The drivers and health risks of the unexpected surface ozone enhancements over the Sichuan basin, China in 2020

Youwen Sun[1,2], Hao Yin[1,2,†], Xiao Lu[3,†], Justus Notholt[4], Mathias Palm[4], Cheng Liu[2], Yuan Tian[5], and Bo Zheng[6]

[1]*Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine*
Mechanics, HFIPS, Chinese Academy of Sciences, Hefei 230031, China

[2]*Key Laboratory of Precision Scientific Instrumentation of Anhui Higher Education Institutes,*
University of Science and Technology of China, Hefei, 230026, China

[3]*School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, 519082, China*

[4]*University of Bremen, Institute of Environmental Physics, P. O. Box 330440, 28334 Bremen,*
Germany

[5]*Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601,*
China

[6]*Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School,*
Tsinghua University, Shenzhen 518055, China

[†]Corresponding authors.

E-mail addresses: Hao Yin (yhyh95@mail.ustc.edu.cn) and Xiao Lu (luxiao25@mail.sysu.edu.cn)

Abstract

After a continuous increase in surface ozone (O_3) level from 2013 to 2019, the overall summertime O_3 concentration across China showed a significant reduction in 2020. In contrast to this overall reduction in surface O_3 across China, unexpected surface O_3 enhancements of 10.2 ± 0.8 ppbv (23.4%) were observed in May-June 2020 vs. 2019 over the Sichuan basin (SCB), China. In this study, we use high resolution nested-grid GEOS-Chem simulation, the eXtreme Gradient Boosting (XGBoost) machine learning method and the exposure−response relationship to determine the drivers and evaluate the health risks of the unexpected surface O_3 enhancements. We first use the XGBoost machine learning method to correct the GEOS-Chem model-to-measurement O_3 discrepancy over the SCB. The relative contributions of meteorology and anthropogenic emissions changes to the unexpected surface O_3 enhancements are then quantified with the combination of GEOS-Chem and XGBoost models. In order to assess the health risks caused by the unexpected O_3 enhancements over the SCB, total premature death mortalities are estimated. The results show that changes in anthropogenic emissions caused 0.9 ± 0.1 ppbv of O_3 reduction and changes in meteorology caused 11.1 ± 0.7 ppbv of O_3 increase in May-June 2020 vs. 2019. The meteorology-induced surface O_3 increase is mainly attributed to significant increases in temperature and downward potential vorticity, and decreases in precipitation, specific humidity and cloud fractions over the SCB and surrounding regions in May-June 2020 vs. 2019. These changes in meteorology combined with the complex basin effect enhance downward transport of O_3 from upper troposphere, enhance biogenic emissions of volatile organic compounds (VOCs) and nitrogen oxides (NO_x), speed up O_3 chemical production, and inhabit the ventilation of O_3 and its precursors, and therefore account for the surface O_3 enhancements over the SCB. The total premature mortality due to the unexpected surface O_3 enhancements over the SCB has increased by 89.8% in May-June 2020 vs. 2019.

1   Keywords: Ozone; Health risk; Emissions; Meteorology; Chemical model; Machine learning

2   **1. Introduction**

3   Surface ozone (O_3) is largely generated from its local anthropogenic (fossil fuel and biofuel

4   combustions) and natural (biomass burning (BB), lightning, and biogenic emissions) precursors

5   such as volatile organic compounds (VOCs), nitrogen oxides (NO_x), and carbon monoxide (CO) via

6   a chain of photochemical reactions (Cooper 2019; Sun et al., 2018). Additional portion of surface

7   O_3 is transported from far away regions or from stratosphere (Wang et al., 2020; Akimoto et al.,

8   2015). Surface O_3 is one of the most harmful air pollutants that threatens human health and corps

9   production (Van Dingenen et al., 2019; Lu et al., 2020; Sun et al., 2018; Fleming et al., 2018).

10   Exposure to ambient O_3 pollution evokes a series of health risks including stroke, respiratory disease

11   (RD), hypertension, cardiovascular disease (CVD), and chronic obstructive pulmonary disease

12   (COPD) (Brauer et al., 2016; Liu et al., 2018; Lelieveld et al., 2013; Li et al., 2015; Wang et al.,

13   2020; Lu et al., 2020). Lu et al. (2020) estimated that the premature RD mortalities attributable to

14   ambient O_3 exposure in 69 Chinese cities in 2019 reached up to 64,370.

15   Surface O_3 variability is sensitive to both emissions and meteorological changes (Liu et al.,

16   2020a; Liu et al., 2020b; Lu et al., 2019a). Meteorological conditions affect surface O_3 variability

17   indirectly through changes in natural emissions of its precursors or directly via changes in wet and

18   dry removal, dilution, chemical reaction rates, and transport flux (Lu et al., 2019b; Li et al., 2019a;

19   Lin et al., 2008; Liu et al., 2020a). A reduction in temperature can lessen O_3 production by slowing

20   down the chemical reaction rates (Lee et al., 2014; Fu et al., 2015; Liu et al., 2020a) or reducing the

21   biogenic VOCs and NO_x emissions (Guenther et al., 2006; Tarvainen et al., 2005; Im et al., 2011).

22   A dryer meteorological condition can result in an increase in surface O_3 level (Kalabokas et al.,

23   2015; He et al., 2017; Liu et al., 2020a). Depending which process dominates the influence of

24   planetary boundary layer height (PBLH) on surface pollutants, a higher PBLH can either reduce

25   surface O_3 level by diluting O_3 and its precursors into a larger volume of air (Sanchez-Ccoyllo et al.,

26   2006; Wang et al., 2020) or increase in surface O_3 level by transporting more O_3 from upper

27   troposphere or lessening NO abundance for O_3 titration (Sun et al., 2010; He et al., 2017; Liu et al.,

28   2020a). Transport of O_3 from stratosphere to troposphere by synoptic scale and mesoscale process,

29   as indicated by an increase in potential vorticity (PV), typically leads to surface O_3 enhancement

30   (Wang et al., 2019; Wang et al., 2020). Precipitation has been verified to decrease surface O_3 level

31   through the wet removal of its precursors, and clouds reduce surface O_3 level by decreasing the

32   oxidative capacity of the atmosphere and enhancing scavenging of atmospheric oxidants (Lelieveld

33   and Crutzen, 1990; Liu et al., 2020b; Seinfeld and Pandis, 2016; Shan et al., 2008). A higher wind

34   speed can decrease surface O_3 level by a fast ventilation of O_3 and its precursors (Lu et al., 2019a;

35   Sanchez-Ccoyllo et al., 2006).

36   Emissions of air pollutants affect surface O_3 variability by perturbing the abundances of

37   hydroperoxyl (HO_2) and alkylperoxyl (RO_2) radicals which are the key atmospheric constituents in

38   formation of O_3 (Liu et al., 2020b). Many previous studies have verified a nonlinear relationship

39   between O_3 and its precursors (e.g., Atkinson, 2000; Wang et al., 2017; Liu et al., 2020b; Sun et al.,

40   2018; Lu et al., 2019). If surface O_3 formation regime lies within the VOCs limited region,

41   reductions in VOCs emissions will result in a reduction in surface O_3 level. Similarly, if surface O_3

42   formation regime lies within the NO_x limited region, reductions in NO_x emissions will result in a

1 reduction in surface O_3 level (Atkinson, 2000; Wang et al., 2017). If surface O_3 formation regime

2 lies within transitional region, reductions in either VOC or NO_x emissions will result in a reduction

3 in surface O_3 level. Atmospheric aerosols can affect surface O_3 level through either heterogeneous

4 reactions of reactive gases (Lu et al., 2012; Li et al., 2018; Stadtler et al., 2018; Lou et al., 2014) or

5 affecting the solar flux for gases photolysis and oxidation (Li et al., 2011; Xing et al., 2017).

6   Understanding the drivers of surface O_3 variability has a strong implication for O_3 mitigation

7 purpose (Sun et al., 2018; Lu et al., 2019a). China has experienced a continuous increase in surface

8 O_3 level despite the implementation of control measures on NO_x since 2013 (Liu et al., 2020a, 2020b;

9 Lu et al., 2018, 2020). Many studies have attempted to determine the drivers of high-O_3 events

10 occurred in specific region and time across China. Most of these studies focus on the most densely

11 populated and highly industrialized areas in eastern China, whereas the studies in the rest part of

12 China are still limited (Liu et al., 2020a; Liu et al., 2020b; Lu et al., 2018; Sun et al., 2018; Wang et

13 al., 2017). As China has a vast territory with a wide range of emission levels and meteorological

14 conditions, O_3 variability and its drivers may vary both temporally and geographically, so the results

15 from one region are not likely to be applicable nationally. In addition, previous studies typically use

16 state-of-the-art chemical transport models (CTMs) with sensitivity simulations to quantify the

17 drivers of O_3 variability, e.g., fixed meteorology but varied emission levels to quantify the influences

18 of emission changes or vice versa (Lu et al., 2019a; Liu et al., 2020a; Liu et al., 2020b). However,

19 uncertainties in local meteorological fields, emission estimates, and model mechanism can lead to

20 discrepancy in CTMs that may affect the accuracy of O_3 predictions and their sensitivities to changes

21 in emissions and meteorology (Lu et al., 2019a; Young et al., 2018). This is in particular for the

22 Sichuan basin (SCB), one of the most industrialized and populated cities cluster in western China,

23 where large discrepancies between measured and modelled surface O_3 are found due to the complex

24 terrain (Lu et al, 2019a; Wang et al, 2020).

25   After a continuous increase in surface O_3 level from 2013 to 2019, the summertime (May-

26 August) O_3 concentration across China showed a significant reduction in 2020 (Figure 1 (d)). In this

27 study, we use high resolution nested-grid GEOS-Chem simulation, the eXtreme Gradient Boosting

28 (XGBoost) machine learning method and the exposure−response relationship to determine the

29 drivers and evaluate the health risks of the unexpected surface O_3 enhancements. We first use the

30 XGBoost machine learning method to correct the GEOS-Chem model-to-measurement O_3

31 discrepancy over the SCB. The relative contributions of meteorology and anthropogenic emissions

32 changes to the unexpected surface O_3 enhancements are then quantified with the combination of

33 GEOS-Chem and XGBoost models. In order to assess the health risks caused by the unexpected O_3

34 enhancements over the SCB, total premature death mortalities are also estimated.

35 **2. Methods and data**

36 **2.1 Surface O_3 data and auxiliary data over the SCB**

37   China has identified nine cities clusters that lead the populations and developments of economy,

38 society, and culture across China. The SCB contains the fourth largest cities cluster in China after

39 the Yangtze River Delta (YRD), the Pearl River Delta (PRD), and the Beijing-Tianjin-Hebei (BTH)

40 cities clusters. The location of the SCB cities cluster is shown in Figure S1. With Chongqing and

41 Chengdu as the dual cores, more than a dozen cities including Mianyang, Deyang, Yibin, Nanchong,

42 Dazhou, and Luzhou over the SCB have achieved rapid economic development and industrial

1   upgrading. As the region with the strongest economic strength and best economic foundation in
2   western China, the SCB region has many industries such as energy and chemical industry, electronic
3   information, food processing, equipment manufacturing, eco-tourism, and modern finance. As one
4   of the most densely populated and highly industrialized region in China combined with the basin
5   terrain which is easy to trap air pollutants, the SCB is a newly emerging severe air pollution region
6   in China.

7   Surface O_3 measurements over the SCB are available from the China National Environmental
8   Monitoring Center (CNEMC) network (http://www.cnemc.cn/en/, last access: 2 July 2021). The
9   CNEMC network has routinely measured the concentrations of CO, O_3, NO_2, SO_2, PM_{10}, and $PM_{2.5}$
10 at 122 sites in 22 key cities over the SCB since 2015. The mean geolocation, population, the number
11 of measurement site, data range of each city are summarized in Table 1. The altitude of these cities
12 ranges from 0.3 to 4.3 km (above sea level, a.s.l.) and the population ranges from 822 to 32,054
13 thousand. The number of measurement site in each city ranges from 2 to 21. Surface O_3
14 measurements at all measurement sites are based on similar differential absorption ultraviolet (UV)
15 analyzers. The hourly mean time series of surface O_3 concentrations have covered the period from
16 January 2015 to present at all measurement sites in the 22 cities. After removing unreliable
17 measurements with the filter criteria used in Lu et al. (2020), we average the O_3 concentrations at
18 all measurement sites in each city to form a city representative O_3 dataset. O_3 metric used in this
19 study is on maximum 8-h average (MDA8) basis.

20   Since the vertical distributions of tropospheric HCHO and NO_2 are weighted heavily toward
21 the lower troposphere over polluted regions, many studies have used tropospheric column
22 measurements of these gases to represent near-ground conditions (Streets et al., 2013; Sun et al.,
23 2018; Sun et al., 2021). In this study, the tropospheric NO_2 and HCHO columns used for
24 investigating the changes in O_3 precursors in May-June 2020 vs. 2019 are prescribed from the
25 TROPOMI Level 3 products. TROPOMI overpasses China at approximately 13:30 local time (LT)
26 with a ground pixel size of 7 km × 7 km. Pixels with quality assurance values of less than 50% for
27 HCHO and 75% for NO_2 are not included in present work.

28 **2.2 GEOS-Chem nested-grid simulation**

29   We use the high resolution nested-grid GEOS-Chem model version 12.2.1 to simulate surface
30 O_3 over the SCB (Bey et al., 2001). Simulations are conducted at a horizontal resolution of 0.25 °×
31 0.3125 ° over the nested domain (70 °-140 °E, 15 °-55 °N) covering China and surrounding regions.
32 The boundary conditions for the nested-grid GEOS-Chem simulation are archived from the global
33 simulation at 2 °× 2.5 ° resolution (Yin et al., 2019; Yin et al., 2020; Sun et al., 2021). We spun up
34 the model for one year to remove the influence of the initial conditions. We first run global
35 simulation at 2 °× 2.5 ° resolution and then interpolate the restart file on 1 January 2018 into high
36 resolution (0.25 °× 0.3125 °) for the nested domain to initialize the nested model simulation from
37 January 2019 to June 2020.

38   The simulations were driven by GEOS-FP meteorological field at the native resolution of 0.25°
39 × 0.3125° and 47 layers from surface to 0.01 hPa at the top. The PBLH and surface meteorological
40 variables are implemented in 1-hour interval and other meteorological variables are in 3-hour
41 interval. The time step applied in the model for transport is 5 minutes and for chemistry and
42 emissions is 10 minute (Lu et al., 2019; Philip et al., 2016). The non-local scheme for the boundary

1   layer mixing process is from Lin et al. (2010), wet deposition is from Liu et al. (2001), and dry
2   deposition is generated with the resistance-in-series algorithm (Wesely, 1989; Zhang et al., 2001).
3   The photolysis rates are from the FAST-JX v7.0 photolysis scheme (Bian and Prather, 2002).
4   Chemical mechanism follows the universal tropospheric-stratospheric Chemistry (UCX)
5   mechanism (Eastham et al., 2014). The GEOS-Chem simulation outputs 47 layers of O_3 and other
6   atmospheric constituents over China with a temporal resolution of 1 hour.

7   We use the Community Emissions Data System (CEDS) inventory for global anthropogenic
8   emissions at the latest 2017 level, which is overwritten by the Chinese anthropogenic emissions
9   with the Multi-resolution Emission Inventory (MEIC) in 2019 (Li et al., 2017; Hoesly et al., 2018;
10 Zheng et al., 2018). Anthropogenic emissions are fixed for 2019 and 2020. Global BB and biogenic
11 emissions were from the Global Fire Emissions Database (GFED v4) inventory (Giglio et al. 2013)
12 and the Model of Emissions of Gases and Aerosols from Nature (MEGAN version 2.1) inventory
13 (Guenther et al. 2012), respectively. Natural emissions of BB, biogenic VOCs, lightning NO_x, and
14 soil NO_x are calculated online in the model.

15 **2.3 Correction of GEOS-Chem discrepancy with machine learning method**

16   We used the XGBoost machine learning method to correct the GEOS-Chem model-to-
17 measurement O_3 discrepancy over the SCB. It uses the Gradient Boosting Decision Tree (GBDT)
18 framework to iteratively train the GEOS-Chem model-to-measurement discrepancy to improve the
19 model predictions in a stagewise manner. XGBoost method minimizes the loss function by adding
20 a weak learner for the purpose of optimizing the objective function. The optimization objective
21 function used in XGBoost model is expressed as,

$$L^{(t)} \simeq \sum_{i=1}^{n} [l(y_i, \hat{y}^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t)$$

$$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$$ (1)

$$h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)})$$

22   where g_i and h_i are first and second order gradients of the loss function, respectively. $L^{(t)}$
23 represents the optimization objective function to be solved at the t-th iteration. $l(y_i, \hat{y}^{(t-1)})$ is the
24 loss function representing the difference between the prediction for the i-th sample at the (t-1)-th
25 iteration and the real values y_i. Function $f(t)$ is the change amount at the t-th iteration. Overall, the
26 objective function consists of a two order Taylor approximation expansion of the loss function and
27 the regularization term ($\Omega(f_t)$), which penalize the complexity of the model and prevent overfitting
28 of the model. Compared to traditional GBDT method, XGBoost method has the following
29 advantages: (1) effectively handle missing values; (2) prevent overfitting; (3) reduce computing
30 time by using parallel and distributed computing methods.

31   Since GEOS-Chem model-to-measurement discrepancy is usually site-specific, we train a
32 separate XGBoost model for each site. Similar to the method of Keller et al. (2021), we use a full
33 seasonal cycle of hourly measurements in 2019 at each site as the learning samples, and GEOS-
34 Chem input of emissions and meteorological parameters, output concentrations of atmospheric
35 constituents, and time information as training input data. In order to incorporate emissions and
36 meteorological factors that affect O_3 production properly, we have included the GEOS-Chem

Atmospheric
Chemistry
and Physics
Discussions

Open Access

EGU

1 simulated concentrations of 43 atmospheric chemical constituents, emissions of 21 atmospheric
2 chemical constituents, 10 meteorological parameters, and 4 time parameters (e.g., hour, day, month,
3 and year) into the data training. All these training input data are summarized in Table S1 and have
4 been standardized. We choose a learning rate of 0.35, maximum tree depth of 6, L1 and L2
5 regularization terms of 0 and 1, the loss function of mean square, and evaluation metric of root-
6 mean-square error (RMSE) in the data training.

7 We use k-fold cross-validation method to test the performance of the XGBoost model (k=1 –
8 n). First, all sample data are randomly and uniformly divided into k groups, where one group is
9 taken as the test dataset and the remaining k-1 groups are taken as the training dataset. We then start
10 to train the model and use the test dataset to evaluate the performance of the trained model. We
11 repeated this process for k times by using different groups of dataset as the test data. The training
12 model is finally determined if all the k groups of experiments show similar performances. This
13 method can obtain a stability and robustness of XGBoost model and avoid overfitting. In this study,
14 a10-fold cross-validation method is applied, i.e., we divide the O_3 measurements in 2019 into 10
15 groups of sub data: the training dataset accounts for 90% and the test dataset accounts for the
16 remaining 10% of the total sample data. We also attempted to use 60% and 80% of the sample data
17 as the training dataset and do not find significant influences on the results, indicating the robustness
18 of the XGBoost training model.

19 **2.4 Quantifying meteorological and emissions contributions**

20 We have used the GEOS-Chem only and the combination of GEOS-Chem and XGBoost model
21 (hereafter GEOS-Chem-XGBoost) to quantify the contributions of meteorology and anthropogenic
22 emissions to the unexpected surface O_3 enhancements over the SCB in 2020. For the GEOS-Chem
23 method, since the anthropogenic emissions are fixed, the simulated O_3 differences between 2020
24 and 2019 can be attributed to changes in meteorological conditions, which is calculated as,

25 $$G\_Met = G_{2020} - G_{2019} \qquad (2)$$

26 The contribution of anthropogenic emissions changes can then be quantified as,

27 $$G\_Emis = (Meas_{2020} - Meas_{2019}) - G\_Met \qquad (3)$$

28 where G_Met and G_Emis represent the meteorology and anthropogenic emissions contributions,
29 respectively. $Meas_{2019}$ and $Meas_{2020}$ represent O_3 measurements in 2019 and 2020, respectively.
30 G_{2019} and G_{2020} represent GEOS-Chem O_3 simulations in 2019 and 2020, respectively.

31 Since the GEOS-Chem-XGBoost model has corrected the GEOS-Chem model-to-
32 measurement discrepancy, we assume it can provide accurate predictions to the surface O_3
33 measurements. For predicting O_3 evolutions in 2020, all input parameters except anthropogenic
34 emissions fed into each GEOS-Chem-XGBoost model are updated to match the measurements in
35 2020, but anthropogenic emissions are fixed at the 2019 levels. As a result, the differences between
36 the GEOS-Chem-XGBoost predictions for 2020 and the 2020 measurements are attributed to the
37 changes in anthropogenic emissions (equation (4)). The meteorology induced contributions are then
38 obtained as equation (5) by subtracting the anthropogenic emissions induced contributions.

39 $$XG\_Emis = Meas_{2020} - XG_{2020} \qquad (4)$$

40 $$XG\_Met = (Meas_{2020} - Meas_{2019}) - XG\_Emis \qquad (5)$$

41 where the acronyms are similar to those in equations (1) and (2) but for GEOS-Chem-XGBoost
42 method. By correcting the model-to-measurement discrepancy, GEOS-Chem-XGBoost model is

1 expected to provide a more accurate O_3 sensitivity to changes in both meteorology and
2 anthropogenic emissions.

3 **2.5 Health risks evaluation**

4 We have assessed the total premature mortalities including all nonaccidental causes,
5 hypertension, CVD, RD, COPD, and stroke attributed to ambient O_3 exposure in all cities over the
6 SCB in 2019 and 2020. We first calculated the O_3 induced daily premature mortalities based on the
7 exposure−response relationship described in Cohen et al. (2004), which have been used in many
8 subsequent studies (Anenberg et al., 2010; Liu et al., 2018; Wang et al., 2021). We then added up
9 the daily premature mortalities within May-June or the whole year to get the total O_3 induced
10 premature mortalities in the respective periods. The population data used in this study include all
11 age groups, which may result in higher daily mortalities than expected (Liu et al., 2018; Wang et al.,
12 2021). According to Cohen et al. (2004), the daily premature mortalities attributable to ambient O_3
13 exposure can be estimated by the following equation (Cohen et al., 2004),

$$\Delta x = \begin{cases} 0, & (if \quad C_{meas} - C_{thres} \leq 0) \\ C_{meas} - C_{thres}, & (if \quad C_{meas} - C_{thres} \geq 0) \end{cases} \tag{6}$$

$$\Delta M = y_0[1 - \exp(-\beta \Delta x)] \times Pop \tag{7}$$

14 where ΔM represents the daily premature mortalities due to ambient O_3 exposure. The city
15 representative daily mean O_3 concentration C_{meas} is an average of all measurements in each city.
16 Variable y_0 is the daily baseline mortality rate for each disease averaged from all ages and genders.
17 We follow the method of Wang et al. (2021) and use the daily y_0 value for each disease from the
18 latest China Health Statistical Yearbook in 2018. β represents the increase in daily mortality as a result
19 of each 10 μg/cm^3 (~ 5.1 ppbv) increase in daily O_3 concentration, which is often referred to as the
20 concentration response function (CRF) in previous studies. We collected the CRF values straightly
21 from those used in Yin et al. (2017) and Wang et al. (2021). Δx represents the incremental O_3
22 concentration relative to the threshold concentration C_{thres} of 35.1 ppbv, which are used following
23 Lim et al. (2012) and Liu et al. (2018). Pop represents the population exposed in the ambient O_3
24 pollution, which are available from the seventh nationwide population census in 2020 provided by
25 National Bureau of Statistics of China. The daily y_0 and β values for all non-accidental causes,
26 hypertension, CVD, RD, COPD, and stroke are summarized in Table S2.

27 **3 Unexpected surface O_3 enhancements over the SCB in 2020**

28 Figures 1(a)-(b) show the May-June mean MDA8 O_3 concentrations at all measurement sites
29 over the SCB in 2019 and 2020. The May-June mean MDA8 O_3 concentrations averaged over all
30 cities in the SCB region in 2019 and 2020 are 48.1 ppbv and 58.3 ppbv, which are 11.0 ppbv lower
31 and 1.2 ppbv higher than those averaged over all Chinese cities in the same period, respectively. As
32 the most densely populated and highly industrialized region in western China, the land use,
33 industrialization, infrastructure construction, and urbanization over the SCB have expanded rapidly
34 in recent years, resulting in the highest anthropogenic emissions of O_3 precursors and highest surface
35 O_3 levels in the region (Figure S2). Although the O_3 levels in the SCB cities cluster are lower than
36 those in the three most developed city clusters in eastern China, i.e., the BTH, the Fenwei Plain
37 (FWP), and the YRD city clusters, the SCB region has been classified by the MEE as a newly

1   pollution region for O_3 mitigation (Sun et al., 2021). Situated in the basin with stationary
2   meteorological fields combined with high anthropogenic emissions, the SCB cities cluster is
3   potential to become a new region with frequent high-O_3 events after BTH, FWP, and YRD.
4   We find significant O_3 enhancements by 10.2 ± 0.8 ppbv (23.4%) (mean $\pm 1\sigma$ standard deviation)
5   averaged over all cities in the SCB in May-June 2020 vs. 2019 levels (Figure 1(c)). The largest
6   enhancements are observed in the most densely populated areas around the megacities Chongqing
7   and Chengdu (11.8 ± 0.6 ppbv (26.0%)). These year-to-year O_3 enhancements over the SCB are
8   record high in the 2015-2020 period, following an increasing change rate of 1.2% yr^{-1} from 2015 to
9   2017 and then a decreasing change rate of -0.7% yr^{-1} from 2017 to 2019. These surface O_3
10   enhancements mainly reflect regional emissions and meteorology changes in the SCB and
11   surrounding regions since the lifetimes of O_3 and most of its precursors are too short to undergo
12   long range transport.
13   The significant O_3 enhancements over the SCB in May-June 2020 vs. 2019 are opposite to the
14   overall decrease in surface O_3 levels across China in the same period (Figure 1 (d)). After a
15   continuous increase in surface O_3 levels from 2013 to 2019 by approximately 5% yr^{-1} (Figure 1(d)),
16   the MDA8 O_3 averaged over all cities outside the SCB across China in May-June 2020 vs. 2019
17   levels showed a significant reduction of 5.3 ± 0.5 ppbv (8.3%). Such O_3 reductions are widespread
18   in the eastern China, especially in the BTH, FWP, and YRD regions.

19   **4 Model performance assessment**

20   We use the metrics of normalized root-mean-square error (NRMSE), normalized mean bias
21   (NMB), and Pearson correlation coefficient (R) to assess the performance of the GEOS-Chem-
22   XGBoost model. For each measurement site, we analysed these metrics for both training (blue) and
23   test (red) datasets as shown in Figure S3. We define the NRMSE as the RMSE normalized by the
24   difference between the 95th and 5th percentile concentrations, and NMB as the mean bias
25   normalized by average concentration at the given measurement site. The formulas of above metrics
26   are summarized in Section S1.
27   The GEOS-Chem-XGBoost model predictions for surface O_3 over the SCB show no bias when
28   evaluated against the training data (NMB=0.01), NRMSEs of less than 0.1, and R between 0.93 –
29   1.0. Compared to the training data, the performances on the test data show a higher variability, with
30   an average NMB of –0.04, NRMSE of 0.22, and R of 0.83. We find no significant difference in
31   prediction performance between clean (less than the C_{thres} defined in section 2.5) and polluted
32   measurement sites. A number of factors likely contribute to relative poorer statistical results at some
33   sites such as Ganzizhou, Leshan, and Suining. On the one hand, the training data of these sites may
34   include certain temporal events that are not easily generalizable, such as unusual emissions activity
35   (e.g., BB, fireworks, closure of nearby point source) or weather patterns that are not properly
36   observed, which might be prone to overfitting. In addition, the differences in surface O_3 variabilities
37   between the training data and the test data at these sites are relative larger than other sites, which
38   can contribute to a relative poorer performance.
39   We use the SHapely Additive exPlanations (SHAP) approach to understand how the GEOS-
40   Chem-XGBoost model uses the input variables to make a prediction. The SHAP approach is based
41   on game-theoretic Shapely values and represents a measure of each predictor's responsibility for a
42   change in the model prediction (Lundberg et al., 2017). The SHAP values are computed separately

Atmospheric
Chemistry
and Physics
Discussions

1   for each individual model prediction, which offer detailed insight into the importance of each input
2   variable to this prediction while also consider the role of variables interactions (Lundberg et al.,
3   2020; Keller et al., 2021). Figure 2 shows the SHAP value distribution for all O_3 predictors averaged
4   over all cities in the SCB. The results show that any variables that are expected to be associated with
5   O_3 formation affect model O_3 prediction. Generally, the temperatures (at the surface, 2 m height,
6   and 10 m height) are the most important predictors for the GEOS-Chem model-to-measurement
7   discrepancy over the SCB, followed by the uncorrected GEOS-Chem simulated O_3, reactive
8   nitrogen (e.g., NO_2, Peroxyacetyl nitrate (PAN)), atmospheric oxidants (O_x, hydrogen peroxide
9   (H_2O_2)), fine aerosol, VOCs (Isoprene, C_3H_8), hour of the day, and meteorological variables
10   including horizontal and vertical wind speeds (u10m, v10m). All of these factors have tight
11   connections to surface O_3 formation over the SCB and it is thus not surprising that the GEOS-Chem
12   model-to-measurement discrepancies are most sensitive to them (Seinfeld and Pandis, 2016).
13   We have compared the performances of GEOS-Chem and GEOS-Chem-XGBoost in capturing
14   the measured surface O_3 levels. Figure 3 (a) shows the time series of measured and models predicted
15   O_3 concentrations averaged over all cities in the SCB region. Figure 3 (b) shows histogram of the
16   differences between the GEOS-Chem-XGBoost predictions and the measurements. The GEOS-
17   Chem simulations generally capture the daily variability of MDA8 O_3 over the SCB, but they show
18   high MB of 7.8 ppbv (17.5%) and RMSE of 15 ppbv across all measurement sites within the SCB
19   region. The discrepancy can be mainly attributed to uncertainties in the horizontal transport and
20   vertical mixing schemes simulated by the GEOS-Chem model at a relatively coarse spatial
21   resolution compared to the measurements at a single point, and also associated with the errors in
22   emission estimates, chemical mechanism, and sub-grid-scale local meteorological processes.
23   Especially errors in high SHAP values of O_3 predictors are more likely to result in large model-to-
24   measurement discrepancy. For example, GEOS-Chem model overestimates the correlations
25   between surface O_3 concentration and temperature (Figure S5 (a)), indicating that this
26   overestimation of O_3-to-temperature sensitivity is one of the major factors contributing to higher
27   GEOS-Chem model O_3 predictions.
28   By iteratively training and correcting the GEOS-Chem model-to-measurement discrepancy in
29   O_3-to-temperature sensitivity, the correlations between surface O_3 concentration and temperature
30   predicted by the GEOS-Chem-XGBoost model were in good agreement with the measurements
31   (Figure S5 (a)). With respect to the performance of reproducing the sensitivities of O_3 to other
32   meteorological parameters such as humidity, cloud fraction, and precipitation, the GEOS-Chem-
33   XGBoost model is also better than the GEOS-Chem (Figure S5 (b)-(d)). After correcting the errors
34   in all O_3 predictors, the GEOS-Chem-XGBoost model significantly improves the prediction of
35   surface O_3 concentrations in all cities over the SCB compared to the GEOS-Chem (Figure S6). It
36   shows a MB of 0.5 ppbv and RMSE of 0.3 ppbv for all O_3 measurements in 2019 over the SCB. As
37   a result, the overall GEOS-Chem-XGBoost model performance is acceptable and can support
38   further investigation of the drivers of the unexpected surface O_3 enhancements over the SCB in
39   May-June 2020.

40   **5 Attribution**
41   **5.1 Separation of meteorological and anthropogenic emissions contributions**

42   We attribute quantitatively the surface O_3 enhancements in May-June 2020 over the SCB to

1   changes in anthropogenic emissions and meteorological conditions according to equations (3) and
2   (4). Differences between the measured and GEOS-Chem-XGBoost predicted O_3 in May-June 2020
3   represent the anthropogenic emissions-induced O_3 changes in 2020 vs. 2019, as indicated by the
4   shadings in Figure 4(a). Figure 4(b) summarizes the mean contributions driven by changes in
5   anthropogenic emissions and meteorological conditions. Due to different change rates in
6   anthropogenic emissions in May and June in 2020 (see section 5.3), the changes in anthropogenic
7   emissions caused an overall increase in surface O_3 level in May but a reduction in surface O_3 level
8   in June (Figure 4 (a)). For the May-June mean contributions averaged over all cities in the SCB,
9   changes in anthropogenic emissions caused 0.9 ± 0.1 ppbv of O_3 reduction and changes in
10 meteorology caused 11.1 ± 0.7 ppbv of O_3 increase, which correspond to -8.0% and 108% of
11 relative contributions to the total O_3 enhancement (10.2 ± 0.8 ppbv) over the SCB in May-June 2020,
12 respectively. As a result, the unexpected O_3 enhancements over the SCB in 2020 were attributed to
13 that the anthropogenic emissions induced O_3 reductions are dominantly overwhelmed by the
14 meteorology induced O_3 increases.
15   We compare the meteorology and anthropogenic emissions induced contributions to the
16 unexpected surface O_3 enhancements estimated by the GEOS-Chem-XGBoost model with those by
17 the GEOS-Chem model only (Figure 4 (b)). Both methods agree that changes in meteorology play
18 a significant role in interpreting the O_3 enhancements, while the absolute magnitudes differ slightly
19 with each other. For example, the anthropogenic emissions induced O_3 reduction calculated with
20 the GEOS-Chem model only is 0.94 ppbv, while the value for GEOS-Chem-XGBoost model is 1.36
21 ppbv. By taking the subtraction in equation (1) and the average over all cities, the propagation of
22 systematic model discrepancies that are common to all measurements sites was effectively
23 minimized, which can mitigate the difference in attribution results between the GEOS-Chem and
24 GEOS-Chem-XGBoost methods. However, as demonstrated in Figure S6, model discrepancies may
25 differ from one region to the other and from time to time. Therefore, the GEOS-Chem-XGBoost
26 approach is expected to provide a more accurate and consistent estimate on O_3 change attribution.

27 **5.2 Meteorological contribution**

28   Figure 5 shows the terrain elevations and May-June mean wind fields and surface pressures
29 over the SCB and surrounding regions. The terrain altitudes are at a resolution of 3×3 arc-minute,
30 which indicates a rapid change in altitude from the Tibetan Plateau (4.0 –5.0 km) and Yunan-
31 Kweichou Plateau (2–3 km) to the SCB (0.5 km). The SCB is located in the saddle between the
32 Tibetan and Yunnan-Kweichou Plateau (Chen et al., 2009; Sun et al., 2021b). Figure 5 (b) are the
33 May-June mean wind fields at 500 m overlaid with surface pressure available from GEOS-FP fields
34 at $0.25° \times 0.3125°$ resolution. In May-June, the western Pacific Subtropical High originated from
35 the middle region of the Tibetan Plateau shifts westward to the west of the SCB (Chen et al., 2009).
36 The southwesterly East Asian summer monsoon generates a cyclonic pattern over the southeast part
37 of the SCB. Driven by large scale circulations, southwesterly flow enters the east part of the SCB
38 near the northwest edge of the Yunnan-Kweichou Plateau, while strong northwesterly flow enters
39 the SCB near the east edge of the Tibetan Plateau. The interaction of these two flows results in a
40 convergent zone of northward jet stream over the east part of the SCB due to the westward shift of
41 the Western Pacific Subtropical High and the blocking effect of topography. Furthermore, strong
42 instability of vertical convection could originate over the basin and move toward the east part of the

1 SCB as dry air continuously entered the upper layer over the SCB (Chen et al., 2009). This process
2 will continuously intensify the cyclonic vorticity over the SCB, and make it a favorable region for
3 stationary low-level vortices, which tend to trap air pollutants within the SCB region and is referred
4 to as the Southwest Vortex (Chen et al., 2009; Liu et al., 2003).

5   Figure 6 shows the May-June mean differences in potential vorticity (PV), precipitation,
6 temperature, specific humidity, cloud fraction, and PBLH between 2020 and 2019. In May-June
7 2020, the northwest, central western and southern China experienced anomaly strong or even
8 record-breaking droughts (https://quotsoft.net/air/), leading to significant increases in temperature
9 and decreases in precipitation, specific humidity and cloud fractions compared to the 2019 levels.
10 These meteorological conditions could enhance biogenic VOCs emissions, speed up O_3 chemical
11 production, and the aforementioned basin effect inhabit the ventilation of O_3 and its precursors,
12 which contributed to the O_3 enhancements over the SCB. Although higher PBLH over the SCB in
13 May-June 2020 vs. 2019 could reduce surface O_3 levels by diluting O_3 and its precursors into a larger
14 volume of air, this reduction effect was overwhelmed by its enhancement effect, i.e., higher PBLH
15 enhanced downward transport of O_3 from the upper troposphere. Indeed, we observed an increase
16 in downward potential vorticity (PV) over the SCB in May-June 2020 vs. 2019 (Figure 6 (a)). It is
17 worth noting that, with similar meteorological conditions in May-June 2020 vs. 2019, the O_3
18 enhancements were not observed in the northwest China such as Xinjiang and Inner Mongolia
19 provinces, and southern China such as the Pearl River Delta (PRD) region, which is also one of the
20 nine well-developed city clusters in China with severe air pollution. This can be partly attributed to
21 low anthropogenic emissions of O_3 precursors in northwest China (Lu et al. 2019; Zheng et al. 2018);
22 and that strong exchange between the land and sea in the coastal regions driven by the summer
23 monsoon facilitates the ventilation of O_3 and its precursors in the PRD region. Furthermore, the
24 meteorology induced O_3 enhancements are probably overwhelmed by the anthropogenic emissions
25 induced O_3 reductions in the aforementioned two regions.

26 **5.3 Emissions contribution**

27   To suppress the spread of coronavirus pandemic 2019 (COVID-19) across China and above,
28 the Chinese government sealed off several cities starting in January 2020; this included closing local
29 businesses and halting public transportation at an unprecedented scale (Steinbrecht et al., 2021; Liu
30 et al., 2020). These prevention measures quickly spread nationwide. Although the COVID-19
31 lockdowns in all cities have been removed before May, there are still restrictions on public
32 transportation, businesses, social activities and industrial manufactures, which could cause domestic
33 anthropogenic emissions reductions in both HCHO and NO_x. Furthermore, the MEE has
34 implemented The 2020 Action Plan on VOCs Mitigations in 2020. This Action Plan issues a number
35 of control measures including implementation of stringent VOCs emission standards, replacement
36 of raw and auxiliary materials with low VOCs content, and mitigation of unorganized emissions.
37 Driven by above two factors, the TROPOMI observed tropospheric HCHO and NO_2 over China in
38 May-June 2020 vs. 2019 reduced by 2.0 ± 0.3% (averaged for all Chinese cities) and 1.1 ± 0.2%,
39 respectively. Due to the relative short lifetime of both HCHO and NO_2 in troposphere, these
40 reductions mostly reflect local emissions changes. These reductions in domestic anthropogenic
41 emissions dominated the significant reduction of summertime MDA8 O_3 across China in 2020 vs
42 2019.

1   We have used the HCHO/NO_2 ratios following the method of Sun et al. (2018) to investigate
2 the O_3 production regime over the SCB region. The results show that the satellite observations of
3 HCHO/NO_2 ratios in May-June in most cities over the SCB have indicated a shift toward high values
4 from 2019 to 2020 but the O_3 chemical sensitivity in 2020 still lies within the transitional regime
5 (Jin et al., 2015; Jin et al., 2017; Figure S7). Meanwhile, the O_3 chemical sensitivity in May 2020
6 is similar to that in June, indicating that the O_3 variability in May-June 2020 is sensitive to both NO_x
7 and VOCs. The recently available Chinese anthropogenic emissions statistic data provided by the
8 MEE show that the anthropogenic VOCs over the SCB has decreased by 5.0% and 3.5% in May
9 and June in 2020 relative to the 2019 level, respectively. The anthropogenic NO_x in the same period
10 has increased by 1.5% and decreased by 1.7%, respectively (Zheng et al., 2021). The increase in
11 anthropogenic NO_x in May 2020 vs. 2019 is attributed to an increase in NO_x emission from power
12 plant sector, which was not affected by the post-lockdown restrictions for suppressing the spread of
13 COVID-19 (Table S3). For the May-June aggregation, the anthropogenic VOCs and NO_x over the
14 SCB have decreased by 4.3% and 0.3%, respectively (Zheng et al., 2021). These independent
15 analyses on anthropogenic emissions explain the different predicted O_3 changes due to
16 anthropogenic emissions alone in May (increase) versus June (decrease) in the SCB.

17   In contrast to the widespread reductions in both HCHO and NO_2 across the BTH, FWP, and
18 YRD regions, we find notable increases in both HCHO and NO_2 in the SCB in May-June 2020 vs.
19 2019 levels. The tropospheric HCHO and NO_2 columns averaged over all cities in the SCB region
20 have been increased by (2.8 ± 0.3%) and (5.1 ± 0.5%) in 2020 vs. 2019 levels, respectively. Since
21 both anthropogenic VOCs and NO_x emissions in the SCB showed decreasing change rates in May-
22 June 2020 vs. 2019, these regional increases in both HCHO and NO_2 could thus be attributed to
23 natural emissions enhancements in both VOCs and NO_2 in the SCB. Indeed, natural emissions of
24 biogenic VOCs and soil NO_x calculated online in the GEOS-Chem model show increasing change
25 rates in May-June 2020 vs. 2019 in the SCB and surrounding regions (Figure 7). These enhanced
26 biogenic VOCs and NO_x emissions are most likely driven by the hotter and dryer meteorological
27 conditions in the SCB and surrounding regions (Figure 7).

28   Finally, we concluded that natural emissions enhancements of both NO_x and VOCs induced by
29 the unexpected meteorology could be accounted for the O_3 enhancements in May-June 2020 over
30 the SCB. In present work, we were not able to determine which specific VOCs species are the most
31 effective for O_3 enhancements and cannot quantify the relative contributions of VOCs and NO_x
32 enhancements to the O_3 enhancements in the SCB. A series of sensitivity studies might be able to
33 address this important issue, but this is beyond the scope of present work.

34 **6 Health risks for the O_3 enhancements over the SCB**

35   Figure 8 presents the total premature mortalities from all non-accidental causes, hypertension,
36 CVD, RD, COPD, and stroke attributable to ambient O_3 exposure in all cities over the SCB during
37 May-June in 2019 and 2020. The statistical results for each city in 2019 and 2020 are summarized
38 in Table S4 and S5, respectively. The surface O_3 enhancements over the SCB in May-June 2020 vs.
39 2019 results in dramatically higher health risks. The estimated total premature mortalities from all
40 non-accidental causes due to the surface O_3 enhancements in May-June 2020 over the SCB is 5455,
41 which is 89.8% higher than that in the same period in 2019 (i.e., 2874). All above O_3 induced
42 diseases over the SCB have significant increases in total mortalities in May-June 2020 vs. 2019.

1 The highest health risk among these diseases is from CVD which is 741 in May-June 2019, followed
2 by RD (236), COPD (231), and hypertension (223). This O_3 induced health risk rank over the SCB
3 is consistent with those in the YRD, BTH, and PRD in previous studies (Liu et al., 2018; Lu et al.,
4 2020; Yin et al., 2017; Wang et al., 2021). In May-June 2020, total mortalities from CVD, RD,
5 COPD, hypertension, and stroke over the SCB reached up to 1405, 450, 439, 418, and 46,
6 respectively, due to significant O_3 enhancements. The change rates for these diseases are 89.6, 90.7,
7 90.1, 87.4, and 91.7%, respectively.

8   From a whole year view, the estimated total premature mortalities from all non-accidental
9 causes due to surface O_3 exposure over the SCB in 2019 and 2020 are 16,772 and 18,301,
10 respectively (Tables S4 and S5). All O_3 induced diseases within May-June 2019 account for about
11 ~ 17.0% of those in the whole year 2019, and this percentage reaches up to ~ 30.0% in 2020 (Figure
12 S8). The total premature mortalities from all non-accidental causes due to surface O_3 exposure over
13 the SCB has increased by 1528 in the whole year 2020 vs. 2019 (Figure S9), which is 40.8% lower
14 than that within May-June 2020 vs. 2019 (i.e., 2581). This indicates that the O_3 level over the SCB
15 showed an overall decreasing change rate in all months except May-June in 2020 vs. 2019, which
16 resulted in a decrease (by 1053) in O_3 induced diseases in the period.

17   We further investigated the O_3 induced diseases in the two most densely populated cities over
18 the SCB (i.e., Chengdu and Chongqing) during May-June in 2019 and 2020. The premature
19 mortalities from all O_3 induced diseases in 2020 vs. 2019 in each city are dependent on regional
20 population, surface O_3 level, and enhancement level (equation (6)). With largest populations and
21 highest O_3 enhancements, the estimated total premature mortalities in Chengdu and Chongqing
22 accounted for 46.9% of total O_3 induced mortalities over the SCB during May-June 2020 (Figure 8
23 (b)-(c)). Since the O_3 level and enhancement in Chengdu are larger than those in Chongqing, the
24 total O_3 induced mortalities in Chengdu are larger than those in Chongqing, though the population
25 in Chengdu is lower than that in Chongqing. The change rates for all O_3 induced diseases in Chengdu
26 are about 75% and in Chongqing are about 160% during May-June 2020 vs. 2019, which are much
27 higher than the enhancement percentages in the two cities (29.9 %). In order to reduce the O_3
28 induced health risk, strident O_3 control policies are necessary in densely populated cities.

29 **7 Conclusions**

30   Understanding the drivers and health risks of surface high O_3 events has a strong implication
31 for O_3 mitigation purpose. After a continuous increase in surface O_3 level from 2013 to 2019, the
32 overall summertime O_3 concentration across China showed a significant reduction in 2020. In
33 contrast to this overall reduction in surface O_3 level across China, unexpected surface O_3
34 enhancements of 10.2 ± 0.8 ppbv (23%) were observed in May-June 2020 vs. 2019 over the Sichuan
35 basin (SCB), China. In this study, we have used high resolution nested-grid GEOS-Chem simulation,
36 the eXtreme Gradient Boosting (XGBoost) machine learning method and the exposure−response
37 relationship to determine the drivers and evaluated the health risks of the unexpected surface O_3
38 enhancements.

39   By iteratively training and correcting the GEOS-Chem model-to-measurement discrepancies,
40 the GEOS-Chem-XGBoost model significantly improves the prediction of surface O_3
41 concentrations compared to the GEOS-Chem. It shows a MB of 0.5 ppbv and RMSE of 0.3 ppbv
42 against all O_3 measurements over the SCB. As a result, the overall GEOS-Chem-XGBoost model

Atmospheric
Chemistry
and Physics
Discussions
Open Access
EGU

performance is acceptable and can support further investigation of the drivers of the unexpected surface O_3 enhancements over the SCB in May-June 2020. The results show that changes in anthropogenic emissions caused 0.9 ± 0.1 ppbv of O_3 reduction and changes in meteorology caused 11.1 ± 0.7 ppbv of O_3 increase. The meteorology-induced surface O_3 increase is mainly attributed to significant increases in temperature and downward potential vorticity, and decreases in precipitation, specific humidity and cloud fractions over the SCB and surrounding regions in 2020 vs. 2019 levels. These changes in meteorology combined with the complex basin effect enhance downward transport of O_3 from the upper troposphere and biogenic emissions of VOCs and NO_x, speed up O_3 chemical production, and inhibit the ventilation of O_3 and its precursors, and therefore account for the surface O_3 enhancements over the SCB.

The unexpected surface O_3 enhancements over the SCB in May-June 2020 vs. 2019 result in dramatically higher health risks. The estimated total premature mortalities due to the unexpected surface O_3 enhancements over the SCB in May-June 2020 is 5455, which is 89.8% higher than that in the same period in 2019 (i.e., 2874). We further investigated the O_3 induced diseases in the two most densely populated cities over the SCB (i.e., Chengdu and Chongqing) during May-June in 2019 and 2020. With largest populations and highest O_3 enhancements, the estimated total premature mortalities in Chengdu and Chongqing accounted for 46.9% of total O_3 induced mortalities over the SCB. The change rates for all O_3 induced diseases in Chengdu are about 75% and in Chongqing are about 160% during May-June 2020 vs. 2019, which are much higher than the enhancement percentages in the two cities (29.9 %). In order to reduce the O_3 induced health risks, strident O_3 control policies are necessary in densely populated cities.

Code and data availability. Surface O_3 measurements over the SCB are from http://www.cnemc.cn/en/. All other data are available on request of YS (ywsun@aiofm.ac.cn)

Author contributions. YS designed and wrote the paper. HY carried out the GEOS-Chem simulations and GEOS-Chem-XGBoost training and evaluation. XL designed the concept of health risk evaluation and revised the manuscript. BZ constructed the latest MEIC emission inventory. JN, MP, CL, and YT provided constructive comments.

Competing interests. None.

References

Atkinson, R.: Atmospheric chemistry of VOCs and NO_x, Atmos. Environ., 34, 2063–2101, https://doi.org/10.1016/S1352-2310(99)00460-4, 2000.

Anenberg, S. C., Horowitz, L. W., Tong, D. Q., and West, J. J.: An Estimate of the Global Burden of Anthropogenic Ozone and Fine Particulate Matter on Premature Human Mortality Using Atmospheric Modeling, Environ. Health Persp., 118, 1189–1195, 2010.

Akimoto, H., Mori, Y., Sasaki, K., Nakanishi, H., Ohizumi, T., and Itano, Y.: Analysis of monitoring data of ground-level ozone in Japan for long-term trend during 1990–2010: Causes of temporal and spatial variation, Atmos. Environ., 102, 302–310, https://doi.org/10.1016/j.atmosenv.2014.12.001,

1   2015.

2   Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q. B., Liu, H. G. Y.,
3   Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated
4   meteorology: Model description and evaluation, J. Geophys. Res.-Atmos., 106, 23073-23095, 2001.

5   Bian, H. S., and Prather, M. J.: Fast-J2: Accurate simulation of stratospheric photolysis in global chemical
6   models, J. Atmos. Chem., 41, 281-296, 2002.

7   Cooper O. R. Detecting the fingerprints of observed climate change on surface ozone variability. Sci Bull
8   2019;64:359–60.

9   Cohen, A. J., Anderson, H. R., Ostro, B., Pandey, K. D., Krzyzanowski, M., Künzli, N., Gutschmidt, K.,
10   Pope, A., Romieu, I., Samet, J. M., and Smith, K.: Urban air pollution, in: Comparative
11   quantification of health risks, Global and regional burden of disease attributable to selected major
12   risk factors, Volume1, World Health Organization, Geneva, 2004.

13   Chen, D., Wang, Y., McElroy, M. B., He, K., Yantosca, R. M., and Le Sager, P.: Regional CO pollution
14   and export in China simulated by the high-resolution nested-grid GEOS-Chem model, Atmos. Chem.
15   Phys., 9, 3825-3839, 2009.

16   Eastham, S. D., Weisenstein, D. K., and Barrett, S. R. H.: Development and evaluation of the unified
17   tropospheric-stratospheric chemistry extension (UCX) for the global chemistry-transport model
18   GEOS-Chem, Atmos. Environ., 89, 52-63, 2014.

19   Fu, T. M., Zheng, Y. Q., Paulot, F., Mao, J. Q., and Yantosca, R. M.: Positive but variable sensitivity of
20   August surface ozone to large scale warming in the southeast United States, Nat. Clim. Change, 5,
21   454–458, https://doi.org/10.1038/Nclimate2567, 2015.

22   Fleming, Z. L.; Doherty, R. M.; Von Schneidemesser, E.; Malley, C. S.; Cooper, O. R.; Pinto, J. P.; Colette,
23   A.; Xu, X.; Simpson, D.; Schultz, M. G.; Lefohn, A. S.; Hamad, S.; Moolla, R.; Solberg, S.; Feng,
24   Z. Tropospheric ozone assessment report: Present-day ozone distribution and trends relevant to
25   human health. Elementa: Sci. Anthrop. 2018, 6, 12.

26   Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area
27   using the fourth-generation global fire emissions database (GFED4), J. Geophys. Res.-Biogeo., 118,
28   317-328, 10.1002/jgrg.20042, 2013.

29   Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.:
30   The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended
31   and updated framework for modeling biogenic emissions, Geosci. Model. Dev., 5, 1471-1492, 2012.

32   Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global
33   terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from
34   Nature), Atmos. Chem. Phys., 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

35   Hoesly, R. M., Smith, S. J., Feng, L. Y., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J.,
36   Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J., Li, M.,
37   Liu, L., Lu, Z. F., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750-2014)
38   anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data
39   System (CEDS), Geosci. Model. Dev., 11, 369-408, 2018.

40   He, J., Gong, S., Yu, Y., Yu, L., Wu, L., Mao, H., Song, C., Zhao, S., Liu, H., Li, X., and Li, R.: Air
41   pollution characteristics and their relation to meteorological conditions during 2014– 2015 in major
42   Chinese cities, Environ. Pollut., 223, 484–496, https://doi.org/10.1016/j.envpol.2017.01.050, 2017.

43   Im, U., Markakis, K., Poupkou, A., Melas, D., Unal, A., Gerasopoulos, E., Daskalakis, N., Kindap, T.,
44   and Kanakidou, M.: The impact of temperature changes on summer time ozone and its precursors

1   in the Eastern Mediterranean, Atmos. Chem. Phys., 11, 3847–3864, https://doi.org/10.5194/acp-11-
2   3847-2011, 2011.

3   Jin, X., and T. Holloway (2015), Spatial and temporal variability of ozone sensitivity over China observed
4   from the Ozone Monitoring Instrument, J. Geophys. Res. Atmos., 120, 7229–7246,
5   doi:10.1002/2015JD023250.

6   Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., Tonnesen, G. S.: Evaluating
7   a space-based indicator of surface ozone-NOx-VOC sensitivity over midlatitude source regions and
8   application   to   decadal   trends.   J   Geophys.Res.-Atmos.,   122,   10,   439– 10,
9   461.   https://doi.org/10.1002/2017JD026720, 2017.

10 Kalabokas, P. D., Thouret, V., Cammas, J. P., Volz-Thomas, A., Boulanger, D., and Repapis, C. C.: The
11   geographical distribution of meteorological parameters associated with high and low summer ozone
12   levels in the lower troposphere and the boundary layer over the eastern Mediterranean (Cairo case),
13   Tellus B, 67, 27853, https://doi.org/10.3402/tellusb.v67.27853, 2015.

14 Keller, C. A.; Evans, M. J.; Knowland, K. E.; Hasenkopf, C. A.; Modekurty, S.; Lucchesi, R. A.; Oda, T.;
15   Franca, B. B.; Mandarino, F. C.; Díaz Suárez, M. V.; Ryan, R. G.; Fakes, L. H.; Pawson, S., Global
16   impact of COVID-19 restrictions on the surface concentrations of nitrogen dioxide and ozone.
17   Atmos. Chem. Phys. 2021, 21 (5), 3555-3592.

18 Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, Adv. Neural Inf.
19   Process. Syst., 30, 4768– 4777, 2017.

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

23 Li, M., Zhang, Q., Kurokawa, J., Woo, J. H., He, K. B., Lu, Z. F., Ohara, T., Song, Y., Streets, D. G.,
24   Carmichael, G. R., Cheng, Y. F., Hong, C. P., Huo, H., Jiang, X. J., Kang, S. C., Liu, F., Su, H., and
25   Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international
26   collaboration framework of the MICS-Asia and HTAP, Atmos. Chem. Phys., 17, 935-963, 2017.

27 Liu, H. Y., Jacob, D. J., Bey, I., and Yantosca, R. M.: Constraints from Pb-210 and Be-7 on wet deposition
28   and transport in a global three-dimensional chemical tracer model driven by assimilated
29   meteorological fields, J. Geophys. Res.-Atmos., 106, 12109-12128, 2001.

30 Lu, X., Hong, J. Y., Zhang, L., Cooper, O. R., Schultz, M. G., Xu, X. B., Wang, T., Gao, M., Zhao, Y. H.,
31   and Zhang, Y. H.: Severe Surface Ozone Pollution in China: A Global Perspective, Environ. Sci.
32   Tech. Let., 5, 487-494, 2018.

33 Li, K.; Jacob, D. J.; Liao, H.; Qiu, Y.; Shen, L.; Zhai, S.; Bates, K. H.; Sulprizio, M. P.; Song, S.; Lu, X.;
34   Zhang, Q.; Zheng, B.; Zhang, Y.; Zhang, J.; Lee, H. C.; Kuk, S. K., Ozone pollution in the North
35   China Plain spreading into the late-winter haze season. Proceedings of the National Academy of
36   Sciences 2021, 118 (10), e2015797118.

37 Li, J., Wang, Z., Wang, X., Yamaji, K., Takigawa, M., Kanaya, Y., Pochanart, P., Liu, Y., Irie, H., Hu, B.,
38   Tanimoto, H., and Akimoto, H.: Impacts of aerosols on summertime tropospheric photolysis
39   frequencies and photochemistry over Central Eastern China, Atmos. Environ., 45, 1817–1829,
40   https://doi.org/10.1016/j.atmosenv.2011.01.016, 2011.

41 Liu, H. Y., Jacob, D. J., Bey, I., Yantosca, R. M., Duncan, B. N., and Sachse, G. W.: Transport pathways
42   for Asian pollution outflow over the Pacific: Interannual and seasonal variations, J. Geophys. Res.-
43   Atmos., 108, 2003.

44 Lou, S. J., Liao, H., and Zhu, B.: Impacts of aerosols on surfacelayer ozone concentrations in China

through heterogeneous reactions and changes in photolysis rates, Atmos. Environ., 85, 123– 138, https://doi.org/10.1016/j.atmosenv.2013.12.004, 2014.

Lu, K. D., Rohrer, F., Holland, F., Fuchs, H., Bohn, B., Brauers, T., Chang, C. C., Häseler, R., Hu, M., Kita, K., Kondo, Y., Li, X., Lou, S. R., Nehr, S., Shao, M., Zeng, L. M., Wahner, A., Zhang, Y. H., and Hofzumahaus, A.: Observation and modelling of OH and HO_2 concentrations in the Pearl River Delta 2006: a missing OH source in a VOC rich atmosphere, Atmos. Chem. Phys., 12, 1541–1569, https://doi.org/10.5194/acp-12-1541-2012, 2012.

Li, J., Chen, X. S., Wang, Z. F., Du, H. Y., Yang, W. Y., Sun, Y. L., Hu, B., Li, J. J., Wang, W., Wang, T.,Fu, P. Q., and Huang, H. L.: Radiative and heterogeneous chemical effects of aerosols on ozone and inorganic aerosols over East Asia, Sci. Total Environ., 622, 1327–1342, https://doi.org/10.1016/j.scitotenv.2017.12.041, 2018.

Lelieveld, J. and Crutzen, P. J.: Influences of cloud photochemical processes on tropospheric ozone, Nature, 343, 227–233, https://doi.org/10.1038/343227a0, 1990.

Lee, Y. C., Shindell, D. T., Faluvegi, G., Wenig, M., Lam, Y. F., Ning, Z., Hao, S., and Lai, C. S.: Increase of ozone concentrations, its temperature sensitivity and the precursor factor in South China, Tellus B, 66, 23455, https://doi.org/10.3402/tellusb.v66.23455, 2014.

Lin, J. T., Patten, K. O., Hayhoe, K., Liang, X. Z., and Wuebbles, D. J.: Effects of future climate and biogenic emissions changes on surface ozone over the United States and China, J. Appl. Meteorol. Climatol., 47, 1888–1909, https://doi.org/10.1175/2007jamc1681.1, 2008.

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, P. Natl. Acad. Sci. USA, 116, 422–427, 2019a.

Li, K., Jacob, D. J., Liao, H., Zhu, J., Shah, V., Shen, L., Bates, K. H., Zhang, Q., and Zhai, S.: A two-pollutant strategy for improving ozone and particulate air quality in China, Nat. Geosci., 12, 906– 910, https://doi.org/10.1038/s41561-019-0464-x, 2019b.

Liu, Y. and Wang, T.: Worsening urban ozone pollution in China from 2013 to 2017 – Part 1: The complex and varying roles of meteorology, Atmos. Chem. Phys., 20, 6305–6321, https://doi.org/10.5194/acp-20-6305-2020, 2020a.

Liu, Y. and Wang, T.: Worsening urban ozone pollution in China from 2013 to 2017 – Part 2: The effects of emission changes and implications for multi-pollutant control, Atmos. Chem. Phys., 20, 6323– 6337, https://doi.org/10.5194/acp-20-6323-2020, 2020b.

Lin, J. T., Mcelroy, M. B.:Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing, Atmospheric Environment, 2010, 44(14): 1726-1739.

Lu, X., Zhang, L., Chen, Y. F., Zhou, M., Zheng, B., Li, K., Liu, Y. M., Lin, J. T., Fu, T. M., and Zhang, Q.: Exploring 2016-2017 surface ozone pollution over China: source contributions and meteorological influences, Atmos. Chem. Phys., 19, 8339-8361, 2019a.

Lu, X., Zhang, L., Yue, X., Zhang, J., Jaffe, D. A., Stohl, A., Zhao, Y., and Shao, J.: Wildfire influences on the variability and trend of summer surface ozone in the mountainous western United States, Atmos. Chem. Phys., 16, 14687–14702, https://doi.org/10.5194/acp-16-14687-2016, 2016.

Lu, X.; Zhang, L.; Wang, X.; Gao, M.; Li, K.; Zhang, Y.; Yue, X.; Zhang, Y., Rapid Increases in Warm-Season Surface Ozone and Resulting Health Impact in China Since 2013. Environ. Sci. Tech. Let. 2020, 7 (4), 240-247.

Lu, X., Zhang, L., and Shen, L.: Meteorology and climate influences on tropospheric ozone: A review of natural sources, chemistry, and transport patterns, Curr. Pollution. Rep., 5, 238–260,

1   https://doi.org/10.1007/s40726-019-00118-3, 2019b.

2   Liu, F. , Page, A. , Strode, S. A. , Y Yoshida, Choi, S. , & Zheng, B. , et al. (2020). Abrupt declines in
3   tropospheric nitrogen dioxide over china after the outbreak of covid-19. Science Advances,
4   eabc2992.

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

8   Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adairrohani, H., Almazroa, M. A., Amann,
9   M., Anderson, H. R., and Andrews, K. G.: A comparative risk assessment of burden of disease and
10   injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic
11   analysis for the Global Burden of Disease Study 2010, Lancet, 380, 2224–2260, 2012.

12   Philip, S., Martin, R. V., and Keller, C. A.: Sensitivity of chemistry-transport model simulations to the
13   duration of chemical and transport operators: a case study with GEOS-Chem v10-01, Geosci. Model
14   Dev., 9(5), 1683-1695. https://doi.org/10.5194/gmd-9-1683-2016, 2016.

15   Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P.,
16   Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y.,
17   Lu, Z., Martin, R. V., Pfister, G. G., Pinder, R.W., Salawitch, R. J., and Wecht, K. J.: Emissions
18   estimation from satellite retrievals: A review of current capability, Atmos. Environ., 77, 1011–1042,
19   https://doi.org/10.1016/j.atmosenv.2013.05.051, 2013.

20   Stadtler, S., Simpson, D., Schröder, S., Taraborrelli, D., Bott, A., and Schultz, M.: Ozone impacts of gas–
21   aerosol uptake in global chemistry transport models, Atmos. Chem. Phys., 18, 3147– 3171,
22   https://doi.org/10.5194/acp-18-3147-2018, 2018.

23   Shan, W., Yin, Y., Zhang, J., and Ding, Y.: Observational study of surface ozone at an urban site in East
24   China, Atmos. Res., 89, 252–261, https://doi.org/10.1016/j.atmosres.2008.02.014, 2008.

25   Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics: From Air Pollution to Climate
26   Change, John Wiley & Sons, Hoboken, 2016.

27   Sanchez-Ccoyllo, O. R., Ynoue, R. Y., Martins, L. D., and Andrade, M. D.: Impacts of ozone precursor
28   limitation and meteorological variables on ozone concentration in Sao Paulo, Brazil, Atmos.
29   Environ., 40, 552–562, https://doi.org/10.1016/j.atmosenv.2006.04.069, 2006.

30   Sun, Y., Wang, Y., and Zhang, C.: Vertical observations and analysis of $PM_{2.5}$, O_3, and NO_x at Beijing
31   and Tianjin from towers during summer and Autumn 2006, Adv. Atmos. Sci., 27, 123,
32   https://doi.org/10.1007/s00376-009-8154-z, 2010.

33   Sun, Y., Yin, H., Liu, C., Zhang, L., Cheng, Y., Palm, M., Notholt, J., Lu, X., Vigouroux, C., Zheng, B.,
34   Wang, W., Jones, N., Shan, C., Qin, M., Tian, Y., Hu, Q., Meng, F., and Liu, J.: Mapping the drivers
35   of formaldehyde (HCHO) variability from 2015 to 2019 over eastern China: insights from Fourier
36   transform infrared observation and GEOS-Chem model simulation, Atmos. Chem. Phys., 21, 6365–
37   6387, https://doi.org/10.5194/acp-21-6365-2021, 2021a.

38   Sun, Y. , Liu, C., Palm, M., Vigouroux, C., Notholt, J., Hui, Q. H., Jones, N., Wang, W., Su, W. J., Zhang,
39   W. Q., Shan, C. G., Tian, Y., Xu, X. W., De Maziere, M., Zhou, M. Q., and Liu, J. G.: Ozone seasonal
40   evolution and photochemical production regime in the polluted troposphere in eastern China derived
41   from high-resolution Fourier transform spectrometry (FTS) observations, Atmos. Chem. Phys., 18,
42   14569-14583, 2018.

43   Sun, Y., Yin, H., Cheng, Y., Zhang, Q., Zheng, B., Notholt, J., Lu, X., Liu, C., Tian, Y., and Liu, J.:
44   Quantifying variability, source, and transport of CO in the urban areas over the Himalayas and

1   Tibetan Plateau, Atmos. Chem. Phys., 21, 9201–9222, https://doi.org/10.5194/acp-21-9201-2021,
2   2021b.

3   Sun, Y., Liu, C., Zhang, L., Palm, M., Notholt, J., Hao, Y., Vigouroux, C., Lutsch, E., Wang, W., Shan,
4   C. G., Blumenstock, T., Nagahama, T., Morino, I., Mahieu, E., Strong, K., Langerock, B., De
5   Maziere, M., Hu, Q. H., Zhang, H. F., Petri, C., and Liu, J. G.: Fourier transform infrared time series
6   of tropospheric HCN in eastern China: seasonality, interannual variability, and source attribution,
7   Atmos. Chem. Phys., 20, 5437-5456, 2020.

8   Steinbrecht, W., Kubistin, D., Plass Dülmer, C., Davies, J., Tarasick, D. W., Gathen, P., et al. (2021).
9   COVID-19 crisis reduces free tropospheric ozone across the Northern Hemisphere. Geophysical
10   Research Letters, 48, e2020GL091987. https://doi.org/10.1029/2020GL091987

11 Tarvainen, V., Hakola, H., Hellén, H., Bäck, J., Hari, P., and Kulmala, M.: Temperature and light
12   dependence of the VOC emissions of Scots pine, Atmos. Chem. Phys., 5, 989–998,
13   https://doi.org/10.5194/acp-5-989-2005, 2005.

14 Van Dingenen, R., Dentener, F. J., Raes, F., Krol, M. C., Emberson, L., and Cofala, J.: The global impact
15   of ozone on agricultural crop yields under current and future air quality legislation, Atmos. Environ.,
16   43, 604-618, 2009.

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

20 Wesely, M. L.: Parameterization of Surface Resistances to Gaseous Dry Deposition in Regional-Scale
21   Numerical-Models, Atmos. Environ., 23, 1293-1304, Doi 10.1016/0004-6981(89)90153-4, 1989.

22 Wang, H., Wang W., Huang X., and Ding A.: Impacts of stratosphere-to-troposphere-transport on
23   summertime surface ozone over eastern China. Science Bulletin 65. 4:276-279, 2020.

24 Wang, X. , Wu, Y. , Randel, W. , and Tilmes, S.: Stratospheric contribution to the summertime high
25   surface ozone events over the western united states. Environ. Res. Lett., 15(10), 1040a6, 2020.

26 Wang, P., Qiao X., Zhang, H.: Modeling $PM_{2.5}$ and O_3 with aerosol feedbacks using WRF/Chem over the
27   Sichuan Basin, southwestern China, Chemosphere, 254, 126735, ISSN 0045-6535,
28   https://doi.org/10.1016/j.chemosphere.2020.126735, 2020.

29 Wang, P., Shen, J., Xia, M., Sun, S., Zhang, Y., Zhang, H., and Wang, X.: Unexpected enhancement of
30   ozone exposure and health risks during National Day in China, Atmos. Chem. Phys., 21, 10347–
31   10356, https://doi.org/10.5194/acp-21-10347-2021, 2021.

32 Xing, J., Wang, J., Mathur, R., Wang, S., Sarwar, G., Pleim, J., Hogrefe, C., Zhang, Y., Jiang, J., Wong,
33   D. C., and Hao, J.: Impacts of aerosol direct effects on tropospheric ozone through changes in
34   atmospheric dynamics and photolysis rates, Atmos. Chem. Phys., 17, 9869–9883,
35   https://doi.org/10.5194/acp-17- 9869-2017, 2017.

36 Yin, P., Chen, R., Wang, L., Meng, X., Liu, C., Niu, Y., Lin, Z., Liu, Y., Liu, J., and Qi, J.: Ambient Ozone
37   Pollution and Daily Mortality: A Nationwide Study in 272 Chinese Cities, Environ. Health Persp.,
38   125, 117006, https://doi.org/10.1289/EHP1849, 2017.

39 Yin, H., Sun, Y. W., Liu, C., Zhang, L., Lu, X., Wang, W., Shan, C. G., Hu, Q. H., Tian, Y., Zhang, C. X.,
40   Su, W. J., Zhang, H. F., Palm, M. A., Notholt, J., and Liu, J. G.: FTIR time series of stratospheric
41   NO_2 over Hefei, China, and comparisons with OMI and GEOS-Chem model data, Opt. Express, 27,
42   A1225-A1240, 2019.

43 Yin, H., Sun, Y. W., Liu, C., Lu, X., Smale, D., Blumenstock, T., Nagahama, T., Wang, W., Tian, Y., Hu,
44   Q. H., Shan, C. G., Zhang, H. F., and Liu, J. G.: Ground-based FTIR observation of hydrogen

1 chloride (HCl) over Hefei, China, and comparisons with GEOS-Chem model data and other ground-
2   based FTIR stations data, Opt. Express, 28, 8041-8055, 2020.
3 Zhang, L. M., Gong, S. L., Padro, J., and Barrie, L.: A size-segregated particle dry deposition scheme for
4   an atmospheric aerosol module, Atmos. Environ., 35, 549-560, 2001.
5 Zheng, B., Tong, D., Li, M., Liu, F., Hong, C. P., Geng, G. N., Li, H. Y., Li, X., Peng, L. Q., Qi, J., Yan,
6   L., Zhang, Y. X., Zhao, H. Y., Zheng, Y. X., He, K. B., and Zhang, Q.: Trends in China's
7   anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys.,
8   18, 14095-14111, 10.5194/acp-18-14095-2018, 2018.
9 Zheng, B., Zhang, Q., Geng, G., Chen, C., Shi, Q., Cui, M., Lei, Y., and He, K.: Changes in China's
10   anthropogenic emissions and air quality during the COVID-19 pandemic in 2020, Earth Syst. Sci.
11   Data, 13, 2895–2907, https://doi.org/10.5194/essd-13-2895-2021, 2021.
12

1 **Tables**

2 **Table 1**. Measurement sites in the SCB city clusters. All sites are organised alphabetically. Population statistics are
3 based on the seventh nationwide population census in 2020 provided by National Bureau of Statistics of China.

| Name	Longitude mean	Latitude mean	Altitude mean (km)	Population	Number of sites	Time period
Abazhou	102.21°E	31.91°N	3.5	822,587	3	2015 - present
Bazhong	106.75°E	31.85°N	0.8	2,712,894	4	2015 - present
Chengdu	104.04°E	30.69°N	0.5	20,938,000	10	2015 - present
Chongqing	106.51°E	29.58°N	0.4	32,054,200	21	2015 - present
Dazhou	107.5°E	31.22°N	1.0	5,385,422	5	2015 - present
Deyang	104.39°E	31.12°N	0.5	3,456,161	4	2015 - present
Ganzizhou	101.96°E	30.05°N	3.5	1,107,431	2	2015 - present
Guangan	106.63°E	30.48°N	1.7	3,254,883	6	2015 - present
Guangyuan	105.85°E	32.44°N	2.1	2,305,657	4	2015 - present
Leshan	103.76°E	29.57°N	0.5	3,160,168	4	2015 - present
Liangshanzhou	102.28°E	27.87°N	2.3	4,858,359	5	2015 - present
Luzhou	105.43°E	28.9°N	0.3	4,254,149	4	2015 - present
Meishan	103.85°E	30.07°N	0.8	2,955,219	6	2015 - present
Mianyang	104.73°E	31.48°N	0.7	4,868,243	4	2015 - present
Nanchong	106.09°E	30.8°N	0.3	5,607,565	6	2015 - present
Neijiang	105.05°E	29.59°N	0.5	3,140,678	4	2015 - present
Panzhihua	101.69°E	26.56°N	2.6	1,212,203	5	2015 - present
Suining	105.71°E	30.58°N	0.5	2,814,196	4	2015 - present
Yaan	103.01°E	29.99°N	3.1	1,434,603	4	2015 - present
Yibin	104.62°E	28.78°N	2.0	4,588,804	6	2015 - present
Zigong	104.75°E	29.35°N	0.3	2,489,256	6	2015 - present
Ziyang	104.64°E	30.13°N	0.5	2,308,631	5	2015 - present

4
5

1 **Figures**
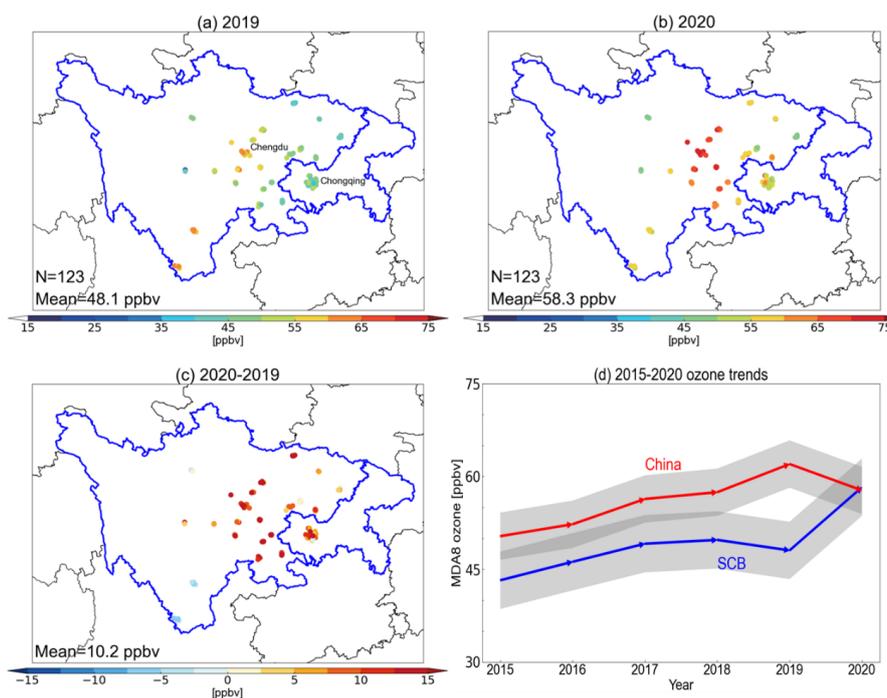
2

3 **Figure 1** Surface O_3 enhancements over the SCB region in May-June 2020 vs. 2019. (a) Spatial distributions of

4 May-June mean O_3 concentrations over the SCB region in 2019. Number (N) denotes available measurement sites

5 for this year. We average the O_3 concentrations at all measurement sites in each city to form a city representative O_3

6 dataset. (b) Same as (a) but for 2020. (c) Differences between 2020 and 2019. (d) Trends in May-June mean ozone

7 concentrations from 2015-2020 averaged for all Chinese cities (red) and for the SCB cities cluster (blue). Grey

8 shadings represent the range of mean value $\pm$ 1σ STD across all cities.
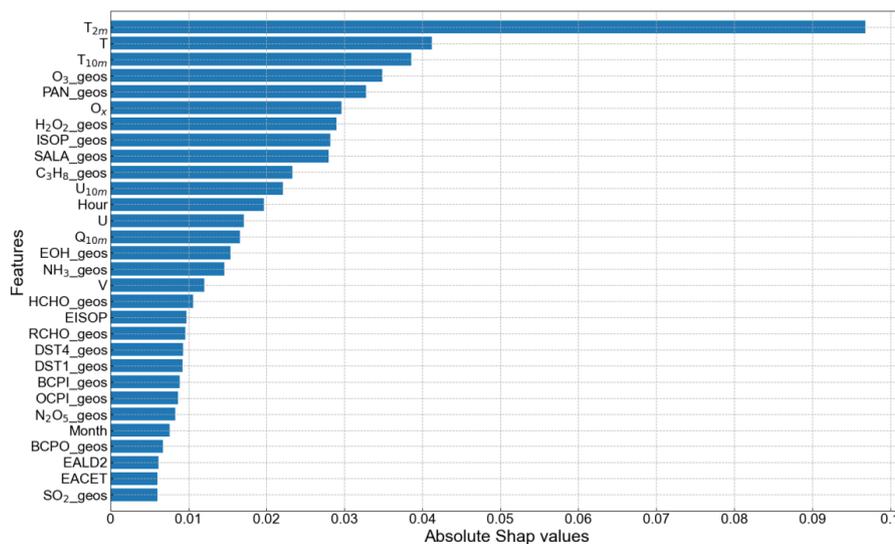
Figure 2 Importance of input variables for the XGBoost model trained to correct the GEOS-Chem model-to-measurement O_3 discrepancy over the SCB. Shown are the distribution of the SHAP values for each variable averaged over all cities in the SCB, ranked by the average importance of each feature. Higher SHAP value indicates higher feature importance. Descriptions for all acronyms are listed in Table A1. For clarity, only the top 30 variables are shown. See Figure S4 for importance of all variables.
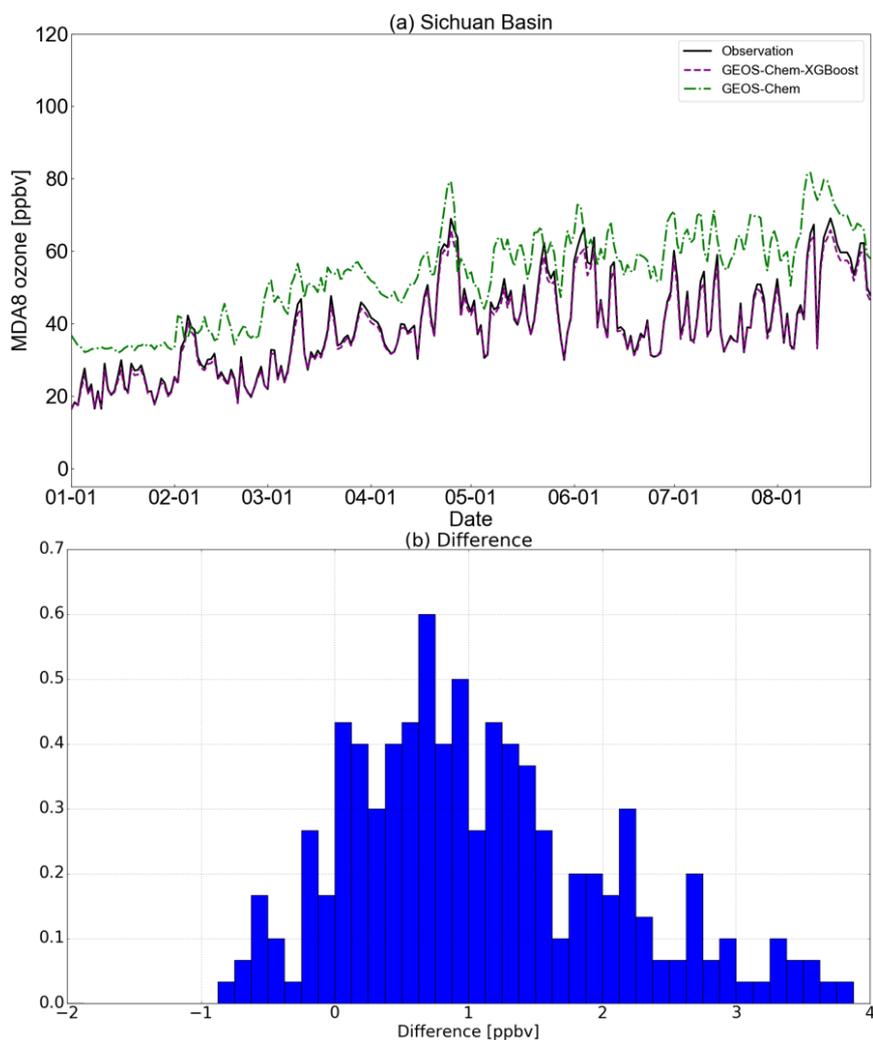
1

2   **Figure 3** Measured and modelled O₃ variabilities over the SCB in 2019 (a). Measured, GEOS-Chem, and GEOS-

3   Chem-XGBoost predicted O₃ values are denoted by black solid, grey dashed, and purple dashed lines, respectively.

4   (b) Histogram of the differences between the GEOS-Chem-XGBoost predictions and the measurements.

1
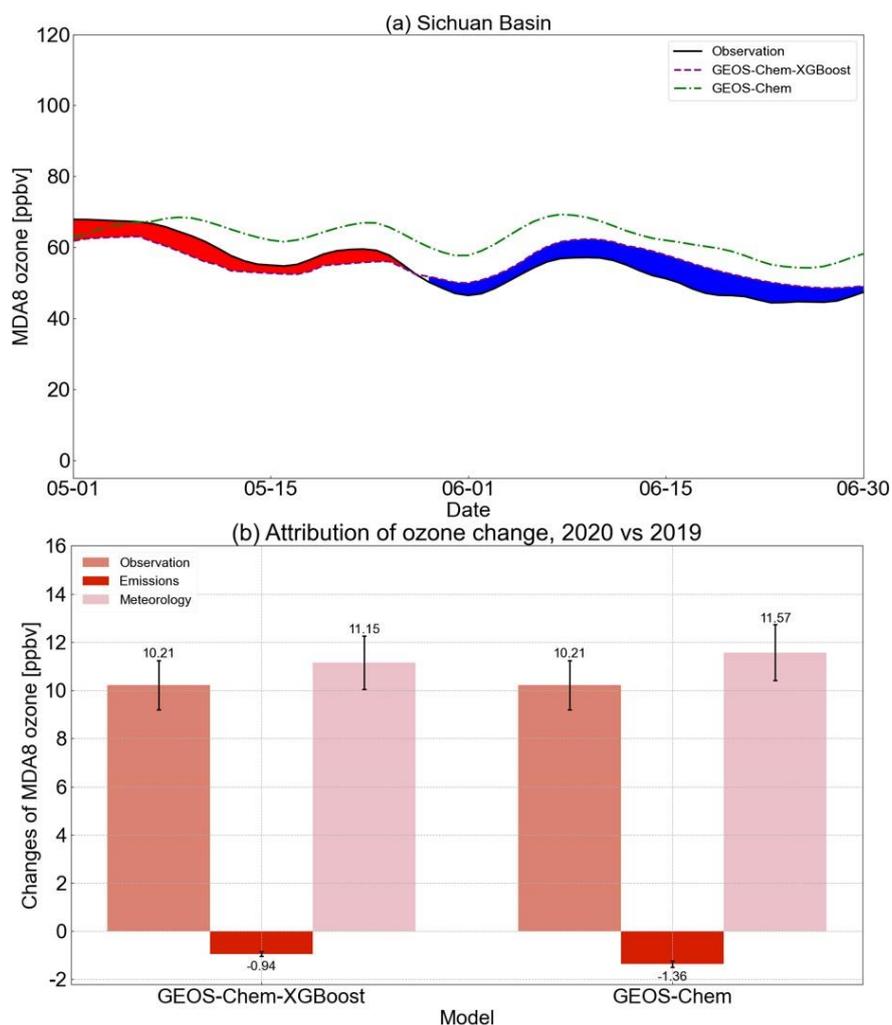
Figure 4 (a) Comparison of the GEOS-Chem-XGBoost O_3 predictions to the 2020 measurements. Red (blue)

shadings represent where GEOS-Chem-XGBoost predictions are higher (lower) than the actual measurements in

2020, indicating that changes in anthropogenic emission lead to O_3 increase (decrease) in 2020. (b) Attribution of

surface O_3 enhancements over the SCB in May-June 2020 vs. 2019. Filled colored bars denote O_3 change as seen

from measurements, and due to changes in anthropogenic emission and meteorological conditions estimated by the

GEOS-Chem-XGBoost model and the GEOS-Chem model. Black vertical bars represent 1σ STD across cities.
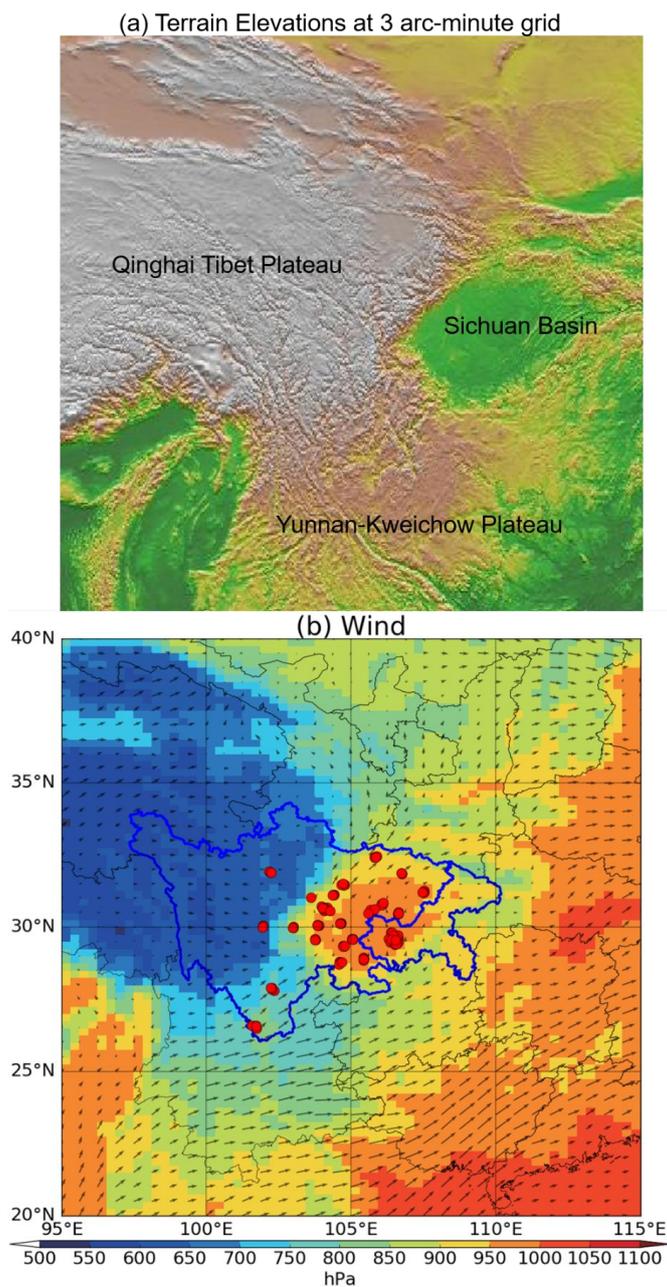
Figure 5 Terrain elevations (a) and surface temperature and wind fields (b) over the SCB and surrounding regions.

The spatial resolutions for (a) and (b) are 3 × 3 arc-minute and 0.25° × 0.25°, respectively. The white area in black

line is Tibetan Plateau (with altitudes of 4–5 km a.s.l.), the yellow area in red line is the Yunnan-Kweichou Plateau

(2–3 km a.s.l), the green area in circle is the SCB (0.5–1 km a.s.l).
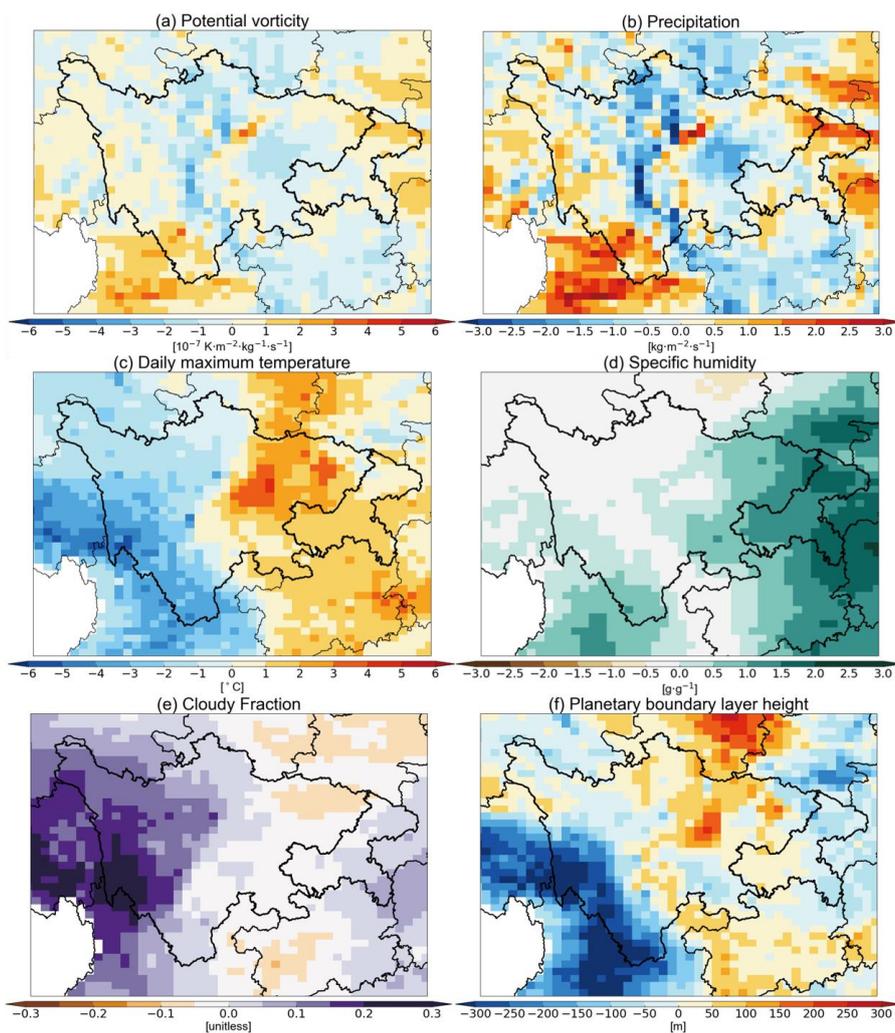
Figure 6 May-June mean differences in PV (a), precipitation (b), temperature (c), specific humidity (d), cloud fraction (e), and PBLH (f) between 2020 and 2019 over the SCB and surrounding regions. All these meteorological parameters are from the GEOS-FP dataset. PV is prescribed at the PBLH and others are at the surface.
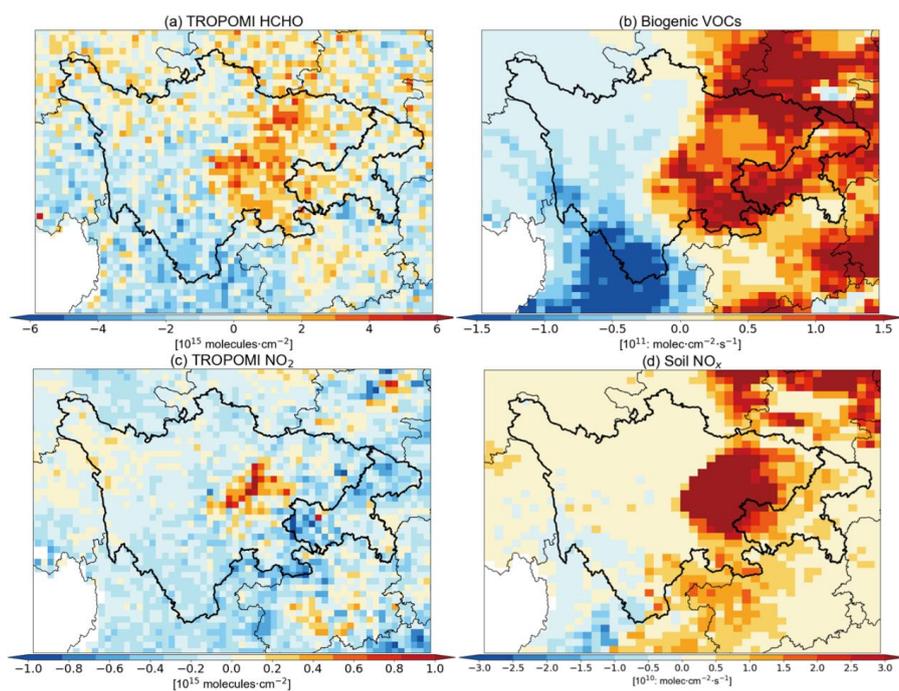
Figure 7 May-June mean differences in O₃ precursors between 2020 and 2019. (a) TROPOMI observed HCHO, (b) biogenic VOCs, (c) TROPOMI observed NO₂, and (d) Soil NOₓ. Biogenic VOCs and soil NOₓ are available from GEOS-Chem model online calculations.
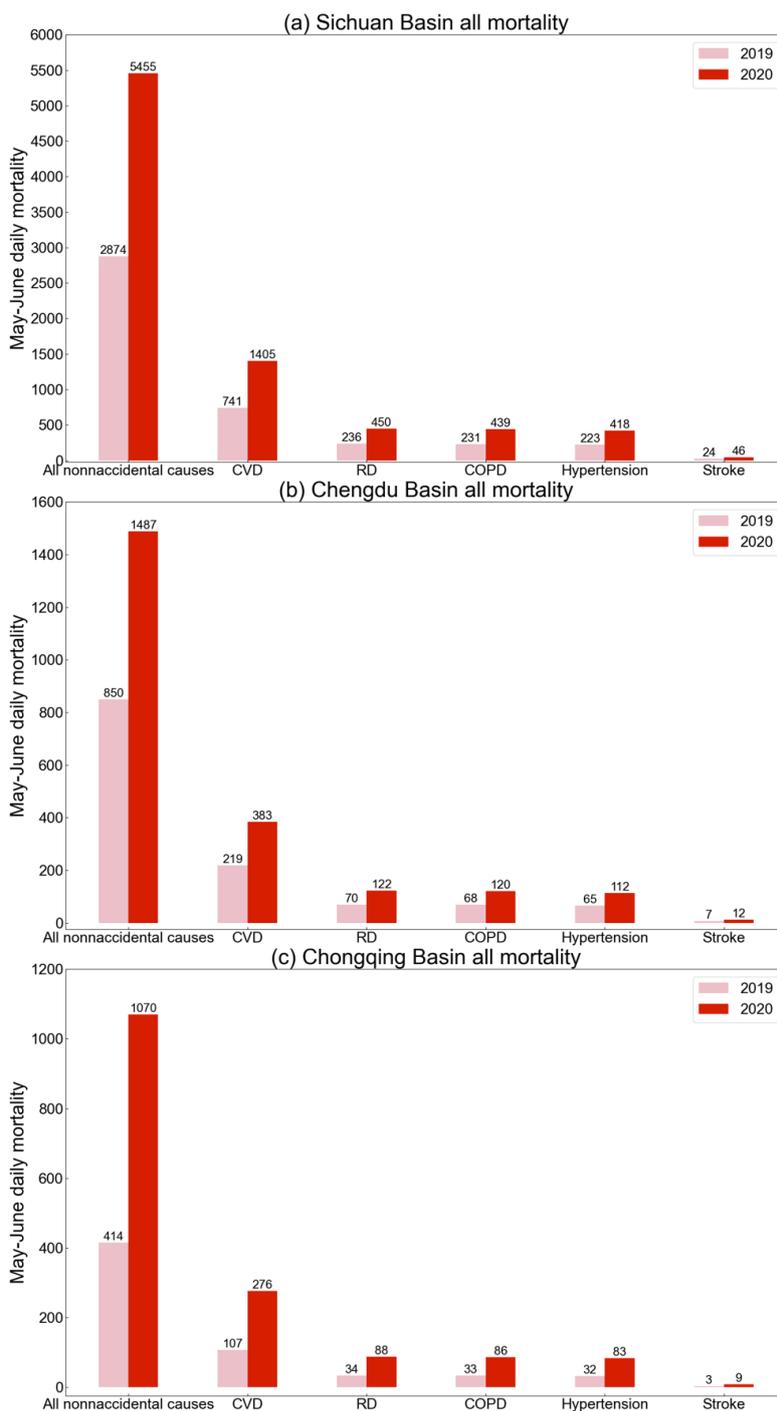
Figure 8 Total daily mortality from all non-accidental causes, CVD, RD, COPD, hypertension, and stroke attributable to ambient O_3 exposure over the SCB during May-June in 2019 and 2020.