https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Atmospheric Chemistry and Physics
Discussions

Ground-level gaseous pollutants across China: daily seamless mapping and long-term spatiotemporal variations

Jing Wei[1*], Zhanqing Li[1*], Jun Wang[2], Can Li[1], Pawan Gupta[3,4], Maureen Cribb[1]

1. Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
2. Department of Chemical and Biochemical Engineering, Iowa Technology Institute, Center for Global and Regional Environmental Research, University of Iowa, USA
3. STI, Universities Space Research Association (USRA), Huntsville, AL, USA
4. NASA Marshall Space Flight Center, Huntsville, AL, USA

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

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Abstract

Gaseous pollutants at the ground level seriously threaten the urban air quality environment and public health. There are few estimates of gaseous pollutants that are spatially and temporally resolved and continuous over long periods in China. This study takes advantage of big data and artificial intelligence technologies to generate seamless daily maps of three major pollutant gases, i.e., NO_2, SO_2, and CO, across China from 2013 to 2020 at a uniform spatial resolution of 10 km. Cross-validation illustrated a high data quality on a daily basis for NO_2, SO_2, and CO, with mean out-of-bag coefficients of determination (root-mean-square errors) of 0.84 (7.99 $\mu g/m^3$), 0.84 (10.7 $\mu g/m^3$), and 0.80 (0.29 mg/m^3), respectively. They have experienced significant declines and then recoveries during and after the COVID-19 lockdown associated with changes in anthropogenic emissions in eastern China, while surface CO recovered faster than SO_2 and NO_2. All gaseous pollutants decreased significantly by 0.23, 2.01, and 49 $\mu g/m^3$ per year ($p < 0.001$) across China during 2013–2020, especially in three urban agglomerations. The declining rates were larger during 2013–2017 but slowed down in recent years. Both the areas and occurrence probabilities of days exceeding air quality standards also gradually shrank and weakened over time, especially for SO_2 and CO, which almost disappeared during 2018–2020, suggesting significant improvements in air quality in China. This reconstructed dataset of surface gaseous pollutants, i.e., ChinaHighNO_2, ChinaHighSO_2, and ChinaHighCO, will benefit future (especially short-term) air pollution and environmental health-related studies.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

1. Introduction

Air pollution has been a major environmental concern, affecting human health, weather and climate (Kan et al., 2012; Kinney, 2008; Li et al., 2017b; Anenberg et al., 2022; Sun et al., 2010; Orellano et al., 2020; Murray et al., 2020), and has thus drawn worldwide attention. The sources of air pollution are complex. They include natural sources such as wildfires and anthropogenic emissions, including pollutants discharged from industrial production (e.g., smoke/dust, sulfur oxides, NO_x, and VOCs), hazardous substances released from burning coal during heating seasons (e.g., dust, SO_2, and CO), and waste gases (e.g., CO, SO_2, and NO_x) generated by transportation, especially in big cities. Among various air pollutants, the followings have been most widely recognized, particulate matters (e.g., $PM_{2.5}$ and PM_{10}), and gaseous pollutants (e.g., O_3, NO_2, SO_2, and CO, among others]. Many countries have built ground-based networks to monitor a variety of conventional pollutants in real-time. China has experienced serious ambient air pollution for a long time, prompting the establishment of a large-scale air quality monitoring network (Mee, 2018). Over the years, much effort has been made to model different species of air pollutants. Many studies on particulate matter (e.g., PM_1, $PM_{2.5}$, and PM_{10}) have been carried out with a focus on China (Ma et al., 2022; Fang et al., 2016; Wei et al., 2019b; Chen et al., 2018; Wei et al., 2021a; Wei et al., 2021b; Li et al., 2017a; Zhang et al., 2018). By contrast, ground-level gaseous pollutants have been much less studied. The global COVID-19 pandemic has motivated many attempts to estimate surface NO_2 concentrations (Tian et al., 2020; Who, 2020) from various satellite-retrieved tropospheric NO_2 products, e.g., OMI and TROPOMI, by adopting different statistical regression (Chi et al., 2021; Qin et al., 2017; Zhang et al., 2018) and artificial intelligence (Chi et al., 2022; Dou et al., 2021; Chen et al., 2019; Zhan et al., 2018; Wang et al., 2021; Liu, 2021) models. In contrast, studies on surface SO_2 and CO with a focus on whole of China are meager, limited by a lack of tropospheric satellite remote-sensing products and weaker signals (Li et al., 2020; Liu et al., 2019; Wang et al., 2021; Han et al., 2022). Such studies still face more challenges, e.g., satellite data gaps and missing values seriously limit their application and the neglect of the spatial and temporal differences in air pollution in the modeling process. In addition, most previous studies mainly focused on studying a single or a few species during relatively short periods of observation.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

As such, here, we aim to reconstruct a long-term daily seamless dataset of three ambient gaseous pollutants (i.e., surface NO_2, SO_2, and CO) in China at a uniform 10 km resolution to study air quality. We adopted the spatiotemporal ensemble-learning model to estimate three surface gaseous pollutants from big data. Using this dataset, the long-term spatiotemporal variations of the three gaseous pollutants and the impacts of implementing environmental protection policies and COVID-19 epidemic are investigated.

2. Materials and methods

2.1 Big data

2.1.1 Ground-based measurements

Major input data employed in the study are hourly routine measurements of the ground-level NO_2, SO_2, and CO concentrations of at approximately 2000 reference-grade ground-based monitoring stations across China from 2013 to 2020. Due to a change in the reference state implemented on 31 August 2018 (Mee, 2018), we first converted the concentrations of the three gaseous pollutants to the uniform standard condition (i.e., 273 K and 1013 hPa) for consistency. Daily values for each air pollutant at each station in each year were then averaged from valid hourly measurements that had undergone additional quality-control checks.

2.1.2 Satellite, reanalysis, and model data

Satellite remote sensing data used here include the daily seamless tropospheric NO_2 products (0.25° × 0.25°) generated by first combining OMI/Aura and Global Ozone Monitoring Experiment–2B retrievals (He et al., 2020), and then gap-filling using CAMS tropospheric NO_2 simulations via machine learning (Wei et al., 2022b), and MODIS monthly NDVI (0.05° × 0.05°), LandScan™ annual population (POP, 1 km) (Bright et al., 2000), and the SRTM digital elevation model (DEM, 90 m). ERA5-Land (0.1° × 0.1°) and ERA5 global reanalysis (0.25° × 0.25°) provided hourly meteorological fields (Muñoz-Sabater et al., 2021; Hersbach et al., 2020). The following eight variables form the reanalysis are employed in our study: 2-m temperature (TEM), precipitation (PRE), evaporation (ET), 10-m u- and v-components of wind (WU and WV), boundary-layer height (BLH), relative humidity (RH), and surface pressure (SP). Besides, model-simulated SO_2 and CO

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

surface mass concentrations were also included from the MERRA-2 and GEOS-FP global reanalysis every 1 and 3 hours at horizontal resolutions of 0.625° × 0.5° and 0.3125° × 0.25°, respectively. CAMS global reanalysis provided three-hour NO_2 simulations modeled on the earth's surface (every 3 hours, with a horizontal resolution of 0.75° × 0.75°) (Inness et al., 2019). Monthly ° × 0.1° anthropogenic emissions, i.e., NO_x, SO_2, and CO, were collected from CAMS global emission inventories. Here, for these fine-temporal-resolution variables, all hourly-level simulations in a day were first averaged for each grid to calculate daily means. All variables were aggregated or resampled into a 0.1° × 0.1° resolution for consistency.

2.2 Pollutant gas modelling

The current study is an extension of our previous work related to O_3 (Wei et al., 2022a), aimed at extending the long-term and full-coverage mapping of multi-type ground-level gaseous pollutants at a uniform grid resolution of 0.1° × 0.1° across China. Thus, the developed Space-Time Extra-Tree (STET) model was extended to estimate three additional species of surface pollutant gases, i.e., NO_2, SO_2, and CO. The uniqueness of this method is that it considers the autocorrelation and differences in air pollution in space and time, improving the model performance on the basis of ensemble learning.

For the surface NO_2 estimation, the STET was applied to the main input variables of NO_2 column amounts, model-simulated surface NO_2 concentrations, and NO_x emissions; together with the ancillary input variables of the aforementioned meteorological, land, and population (POP) variables, as denoted in Equation 1. Limited by long-term and high-resolution satellite tropospheric SO_2 and CO products, model-simulated surface SO_2 and CO concentrations and emissions were used as main predictors along with the same auxiliary variables as NO_2 to construct the STET model for separately estimating surface SO_2 (Equation 2) and CO (Equation 3):

$$NO_{2(ijt)} \sim f_{STET}(MNO_{2(ijt)}, SNO_{2(ijt)}, ENOx_{ijm}, Meteorology_{ijt}, NDVI_{ijm}, DEM_{ijy}, POP_{ijy}, P_s, P_t) \quad (1)$$

$$SO_{2(ijt)} \sim f_{STET}(MSO_{2(ijt)}, ESO_{2(ijm)}, Meteorology_{ijt}, NDVI_{ijm}, DEM_{ijy}, POP_{ijy}, P_s, P_t) \quad (2)$$

$$CO_{ijt} \sim f_{STET}(MCO_{ijt}, ECO_{ijm}, Meteorology_{ijt}, NDVI_{ijm}, DEM_{ijy}, POP_{ijy}, P_s, P_t) \quad (3)$$

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Atmospheric Chemistry and Physics
Discussions

where $NO_{2(ijt)}$, $SO_{2(ijt)}$, and CO_{ijt} indicate daily ground-based NO_2, SO_2, and CO measurements or estimations at one grid (i, j) on the t[th] day of a year; $MNO_{2(ijt)}$, $MSO_{2(ijt)}$, MCO_{ijt}, and $SNO_{2(ijt)}$ indicate daily model-simulated surface NO_2, SO_2, and CO concentrations and the satellite tropospheric NO_2 column amount at one grid (i, j) on the t[th] day of a year; $ENOx_{ijm}$, $ESO_{2(ijm)}$, and ECO_{ijm} indicate monthly anthropogenic NO_x, SO_2, and CO emissions at one grid (i, j) on the m[th] month of a year; $Meteorology_{ijt}$ represents each meteorological variable at one grid (i, j) on the t[th] day of a year; DEM_{ijy} and POP_{ijy} indicate the population at one grid (i, j) of a year; P_s and P_t indicate the space and time terms, represented by the longitudes and latitudes of spatial points and their distances to the center (D_{md}) and each corner of the study domain, and the day of the year (DOY), respectively.

Here, two widely-used 10-fold out-of-sample and out-of-station cross-validation (CV) methods (Wei et al., 2022a; Rodriguez et al., 2010) were employed to assess the data quality. They were performed by randomly dividing data samples and ground monitoring stations into independent training and testing datasets to evaluate the overall accuracy and prediction reliability, i.e., estimates for the samples and predictions for the stations that are excluded from training, respectively. Wei et al. (2022a) provides details about how these two methods work.

3. Results and discussion

3.1 Model performance

Using the constructed STET models, we generated daily 10-km resolution dataset with complete coverage (spatial coverage = 100%) for three ground-level gaseous pollutants from 2013 to 2020 in China, called ChinaHighNO₂, ChinaHighSO₂, and ChinaHighCO, respectively. They are all assembled to the ChinaHighAirPollutants (CHAP) dataset. Figure 1 shows the cross-validation results of all daily estimates and predictions for ground-level NO_2, SO_2, and CO concentrations in China (sample size: $N \approx 3.6$ million). Surface NO_2 and SO_2 concentrations fell between 200 and 500 $\mu g/m^3$, respectively, and daily estimates were highly correlated to observations, with the same coefficient of determination ($R^2 = 0.84$) and slopes (0.86 and 0.84) close to 1, respectively. Average root-mean-square errors (RMSEs) of surface NO_2 and SO_2 estimates were 7.99 and 10.07 $\mu g/m^3$, respectively, and mean absolute errors (MAEs) were 5.34 and 4.68 $\mu g/m^3$, respectively. Most

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

observed CO concentrations were less than 10 mg/m^3, agreeing well (R^2 = 0.80, slope = 0.79) with daily estimates, and the average RMSE (MAE) is 0.29 (0.16) mg/m^3. Compared to estimation accuracies (Figure 1a-c), prediction accuracies slightly decreased, which is acceptable considering the weak signals of trace gases. Daily SO_2, NO_2, and CO predictions (Figure 1d-f) show reasonable agreements with ground measurements (R^2 = 0.70, 0.68, and 0.61, respectively) and their respective RMSE (MAE) values were 14.28 (8.10) μg/m^3, 11.57 (7.06) μg/m^3, and 0.42 (0.24) mg/m^3, which basically represent the accuracy for areas without ground monitoring stations.

[Please insert Figure 1 here]

The performance of our air pollution modeling was also evaluated on an annual basis (Table 2). Our model works well in estimating and predicting the concentrations of different ground-level pollutant gases among different years. The model performance has continuously improved over time, as indicated by increasing correlations and decreasing uncertainties, because of increasing density of ground stations (especially in the suburban areas of cities) and improved quality control of measurements, significantly increasing the number (e.g., from 169 thousand in 2013 to more than 522 thousand in 2020) and quality of data samples.

[Please insert Table 2 here]

Figure 2 shows the individual-site-scale cross-validated accuracy and uncertainty in estimating daily pollutant gases in China. Our model has a strong ability to capture daily surface NO_2 concentrations at most stations in China, with about 80% (83%) of them having CV-R^2 > 0.7 (RMSE < 10 μg/m^3) with reference to ground measurements. By contrast, the model did not perform as well in estimating daily surface SO_2 and CO concentrations. Nevertheless, regarding SO_2, 80% of the stations had CV-R^2 values > 0.6, RMSE values were < 12 μg/m^3. Regarding CO, 83% of the stations had CV-R^2 values > 0.6, and RMSE values were generally < 0.4 mg/m^3.

[Please insert Figure 2 here]

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Figure 3 shows the monthly and yearly composites of ground-level NO_2, SO_2, and CO concentrations as a function of ground measurements from all monitoring stations in China for the years 2013 to 2020. On the monthly scale, we collected a total of ~119,000 matched samples of the three gaseous pollutants. Accuracies significantly improved, with increasing R^2 (decreasing RMSE) values of 0.93 (4.41 μg/m^3), 0.97 (4.03 μg/m^3), and 0.94 (0.13 mg/m^3) for NO_2, SO_2, and CO, respectively. On the annual scale, more than ~10,000 matched samples were collected, showing better agreement with observations (e.g., R^2 = 0.94, 0.98, and 0.97) and lower uncertainties (e.g., RMSE = 3.06 μg/m^3, 2.46 μg/m^3, and 0.07 mg/m^3) for NO_2, SO_2, and CO, respectively. These results illustrate the high quality of our dataset for different gaseous pollutants, illustrating its applicability to the investigations of short-term exposure and long-term variations.

[Please insert Figure 3 here]

3.2 Spatiotemporal variations

3.2.1 Spatial coverage and distribution

Figure 4 shows spatial distributions of the three pollutant gases across China on a typical day (1 January 2018). The spatial patterns of these gaseous pollutants are consistent with those observed on the ground, especially in highly polluted areas, e.g., severe surface NO_2 pollution in the North China Plain (NCP) and high surface SO_2 emissions in Shanxi Province. The unique advantage of our dataset is that it can provide valuable gaseous pollutant information on a daily basis at locations in China where ground measurements are not available. This addresses the major issues of scanning gaps and the numerous missing values in satellite remote sensing retrievals at cloudy locations, e.g., average daily spatial coverage of the OMI tropospheric NO_2 product is only 42%. Our dataset provides spatial-complete coverage, significantly increasing the daily data utilization by 58%.

[Please insert Figure 4 here]

Figure 5 shows seasonal maps for each gas pollutant during the period 2013–2020 across China. Pollutant gases vary significantly in space and time across China, where high surface NO_2 levels are

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

mainly distributed in typical urban agglomerations, e.g., the Beijing-Tianjin-Hebei (BTH) region, the Yangtze River and Pearl River Deltas (YRD and PRD), and scattered large cities with intensive human activities and highly developed transportation systems (e.g., Urumqi, Chengdu, Xi'an, and Wuhan, among others). High surface SO_2 concentrations are mainly observed in northern China (e.g., Shanxi, Hebei, and Shandong Provinces), associated with combustion emissions from anthropogenic sources, and the Yunnan Guizhou Plateau in southwest China, likely associated with emissions from volcanic eruptions. By contrast, except in some areas in central China (e.g., Shanxi and Hebei), surface CO concentrations are overall low.

Over time, significant differences in spatial patterns were seen. Surface NO_2, SO_2, and CO in summer (average = 15.9 ± 4.7 µg/m^3, 22.9 ± 13.4 µg/m^3, and 1.1 ± 0.3 mg/m^3, respectively) were the lowest, thanks to favorable meteorological conditions. Pollution levels were highest in winter, with average values increasing by ~1.5–1.9 times those in summer. This difference was much larger in central and eastern China, e.g., 2.3–3.4 times higher in the BTH, due to large amounts of direct NO_x, SO_2, and CO emissions from burning coal for heating in winter in northern China. Spring and autumn show relatively similar spatial patterns among three gaseous pollutants.

[Please insert Figure 5 here]

3.2.2 Short-term epidemic effects on air quality

The unique advantage of seamless day-to-day gaseous pollutant maps allows us to investigate the COVID-19 effects (Who, 2020) in China. Here, we compared the relative differences of each air pollutant from February to April between 2020 and 2019 in eastern China (Figure 6). In February, surface NO_2 sharply reduced by more than 30% in eastern China, especially in key urban agglomerations and megacities (relative change > 50%). A significant decrease in surface SO_2 (> 40%) was observed in northern areas where heavy industry is the mainstay in China, while there was little change in southern China. Surface CO also showed drastic decreases, especially in southeast China, but the amplitude was smaller than those of the other two gaseous pollutants. These were attributed to extensive plant closures and traffic controls due to the lockdown, which started at the end of January 2020, significantly reducing anthropogenic NO_x, SO_2, and CO

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

emissions (Ding et al., 2020; Zheng et al., 2021).

In March, surface NO_2 concentrations were still largely lower than typical levels in most areas, especially in areas hit hardest by the epidemic, i.e., Hubei province and its surrounding areas (relative change > 30%). By comparison, the decline in surface SO_2 largely slowed by more than two times in the NCP and Shanxi Province. However, surface CO returned to normal levels the fastest. In April, surface NO_2 and SO_2 were comparable to historical concentrations (within ± 10%) or alternately changed across eastern China due to rebounding anthropogenic emissions (Ding et al., 2020), indicating that they were almost recovered. In addition, the three air pollutants fell within ± 10% around Wuhan, Hubei Province, indicating that people there had returned to normal life.

[Please insert Figure 6 here]

3.2.3 Long-term trends and policy implications

To better investigate the spatiotemporal variations of surface air pollution, we calculated linear trends and significance levels using monthly anomalies by removing seasonal cycles (Wei et al., 2019a). Given that monitoring stations were sparse and unevenly distributed in western China, especially in earlier years, we will focus on eastern part of the country for the trend analysis. Figure 7 shows annual mean maps of the three gaseous pollutants for each year from 2013 to 2020 in eastern China. They all have changed greatly in the past eight years across China, peaking around 2017 and have declined to their lowest levels in 2020, at annual mean surface decreasing rates of -12%, -55%, and -31% across China for NO_2, SO_2, and CO concentrations, respectively.

[Please insert Figure 7 here]

Most of eastern China showed more significant decreasing trends, especially in three urban agglomerations (trend = -0.51~-1.21 μg/m^3/yr, $p < 0.001$), as well as in other large cities (e.g., Wuhan and Chengdu) (Figure 8, and Table 3). The largest downward trends mainly occurred in northern and central China, especially in the BTH (trend = -6.01 and 109 μg/m^3/yr, $p < 0.001$, respectively). This is mainly due to the change in fuel for heating from coal to gas widespread

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

across China in winter (Wang et al., 2020), greatly reducing the emissions of precursor gases (Koukouli et al., 2018). Increasing trends were, however, also found in Ningxia and Shanxi in central China.

We divided the study into four periods to investigate the impact of implemented major emission control measures taken in China. During the Clear Air Action Plan (CAAP, 2013–2017), surface SO_2 and CO concentrations significantly decreased in most parts of eastern China, while surface NO_2 reductions are limited to some places. The rates of decreases for all three pollutants accelerated in recent years, especially for NO_2 in southeast China (trend > 2 $\mu g/m^3/yr$, $p < 0.05$). These are thanks to the dramatic reductions in all main pollutant emissions (e.g., PM, SO_2, and NO_x) in key regions (especially urbans) through the upgrading of key industries, industrial structure adjustment, and coal-fired boiler remediation.

During the 13rd Five-Year-Plan (FYP, 2016–2020), surface NO_2 decreased at a rate of -0.46 $\mu g/m^3$ per year ($p < 0.001$) across China, with larger decreases in the BTH region and Hunan Province (trend > 2 $\mu g/m^3/yr$, $p < 0.05$). More striking decreasing trends were found in southeast China. Surface SO_2 also significant decreased but slowed down in eastern China. However, a greater downward trend was observed in Shanxi Province, mainly due to the reduction in coal consumption. Surface CO also continuously decreased, more rapidly in central China but less rapidly elsewhere. The continuous decline in all pollutants is due to the binding reduction by 10–15% of total emissions of major pollutants like COD, SO_2, and NO_x.

During the Blue Sky Defense War (2018–2020), the majority of pollutants have dropped considerably, of which surface SO_2 changed the most in China (average = -27%), especially in central and eastern China (relative change > 50%). In addition, SO_2 decreased by 23–44% in three key regions. Followed by surface CO with the national and regional mean reductions of 17% and 10–24%, respectively. By contrast, surface NO_2 had the least reduction at an average decreasing rate of 11% across China and 9–16% for the three key regions. The improvement in air quality benefited from continuous reductions (by 2–3%) in total air pollutant emissions, coordinated reductions in greenhouse gas emissions, and the impacts of COVID-19.

[Please insert Figure 8 and Table 3 here]

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions

With the daily seamless datasets, we can investigate the number of days exceeding their respective air quality standard (Level 2 limitation) in a given year to evaluate the distribution and variations of short-term pollution exposure (Figure 9). The areal extent of regions exposed to unacceptably high pollutant-gas levels (i.e., daily NO_2 > 80 μg/m^3, SO_2 > 150 μg/m^3, and CO > 4 mg/m^3) were usually small. NO_2 pollution was mainly found in the NCP and a handful of big cities (e.g., Xi'an, Wuhan, Guangzhou, and Shanghai), changing little over time. Surface SO_2 pollution was mainly observed in in central China (e.g., Hebei, Shandong, and Shanxi), with the areal extent of polluted regions gradually decreasing over time until almost disappearing by 2020. The same was seen for surface CO pollution, being worst in the BTH region and its surrounding areas before 2018, then almost disappearing by 2020. Surface NO_2 pollution was mainly observed in developed urban areas (e.g., Beijing, Tianjin, Shijiazhuang, and Wuhan), with 15% of the days exceeding the acceptable standard in the early part of the study period, then decreasing to below 5% afterward.

[Please insert Figure 9 here]

Regionally, significant differences in the percentage of days with pollution levels exceeding national standards are seen (Figure 10). For example, the BTH region was the only region experiencing a high NO_2 exposure risk, which gradually lessened from 2013 to 2018, when the exposure risk reached zero. For surface SO_2, no high exposure days (daily mean > 150 μg/m^3) were observed regionally, but a large number of days exceeding the Level 1 limitation (i.e., daily mean > 50 μg/m^3) were found in the BTH region. However, the number of days gradually decreased until reaching zero after 2017. For surface CO, less than 1% of highly polluted days are found in the BTH region in some individual years. These results suggest that with the unrelenting efforts of air pollution prevention and control, the number of days with high pollution has been significantly reduced across China, indicating significant improvements in air quality.

[Please insert Figure 10 here]

© Author(s) 2022. CC BY 4.0 License.

3.3 Model comparison

Long-term datasets (at least 2 years) of different gaseous pollutants generated with different developed models in previous studies focusing on the whole of China are compared here. Only those studies applying the same validation approach were selected (Table 4). Most generated surface NO_2 datasets had low spatial resolutions (~0.125°–0.25°) with numerous missing values limited by OMI satellite observations (Zhan et al., 2018; Dou et al., 2021; Chen et al., 2019; Xu et al., 2019; Chi et al., 2021). Some studies improved the spatial resolutions by introducing NO_2 data from the recently launched Sentinel-5 TROPOMI satellite but can only provide dataset after October 2018 (Chi et al., 2022; Liu, 2021; Wang et al., 2021; Wei et al., 2022b). Surface SO_2 estimated from a SO_2 emission inventory and surface CO from the MOPITT and TROPOMI retrievals (Li et al., 2020; Liu et al., 2019; Wang et al., 2021) have a much lower data quality (Li et al., 2020; Li et al., 2017b; Wang et al., 2021). Overall, our gaseous pollutant datasets are superior to those from the studies listed in Table 4 in terms of either overall accuracy, or spatial coverage, or length of data records.

[Please insert Table 4 here]

3.4 Successful applications

Our surface gaseous pollutant datasets have been freely available to the public online since March 2021 and have now been successfully employed for various application studies in environment and health. Strong associations and negative effects between ambient gaseous pollution (e.g., NO_2, SO_2, and CO) and a variety of diseases has been demonstrated for people of all ages through multi-regional and national cohort studies in China. These diseases include general mortality (Zhang et al., 2022), cause-specific cardiovascular disease (Xu et al., 2022a), ischemic and hemorrhagic stroke (Xu et al., 2022b; Wu et al., 2022b; Cai et al., 2022; He et al., 2022), dementia mortality (Liu et al., 2022), blood pressure (Song et al., 2022; Wu et al., 2022a), renal function (Li et al., 2022a), neurodevelopmental delay (Su et al., 2022), serum liver enzymes (Li et al., 2022b), overweight and obesity (Chen et al., 2022b), insomnia (Xu et al., 2021), subjective sleep quality (Wang et al., 2022), and visual impairment (Chen et al., 2022a). These studies attest well to the value of the

CHAP dataset with some unique merits for the sake of public health, among others, now and in the future.

4. Summary and conclusions

Exposure to air pollution is detrimental to human health, which has been a major public concern in heavy polluted regions like China where ground-based observation of pollutants is relatively short and sparse than major developed countries. Moreover, the pollutants to travel long distances to affect large down steam regions. To remedy the limitations of ground-based air pollution observation, this study applies the machine learning model of Space-Time Extra-Tree to estimate ambient gaseous pollutants across China, with extensive input variables measured on the monitors, satellites, and models. The estimated quantities are daily 10 km resolution (about 0.1 degrees) seamless (spatial coverage = 100%) dataset for ground-level NO_2, SO_2, and CO concentrations in China since 2013. This dataset was cross-evaluated in terms of overall accuracy and predictive ability at different spatiotemporal levels. Daily estimates (predictions) of surface NO_2, SO_2, and CO from 2013 to 2020 across China are highly consistent with observations with average sample-based (station-based) CV-R^2 of 0.84 (0.68), 0.84 (0.7), and 0.8 (0.61), and RMSEs of 7.99 (11.57) $\mu g/m^3$, 10.7 (14.28) $\mu g/m^3$, and 0.29 (0.42) mg/m^3y.

Pollutant gas concentrations varied significantly in the region, where high levels were mainly found in the Northern China, especially in winter. All gaseous pollutants sharply declined in eastern China during the COVID-19 outbreak, then gradually returned to historical levels. The recovery speed of surface CO was faster than for NO_2 and SO_2. Over time and at the national scale, they have significantly ($p < 0.001$) decreased by 0.23, 2.01, and 43 $\mu g/m^3$ per year during 2013–2020. Larger reductions were found at the regional scale, especially the BTH. Improvements in air quality were larger in the last decade or so but have slowed down in recent years. In particular, the areal extents of regions experiencing air pollution and the probability of air pollution occurring have also gradually decreased over time, especially during the period 2018–2020. This may be related with the implementation of a series of tough environmental protection policies, which greatly reduced anthropogenic emissions and significantly improved air quality. This high-quality daily seamless dataset will benefit future environmental and health-related studies focused on China, especially

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

studies investigating short-term air pollution exposure.

Data availability

CNEMC gaseous pollutants measurements are available at http://www.cnemc.cn; Reconstructed OMI/Aura tropospheric NO_2 product is available at https://doi.org/10.6084/m9.figshare.13126847; MODIS series products and MERRA2 reanalysis are available at https://search.earthdata.nasa.gov/; SRTM DEM is available at https://www2.jpl.nasa.gov/srtm/; LandScan™ population is available at https://landscan.ornl.gov/; ERA5 reanalysis is available at https://cds.climate.copernicus.eu/; GEOS CF data is available at https://portal.nccs.nasa.gov/datashare/gmao/; CAMS reanalysis and emission inventory are available at https://ads.atmosphere.copernicus.eu/.

CHAP dataset availability

The ChinaHighAirPollutants (CHAP) dataset is open access and freely available at https://weijing-rs.github.io/product.html. The ChinaHighNO_2 dataset is available at https://doi.org/10.5281/zenodo.4641542, the ChinaHighSO_2 dataset is available at https://doi.org/10.5281/zenodo.4641538, and the ChinaHighCO dataset is available at https://doi.org/10.5281/zenodo.4641530.

Acknowledgments

JW, and ZL, and JW were supported by the NASA grants (80NSSC21K1980 and 80NSSC19K0950). PG was supported by NASA's Research Opportunities in Space and Earth Science (ROSES-2020), Program Element A.38: Health and Air Quality Applied Sciences Team.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions
Open Access

References

Anenberg, S. C., Mohegh, A., Goldberg, D. L., Kerr, G. H., Brauer, M., Burkart, K., Hystad, P., Larkin, A., Wozniak, S., and Lamsal, L.: Long-term trends in urban NO₂ concentrations and associated paediatric asthma incidence: estimates from global datasets, The Lancet Planetary Health, 6, e49-e58, ttps://doi.org/10.1016/S2542-5196(21)00255-2, 2022.

Bright, E. A., Coleman, P. R., Dobson, J. E. J. P. E., and Sensing, R.: LandScan: A Global Population Database for Estimating Populations at Risk, 66, 849-858, 2000.

Cai, M., Zhang, S., Lin, X., Qian, Z., McMillin, S. E., Yang, Y., Zhang, Z., Pan, J., and Lin, H.: Association of Ambient Particulate Matter Pollution of Different Sizes With In-Hospital Case Fatality Among Stroke Patients in China, Neurology, 10.1212/WNL.0000000000200546, https://doi.org/10.1212/WNL.0000000000200546, 2022.

Chen, G., Wang, Y., Li, S., Cao, W., Ren, H., Knibbs, L. D., Abramson, M. J., and Guo, Y.: Spatiotemporal patterns of PM10 concentrations over China during 2005–2016: A satellite-based estimation using the random forests approach, Environmental Pollution, 242, 605-613, https://doi.org/10.1016/j.envpol.2018.07.012, 2018.

Chen, L., Wei, J., Ma, T., Gao, D., Wang, X., Wen, B., Chen, M., Li, Y., Jiang, J., Wu, L., Li, W., Liu, X., Song, Y., Guo, X., Dong, Y., and Ma, J.: Ambient gaseous pollutant exposure and incidence of visual impairment among children and adolescents: findings from a longitudinal, two-center cohort study in China, Environmental Science and Pollution Research, https://doi.org/10.1007/s11356-022-20025-3, 2022a.

Chen, L., Gao, D., Ma, T., Chen, M., Li, Y., Ma, Y., Wen, B., Jiang, J., Wang, X., Zhang, J., Chen, S., Wu, L., Li, W., Liu, X., Guo, X., Huang, S., Wei, J., Song, Y., Ma, J., and Dong, Y.: Could greenness modify the effects of physical activity and air pollutants on overweight and obesity among children and adolescents?, Science of The Total Environment, 832, 155117, https://doi.org/10.1016/j.scitotenv.2022.155117, 2022b.

Chen, Z.-Y., Zhang, R., Zhang, T.-H., Ou, C.-Q., and Guo, Y.: A kriging-calibrated machine learning method for estimating daily ground-level NO2 in mainland China, Science of The Total Environment, 690, 556-564, https://doi.org/10.1016/j.scitotenv.2019.06.349, 2019.

Chi, Y., Fan, M., Zhao, C., Sun, L., Yang, Y., Yang, X., and Tao, J.: Ground-level NO2 concentration estimation based on OMI tropospheric NO2 and its spatiotemporal characteristics in typical regions of China, Atmospheric Research, 264, 105821, https://doi.org/10.1016/j.atmosres.2021.105821, 2021.

Chi, Y., Fan, M., Zhao, C., Yang, Y., Fan, H., Yang, X., Yang, J., and Tao, J.: Machine learning-based estimation of ground-level NO2 concentrations over China, Science of The Total Environment, 807, 150721, https://doi.org/10.1016/j.scitotenv.2021.150721, 2022.

Ding, J., van der A, R. J., Eskes, H. J., Mijling, B., Stavrakou, T., van Geffen, J. H. G. M., and Veefkind, J. P.: NOx Emissions Reduction and Rebound in China Due to the COVID-19 Crisis, 47, e2020GL089912, https://doi.org/10.1029/2020GL089912, 2020.

Dou, X., Liao, C., Wang, H., Huang, Y., Tu, Y., Huang, X., Peng, Y., Zhu, B., Tan, J., Deng, Z., Wu, N., Sun, T., Ke, P., and Liu, Z.: Estimates of daily ground-level NO2 concentrations in China based on Random Forest model integrated K-means, Advances in Applied Energy, 2, 100017, https://doi.org/10.1016/j.adapen.2021.100017, 2021.

Fang, X., Zou, B., Liu, X., Sternberg, T., and Zhai, L.: Satellite-based ground PM2.5 estimation using timely structure adaptive modeling, Remote Sensing of Environment, 186, 152-163,

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions
Open Access
EGU

https://doi.org/10.1016/j.rse.2016.08.027, 2016.

Han, W., He, T. L., Tang, Z., Wang, M., Jones, D., and Jiang, Z.: A comparative analysis for a deep learning model (hyDL-CO v1.0) and Kalman filter to predict CO concentrations in China, Geosci. Model Dev., 15, 4225-4237, https://doi.org/10.5194/gmd-15-4225-2022, 2022.

He, F., Wei, J., Dong, Y., Liu, C., Zhao, K., Peng, W., Lu, Z., Zhang, B., Xue, F., Guo, X., and Jia, X.: Associations of ambient temperature with mortality for ischemic and hemorrhagic stroke and the modification effects of greenness in Shandong Province, China, Science of The Total Environment, 158046, https://doi.org/10.1016/j.scitotenv.2022.158046, 2022.

He, Q., Qin, K., Cohen, J. B., Loyola, D., Li, D., Shi, J., and Xue, Y.: Spatially and temporally coherent reconstruction of tropospheric NO₂ over China combining OMI and GOME-2B measurements, Environmental Research Letters, 15, 125011, https://doi.org/10.1088/1748-9326/abc7df, 2020.

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

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A. M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V. H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515-3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Kan, H., Chen, R., and Tong, S.: Ambient air pollution, climate change, and population health in China, Environment International, 42, 10-19, https://doi.org/10.1016/j.envint.2011.03.003, 2012.

Kinney, P. L.: Climate Change, Air Quality, and Human Health, American Journal of Preventive Medicine, 35, 459-467, https://doi.org/10.1016/j.amepre.2008.08.025, 2008.

Koukouli, M. E., Theys, N., Ding, J., Zyrichidou, I., Mijling, B., Balis, D., and van der A, R. J.: Updated SO2 emission estimates over China using OMI/Aura observations, Atmos. Meas. Tech., 11, 1817-1832, https://doi.org/10.5194/amt-11-1817-2018, 2018.

Li, R., Cui, L., Liang, J., Zhao, Y., Zhang, Z., and Fu, H.: Estimating historical SO2 level across the whole China during 1973–2014 using random forest model, Chemosphere, 247, 125839, https://doi.org/10.1016/j.chemosphere.2020.125839, 2020.

Li, S., Meng, Q., Laba, C., Guan, H., Wang, Z., Pan, Y., Wei, J., Xu, H., Zeng, C., Wang, X., Jiang, M., Lu, R., Guo, B., and Zhao, X.: Associations between long-term exposure to ambient air pollution and renal function in Southwest China: The China Multi-Ethnic Cohort (CMEC) study, Ecotoxicology and Environmental Safety, 242, 113851, https://doi.org/10.1016/j.ecoenv.2022.113851, 2022a.

Li, T., Shen, H., Yuan, Q., Zhang, X., and Zhang, L.: Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach, 44, 11,985-911,993, https://doi.org/10.1002/2017GL075710, 2017a.

Li, Y., Yuan, X., Wei, J., Sun, Y., Ni, W., Zhang, H., Zhang, Y., Wang, R., Xu, R., Liu, T., Yang, C.,

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Chen, G., Xu, J., and Liu, Y.: Long-term exposure to ambient air pollution and serum liver enzymes in older adults: A population-based longitudinal study, Annals of Epidemiology, 74, 1-7, https://doi.org/10.1016/j.annepidem.2022.05.011, 2022b.
Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., and Zhu, B.: Aerosol and boundary-layer interactions and impact on air quality, National Science Review, 4, 810-833, https://doi.org/10.1093/nsr/nwx117, 2017b.
Liu, D., Di, B., Luo, Y., Deng, X., Zhang, H., Yang, F., Grieneisen, M. L., and Zhan, Y.: Estimating ground-level CO concentrations across China based on the national monitoring network and MOPITT: potentially overlooked CO hotspots in the Tibetan Plateau, Atmos. Chem. Phys., 19, 12413-12430, https://doi.org/10.5194/acp-19-12413-2019, 2019.
Liu, J.: Mapping high resolution national daily NO2 exposure across mainland China using an ensemble algorithm, Environmental Pollution, 279, 116932, https://doi.org/10.1016/j.envpol.2021.116932, 2021.
Liu, T., Zhou, Y., Wei, J., Chen, Q., Xu, R., Pan, J., Lu, W., Wang, Y., Fan, Z., Li, Y., Xu, L., Cui, X., Shi, C., Zhang, L., Chen, X., Bao, W., Sun, H., and Liu, Y.: Association between short-term exposure to ambient air pollution and dementia mortality in Chinese adults, Science of The Total Environment, 849, 157860, https://doi.org/10.1016/j.scitotenv.2022.157860, 2022.
Ma, Z., Dey, S., Christopher, S., Liu, R., Bi, J., Balyan, P., and Liu, Y.: A review of statistical methods used for developing large-scale and long-term PM2.5 models from satellite data, Remote Sensing of Environment, 269, 112827, https://doi.org/10.1016/j.rse.2021.112827, 2022.
MEE: Revision of the Ambien air quality standards (GB 3095-2012) (in Chinese), Ministry of Ecology and Environment, available at: http://www.mee.gov.cn/xxgk2018/xxgk/xxgk2001/201808/t20180815_20629602.html, 2018.
Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J. N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349-4383, 10.5194/essd-13-4349-2021, 2021.
Murray, C. J. L., Aravkin, A. Y., Zheng, P., and al., e.: Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, The Lancet, 396, 1223-1249, https://doi.org/10.1016/S0140-6736(20)30752-2, 2020.
Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., and Ciapponi, A.: Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis, Environment International, 142, 105876, https://doi.org/10.1016/j.envint.2020.105876, 2020.
Qin, K., Rao, L., Xu, J., Bai, Y., Zou, J., Hao, N., Li, S., and Yu, C.: Estimating Ground Level NO2 Concentrations over Central-Eastern China Using a Satellite-Based Geographically and Temporally Weighted Regression Model, 9, 950, 2017.
Rodriguez, J. D., Perez, A., and Lozano, J. A.: Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32, 569-575, 10.1109/TPAMI.2009.187, 2010.
Song, J., Du, P., Yi, W., Wei, J., Fang, J., Pan, R., Zhao, F., Zhang, Y., Xu, Z., Sun, Q., Liu, Y.,

© Author(s) 2022. CC BY 4.0 License.

Chen, C., Cheng, J., Lu, Y., Li, T., Su, H., and Shi, X.: Using an Exposome-Wide Approach to Explore the Impact of Urban Environments on Blood Pressure among Adults in Beijing–Tianjin–Hebei and Surrounding Areas of China, Environmental Science & Technology, https://doi.org/10.1021/acs.est.1c08327, 2022.

Su, X., Zhang, S., Lin, Q., Wu, Y., Yang, Y., Yu, H., Huang, S., Luo, W., Wang, X., Lin, H., Ma, L., and Zhang, Z.: Prenatal exposure to air pollution and neurodevelopmental delay in children: A birth cohort study in Foshan, China, Science of The Total Environment, 816, 151658, https://doi.org/10.1016/j.scitotenv.2021.151658, 2022.

Sun, Q., Hong, X., and Wold, L. E.: Cardiovascular Effects of Ambient Particulate Air Pollution Exposure, 121, 2755-2765, https://doi.org/10.1161/CIRCULATIONAHA.109.893461, 2010.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U. G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O. N., Yang, R., Grenfell, B. T., Pybus, O. G., and Dye, C.: An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China, Science, 368, 638-642, https://doi.org/10.1126/science.abb6105, 2020.

Wang, L., Zhang, J., Wei, J., Zong, J., Lu, C., Du, Y., and Wang, Q.: Association of ambient air pollution exposure and its variability with subjective sleep quality in China: A multilevel modeling analysis, Environmental Pollution, 312, 120020, https://doi.org/10.1016/j.envpol.2022.120020, 2022.

Wang, S., Su, H., Chen, C., Tao, W., Streets, D. G., Lu, Z., Zheng, B., Carmichael, G. R., Lelieveld, J., Pöschl, U., and Cheng, Y.: Natural gas shortages during the “coal-to-gas” transition in China have caused a large redistribution of air pollution in winter 2017, 117, 31018-31025, https://doi.org/10.1073/pnas.2007513117, 2020.

Wang, Y., Yuan, Q., Li, T., Zhu, L., and Zhang, L.: Estimating daily full-coverage near surface O3, CO, and NO2 concentrations at a high spatial resolution over China based on S5P-TROPOMI and GEOS-FP, ISPRS Journal of Photogrammetry and Remote Sensing, 175, 311-325, https://doi.org/10.1016/j.isprsjprs.2021.03.018, 2021.

Wei, J., Peng, Y., Mahmood, R., Sun, L., and Guo, J.: Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products, Atmos. Chem. Phys., 19, 7183-7207, https://doi.org/10.5194/acp-19-7183-2019, 2019a.

Wei, J., Li, Z., Guo, J., Sun, L., Huang, W., Xue, W., Fan, T., and Cribb, M.: Satellite-Derived 1-km-Resolution PM1 Concentrations from 2014 to 2018 across China, Environmental Science & Technology, 53, 13265-13274, https://doi.org/10.1021/acs.est.9b03258, 2019b.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1-km-resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sensing of Environment, 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021a.

Wei, J., Li, Z., Xue, W., Sun, L., Fan, T., Liu, L., Su, T., and Cribb, M.: The ChinaHighPM10 dataset: generation, validation, and spatiotemporal variations from 2015 to 2019 across China, Environment International, 146, 106290, https://doi.org/10.1016/j.envint.2020.106290, 2021b.

Wei, J., Li, Z., Li, K., Dickerson, R. R., Pinker, R. T., Wang, J., Liu, X., Sun, L., Xue, W., and Cribb, M.: Full-coverage mapping and spatiotemporal variations of ground-level ozone (O3) pollution from 2013 to 2020 across China, Remote Sensing of Environment, 270, 112775, https://doi.org/10.1016/j.rse.2021.112775, 2022a.

© Author(s) 2022. CC BY 4.0 License.

Wei, J., Liu, S., Li, Z., Liu, C., Qin, K., Liu, X., Pinker, R. T., Dickerson, R. R., Lin, J., Boersma, K. F., Sun, L., Li, R., Xue, W., Cui, Y., Zhang, C., and Wang, J.: Ground-Level NO2 Surveillance from Space Across China for High Resolution Using Interpretable Spatiotemporally Weighted Artificial Intelligence, Environmental Science & Technology, 10.1021/acs.est.2c03834, 2022b.

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

Wu, H., Zhang, Y., Zhao, M., Liu, W., Magnussen, C. G., Wei, J., and Xi, B.: Short-term effects of exposure to ambient PM1 on blood pressure in children and adolescents aged 9 to 18 years in Shandong Province, China, Atmospheric Environment, 283, 119180, https://doi.org/10.1016/j.atmosenv.2022.119180, 2022a.

Wu, H., Lu, Z., Wei, J., Zhang, B., Liu, X., Zhao, M., Liu, W., Guo, X., and Xi, B.: Effects of the COVID-19 Lockdown on Air Pollutant Levels and Associated Reductions in Ischemic Stroke Incidence in Shandong Province, China, 10, https://doi.org/10.3389/fpubh.2022.876615, 2022b.

Xu, H., Bechle, M. J., Wang, M., Szpiro, A. A., Vedal, S., Bai, Y., and Marshall, J. D.: National PM2.5 and NO2 exposure models for China based on land use regression, satellite measurements, and universal kriging, Science of The Total Environment, 655, 423-433, https://doi.org/10.1016/j.scitotenv.2018.11.125, 2019.

Xu, J., Zhou, J., Luo, P., Mao, D., Xu, W., Nima, Q., Cui, C., Yang, S., Ao, L., Wu, J., Wei, J., Chen, G., Li, S., Guo, Y., Zhang, J., Liu, Z., and Zhao, X.: Associations of long-term exposure to ambient air pollution and physical activity with insomnia in Chinese adults, Science of The Total Environment, 792, 148197, https://doi.org/10.1016/j.scitotenv.2021.148197, 2021.

Xu, R., Shi, C., Wei, J., Lu, W., Li, Y., Liu, T., Wang, Y., Zhou, Y., Chen, G., Sun, H., and Liu, Y.: Cause-specific cardiovascular disease mortality attributable to ambient temperature: A time-stratified case-crossover study in Jiangsu province, China, Ecotoxicology and Environmental Safety, 236, 113498, https://doi.org/10.1016/j.ecoenv.2022.113498, 2022a.

Xu, R., Wang, Q., Wei, J., Lu, W., Wang, R., Liu, T., Wang, Y., Fan, Z., Li, Y., Xu, L., Shi, C., Li, G., Chen, G., Zhang, L., Zhou, Y., Liu, Y., and Sun, H.: Association of short-term exposure to ambient air pollution with mortality from ischemic and hemorrhagic stroke, n/a, https://doi.org/10.1111/ene.15343, 2022b.

Zhan, Y., Luo, Y., Deng, X., Zhang, K., Zhang, M., Grieneisen, M. L., and Di, B.: Satellite-Based Estimates of Daily NO2 Exposure in China Using Hybrid Random Forest and Spatiotemporal Kriging Model, Environmental Science & Technology, 52, 4180-4189, 10.1021/acs.est.7b05669, 2018.

Zhang, Y., Li, Z., Wei, J., Zhan, Y., Liu, L., Yang, Z., Zhang, Y., Liu, R., and Ma, Z.: Long-term exposure to ambient NO2 and adult mortality: A nationwide cohort study in China, Journal of Advanced Research, https://doi.org/10.1016/j.jare.2022.02.007, 2022.

Zhang, Z., Wang, J., Hart, J. E., Laden, F., Zhao, C., Li, T., Zheng, P., Li, D., Ye, Z., and Chen, K.: National scale spatiotemporal land-use regression model for PM2.5, PM10 and NO2 concentration in China, Atmospheric Environment, 192, 48-54, https://doi.org/10.1016/j.atmosenv.2018.08.046, 2018.

Zheng, B., Zhang, Q., Geng, G.-n., Chen, C., Shi, Q., Cui, M., Lei, Y., and He, K. J. E. S. S. D.: Changes in China's anthropogenic emissions and air quality during the COVID-19 pandemic in 2020, 2021.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics
Discussions

Figures

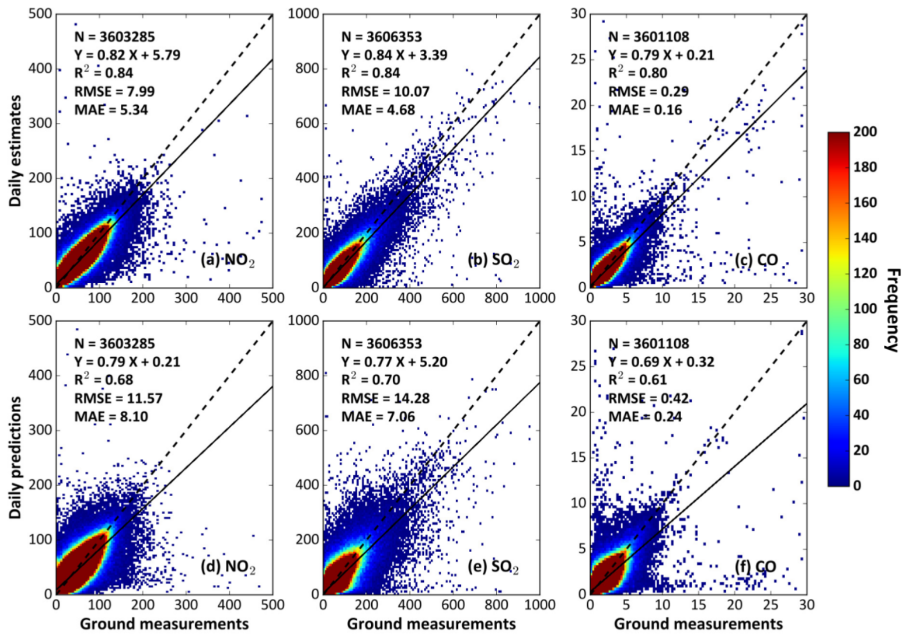


Figure 1. Density plots of daily (a-c) estimates and (d-f) predictions of ground-level NO_2 (μg/m^3), SO_2 (μg/m^3), and CO (mg/m^3) concentrations as a function of ground measurements in China from 2013 to 2020 using the out-of-sample (top panels) and out-of-station (bottom panels) cross-validation methods.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

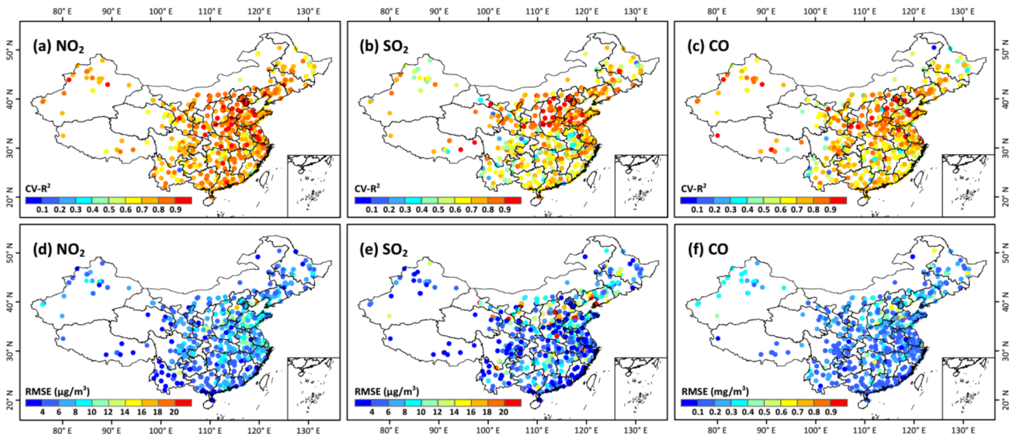


Figure 2. Validation of daily ground-level three NO_2 (μg/m^3), SO_2 (μg/m^3), and CO (mg/m^3) estimates at each individual monitoring station in China from 2013 to 2020: (a-c) accuracy (i.e., CV-R^2) and (d-f) uncertainty (i.e., RMSE).

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

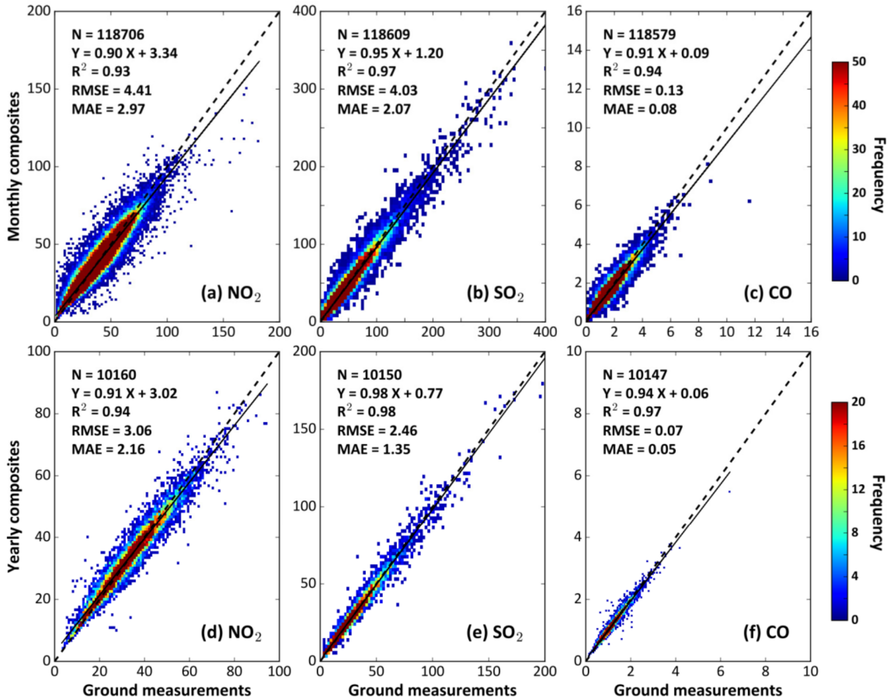
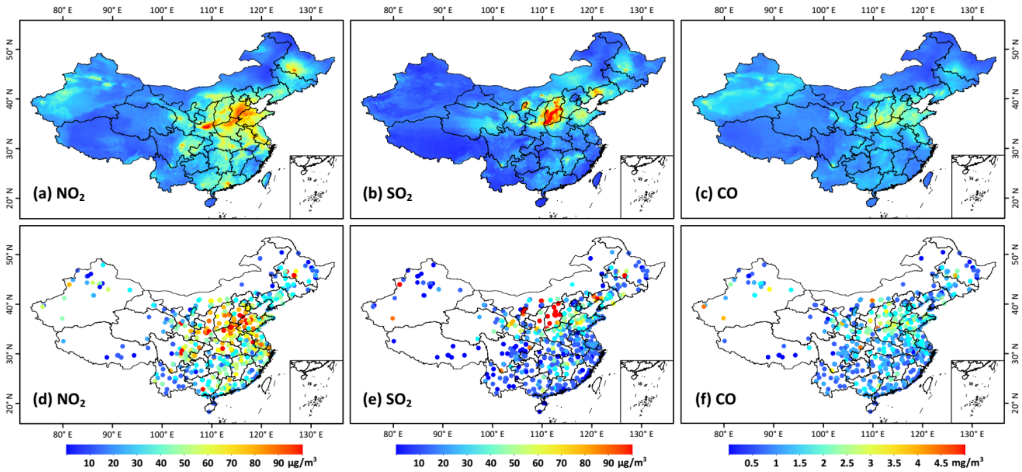


Figure 3. Validation results of (a-c) monthly and (d-f) yearly composites of ground-level NO_2 ($\mu g/m^3$), SO_2 ($\mu g/m^3$), and CO (mg/m^3) against ground measurements from all monitoring stations in China for the years 2013 to 2020. Black lines are best-fit lines from linear regression, and dashed lines are 1:1 lines.



644

645 **Figure 4.** Comparisons between (a-c) big-data-derived (horizontal resolution = 10 km) seamless
646 ground-level NO_2 ($\mu g/m^3$), SO_2 ($\mu g/m^3$), and CO (mg/m^3) concentrations and (d-f) corresponding
647 ground measurements on 1 January 2018 in China.

648

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Atmospheric Chemistry and Physics
Discussions

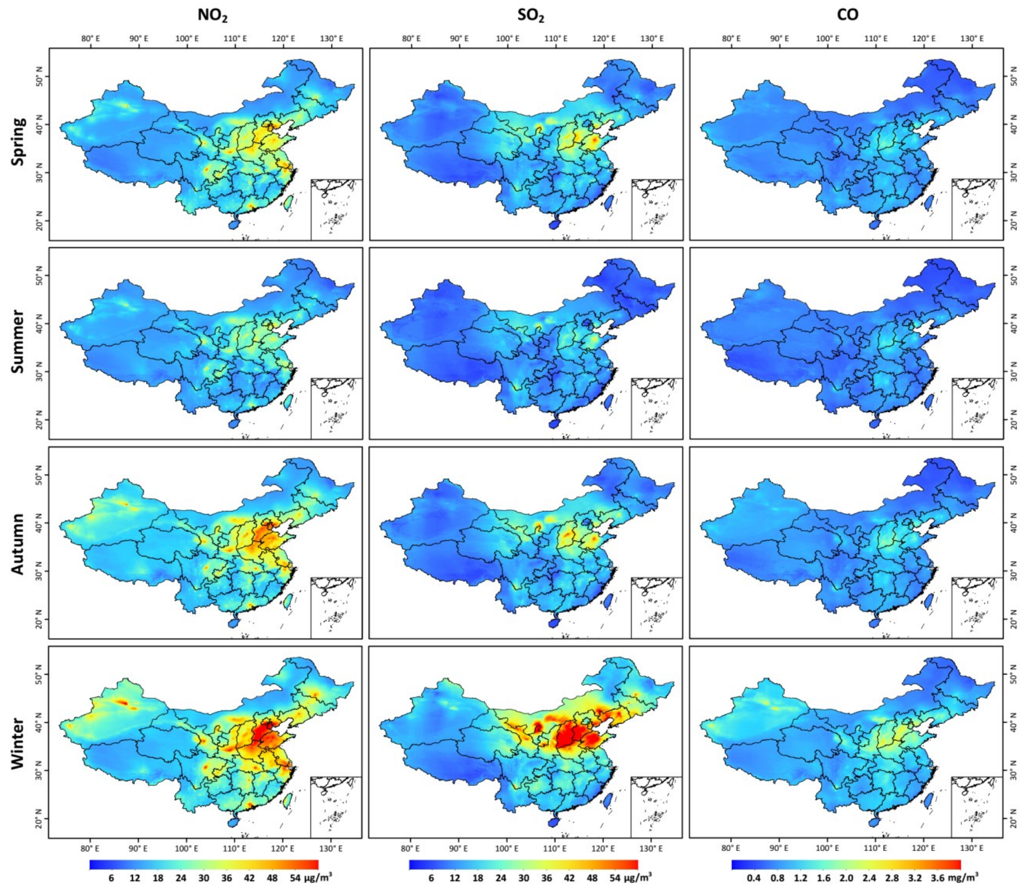


Figure 5. Seasonal mean maps (horizonal resolution = 10 km) of ground-level NO_2 ($\mu g/m^3$), SO_2 ($\mu g/m^3$), and CO (mg/m^3) concentrations averaged for 2013–2020 in China.

© Author(s) 2022. CC BY 4.0 License.

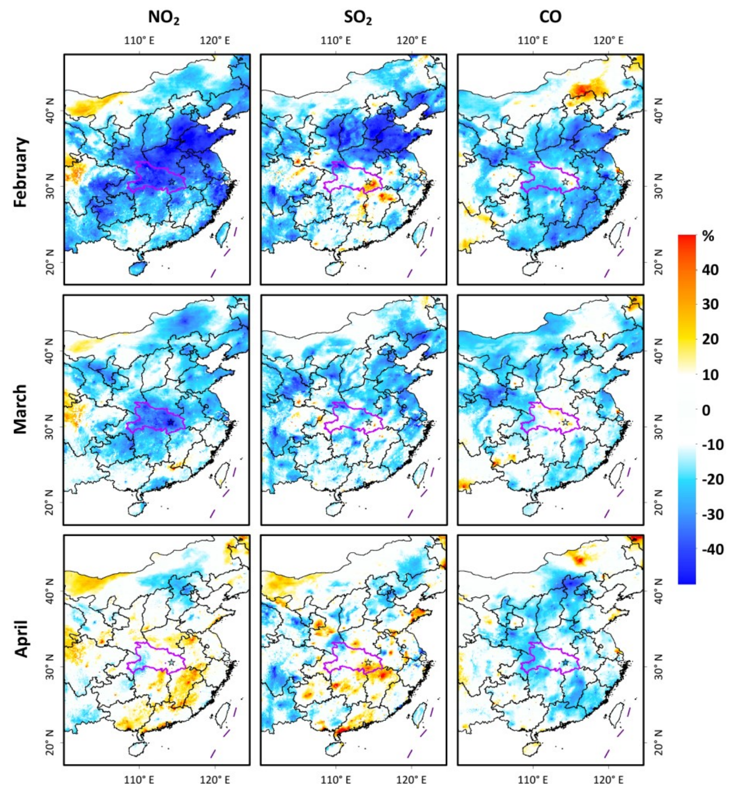


Figure 6. Relative changes (%) in mean ground-level NO_2, SO_2, and CO concentrations (μg/m^3) in February, March, and April between 2019 and 2020 during the COVID-19 epidemic across the East China. The area outlined in magenta and the star indicate Hubei Province and Wuhan City, respectively.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Atmospheric Chemistry and Physics
Open Access
Discussions
EGU

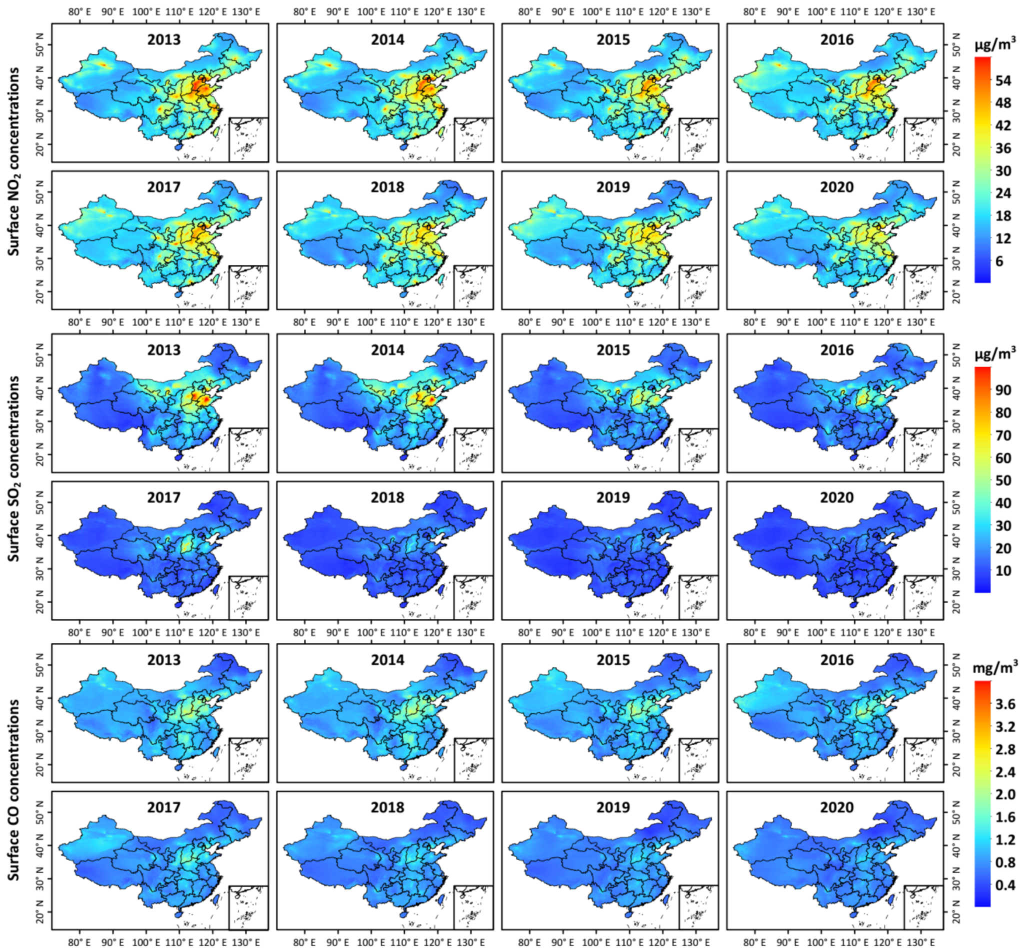


Figure 7. Spatial distributions of annual mean (horizonal resolution = 10 km) of ground-level NO_2 (μg/m^3), SO_2 (μg/m^3), and CO (mg/m^3) concentrations for each year from 2013 to 2020 in China.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

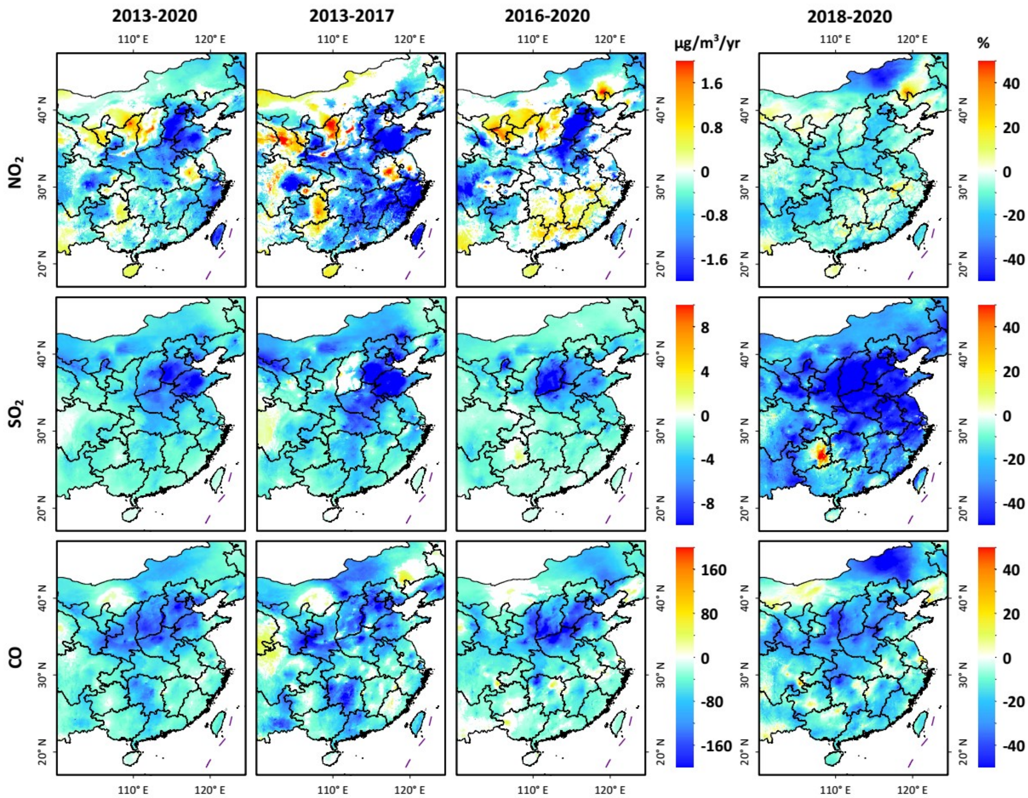


Figure 8. Temporal trends (μg/m^3/yr) of ground-level NO_2, SO_2, and CO concentrations in eastern China during the whole period (2013–2020), Clean Air Action Plan (2013–2017), 13rd Five-Year Plan (2016–2020), where only regions with trends that are significant at the 95% ($p < 0.05$) confidence level are shown, and relative changes (%) during the Blue Sky Defense War (2018–2020).

© Author(s) 2022. CC BY 4.0 License.

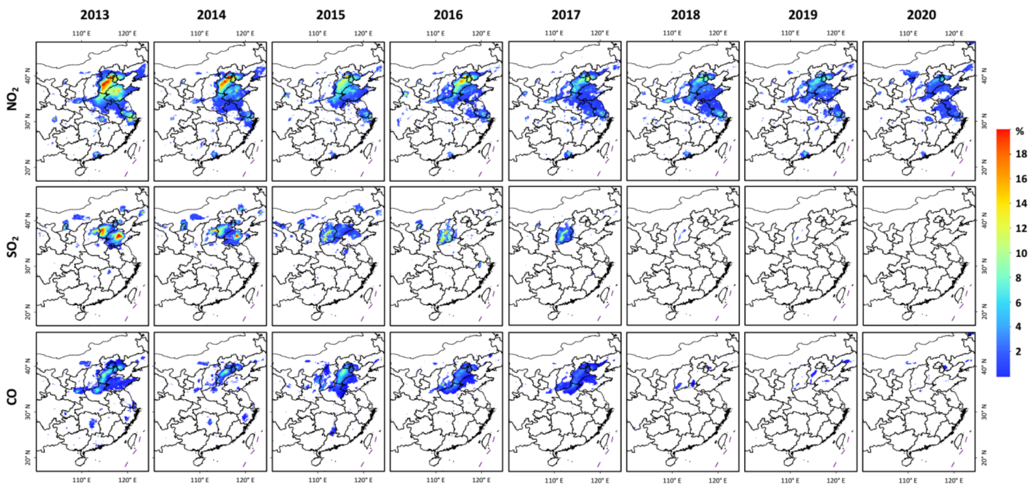


Figure 9. Spatial distributions of the percentage of polluted days exceeding air quality standards for ground-level NO_2 (daily mean > 80 μg/m^3), SO_2 (daily mean > 150 μg/m^3), and CO (daily mean > 4 mg/m^3) for each year from 2013 to 2020 in eastern China.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

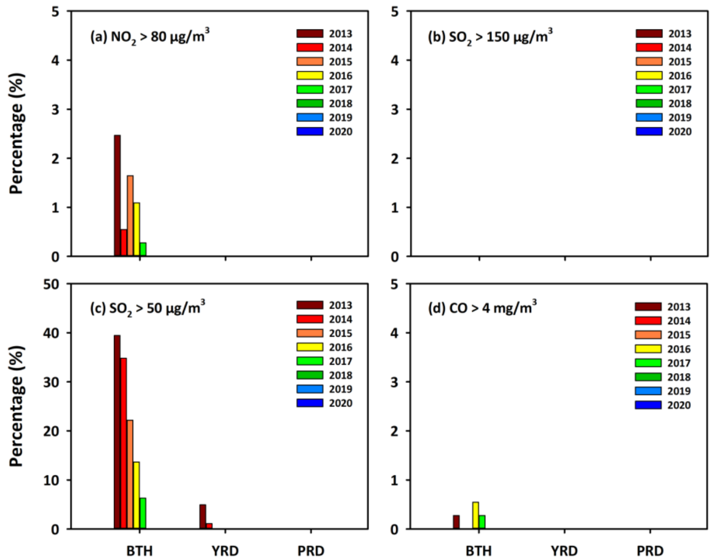


Figure 10. Percentage of days exceeding the air quality standards for ground-level (a) NO_2 (daily mean > 80 μg/m^3), (b) SO_2 (daily mean > 150 μg/m^3), (c) SO_2 (daily mean > 50 μg/m^3), and (d) CO (daily mean > 4 mg/m^3) for each year from 2013 to 2020 in three typical urban agglomerations in China.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions
Open Access EGU

Tables

Table 1. Summary of big data used in this study.

Category	Scientific Dataset	Abbreviation	Spatial Resolution	Temporal Resolution	Time Period	Data Source
Measurements	NO_2, SO_2, CO	-	In-situ	Hourly	2013–2020	MEE
Satellite remote sensing products	Tropospheric NO_2 column	NO_2	0.25°×0.25°	Daily	2013–2020	(He et al., 2020)
	Normalized difference vegetation index	NDVI	0.05°×0.05°	Monthly	2013–2020	MOD13C2
	Surface elevation	DEM	90 m	-	-	SRTM
	Population distribution	POP	1 km	Annual	2013–2020	LandScan^{TM}
Model simulation	2-m air temperature	TEM	0.1°×0.1°	Hourly	2013–2020	ERA5 reanalysis
	Precipitation	PRE				
	Evaporation	ET				
	Surface pressure	SP				
	10-m u-component of wind	WU				
	10-m v-component of wind	WV				
	Boundary-layer height	BLH	0.25°×0.25°			
	Relative humidity	RH				
	SO_2 surface mass concentration	SO_2	0.3125°×0.25°	3-hour	2015–2020	GEOS-FP reanalysis
	CO surface concentration	CO				
	SO_2 surface mass concentration	SO_2	0.625°×0.5°	Hourly	2013–2020	MERRA2 reanalysis
	CO surface concentration	CO				
	NO_2 surface concentration	NO_2	0.75°×0.75°	3-hour	2013–2020	CAMS reanalysis
	Carbon monoxide	CO	0.1°×0.1°	Monthly	2013–2020	CAMS emission
	Nitrogen oxides	NO_x				
	Sulphur dioxide	SO_2				

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions
Open Access
EGU

682 **Table 2.** Statistics of the overall accuracies and predictive abilities of ambient gaseous pollutants for
683 each year in China from 2013 to 2020.

Year	Sample size	Overall accuracy						Predictive ability					
		NO_2		SO_2		CO		NO_2		SO_2		CO	
	N (10^3)	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
2013	169	0.77	12.48	0.83	17.97	0.80	0.56	0.53	18.16	0.68	25.04	0.60	0.78
2014	324	0.76	10.97	0.83	15.87	0.77	0.38	0.54	15.56	0.66	22.45	0.51	0.57
2015	518	0.79	9.34	0.80	13.71	0.74	0.38	0.61	13.10	0.61	19.49	0.50	0.55
2016	516	0.82	8.59	0.83	11.26	0.76	0.34	0.64	12.20	0.65	16.28	0.57	0.46
2017	527	0.86	7.57	0.86	7.79	0.82	0.24	0.72	10.67	0.74	10.80	0.70	0.32
2018	513	0.87	6.92	0.83	5.61	0.82	0.20	0.76	9.33	0.68	7.80	0.69	0.26
2019	515	0.87	6.78	0.81	4.84	0.82	0.20	0.77	9.23	0.66	6.63	0.70	0.25
2020	522	0.89	5.78	0.80	4.02	0.82	0.17	0.79	8.04	0.62	5.57	0.69	0.23

684

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
Atmospheric Chemistry and Physics Discussions

Table 3. Statistics of temporal trends (μg/m^3/yr) and relative changes (%) of ground-level NO_2, SO_2, and CO concentrations during the whole period (T_{All}, 2013–2020), the Clear Air Action Plan (T_{CAAP}, 2013–2017), the 13[rd] Five-Year-Plan (T_{FYP}, 2016–2020), and the Blue Sky Defense War (Δ_{BSDW}, 2018–2020) in China and three typical regions.

Region	NO_2				SO_2				CO			
	T_{All}	T_{CAAP}	T_{FYP}	Δ_{BSDW}	T_{All}	T_{CAAP}	T_{FYP}	Δ_{BSDW}	T_{All}	T_{CAAP}	T_{FYP}	Δ_{BSDW}
China	-0.23***	-0.06	-0.46***	-11	-2.01***	-2.28***	-1.54***	-27	-49***	-43***	-50***	-17
BTH	-1.21***	-1.04***	-1.43***	-15	-6.01***	-7.78***	-3.78***	-44	-109***	-114***	-97***	-24
YRD	-0.58***	-0.88***	-0.33	-9	-3.13***	-3.53***	-2.57***	-44	-40***	-50***	-30***	-10
PRD	-0.51***	-0.93**	-0.21	-16	-2.01***	-3.11***	-0.80***	-23	-58***	-78***	-23***	-14

Note: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

https://doi.org/10.5194/acp-2022-627
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.

Table 4. Comparison with long-term datasets of different gaseous pollutants focusing on the whole of China generated in previous related studies.

Species	Model	Missing values	Spatial resolution	Main input	Validation period	CV-R^2	RMSE	Literature
NO_2	RF-STK	Yes	0.25°	OMI	2013–2016	0.62	13.3	(Zhan et al., 2018)
	RF-K	Yes	0.25°	OMI	2013–2018	0.64	11.4	(Dou et al., 2021)
	KCS	Yes	0.125°	OMI	2014–2016	0.72	7.9	(Chen et al., 2019)
	LUR	Yes	0.125°	OMI	2014–2015	0.78	-	(Xu et al., 2019)
	LME	Yes	0.1°	OMI	2014–2020	0.65	7.9	(Chi et al., 2021)
	XGBoost	Yes	0.125°	TROPOMI	2018–2020	0.67	6.4	(Chi et al., 2022)
	XGBoost	Yes	0.05°	TROPOMI	2018–2019	0.83	7.6	(Liu, 2021)
	LightGBM	No	0.05°	TROPOMI	2018–2020	0.83	6.6	(Wang et al., 2021)
	SWDF	No	0.01°	TROPOMI	2019–2020	0.93	4.9	(Wei et al., 2022b)
	STET	No	0.1°	Big data	2013–2020	0.84	8.0	This study
SO_2	RF	No	0.25°	Emissions	2013–2014	0.64	17.1	(Li et al., 2020)
	STET	No	0.1	Big data	2013–2020	0.84	10.1	This study
CO	RF–STK	Yes	0.1	MOPITT	2013–2016	0.51	0.54	(Liu et al., 2019)
	LightGBM	No	0.07°	TROPOMI	2018–2020	0.71	0.26	(Wang et al., 2021)
	STET	No	0.1°	Big data	2013–2020	0.80	0.29	This study

KCS: kriging-calibrated satellite method; LightGBM: light gradient boosted model; LME: linear mixed effect model; LUR: land use regression; MOPITT: Measurements of Pollution in the Troposphere; OMI: Ozone Monitoring Instrument; RF: random forest; RF-K: random forest integrated with K-means; RF-STK: random-forest-spatiotemporal-kriging model; STET: space-time extremely randomized tree; SWDF: spatiotemporally weighted deep forest; TROPOMI: TROPOspheric Monitoring Instrument; XGBoost: extreme gradient boosting