Science Bulletin 66 (2021) 612–620

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Article

Mapping anthropogenic emissions in China at 1 km spatial resolution and its application in air quality modeling

Bo Zheng [a], Jing Cheng [b], Guannan Geng [c], Xin Wang [d], Meng Li [b,*], Qinren Shi [c], Ji Qi [e], Yu Lei [f], Qiang Zhang [b], Kebin He [c,*]

[a] Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[b] Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China
[c] State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
[d] China National Environmental Monitoring Center, Beijing 100012, China
[e] Center for Environmental Risk and Damage Assessment, Chinese Academy of Environmental Planning, Beijing 100012, China
[f] Center for Regional Air Quality Simulation and Control, Chinese Academy of Environmental Planning, Beijing 100012, China

ARTICLE INFO

Article history:
Received 22 April 2020
Received in revised form 2 September 2020
Accepted 3 September 2020
Available online 9 December 2020

Keywords:
Anthropogenic emission inventory
High-resolution mapping
Air quality modeling

ABSTRACT

New challenges are emerging in fine-scale air quality modeling in China due to a lack of high-resolution emission maps. Currently, only a few emission sources have accurate geographic locations (point sources), while a large part of sources, including industrial plants, are estimated as provincial totals (area sources) and spatially disaggregated onto grid cells based on proxies; this approach is reasonable to some extent but is highly questionable at fine spatial resolutions. Here, we compile a new comprehensive point source database that includes nearly 100,000 industrial facilities in China. We couple it with the frame of Multi-resolution Emission Inventory for China (MEIC), estimate point source emissions, combine point and area sources, and finally map China's anthropogenic emissions of 2013 at the spatial resolution of 30″×30″ (~1 km). Consequently, the percentages of point source emissions in the total emissions increase from less than 30% in the MEIC up to a maximum of 84% for SO_2 in 2013. The new point source-based emission maps show the uncoupled distribution of emissions and populations in space at fine spatial scales, however, such a pattern cannot be reproduced by any spatial proxy used in the conventional emissions mapping. This new accurate high-resolution emission mapping approach reduces the modeled biases of air pollutant concentrations in the densely populated areas compared to the raw MEIC inventory, thus improving the assessment of population exposure.

1. Introduction

Surface emissions are a fundamental input for chemical transport models (CTMs) to solve the continuity equation [1], and the quality of emissions data is one of the key factors controlling the model performance. The simulation of air pollutants, especially short-lived species (e.g., aerosol [2] and ozone [3,4]) whose concentrations vary considerably across space, particularly relies on the accuracy of emission distributions [5,6]. Recently, the emerging demand for kilometer-scale air pollution maps for human and ecosystem health assessments has stimulated the need for developing accurate high-resolution emission maps towards fine spatial resolutions.

The emissions data used by CTMs are primarily compiled using a bottom-up approach, which estimates emissions by multiplying activity data by emission factors and shapes the emissions totals to the gridded format. Point sources distribute their emissions onto the grid where they are located, while area and line sources estimated at the state- or city-scale usually allocate emissions based on other gridded proxies, such as population, nighttime light, and road network, with the assumption that the spatial intensity of a proxy approximates emissions distribution in space. However, this assumption is highly questionable, and the selection of spatial proxies is rather arbitrary in practice over regions without detailed local information of emission distributions, which causes large uncertainties in kilometer-scale emissions mapping [6,7].

To constrain the uncertainties, it is essential to increase the proportion of point sources and reduce the arbitrary use of the proxy-based emissions distribution method. However, only a small share of anthropogenic emissions is currently estimated as point sources

* Corresponding authors.
E-mail addresses: meng.li.atm@gmail.com (M. Li), hekb@tsinghua.edu.cn (K. He).

https://doi.org/10.1016/j.scib.2020.12.008
2095-9273/

in bottom-up inventories. Most industrial emission sources, except for very few high energy-consuming plants (e.g., power plants and metal smelters) [8,9], are distributed onto grids based on spatial proxies, which introduce large uncertainties into gridded emission maps, illustrated by the satellite imagery that plenty of local emissions hotspots are missing in such bottom-up gridded inventories [10–13].

High-resolution emissions inventories built upon sufficient point sources are scarce in China, the largest emitter of anthropogenic pollutants worldwide. The widely used emissions inventory, the Multi-resolution Emission Inventory for China (MEIC) developed by the same research team of this study, includes only coal-fired power plants as point sources due to the lack of data availability when it was initially built. The MEIC data achieve good performance in regional CTMs running at tens of kilometers but have difficulty fitting fine-scale modeling, such as in cities [6], which has to be complemented by local inventories [14]. We have been working continuously on new methods to improve the MEIC model by including more industrial infrastructures and have successfully applied the methods to several provinces [15] and national carbon accounting [16].

Herein, this study extends the methods explored by our previous studies and develops a new comprehensive approach to characterize emission distribution patterns at kilometer resolutions and to evaluate the resultant performance in CTMs. For the first time, we generate 30″×30″ (~1 km) emissions maps of China's anthropogenic sources for 2013, which are denoted as the Multi-resolution Emission Inventory for China - High Resolution (MEIC-HR) in this study. Several newly available state-of-the-art point source datasets are harmonized to form a comprehensive database that includes nearly 100,000 industrial infrastructures in China. We merge all the data streams of point, area, and line sources under the framework of the MEIC model and generate emissions maps of MEIC-HR for the nested domains of a CTM, which is finally used to evaluate the accuracy of high-resolution emissions data and examine the effect on atmospheric chemistry modeling and aerosol exposure assessment.

2. Data and methods

2.1. Emissions model framework

We combine three industrial datasets (Text S1 online) that provide facility- or factory-level data to generate a synthesized industrial point-source database for China. The core concept of this study is to couple this new point-source database with the MEIC model. The power, industrial combustion, and industrial process sources are all estimated as point sources, and the other source sectors are area sources estimated at the province- or county-level by the MEIC emission model. Total emissions are merged and disaggregated to 30″ × 30″ using the coordinates of point sources and the spatial proxies for nonpoint sources. Table S1 (online) summarizes the data processing chain.

We have two principles for the coupling of the point source database with the MEIC model. First, we use the activity data in MEIC as a total constraint. The summation of activity data of all the point sources tends to be slightly underestimated because small factories are possibly omitted in the factory-level statistics. The MEIC model is built upon provincial statistics, accounting for the balance among production, consumption, import, export, and stock change of energy and industrial products, which are considered more accurate. We scale the activity data of each source covered by the point source database to be consistent with the MEIC national totals. Second, the facility-level database has a higher priority to provide calculation parameters related to emission rates, such as emission factors and pollution control efficiency, which reflect the heterogeneous emission levels differing considerably from one another. The province-level emission rates in the raw MEIC are not used unless such parameters are not available in the point source database.

2.2. Point sources

The emissions from each industrial facility are estimated using

$$E_s = A \times EF_s \times \prod_n \left(1 - \eta_{n,s}\right), \tag{1}$$

where s represents pollutants, n represents pollution control devices, E is the emissions value (g), A is the activity data (kg), EF is the emissions factor (g kg^{-1}), and η is the pollution removal efficiency (%). The pollutants estimated in this study include sulfur dioxide (SO_2), nitrogen oxides (NO_x), carbon monoxide (CO), non-methane volatile organic compounds (NMVOCs), ammonia (NH_3), carbon dioxide (CO_2), and particulate matter (PM), including black carbon (BC) and organic carbon (OC).

The EF_s of SO_2 and PM from coal combustion are estimated with [17]

$$EF_{SO_2} = 2 \times SCC \times (1 - Sr), \tag{2}$$

$$EF_{PM,d} = ACC \times (1 - Ar) \times f_d, \tag{3}$$

where d represents the aerodynamic diameter of PM (2.5 or 10 μm), SCC (ACC) is the sulfur (ash) content of coal (g kg^{-1}), Sr (Ar) is the mass fraction of sulfur (ash) retained in the bottom ash, and f_d is the mass fraction of PM with an aerodynamic diameter smaller than d.

Three industrial point-source datasets (Text S1 online) are combined to provide the parameters described in the above equations. The Environmental Statistics (ES) database and the 1st National Census on Pollution Sources (NCPS) are harmonized to establish a complete database that includes more than 90,000 industrial facilities and contains the information on the geographic coordinates, time each plant came online, industrial category, combustion/manufacturing technology, production capacity, industrial product output, fuel consumption, coal sulfur and ash content, and pollution control devices. The field investigation data collected by the Ministry of Ecology and Environment (MEE) are also used to provide information on end-of-pipe pollution control devices. The data harmonization method is described in Text S1 (online) and in our previous studies [6,15,16].

The point sources are mapped to the MEIC source classification and incorporated into the MEIC model to replace the corresponding emission sources. The activity data A of each facility is derived from the point source database with the summation of national totals scaled to the MEIC data (e.g., the scaling factor is 1.02 for the coal combustion in the industrial sector). The EF_s of SO_2 and PM are estimated using the unit-level SCC and ACC and the technology-specific parameters Sr, Ar, and f_d from the MEIC. The EF_s of other species and the pollution removal efficiency η are derived from the source- and technology-based database in the MEIC model combined with the information of fuel and technology types of each industrial site.

2.3. Nonpoint sources

The nonpoint sources are calculated by the MEIC model, with the on-road emissions estimated at the county level [18] and the other sources estimated at the province level [19]. The provincial emissions of each area source are estimated using

$$E_s = A \times \sum_m \left(X_m \times EF_{m,s} \times \sum_n \left(C_{m,n} \times \left(1 - \eta_{n,s} \right) \right) \right), \quad (4)$$

where s represents the pollutant, m represents combustion or manufacturing technologies, n represents the pollution control technologies, A is the activity data of provincial totals (kg), EF is the emissions factor (g kg^{-1}), X and C are the mass fraction of A using one specific technology, and η is the pollution removal efficiency (%). The EF_s of SO_2 and PM are estimated using Eqs. (2) and (3). The provincial parameters in Eqs. (2)–(4) are drawn from a wide range of statistical reports and literature. The activity data of fuel burnt and solvents and products used are collected from various statistical yearbooks with necessary adjustments on rural energy consumption based on field surveys [20]. X and C are derived from the technology turnover models in the MEIC, which have been tuned to reflect historical trends and drivers [21–25]. The MEE investigation data are also used to adjust the technology turnover models [19]. The EF_s of different species and the η of different technologies primarily rely on local measurements [26] or are compiled from previous inventories [21,23,27–29] when local data are not available.

The on-road emissions are estimated using the county-level activity data and emissions factors [18]. The activity data and technology distributions are modeled for each county to reflect the influence of socioeconomic development on vehicle ownership and fleet turnover rates. The emissions factors that are sensitive to ambient conditions are modeled by an emission factor model combined with county-level meteorological data. Please refer to our previous paper [18] for details.

2.4. Emissions mapping

The point source database has the geographic coordinates of each factory. To ensure that the coordinates are accurate, we visually check the locations of large point sources through Google Earth and fix the wrong coordinates, which include all of the power plants, iron and steel plants, cement plants, glass plants, and the other industrial facilities that account for more than 90% of the remaining industrial emissions in 2013. The nonpoint sources are spatially downscaled using source-specific spatial proxies (Table S1 online) via two steps: (1) disaggregation from province to county and (2) distribution from county to 30″×30″ grid cells. The spatial proxies used include the industrial GDP, population density, road network, and other proxies associated with the related sources [6]. The emissions of each point source are put into the grid cell that contains its coordinates and combined with the nonpoint sources to generate total emission maps.

2.5. Chemical transport model

We use Community Multiscale Air Quality Version 5.2 (CMAQv5.2) to perform the modeling, with the meteorological fields generated by Weather Research and Forecasting Version 3.9 (WRFv3.9). We use nested model domains with horizontal resolutions of 36, 12, and 4 km, which cover the Chinese mainland, central and eastern China, and the four most densely populated regions (Fig. S1 online), respectively. We conduct 13-month simulations from Dec. 2012 to Dec. 2013, with the first month as the model spin-up. Two sets of simulations are performed separately using the new high-resolution MEIC-HR and raw MEIC emissions maps to represent China's anthropogenic emissions. Emission magnitudes of the MEIC are scaled consistently with our new inventory MEIC-HR by the source sector. The other configurations of the modeling system follow our previous studies [6,14,30]. The modeled annual and seasonal daily average concentrations of SO_2, nitrogen dioxide (NO_2), ozone (O_3), and $PM_{2.5}$ within all of the nested domains are evaluated against ground-based *in situ* measurements over China (http://106.37.208.228:8082/).

3. High-resolution emissions mapping

3.1. Spatial distribution pattern

China's anthropogenic emissions in 2013 are estimated to be 26.1 Tg SO_2, 27.8 Tg NO_x, 170.4 Tg CO, 28.5 Tg NMVOCs, 10.6 Tg NH_3, 10350.5 Tg CO_2, 11.8 Tg $PM_{2.5}$, 15.8 Tg PM_{10}, 28.5 Tg TSP, 1.6 Tg BC, and 2.9 Tg OC. These values are quite close to the MEIC estimates [19], with a slight difference of less than 5% for gaseous species and less than 10% for particulate matter, both of which are within the typical uncertainty range of bottom-up emissions inventories. The emissions estimates are highly consistent because the MEIC model has referred to the MEE investigation to tune technology turnover models and pollution control levels in the industry sector.

We present the spatial distribution of all the point sources with their emissions of SO_2 (Fig. 1), NO_x (Fig. S2 online), $PM_{2.5}$ (Fig. S3 online), and CO_2 (Fig. S4 online) in maps. The point sources are unevenly distributed in space and are primarily located in East and South China. Their emissions span a wide range across orders of magnitude (dot sizes in figures), and a small number of large point sources dominate the total emissions budget. The top 1000 plants contribute more than half of the industrial emissions of all the species. These emissions hotspots mainly come from the top five industrial categories (dot colors in Figs. 1 and S2–S4 online), which account for 83%, 89%, 89%, and 90% of the total industrial emissions of SO_2, NO_x, $PM_{2.5}$, and CO_2, respectively, and contribute 87%–94% of the emissions of other species. These five industrial categories emit 13%–72% of the total emissions. Electricity and heat power plants and the manufacturers of ferrous metals are major emitters of SO_2 and NO_x, and the manufacturers of nonmetal mineral products contribute the most to $PM_{2.5}$.

We combine the point sources with nonpoint sources and generate 30″×30″ emission maps. Fig. 2 shows the gridded maps of SO_2 emissions and Figs. S5–S7 (online) present the emission maps of NO_x, $PM_{2.5}$, and CO_2, respectively. These 30″×30″ emission maps resolve the spatial variation in emission distributions at local scales, identifying plenty of hotspots that occupy small areas and illustrate sharp emissions gradients. Most of these emissions hotspots are large cities in China. We focus on the city that emits the most SO_2 in 2013 and zoom in on the map to simultaneously show its total emissions (Fig. 2c), nonpoint sources (Fig. 2d), and point sources (Fig. 2e).

The selected city is Tangshan, which produces the most iron and steel in China. Its SO_2 emissions in 2013 are estimated to be 407 Gg, with 375 Gg SO_2 emitted from point sources. The nonpoint sources account for only 32 Gg SO_2, and since the nonpoint source emissions mainly come from residential sources that are downscaled based on the population distribution map, the spatial pattern of nonpoint source emissions mainly follows the population distributions (Fig. 2d) and the spatial extent of urban built-up area. In contrast, industrial point sources (e.g., iron and steel plants) are located outside urban areas (Fig. 2e) to reduce population exposure to air pollutants. The total SO_2 emissions within Tangshan city are consequently uncoupled from the spatial distribution of population in space, with less than 10% of emissions allocated at the urban center but more than 90% of emissions located outside, which generates hot emission pixels that are widely distributed in the suburban and rural areas (Fig. 2c).

3.2. Role of point sources

The most significant improvement compared to MEIC is the increased percentage of point sources, especially for the pollutants

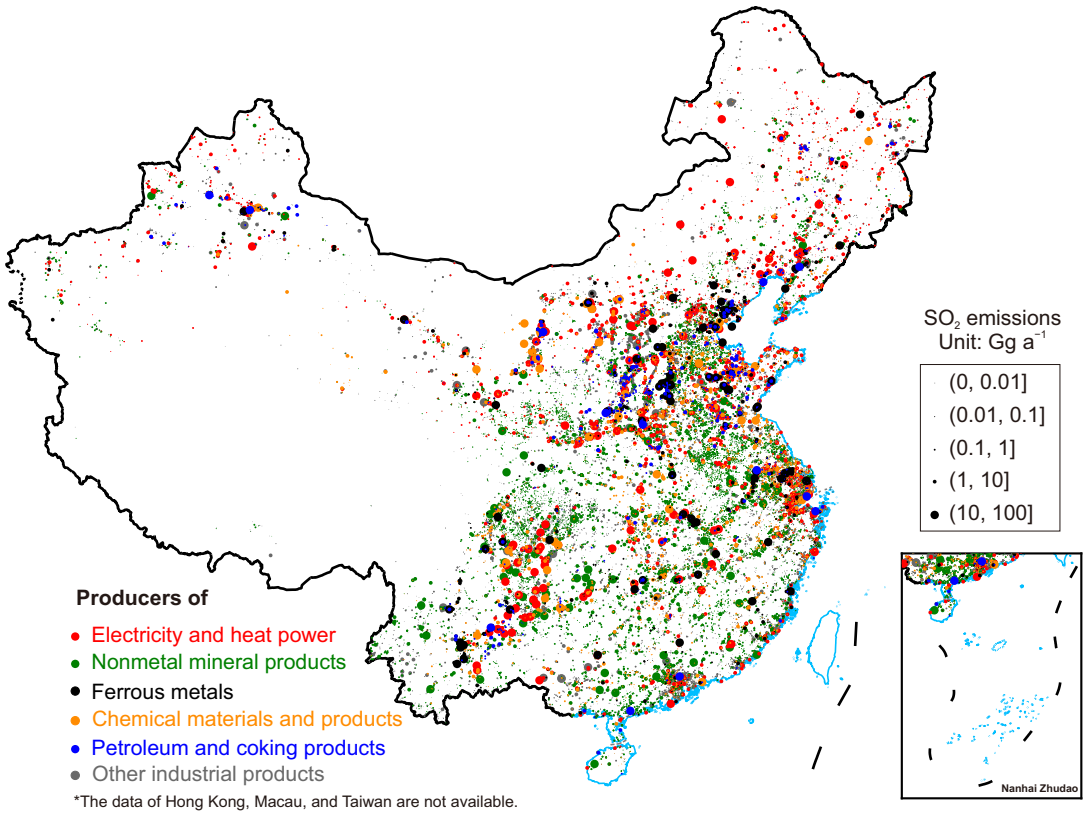


Fig. 1. Point sources of anthropogenic SO_2 emissions. Each dot represents an industrial point source with its SO_2 emissions (dot size) and the industrial category (dot color). The industrial categories shown here are the top five emitters of SO_2 in the industry sector in 2013.

dominated by industrial combustion sources (Fig. 3). The percentages of point source emissions increase from a maximum of 30% of the total emissions in the MEIC to 84% for SO_2, 77% for TSP, 71% for CO_2, 62% for PM_{10}, 58% for NO_x, 55% for $PM_{2.5}$, 41% for CO, 27% for BC, 17% for NMVOCs, and 15% for OC in the MEIC-HR. The species SO_2, CO_2, NO_x, and PM reveal the largest growth in the point sources because they are primarily emitted from industrial combustion sources, which all become point sources in this study. TSP and PM_{10} observe larger increases than $PM_{2.5}$ because the industrial sources emit more coarse particles than fine particles. The other species present moderate improvements in the share of point source emissions because they are also contributed by residential (e.g., CO, BC, and OC), transport (e.g., CO and NMVOCs), or fugitive sources such as solvent use (e.g., NMVOCs), which are currently nonpoint sources.

The increase in point source emissions is evident not only for the whole country (Fig. 3) but also for most of the Chinese cities (Fig. S8 online). Half of China's SO_2 (Fig. S8a online), NO_x (Fig. S8b online), $PM_{2.5}$ (Fig. S8c online), and CO_2 (Fig. S8d online) emissions are distributed in the cities where the percentages of point sources reach at least 87%, 59%, 56%, and 73%, respectively, which are substantially improved from the raw MEIC levels of 20% for SO_2, 26% for NO_x, 6% for $PM_{2.5}$, and 31% for CO_2. Since industrial infrastructures dominate city emissions of these species, we can largely reduce dependence on the spatial proxies to downscale city emissions to fine-scale grid cells, which indeed improves the spatial accuracy of emission distributions within a city.

We examine the influence of adding more point sources on the spatial distribution of emissions by comparing the MEIC and the MEIC-HR. To eliminate the effects of slightly different magnitudes, we scale the MEIC emissions consistent with the MEIC-HR estimates by sector and by species. We calculate the spatial correlation coefficient and the sum of absolute difference of gridded emissions at the horizontal resolutions spanning from 0.05° to 2° (Table S2 online). The correlation coefficient is lower than 0.5 for most of the species at 0.05°, while it increases larger than 0.5 and 0.9 at 0.25° and 2°, respectively, with the summed absolute differences of emissions declining simultaneously. This result suggests that emission distribution patterns are broadly consistent between the MEIC-HR and the MEIC at coarse spatial resolutions, while the emission patterns at fine resolutions diverge after adding more point sources in the MEIC-HR, which is more evident for the species that include higher percentages of point sources (e.g., SO_2 and PM).

To elucidate how those additional point sources affect emissions mapping, we perform a joint analysis of emissions and population densities at different resolutions (Fig. 4). Overall, the MEIC-HR tends to allocate fewer emissions to densely populated grid cells than the MEIC. The most densely populated areas (i.e., urban) that contribute 25% of SO_2 in the MEIC now account for only 13% of SO_2 in the MEIC-HR at the resolution of 0.05° (Fig. 4a). The reduced emissions are reallocated to the less populated grid cells in suburban and rural areas. The turning point for SO_2 at 0.05° is approximately 1700 people km^{-2} (Fig. 4b), below which the grid cells in the MEIC-HR tend to have higher emissions than the MEIC. The redistribution of emissions is also evident for the other pollutant species dominated by point sources such as NO_x (Fig. S9a, d online), $PM_{2.5}$ (Fig. S9b, e online), and CO_2 (Fig. S9c, f online), and the turning points of population densities that reflect the transition of differences in emission spatial distributions between MEIC-HR and MEIC agree with that of SO_2.

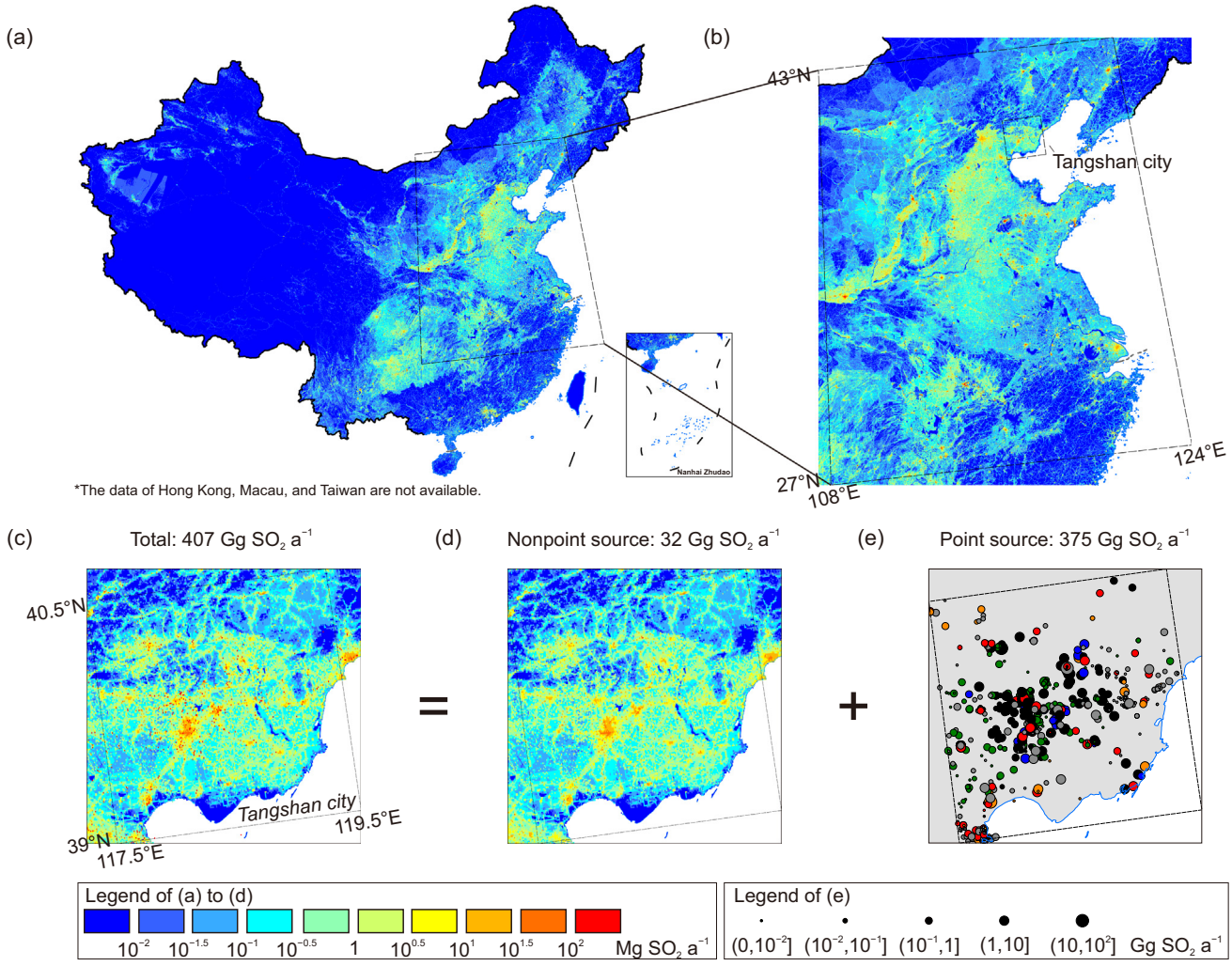


Fig. 2. High-resolution SO_2 emission maps at the horizontal resolution of 30″ × 30″. The emission maps are for the Chinese mainland (a), East China (b), and the city of Tangshan (c–e). Tangshan is the city with the largest SO_2 emissions in 2013. (a)–(c) present total SO_2 emissions, while (d) shows the emissions from nonpoint sources. (e) shows the point sources, with the dot size representing the emission magnitude and the dot color representing the industrial category, as shown in Fig. 1.

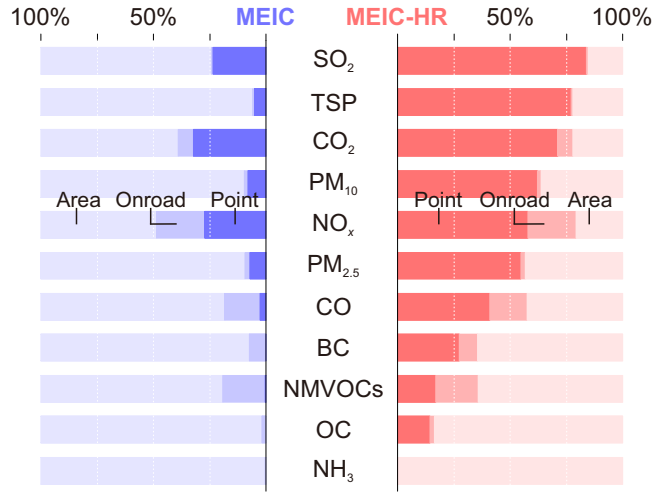


Fig. 3. Percentages of point, on-road, and area sources in China's anthropogenic emissions. The results are compared between the MEIC (left, blue) and MEIC-HR (right, red) inventories for 2013.

The MEIC-HR inventory allocates more industrial emissions to suburban and rural areas because the industrial plants are distant from populations in urban areas (see Tangshan in Fig. 2). With the emissions spatially disaggregated based on population densities, the MEIC industrial emissions are overestimated in urban areas. This systematic bias cannot be reduced by using other spatial proxies because the MEIC-HR emissions show low spatial correlation with the gridded datasets of population, road network, and nighttime light at high spatial resolutions (Figs. 4c and S10 online). None of these commonly used proxies can represent the spatial decoupling of industrial emissions and populations at fine scales. When the size of grid cells increases enough to include both urban and suburban areas, those proxies show a high spatial correlation with the MEIC-HR emissions. Therefore, the differences between the MEIC and the MEIC-HR become smoother at coarse resolutions.

4. Impacts on air quality modeling

4.1. Modeling biases compared to observations

Fig. 5 evaluates the WRF-CMAQ simulations driven by the MEIC and MEIC-HR emissions data. The evaluation focuses on all of the

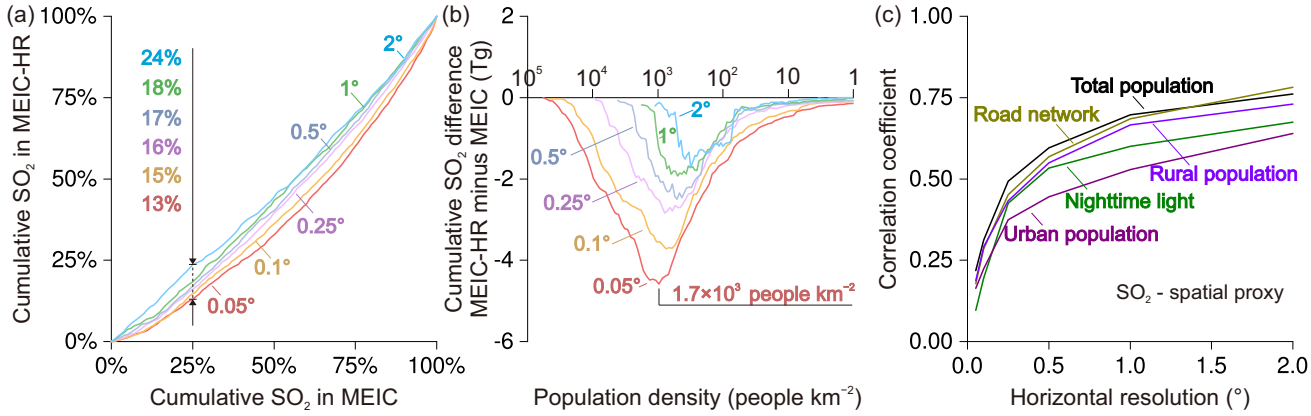


Fig. 4. Spatial pattern analysis among MEIC-HR emissions, MEIC emissions, and different spatial proxies. Cumulative SO_2 emissions are compared between the MEIC inventory (*x*-axis in (a)) and the MEIC-HR inventory (*y*-axis in (a)) at resolutions from 0.05° to 2°. The cumulative emissions are calculated based on the descending order of populations. Each dot in (b) is plotted according to the population density (*x*-axis) and the difference of cumulative SO_2 emissions (MEIC-HR minus MEIC) for all of the grid cells that have higher population densities (*y*-axis). (c) presents the spatial correlation coefficients between the MEIC-HR SO_2 emissions and different spatial proxies from 0.05° to 2°.

cities covered by the four 4 km domains (Fig. S1 online). The annual and seasonal daily average concentrations of air pollutants modeled at the 36, 12, and 4 km domains are separately compared with ground-based observations. Theoretically, high-resolution grid cells should be more consistent with surface measurement sites due to reduced representation errors. However, we find that the increased model resolutions tend to increase simulation biases, especially using the MEIC inventory (blue bars). Compared to ground observations, the normalized mean biases of modeled SO_2, NO_2, O_3, and $PM_{2.5}$ based on MEIC emissions data increase from 37% to 212%, 4% to 46%, −9% to −35%, and −2% to 27%, respectively, with the horizontal resolutions of the model domain increasing from 36 to 4 km. The root mean square error of the modeled pollutant concentrations also increases from the 36 km simulation to the 4 km simulation (Table S3 online).

The modeling with the MEIC-HR inventory matches better with ground-based observations than that with the MEIC inventory. The normalized mean biases (red bars) are 77% for SO_2, 31% for NO_2, −28% for O_3, and 5% for $PM_{2.5}$ at the horizontal resolution of 4 km, much smaller than the results of the MEIC-based modeling. The root mean square errors are also smaller in the MEIC-HR-based simulations (Table S3 online). Since emissions totals are scaled consistently before running the WRF-CMAQ model, the larger biases of the MEIC-based simulations at fine spatial resolutions are mainly caused by errors in the emission mapping. Given that the observation data used here are mostly measured in urban areas, the above evaluation results suggest that the urban emissions in the MEIC model are probably overestimated by a large extent. Such errors are reduced by our new point-source-based emissions maps, which reduce the simulation biases of WRF-CMAQ. The remaining biases of the MEIC-HR-based simulations at the 4 km domains are probably due to uncertainties in physical and chemical processes of the meteorological and air quality models.

Figs. S11–S14 (online) present station-scale model evaluation results, which include spatial correlations between modeled and observed average concentrations at the annual (Fig. S11 online) and seasonal (Fig. S13 online) scales and slope of regression lines

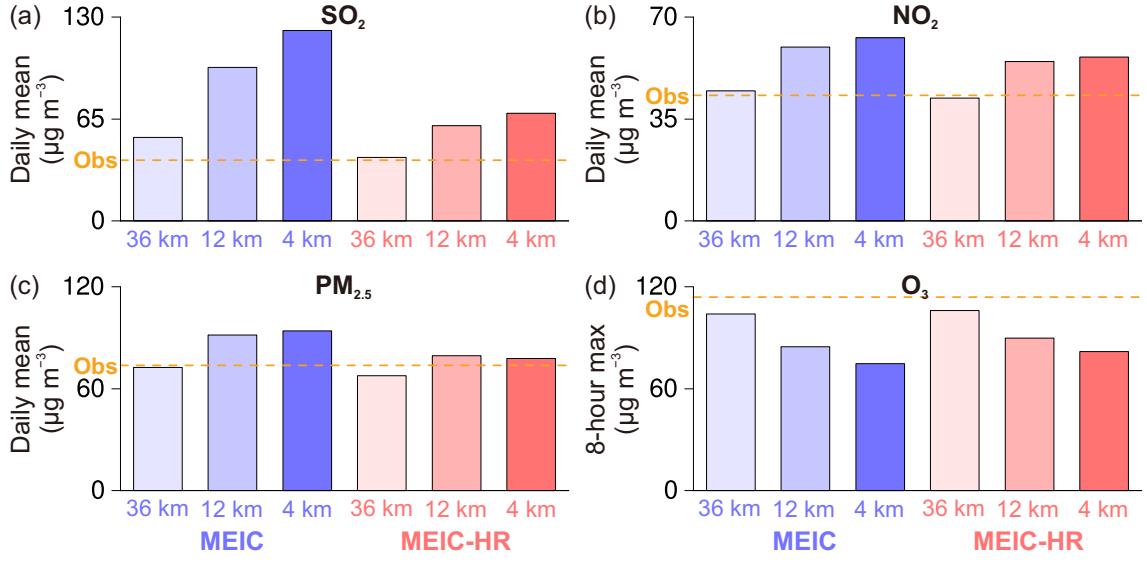


Fig. 5. Comparison of modeled air pollutants with ground-based *in situ* observations. The comparison focuses on the average data of the 58 cities located in the four 4-km modeling domains (Fig. S1 online). The modeled annual daily average concentrations of SO_2 (a), NO_2 (b), and $PM_{2.5}$ (c) are calculated based on MEIC-based (blue bars) and MEIC-HR-based (red bars) simulations, respectively, and are compared with the observation data (orange dashed lines). For O_3 (d), the summertime average of daily 8-hour maximum concentrations is used in the comparison. All of the simulation results used here are sampled at the location and time of observation data.

between modeled and observed values (Figs. S12 and S14 online). Among the 371 observational stations, spatial correlations between modeled and observed annual average concentrations at 4 km domains improve from 0.37 to 0.56 for SO_2, from 0.59 to 0.61 for NO_2, from 0.58 to 0.67 for $PM_{2.5}$, and from 0.26 to 0.32 for O_3 comparing the MEIC-based to MEIC-HR-based simulations. The slopes of fitted lines between modeled and observed values change from 1.87 to 1.45 for SO_2, from 1.34 to 1.19 for NO_2, from 1.16 to 1.10 for $PM_{2.5}$, and from 0.21 to 0.22 for O_3. The improved spatial correlations and the slopes closer to 1 at 4 km domains both suggest that the MEIC-HR emission inventory has improved the simulations of air pollutant spatial distributions at the local scales. The other evaluation results that focus on seasonal means (Figs. S13 and S14 online) and over 36 and 12 km domains also confirm the improved WRF-CMAQ model simulations based on the MEIC-HR emission inventory.

Short-lived pollutants show high concentrations near sources but decay rapidly with distance, which is highly correlated with emission sources distributions in space. The decoupling of emissions and population introduces large errors in proxy-based emission maps, such as MEIC, and thus affects the subsequent modeling of short-lived species. The influence of emission mapping errors on CTM simulations can be reduced by using coarse-resolution model grids because the emissions are averaged over a large area and only a small part of pollutants can be transported out of the coarse grid cell where they are emitted. The MEIC-based modeling results show much smaller normalized mean biases at 36 km (−9% to 37%) than at 12 km (−26% to 153%) and at 4 km (−35% to 212%) (Fig. 5) because the 36 km model grids almost eliminate errors in emission mapping, while the 4 km resolution grids reveal all of these errors.

The point-source-based MEIC-HR inventory successfully represents the decoupling of emissions and population in space and thus constrains the modeling bias at fine spatial scales. The remaining biases in the MEIC-HR inventory are caused by emissions sources that are still spatially disaggregated using proxies, such as residential stoves (SO_2), transport (NO_x), or fugitive solvent use (NMVOCs). Emissions from these nonpoint sources are estimated using provincial-level emission parameters and are allocated to grid cells in space based on population distributions, which tend to overestimate the emissions and the modeled pollutant concentrations in urban areas at 4 km domains (Fig. 5). To further reduce modeling biases at fine spatial scales, the percentage of point sources for all of the species and source sectors needs to be continuously improved.

4.2. $PM_{2.5}$ exposure assessment

The four 4 km domains (Fig. S1 online) include 0.8 billion people, who account for 62% of China's total population. The population-weighted annual average $PM_{2.5}$ of these 0.8 billion people are estimated as 66.0, 74.2, and 70.3 μg m^{-3} in the 36, 12, and 4 km simulations, respectively, with the MEIC inventory. The MEIC-HR-based simulations estimate slightly lower values of 64.3, 70.7, and 65.4 μg m^{-3} at 36, 12, and 4 km, respectively. The largest difference is found at the 4 km resolution, where the MEIC-HR-based estimate is 7% lower than the MEIC estimate.

We analyzed how human exposure to $PM_{2.5}$ varies with population density (Fig. 6). The simulations using MEIC (Fig. 6a) and MEIC-HR (Fig. 6b) both suggest that the population-weighted $PM_{2.5}$ concentrations increase linearly with the population density increasing to 500 people km^{-2}, remain at approximately 70 μg m^{-3} between 500 and 2000 people km^{-2}, and increase again when the population density gets larger. The major discrepancy between the two simulations occurs at grid cells with more than 2000 people km^{-2} (Fig. 6c). The population-weighted $PM_{2.5}$ reaches 132 μg m^{-3} in the most densely populated areas in the MEIC-based 4 km simulations, while the largest concentration estimated by the MEIC-HR-based simulations is only 95 μg m^{-3} in the 4 km domain.

In addition, the highest concentration in the MEIC-HR simulations is not found in the most densely populated areas, unlike the MEIC-based simulations. This result is due to a slight decrease in modeled $PM_{2.5}$ when the population density is larger than approximately 10,000 people km^{-2} (Fig. 6b), which reflects the decoupling of emissions and population at fine spatial scales, especially in areas with high numbers of people. Over the densely populated urban areas, the MEIC-based simulation estimates 46% higher population-weighted $PM_{2.5}$ than the MEIC-HR (equivalent to 42 μg m^{-3}) at the 4 km resolution. These results suggest that we may overestimate the population exposure to $PM_{2.5}$ by nearly 50% if using the proxy-based emission maps to perform kilometer-scale air quality modeling in the densely populated urban areas.

5. Implications for future research

The accuracy of air quality modeling at kilometer scales is quite sensitive to the method of how we map emissions in space. If we do not have enough information on point sources, some spatial proxy must be used to disaggregate emissions in space. Although it is reasonable to some extent, we must evaluate whether this method fails to represent emission distribution patterns at kilometer scales, which could introduce large errors into air quality modeling. Our regional study over China suggests that the proxy-based emission maps worsen the WRF-CMAQ model performance at 4 km compared to that at 36 km because such inventory tends to misallocate the industrial emissions that are located in suburban and rural areas to the densely populated urban areas. With the proxy-based emission maps, it would be safer to run CTMs at coarser resolutions of tens of kilometers to achieve a satisfactory model performance.

Fixing this problem requires a point-source-based emissions inventory with facility-level emission estimates and exact geographic coordinates. Through integrating industrial point sources, we map China's anthropogenic emissions at a horizontal resolution of 1 km, which adequately represents the emissions hotspots distributed widely in suburban and rural areas and improves the emission mapping accuracy over densely populated urban areas. Modeling with this new point-source-based inventory reduces the bias in the modeled urban concentrations of air pollutants at a resolution of 4 km and improves the assessment of population exposures to ambient aerosols. Such a new high-resolution inventory will allow finer-scale simulations than the model community ever did before and will also facilitate the comparison with the satellite-constrained city and point source emissions, which are progressing rapidly towards kilometer-scale resolutions [31,32].

To further improve the accuracy of emission mapping and air quality modeling, the species that still have few point sources (e.g., NH_3 and NMVOCs) will need more attention from emissions inventory developers. These species are mainly generated by sources that are difficult to treat as point sources (e.g., industrial solvent use, residential, and agriculture) due to a lack of data availability. We need more unit-level information on industrial solvent use and storage facilities, residential and agricultural infrastructures, such as household cookstoves and livestock farms. It is valuable to conduct a sample survey through field research to characterize the activity levels and emission distribution patterns [20,33]. Another possibility is to integrate the top-down information into the bottom-up estimated emission maps. Satellite-observed NH_3 columns have been used to quantify NH_3 emissions from large point sources using the oversampling technique and the

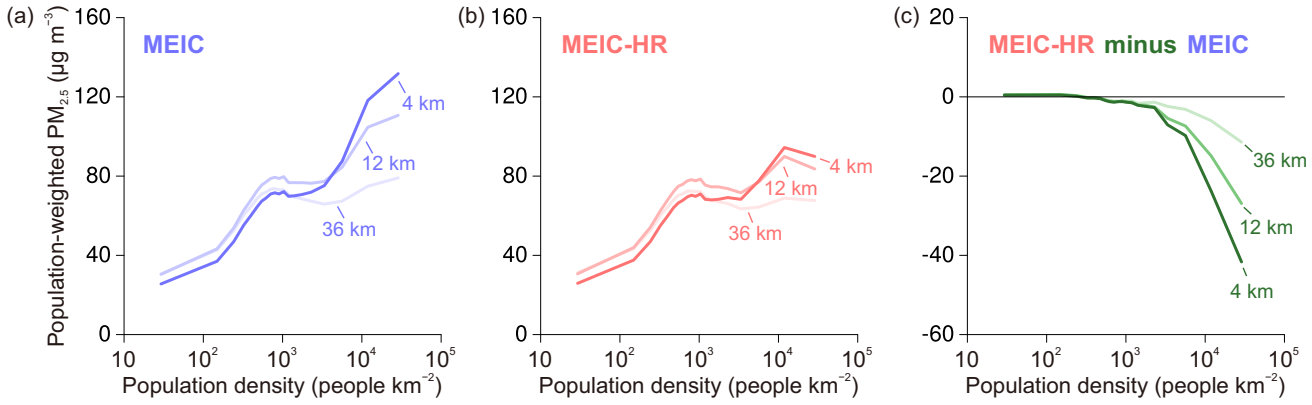


Fig. 6. The variation in population-weighted $PM_{2.5}$ concentrations with population density. The analysis focuses only on the four 4-km modeling domains (Fig. S1 online). We first sort all of the 4-km grid cells from high to low according to the population in each grid, divide these grid cells into 20 groups, and ensure that each group has an equivalent 5% of the total population (approximately 42 million). Finally, we calculate the average population density and the population-weighted $PM_{2.5}$ using each group of data and plot these data according to the population density in the *x*-axis and the population-weighted $PM_{2.5}$ in the *y*-axis. The MEIC-based simulation results (blue curves) are plotted in (a), the MEIC-HR-based simulation results (red curves) are plotted in (b), and the difference between the MEIC-HR and MEIC estimates (green curves) are plotted in (c).

Gaussian fitting model [12,34]. Such a method can identify the missing point sources in bottom-up inventories, and a combination of bottom-up inventories with satellite-derived point sources has the potential to achieve a more comprehensive emission map. These high-resolution satellite observations, as well as ground-based dense air quality sensor networks, can also be combined with fine-scale chemical transport models to evaluate the spatial distribution patterns of surface emissions at a fine spatial scale, which could support the development and refinement of kilometer-scale emission inventories in the future. More detailed statistics and observation constraints should be involved in parallel to further improve the bottom-up emission model framework towards higher spatial resolutions.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (91744310, 41625020 and 41921005) and the National Research Program for Key Issues in Air Pollution Control (DQGG0201). Maps in this article were reviewed by Ministry of Natural Resources of the People's Republic of China (GS(2020) 6011)).

Author contributions

Bo Zheng and Qiang Zhang designed the study. Bo Zheng developed the high-resolution emission inventory for China and Jing Cheng performed the WRF-CMAQ simulations. Meng Li built the database of VOCs emission factors and Qinren Shi processed the industrial point source datasets. Qiang Zhang, Xin Wang, Ji Qi, and Yu Lei provided the data resources necessary for this study. Bo Zheng interpreted the results and wrote the paper with contributions and suggestions from Guannan Geng, Qiang Zhang, and Kebin He.

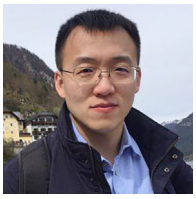
Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2020.12.008.

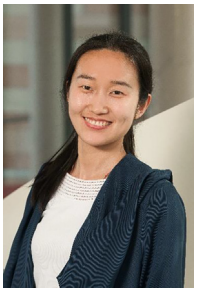
References

[1] Brasseur GP, Jacob DJ. In: Modeling of atmospheric chemistry. Cambridge: Cambridge University Press; 2017. p. 399–435.
[2] Li X, Zhang Q, Zhang Y, et al. Attribution of $PM_{2.5}$ exposure in Beijing–Tianjin–Hebei region to emissions: implication to control strategies. Sci Bull 2017;62:957–64.
[3] Lu X, Zhang L, Zhao Y, et al. Surface and tropospheric ozone trends in the Southern Hemisphere since 1990: possible linkages to poleward expansion of the Hadley circulation. Sci Bull 2019;64:400–9.
[4] Tan Z, Lu K, Dong H, et al. Explicit diagnosis of the local ozone production rate and the ozone-NO_x-VOC sensitivities. Sci Bull 2018;63:1067–76.
[5] Wang R, Tao S, Balkanski Y, et al. Exposure to ambient black carbon derived from a unique inventory and high-resolution model. Proc Natl Acad Sci USA 2014;111:2459–63.
[6] Zheng B, Zhang Q, Tong D, et al. Resolution dependence of uncertainties in gridded emission inventories: a case study in Hebei, China. Atmos Chem Phys 2017;17: 921–933.
[7] Geng G, Zhang Q, Martin RV, et al. Impact of spatial proxies on the representation of bottom-up emission inventories: a satellite-based analysis. Atmos Chem Phys 2017;17:4131–4145.
[8] Crippa M, Guizzardi D, Muntean M, et al. Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. Earth Syst Sci Data 2018;10:1987–2013.
[9] Hoesly RM, Smith SJ, Feng L, et al. Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the community emissions data system (CEDS). Geosci Model Dev 2018;11:369–408.
[10] McLinden CA, Fioletov V, Shephard MW, et al. Space-based detection of missing sulfur dioxide sources of global air pollution. Nat Geosci 2016;9:496–500.
[11] Liu F, Choi S, Li C, et al. A new global anthropogenic SO_2 emission inventory for the last decade: a mosaic of satellite-derived and bottom-up emissions. Atmos Chem Phys 2018;18:16571–86.
[12] Van Damme M, Clarisse L, Whitburn S, et al. Industrial and agricultural ammonia point sources exposed. Nature 2018;564:99–103.
[13] Kong H, Lin J, Zhang R, et al. High-resolution (0.05° × 0.05°) NO_x emissions in the Yangtze River Delta inferred from OM. Atmos Chem Phys 2019;19:12835–56.
[14] Cheng J, Su J, Cui T, et al. Dominant role of emission reduction in $PM_{2.5}$ air quality improvement in Beijing during 2013–2017: a model-based decomposition analysis. Atmos Chem Phys 2019;19:6125–46.
[15] Qi J, Zheng B, Li M, et al. A high-resolution air pollutants emission inventory in 2013 for the Beijing-Tianjin-Hebei region, China. Atmos Environ 2017;170:156–68.
[16] Zheng B, Zhang Q, Davis SJ, et al. Infrastructure shapes differences in the carbon intensities of Chinese cities. Environ Sci Technol 2018;52:6032–41.
[17] Liu F, Zhang Q, Tong D, et al. High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010. Atmos Chem Phys 2015;15:13299–317.
[18] Zheng B, Huo H, Zhang Q, et al. High-resolution mapping of vehicle emissions in China in 2008. Atmos Chem Phys 2014;14:9787–805.
[19] Zheng B, Tong D, Li M, et al. Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. Atmos Chem Phys 2018;18:14095–111.
[20] Peng L, Zhang Q, Yao Z, et al. Underreported coal in statistics: a survey-based solid fuel consumption and emission inventory for the rural residential sector in China. Appl Energy 2019;235:1169–82.

[21] Zhang Q, Streets DG, He K, et al. NO_x emission trends for China, 1995–2004: the view from the ground and the view from space. J Geophys Res Atmos 2007;112:1–18.
[22] Zhang Q, Streets DG, Carmichael GR, et al. Asian emissions in 2006 for the NASA INTEX-B mission. Atmos Chem Phys 2009;9:5131–53.
[23] Lei Y, Zhang Q, He KB, et al. Primary anthropogenic aerosol emission trends for China, 1990–2005. Atmos Chem Phys 2011;11:931–54.
[24] Li M, Zhang Q, Kurokawa JI, et al. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmos Chem Phys 2017;17:935–63.
[25] Li M, Zhang Q, Zheng B, et al. Persistent growth of anthropogenic non-methane volatile organic compound (NMVOC) emissions in China during 1990–2017: drivers, speciation and ozone formation potential. Atmos Chem Phys 2019;19:8897–913.
[26] Li M, Liu H, Geng G, et al. Anthropogenic emission inventories in China: a review. Natl Sci Rev 2017;4:834–66.
[27] Streets DG, Zhang Q, Wang L, et al. Revisiting China's CO emissions after the transport and chemical evolution over the pacific (TRACE-P) mission: synthesis of inventories, atmospheric modeling, and observations. J Geophys Res Atmos 2006;111:D14306.
[28] Lu Z, Streets DG, Zhang Q, et al. Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000. Atmos Chem Phys 2010;10:6311–31.
[29] Li M, Zhang Q, Streets DG, et al. Mapping Asian anthropogenic emissions of non-methane volatile organic compounds to multiple chemical mechanisms. Atmos Chem Phys 2014;14:5617–38.
[30] Zheng B, Zhang Q, Zhang Y, et al. Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China. Atmos Chem Phys 2015;15:2031–49.
[31] Reuter M, Buchwitz M, Schneising O, et al. Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. Atmos Chem Phys 2019;19:9371–83.
[32] Goldberg DL, Lu Z, Streets DG, et al. Enhanced capabilities of TROPOMI NO_2: estimating NO_x from North American cities and power plants. Environ Sci Technol 2019;53:12594–601.
[33] Tao S, Ru MY, Du W, et al. Quantifying the rural residential energy transition in China from 1992 to 2012 through a representative national survey. Nat Energy 2018;3:567–73.
[34] Dammers E, McLinden CA, Griffin D, et al. NH_3 emissions from large point sources derived from CrIS and IASI satellite observations. Atmos Chem Phys 2019;19:12261–93.



Bo Zheng is an assistant professor at Tsinghua Shenzhen International Graduate School, Tsinghua University. He received his bachelor's degree in 2011 and his Ph.D. degree in 2016 from Tsinghua University. He has been then a postdoctoral fellow at the Laboratory for Sciences of Climate and Environment in France since 2016. His research interest focuses on sources and sinks of atmospheric trace gases, trends and drivers of atmospheric compositions, and the policies to mitigate climate change and reduce air pollution.



Meng Li received her Ph.D. degree in Environmental Science and Engineering from Tsinghua University in 2016. She worked at Max-Planck Institute for Chemistry in Germany as a postdoctoral researcher (2017–2019), and then at National Oceanic and Atmospheric Administration in the United States as a research scientist from 2019. Her research interest focuses on emission inventory development, volatile organic compounds (VOCs) speciation, and top-down model evaluations.

Kebin He is a professor at School of Environment, Tsinghua University. He has studied air pollution control in China for more than 30 years. He led the development of the Multi-resolution Emission Inventory for China and established a comprehensive approach of air pollution source apportionment that integrates emission inventory, field observation, and model simulation.