Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____   _____
Muwu Xu   Date

Using TROPOMI-Based Estimation of Daily Ozone Ground Levels to Assess the Impact of COVID-

19 on Ozone Concentrations in China

By

Muwu Xu

Master of Public Health

Epidemiology

_____

Venkat Narayan, PhD

Committee Chair

_____

Yang Liu, PhD

Committee Member

Using TROPOMI-Based Estimation of Daily Ozone Ground Levels to Assess the Impact of COVID-19 on Ozone Concentrations in China

By

Muwu Xu

MBBS

Sun Yat-sen University

2017

Thesis Committee Chair: Venkat Narayan, PhD

Committee Member: Yang Liu, PhD

An abstract of

a thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

Abstract

Using TROPOMI-Based Estimation of Daily Ozone Ground Levels to Assess the Impact of COVID-19 on Ozone Concentrations in China

By Muwu Xu

Background: While China's strict quarantine policy during the COVID-19 pandemic reduced the formation of ozone precursors from transportation and industrial sectors, ground observations have reported an increase in ozone levels in many Chinese cities. However, few studies have been able to evaluate the ozone level change in the entire country.

Methods and Materials: We developed a machine learning model using nadir ground level ozone column from the TROPOspheric Monitoring Instrument (TROPOMI), ozone profiles from the Ozone Monitoring Instrument (OMI), metrological parameters, and land use data to estimate full-coverage daily ground ozone concentrations across China at 0.05° spatial resolution.

Results: We built two separate models for the pandemic year (11/2019 – 04/2020) and the reference year (11/2018 – 04/2019), respectively. There were 209, 654 daily measurements from a total of 1, 500 AQS monitor during the study period. The out of bag R^2 was 86.7% in the reference year model and 90.06% in the pandemic year model. During the phase of lockdown in Covid-19 (Jan/23/2020-Feb/13/2020) , defined as high level quarantine phase, a significant increase of concentration of ozone took place comparing to the concentration in pre-lockdown episode (Jan/1/2020to Jan/22/2020) in China (95% CI: 9.80 μg/m3, 9.88 μg/m3; p < 0.0001).

Conclusion: Our study demonstrates the possibility and utilization of TROPOMI product for modelling Ozone at a fine spatial and temporal resolution, which will allow us for construction of long-term daily Ozone measurements at 5km^2 spatial resolution and support further epidemiological and environmental studies about ground Ozone.

Acknowledgement

I could not pursue my dream in America without the love and support from my parents and my friends. I am grateful to the love from my fellowship from my Christian church community and without the support I would not be able to overcome the difficulties and challenges in the journey.

I am grateful to Prof Yang Liu as a great advisor, who gave me a lot of direction and guidance since I step into the field of air pollution modeling and public health research. Working with him was a joy and a privilege, his knowledge, patience and hardworking gave me a lot of inspiration and support to build myself strong confidence and arsenal for the future research project. Meanwhile, his diligence, enthusiasm and so many other great features stimulate my ambition to be better and better.

Thanks to Prof Venkat Narayan to be my faculty thesis advisor and gave me many supports and encouragement to me to overcome the difficulties so that I could make it finally.

Thanks to my friend Qingyang Zhu, who gave me plenty of useful tips and resources to develop my coding skills, ArcGIS Pro tech and presenting skills.

To the faculty and staff of the Emory Rollins School of Public Health Department of Epidemiology and Environmental Health, thank you for the resources, lectures and support.

Table of Contents

1. Introduction:

Ground level Ozone (O_3) is nowadays a major air pollutant that influence the health of human beings a lot. Many adverse effects such as respiratory symptoms or cardiovascular dysfunction are attributed to high concentrations of ground level ozone [1]. It is reported in China that a 10 μg/m3 increase of the maximum 8-hour average concentration of O3 resulted in increases in percent change for non-accidental mortality, cardiovascular mortality and respiratory mortality of 0.42% (95%CI, 0.32-0.52%), 0.44% (95%CI, 0.17-0.70%) and 0.50% (95%CI, 0.22-0.77%), respectively *[2]*. Studies about air pollutions in China remain mostly to PM2.5, NO2 and somehow are only conducted in certain regions[3-6]. Moreover, current studies on ozone predictions in China either don't have a full profile of ground level ozone or not very accuracy and high resolution[7, 8].

During the serious period of 2019 novel coronavirus (COVID-19), nationwide quarantines and restrictions on the movement of population are implemented by China to stop COVID-19 from spreading after the Chinese New Year of 2020, which leads to large reduction of industrial and transportation activities associated to the ozone emissions[9-11]. According to previous research on the formation and fate of ozone, these activities were essential and usually good for ozone formation[12, 13]. The COVID-19 pandemic in China allowed us to investigate the association between industrial and transportation emissions and ground level ozone concentration in a very large scale, considering almost all the cities and people reduced their level of activities that could increase the ozone[9, 14]. Considering the lockdown policies implementation time point, we choose 4 different phases in out study as for lockdown phases and reference phases: pre-lockdown phase: Jan/1/2020 to Jan/22/2020; lockdown phase:

Jan/23/2020 to Feb/13/2020; 2019 Lunar calendar reference: Feb/4/2019 to Feb/25/2019;

2019 Gregorian calendar reference: Jan/23/2019 to Feb/13/2019.

Because O3 is an important air pollutant as it connects air quality and public health, it draws a

growing attention from environmental science community[15-17]. However, it is very difficult

to catch comprehensive ground ozone due to the lack of enough deployment of ground

monitors. There is only a few research and publication for China[18], using a monitoring

network covering major cities across whole country which was completed in 2012. However,

due to the sparse monitor located merely on major cities, the resolution and

comprehensiveness of the ground level ozone are large limitations for directly utilizing ground-

level air monitoring data. Recently, the satellite remote sensing techniques are a widely used

tools and provide useful way to estimate and investigate the ground-level air pollutants[15, 19,

20]. For ozone pollution, remote sensing data have been utilized also, however, previous study

for China ozone by the use of space-based Ozone Monitoring Instrument data doesn't give a

great result of estimation with the correlation coefficient ranging from 0.3 to 0.6 for sites in

southern China and ranging from 0.1 to 0.3 for sites in northern China during the summertime

in daily level from 2013 to 2017[21]. In our study, we try to discover some new remote sensing

data source to have a better retrieve of ground level ozone. In our study, TROPOMI turns out to

be a very useful data source showing a great result of estimation, with a R^2 of around 0.8. And

because of its high resolution, a ground level ozone prediction map is built at a resolution of

0.5°, which by far is one of the finest prediction maps of ground level ozone across all China.

In this analysis, our model implemented elevation, land-use type, atmospheric chemistry components and meteorological factors using machine learning technique to estimate the daily surface ozone at a spatial resolution of 0.5°. With these models, we can evaluate the actual change of ground level-ozone in the period of COVID-19 and the reference period of the last year to see how the COVID-19 policies of restrictions change the ground level ozone in different regions.

2.Data and methods:

2.1 Study Design

In our study, monthly average MDA8 ozone concentrations in China were predicted from 2018/11/01 to 2019/04/30 and 2019/11/01 to 2020/04/30 using satellite remote sensing data. The workflow design of our study is shown in Fig.1. At first, the fraction of the ground-layer ozone column were extracted from the SAO OMI ozone profile. And we imputed the missing values by daily random forest models with the application of MERRA-2 meteorological fields and other fields like surface flux. Then the full coverage ground level column ozone could be got by multiplying the gap-filled fraction of the ground-layer ozone with the TROPOMI total column ozone. Then we could get the final daily average MDA8 ozone predictions with the application of another random forest model, which is trained with multiple features, such as the TROPOMI full profile ground-level ozone, land-use terms, population, meteorological fields, and chemistry fields.

2.2 Datasets:

2.2.1 Ozone monitoring data

The study domain covers mainland China, Hong Kong special administrative region, and Taiwan.

The China map is made through ArcGIS Pro and the province outlines is downloaded from

Resource and Environment Science and Data Center[22]. Then a 0.5° modeling grid was

constructed covering the study domain for data integration.

Ozone measurement: Concentrations of O3 at around ~ 1590 stations are from air monitoring

data center of China. Measurements are published by the China National Environmental

Monitoring Center (http://www.cnemc.cn/). We calculated the maximum daily 8-hour average

(MDA8) from the measurements. Then MDA8 Ozone measurements from stations located in

the same grid cell were averaged, and there were 1232 grid cells with Ozone measurement

totally.

2.2.2 SAO OMI ozone profile

The Smithsonian Astrophysical Observatory (SAO) Ozone Monitoring Instrument (OMI) using

the NASA Aura satellite is a nadir viewing spectrometer and its ozone profiles is from the 0.22

to 261 hPa and so can be used to estimate the ground layer fraction ozone layer[23].

2.2.3 TROPOMI data:

The TROPOspheric Monitoring Instrument (TROPOMI) is a nadir-viewing spectrometer data

from sentinel-5p satellites at a resolution of 0.07° * 0.035° resolution, which can be

downloaded from Goddard Earth Sciences Data and Information Services Center[24]. To have a

full ground ozone profile, we use the gap-filled ground layer fraction from the SAO OMI ozone profile and then use inverse distance weighting (IDW) technique to impute the missing values of TROPOMI ground layer ozone.

2.2.4 Meteorological variables

Meteorological parameters in our study period were from the Goddard Earth Observing System Data Assimilation System Forward Processing (GEOS-FP) at resolution of 0.25° latitude * 0.3125° longitude resolution[25]. Then they are upscaled to 0.5° grid cell by inverse distance weighting (IDW)[26].

2.2.5 Aerosol chemistry

Trace chemistry and gases data are retrieved from the Goddard Earth Observing System Data Assimilation System GEOS-5 Composition Forecasting systems (GEOS-CF)[27]. We use the time-averaged two-dimensional collections (tavg) as our input data sources, whose resolution is 0.25°. Also, we regrided the data into 0.5° grid cell by inverse distance weighting (IDW). Among the different chemistry and gases components, we selected C2H6, CO, HCHO, ISOP, NO2, NO and BENZ, according to previous research discussed the precursors by P.S.Monks et.al[12].

2.2.6 Land use, population density, road length and digital elevation

As we know, the parameters, such as land use, population density, road length and digital elevation could affect the emission, transportation, and chemistry reaction. Below are these data sources.

We grabbed the land use data from the European Space Agency (ESA) Climate Change Initiative Global Land Cover Maps Version 2.0.7., which contains 23 types of land covering based on satellite observations at a spatial resolution of 300m.

LandScan global population data were grabbed from the Oak Ridge National Laboratory. The spatial resolution of the dataset is 30 arc-seconds and we choose the 2018 version as the estimate of the population density[28].

Also, from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM), version 3, we grabbed the 30-m elevation data, which is the latest version with an increased accuracy comparing to previous versions[29].

2.3 Data integration

0.05° grid cells were created across whole china for integration. The input features were resampled to the cells by a few methods. At first, IDW was used to interpolate and resample data which spatial resolutions are rougher than 0.05. These data included ground-level column TROPOMI ozone, GEOS-FP meteorological fields, GEOS-CF chemistry fields. For the land use data, which is categorical, we further categorized them into 7 categories (cropland, urban, vegetation, water, bare, wetland and ice) and then made a calculation of the percentage coverage of these categories within each cell. Moreover, the average elevation, road length, and population density were calculated and integrated within each cell. After integrating the feature data, the training dataset were built by selecting grid cells that contained the air quality monitoring stations.

2.4 Model construction and validation

We trained two separate random forest models within the same set of parameters for the two study phases (ref: 2018/11/01 to 2019/04/30; covid: 2019/11/01 to 2020/04/30), because we believe that because of the COVID 19 quarantine, the sources of the ground-level ozone are shifted and so the rank of importance of different input features would be changed. The detailed list and the ranking of the variables in the random forest models could be checked in Figure 4.

For the validation, a 10-fold cross-validation method was used. Randomly, the original dataset was divided into ten equal fold subsets. Each time we used 9 folds of them to train the model and make prediction on the one left. After 10 times, each fold would have a prediction value corresponding to the original ground-level ozone concentration. Then we could calculate the R^2 and root-mean-square error (RMSE) out of them, which could be used to evaluate the model performance. Furthermore, a spatial 10-fold CV was used to test the generalizability of our model prediction in different regions. Likely, the original dataset was divided randomly but based on the geospatial location of the data points. On such condition, the predictions on a given site would be output by a trained model from other locations.

R version 4.0.3 was used to process and perform statistical analysis. Package 'randomForest' and 'ranger' were utilized for training and testing the random forest models.

2.4.1 Random forest model

A random forest (RF) model was used to fit 26 predictors to over 1500 ground measurements. There are several reasons we choose RF model: 1. It has high accuracy in learning and

classifying features; 2. We can put a large number of input variables as predictors; 3. It will output the variable importance so we could tell what is of high ranking in our input predictors. As we know, random forest is a supervised machine learning model involving a bagged classifier built on decision tree that proposed by Breiman[30].By averaging a set of decision tress with a subset of, the RF model will give you the best predictions. In the process, a random subset of samples will be selected from all observations with replacement, and gradually RF model will select the best set of predictors that provides the best split at each node. There are two main parameters for RF model, which are the number of predictors sampled for each node (m_{try}) and the number of tress to be averaged (n_{tree}). The parameters we used in our model are m_{try} 8 and n_{tree} 500 after a comparison of results with different settings to achieve the best prediction accuracy. The 26 predictors used in our model includes are shown in Figure 4.

3. Results and discussion:

3.1 Descriptive statistics of the training dataset:

After gap-filling for the TROPOMI data, only 9383 out of 222392 (4%) site-days samples and 2838 out of 243043 (1%) site-days samples of near ground TROPOMI for 2018-2019 and 2019-2020 model are missing, respectively. For ground monitor: 3494/222392 (1%) site-days samples and 13480/243043 (5%) site-days samples of ground monitor MDA8 ozone, in 2018-2019 and 2019-2020 are missing, respectively. The mean concentrations ($\mu g\ m^{-3}$) of MDA-8 of ozone in 2018-2019 and 2019-2020 are 72.82 and 79.16, respectively.

3.2 Random forest model performance and cross-validation:

From figure 3 and table 1, we can see the CV results of the random forest models at datily level

generated for 2018-2019 and 2019-2020 respectively, consisting of the R^2 and RMSE between

the estimates and observations. The random forest models' out-of-bag R2 are 86.6% and

90.0%, showing a relatively good agreements between the predictions and the observations,

respectively. Though their performance were different, we can assume they both had a pretty

good performance comparing to models in previous report [31], the reasons are mostly

because of the improvement of the data quality of ground monitors and the improvement of

the data quality of our predictors.

Our general CV performance of our models were basically same comparing to their out-of-

bags'. In terms of spatial CV, the results are still comparable, with only a 1.9 and 1.7 slight

decrease comparing to the correspoinding out-of-bags' R^2, indicating that our models' stability

and efficiency in terms of spatial variability capturing ability of the surface ozone

concentrations, which will be a great help for us to detect the different spatial patterns in the

phases of lockdown in different regions across China.

3.3 Importance ranking of model predictors

Figure 4 shows us the predictors' importance rankings used in models development accroding

to their percent change of mean square error (MSE) for each year, respectively. From the

ranking plot, we can tell that the TROPOMI ozone profile played an important roles in the two

models with a high importance rankings. Meterological fields, such as wind component (U,V)

surface incoming shortwave flux (SWGDN) and surface pressure (PS), were among the highest

ranking predictors. The VOC components, such as C2H6, HCHO, CO and ISOP, were also among

the high important predictors. These predicators of high ranking are related to the production and transition of ozone[12]. There are some difference between the two models considering the importrance rankings (figure S1), in 2018-2019 model, the wind components, TROPOMI ozone profile, C2H6 and surface pressure are among the top 5 important. On the other hand, in the 2019-2020 model, the CO, C2H6, surface incoming shortwave flux, wind component and air temperature are among the highest top 5. However, for the overall ranking plot, the difference is generally slight and the performance of each model is good. Among the top 15 important predictors in different models, only one predictor was different, which is the cropland land cover type ranked 10th in 2018-2019 vs urban land cover type ranked 14th in 2019-2020.

3.4 Spatial and temporal trends of O3 predictions during COVID:

To control the spread of the COVID-19 virus among human beings, a major preventive lockdown was firstly implemented on 23 January in Wuhan, Hubei and the other provinces and cities followed the similar regulation for at least 3 weeks[32]. Thus, we set the period from 23 January to 13 February 2020 as lock down period (LD), which consisted of a 7-day national holiday (Spring Holiday), during which previous articles have reported that reductions happened as in anthropogenic emissions in eastern china regions[33]. There are three reference period we set for comparing. Firstly, the previous 3 weeks before the lock down (pre-LD); secondly, the same time periods as LD in 2018-2019 in the Chinese lunar calendar (LNY); thirdly, the same time periods as LD in 2018-2019 in the Gregorian calendar (ref). In the national scale, we could see that there are a nonnegligible amount of change in ground-level ozone concentration between the After-quarantine and Before-quarantine (Figure 6). In the high population density, while a notable decrease happened in Pearl River Delta, the other

regions, like the North China Plain, Yangtze River Delta and Wuhan city are all having an increasing level of ground-level ozone concentrations after quarantine (Table 2).

There are four regions or cities (NCP: North China Plain; YRD: Yangtze River Delta; Wuhan; PRD: Pearl River Delta) were selected to focus to tell how the ozone were changed and why such changes happened in terms of the aerosol chemical, meteorological and other factors. Calculation of the mean values of ozone concentrations in the four different periods among the 4 regions we selected to study the COVID-19 lock down effect were shown as figure 5. For Wuhan, which is the first city of locking down in China, the ozone concentration after quarantine was the higher than 3 weeks means before the LD period as well as the means concentration of the reference periods in the last year. Particularly, ozone was increased by $24.61 \mu g/m^3$ (48.62%), $19.38 \mu g/m^3$ (34.70%) and $15.73 \mu g/m^3$ (26.44%) compared with the Before-quarantine, the Lunar Calendar reference (LNY) and Gregorian Calendar reference (ref). Similarly, Yangtze River Delta and NCP showed the same changing patterns where the ground-level ozone concentrations after quarantine were the highest comparing to the other three reference periods (Before-quarantine, Ref Lunar, Ref Gregorian), where the differences are $19.56 \mu g/m^3$ (33.82%), $7.24 \mu g/m^3$ (10.32%) and $7.80 \mu g/m^3$ (11.35%) for YRD and $20.86 \mu g/m^3$ (36.00%), $2.48 \mu g/m^3$ (3.25%) and $10.58 \mu g/m^3$ (15.51%). However, in Pearl River Delta, the ground-level ozone decreased after quarantine comparing to the ozone level before quarantine ($70.49 \mu g/m^3$ vs $84.88 \mu g/m^3$), showing another change pattern comparing to the other three northern regions.

The different patterns of ground-level ozone changes were possibly due to following reasons: 1. During the quarantine, in the Northern China, the transportation is limited and thus the combustion and the resulting NOx production became dramatically less than usual, the decrease of NOx is much more significant than the decrease of the VOCs, which means the titration effect of the NOx on ozone impaired and thus the ground-level ozone increased[34]. 2. In the Southern China, because the temperature was higher and the plant area were larger, so the biological sources of VOCs production was great and kept the ground-level ozone in a relative high level[35]. 3. In the Northern China high population density area, due to the quarantine happened in the time of Chinese Spring Festival when family gathered and a lot of biomass fuel, such as coal or wood were burnt, which may also increase the level of VOCs and kept the ground-level ozone in a high level.

3.5 Strengths and Limitations

Our Ozone model is the first advanced model during the COVID-19 period in China to incorporate both TROPOMI satellite remote sensing data and other important features to provide daily ground measurement at 5-km^2 resolution, to help examine the ozone pollution level and aid in epidemiology studies in the future. A very important major of the study is that it provides a high resolution of 5-km^2 through whole china by the application of TROPOMI gap-filled Ozone profile data. It is able to seize the spatial temporal change trends to further investigate epidemiological health studies requiring daily measurements, especially the investigation about whether the ozone concentration would help fight against the COVID 19 virus[36]. Moreover, an emerging ensemble classifier, the random forest model, is used to generate the estimates with a very high accuracy, which is very inspiring and tell a story that possible the machine learning approaches or more cutting edges models could be very useful in the air pollution assessment in the future.

Also, our work was subject to several limitations. 1. Our study period is winter and spring where the concentration of ground-level ozone is relatively low and not thought as a regular air pollution. To improve that, we could further use the model to investigate the whole year range of ground-level ozone concentrations in China to see whether the model is still reliable. 2. The study can't tell precisely which input features or sources were essential to the ozone production during the different study periods we chose. Although the model performances are pretty trust-worthy, the mechanisms and pathways of the ground-level ozone concentrations were unclear and request further efforts and investigations to understand.

4.Figure

Figure 1.

Figure 2.

Figure 3.

Random Forest 2018-2019
CV

N=224078
R^2=0.87
RMSE=12.56μg/m^3
y=1.06x-4.5

Random Forest 2019-2020
CV

N=242340
R^2=0.90
RMSE=10.96μg/m^3
y=1.05x-3.9
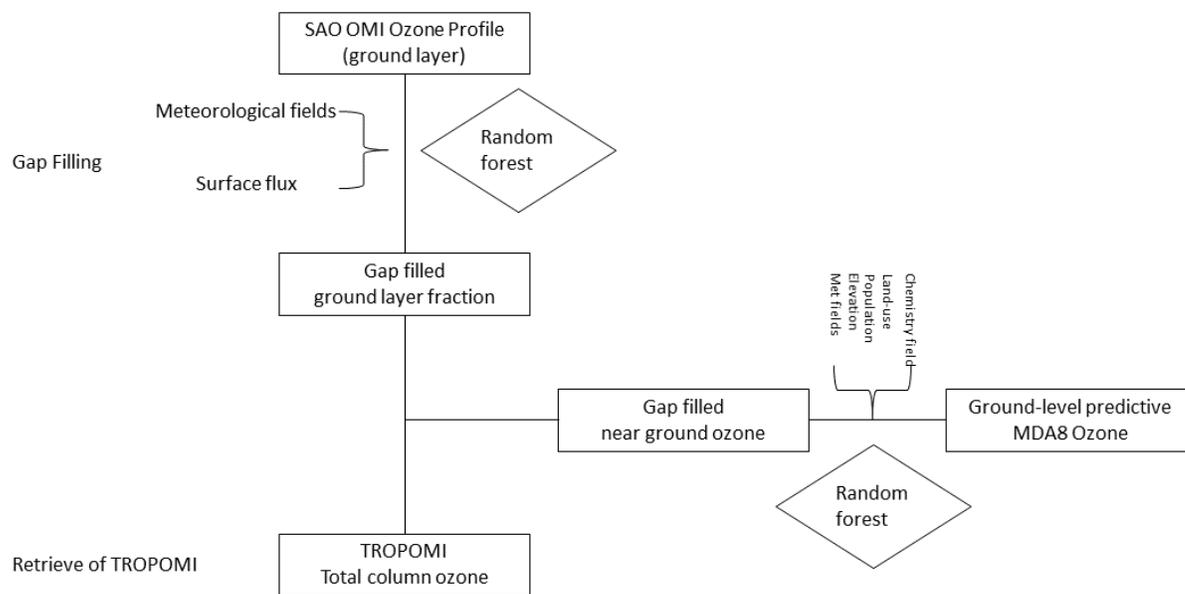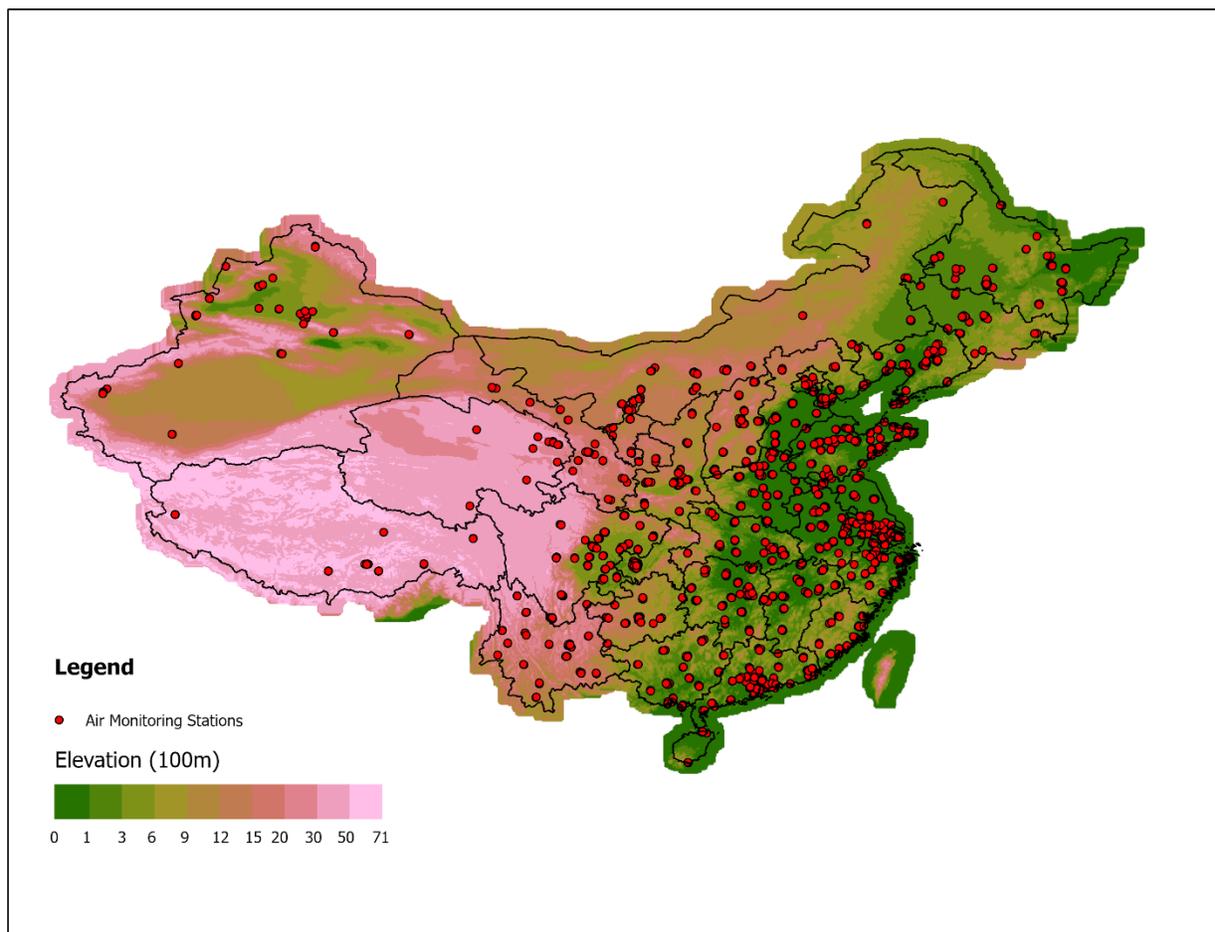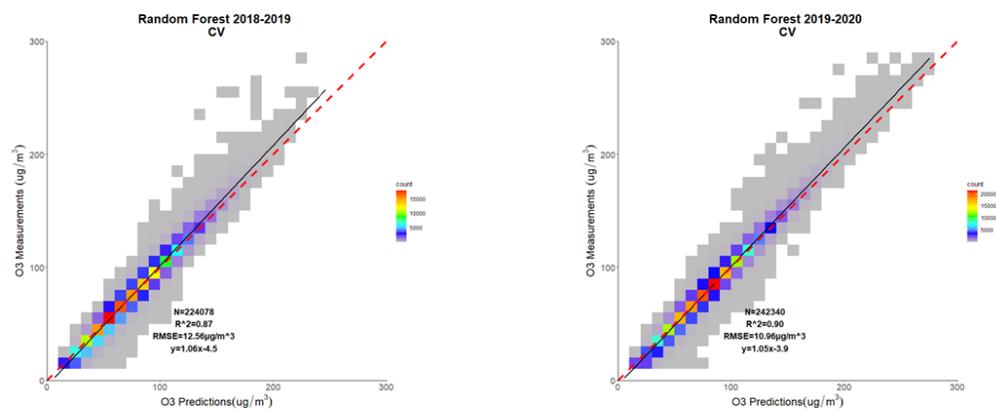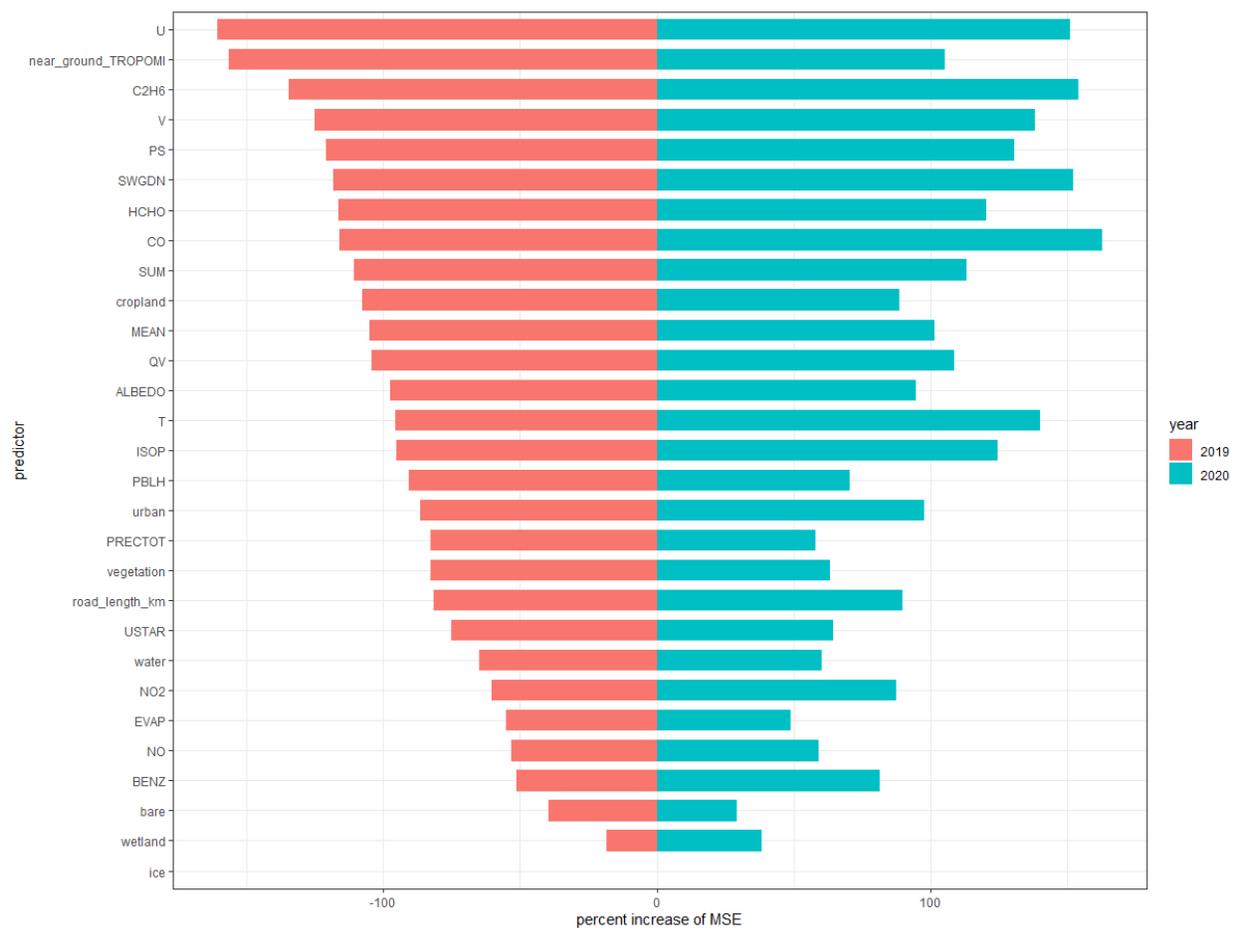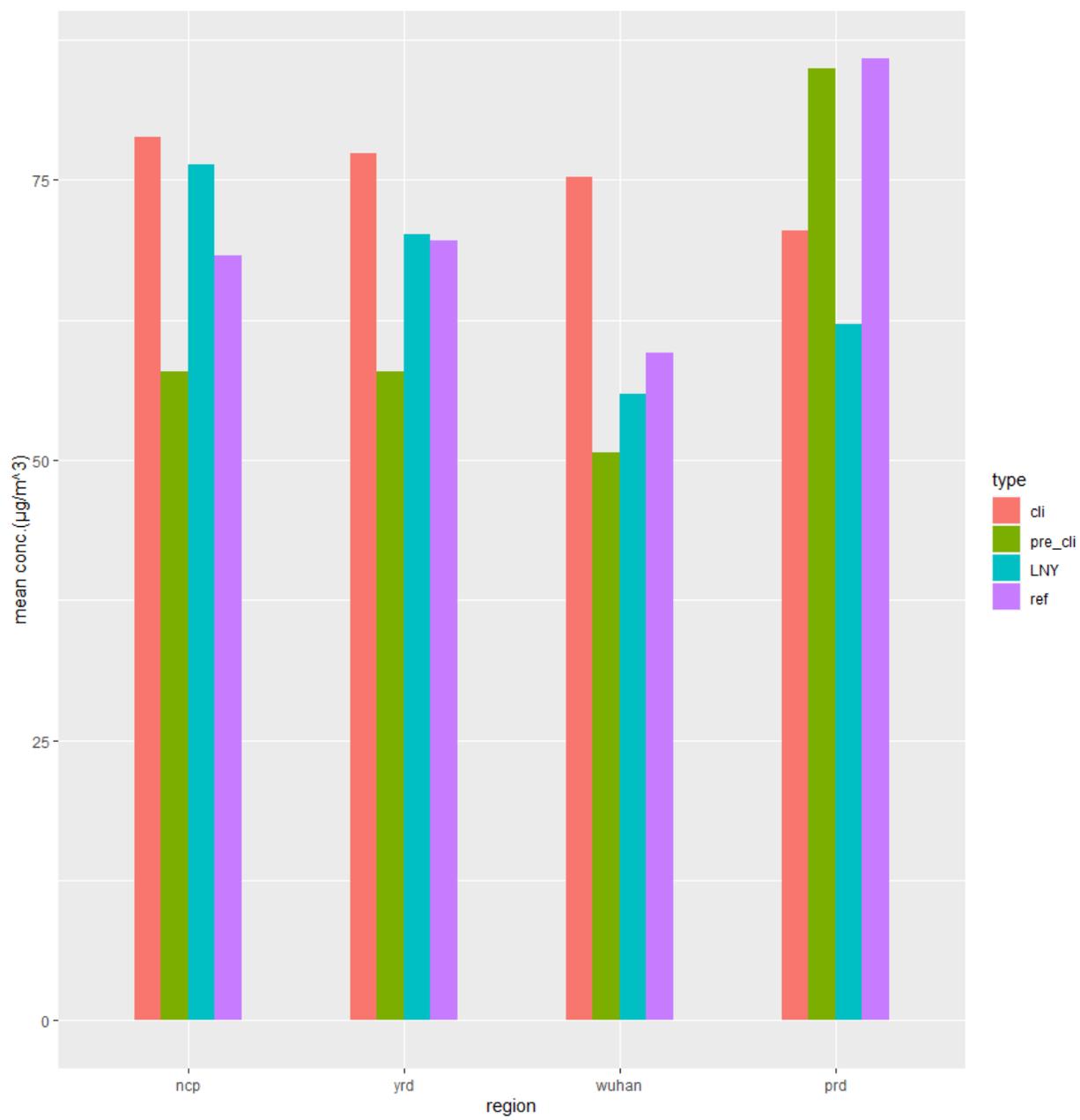
Figure 4.

Figure 5.

Figure 6.

Average Concentration of Ozone after Quarantine

Chinese 9 line dash
— Chinese 9 line dash

Province Border
▭ Province Border

Ozone Conc. (ug/m³)
- ≤20
- ≤35
- ≤40
- ≤50
- ≤55
- ≤60
- ≤65
- ≤70
- ≤75
- ≤80
- ≤85
- ≤90
- ≤105
- ≤120

0   330   660 Km

Average Concentration of Ozone before Quarantine

Chinese 9 line dash
— Chinese 9 line dash

Province Border
▭ Province Border

Ozone Conc. (ug/m³)
- ≤20
- ≤35
- ≤50
- ≤55
- ≤60
- ≤65
- ≤70
- ≤75
- ≤80
- ≤85
- ≤90
- ≤105
- ≤120

0   330   660 Km

Average Concentration Change of Ozone between After-quarantine
and Before-quarantine

N

Chinese 9 line dash
—— Chinese 9 line dash

Province Border
☐ Province Border

average
- ≤-20
- ≤-10
- ≤0
- ≤5
- ≤10
- ≤15
- ≤20
- ≤25
- ≤30
- ≤40
- ≤50
- ≤60

0   330   660 Km
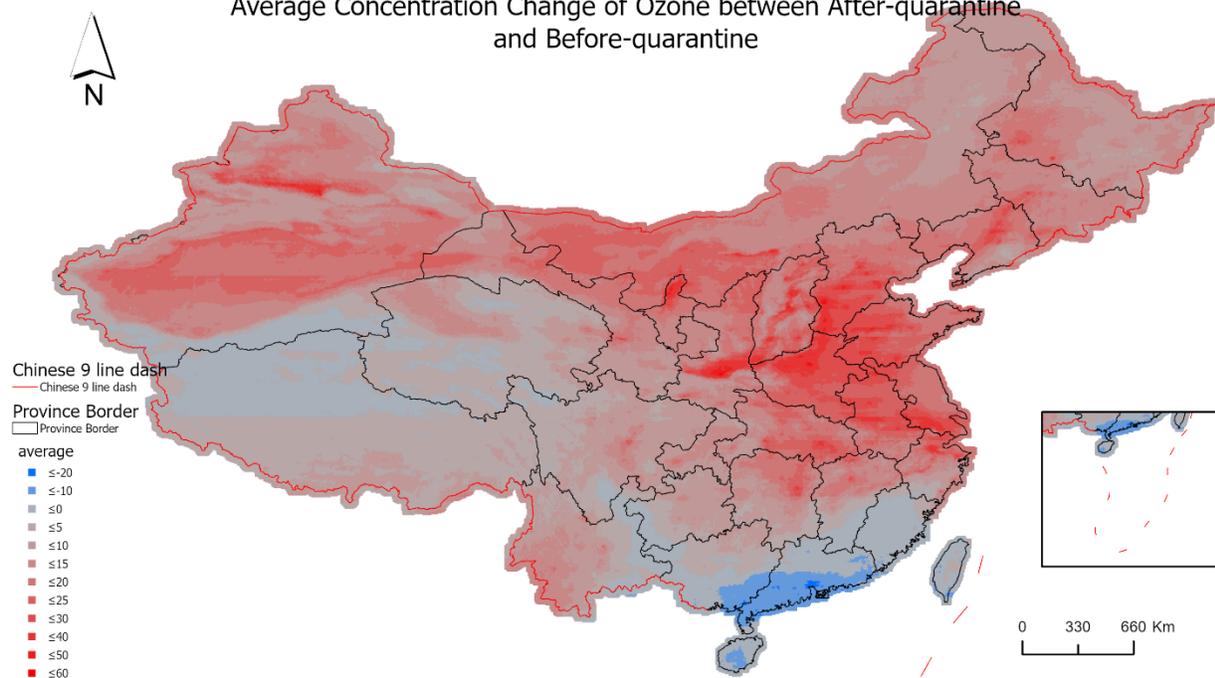
5.Table

Table 1.

|	Out-of-bag		General CV		Spatial CV	
Year	R^2	RMSE	R^2	RMSE	R^2	RMSE
2018-2019	86.6	12.56	86.6	12.78	84.7	13.57
2019-2020	90.0	10.97	89.9	11.20	88.2	12.06

Table 1. R^2 and RMSE (μg m^{-3}) values of out-of-bag predictions, general CV, spatial CV for 2018-2019 and 2019-2020 model.

Table 2.

Ozone(μg/m³) / Region	After-quarantine	Before-quarantine	Ref Lunar	Ref Gregorian
NCP	78.79	57.93	76.31	68.21
YRD	77.39	57.83	70.15	69.59
Wuhan	75.23	50.62	55.85	59.50
PRD	70.49	84.88	62.09	85.76

Supplemental:

Figure S1:

rf_2019

rf_2020
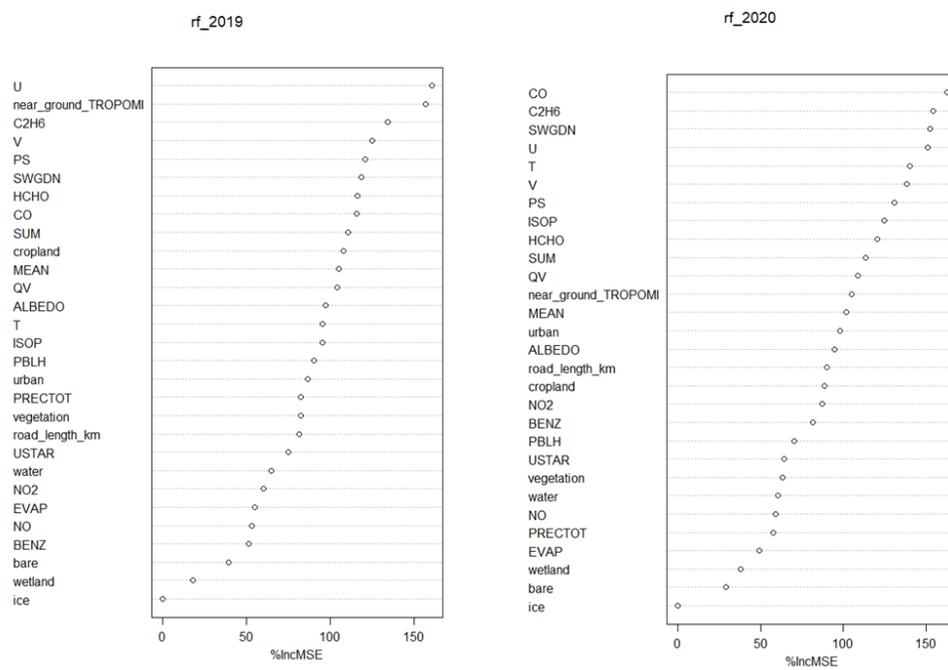
1.   Ji, M., D.S. Cohan, and M.L. Bell, *Meta-analysis of the Association between Short-Term Exposure to Ambient Ozone and Respiratory Hospital Admissions.* Environ Res Lett, 2011. **6**(2).

2.   Yan, M., et al., *Meta-analysis of the Chinese studies of the association between ambient ozone and mortality.* Chemosphere, 2013. **93**(6): p. 899-905.

3.   Si, Y., et al., *Assessment of satellite-estimated near-surface sulfate and nitrate concentrations and their precursor emissions over China from 2006 to 2014.* Sci Total Environ, 2019. **669**: p. 362-376.

4.   Tian, X., et al., *Long-term observations of tropospheric NO2, SO2 and HCHO by MAX-DOAS in Yangtze River Delta area, China.* J Environ Sci (China), 2018. **71**: p. 207-221.

5.   Zhong, M., F. Chen, and E. Saikawa, *Sensitivity of projected PM2.5- and O3-related health impacts to model inputs: A case study in mainland China.* Environ Int, 2019. **123**: p. 256-264.

6.   Zhu, Y., et al., *Spatiotemporally mapping of the relationship between NO2 pollution and urbanization for a megacity in Southwest China during 2005-2016.* Chemosphere, 2019. **220**: p. 155-162.

7.   Zhang, W., et al., *Quantification of ozone exposure- and stomatal uptake-yield response relationships for soybean in Northeast China.* Sci Total Environ, 2017. **599-600**: p. 710-720.

8.   Yang, G., Y. Liu, and X. Li, *Spatiotemporal distribution of ground-level ozone in China at a city level.* Sci Rep, 2020. **10**(1): p. 7229.

9.   Shen, J., et al., *Prevention and control of COVID-19 in public transportation: Experience from China.* Environ Pollut, 2020. **266**(Pt 2): p. 115291.

10.   Shen, Y., et al., *Community Outbreak Investigation of SARS-CoV-2 Transmission Among Bus Riders in Eastern China.* JAMA Intern Med, 2020. **180**(12): p. 1665-1671.

11. Chen, K., et al., *Air pollution reduction and mortality benefit during the COVID-19 outbreak in China.* Lancet Planet Health, 2020. **4**(6): p. e210-e212.

12. Monks, P.S., et al., *Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer.* Atmos. Chem. Phys., 2015. **15**(15): p. 8889-8973.

13. Zhang, J.J., Y. Wei, and Z. Fang, *Ozone Pollution: A Major Health Hazard Worldwide.* Front Immunol, 2019. **10**: p. 2518.

14. Cui, Q., et al., *The impacts of COVID-19 pandemic on China's transport sectors based on the CGE model coupled with a decomposition analysis approach.* Transp Policy (Oxf), 2021. **103**: p. 103-115.

15. Requia, W.J., et al., *An Ensemble Learning Approach for Estimating High Spatiotemporal Resolution of Ground-Level Ozone in the Contiguous United States.* Environ Sci Technol, 2020. **54**(18): p. 11037-11047.

16. Paulin, L.M., et al., *Association of Long-term Ambient Ozone Exposure With Respiratory Morbidity in Smokers.* JAMA Intern Med, 2019.

17. Wang, M., et al., *Association Between Long-term Exposure to Ambient Air Pollution and Change in Quantitatively Assessed Emphysema and Lung Function.* JAMA, 2019. **322**(6): p. 546-556.

18. Wang, W.N., et al., *Assessing Spatial and Temporal Patterns of Observed Ground-level Ozone in China.* Sci Rep, 2017. **7**(1): p. 3651.

19. Vu, B.N., et al., *Developing an Advanced PM2.5 Exposure Model in Lima, Peru.* Remote Sens (Basel), 2019. **11**(6).

20. Zhu, Q., et al., *Predicting gestational personal exposure to PM2.5 from satellite-driven ambient concentrations in Shanghai.* Chemosphere, 2019. **233**: p. 452-461.

21.   Shen, L., et al., *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., 2019. **19**(9): p. 6551-6560.

22.   *Resource and Environment Science and Data Center*. Available from: http://www.resdc.cn/DataSearch.aspx.

23.   Huang, G., et al., *Validation of 10-year SAO OMI ozone profile (PROFOZ) product using Aura MLS measurements.* Atmos. Meas. Tech., 2018. **11**(1): p. 17-32.

24.   *Goddard Earth Sciences Data and Information Services Center*. Available from: https://disc.gsfc.nasa.gov/.

25.   *Goddard Earth Observing System Data Assimilation System Forward Processing.*

26.   Lu, G.Y. and D.W. Wong, *An adaptive inverse-distance weighting spatial interpolation technique.* Computers & Geosciences, 2008. **34**(9): p. 1044-1055.

27.   *Goddard Earth Observing System Data Assimilation System GEOS-5 Composition Forecasting systems (GEOS-CF). .* Available from: https://gmao.gsfc.nasa.gov/weather_prediction/.

28.   *LandScan Global Population Database.*

29.   *Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM).* Available from: https://asterweb.jpl.nasa.gov/.

30.   Breiman, L., *Random Forests.* Machine Learning, 2001. **45**(1): p. 5-32.

31.   Liu, R., et al., *Spatiotemporal distributions of surface ozone levels in China from 2005 to 2017: A machine learning approach.* Environ Int, 2020. **142**: p. 105823.

32.   Le, T., et al., *Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China.* Science, 2020. **369**(6504): p. 702-706.

33.   Gong, D.-Y., et al., *Observed holiday aerosol reduction and temperature cooling over East Asia.* Journal of Geophysical Research: Atmospheres, 2014. **119**(11): p. 6306-6324.

34.   Di, Q., et al., *Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging.* Environ Sci Technol, 2020. **54**(3): p. 1372-1384.

35.   Fitzky, A.C., et al., *The Interplay Between Ozone and Urban Vegetation—BVOC Emissions, Ozone Deposition, and Tree Ecophysiology.* Frontiers in Forests and Global Change, 2019. **2**(50).

36.   Blanco, A., et al., *Ozone potential to fight against SAR-COV-2 pandemic: facts and research needs.* Environ Sci Pollut Res Int, 2021.