Impact of spatial proxies on the representation of bottom-up emission inventories: A satellite-based analysis

Guannan Geng[1], Qiang Zhang[1], Randall V. Martin[2,3], Jintai Lin[4], Hong Huo[5], Bo Zheng[6], Siwen Wang[7], and Kebin He[6]

[1]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China
[2]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada
[3]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
[4]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China
[5]Institute of Energy, Environment and Economy, Tsinghua University, Beijing, China
[6]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China
[7]Max-Planck-Institute for Chemistry, Mainz, Germany

Correspondence to: Q. Zhang (qiangzhang@tsinghua.edu.cn)

Abstract. Spatial proxies used in bottom-up emission inventories to derive the spatial distributions of emissions are usually empirical and involve additional levels of uncertainty. Although uncertainties in current emission inventories have been discussed extensively, uncertainties resulting from improper spatial proxies have rarely been evaluated. In this work, we investigate the impact of spatial proxies on the representation of gridded emissions by comparing six gridded NO_x emission datasets over China developed from the same magnitude of emissions and different spatial proxies. GEOS-Chem modeled tropospheric NO_2 vertical columns simulated from different gridded emission inventories are compared with satellite-based columns. The results show that differences between modeled and satellite-based NO_2 vertical columns are sensitive to the spatial proxies used in the gridded emission inventories. The total population density is less suitable for allocating NO_x emissions than nighttime light data because population density tends to allocate more emissions to rural areas. Determining the exact locations of large emission sources could significantly strengthen the correlation between modeled and observed NO_2 vertical columns. Using vehicle population and an updated road network for the on-road transport sector could substantially enhance urban emissions and improve the model performance. When further applying industrial gross domestic product (IGDP) values for the industrial sector, modeled NO_2 vertical columns could better capture pollution hotspots in urban areas and exhibit best performance of the six cases comparing to satellite-based NO_2 vertical columns (slope = 1.01 and R^2 = 0.85). This analysis provides a framework for information from satellite observations to inform bottom-up inventory development. In the future, more effort should be devoted to the representation of spatial proxies to improve spatial patterns in bottom-up emission inventories.

1 Introduction

Emission inventories are essential for predicting spatial and temporal variations in air pollutants and for helping policy makers develop pollution control strategies. The traditional way of developing an emission inventory is the bottom-up approach whereby activity rates and emission factors are aggregated for all known sources (e.g., Streets et al., 2003; Zhang et al., 2009). Emission inventories are most commonly estimated as emission totals of municipal districts (e.g., counties, provinces, or countries) because activity data from statistical yearbooks are typically available for such districts (e.g., Woo et al., 2003; Ohara et al., 2007; Kurokawa et al., 2013). However, gridded emissions are needed to apply inventories in chemical transport models.

Many methods for allocating regional emission totals to grids are available. The most accurate approach involves allocating emissions for which the actual latitude/longitude coordinates of the emitting facilities (e.g., power plants or cement plants) are available. For mobile and area sources (i.e., sources for which the exact emission locations are unknown), parameters of so-called spatial proxies must be used to represent the spatial distributions of emissions. For example, emissions from road transportation sources can be allocated based on road networks, and residential emissions that are strongly related to human activities can be gridded using population densities or nighttime lights (Streets et al., 2003; Woo et al., 2003; Ohara et al., 2007; Zhang et al., 2009; Oda et al., 2011). In many cases, industrial emissions are also allocated by proxies because of the limited information available.

The selection of such spatial proxies is empirical, and their representations of real-world spatial emissions patterns are of considerable concern (Zhou and Gurney, 2011; Andres et al., 2012). Recent efforts have been made to reveal the uncertainties of spatial proxies used in bottom-up CO_2 inventories (Rayner et al., 2010; Zhou and Gurney, 2011; Andres et al., 2012; Gately et al., 2013; Gately et al., 2015; Andres et al., 2016), and sophisticated method of allocating emissions to high-resolution grids have been formulated (Gurney et al., 2009; Rayner et al., 2010; Nassar et al., 2013; Asefi-Najafabady et al., 2014). Using population density to downscale fossil fuel emissions can induce biases when analyses are conducted at sub-national scales (Rayner et al., 2010). A high-resolution fossil fuel CO_2 emission inventory for the United States further confirms that source heterogeneities are significant and vary by region and sector, indicating that population density is a biased spatial proxy below the state level (Zhou and Gurney, 2011). The above studies suggest that using population density as a spatial proxy may not be appropriate under the hidden assumption that per capita emissions are homogeneous within a region. However, population density remains one of the most widely used spatial proxies in global and regional emission inventories (Zhang et al., 2009; Lu et al., 2011; Kurokawa et al., 2013), and uncertainties transmitted from improper spatial proxies to chemical transport models have rarely been evaluated.

Recent remarkable development in satellite-based remote sensing instruments, or the so-called top-down approach, provides additional constraints to evaluate and improve the existing understanding of emission inventories (Martin et al., 2008; Streets et al., 2013). Tropospheric column densities of important trace gases, such as NO_2, SO_2, CO, and HCHO, derived from

satellite instruments generate an abundance of useful information on the emission sources of these gases (e.g., Duncan et al., 2010; Boeke et al., 2011; Lin, 2012; Pechony et al., 2013; Stavrakou et al., 2015; Wang et al., 2015; Liu et al., 2016), despite the biases intrinsic to satellite retrievals (Boersma et al., 2008; Lin et al., 2014). Many studies have compared model-simulated column densities with satellite-derived columns to validate the accuracy of bottom-up emissions (e.g., van Noije et al., 2006; Uno et al., 2007; Kim et al., 2009; Sheel et al., 2010; Lin et al., 2010; Itahashi et al., 2014; Han et al., 2015) and have attributed discrepancies between modeled and satellite-based column densities to errors in the magnitudes and/or spatial distributions of the emission inventories used in their models. Inverse modeling techniques can further derive "top-down" emission inventories with optimized magnitudes and emission spatial distributions (e.g., Martin et al., 2003; Jaeglé et al., 2005; Martin et al., 2006; Wang et al., 2007; Lin, 2012; Tang et al., 2013; Stavrakou et al, 2015). For example, Lin (2012) found that the widely used Intercontinental Chemical Transport Experiment-Phase B (INTEX-B) inventory may underestimate NO_x emissions in polluted urban areas or near large point sources. Lamsal et al. (2013) demonstrated that urban NO_2 pollution is a power law scaling function of population size. The exponent values vary by region, reflecting regional differences in industrial development and per capita emissions. Although these top-down studies have identified uncertainties in the spatial representation of current NO_x inventories and have provided correction factors, these factors are difficult to incorporate into bottom-up inventories or apply to emissions of other species because less attention has been paid to gridding processes in bottom-up approaches, which vary by sector and are shared across different species.

In this work, we use NO_x as a case to study the influence of spatial proxies on spatial distributions of bottom-up emission inventories because NO_2 satellite retrieval is less uncertain and the spatial distributions of tropospheric NO_2 vertical columns are similar to those of surface NO_x emissions, especially in the summer when the lifetime of NO_x is short (Richter et al., 2005; Beirle et al., 2003; Lin, 2012). Based on the same magnitude of NO_x emissions, we develop six sets of gridded emission data using different spatial proxies. We then use these gridded emissions and the nested GEOS-Chem to simulate tropospheric NO_2 vertical columns and compare them with satellite-based observations. The effects of spatial proxies on the modeled NO_2 vertical columns and representations of different spatial proxies are evaluated and discussed.

2 Methods and Data

2.1 Review of spatial proxies

We first review the sector-specific spatial proxies utilized in several widely used regional NO_x emission inventories covering China, including TRACE-P (Streets et al., 2003), INTEX-B (Zhang et al., 2009), Regional Emission inventory in Asia (REAS) version 1 (Ohara et al., 2007), and REAS version 2 (Kurokawa et al., 2013), as shown in Table 1. In bottom-up inventories, spatial proxies are usually sector dependent and shared across different species. In general, all of these inventories rely on similar approaches to allocate emissions from combustion sources. Emissions from large-capacity power

generation units are typically allocated according to their latitude/longitude information, which is the most accurate way to distribute emissions, whereas the population density is used to distribute emissions from small power plants whose locations are unknown (Streets et al., 2003; Ohara et al., 2007; Zhang et al., 2009). Recently, power plant locations from the Carbon Monitoring for Action (CARMA) database (Wheeler and Ummel, 2008) are used to locate emissions (Kurokawa et al., 2013),

5 providing a larger dataset of power plants than previous work. However, the accuracy of the CARMA database is still uncertain (Oda et al., 2011).

For industrial and residential combustion, population density is the most frequently used spatial proxy because such emissions are believed to be highly correlated with human activities. Total population density is applied for industrial combustion in the TRACE-P and REAS emission inventories (Streets et al., 2003; Ohara et al., 2007; Kurokawa et al., 2013),

10 which may allocate a large fraction of industrial emissions to rural areas because, in China, the rural population exceeds the urban population (China Statistical Yearbook, National Bureau of Statistics, 2007). In the INTEX-B inventory (Zhang et al., 2009), the urban population is used rather than the total population because industrial activities occur more often in urban areas; however, this strategy is still based on an assumption that per capita industrial emissions are the same across regions within a country. Per capita emissions among regions can differ widely in terms of the regions' levels of economic

15 development and industrial activity or overall standards of living. Lamsal et al. (2013) found that in different urban areas worldwide, per capita emissions varied significantly because of differing energy consumption rates and energy production infrastructure.

For the transportation sector, road networks are widely used to spatially distribute on-road vehicle emissions based on the assumption that traffic volumes remain the same on different types of roads (e.g., Streets et al., 2003; Ohara et al., 2007).

20 However, this is not true because of the existence of varied vehicle populations and road capacities. Commonly used road networks in the above inventories are extracted from the Digital Chart of the World (DCW) (DMA, 1993), which has not been updated since 1992. In this case, using DCW road networks to allocate on-road emissions in China may create significant biases in the spatial distributions of emissions because road construction has occurred continuously over the past two decades.

25 Based on our review of the spatial proxies used in bottom-up emission inventories for China, it can be concluded that such spatial proxies are empirical and may have introduced considerable uncertainties into the spatial distributions of emissions. Studies on fossil fuel CO_2 inventories have made tremendous efforts to improve the spatial distribution of emissions based on satellite-derived nighttime light (Oda et al., 2011), fuel sales and traffic data (McDonald et al., 2014), multivariate regressions (Wang et al., 2013; Gately et al., 2015), or combined fossil fuel data assimilation systems (Rayner et al., 2010;

30 Asefi-Najafabady et al., 2014), shedding new light on ways to improve the spatial distribution of air pollutant inventories.

2.2 Gridded NO$_x$ emission inventory

The bottom-up NO$_x$ emission inventory evaluated in this work is obtained from the Multi-resolution Emission Inventory for China (MEIC, http://www.meicmodel.org) for 2006. The MEIC inventory is developed using a technology-based methodology which estimates anthropogenic emissions in China from ~700 emitting sources (Zhang et al., 2007, 2009; Lei et al., 2011). Table 2 presents the 2006 anthropogenic NO$_x$ emissions estimated by the MEIC for China.

We then develop six sets of gridded NO$_x$ emission data using the same magnitude of emissions from the MEIC and different spatial proxies, as presented in Table 3. Many geographic information system (GIS) grid-based spatial proxies (e.g., population density and road networks) are at a resolution of 1 km × 1 km. In this work, emissions are first gridded at a resolution of 1 km × 1 km and then regridded to a resolution of 0.667° lon × 0.5° lat to fit the GEOS-Chem model. The first two emission datasets evaluate two types of common spatial proxy maps: population distributions and nighttime lights. In the first gridded emission dataset (S1), all emissions are allocated based on population densities obtained from the Landscan population database (ORNL, 2006). The Gridded Population of the World (GPW) population map is also frequently used to allocate emissions, but we do not include it in our analysis because the uncertainties introduced by differences in population maps are minor (Andres et al., 2016). In the second dataset (S2), all emissions are allocated based on the nighttime lights map drawn from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OSL) satellite (http://www.ngdc.noaa.gov/dmsp/download_rad_cal_96-97.html). These two scenarios, which involve single type of spatial proxy for all sectors, are simplifications of the common practice utilized in current emission inventories. We use S1 as the base case and make modifications upon it to evaluate different spatial proxies used in different sectors. Figure 1 compares the spatial distributions of the total population and nighttime light data. The nighttime light data present more significant urban-rural gradients than the total population density data and, thus, may better represent differences in economic development levels between urban and rural areas.

Our previous study revealed that the locations of large point sources in emission inventories significantly affect the prediction accuracy of the chemical transport models (Wang et al., 2012). A third dataset (S3) is used to investigate the effects of using the exact locations of emissions from large point sources. S3 is based on S1 but consists of a unit-based power plant emission dataset including the locations of ~6,400 power generation units across China (Liu et al., 2015) to override power sector emissions.

The fourth dataset (S4) is based on S3 but uses DCW road networks to allocate on-road transportation emissions, which have been applied in several widely used regional emission inventories for China. In S4, emissions are allocated according to the road's length, neglecting the distinctions of road classes and capacities. An improved approach of allocating on-road transport emissions is implemented in the fifth dataset (S5). We use the county-level vehicle population as a spatial proxy to distribute provincial emissions of on-road transportation to each county (an administrative unit one level lower than city). The county-level vehicle population is simulated by the Gompertz Function ($V = V^* \times e^{\alpha e^{\beta E}}$), where V and V^* represents

actual and saturation level of total vehicle ownership (vehicles/1000 people) separately and E represents per-capita GDP (more information is provided in Zheng et al., 2014). The new proxy overrides the assumption of linear relationships between vehicle ownership and population density or road density. Then, county-level emissions are mapped into grids using GIS-based road networks, which includes various road types (highway, national, provincial and county roads) and the total

5   vehicle kilometers traveled data is used as allocation weights on different road types. As stated in Sect. 2.1, the DCW road networks have not been updated since 1992 and, thus, represent the road conditions of the early 1990s (Fig. 2b). Hence, we update the road network data to a new version (China Digital Road-network Map [CDRM] data developed in 2010 by the National Administration of Surveying, Mapping and Geoinformation of China), which reflects the current road conditions (Fig. 2c). Comparisons between county-level vehicle populations, the total population and these two types of road networks

10   are shown in Fig. 2(d-f). Densely populated regions typically have larger vehicle populations and, therefore, greater road network demands. However, the DCW road network fails to provide detailed accounts of the roads in urban areas, and as a result, the emissions allocated to these regions are underestimated. The updated CDRM road networks can help resolve this shortcoming.

In the last dataset (S6), further modifications of S5 are made to better represent the spatial patterns of the industrial sector.

15   We use the Industrial Gross Domestic Product (IGDP) as a first-step spatial proxy instead of population density. Fig. 3 shows the correlations between the normalized industrial emissions and three factors (total population, urban population and IGDP) at the provincial level. IGDP is more closely correlated with emissions ($R^2 = 0.72$) than other factors, indicating that IGDP can better represent the spatial patterns of industrial activities than population density. Using population density as a spatial proxy to allocate industrial emissions assumes that per capita industrial emissions are the same across regions, which

20   may result in underestimations for industrialized regions and overestimations for rural regions. In S6, provincial emissions are first distributed into counties according to the IGDP of each county, and then county-level emissions are allocated to grids based on the population densities drawn from Landscan.

2.3 GEOS-Chem model

The GEOS-Chem model is a global, three-dimensional (3-D) model of atmospheric chemistry that includes > 80 species and >

25   300 reactions (Bey et al., 2001; Park et al., 2004). The nested-grid GEOS-Chem model developed by Chen et al. (2009) is used in this work. It has a horizontal resolution of 0.667° lon × 0.5° lat with 47 vertical layers and a nested-grid domain that covers China and most of its neighboring countries (70°E-150°E, 11°S-55°N). The global model, which has a spatial resolution of 2.5° lon × 2° lat, provides time-varying boundary conditions via the one-way nested approach. Both global and nested simulations are driven by the 3-D meteorological fields of GEOS5 assimilated by the Goddard Earth Observing

30   System (GEOS) at the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO; http://gmao.gsfc.nasa.gov/). Mixing in the planetary boundary layer follows a non-local scheme (Lin and McElroy,

2010) that improves upon previous assumptions of a fully mixed boundary layer. Convection occurs according to a modified Relaxed Arakawa-Schubert scheme (Rienecker et al., 2008).

In this study, we use GEOS-Chem version 09-01-02 to simulate tropospheric NO_2 vertical columns over China for 2006. The EDGAR v3 emission inventory (Olivier and Berdowski, 2001) is used for global anthropogenic emissions, and the East/Southeast Asia region is replaced with the INTEX-B inventory (Zhang et al., 2009). We further override the anthropogenic NO_x emission inventory for China using the six datasets described in Sect. 2.2. To remove the effects of the initial concentration fields, a 1-year spin up is conducted. We use averaged summer (June, July and August [JJA]) NO_2 columns for this evaluation because the short lifetime of NO_2 in the summer favors NO_2 column linkage with local NO_x emissions. We average daily 2-h modeled tropospheric NO_2 vertical columns at a local time of 1300-1500 and sample the model at grids coincident with the daily satellite pixels used in the final average columns.

Grids in the nested-grid GEOS-Chem model are 0.667° lon × 0.5° lat, and their areas range from 2,500-4,000 km^2 over eastern China, comparable to the mean size of a county (~3,000 km^2) in this region. Thus, the spatial pattern of emissions evaluated in this work can represent the spatial variations of emissions at the county level. To draw comparisons with other county-level indicators, gridded NO_2 column densities simulated using this model are resampled to county averages by area weights. In this work, a total of 2,364 county-level districts are covered, including both counties and municipal districts across China.

2.4 Satellite data

The satellite data used in this work comes from the Ozone Monitoring Instrument (OMI) aboard the Aura satellite (Levelt et al., 2006). NO_2 slant column densities are derived using a Differential Optical Absorption Spectroscopy (DOAS) algorithm (Platt, 1994; Boersma et al., 2002; Bucsela et al., 2006). The tropospheric slant NO_2 column densities used in this work are drawn from the Dutch OMI NO_2 (DOMINO) product (version 2, collection 3) (Boersma et al., 2011) available from the Tropospheric Emission Monitoring Internet Service (TEMIS) (http://www.temis.nl/). The air mass factor (AMF) is a multiplicative factor used to convert slant columns into vertical columns (Palmer et al., 2001). The retrieved tropospheric vertical NO_2 column is sensitive to the NO_2 vertical profile used during the AMF calculation. Following Lamsal et al. (2010), we revise the AMF by replacing the original NO_2 profile with that generated from the nested-grid GEOS-Chem model using S1-S6 emissions described above. The new NO_2 vertical profiles have a finer spatial resolution of 0.667° lon × 0.5° lat compared to the original resolution of 3° lon × 2° lat. Six scenario-specific OMI NO_2 products are generated for comparison with corresponding model results, and higher resolution prior profiles could reduce the bias by 3%-6% between modeled and satellite data for different scenarios. In this work, we restrict the use of OMI pixels to those at a solar zenith angle of ≤ 70° and a cloud fraction of ≤ 0.3 in the final averaged columns. Pixels at swath edges (five pixels on each side) are rejected to reduce spatial averaging. Finally, each OMI pixel is allocated to 0.667° lon × 0.5° lat grids by area weights with corner

coordinate information to create daily tropospheric vertical NO_2 column maps. The retrieved NO_2 columns are also resampled from pixels to county averages to draw comparisons with indicators at the county level.

3 Results and Discussion

3.1 Results

5   The spatial distributions of tropospheric NO_2 vertical columns over China in the summer simulated from six gridded inventories are presented and compared with scenario-specific satellite-observed NO_2 vertical columns in Fig. 4. Because all the inventories in model simulations have the same emissions totals, differences among S1-S6 reflect differences in the spatial allocations of the total emissions. In general, modeled and observed NO_2 vertical columns exhibit similar patterns but different fine structures. All modeled cases can reproduce highly polluted areas over the North China Plain and Yangtze

10   River Delta, whereas many pollution hotspots in these regions are underestimated compared to the satellite data in S1. As discussed above, using the total population distribution as a spatial proxy may have misrepresented the urban-rural gradients of economic development levels and allocated disproportionately large fractions of emissions to rural regions. Consequently, urban emission hotspots may have been underestimated because, in China, the rural population exceeds the urban population. Figure 5 compares modeled and satellite-retrieved tropospheric NO_2 vertical columns by county in China for the analyzed

15   six cases. The first column in Fig. 5 compares the model and satellite data for all districts and counties in China for the summer of 2006. Modeled NO_2 columns are generally in good agreement with OMI NO_2 columns, with regression slopes varying from 0.78~1.01 and R^2 values varying from 0.75~0.86. Simulations obtained using S1 substantially underestimate NO_2 columns compared to satellite-based columns, especially in densely populated regions. Tropospheric NO_2 vertical columns simulated using S2 present more pollution hotspots compared to S1, particularly in economically developed regions,

20   such as the capital cities in each province. Using nighttime lights as a spatial proxy (S2) instead of the total population can improve the model's performance. In this case, the slope and R^2 increase to 0.94 and 0.81, respectively, and the normalized mean bias (NMB) increases from -11.1% to -6.0%. These results indicate that the nighttime light map may serve as a better indicator for NO_x emissions than the total population because it can better represent a region's economic development level than the total population.

25   When using the exact positions of power plants instead of total population density for the power sector (S3 vs. S1), most hotspots in the simulated tropospheric NO_2 vertical columns are enhanced, and the discrepancies between the modeled and observed columns are reduced. Model simulations based on S3 correlate better with satellite observations (slope = 0.87 and $R^2 = 0.83$) than those based on S1, proving the importance of determining the positions of large point sources. NO_2 columns simulated using S4 have more bias than those generated using S3, mainly because of underestimations of on-road

30   transportation emissions resulting from the use of outdated DCW road networks as a proxy, as discussed above. Model

simulations based on S5 agree better with satellite-based NO_2 columns (slope = 0.95 and R^2 = 0.86) than those based on S3 and S4, which shows that using vehicle population and CDRM road networks can better represent transportation emissions. When using IGDP to constrain industrial emissions, modeled NO_2 vertical columns further improved (S6 vs. S5). Results indicate that improvements in the transportation sector have more significant effects on modeled NO_2 columns than that in

5   the industrial sector. Finally, the simulations based on S6 exhibit the best performance for all six cases compared to satellite observations (slope = 1.01 and R^2 = 0.85).

To further understand the biases in the model simulations, we divide all counties into three categories—counties in four municipalities (i.e., Beijing, Tianjin, Shanghai and Chongqing), urban counties, and suburban counties according to administrative district definition—and compare the modeled and observed NO_2 columns for these three categories (Fig. 5).

10   For the four Chinese municipalities studied, emission totals are allocated from cities to counties; however, for other provinces, emissions are allocated from province to counties. Municipalities are defined as a separate category in the following analysis.

Model underestimation in S1 mainly occurs for urban counties (Fig. 5c, slope = 0.53 and NMB = -20.3%). This approach, which uses the total population as a spatial proxy, assumes that the per capita emissions in different regions in a province are

15   the same. However, because of varied industrial levels and economic patterns, per capita emissions can be very different across regions. According to Lamsal et al. (2013), the tropospheric NO_2 column is a power law scaling function of the population size, and the exponent is affected by regional differences in per capita emissions. Using population density as a spatial proxy in these areas significantly underestimates the emissions in urban areas, as shown in Fig. 5c. Using nighttime lights to allocate emissions can significantly improve the model's performance for urban areas (Fig. 5g, slope = 1.02 and

20   NMB = -1.8%), although overestimations are identified in a few counties. This finding demonstrates the feasibility of using nighttime lights alone as a spatial proxy when more complex indicators are not available. After determining the exact positions of power plant emissions, the model simulations are substantially improved for all types of regions (Fig. 5j-5l); however, the urban emissions are still underestimated (Slope = 0.69 and NMB = -13.2%), possibly because of underestimations of the industrial and transportation emissions in urban regions when the population density is applied as a

25   spatial proxy. For S4, the model performances are slightly worse than those for S3 for urban regions. In S4, the vehicle populations of different counties are assumed to be linearly correlated with the outdated DCW road networks, and as a result, the on-road transportation emissions in urban areas may be substantially underestimated. Using vehicle population and the updated road networks can significantly improve this issue (S5) and NMB decreases to -4.6%. Finally, for S6, the model performances are further improved for urban regions. Thus, using IGDP and the updated road network as spatial proxies can

30   better represent the emission sources for urban areas.

For the four municipalities studied, total populations and nighttime lights alone cannot effectively represent emission patterns because these proxies are concentrated in city centers, whereas large emission sources, such as power plants and

industrial activities, have largely been relocated away from urban regions in municipalities. When the locations of power emissions are considered, the correlation between the modeled and observed NO_2 columns is significantly improved (Fig. 5j, 5n, and 5r). However, using the IGDP and updated road network as spatial proxies disproportionately concentrates emissions in urban areas, resulting in an overestimation of modeled NO_2 columns.

5   Fig. 6 presents the distributions of ratios between simulated and satellite-based county-level NO_2 column densities for the six cases. We remove those counties with OMI NO_2 columns of less than 3×10^{15} molecules cm^{-2} to avoid the influence of the background areas with more uncertain retrieved columns. After rejecting the background regions, ~770 counties (33%) covering much of eastern China remain. Model simulations based on S6 exhibit the best performance. Differences between the modeled and satellite-based NO_2 columns are within 20% for 391 counties in S6 compared to 310 and 331 counties in S1

10 and S2, respectively. Model simulations of S1 present large negative biases compared to satellite-based NO_2 columns, with 119 counties underestimated by over 50% (66 counties for S6). However, for S2, positive biases between the modeled and satellite-based NO_2 columns exceed 50% for 28 counties (10 counties for S5), indicating that using nighttime light maps may overestimate urban emissions in certain regions.

3.2 Uncertainties

15 This work is subject to several uncertainties. Biases between model simulations and satellite data come not only from emission inventories but also from the model itself or satellite retrievals. Potential errors in nested-grid GEOS-Chem model simulations are compounded by errors in GEOS-5 meteorological fields, PBL heights, and a variety of chemical parameters selected in a given model. Model simulation errors for eastern China are estimated to present a negative systematic bias of 10–20% (season dependent) plus a random error of 30% according to previous work (Martin et al., 2003; Lin and McElroy,

20 2011; Lin et al., 2012). Sensitivity simulations (Lin et al., 2012) of the above model factors show that none of them can fully explain the bias between model simulations and satellite observations. Combining all these modifications can achieve better agreement with satellite observations but cannot eliminate the negative biases associated with extremely polluted locations (Lin et al., 2012), suggesting that model errors are not the primary cause of model-satellite biases for urban areas.

The stated uncertainties of individual DOMINO v2.0 NO_2 column retrievals are estimated at 1.0×10^{15} molecules cm^{-2} + 25%

25 (Boersma et al., 2011), which demonstrates the dominance of errors arising during the calculation of AMF in polluted areas (Boersma et al., 2007). In particular, DOMINO v2.0 OMI products do not explicitly account for the effects of aerosols on solar radiation, which are important for the calculation of AMF and particularly significant for eastern China because of its high aerosol loadings. In Lin et al. (2014), explicitly including aerosol scattering and absorption exerts either positive or negative effects on retrieved NO_2, with a mean effect of 14%. However, aerosol effects cannot fully explain the large

30 discrepancies observed between model and satellite results for urban areas.

Biases in the satellite NO_2 data could affect the comparisons for different scenarios. Previous studies indicated that the OMI NO_2 columns could be biased high for 20% over China (Irie et al., 2008; Lin et al., 2014). In this case, the best scenario of spatial proxies would be S4 instead of S6. If the OMI data is biased low instead of biased high, the best scenario will remain the same but with less agreement compared to satellite observations. However, given the high sensitivity of modeled NO_2 columns to spatial proxies, we can still conclude that the spatial proxies used in gridded emission inventories could affect the representation of bottom-up emission inventory significantly.

Other factors, such as the resolution, may also introduce uncertainty. The spatial proxies in S6 are quite good at the resolution of 0.667° lon × 0.5° lat used in this work, which roughly corresponds to the county level in eastern China. However, they may not be suitable at other resolutions. Further work based on models with finer grids should be performed to explore appropriate spatial proxies at finer resolutions. Natural emissions from lightning and soil sources are not discussed in this work, although they are suggested to be underestimated by approximately 16% for China for 2006; they account for less than 3% and 6% of anthropogenic emissions, respectively, and even less in highly polluted regions (Lin, 2012).

3.3 Discussion

In this work, we use NO_x emissions to relate the biases between model simulations and observations to local emissions and evaluate the impacts of spatial proxies on the distributions of bottom-up emission inventories at the county level using satellite constraints. Insight obtained from this work can be applied to other species generated from fossil fuel combustion (e.g., SO_2 and CO_2) because they typically come from the same sources. Our method represents a feasible approach to studying species that are difficult to validate directly because suitable observation data are not available and/or lifetimes are long.

As shown in Sect. 2.1 and described in this work, regardless of how spatial proxies are adjusted in a bottom-up inventory, they are always empirical and contribute uncertainties to the spatial representation of emission inventories. A companion paper of this work found that large uncertainties existed in proxy-based inventories on urban scale (less than 0.25°) by comparing the proxy-based inventory with an inventory developed from exact locations of emitting facilities in Hebei, China (Zheng et al., 2017). Critical evaluations must be conducted to ensure the accuracy of these proxies. Our work presents a practical means to diagnose this problem and involves using satellite observations as an indicator of ground emissions to determine the relationships between emissions and local parameters. Our approach also has the attribute of propagating information from satellite observations into bottom-up inventory developments.

The approach presented here can also be expanded to other regions because a universal relationship exists between emissions and the economy. Spatial proxies that work well for China may not be suitable for other regions because of regional differences in energy consumption, industrial development and living standards, as demonstrated in Lamsal et al. (2013).

However, by integrating local satellite observation data, a better understanding of the spatial distribution of emissions and their relationships to local parameters can be obtained, which has implications for the local selection of spatial proxies.

4 Concluding Remarks

The spatial proxies used when developing gridded emissions inventories are empirical and can introduce uncertainties in bottom-up emissions inventories. This issue has rarely been evaluated. In this work, we evaluate the effects of spatial proxies on the representations of spatial distributions of emissions using an integrated framework of bottom-up emission inventories, a chemical transport model, and satellite observations. We first develop six sets of gridded NO_x emissions for China using the same magnitude of emissions from the MEIC and different spatial proxies. The spatial proxies considered in this study include the following: the total population, nighttime lights, the locations of power plants, IGDP, vehicle populations and two different road network datasets. The nested-grid GEOS-Chem model is then used to simulate tropospheric NO_2 vertical columns using the six gridded emissions, and modeled NO_2 columns are compared to satellite-based NO_2 columns derived from OMI data.

We found that the spatial proxies used in gridded emission inventories significantly affect simulated NO_2 columns. The model performance is largely dependent on the representations of urban emissions in the bottom-up inventory, which are very sensitive to spatial proxies. Using the total population density tends to allocate more emissions to rural areas and to underestimate NO_2 columns compared to satellite observations. Nighttime lights represent urban emissions better than population density because they correlate more closely with economic development levels. When using sophisticated combinations of different proxies to represent urban emissions (i.e., positions of large point sources, IGDP, vehicle populations, and the most recent road network), modeled NO_2 columns agree better with satellite observations, indicating that improving the spatial representation of emissions could significantly increase the accuracy of emission inventories.

The results of this work emphasize the importance of spatial proxies for bottom-up emission inventory development. Discrepancies between models and observations should be attributed to not only errors in the magnitude of total emission estimates but also spatial proxies. Although the selection of spatial proxies in this work is still empirical and may not represent the best case, we illustrate methods for improving gridded emission inventories by carefully selecting spatial proxies. This study provides a framework to apply information from satellite observations to inform bottom-up inventory development. The approach used here could be further extended to other species and regions, and more advanced optimized approaches could be introduced into the development of emission inventories of different air pollutants (Asefi-Najafabady et al., 2014; Gately et al., 2013; Nassar et al., 2013). More efforts should be made to improve the spatial distributions of bottom-up emission inventories in the future.

Acknowledgements. This work was supported by China's National Basic Research Program (2014CB441301), the National Science Foundation of China (41222036, 41275026, and 21221004), and the MarcoPolo project of the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement number 606953. We acknowledge our free use of tropospheric NO_2 column data from OMI: www.temis.nl.

5 References

Andres, R. J., Boden, T. A., Breon, F. M., Ciais, P., Davis, S., Erickson, D., Gregg, J. S., Jacobson, A., Marland, G., Miller, J., Oda, T., Olivier, J. G. J., Raupach, M. R., Rayner, P., and Treanton, K.: A synthesis of carbon dioxide emissions from fossil-fuel combustion, Biogeosciences, 9, 1845–1871, doi:10.5194/bg-9-1845-2012, 2012.

Andres, R. J., Boden, T. A., and Higdon, D. M.: Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC
10 example, Atmos. Chem. Phys. Discuss., 2016, 1-56, 2016.

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global gridded fossil fuel CO_2 emission data product: Evaluation and analysis of results, J. Geophys. Res., 119, 2013JD021296, 2014.

Beirle, S., Platt, U., Wenig, M., and Wagner, T.: Weekly cycle of NO_2 by GOME measurements: a signature of
15 anthropogenic sources, Atmos. Chem. Phys., 3, 2225-2232, 2003.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q. B., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23073–23095, doi:10.1029/2001JD000807, 2001.

Boeke, N. L., Marshall, J. D., Alvarez, S., Chance, K. V., Fried, A., Kurosu, T. P., Rappenglueck, B., Richter, D., Walega, J.,
20 Weibring, P., and Millet, D. B.: Formaldehyde columns from the Ozone Monitoring Instrument: Urban versus background levels and evaluation using aircraft data and a global model, J. Geophys. Res., 116, D05303, doi:10.1029/2010JD014870, 2011.

Boersma, K. F., Bucsela, E., Brinksma, E., and Gleason, J. F.: NO_2, in OMI Algorithm Theoretical Basis Document, vol. IV, Trace Gas Algorithms, edited by: Chance K., NASA Goddard Space Flight Cent., Greenbelt, Md, 15–36, 2002.

25 Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmos. Chem. Phys., 7, 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Jacob, D. J., Bucsela, E. J., Perring, A. E., Dirksen, R., van der A, R. J., Yantosca, R. M., Park, R. J., Wenig, M. O., Bertram, T. H., and Cohen, R. C.: Validation of OMI tropospheric NO_2 observations during INTEX-B and
30 application to constrain NO_x emissions over the eastern United States and Mexico, Atmos. Environ., 42, 4480-4497, 2008.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmos. Meas. Tech., 4, 1905-1928, doi:10.5194/amt-4-1905-2011, 2011.

Bucsela, E. J., Celarier, E. A., Wenig, M. O., Gleason, J. F., Veefkind, J. P., Boersma, K. F., and Brinksma, E. J.: Algorithm
5 for NO_2 vertical column retrieval from the Ozone Monitoring Instrument, IEEE T. Geosci. Remote Sens., 44, 1245–1258, doi:10.1109/TGRS.2005.863715, 2006.

CIESIN (Center for International Earth Science Information Network), Columbia University, United Nations Food and Agriculture Programme (FAO), and Centro Internacional de Agricultura Tropical (CIAT): Gridded Population of the World, Version 3 (GPWv3), Population Count Grid, Future Estimates, Palisades, NY, NASA Socioeconomic Data and Applications
10 Center (SEDAC), available at: http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-count-future-estimates (last access: 10 November 2013), 2005.

CIESIN (Center for International Earth Science Information Network), Columbia University, International Food Policy Research Institute (IFPRI), The World Bank, and Centro Internacional de Agricultura Tropical (CIAT): Global Rural-Urban Mapping Project, Version 1 (GRUMPv1), Population Count Grid, Palisades, NY, NASA Socioeconomic Data and
15 Applications Center (SEDAC), available at: http://sedac.ciesin.columbia.edu/data/set/grump-v1-population-count (last access: 10 November 2013), 2011.

Chen, D., Wang, Y., McElroy, M. B., He, K., Yantosca, R. M., and Le Sager, P.: Regional CO pollution and export in China simulated by the high-resolution nested-grid GEOS-Chem model, Atmos. Chem. Phys., 9, 3825–3839, doi:10.5194/acp-9-3825-2009, 2009.

20 Defense Mapping Agency (DMA): Digital Chart of the World Database, Washington DC, USA, 1993.

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J., and Crawford, J. H.: Application of OMI observations to a space-based indicator of NO_x and VOC controls on surface ozone formation, Atmos. Environ., 44, 2213–2223, 2010.

Gately, C. K., Hutyra, L. R., Wing, I. S., and Brondfield, M. N.: A Bottom up Approach to on-Road CO_2 Emissions
25 Estimates: Improved Spatial Accuracy and Applications for Regional Planning, Environ. Sci. Technol., 47, 2423-2430, 10.1021/es304238v, 2013.

Gately, C. K., Hutyra, L. R., and Sue Wing, I.: Cities, traffic, and CO_2: A multidecadal assessment of trends, drivers, and scaling relationships, Proc. Natl. Acad. Sci., 112, 4999-5004, 10.1073/pnas.1421723112, 2015.

Graedel, T. F., Bates, T. S., Bouman, A. F., Cunnold, D., Dignon, J., Fung, I., Jacob, D. J., Lamb, B. K., Logan, J. A.,
30 Marland, G., Middleton, P., Pacyna, J. M., Placet, M., and Veldt, C.: A compilation of inventories of emissions to the atmosphere, Global Biogeochem. Cycle, 7, 1–26, doi: 10.1029/92GB02793, 1993.

Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S., and de la Rue du Can, S.: High Resolution Fossil Fuel Combustion CO_2 Emission Fluxes for the United States, Environ. Sci. Technol., 43, 5535-5541, 10.1021/es900806c, 2009.

Han, K. M., Lee, S., Chang, L. S., and Song, C. H.: A comparison study between CMAQ-simulated and OMI-retrieved NO_2 columns over East Asia for evaluation of NO_x emission fluxes of INTEX-B, CAPSS, and REAS inventories, Atmos. Chem. Phys., 15, 1913-1938, 2015.

Irie, H., Kanaya, Y., Akimoto, H., Tanimoto, H., Wang, Z., Gleason, J. F., and Bucsela, E. J.: Validation of OMI tropospheric NO_2 column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment 2006, Atmos. Chem. Phys., 8, 6577-6586, 2008.

Itahashi, S., Uno, I., Irie, H., Kurokawa, J. I., and Ohara, T.: Regional modeling of tropospheric NO_2 vertical column density over East Asia during the period 2000–2010: comparison with multisatellite observations, Atmos. Chem. Phys., 14, 3623-3635, 2014.

Jaeglé, L., Steinberger, L., Martin, R. V., and Chance, K.: Global partitioning of NO_x sources using satellite observations: Relative roles of fossil fuel combustion, biomass burning and soil emissions, Faraday Discus., 130, 407–423, doi:10.1039/b502128f, 2005.

Kim, S. W., Heckel, A., Frost, G. J., Richter, A., Gleason, J., Burrows, J. P., McKeen, S., Hsie, E. Y., Granier, C., and Trainer, M.: NO_2 columns in the western United States observed from space and simulated by a regional chemistry model and their implications for NO_x emissions, J. Geophys. Res., 114, D11301, doi:10.1029/2008JD011343, 2009.

Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima, K., and Akimoto, H.: Emissions of air pollutants and greenhouse gases over Asian regions during 2000–2008: Regional Emission inventory in ASia (REAS) version 2, Atmos. Chem. Phys., 13, 11019-11058, 10.5194/acp-13-11019-2013, 2013.

Lamsal, L. N., Martin, R. V., van Donkelaar, A., Celarier, E. A., Bucsela, E. J., Boersma, K. F., Dirksen, R., Luo, C., and Wang, Y.: Indirect validation of tropospheric nitrogen dioxide retrieved from the OMI satellite instrument: Insight into the seasonal variation of nitrogen oxides at northern midlatitudes, J. Geophys. Res., 115, D05302, doi:10.1029/2009JD013351, 2010.

Lamsal, L.N., Martin, R.V., Parrish, D.D., and Krotkov, N.A.: Scaling relationship for NO_2 pollution and urban population size: A satellite perspective, Environ. Sci. Technol., 47, 7855–7861, 2013.

Lei, Y., Zhang, Q., He, K. B., and Streets, D. G.: Primary anthropogenic aerosol emission trends for China, 1990–2005, Atmos. Chem. Phys., 11, 931-954, 2011.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., M¨alkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, IEEE T. Geosci. Remote Sens., 44, 1093–1101, doi:10.1109/TGRS.2006.872333, 2006.

Lin, J.-T.: Satellite constraint for emissions of nitrogen oxides from anthropogenic, lightning and soil sources over East China on a high-resolution grid, Atmos. Chem. Phys., 12, 2881-2898, doi:10.5194/acp-12-2881-2012, 2012.

Lin, J.-T., R. V. Martin, K. F. Boersma, M. Sneep, P. Stammes, R. Spurr, P. Wang, M. Van Roozendael, K. Clémer, and H. Irie: Retrieving tropospheric nitrogen dioxide from the Ozone Monitoring Instrument: Effects of aerosols, surface reflectance anisotropy, and vertical profile of nitrogen dioxide, Atmos. Chem. Phys., 14, 1441-1461, doi:10.5194/acp-14-1441-2014, 2014

Lin, J.-T., McElroy, M. B., and Boersma, K. F.: Constraint of anthropogenic NO_x emissions in China from different sectors: a new methodology using multiple satellite retrievals, Atmos. Chem. Phys., 10, 63–78, doi:10.5194/acp-10-63-2010, 2010.

Lin, J.-T. and McElroy, M. B.: Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing, Atmos. Environ., 44, 1726-1739, 2010.

Liu, F., Beirle, S., Zhang, Q., Dörner, S., He, K., and Wagner, T.: NO_x lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations, Atmos. Chem. Phys., 16, 5283-5298, 2016.

Liu, F., Zhang, Q., Tong, D., Zheng, B., Li, M., Huo, H., and He, K. B.: High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010, Atmos. Chem. Phys., 15, 13299-13317, 2015.

Lu, Z., Zhang, Q., and Streets, D. G.: Sulfur dioxide and primary carbonaceous aerosol emissions in China and India, 1996-2010, Atmos. Chem. Phys., 11, 9839–9864, doi:10.5194/acp-11-9839-2011, 2011.

Martin, R. V.: Satellite remote sensing of surface air quality, Atmos. Environ., 42, 7823-7843, 2008.

Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, J. Geophys. Res., 108, 4537, doi:10.1029/2003JD003453, 2003.

Martin, R. V., Sioris, C. E., Chance, K., Ryerson, T. B., Bertram, T. H., Wooldridge, P. J., Cohen, R. C., Andy Neuman, J., Swanson, A., and Flocke, F. M.: Evaluation of space-based constraints on global nitrogen oxide emissions with regional aircraft measurements over and downwind of eastern North America, J. Geophys. Res., 111, D15308, doi:10.1029/2005JD006680, 2006.

McDonald, B. C., McBride, Z. C., Martin, E. W., and Harley, R. A.: High-resolution mapping of motor vehicle carbon dioxide emissions, J. Geophys. Res., 119, 2013JD021219, 2014.

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., and Deng, F.: Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets, J. Geophys. Res., 118, 917–933, doi:10.1029/2012JD018196, 2013.

National Bureau of Statistics (NBS): China Statistical Yearbook, China Statistics Press, Beijing, China, 4, 2007.

Oak Ridge National Laboratory (ORNL): LandScan Global Population Database, Oak Ridge National Laboratory, Oak Ridge, Tenn., USA, 1999, 2001, 2006, 2008, 2009.

Oda, T., and Maksyutov, S.: A very high-resolution (1 km x 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543–556, doi:10.5194/acp-11-543-2011, 2011.

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, Atmos. Chem. Phys., 7, 4419–4444, doi:10.5194/acp-7-4419-2007, 2007.

Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in: The Climate System, Lisse, The Netherlands, 33–78, 2001.

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106, 14539-14550, 10.1029/2000jd900772, 2001.

Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, J. Geophys. Res., 109, D15204, doi:10.1029/2003JD004473, 2004.

Pechony, O., Shindell, D. T., and Faluvegi, G.: Direct top-down estimates of biomass burning CO emissions using TES and MOPITT versus bottom-up GFED inventory, J. Geophys. Res., 118, 8054-8066, 2013.

Platt, U.: Differential Optical Absorption Spectroscopy (DOAS), in: Air Monitoring by Spectroscopic Techniques, John Wiley, New York, USA, 27–84, 1994.

Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., and Koffi, E.: A new global gridded dataset of CO_2 emissions from fossil fuel combustion: Methodology and evaluation, J. Geophys. Res., 115, D19306, doi:10.1029/2009JD013439, 2010.

Richter, A., Burrows, J. P., N¨uß, H., Granier, C., and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129–132, 2005.

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. D., Gelaro, R., Stajner, I., and Nielsen, E.: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 9.1.0, and 5.2.0, NASA, 2008.

Shah, J., Nagpal, T., Johnson, T., Amann, M., Carmichael, G., Foell, W., Green, C., Hettelingh, L. P., Hordijk, L., Li, J., Peng, C., Pu, Y. F., Ramankutty, R., and Streets, D.: Integrated analysis for acid rain in Asia: Policy implications and results of RAINS-ASIA model, Annu. Rev. Energy Environ., 25, 339–375, doi:10.1146/annurev.energy.25.1.339, 2000.

Sheel, V., Lal, S., Richter, A., and Burrows, J. P.: Comparison of satellite observed tropospheric NO_2 over India with model simulations, Atmos. Environ., 44, 3314–3321, 2010.

Stavrakou, T., Müller, J. F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P. F., and Guenther, A.: How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?, Atmos. Chem. Phys., 15, 11861-11884, 2015.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res., 108, 8809, doi:10.1029/2002JD003093, 2003.

Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P., Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y., Lu, Z., Martin, R. V., Pfister, G. G., Pinder, R. W., Salawitch, R. J., and Wecht, K. J.: Emissions estimation from satellite retrievals: A review of current capability, Atmos. Environ., 77, 1011-1042, 2013.

Tang, W., Cohan, D. S., Lamsal, L. N., Xiao, X., and Zhou, W.: Inverse modeling of Texas NO_x emissions using space-based and ground-based NO_2 observations, Atmos. Chem. Phys., 13, 11005-11018, 2013.

Uno, I., He, Y., Ohara, T., Yamaji, K., Kurokawa, J.-I., Katayama, M., Wang, Z., Noguchi, K., Hayashida, S., Richter, A., and Burrows, J. P.: Systematic analysis of interannual and seasonal variations of model-simulated tropospheric NO_2 in Asia and comparison with GOME-satellite data, Atmos. Chem. Phys., 7, 1671–1681, doi:10.5194/acp-7-1671-2007, 2007.

van Noije, T. P. C., Eskes, H. J., Dentener, F. J., Stevenson, D. S., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Boersma, K. F., Butler, T., Cofala, J., Drevet, J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Krol, M. C., Lamarque, J. F., Lawrence, M. G., Martin, R. V., Montanaro, V., Mueller, J. F., Pitari, G., Prather, M. J., Pyle, J. A., Richter, A., Rodriguez, J. M., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., and van Roozendael, M.: Multi-model ensemble simulations of tropospheric NO_2 compared with GOME retrievals for the year 2000, Atmos. Chem. Phys., 6, 2943-2979, 2006.

Wang, R., Tao, S., Ciais, P., Shen, H. Z., Huang, Y., Chen, H., Shen, G. F., Wang, B., Li, W., Zhang, Y. Y., Lu, Y., Zhu, D., Chen, Y. C., Liu, X. P., Wang, W. T., Wang, X. L., Liu, W. X., Li, B. G., and Piao, S. L.: High-resolution mapping of combustion processes and implications for CO_2 emissions, Atmos. Chem. Phys., 13, 5189-5203, doi:10.5194/acp-13-5189-2013, 2013.

Wang, S., Zhang, Q., Martin, R. V., Philip, S., Liu, F., Li, M., Jiang, X., and He, K.: Satellite measurements oversee China's sulfur dioxide emission reductions from coal-fired power plants, Environ. Res. Lett., 10, 2015.

Wang, S. W., Zhang, Q., Streets, D. G., He, K. B., Martin, R. V., Lamsal, L. N., Chen, D., Lei, Y., and Lu, Z.: Growth in NO_x emissions from power plants in China: bottom-up estimates and satellite observations, Atmos. Chem. Phys., 12, 4429–4447, doi:10.5194/acp-12-4429-2012, 2012.

Wang, Y., McElroy, M. B., Martin, R. V., Streets, D. G., Zhang, Q., and Fu, T.-M.: Seasonal variability of NO_x emissions over east China constrained by satellite observations: Implications for combustion and microbial sources, J. Geophys. Res., 112, D06301, doi:10.1029/2006JD007538, 2007.

Wheeler, D. and Ummel, K.: Calculating CARMA: Global Estimation of CO_2 Emissions From the Power Sector, http://www.cgdev.org/files/16101 file Calculating CARMA FINAL.pdf, 2008.

Woo, J.-H., Baek, J. M., Kim, J.-W., Carmichael, G. R., Thongboonchoo, N., Kim, S. T., and An, J. H.: Development of a multi-resolution emission inventory and its impact on sulfur distribution for Northeast Asia, Water Air Soil Poll., 148, 259–278, 2003.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131–5153, doi:10.5194/acp-9-5131-2009, 2009.

Zhang, Q., Streets, D. G., He, K., Wang, Y., Richter, A., Burrows, J. P., Uno, I., Jang, C. J., Chen, D., Yao, Z., and Lei, Y.: NO_x emission trends for China, 1995–2004: The view from the ground and the view from space, J. Geophys. Res., 112, D22306, 2007.

Zheng, B., Huo, H., Zhang, Q., Yao, Z. L., Wang, X. T., Yang, X. F., Liu, H., and He, K. B.: High-resolution mapping of vehicle emissions in China in 2008, Atmos. Chem. Phys., 14, 9787-9805, 10.5194/acp-14-9787-2014, 2014.

Zheng, B., Zhang, Q., Tong, D., Chen, C., Hong, C., Li, M., Geng, G., Lei, Y., Huo, H., and He, K.: Resolution dependence of uncertainties in gridded emission inventories: a case study in Hebei, China, Atmos. Chem. Phys., 17, 921-933, 2017.

Zhou, Y., and Gurney, K. R.: Spatial relationships of sector-specific fossil fuel CO_2 emissions in the United States, Global Biogeochem. Cycles, 25, GB3002, doi:10.1029/2010GB003822, 2011.

Table 1. Review of the spatial proxies used in regional bottom-up NO$_x$ emission inventories covering China.

Inventories	Sectors	Spatial proxies	Data sources
TRACE-P (Streets et al., 2003)	Large power plants	Location	RAINS-Asia (Shah et al., 2000) and GEIA inventory (Graedel et al., 1993)
	Small power plants	Total population	LandScan[1]
	Industrial combustion	Total population	LandScan
	Residential fossil fuel	Total population	LandScan
	Residential biofuel	Rural population	LandScan
	On-road transport	Road networks	DCW[2]
	Off-road transport	Total population	LandScan
REAS v.1.1 (Ohara et al., 2007)	Large power plants	Location	China State Grid Company
	Small power plants	Total population	LandScan
	Industrial combustion	Total population	LandScan
	Residential fossil fuel	Total population	LandScan
	Residential biofuel	Rural population	LandScan
	On-road transport	Road networks	DCW
	Off-road transport	Total population	LandScan
REAS v.2 (Kurokawa et al., 2013)	Large power plants	Location	CARMA (Wheeler and Ummel, 2008)
	Small power plants	Total population	GPWv3[3]
	Industrial combustion	Total population	GPWv3
	Residential fossil fuel	Total population	GPWv3
	Residential biofuel	Rural population	GPWv3 and GRUMPv1[4]
	On-road transport	Road networks	DCW
	Off-road transport	Total population	GPWv3
INTEX-B (Zhang et al., 2009)	Large power plants	Location	Ministry of Environmental Protection
	Small power plants	Total population	LandScan
	Industrial combustion	Urban/rural population	LandScan
	Residential fossil fuel	Total population	LandScan
	Residential biofuel	Rural population	LandScan
	On-road transport	Road networks	DCW
	Off-road transport	Total population	LandScan

[1]LandScan Global Population database (ORNL, 1999, 2001, 2006)

[2]DCW, Digital Chart of the World (DMA, 1993)

[3]GPWv3, Gridded Population on the World (CIESIN et al., 2005, 2011)

5 [4]GRUMPv1, Global Rural-Urban Mapping Project (CIESIN et al., 2005, 2011)

Table 2. Anthropogenic NO$_x$ emissions by sector in China for 2006 from the MEIC inventory.

Sector	Annual Emissions (Tg)
Power plants	8.31
Industry	7.34
On-road transport	4.51
Off-road transport	1.88
Residential (biofuel + fossil fuel)	1.00
Total	23.04

Table 3. Spatial proxies used in the gridding process for the six emission scenarios developed in this study.

Gridding process	Sectors	S1	S2	S3	S4	S5	S6
Province to county	Power Plant	TP[a]	NL[b]	N/A	N/A	N/A	N/A
	Industry	TP	NL	TP	TP	TP	IGDP[c]
	Residential	TP	NL	TP	TP	TP	TP
	On-road transport	TP	NL	TP	DCW[d]	VP[e]	VP
	Off-road transport	TP	NL	TP	TP	TP	TP
County to grid	Power Plant	TP	NL	PS[f]	PS	PS	PS
	Industry	TP	NL	TP	TP	TP	TP
	Residential	TP	NL	TP	TP	TP	TP
	On-road transport	TP	NL	TP	DCW	CDRM[g]	CDRM
	Off-road transport	TP	NL	TP	TP	TP	TP

[a]TP: Total population from the Landscan population database (ORNL, 2006);

[b]NL: Nighttime light from the DMSP-OSL satellite (http://www.ngdc.noaa.gov/dmsp/download_rad_cal_96-97.html);

[c]IGDP: Industrial GDP (China Statistical Yearbook, National Bureau of Statistics, 2007);

[d]DCW: Road networks from the Digital chart of the world (DMA, 1993);

[e]VP: Vehicle population (Zheng et al., 2014);

[f]PS: Coordinates of point sources (Liu et al., 2015);

[g]CDRM: Road networks from the China Digital Road Map developed by the National Administration of Surveying, Mapping and Geoinformation of China.
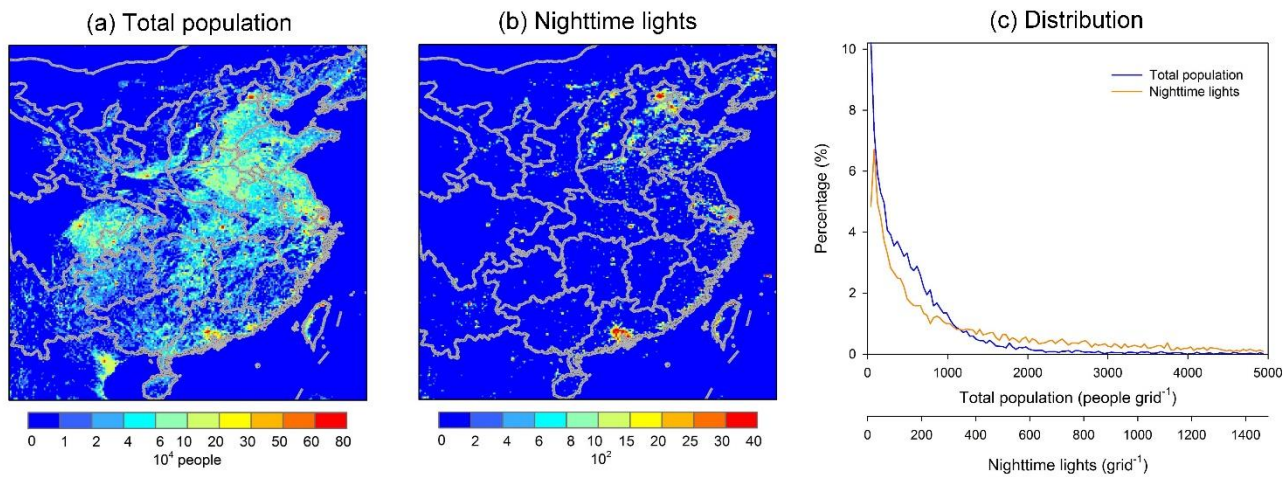
Fig. 1. Spatial patterns of (a) the total population density and (b) nighttime lights for eastern China, and (c) the distributions of total population density and nighttime lights in China at a resolution of $0.1° \times 0.1°$.
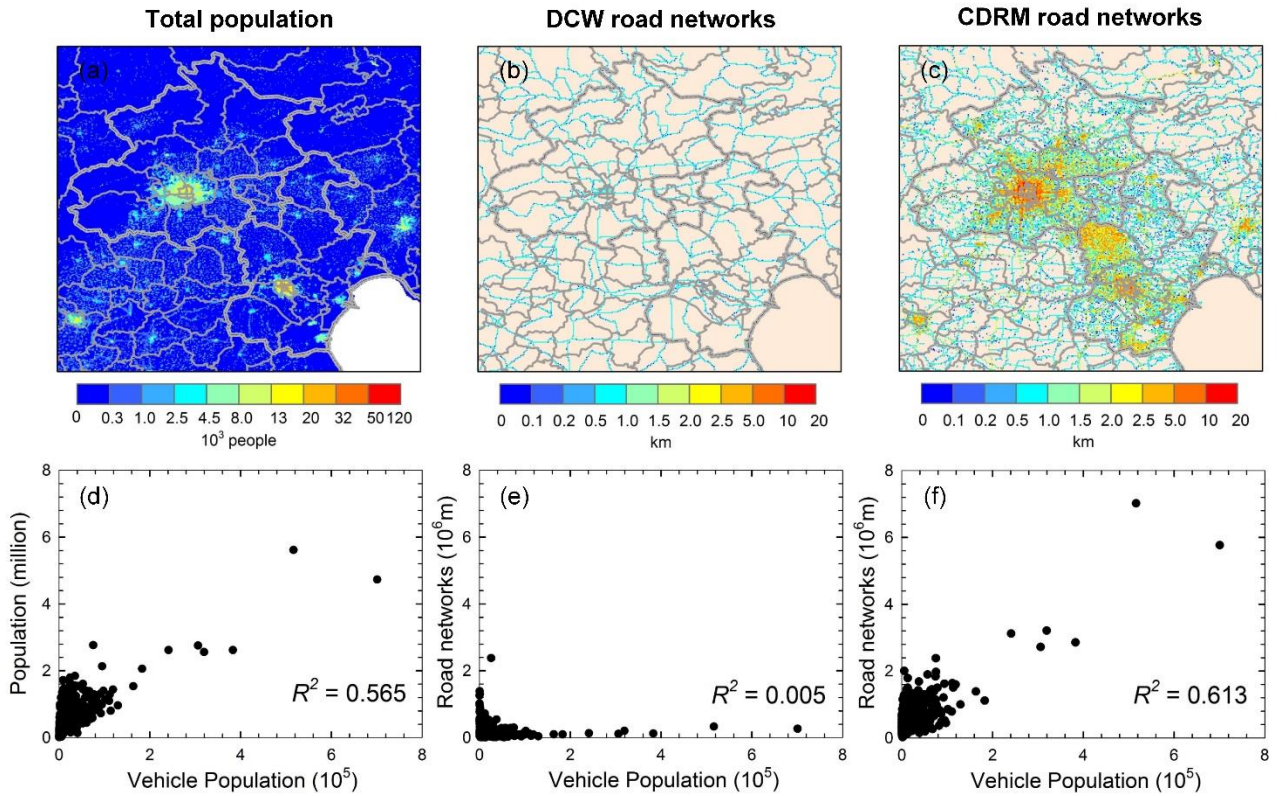
Fig. 2. Comparisons between the total population (left column), DCW road networks (middle column) and CDRM road networks (right column). (a-c) Maps of the total population and two road networks in the Beijing-Tianjin region as an example. (d-f) Comparison with county-level vehicle populations.
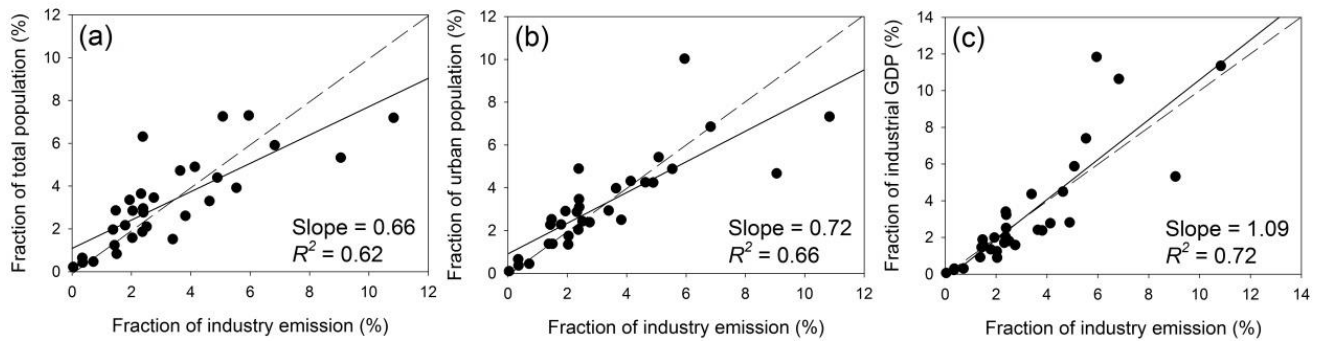
Fig. 3. Correlations between normalized provincial industrial emissions and three types of spatial proxies: (a) total population, (b) urban population and (c) IGDP data.
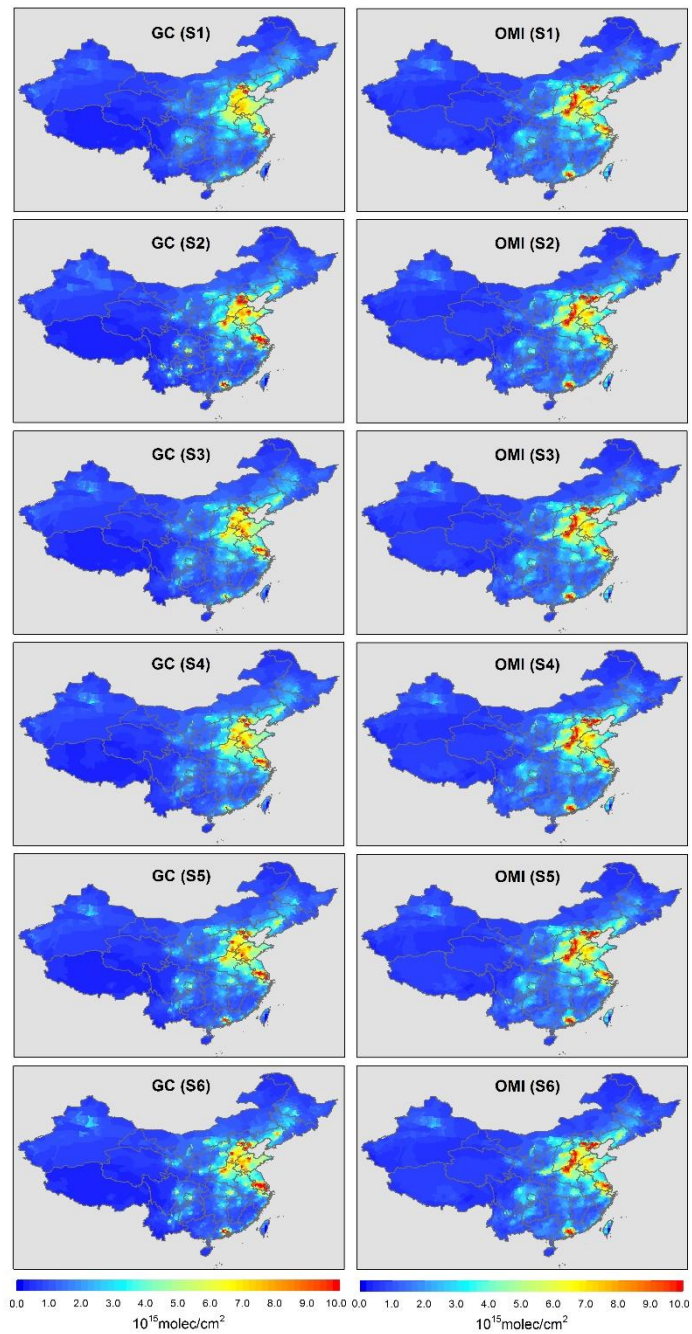
Fig. 4. Spatial distributions of summer averaged tropospheric NO₂ vertical columns modeled by GEOS-Chem based on S1-S6 emissions and scenario-specific data retrieved from the OMI for 2006.
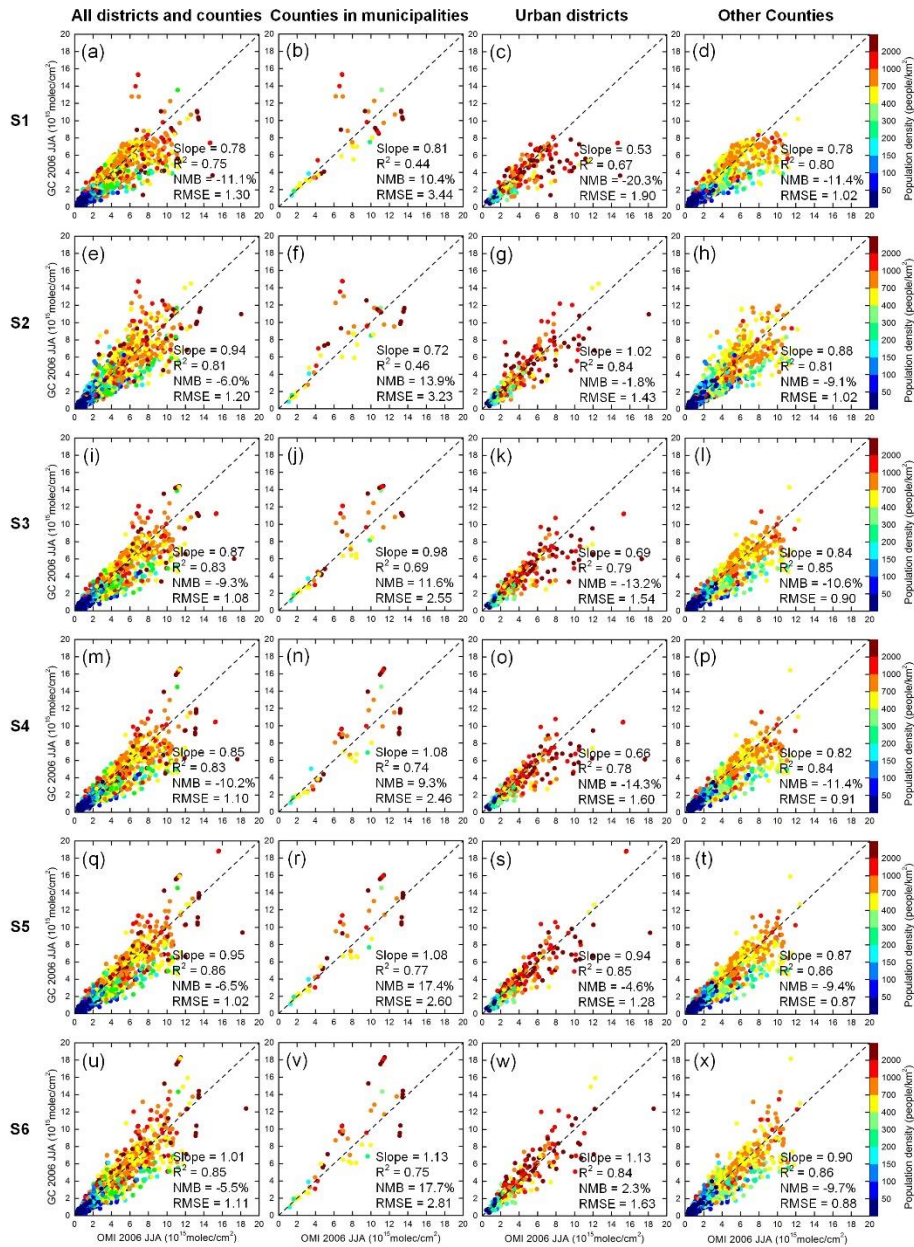
Fig. 5. Comparisons between county-level model simulations from the six emission cases and OMI NO₂ vertical columns for all counties (first column), urban districts in municipalities (second column), urban districts in other cities (third column) and other counties (fourth column). The color of each symbol corresponds to the population density in the county specified by that symbol. The dotted line has a slope of 1.
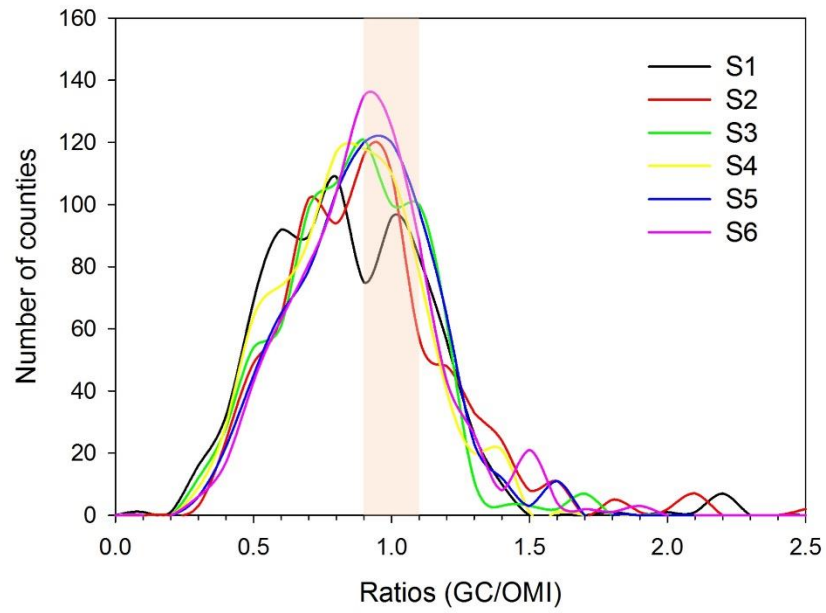
Fig. 6. Distributions of the ratios between the county-level model simulations and satellite observations from the six emission inventories. The shaded region indicates a range of 0.9~1.1.