Magzamen et al., 2017). Moreover, surface ozone of high concentrations could damage the leaf cell structure of plants and thus decrease natural vegetation productivity, crop yield and quality (Cooper et al., 2014; Giles, 2005; Lu et al., 2018; Tian et al. 2016).

In the past decades, the number of ozone pollution events has increased significantly, particularly in highly populated and developed regions (Huang et al., 2018; Ma et al., 2016; Maji and Namdeo, 2021; Sahu et al., 2021). Real-time surface ozone monitoring networks have been established on a regional basis around the world (Chang et al., 2017). But their coverage is still insufficient in both space and time, due to uneven distribution of monitoring sites and lack of mid- to long-term continuous records in the majority of the world (Chang et al., 2017; Lu et al., 2018). In contrast, satellite remote sensing can monitor the spatial and temporal variability of ozone at regional to global scales. For instance, the Ozone Monitoring Instrument (OMI) on the Aura satellite, launched in 2004, provides global daily total column ozone retrievals. Nonetheless, satellite-based estimates of surface ozone concentrations are not available at high spatial and temporal resolutions (Liu et al., 2010; Shen et al., 2019). Hence, various models have been developed to extrapolate site observations, refine satellite retrievals, or fuse them to generate long-term, high-quality surface ozone datasets (e.g., Liu et al., 2020; Wei et al., 2022).

These models, according to their underlying principles, can be generally grouped into chemical transport models (CTMs), geostatistical models, and machine learning models. CTMs are physics-based, accounting for atmospheric chemical reactions, emission inventories, meteorological conditions and transport of atmospheric pollutants, but

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

usually are prone to high uncertainties in emission inventories and model assumptions (Liu et al., 2018; Sun et al., 2019; Travis et al., 2016). Geostatistical models, such as Kriging interpolation (Adam-Poupart et al., 2014), land-use regression (LUR), Bayesian maximum entropy (BME; Chen et al., 2020a), and geographically weighted regression (GWR; Zhang et al., 2020), estimate surface ozone by fitting its relationships with the influential factors. However, collinearity (the non-independence of predictor variables) in these geostatistical models usually makes them difficult to estimate accurately (Jumin et al., 2020; Liu et al., 2020). Machine learning models, such as random forest (Wei et al., 2022) and Extreme Gradient Boosting (Liu et al., 2020), have also been widely used due to their strong data-mining ability. Yet ozone concentrations were not independent at individual time points, multiple predictor variables were correlated at different time points, which was neglected in long time series predictions of ozone.

In recent years, surface ozone pollution in China has become increasingly serious, with frequent large-scale high ozone pollution events (Li et al., 2017a; Mousavinezhad et al., 2021; Wang et al., 2017b). Since 2013, China has established a national ozone observation network (Lu et al., 2018), utilizing which several gridded surface ozone products were generated (Li et al., 2021; Xue et al., 2020). However, gridded products covering an hourly time-step are still lacking in China. Such a data gap impedes accurate assessment of environmental and human health impacts of surface ozone. For example, in estimating ozone damage to vegetation productivity, hourly ozone data is usually required for stomatal ozone flux models (e.g., Feng et al., 2012) or generating ozone exposure index (Ren et al., 2007; Mills et al., 2011). Moreover, hourly ozone data is advantageous over that at coarser temporal resolution in determining ozone exposure of humans (Kim et al., 2011; Niu et al., 2022).

To address the issue, here we developed a deep learning model based on the Long Short-

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Term Memory (LSTM) recurrent neural networks to generate hourly surface ozone data (HrSOD) at a spatial resolution of 0.1°×0.1° from 2005 to 2020 over China. The paper is organized as follows: the data and methods are introduced in Sect. 2; the results regarding model validation, spatiotemporal variations of surface ozone across China, and surface ozone changes in key regions are presented in Sect. 3; comparison of HrSOD with previous studies and the key variables determining surface ozone dynamics are discussed in Sect. 4; data availability is described in Sect. 5; and the conclusions are summarized in Sect. 6.

2 Data and methods

2.1 Data

2.1.1 Surface ozone observation data

Over six million records of hourly surface ozone concentration measurements during 2015–2020 were obtained from the real-time air quality monitoring platform of the China National Environmental Monitoring Centre (CNEMC; http://www.cnemc.cn/, last access: 20 December, 2021). The monitoring network was expanded to more than 1500 monitoring sites from 2013 to 2020, covering 31 provinces and 368 cities across mainland China. However, these monitoring sites are mainly located in the eastern region of China, with a much lower site distribution density in the northwest (Fig. 1).

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

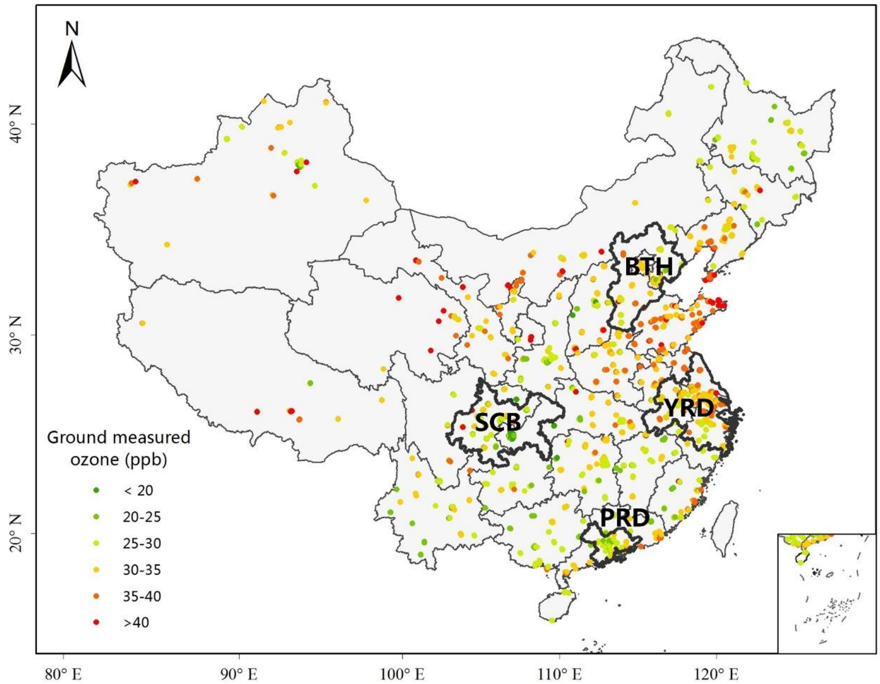

Figure 1. Spatial distribution of surface ozone observation sites in China. The color indicates the annual mean surface ozone concentrations at each site during 2015–2020. The bold black lines indicate the boundaries of four megacity clusters of China, namely the Beijing-Tianjin-Hebei (BTH) region, the Pearl River Delta (PRD), the Sichuan Basin (SCB), and the Yangtze River Delta (YRD).

Hourly ozone concentrations are measured at all monitoring sites by continuous monitoring instruments, and the unit of ozone reported by CNEMC is μg m^{-3} (standard atmospheric conditions at a temperature of 273 K and a pressure of 1013.25 hPa; 1 μg m^{-3} = 0.467 ppb). According to the Ambient Air Quality Standard (GB3095-2012; MEPC, 2012) set by the Ministry of Environmental Protection of China (MEPC) for ozone concentration data norms and standards, the ozone data were screened by eliminating outliers and null values. The annual mean hourly ozone concentrations ranged from 14–48 ppb in the period 2015–2020 in China, with areas of high ozone

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

concentrations mainly in eastern China, especially in the four densely populated megacity clusters of China, i.e., the Beijing-Tianjin-Hebei (BTH) region, the Pearl River Delta (PRD), the Sichuan Basin (SCB) and the Yangtze River Delta (YRD; Fig. 1).

2.1.2 Predictor variables

The predictor variables include meteorological factors, land use, population, gross domestic product, remote sensed total column ozone products, and surface ozone concentration observations (see Table 1).

(1) Climate data

A total of seven climatic variables (solar radiation intensity, temperature, relative humidity, pressure, horizontal wind velocity, vertical wind velocity, and precipitation) were obtained from the ERA5-Land reanalysis dataset (Table 1). The dataset has a spatial resolution of 0.1° × 0.1° (about 9 km) and an hourly time-step and was produced by the European Centre for Medium-Range Weather Forecasts (ECMWF; https://www.ecmwf.int/en/forecasts, last access: 20 December, 2021). The ERA5 reanalysis data combines land surface model simulations with ground and satellite observations (Albergel et al., 2018; Hersbach et al., 2020), and has been widely used across the world (Muñoz-Sabater et al., 2021). It has also been validated in China, showing good performance in air temperature (Zou et al., 2022), solar radiation (Jiang et al., 2020), and precipitation (Jiang et al., 2021).

(2) Remote sensing data

We collected remote sensing data including OMI Level 3 global daily total ozone grid product ($OMITO_3G$; in DU; Pawan, 2012) and ozone profile products (PROFOZ; v0.9.3, level 2) measured by the OMI, which is carried by the Earth Observing System (EOS) Aura satellite. The OMI provided daily and near-global column concentration

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

data (0.25° × 0.25°) and profiles (13 km × 24 km) of O_3, NO_2, SO_2, HCHO. The ozone profile product contained 24 vertical ozone layers (Mcpeters et al., 2008), of which the first layer was selected to represent surface ozone in this study.

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Open Access Earth System Science Data Discussions

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

Category	Variable	Unit	Spatial resolution	Temporal resolution	Source
Ground measurements	O_3: Ozone	μ g m^{-3} convert to ppb	-	Hourly	CNEMC
Climate data	DSR: Downward shortwave radiation	W m^{-2}	0.1° × 0.1°	Hourly	ERA5-Land reanalysis
	RH: Relative humidity	%			
	TEM: 2-m air temperature	K			
	SP: Surface pressure	hPa			
	WU: 10m u-component of wind	m s^{-1}			
	WV: 10m v-component of wind	m s^{-1}			
	PRE: Precipitation	mm			
Socio-economic data	POP: Population	One person per grid	1 km × 1 km	5 years	Resource and Environment Science and Data Center
	GDP: Gross Domestic Production	10000 yuan per grid			
Satellite remote sensing data	TO_3: Total column ozone	DU	0.25° × 0.25°	Daily	OMI/Aura products
	SFO_3: Surface ozone concentrations	DU	13 km × 24 km	Hourly	
	Land use: Forest \ Grassland \ Urban land \ Cropland		0.05°× 0.05°	Annual	MODIS products

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Open Access Earth System Science Data Discussions

(3) Auxiliary data

Socio-economic data reflect human living and production activities, which are major sources of ozone precursors (VOCs and NOx). Thus, it is also an important factor for ozone simulation. We obtained population distribution data and Gross Domestic Product (GDP) data with 1 km spatial resolution from the Resource and Environmental Science and Data Center, Chinese Academy of Sciences (Xu et al., 2017). The data has a time interval of five years and thus it is available in four years (2005, 2010, 2015, and 2020) during the study period. The nationwide land use data was derived from the Moderate Resolution Imaging Spectroradiometer (MODIS; Friedl and Sulla-Menashe, 2015) product (at a resolution of 0.05°).

(4) Data processing

We constructed a 0.1° × 0.1° grid over China and averaged concurrent surface ozone measurements of monitoring sites within a grid cell to obtain grid-level surface ozone concentrations. In addition, all predictor variables (including climate data, land use data, population distribution, GDP data, and remote sensing data) were aggregated or resampled to the targeted grid resolution of 0.1° × 0.1° using the nearest neighbor interpolation and bilinear interpolation approach (Fig. 2).

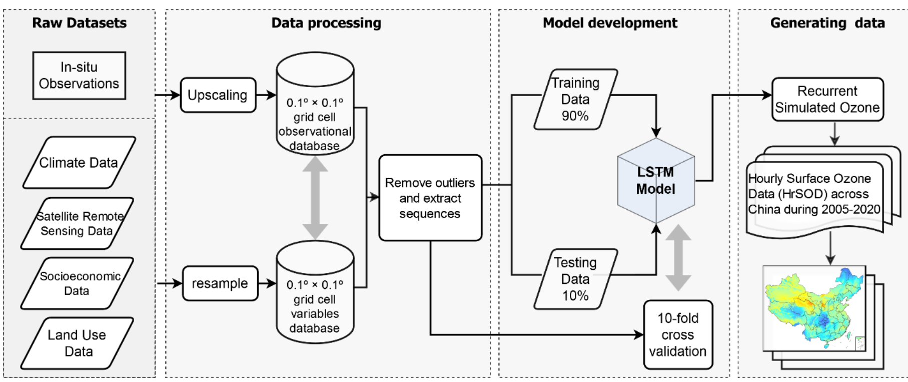

Figure 2. Flowchart for generating hourly surface ozone data (HrSOD) across China.

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Earth System Science Data Discussions
Open Access

2.2 Model development

2.2.1 The long short-term memory network model

The long short-term memory network is a special type of recurrent neural networks (RNNs) that differs from traditional neural networks. The traditional artificial neural network (ANN) is fully connected between layers and has no connection within a specific layer, whereas the hidden layers of RNNs are connected (Hochreiter et al., 1997). The output of an RNN is not only affected by the current input features but also influenced by the output of the previous or next moment, hence RNNs have better performance in estimating time-series and have been widely used to proceed sequence data (Goodfellow et al., 2016).

The LSTM can further overcome the limitations of conventional RNNs that they could be trapped by vanishing gradient or exploding gradient during training (Bengio et al., 1994; Razvan et al., 2013). It excels through integrating input gates, forgetting gates, and output gates into the cell structure. The input gates control whether a cell value can be added to a memory cell, the forgetting gates determine the weight of the value, and the output gates determine which information eventually is output from the cell. The LSTM has a long-term memory capability, which is ideal to predict long time-series of historical ozone concentrations.

Specifically, based on LSTM, we built a five-layer neural network model for surface ozone concentration prediction. It consists of an input layer, two LSTM layers, one Dense layer (also called the fully connected layers), and an output layer (Table 2). The data specification for the model input layer is in a 3-dimensional format (*n_samples, n_time_steps, n_features*), *n_samples* represents the batch size for training, *n_time_steps* is the time window of 24 hours, representing the first 24 hours' O_3 sequence to predict the O_3 at the 25th hour, and *n_features* is the number of 12 variables in the training set. The number of neurons in each hidden layer is 50, and we used mean

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

absolute error (MAE) as the loss function and the Adaptive moment estimation (Adam) as the optimization algorithm. The model was trained for 50 epochs with a batch size of 3000. The CNEMC ground measurements were used as the target for the model training and validation.

Table 2. Detailed configuration of the neural network

Configuration	Value
Training algorithm	Long Short-Term Memory (LSTM)
Number of hidden layers	3
Number of neurons in a hidden layer	50
Number of input variables	12
Number of output variables	1
Training data percentage	90 %
Validation data percentage	10 %
Data normalization	Minmax
Loss function	Mean absolute error (MAE)
Optimization algorithm	Adaptive moment estimation (Adam)

2.2.2 Model evaluation

The 10-fold cross-validation (CV) approach was utilized to evaluate the performance of the LSTM model, with three sampling strategies, namely sample-based CV, site-based CV and by-year CV, corresponding to the model's performances on capturing overall, spatial and temporal patterns, respectively. In each strategy, 90 % of the total surface ozone observations were randomly sampled for training, and the rest 10 % was used for validation, the process of which was repeated 10 times. The overall adjusted coefficient of determination (R^2), root mean square error (RMSE), linear regression slope, and intercept were calculated to evaluate the performance of the model.

3 Results

3.1 Model validation

At the hourly time-scale, the LSTM model obtained R^2 values of 0.72, 0.65, 0.71 using three CV sampling methods (sample-based, site-based and by-year), respectively, and the corresponding RMSE values were 11.71 ppb, 12.81 ppb, 11.14 ppb (Figs. 3a–c). At the daily time-step, the model's performance improved with R^2 values being 0.71, 0.63, 0.71 (sample-based, site-based, and by-year) and RMSE values being 8.53 ppb, 9.61 ppb, and 7.97 ppb (Figs. 3d–f). The predictive ability of the model further improved at the monthly time-step, with higher R^2 values of 0.82, 0.72, 0.84 (sample-based, site-based, and by-year) and smaller RMSE values of 5.14 ppb, 6.54 ppb, 4.39 ppb (sample-based, site-based, and by-year (Figs. 3g–i).

Among the three CV sampling strategies, the site-based CV (Figs. 3b, e, h) R^2 values were slightly lower than the sample-based CV (Figs. 3a, d, g) R^2 values and by-year CV (Figs. 3c, f, i) R^2 values, while the RMSE values were slightly higher than the sample-based CV RMSE values and by-year CV RMSE values. It is noted that the model underestimated surface ozone when it was at high concentrations, but this bias was largely ameliorated at the monthly time-step (Figs. 3g–i).

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

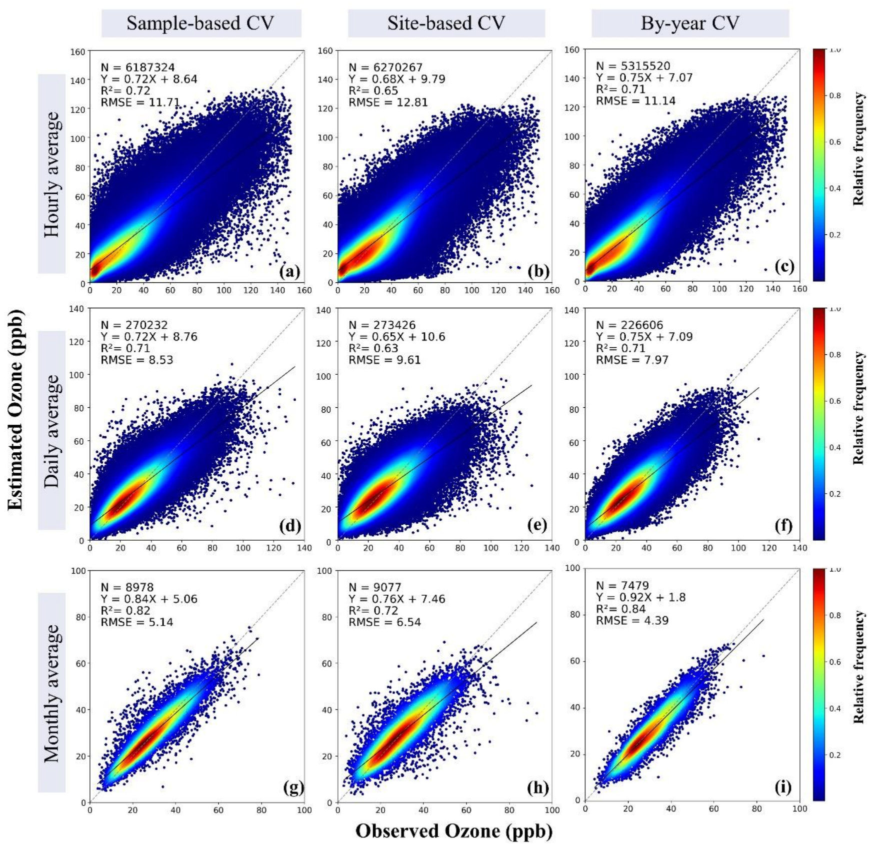

Figure 3. Comparisons between model estimated surface ozone concentrations and observations across China. The panels are sample-based cross validations at hourly, daily and monthly time-steps **(a, d, g)**, site-based cross validations at hourly, daily and monthly time-steps **(b, e, h)**, and by-year cross validations at hourly, daily and monthly time-steps **(c, f, i)**. The dashed and black lines represent the 1:1 lines and the linear regression lines, respectively.

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

3.2 Spatiotemporal variations of surface ozone across China

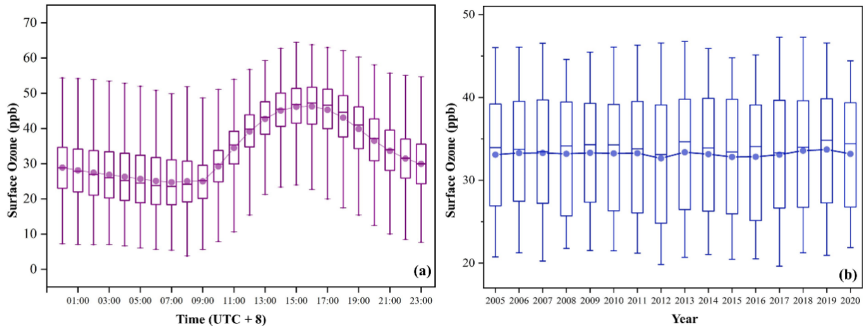

Figure 4. Diurnal **(a)** and interannual **(b)** variations of mean surface ozone concentrations in China during 2005–2020. Boxplots indicate the median (horizontal line) and interquartile ranges (boxes) and the whiskers specify the maximum and minimum values.

Fig. 4a shows the diurnal variations of mean hourly surface O_3 concentrations across China during 2005-2020. The diurnal variation presented a unimodal curve, which started to increase at around 9:00–10:00 (UTC + 8) and peaked at around 15:00 (UTC + 8) reaching about 46 ppb. After that, the hourly mean O_3 concentrations gradually declined to about 25–28 ppb. The annual average O_3 concentration across China (Fig. 4b) ranged from 32.56 ± 7.59 ppb to 33.61 ± 7.16 ppb during 2005–2020. During the first part of the period (2005–2016), ozone concentrations remained stable. However, after 2016 the O_3 concentration showed a significantly increasing trend from 2016 (32.75 ± 7.17 ppb) to 2019 (33.61 ± 7.16 ppb), and then an abrupt decrease by 2020 (33.09 ± 6.93 ppb).

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

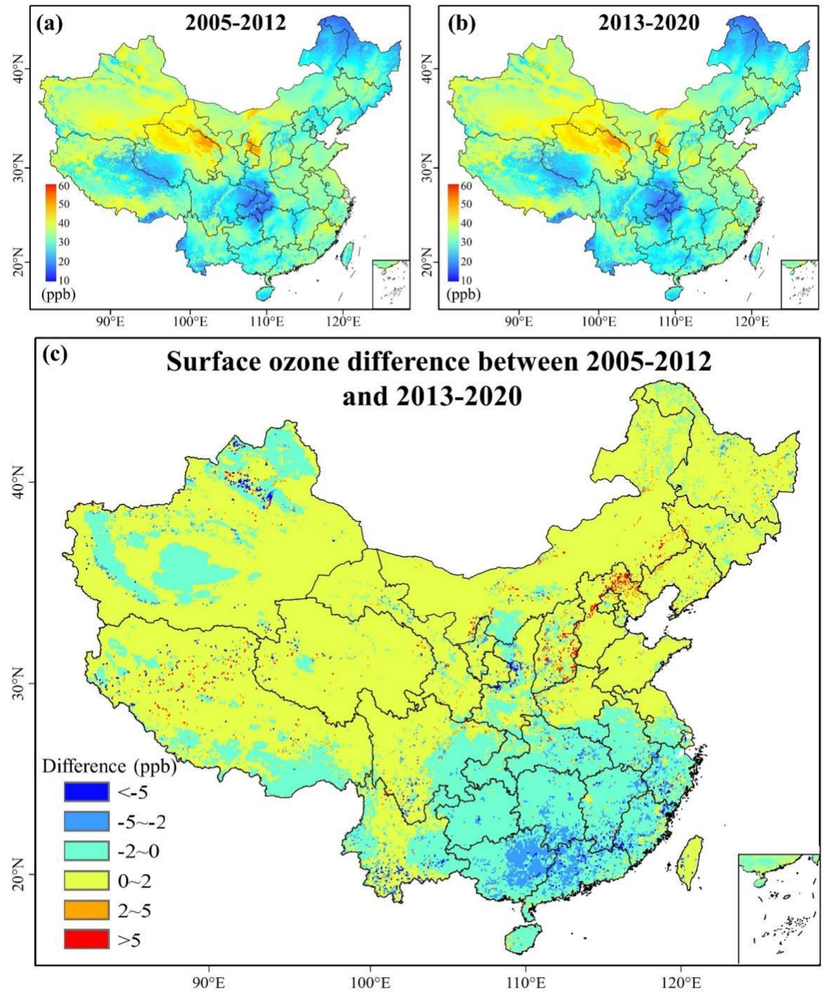

Figure 5. Mean annual O_3 concentrations during the periods 2005–2012 **(a)**, and 2013–2020 **(b)**, with their difference also shown **(c)**.

Fig. 5a and Fig. 5b show the mean annual surface ozone concentrations in China from 2005 to 2012 and from 2013 to 2020, respectively. The spatial distribution of surface ozone concentrations was similar and did not change significantly over the 16-year period. High O_3 concentrations were primarily in the northwest of China, while areas with low O_3 concentrations were mainly located in the Sichuan Basin and the northern

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

part of northeast China. Compared to the first eight years, mean annual surface ozone concentrations in the last eight years increased by 0–2 ppb in most northern regions of China (Fig. 5c). In particular, the BTH region experienced a faster increase of 2–5 ppb. In contrast, in southern China, surface O_3 concentrations had generally decreased, especially in the southern coastal regions, which decreased by 2–5 ppb.

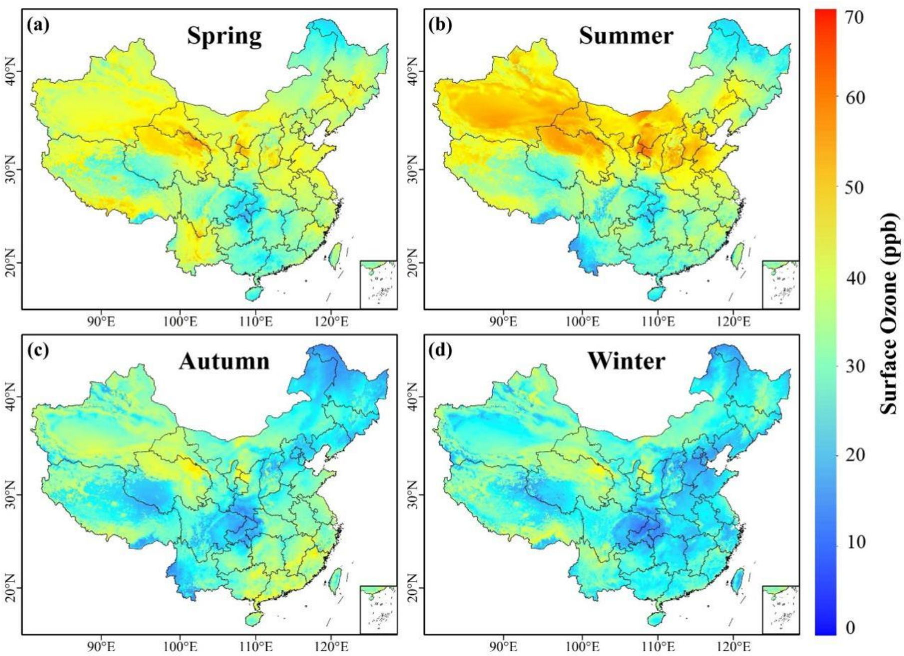

Figure 6. Seasonal average surface ozone concentrations from 2005 to 2020 across China in spring **(a)**, summer **(b)**, autumn **(c)**, and winter **(d)**.

The multi-year mean seasonal O_3 concentrations were predicted to be 37.64 ± 3.35, 39.16 ± 2.37, 28.40 ± 3.17, and 25.07 ± 2.60 ppb in spring (March–May), summer (June–August), autumn (September–November), and winter (December, January, and February), respectively (Fig. 6). Surface ozone concentrations in spring were higher in northern and eastern China. In summer, the areas with high ozone concentrations were North China, the Northwestern District of China and southern Inner Mongolia. The

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

hotspot areas with high O_3 concentrations in autumn shrink sharply and spread to the southeast coast. During winter, the areas of high O_3 concentrations almost disappeared in southeastern China.

3.3 Surface ozone changes in key regions

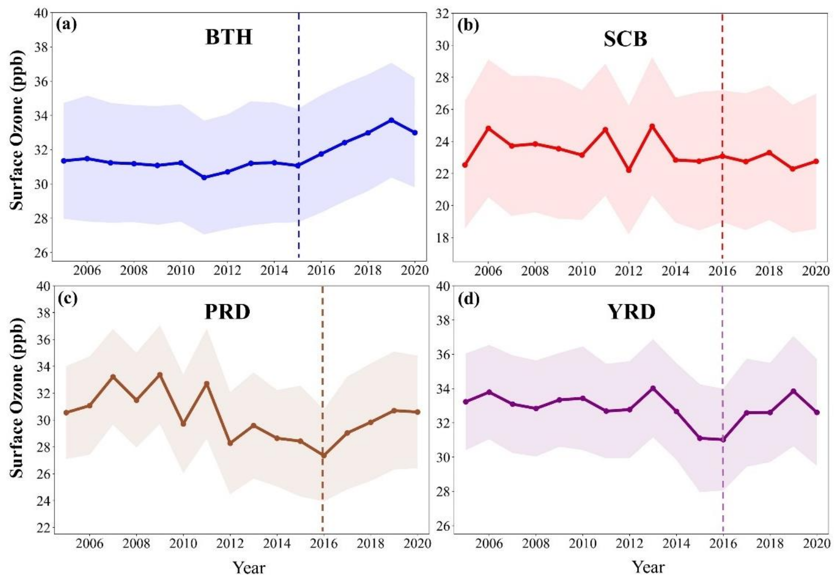

Figure 7. Temporal dynamics of mean annual mean surface O_3 concentrations in the BTH **(a)**, SCB **(b)**, PRD **(c)**, and YRD **(d)** regions. BTH: Beijing-Tianjin-Hebei region; SCB: Sichuan Basin; PRD: Pearl River Delta; YRD: Yangtze River Delta.

Among the four megacity clusters, mean annual surface ozone concentrations in BTH (mean = 31.63 ppb), YRD (mean = 32.85 ppb), and PRD (mean = 30.28 ppb) regions were higher than in the SCB (mean = 23.33 ppb) region during 2005–2020 (Fig. 7). In the BTH region (Fig. 7a), surface ozone concentrations remained stable from 2005 to 2015. However, after 2015 it showed a continuous and noticeable increase from 31.07 ± 3.31 ppb in 2015 to 33.72 ± 3.34 ppb in 2019, before decreasing to 32.99 ± 3.19 ppb in 2020. In the PRD and YRD regions (Figs. 7c and d), the annual ozone concentrations showed an obvious decline from 29.59 ± 3.94 and 34.01 ± 2.85 ppb in 2013 to 27.35 ±

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

3.41 ppb and 31.01 ± 2.96 ppb in 2016, respectively, and then increased from 2016 to 30.70 ± 4.39 ppb and 33.84 ± 3.21 ppb in 2019. Same with BTH, both regions experienced a decrease in ozone concentrations in 2020. In contrast, annual surface ozone concentrations in the SCB region were constantly low and had a slow decreasing trend.

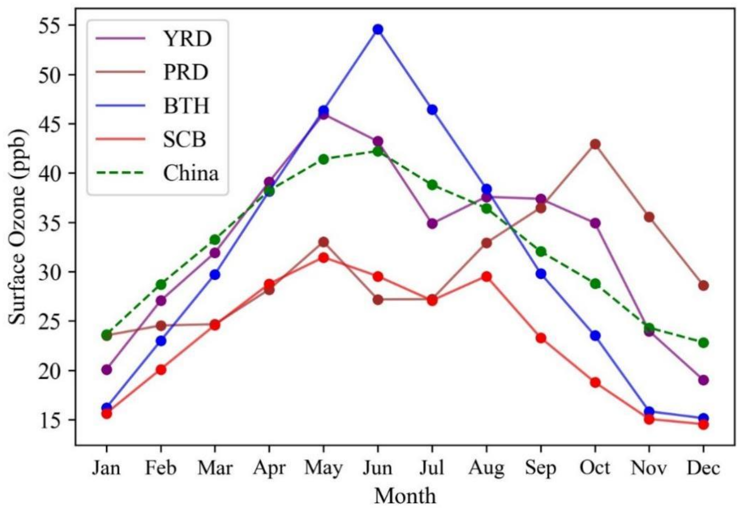

Figure 8. Mean monthly surface ozone concentrations in China and the four hotspot regions of BTH **(a)**, SCB **(b)**, PRD **(c)**, and YRD **(d)**. BTH: Beijing-Tianjin-Hebei region; SCB: Sichuan Basin; PRD: Pearl River Delta; YRD: Yangtze River Delta.

The seasonal patterns of surface ozone concentrations were different in four key regions (Fig. 8). The monthly mean ozone concentrations were higher than 38.00 ppb across China from April to July and less than 24.00 ppb in January, November, and December. The ozone concentrations in BTH were unimodal distribution and gradually increased with time, peaking in June (55.58 ppb) and then began to decline, reaching the lowest value in December (15.22 ppb). Unlike BTH, the other three regions (YRD, SCB, and PRD) showed a bimodal pattern. The first peak of ozone concentrations occurred in May (45.29 ppb in YRD, 31.76 ppb in PRD, and 30.99 ppb in SCB), and the ozone

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Open Access Earth System Science Data Discussions

concentrations in the YRD, PRD, and SCB reached the second peak in September (38.53 ppb), October (40.55 ppb), and August (29.90 ppb), respectively. The lowest surface ozone concentrations were found to be 18.35 ppb (YRD in December), 23.03 ppb (PRD in January), and 14.11 ppb (SCB in December).

4 Discussion

4.1 Comparison with previous studies

To make comparisons with previous studies, we also generated maximum daily average 8-hour ozone (MDA8) concentration using the HrSOD data. The HrSOD MDA8 O_3 concentrations compared equally (Table 3) with previous products produced on the daily or monthly scales, indicating that HrSOD is reliable in deriving ozone exposure indexes at longer time-scales. The mean annual HrSOD MDA8 O_3 concentrations (2005–2020) was 43.56 ppb (Figure S1) and it showed no significant trend during the period 2005 to 2015 but tended to increase after 2016 with a growth rate of 0.44 ppb yr^{-1} ($p < 0.005$), similar to previous reports (Liu et al. 2020; Wei et al. 2022; Xue et al. 2020; Table S1). The spatial distributions of mean annual and seasonal HrSOD MDA8 ozone concentrations were also consistent with existing research (Liu et al. 2020; Meng et al. 2022; Figs. S2 and S3). Areas with higher ozone concentrations in summer were concentrated in the North China Plain (NCP), due to decreasing $PM_{2.5}$ concentrations and high NOx emissions, as well as the influence of rapidly increasing temperatures (2017–2019) and foehn winds (Li et al., 2020). Pollution in southeastern China is more severe in autumn, which may be related to the Asian summer monsoon, tropical cyclones, and sea-land winds (Liu et al. 2020).

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Table 3. Comparison with the model performance of previous studies in predicting surface ozone in China.

Method	Time Range	Metric	Accuracy (R^2/ RMSE) (ppb)			Reference
			Hourly	**Daily**	**Monthly**	
GWR	2014	Monthly	–	–	0.81/–	Zhang et al. (2020)
Data fusion	2013–2017	[O_3] MDA8	–	0.70/12.23	0.69 / 9.01	Xue et al. (2020)
XGBoost	2005–2017	[O_3] MDA8	–	0.78/10.03	0.90/5.17	Liu et al. (2020)
RF	2015	[O_3] MDA8	–	0.69/12.42	0.71/8.87	Zhan et al. (2018)
STET	2013–2020	[O_3] MDA8	–	0.83/8.82	0.90/5.80	Wei et al. (2022)
RF	2013–2019	[O_3] MDA8	–	0.80/9.77	0.83/6.612	Meng et al. (2022)
LUR/BME	2015	[O_3] MDA8	–	0.80/10.97	–	Chen et al. (2020a)
LSTM	**2005–2020**	**[O_3] MDA8 (derived from HrSOD)**	–	**0.73/11.37**	**0.82/6.85**	This study
LSTM	**2005–2020**	**Hourly**	**0.72/11.71**	**0.71/8.53**	**0.82/5.14**	This study

LSTM: Long Short-Term Memory; XGBoost: Extreme Gradient Boosting; RF: Random Forest; STET: Space-Time extremely randomized trees; LUR/BME: land-use regression/Bayesian maximum entropy; GWR: geographically weighted regression.

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

4.2 Key variables in estimating surface ozone concentrations

The importance of different variables in the LSTM model was calculated by the permutation importance method (François et al., 2006). Specifically, the feature importance was determined by the degree of decline in the performance score of the model after the random rearrangement of different features. As shown in Fig. S4, air temperature, surface pressure, relative humidity, day of year (DOY), and downwelling surface radiation were the top five factors affecting the spatiotemporal variability of surface ozone concentrations in China, consistent with previous studies (e.g., Wei et al. 2022).

The importance of air temperature in predicting surface ozone concentrations is mainly reflected in two aspects. On the one hand, temperature and UV radiation intensity has a strong correlation and can be used to characterize UV radiation intensity; on the other hand, high temperatures contribute to the volatilization of ozone precursors (such as biogenic volatile organic compounds) and accelerate the rate of photochemical reactions (Xu et al., 2011). There is a significantly negative correlation between ozone concentrations and atmospheric pressure, as changes in air pressure are usually correlated with temperature. For example, low pressure corresponds to higher temperatures, and when the near-surface is controlled by low pressure, pollutants from surrounding areas converge towards the center, driven by high-pressure air masses, resulting in a sharp increase in ozone concentrations in the center of the low pressure (Kovač-Andrić et al., 2009). The relative humidity is negatively correlated with O_3 concentrations because the high relative humidity generally corresponds to precipitation, fog, and other weathers that do not have strong UV radiation, which is not conducive to the occurrence of photochemical reactions and the further development of O_3 pollution. Furthermore, precipitation facilitates the removal of pollutants such as O_3 (Chen et al., 2020b; Li et al., 2017b). The importance of DOY indicates there is temporal autocorrelation in ozone concentrations of neighboring days. GDP and population data show levels of urbanization, fossil fuel consumption, and other socio-economic activities, which are closely related to emissions of ozone

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

precursors such as NOx and VOCs (Trainer et al., 2000). Although DSR is not the most important variable, it also makes a great contribution to simulating O_3 concentrations, which is necessary for the photochemical reactions of O_3 generation (Chen et al., 2020b).

4.3 The driving factors for surface ozone concentrations in key regions

The four city clusters (BTH, PRD, SCB, and YRD) are key areas for the Chinese government to combat air pollution. However, they have different meteorological conditions, climatic backgrounds, and levels of economic development, resulting in differences in surface ozone concentration variabilities (Yan et al., 2021). The increase in surface ozone concentrations in the BTH region was mainly attributable to anthropogenic emissions (VOCs and NOx), as well as a rapid reduction in $PM_{2.5}$, which slowed down the sink of hydrogen peroxide radicals and thus accelerated ozone production. The meteorological conditions of lower humidity and strong solar radiation also contributed to ozone pollution (Li et al., 2019; Mousavinezhad et al., 2021; Wei et al., 2021). In the YRD, PRD and SCB, meteorological factors had a greater influence on surface ozone concentration changes after 2016, with the increase in ozone concentrations in the YRD attributed to increased solar radiation and temperature, and lower atmospheric pressure. Simulation results from the three-dimensional air quality model and system also indicated that greenhouse gas emissions caused changes in meteorological factors that led to increased O_3 concentrations in the YRD region (Xie et al., 2017). Weaker meridional winds, lower relative humidity, and higher temperatures escalated ozone pollution in the PRD region from 2016 to 2018. A decrease in solar radiation and planetary boundary layer height accelerated the decrease in ozone concentrations in the SCB region after 2017 (Mousavinezhad et al., 2021). In addition, ozone concentrations in three areas (BTH, YRD, and PRD) decreased in 2020, mainly due to COVID-19 lockdown (Wei et al., 2022).

4.4 Uncertainties and limitations

In this research, uncertainties exist in several aspects. First, the monitoring stations were

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

mainly concentrated in the central-eastern region of China, which potentially could not fully capture the relationship between surface ozone concentrations and environmental factors in western China. Moreover, most monitoring stations were in urban areas, resulting in limitations of the model in estimating surface ozone concentrations in natural and agricultural ecosystems. Second, the input data may also cause uncertainty. For example, ERA5 reanalysis data underestimates surface temperatures in the coastal urban agglomerations of southeast China and the Tibetan Plateau region (Li et al., 2022; Zou et al., 2022), which may lead the model to underestimate ozone concentrations. To further improve ozone estimation accuracy, it is necessary to improve the accuracy of meteorological data, land use maps, as well as socio-economic data. In addition, the mismatch in temporal resolution between OMI remote sensing data and ozone measurements may also affect the final estimation accuracy. Although the temporal trends of surface ozone concentrations are well captured by the LSTM networks, spatial information, such as the changes in pollutant concentrations due to the emission and transport of surrounding pollutants are not fully considered. Therefore, the current deep learning model can be further improved by combining other algorithms. For instance, CNN networks have powerful feature extraction capability and can be combined with LSTM to generate an integrated CNN-LSTM model, which makes better use of the temporal memory strengths and feature representation capability in prediction.

4.5 Potential applications of HrSOD

Compared to current available surface ozone products in China, HrSOD covers a longer time range and has a higher temporal resolution. This enables it to support more robust historical environmental impact and human health risk assessments. HrSOD can be used to derive various ozone exposure indicators, such as seasonal 7-h mean O_3 concentrations (M7), seasonal 12-h mean O_3 concentrations (M12; Legge et al., 1995), sum of all hourly average concentrations > 60 μg kg^{-1} (SUM06; Lefohn and Foley, 1992), cumulative ozone exposure index based on sigmoid-weighted daytime O_3 concentrations (W126; Fuhrer et al., 1997), accumulated hourly O_3 concentration over a threshold of X μg kg^{-1} during daylight hours (AOTX; Fuhrer et al., 1997). Therefore,

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

HrSOD can meet various requirements by ozone impact models, providing flexibility for assessing ozone effects on ecosystem (Ren et al., 2007) and epidemiological studies (Huangfu and Atkinson, 2020).

5 Data availability

The HrSOD dataset is available on the Zenodo repository at https://doi.org/10.5281/zenodo.7415326 (Zhang et al., 2022). The gridded ozone concentration data are provided in NetCDF format at 0.1° spatial resolution and hourly temporal resolution during 2005–2020 in ppb. The file size is 40 GB. The daily data is a NetCDF file and the file is named "YYYYMMDD.nc", where "YYYY", "MM" and "DD" refer to the year, month, and day of the file.

6 Conclusions

In this study, we trained and validated a LSTM model to estimate HrSOD during 2005–2020 across China. The predictor variables included meteorological factors, remote sensing data, socio-economic data, and land use data, and more than six million ground station monitoring records were collected as reference data. Compared with observations, the model showed good performances at diurnal, seasonal to annual time-scales and site to regional levels. HrSOD showed that surface O_3 concentrations in China tended to increase from 2016 to 2019 due to anthropogenic and meteorological factors such as temperature, humidity, and radiation intensity, despite a decrease in 2020 due to COVID-19 lockdown. In summary, HrSOD had high spatial and temporal accuracies, long time ranges and high temporal resolution, enabling it to be easily converted to various evaluation indicators for ecosystem and human health assessments.

Author contributions

WZ and DL contributed equally, with WZ performing the data curation, modelling, validation, and writing the original draft of the paper. DL was responsible for conceptualization, providing data, and reviewed the manuscript. HS was the lead and corresponding author of this work, supported and supervised the study and reviewed

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

the paper. HT and SW also conceptualized the project and supervised the simulations and analyses. RY, WT and HM provided coding and developed the model. HT, NP, JY, FL, BD and SW contributed to the writing of the manuscript.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Earth System Science Data.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgments

The authors would like to thank members of China National Environmental Monitoring Centre (CNEMC) for their production and maintenance, and providing the in situ surface ozone observation data. Furthermore, authors would also like to thank team members Prof. Dr. Jun Wang from Nanjing University for their contributions to data preparation and the ERA5 team, the MODIS team, and the NASA team for archiving and providing the data used in this study. The authors would like to give special thanks to Haiyang Liu; Xiaoyu Qin; Yihang Li; Minghui Zhang and Shanming Fang for their help in this study.

Financial support

This study is supported by the National Key Research and Development Program of China (Grant No. 2018YFA0606001). H.T. acknowledges funding support from the National Science Foundation of the United States (Grant No. 1903722) and Andrew Carnegie fellowship Program (Grant No. G-F-19-56910).

References

Adam-Poupart, A., Brand, A., Fournier, M., Jerrett, M., and Smargiassi, A.:

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

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, 970–976, https://doi.org/10.1289/ehp.1306566, 2014.

Ainsworth, E. A., Yendrek, C. R., Sitch, S., Collins, W. J., and Emberson, L. D.: 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, 2012.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., de Rosnay, P., and Balsamo, G.: ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better?, Hydrol. Earth Syst. Sci., 22, 3515–3532, https://doi.org/10.5194/hess-22-3515-2018, 2018.

Bengio, Y., Simard, P., and Frasconi, P.: Learning long-term dependencies with gradient descent is difficult, IEEE Transactions on Neural Networks, 5, 157 – 166, https://doi.org/10.1109/72.279181, 1994.

Berman, J. D., Fann, N., Hollingsworth, J. W., Pinkerton, K. T., Rom, W. N., Szema, A. M., Breysse, P. N., White, R. H., and Curriero, F. C.: Health benefits from large-scale ozone reduction in the United States, Environ. Health Perspect., 120, 1404–1410, https://doi.org/10.1289/ehp.1104851, 2012.

Chang, K.-L., Petropavlovskikh, I., Cooper, O. R., Schultz, M. G., and Wang, T.: Regional trend analysis of surface ozone observations from monitoring networks in eastern North America, Europe and East Asia, Elementa: Science of the Anthropocene, 5, https://doi.org/10.1525/elementa.243, 2017.

Chen, L., Liang, S., Li, X., Mao, J., and Azzi, M.: 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, 2020a.

Chen, Z., Li, R., Chen, D., Zhuang, Y., Gao, B., Yang, L., and Li, M.: Understanding the causal influence of major meteorological factors on ground ozone concentrations across China, J. Cleaner Prod., 242, 118498,

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

https://doi.org/10.1016/j.jclepro.2019.118498, 2020b.

Cooper, O. R., Parrish, D. D., and Ziemke, J. R.: Global distribution and trends of tropospheric ozone: an observation-based review, Elementa-Science. Anthrop., 2, 1–28, https://doi.org/10.12952/journal.elementa.000029, 2014.

Feng, Z., Tang, H., Uddling, J., Pleijel, H., Kobayashi, K., Zhu, J., Oue, H., and Guo, W.: A stomatal ozone flux-response relationship to assess ozone-induced yield loss of winter wheat in subtropical China, Environ. Pollut., 164, 16－23, https://doi.org/10.1016/j.envpol.2012.01.014, 2012.

François, D., Wertz, V., and Verleysen, M.: The permutation test for feature selection by mutual information, in: Proceedings of European Symposium on Artificial Neural Networks, Bruges, Belgium, 26–28 April 2006, 239–244, 2006.

Friedl, M., and Sulla-Menashe, D.: MCD12C1 MODIS/Terra + Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006, NASA EOSDIS Land Processes DAAC [data set], https://doi.org/10.5067/MODIS/MCD12C1.006, 2015.

Fuhrer, J., Skärby, L., and Ashmore, M. R.: Critical levels for ozone effects on vegetation in Europe, Environ. Pollut., 97, 91–106, https://doi.org/10.1016/S0269-7491(97)00067-5, 1997.

Giles, J.: Hikes in surface ozone could suffocate crops, Nature, 435, 7, https://doi.org/10.1038/435007a, 2005.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning. MIT Press, Cambridge, MA, USA, 2016.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Comput., 9, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

Huang, J., Pan, X., Guo, X., and Li, G.: Health impact of China's air pollution prevention and control action plan: an analysis of national air quality monitoring and mortality data, Lancet Planet. Health, 2, e313−e323, https://doi.org/10.1016/S2542-5196(18)30141-4, 2018.

Huangfu, P. and Atkinson, R.: Long-term exposure to NO_2 and O_3 and all-cause and respiratory mortality: a systematic review and meta-analysis, Environ. Int., 144, 105998, https://doi.org/10.1016/j.envint.2020.105998, 2020.

Jiang, H., Yang, Y., Bai, Y., and Wang, H.: Evaluation of the total, direct, and diffuse solar radiations from the ERA5 reanalysis data in China, IEEE Geosci. Remote. Sens. Lett., 17, 47–51, https://doi.org/10.1109/LGRS.2019.2916410, 2020.

Jiang, Q., Li, W., Fan, Z., He, X., Sun, W., Chen, S., Wen, J., Gao, J., and Wang, J.: Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland, Journal of Hydrology, 595, 125660, https://doi.org/10.1016/j.jhydrol.2020.125660, 2021.

Jumin, E., Zaini, N., Ahmed, A. N., Abdullah, S., Ismail, M., Sherif, M., Sefelnasr, A., and El-Shafie, A.: Machine learning versus linear regression modelling approach for accurate ozone concentrations prediction, Eng. Appl. Comput. Fluid Mech., 14, 713–725, https://doi.org/10.1080/19942060.2020.1758792, 2020.

Kim, C. S., Alexis, N. E., Rappold, A. G., Kehrl, H., Hazucha, M. J., Lay, J. C., Schmitt, M. T., Case, M., Devlin, R. B., Peden, D. B., and Diaz-Sanchez, D.: Lung function and inflammatory responses in healthy young adults exposed to 0.06 ppm ozone for 6.6 hours, Am. J. Respir. Crit. Care Med., 183, 1215–1221, https://doi.org/10.1164/rccm.201011-1813OC, 2011.

Kovač-Andrić, E., Brana, J., and Gvozdić, V.: Impact of meteorological factors on ozone concentrations modelled by time series analysis and multivariate statistical methods, Ecol. Inform., 4, 117–122, https://doi.org/10.1016/j.ecoinf.2009.01.002, 2009.

Lefohn, A. S. and Foley, J. K.: NCLAN results and their application to the standard-setting process: protecting vegetation from surface ozone exposures, J. Air Waste

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Manage. Assoc., 42, 1046–1052, https://doi.org/10.1080/10473289.1992.10467049, 1992.

Legge, A. H., Grünhage, L., Nosal, M., Jäger, H. J., and Krupa, S. V.: Ambient ozone and adverse crop response: an evaluation of North American and european data at they relate to exposure indices and critical levels, Journal of Applied Botany and Food Quality, 69, 192–205,1995.

Li, G., Bei, N., Cao, J., Wu, J., Long, X., Feng, T., Dai, W., Liu, S., Zhang, Q., and Tie, X.: Widespread and persistent ozone pollution in eastern China during the non-winter season of 2015: observations and source attributions, Atmos. Chem. Phys., 17, 2759–2774, https://doi.org/10.5194/acp-17-2759-2017, 2017a.

Li, H., Wu, S., Pan, L., Xu, J., Shan, J., Yang, X., Dong, W., Deng, F., Chen, Y., Shima, M., and Guo, X.: Short-term effects of various ozone metrics on cardiopulmonary function in chronic obstructive pulmonary disease patients: Results from a panel study in Beijing, China, Environ. Pollut., 232, 358–366, https://doi.org/10.1016/j.envpol.2017.09.030, 2018.

Li, K., Chen, L., Ying, F., White, S. J., Jang, C., Wu, X., Gao, X., Hong, S., Shen, J., Azzi, M., and Cen, K.: Meteorological and chemical impacts on ozone formation: A case study in Hangzhou, China, Atmospheric Research, 196, 40–52, https://doi.org/10.1016/j.atmosres.2017.06.003, 2017b.

Li, K., Jacob, D. J., Liao, H., Shen, L., Zhang, Q., and Bates, K. H.: Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China, Proceedings of the National Academy of Sciences, 116, 422–427, https://doi.org/10.1073/pnas.1812168116, 2019.

Li, K., Jacob, D. J., Shen, L., Lu, X., De Smedt, I., and Liao, H.: Increases in surface ozone pollution in China from 2013 to 2019: anthropogenic and meteorological influences, Atmos. Chem. Phys., 20, 11423–11433, https://doi.org/10.5194/acp-20-11423-2020, 2020.

Li, M., Yu, S., Chen, X., Li, Z., Zhang, Y., Wang, L., Liu, W., Li, P., Lichtfouse, E., Rosenfeld, D., and Seinfeld, J. H.: Large scale control of surface ozone by relative humidity observed during warm seasons in China, Environmental Chemistry

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Letters, 19, 3981–3989, 10.1007/s10311-021-01265-0, 2021.

Li, Y. Z., Qin, X., Liu, Y. S., Jin, Z. Z., Liu, J., Wang, L. H., and Chen, J. Z.: Evaluation of long-term and high-resolution gridded precipitation and temperature products in the Qilian Mountains, Qinghai-Tibet Plateau, Front. Environ. Sci., 10, 10.3389/fenvs.2022.906821, 2022.

Liu, H., Liu, S., Xue, B., Lv, Z., Meng, Z., Yang, X., Xue, T., Yu, Q. and He, K.: Ground-level ozone pollution and its health impacts in China, Atmos. Environ., 173, 223‒230, https://doi.org/10.1016/j.atmosenv.2017.11.014, 2018.

Liu, R., Ma, Z., Liu, Y., Shao, Y., Zhao, W., and Bi, J.: 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, 2020.

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10, 2521–2537, https://doi.org/10.5194/acp-10-2521-2010, 2010.

Lu, X., Hong, J., Zhang, L., Cooper, O. R., Schultz, M. G., Xu, X., Wang, T., Gao, M., Zhao, Y., and Zhang, Y.: Severe surface ozone pollution in China: a global perspective, Environ. Sci. Technol. Lett., 5, 487–494, https://doi.org/10.1021/acs.estlett.8b00366, 2018.

Ma, Z., Xu, J., Quan, W., Zhang, Z., Lin, W., and Xu, X.: Significant increase of surface ozone at a rural site, north of eastern China, Atmos. Chem. Phys., 16, 3969–3977, https://doi.org/10.5194/acp-16-3969-2016, 2016.

Magzamen, S., Moore, B. F., Yost, M. G., Fenske, R. A., and Karr, C, J.: Ozone-related respiratory morbidity in a low-pollution region, Occup. Environ. Med., 59, 624–630, https://doi.org/10.1097/JOM.0000000000001042, 2017.

Maji, K. J. and Namdeo, A.: Continuous increases of surface ozone and associated premature mortality growth in China during 2015–2019, Environ Pollut., 269, 116183, https://doi.org/10.1016/j.envpol.2020.116183, 2021.

Maji, K. J., Ye, W.-F., Arora, M., and Nagendra, S. M. S.: Ozone pollution in Chinese cities: Assessment of seasonal variation, health effects and economic burden, Environ. Pollut., 247, 792–801, https://doi.org/10.1016/j.envpol.2019.01.049,

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

2019.

McPeters, R., Kroon, M., Labow, G., Brinksma, E., Balis, D., Petropavlovskikh, I., Veefkind, J. P., Bhartia, P. K., and Levelt, P. F.: Validation of the Aura ozone monitoring instrument total column ozone product, J. Geophys. Res. Atmos., 113, https://doi.org/10.1029/2007JD008802, 2008.

Meng, X., Wang, W., Shi, S., Zhu, S., Wang, P., Chen, R., Xiao, Q., Xue, T., Geng, G., Zhang, Q., Kan, H., and Zhang, H.: Evaluating the spatiotemporal ozone characteristics with high-resolution predictions in mainland China, 2013–2019, Environ. Pollut., 299, 118865, https://doi.org/10.1016/j.envpol.2022.118865, 2022.

MEPC (Ministry of Environmental Protection of China): Ambient Air Quality Standard (GB3095-2012), available online: https://www.mee.gov.cn/, last access: 20 March, 2022, 2012.

Mills, G., Hayes, F., Simpson, D., Emberson, L., Norris, D., Harmens, H., and Bueker, P.: Evidence of widespread effects of ozone on crops and (semi-) natural vegetation in Europe (1990–2006) in relation to AOT40- and flux- based risk maps, Global Change Biol., 17, 592–613, https://doi.org/10.1111/j.1365-2486.2010.02217.x, 2011.

Mousavinezhad, S., Choi, Y., Pouyaei, A., Ghahremanloo, M., and Nelson, D. L.: A comprehensive investigation of surface ozone pollution in China, 2015–2019: Separating the contributions from meteorology and precursor emissions, Atmos. Res., 257, 105599, https://doi.org/10.1016/j.atmosres.2021.105599, 2021.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Niu, Y., Yang, T., Gu, X., Chen, R., Meng, X., Xu, J., Yang, L., Zhao, J., Zhang, X., Bai, C., Kang, J., Ran, P., Shen, H., Wen, F., Huang, K., Chen, Y., Sun, T., Shan, G.,

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Lin, Y., Wu, S., Zhu, J., Wang, R., Shi, Z., Xu, Y., Ye, X., Song, Y., Wang, Q., Zhou, Y., Ding, L., Yang, T., Yao, W., Guo, Y., Xiao, F., Lu, Y., Peng, X., Zhang, B., Xiao, D., Wang, Z., Zhang, H., Bu, X., Zhang, X., An, L., Zhang, S., Cao, Z., Zhan, Q., Yang, Y., Liang, L., Cao, B., Dai, H., Wu, T., He, J., Li, H., Kan, H., Wang, C., and China Pulmonary Health study group: long-term ozone exposure and small airway dysfunction: the China Pulmonary Health (CPH) study, Am. J. Respir. Crit. Care Med., 205, 450–458, https://doi.org/10.1164/rccm.202107-1599OC, 2022.

Norval, M., Cullen, A. P., de Gruijl, F. R., Longstreth, J., Takizawa, Y., Lucas, R. M., Noonan, F. P., and van der Leun, J. C.: The effects on human health from stratospheric ozone depletion and its interactions with climate change, Photochem. Photobiol. Sci., 6, 232-251, https://doi.org/10.1039/b700018a, 2007.

Pawan, K.: OMI/Aura Ozone (O_3) total column daily L2 global ridded 0.25 degree × 0.25 degree, GESDISC [data set], https://disc.gsfc.nasa.gov, 2012.

Razvan, P., Mikolov, T., and Bengio, Y.: On the difficulty of training recurrent neural networks, in: Proceedings of the 30th International Conference on Machine Learning, Atlanta, USA, 28, 21 November 2012, 1310–1318, https://doi.org/10.48550/arXiv.1211.5063, 2013.

Ren, W., Tian, H., Liu, M., Zhang, C., Chen, G., Pan, S., Felzer, B., and Xu, X.: Effects of tropospheric ozone pollution on net primary productivity and carbon storage in terrestrial ecosystems of China, J. Geophys. Res., 112, https://doi.org/10.1029/2007JD008521, 2007.

Sahu, S. K., Liu, S., Liu, S., Ding, D., and Xing, J.: Ozone pollution in China: background and transboundary contributions to ozone concentration & related health effects across the country, Sci. Total Environ., 761, 144131, https://doi.org/10.1016/j.scitotenv.2020.144131, 2021.

Shen, L., Jacob, D. J., Liu, X., Huang, G., Li, K., Liao, H., and Wang, T.: 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, https://doi.org/10.5194/acp-19-6551-2019, 2019.

Slaper, H., Velders, G. J., Daniel, J. S., de Gruijl, F. R., and van der Leun, J. C.:

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

Estimates of ozone depletion and skin cancer incidence to examine the Vienna Convention achievements, Nature, 384, 256–258, https://doi.org/10.1038/384256a0, 1996.

Sun, L., Xue, L., Wang, Y., Li, L., Lin, J., Ni, R., Yan, Y., Chen, L., Li, J., Zhang, Q., and Wang, W.: Impacts of meteorology and emissions on summertime surface ozone increases over central eastern China between 2003 and 2015, Atmos. Chem. Phys., 19, 1455–1469, https://doi.org/10.5194/acp-19-1455-2019, 2019.

Tian, H., Ren, W., Tao, B., Sun, G., Chappelka, A., Wang, X., Pan, S., Yang, J., Liu, J., S. felzer, B., M. melillo, J., and Reilly, J.: Climate extremes and ozone pollution: a growing threat to china's food security, Ecosystem Health and Sustainability, 2, e01203, https://doi.org/10.1002/ehs2.1203, 2016.

Trainer, M., Parrish, D. D., Goldan, P. D., Roberts, J., and Fehsenfeld, F. C.: Review of observation-based analysis of the regional factors influencing ozone concentrations, Atmos. Environ., 34, 2045–2061, https://doi.org/10.1016/S1352-2310(99)00459-8, 2000.

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., Yu, K., Miller, C. C., Yantosca, R. M., Sulprizio, M. P., Thompson, A. M., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Cohen, R. C., Laughner, J. L., Dibb, J. E., Hall, S. R., Ullmann, K., Wolfe, G. M., Pollack, I. B., Peischl, J., Neuman, J. A., and Zhou, X.: Why do models overestimate surface ozone in the Southeast United States?, Atmos. Chem. Phys., 16, 13561–13577, https://doi.org/10.5194/acp-16-13561-2016, 2016.

van der Leun, J., Tang, X., and Tevini, M.: Environmental effects of ozone depletion and its interactions with climate change: 2002 assessment, Photochem. Photobiol. Sci., 2, vii–vii, https://doi.org/10.1039/B211913G, 2003.

Wang, T., Xue, L., Brimblecombe, P., Lam, Y. F., Li, L., and Zhang, L.: Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects, Sci. Total Environ., 575, 1582–1596, https://doi.org/10.1016/j.scitotenv.2016.10.081, 2017a.

Wang, W.-N., Cheng, T.-H., Gu, X.-F., Chen, H., Guo, H., Wang, Y., Bao, F.-W., Shi,

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

S.-Y., Xu, B.-R., Zuo, X., Meng, C., and Zhang, X.-C.: Assessing spatial and temporal patterns of observed ground-level ozone in China, Sci. Rep., 7, 3651, https://doi.org/10.1038/s41598-017-03929-w, 2017b.

Wang, Y., Shen, L., Wu, S., Mickley, L., He, J., and Hao, J.: Sensitivity of surface ozone over China to 2000–2050 global changes of climate and emissions, Atmos. Environ., 75, 374–382, https://doi.org/10.1016/j.atmosenv.2013.04.045 , 2013.

Wei, J., Li, Z., Li, K., Dickerson, R., Pinker, R., Wang, J., Liu, X., Sun, L., Xue,W., and Cribb, M.: 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, 2022.

Wei, W., Fang, Y., and Zhou, Y.: Synoptic and meteorological drivers of regional ozone pollution events in China, Environ. Res. Commun., 3, 055004, https://doi.org/10.1088/2515-7620/abfe9c, 2021.

Xie, M., Shu, L., Wang, T., Liu, Q., Gao, D., Li, S., Zhuang, B., Han, L., Li, M., and Chen, P.: Natural emissions under future climate condition and their effects on surface ozone in the Yangtze River Delta region, China, Atmos. Environ., 150, 162–180, https://doi.org/10.1016/j.atmosenv.2016.11.053, 2017.

Xu, W., Zhao, C., Ran, L., Deng, Z., Liu, P., Ma, N., Lin, W., Xu, X., Yan, P., He, X., Yu, J., Liang, W., and Chen, L.: Characteristics of pollutants and their correlation to meteorological conditions at a suburban site in the North China Plain, Atmos. Chem. Phys., 11, 4353–4369, https://doi.org/10.5194/acp-11-4353-2011, 2011.

Xu, X.: Spatial distribution of GDP in China with km grid dataset, RESDC [data set], http://www.resdc.cn/DOI, 2017.

Xue, T., Zheng, Y., Geng, G., Xiao, Q., Meng, X., Wang, M., Li, X., Wu, N., Zhang, Q., and Zhu, T.: Estimating spatiotemporal variation in ambient ozone exposure during 2013–2017 using a data-fusion model, Environ. Sci. Technol., 54, 14877–14888, https://doi.org/10.1021/acs.est.0c03098, 2020.

Yan, Y., Zheng, H., Kong, S., Lin, J., Yao, L., Wu, F., Cheng, Y., Niu, Z., Zheng, S., Zeng, X., Yan, Q., Wu, J., Zheng, M., Liu, M., Ni, R., Chen, L., Chen, N., Xu, K., Liu, D., Zhao, D., Zhao, T., and Qi, S.: On the local anthropogenic source

https://doi.org/10.5194/essd-2022-428
Preprint. Discussion started: 21 December 2022

diversities and transboundary transport for urban agglomeration ozone mitigation, Atmos. Environ., 245, 11, https://doi.org/10.1016/j.atmosenv.2020.118005, 2021.

Zhan, Y., Luo, Y., Deng, X., Grieneisen, M.L., Zhang, M., and Di, B.: Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment, Environ. Pollut., 233, 464, https://doi.org/10.1016/j.envpol.2017.10.029, 2018.

Zhang, W., Liu, D., and Shi, H.: Hourly Surface Ozone data (HrSOD) across China during 2005–2020, Zenodo [data set], https://doi.org/10.5281/zenodo.7415326, 2022.

Zhang, X., Zhao, L., Cheng, M., and Chen, D.: Estimating ground-level ozone concentrations in eastern China using satellite-based precursors, IEEE Trans. Geosci. Remote Sens., 58, 4754–4763, https://doi.org/10.1109/TGRS.2020.2966780, 2020.

Zou, J., Lu, N., Jiang, H., Qin, J., Yao, L., Xin, Y., and Su, F.: Performance of air temperature from ERA5-Land reanalysis in coastal urban agglomeration of Southeast China, Sci. Total Environ., 828, 154459, https://doi.org/10.1016/j.scitotenv.2022.154459, 2022.