Satellite-based long-term spatiotemporal trends in ambient NO_2 concentrations and attributable health burdens in China from 2005 to 2020

Keyong Huang[1,2], Qingyang Zhu[3], Xiangfeng Lu[1,2*], Dongfeng Gu[1,2,4], Yang Liu[3*]

[1]Department of Epidemiology, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishi Road, Xicheng District, Beijing 100037, China

[2]Key Laboratory of Cardiovascular Epidemiology, Chinese Academy of Medical Sciences, Beijing, China

[3]Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

[4]School of Medicine, Southern University of Science and Technology, Shenzhen 518055, China

***Correspondence to:**

Yang Liu, PhD

Professor and Chair

Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Rd., Atlanta, GA 30322 USA.

Email: yang.liu@emory.edu

Xiangfeng Lu, PhD

Department of Epidemiology, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishi Road, Xicheng District, Beijing 100037, China

Email: xiangfenglu@sina.com

Key Points:

- We developed an ensemble machine learning model on NO_2 levels using column densities from OMI satellite as main predictors.
- Our model obtained overall, temporal and spatial cross-validation R2 of 0.88, 0.82 and 0.73, respectively.
- The estimated annual mortality burden attributable to chronic NO_2 exposure ranged from 305 thousand to 416 thousand in China.

ABSTRACT

Limited research has assessed the spatio-temporal distribution and chronic health effects of NO_2 exposure, especially in developing countries, due to the lack of historical NO_2 data. A gap-filling model was first adopted to impute the missing NO_2 column densities from satellite, then an ensemble machine learning model incorporating three base learners was developed to estimate the spatiotemporal pattern of monthly mean NO_2 concentrations at 0.05° spatial resolution from 2005 to 2020 in China. Further, we applied the exposure dataset with epidemiologically derived exposure response relations to estimate the annual NO_2 associated mortality burdens in China. The coverage of satellite NO_2 column densities increased from 46.9% to 100% after gap-filling. The ensemble model predictions had good agreement with observations, and the overall, temporal and spatial cross-validation (CV) R^2 were 0.88, 0.82 and 0.73, respectively. In addition, our model can provide accurate historical NO_2 concentrations, with both by-year CV R^2 and external separate year validation R^2 achieving 0.80. The estimated national NO_2 levels showed a increasing trend during 2005-2011, then decreased gradually until 2020, especially in 2012-2015. The estimated annual mortality burden attributable to long-term NO_2 exposure ranged from 305 thousand to 416 thousand, and varied considerably across provinces in China. This satellite-based ensemble model could provide reliable long-term NO_2 predictions at a high spatial resolution with complete coverage for environmental and epidemiological studies in China. Our results also highlighted the heavy disease burden by NO_2 and call for more targeted policies to reduce the emission of nitrogen oxides in China.

Plain Language Summary

This study developed a satellite-based ensemble machine learning model to predict 16-year NO_2 levels and identified high mortality burden attributed to NO_2 in China with great implications for environmental policy making.

1 Introduction

Ambient nitrogen dioxide (NO_2) is a major air pollutant, mainly originating from traffic and fuel combustion emissions. Several epidemiological studies have found that exposure to ambient NO_2 was associated with decreased lung function, cardiopulmonary diseases, and premature deaths independent of other air pollutants (S Huang *et al.*, 2021b; Meng *et al.*, 2021; Strassmann *et al.*, 2021). In addition, ambient NO_2 is a key precursor of a series of secondary pollutants, such as ozone and fine particulate matter ($PM_{2.5}$). In 2021, the World Health Organization (WHO) tightened the air quality guideline, reducing the annul NO_2 standard level from 40 μg/m^3 to 10 μg/m^3. However, mainly due to the lack of historical surface NO_2 data before 2013, the long-term spatiotemporal trend of NO_2 and its chronic health effects were rarely reported in China (Yin *et al.*, 2020).

With a high spatiotemporal coverage, satellite remote sensing technology has become a promising tool to estimate surface air pollutants, and shown potential to fill the gap left by ground fixed monitors, especially in regions with sparse monitoring (Cooper *et al.*, 2022; Di *et al.*, 2020; K Huang *et al.*, 2018). Correspondingly, several statistical models have been developed to convert satellite data to surface air pollutants, such as land use regression, geographically weighted regression and machine learning algorithms (Geddes *et al.*, 2016; C Huang *et al.*, 2021a; Song *et al.*, 2019; Zhan *et al.*, 2018). For example, Geddes et al. estimated the global NO_2 concentrations at 10-km resolution from 1996 to 2012 using NO_2 tropospheric column densities from satellite intruments (Geddes *et al.*, 2016). In China, Zhan et al. developed a hybrid random forest and spatiotemporal kriging model using NO_2 column densities from satellite Ozone Monitoring Instrument (OMI), and predicted surface NO_2 levels at 10-km resolution from 2013 to 2016 (Zhan *et al.*, 2018). Despite the valuable NO_2 predictions from previous studies, there are still some aspects to be improved to promote epidemiological study and disease burden estimation in China. First, most of the national or regional NO_2 estimations were conducted at a coarse spatial resolution in China (e.g., 10 km) (Qin *et al.*, 2020; Y Wu *et al.*, 2021b; Xu *et al.*, 2021), while studies having a high spatial resolution were often constrained to a relatively short period (Wei *et al.*, 2022). To the best of our knowledge, no existing study has predicted ambient NO_2 concentrations in China, simultaneously achieving high spatial resolution, and high spatiotemporal coverage to support large-scale assessments of chronic NO_2 exposure related adverse health effects. Second, the cloud cover and bright surfaces usually lead to the non-random missing of satellite NO_2 column densities (Li and Wu, 2021). Without considering the non-random missing values, it may lead to exposure misclassification and bias the health effects of NO_2 exposure in epidemiological studies. However, most existing studies simply excluded or use the nearby observations to interpolate the missing values (He *et al.*, 2022; Xu *et al.*, 2021).

In the current study, we aimed to develop an ensemble machine-learning model integrating random forest, extreme gradient boosting (XGBoost), and Gradient boosting machine (GBM) algorithms to assess the monthly NO_2 levels at 0.05° × 0.05° spatial resolution from 2005 to 2020, and to evaluate mortality burden attributable to NO_2 exposure at the provincial level in China. We first developed a gap filling approach to impute the missing OMI NO_2 column densities using meteorology, cloud cover, and Copernicus Atmosphere Monitoring Service (CAMS) nitrogen oxides assimilation results. Based on the gap-filled OMI data, we then trained three separate machine learning models and combined them by a generalized additive model (GAM). We finally estimated the mortality burden related to NO_2 in each province of China based on the satellite-derived high resolution NO_2 dataset and the epidemiologically derived exposure response relations.

2 Material and methods

2.1 Study area

Our study domain covered mainland China, Hong Kong, Macao, and Taiwan (Fig. 1). To ensure prediction accuracy at the border area, a 50-km buffer region was created around the national boundary. We constructed a 0.05° resolution modeling grid over this study domain for data integration and model training, totaling 399,513 grid cells.

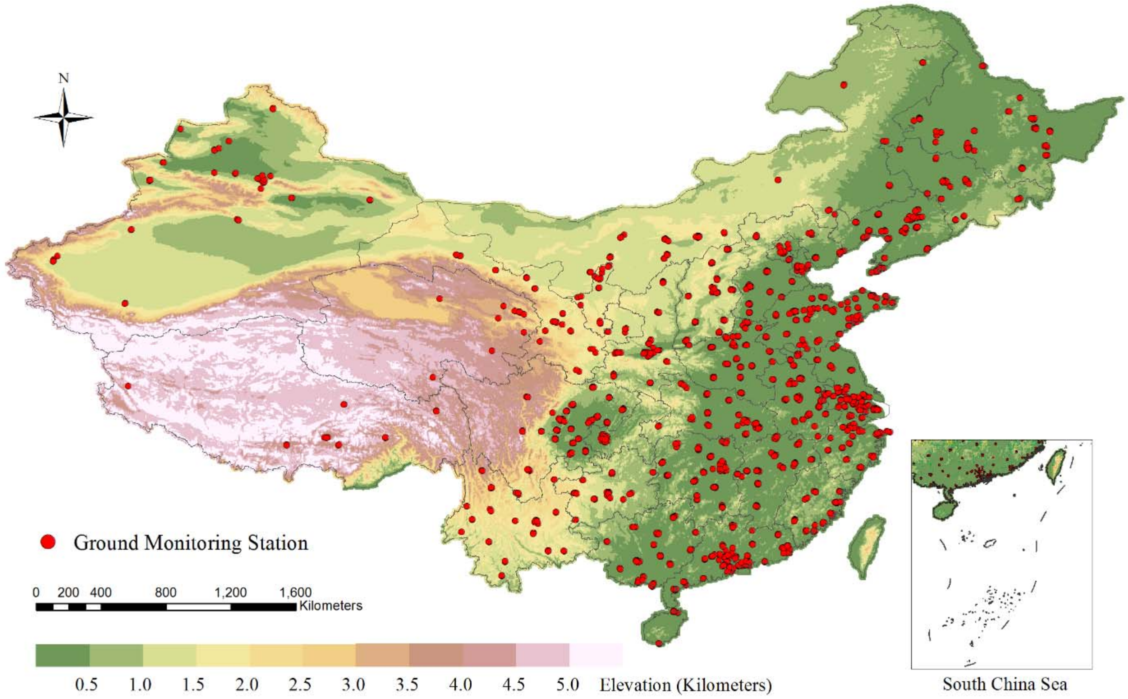

Fig. 1. Study domain plus 50-km buffer and distributions of air quality monitoring stations.

2.2 Ground NO_2 measurements

Ground-level hourly NO_2 measurements were collected from ~1500 air quality monitoring stations administered by the China National Environmental Monitoring Center (http://www.cnemc.cn/) (Fig. 1). We removed repeated identical NO_2 values for at least eight continuous hours because these measurements were likely caused by instrument malfunction. Days with less than 18 (75%) hourly measurements were excluded. In addition, those months with less than 20 days of valid NO_2 measurements were also removed. We multiplied the NO_2 values before September, 2018 by 0.92, since the monitoring condition was amended from standard atmospheric state (273 K, 101.325 kPa) to reference state (298.15 K, 101.325 kPa) for gaseous pollutants (Y Wu *et al.*, 2021b). Finally, daily mean NO_2 concentrations from each station were aggregated to the monthly level. We used the data during 2014-2019 for model training and data during 2020 for external validation.

2.3 Satellite retrieved NO_2 data

We obtained the satellite retrieved tropospheric vertical column density (VCD) of NO_2 during 2005-2020 from OMI NO_2 level-3 data product (OMNO2d, 0.25°×0.25° resolution) (Krotkov *et al.*, 2017). The OMI instrument onboard the Aura satellite is mainly used to observe the ozone profile, air quality and climate change, with a nearly global coverage on a daily basis since July, 2004. To ensure the data quality, we removed those pixels with a cloud fraction >30%. In the current study, we firstly imputed the missing NO_2 VCD, then interpolated it to the 0.05° resolution grid by inverse distance weighting (IDW).

2.4 Meteorological and land use data

Meteorological parameters in 2005-2020 were obtained from the fifth generation European Center for Medium-Range Weather Forecasts (ECMWF) atmospheric re-analysis (ERA5) (Hersbach *et al.*, 2020). It has been shown that meteorological conditions play crucial roles in formatting ambient NO_2 levels. We included air temperature, relative humidity, wind speed, planetary boundary layer height, total precipitation, surface solar radiation and thermal radiation, etc. as predictors in this study. It has a spatial resolution of 0.25°×0.25°, and were downscaled to 0.05° grid by IDW method.

We downloaded annual land cover maps at 300 m resolution from the European Space Agency Climate Change Initiative (CCI) for 2005-2015 (https://vest.agrisemantics.org/content/land-cover-cci-product-user-guide) and the Copernicus Climate Change Service Climate Data Store (CDS) for 2016-2020 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover?tab=overview). From these two products, we extracted the area of urban cover, forest cover, shrub, grass, wetland, cropland, water body and bare land in each grid cell. In addition, we calculated the highway versus non-highway lengths in each grid based on the Global Roads Open Access Data Set (https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1).

2.5 Other predictors

Additional predictors were used in this study to improve the NO_2 model prediction accuracy, including the lightning flash density, simulations from the Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2), monthly total emissions of nitrogen oxides, elevation and population density.

Lightning flash is an important contributor to the formation of NOx, and we obtained monthly lightning flash density data from the Harvard-NASA Emissions Component at 0.5°×0.625° resolution (Murray *et al.*, 2012). MERRA-2 is the latest NASA atmospheric reanalysis at 0.5°×0.625° resolution, and monthly sulfate ion, organic carbon, black carbon, dust and seasalt simulations were extracted (Gelaro *et al.*, 2017). Monthly inventories of total emissions of nitrogen oxides at 0.1°×0.1° resolution were obtained from CAMS (https://ads.atmosphere.copernicus.eu/cdsapp#!/home). Annual population density data from 2005 to 2020 were obtained from the Oak Ridge National Laboratory at 1-km resolution (https://landscan.ornl.gov/). We extracted the elevation data at 30-m resolution from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM), version 3.

2.6 Satellite VCD of NO_2 gap-filling model

In this study, we employed a random forest model, including CAMS total column nitrogen dioxide, total column nitrogen monoxide, cloud fraction, air temperature, dew point temperature, elevation and spatiotemporal trends, to impute the missing OMI VCD values. To consider the temporal trends of OMI VCD, a rolling 3-day sampling window was used to impute the missing OMI VCD values on the middle day. There are two main hyperparameters in the random forest model, i.e., the number of predictors selected for each split (mtry) and the number of trees grown in the forest (ntree). After comparing the performance of different settings, we set mtry and ntree as 5 and 500, respectively.

2.7 Ground level NO_2 prediction model

We used an ensemble model to estimate monthly mean ambient NO_2 concentrations in

China. We first trained three independent machine learning algorithms, including random forest, XGBoost, and GBM. According to Wu et al., we used the scaling factors (i.e., ratios of the surface NO_2 concentrations to the OMI VCD) as the dependent variable rather than the surface NO_2 concentrations (Y Wu *et al.*, 2021b). Then, we multiplied the scaling factors and OMI VCD to obtain the ground level NO_2 values. The scaling factors measure the vertical proportion of NO_2, and have been reported to improve the prediction accuracy of historical NO_2 levels (Y Wu *et al.*, 2021b). Finally, we combined the NO_2 predictions from three individual machine learning models by GAM into an ensemble model. Comprehensively considering the training efficiency and model performance, we set the key parameters as the following: the number of trees was 500 for random forest, XGBoost, and GBM. The number of variables per split was 15 for random forest. The maximum tree depth was 10 for XGBoost and GBM.

We conducted overall 10-fold cross validation (CV) to evaluate the model performance. Correspondingly, the entire dataset was randomly split into 10 groups, with each group containing 10% of the data. In each round of CV, nine groups of data were selected to fit the model, which was then used to predict on the remaining group. This process was repeated 10 times until every group was predicted. In addition, we also performed the 10-fold spatial and temporal CV to evaluate the model prediction accuracy at unmonitored site and time. For spatial CV, we randomly selected 90% of the locations to fit the model and made predictions on the remaining locations. Similarly, for temporal CV, we selected 90% of the months to fit the model and made predictions on the remaining months. Additionally, we conducted by-year CV to evaluate the model's hindcast performance, in which one year was used for testing and remaining years for training. Furthermore, a separate time period, January 2020 to December 2020, was used to characterize the hindcast prediction error. Statistical indicators, such as coefficient of determination (R^2) and root mean squared prediction error (RMSE), were calculated to evaluate the model performance.

2.8 Mortality burden attributable to NO_2 exposure

We calculated the mortality burden caused by NO_2 exposure at provincial level from 2005 to 2020 in China. The annual provincial population and mortality data from 2005 to 2019 were downloaded from the China Statistical Year Book (http://www.stats.gov.cn/). Since the data were still not available for 2020, we used the population and mortality data in 2019 instead. The attributable deaths were calculated using the following equation.

$$RR_C = RR^{\frac{(NO_2-Ref)}{10}}$$

$$\mathrm{AD}_{ij} = (RR_C - 1)/RR_C \times Pop_{ij} \times I_{ij}$$

where RR_C is the relative risk for all cause mortality related to NO_2 exposure above the reference value. Recommended by the recent World Health Organization (WHO) air quality guidelines, we used the RR value of 1.02 (95% CI: 1.01-1.04) for all cause mortality related to per 10 μg/m^3 increase in NO_2. Since no obvious threshold NO_2 value was reported in previous epidemiological studies, we used the counterfactual zero level as the reference (Meng *et al.*, 2021). AD_{ij} is the attributed deaths in province i at year j. Pop_{ij} and I_{ij} is the total population and baseline mortality rate in province i at year j. In addition, to eliminate the motality burden attributable to population growth, the annual NO_2 related deaths were recalculated using the population and mortality data in 2005 as the reference.

3 Results

3.1 Descriptive Statistics

The average ground NO_2 concentrations during 2014-2019 was 30.4 μg/m^3, with standard deviation of 14.6 μg/m^3. The annual mean NO_2 levels decreased by 7 μg/m^3 from 2014 to 2015, and then keep relatively constant from 2015 to 2019 (Supplementary Table S1). The NO_2 levels in China were much higher than the annual NO_2 limit (10 μg/m^3) of WHO air quality guidelines.

3.2 OMI VCD gap-filling by random forest

The mean coverage of OMI VCD of NO_2 in China from 2005 to 2019 was 46.9%. The north of China (~65%) has a higher coverage than the south (<35%), especially in the southwest (Supplementary Table S2 and Fig S1). After imputation, the coverage of OMI VCD increased to 100%. The daily gap filling model achieved an average out-of-bag R^2 of 0.91, with interquartile ranges from 0.88 to 0.94. The spatial distribution of OMI VCD after imputation was almost consistent with that before imputation (Supplementary Fig S2), with much higher values observed in north and east of China. The average OMI VCD increased after gap filling (Supplementary Table S2), possibly because areas where NO_2 is more often missing are more polluted (Supplementary Fig S1 and Fig S2).

3.3 Ground NO_2 prediction model performance

The validation results of the three separate machine learning models and ensemble model were shown in Table 1. Among the individual machine learning models, the XGboost had the highest CV R^2 and the lowest CV RMSE, followed by GBM. The ensemble model outperformed three individual machine learners (R^2, random forest: 0.85, XGboost: 0.87, GBM: 0.87). The overall CV R^2 and RMSE of the ensemble model for monthly NO_2 were 0.88 and 5.14 μg/m^3, respectively, implying a relatively good agreement between model predictions and ground measurements. The temporal CV had a slightly lower R^2 of 0.82 and a higher RMSE of 6.14 μg/m^3. The spatial CV had a less satisfying model performance with a R^2 of 0.73 and a RMSE of 7.57 μg/m^3.

To further evaluate the model's hindcast capability, we conducted by-year CV and external validation using data from January to December, 2020 (Fig. 2). In the by-year CV, the ensemble model predictions matched well with ground observations, with a R^2 of 0.80 and a RMSE of 6.48 μg/m^3. In the external validation, the results shown that our ensemble model fitted by data of 2014-2019 can predict NO_2 levels in 2020 with high accuracy (R^2=0.80, RMSE=5.60 μg/m^3).

Table 1. Model performance for individual machine learning model and ensemble model.

R^2 (RMSE, μg/m^3)	Individual model			Ensemble model
	Random forest	XGboost	GBM	
Overall CV	0.85 (5.74)	0.87(5.20)	0.87 (5.36)	0.88 (5.14)
Temporal CV	0.78 (6.85)	0.82 (6.27)	0.81 (6.32)	0.82 (6.14)
Spatial CV	0.75 (7.36)	0.72 (7.70)	0.72 (7.69)	0.73 (7.57)

RMSE, root mean squared prediction error; XGBoost, extreme gradient boosting; GBM, Gradient boosting machine; CV, cross validation.

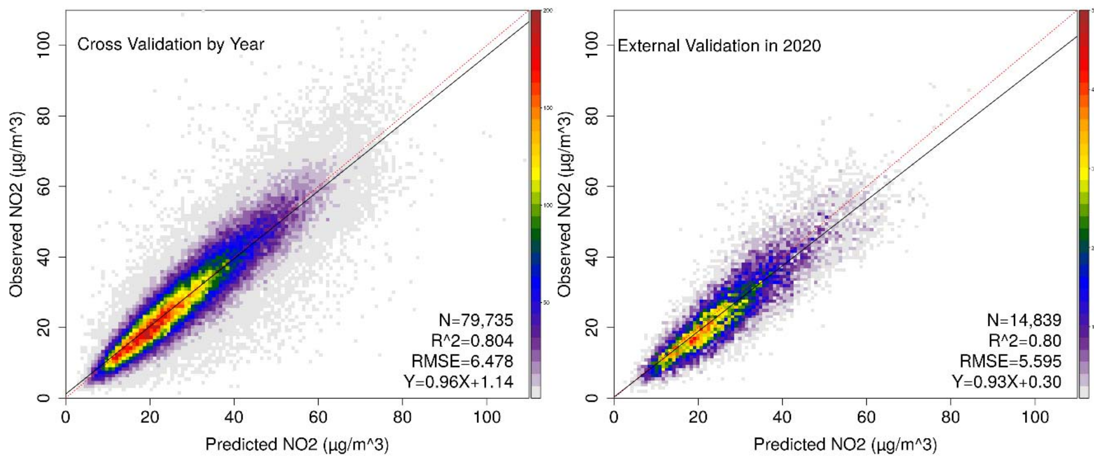

Fig. 2. By-year cross validation and external validation using data from year 2020.

3.4 Spatial and temporal distribution of NO_2

Based on the ensemble model, we obtained monthly ground NO_2 levels in China from 2005 to 2020. We estimated that almost entire population (99.4%) of China lived in areas where NO_2 levels exceeded the 2021 WHO air quality guideline, set at 10 μg/m^3, with 18.3% (255.6 million people) exceeding the WHO interim targets 1 (40 μg/m^3). For the main economic and population concentrated areas, the proportions exceeding 40 μg/m^3 in Beijing-Tianjin-Hebei (BTH) area, Yangtze River Delta (YRD), Pearl River Delta (PRD) and Fenwei Plain (FWP) were 71.6%, 21.7%, 45.8% and 29.9%, respectively.

Fig. 3 displayed the time series of annual population weighted NO_2 for China and four subregions. The national population weighted NO_2 levels began to increase from 2005 to 2007, and experienced the first decrease in 2008. Then it continued to increase until its highest level in 2011, reaching 32.5 μg/m^3. After 2011, it decreased gradually until 2020, especially in 2012-2015. When stratified by region, NO_2 levels in BTH, YRD, PRD and FWP were all higher than the national average, with the highest observed in BTH. Similar to the national trend, the NO_2 in BTH, YRD and FWP reached the peak around 2011-2012, then decrease until 2020. However, the NO_2 in PRD generally shown a continuous downward trend in our study period. During the lock down period due to Covid-19, we observed a significant decrease of NO_2

concentrations. Compared with the data of the same period in 2016-2019, the NO_2 levels in China and Wuhan city decreased by 16.1% and 28.8% from January to April 2020 in lock down (Supplementary Fig S3).

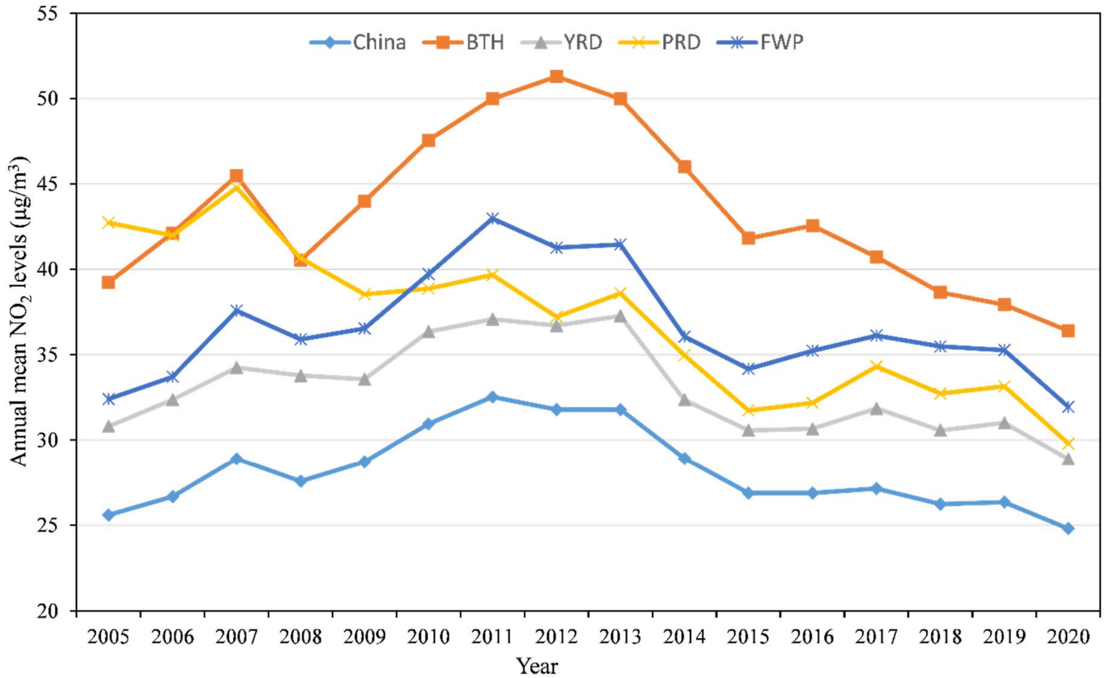

Fig. 3. National and sub-regional annual population weighted NO_2 concentrations from 2005 to 2020. BTH, Beijing-Tianjin-Hebei area; YRD, Yangtze River Delta; PRD, Pearl River Delta; FWP, Fenwei Plain.

The spatial distribution of NO_2 levels in China by season were shown in Fig. 4. The NO_2 levels peaked in winter and were the lowest in summer. The population weighted NO_2 concentrations in China were predicted to be 27.4, 19.6, 29.6 and 36.3 μg/m^3 in spring, summer, autumn and winter, respectively. Spatial trends of NO_2 levels over China were comparable in four seasons, with relatively higher pollution in Beijing, Tianjin, southern Hebei, and northern Henan. Other NO_2 hot spots included the Yangtze River delta, and the Pearl River delta.

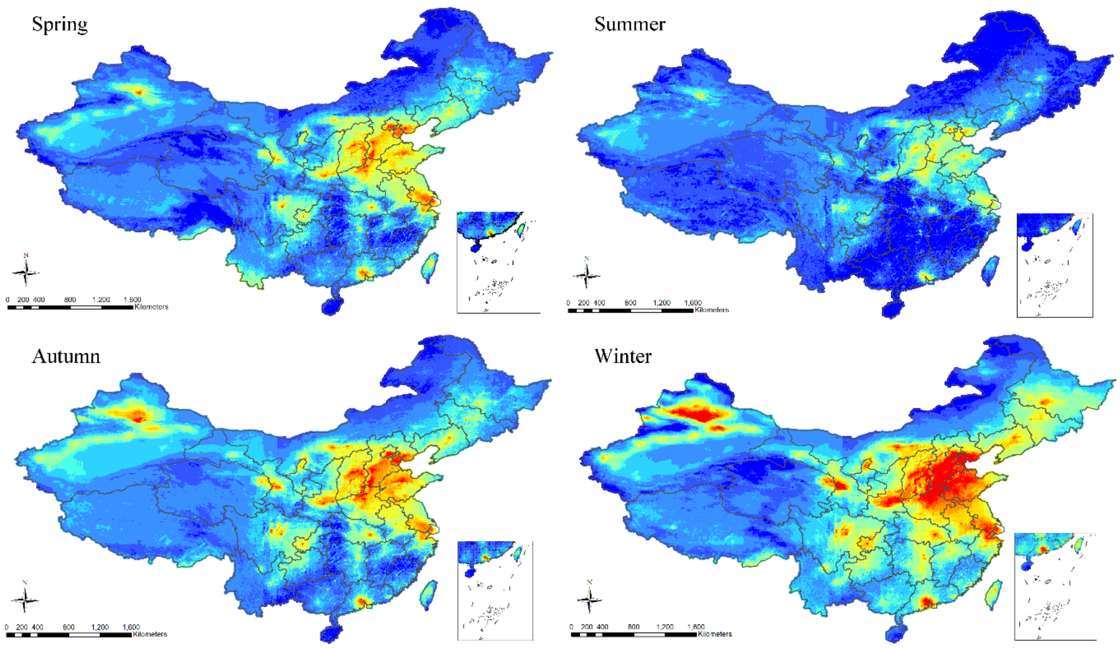

Fig. 4. Spatial distributions of seasonal ambient NO_2 concentrations in China from 2005 to 2020.

3.5 Mortality burden attributable to ambient NO_2

As shown in Fig. 5, the annual mortality burden attributable to NO_2 exposure in China ranged from 305 thousand (2005) to 416 thousand (2012). Overall, it shows a trend of rising first and then declining before 2015, then keeping relatively stable during 2016-2019, followed by a reduction in 2020. If we further subtract the disease burden attributable to population growth by using the population data in 2005, we can still observe a similar trend across the years (Supplementary Fig S4). The estimated number of deaths per 100,000 persons related to NO_2 at the provincial level was shown in Supplementary Fig S5. The per-capita deaths were higher in eastern China, especially in Tianjin, Shandong, Jiangsu, Hebei and Shanghai (39.1 to 48.1 per 100,000 persons), and lower in Hainan, Tibet, and Xinjiang (10.8 to 14.5 per 100,000 persons). We also calculated the provincial absolute number of deaths caused by ambient NO_2 pollution from 2005 to 2020, and found that the provinces with higher 16-year total NO_2 related mortality burden included Shandong, Henan, and Jiangsu province (538 thousand to 668 thousand), and lower in Macao, Tibet and Qinghai (2 thousand to 14 thousand).

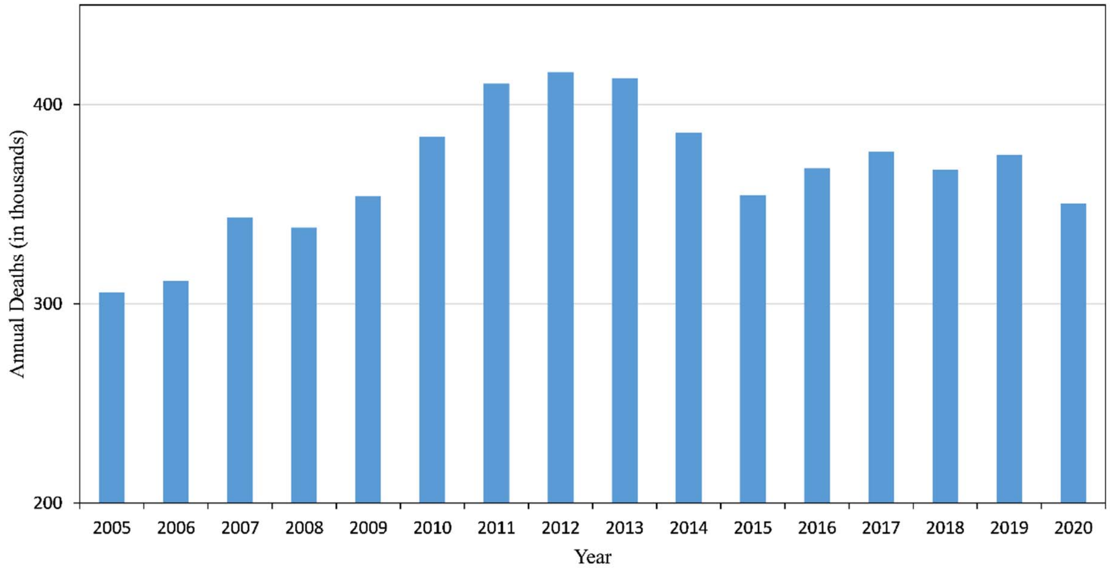

Fig. 5. Annual mortality burden attributable to long-term NO_2 exposure in China from 2005 to 2020.

4 Discussion

In the current study, we filled the research gap by estimating the 16-year spatiotemporal trends in NO_2 and associated mortality burden across provinces in China. We firstly produced a high spatial resolution (0.05° × 0.05°) and long-term (2005-2020) datasets of surface NO_2 concentrations, which enable us to resolve NO_2 variations at small scale. In addition, we found almost entire population of China lived in regions exceeding 2021 WHO guideline annual average NO_2 levels (10 μg/m^3), resulting in 305 thousand to 416 thousand deaths annually from 2005 to 2020.

Accurately estimating the NO_2 concentrations is a crucial step toward epidemiological studies and disease burden estimation. In this study, we developed a high performance ensemble model to predict surface NO_2 levels, and obtained a relatively high prediction accuracy (overall, temporal, and spatial CV R^2: 0.877, 0.824, and 0.732), which was better than most existing models (Chi *et al.*, 2021; Z Huang *et al.*, 2022; Qin *et al.*, 2017; Zhan *et al.*, 2018). Our model has several advantages over previous studies. First, in contrast to most previous studies which trained a single algorithm (Z Huang *et al.*, 2022; Zhan *et al.*, 2018), we ensembled multiple machine learners by GAM algorithm and obtained better model performance compared to the individual learner. Second, our model exhibited a good capability of predicting historical NO_2 estimates at high spatial resolution. Due to lack of routine NO_2 monitoring data before 2013 in China, it's essential to develop NO_2 prediction models which can provide accurate historical exposure estimates. However, most existing studies focused on a limited time period and ignored the model's capability of predicting historical NO_2 concentrations. Only a few studies have evaluated their model's hindcast performance, showing unsatisfying accuracy. For example, Huang et al. estimated daily NO_2 exposure during 2013-2019 in China using an ensemble model, and obtained a CV R^2 of 0.72 and a RMSE of 10.61 μg/m^3. However, the model's accuracy decreased when predicting historical NO_2 concentrations (by-year CV R^2: 0.68, RMSE:11.43 μg/m^3) (C Huang *et al.*, 2021a). In addition, atmospheric NO_2 mainly comes from traffic and industrial emissions and has a short lifetime, likely forming local pollution hotspots. Thus, previous models with coarse spatial resolution may not be able to resolve the NO_2 variations at small scale, leading to bias in exposure assessment. In the current study, our ensemble model displayed a high accuracy of predicting historical NO_2 levels at 0.05° × 0.05° resolution, obtaining by-year CV R^2 of 0.80 (RMSE: 6.5 μg/m^3) and a separate time (year 2020) validation R^2 of 0.80 (RMSE: 5.6 μg/m^3). Third, we proposed a gap filling method to impute the missing OMI VCD values and achieved a 100% spatiotemporal coverage in NO_2 estimation. Most existing studies simply excluded or use the nearby observations to interpolate the missing values (He *et al.*, 2022; S Wu *et al.*, 2021a; Xu *et al.*, 2021). This may lead to significant exposure bias in epimiological studies when there is a high missing rate. Some studies have tried the linear regression, mixed model, or temporal convolution methods to impute the missing OMI VCD values, but obtained limited prediction accuracy and the generalizability is uncertain (de Hoogh *et al.*, 2019; Li and Wu, 2021; Y Wu *et al.*, 2021b). For instance, de Hoogh et al used a liner mixed effect model to impute the missing OMI data and obtained CV R^2 of 0.68 (de Hoogh *et al.*, 2019). In our study, we built a random forest gap filling model to impute OMI NO_2 data incorporating several publicly available covariates, and obtained a high prediction accuracy with R^2 of 0.91, which will greatly reduce the bias in NO_2 exposure and related health impact assessment.

We observed a first increase then decrease trend of NO_2 concentrations from 2005 to 2020 in China, which was consistent with trend of NOx emissions in China (Jiang *et al.*, 2020). The rapid growth of industrial and vehicle NOx emissions lead to the increasing trend of NO_2 levels during 2005-2011. After 2011, the NO_2 levels showed a decreasing trend, especially in 2012-2015. The declined trend during this period was probably due to environmental protection policies in China. In 2011-2015, the Chinese government initiated the 12th Five-Year-Plan (FYP) and set a stringent target to reduce the NOx emissions (Jiang *et al.*, 2020). Through increasing the use of clean energy and installing denitrification facilities, China has successfully reduced the NOx emissions by 18.6% in 2011-2015. After 2015, the downward trend of NO_2 concentrations in China tend to be flat, which may be contributed by the vehicle NOx emissions from sharp increase of private cars (Jiang *et al.*, 2020). Similar with recent findings (Cooper *et al.*, 2022), our ensemble model also observed a significant reduction of NO_2 levels during Covid-19 lock down period in Wuhan. It demonstrated the impact of emission reduction policy on NO_2 pollution.

A series of studies have reported the health burden attributable to air pollution in China, but most of them focused on $PM_{2.5}$ and ozone (Cohen *et al.*, 2017; Liang *et al.*, 2020; Yin *et al.*, 2020). As one of the major NOx emission countries world, China has experienced serious air pollution in recent years. However, very few studies have assessed the NO_2 related disease burden in China. Based on sparse monitoring data in single year of 2016, Zhao et al. estimated 388.5 thousand deaths caused by NO_2 exposure above 5 μg/m^3 in China (Zhao *et al.*, 2020). However, the fixed monitors tend to have higher NO_2 levels, since they are mainly located in eastern urban areas. In addition, the Chinese government has implemented several policies to mitigate the air pollution in the last decades (Lu *et al.*, 2020). No studies have assessed the long-term spatiotemporal distribution of NO_2 attributable disease burden in this process, despite the great implications for future policy making. In the current study, we found NO_2 pollution is also an important risk factor for mortality burden in China, resulting in 305 thousand to 416 thousand annual deaths from 2005 to 2020. In addition, the mortality burden caused by NO_2 has dramatically declined since the 12th FYP implementation (2011-2015), demonstrating the controlling efficacy. However, this trend was not evident since 2016, possibly due to fasting growing vehicles (Jiang *et al.*, 2020). Considering the non-negligible mortality burden by NO_2, targeted policies focusing on vehicle emission control should be strengthened in China.

Our study has some limitations. First, the spatial resolution of our model can be further increased by using Tropospheric Monitoring Instrument (TROPOMI) satellite retrievals at 3.5 × 7.0 km resolution (Veefkind *et al.*, 2012). However, TROPOMI was lauched in October 2017, thus could not meet the needs of long-term health effects epidemiological studies and health impact assessment. In addition, although the exposure response relations between NO_2 and mortality we used was from a recent large-scale meta-analysis adopted by the 2021 WHO air quality guideline (Huangfu and Atkinson, 2020), it mainly relied on studies from the USA and Europe, which may bias the disease burden estimates. More studies on the chronc health effects of NO_2 from China and other low and middle income countires are still needed.

5 Conclusions

In the current study, we presented an ensemble machine learning model to estimate long-term NO_2 concentrations at high spatial resolution in China. Based on this model, we produced reliable historical monthly mean NO_2 estimations at 0.05° resolution, which will greatly enhance epidemiological studies and health impact assessment of chronic NO_2 exposure in China. Furthermore, our results also indicated that exposure to NO_2 contributed to heavy mortality burden in China. Targeted policies reducing the emission of nitrogen oxides and increasing public awareness of the adverse health effects by NO_2 pollution should be strengthened in China.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2018YFE0115300) and the National Natural Science Foundation of China (82003549). This work was partially supported by the National Institute of Environmental Health Sciences (NIEHS) of the National Institutes of Health (Grant 1R01ES032140). The content is solely the responsibility of the authors and does not necessarily representthe official views of NIEHS.

Data Availability Statement

The related data can be accessed from the corresponding author upon reasonable request.

References

Breiman L. Random forests. Machine learning, 2001, 45(1): 5-32.

Ministry of Ecology and Environment of the People's Republic of China. Announcement on Issuing the Amendments to “Ambient Air Quality Standards” (GB 3095-2012). https://www.mee.gov.cn/gkml/sthjbgw/sthjbgg/201808/t20180815_451398.htm (accessed on Aug 3, 2022), edited.

State Council, China’s 13th Five-Year Plan for Ecological & Environmental Protection (2016-2020), State Council. http://www.gov.cn/zhengce/content/2016-12/05/content_5143290.htm (accessed on Aug 3, 2022), edited.

WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization; 2021.

World Health Organization. (2021). WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. https://apps.who.int/iris/handle/10665/345329.

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

Cohen, A. J., et al. (2017), Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, *Lancet*, *389*(10082), 1907-1918, doi:10.1016/S0140-6736(17)30505-6.

Cooper, M. J., R. V. Martin, M. S. Hammer, P. F. Levelt, P. Veefkind, L. N. Lamsal, N. A. Krotkov, J. R. Brook, and C. A. McLinden (2022), Global fine-scale changes in ambient NO2 during COVID-19 lockdowns, *Nature*, *601*(7893), 380-387, doi:10.1038/s41586-021-04229-0.

de Hoogh, K., A. Saucy, A. Shtein, J. Schwartz, E. A. West, A. Strassmann, M. Puhan, M. Roosli, M. Stafoggia, and I. Kloog (2019), Predicting Fine-Scale Daily NO2 for 2005-2016 Incorporating OMI Satellite Data Across Switzerland, *Environ. Sci. Technol.*, *53*(17), 10279-10287, doi:10.1021/acs.est.9b03107.

Di, Q., et al. (2020), Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging, *Environ. Sci. Technol.*, *54*(3), 1372-1384, doi:10.1021/acs.est.9b03358.

Geddes, J. A., R. V. Martin, B. L. Boys, and A. van Donkelaar (2016), Long-Term Trends Worldwide in Ambient NO2 Concentrations Inferred from Satellite Observations, *Environ. Health Perspect.*, *124*(3), 281-289, doi:10.1289/ehp.1409567.

Gelaro, R., et al. (2017), The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *J Clim*, *30*(Iss 13), 5419-5454, doi:10.1175/JCLI-D-16-0758.1.

He, S., H. Dong, Z. Zhang, and Y. Yuan (2022), An Ensemble Model-Based Estimation of Nitrogen Dioxide in a Southeastern Coastal Region of China, *Remote Sensing*, *14*(12), 2807, doi:10.3390/rs14122807.

Hersbach, H., et al. (2020), The ERA5 global reanalysis, *Quart J Royal Meteoro Soc*, *146*(730), 1999-2049, doi:10.1002/qj.3803.

Huang, C., K. Sun, J. Hu, T. Xue, H. Xu, and M. Wang (2021a), Estimating 2013-2019 NO2 exposure with high spatiotemporal resolution in China using an ensemble model, *Environ. Pollut.*, 118285, doi:10.1016/j.envpol.2021.118285.

Huang, K., Q. Xiao, X. Meng, G. Geng, Y. Wang, A. Lyapustin, D. Gu, and Y. Liu (2018), Predicting monthly high-resolution PM2.5 concentrations with random forest model in the North China Plain, *Environ. Pollut.*, *242*(Pt A), 675-683, doi:10.1016/j.envpol.2018.07.016.

Huang, S., H. Li, M. Wang, Y. Qian, K. Steenland, W. M. Caudle, Y. Liu, J. Sarnat, S. Papatheodorou, and L. Shi (2021b), Long-term exposure to nitrogen dioxide and mortality: A systematic review and meta-analysis, *Sci. Total Environ.*, *776*, 145968, doi:10.1016/j.scitotenv.2021.145968.

Huang, Z., X. Xu, M. Ma, and J. Shen (2022), Assessment of NO2 population exposure from 2005 to 2020 in China, *Environ. Sci. Pollut. Res. Int.*, *in press*, doi:10.1007/s11356-022-21420-6.

Huangfu, P., and R. Atkinson (2020), Long-term exposure to NO2 and O3 and all-cause and respiratory mortality: A systematic review and meta-analysis, *Environ. Int.*, *144*, 105998, doi:10.1016/j.envint.2020.105998.

Jiang, L., Y. Chen, H. Zhou, and S. He (2020), NOx emissions in China: Temporal variations, spatial patterns and reduction potentials, *Atmos. Pollut. Res*, *11*(9), 1473-1480, doi:10.1016/j.apr.2020.06.003.

Krotkov, N. A., L. N. Lamsal, E. A. Celarier, W. H. Swartz, S. V. Marchenko, E. J. Bucsela, K. L. Chan, M. Wenig, and M. Zara (2017), The version 3 OMI NO2 standard product, *Atmos. Meas. Tech.*, *10*(9), 3133-3149, doi:10.5194/amt-10-3133-2017.

Li, L., and J. Wu (2021), Spatiotemporal estimation of satellite-borne and ground-level NO2 using full residual deep networks, *Remote Sens. Environ.*, *254*, 112257, doi:10.1016/j.rse.2020.112257.

Liang, F., Q. Xiao, K. Huang, X. Yang, F. Liu, J. Li, X. Lu, Y. Liu, and D. Gu (2020), The 17-y spatiotemporal trend of PM2.5 and its mortality burden in China, *Proc. Natl. Acad. Sci. U. S. A.*, *117*(41), 25601-25608, doi:10.1073/pnas.1919641117.

Lu, X., S. Zhang, J. Xing, Y. Wang, W. Chen, D. Ding, Y. Wu, S. Wang, L. Duan, and J. Hao (2020), Progress of Air Pollution Control in China and Its Challenges and Opportunities in the Ecological Civilization Era, *Engineering*, *6*(12), 1423-1431, doi:10.1016/j.eng.2020.03.014.

Meng, X., et al. (2021), Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities, *BMJ*, *372*, n534, doi:10.1136/bmj.n534.

Murray, L. T., D. J. Jacob, J. A. Logan, R. C. Hudman, and W. J. Koshak (2012), Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, *J. Geophys. Res.*, *117*(D20), D20307, doi:10.1029/2012jd017934.

Qin, K., X. Han, D. Li, J. Xu, D. Loyola, Y. Xue, X. Zhou, D. Li, K. Zhang, and L. Yuan (2020), Satellite-based estimation of surface NO2 concentrations over east-central China: A comparison of POMINO and OMNO2d data, *Atmos. Environ.*, *224*, 117322, doi:10.1016/j.atmosenv.2020.117322.

Qin, K., L. Rao, J. Xu, Y. Bai, J. Zou, N. Hao, S. Li, and C. Yu (2017), Estimating Ground Level NO2 Concentrations over Central-Eastern China Using a Satellite-Based Geographically and Temporally Weighted Regression Model, *Remote Sensing*, *9*(9), 950, doi:10.3390/rs9090950.

Song, W., H. Jia, Z. Li, D. Tang, and C. Wang (2019), Detecting urban land-use configuration effects on NO2 and NO variations using geographically weighted land use regression, *Atmos. Environ*, *197*, 166-176.

Strassmann, A., K. de Hoogh, M. Roosli, S. R. Haile, A. Turk, M. Bopp, M. A. Puhan, and G. Swiss National Cohort Study (2021), NO2 and PM2.5 Exposures and Lung Function in Swiss Adults: Estimated Effects of Short-Term Exposures and Long-Term Exposures with and without Adjustment for Short-Term Deviations, *Environ. Health Perspect.*, *129*(1), 17009, doi:10.1289/EHP7529.

Veefkind, J. P., et al. (2012), TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, *Remote Sens. Environ.*, *120*, 70-83, doi:10.1016/j.rse.2011.09.027.

Wei, J., et al. (2022), Ground-Level NO2 Surveillance from Space Across China for High Resolution Using Interpretable Spatiotemporally Weighted Artificial Intelligence, *Environ. Sci.*

Technol., doi:10.1021/acs.est.2c03834.

Wu, S., B. Huang, J. Wang, L. He, Z. Wang, Z. Yan, X. Lao, F. Zhang, R. Liu, and Z. Du (2021a), Spatiotemporal mapping and assessment of daily ground NO2 concentrations in China using high-resolution TROPOMI retrievals, *Environ. Pollut.*, *273*, 116456, doi:10.1016/j.envpol.2021.116456.

Wu, Y., et al. (2021b), A robust approach to deriving long-term daily surface NO2 levels across China: Correction to substantial estimation bias in back-extrapolation, *Environ. Int.*, *154*, 106576, doi:10.1016/j.envint.2021.106576.

Xu, J., H. Lindqvist, Q. Liu, K. Wang, and L. Wang (2021), Estimating the spatial and temporal variability of the ground-level NO2 concentration in China during 2005–2019 based on satellite remote sensing, *Atmospheric Pollution Research*, *12*(2), 57-67, doi:10.1016/j.apr.2020.10.008.

Yin, P., et al. (2020), The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990-2017: an analysis for the Global Burden of Disease Study 2017, *Lancet Planet Health*, *4*(9), e386-e398, doi:10.1016/S2542-5196(20)30161-3.

Zhan, Y., Y. Luo, X. Deng, K. Zhang, M. Zhang, M. L. Grieneisen, and B. Di (2018), Satellite-Based Estimates of Daily NO2 Exposure in China Using Hybrid Random Forest and Spatiotemporal Kriging Model, *Environ. Sci. Technol.*, *52*(7), 4180-4189, doi:10.1021/acs.est.7b05669.

Zhao, S., S. Liu, Y. Sun, Y. Liu, R. Beazley, and X. Hou (2020), Assessing NO2-related health effects by non-linear and linear methods on a national level, *Sci. Total Environ.*, *744*, 140909, doi:10.1016/j.scitotenv.2020.140909.