Supporting Information

Considerable unaccounted local sources of NOx emissions in China revealed from satellite

*Hao Kong[1], Jintai Lin[1, *], Lulu Chen[1], Yuhang Zhang[1], Yingying Yan[2], Mengyao Liu[3], Ruijing Ni[1], Zehui Liu[1], Hongjian Weng[1]*

[1]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

[2]Department of Atmospheric Science, School of Environmental Sciences, China University of Geosciences (Wuhan), Wuhan, 430074, China

[3]R&D Satellite Observations Department, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands

The Supporting Information includes 25 pages, 14 figures, and 3 tables.

Table of Contents

Figures

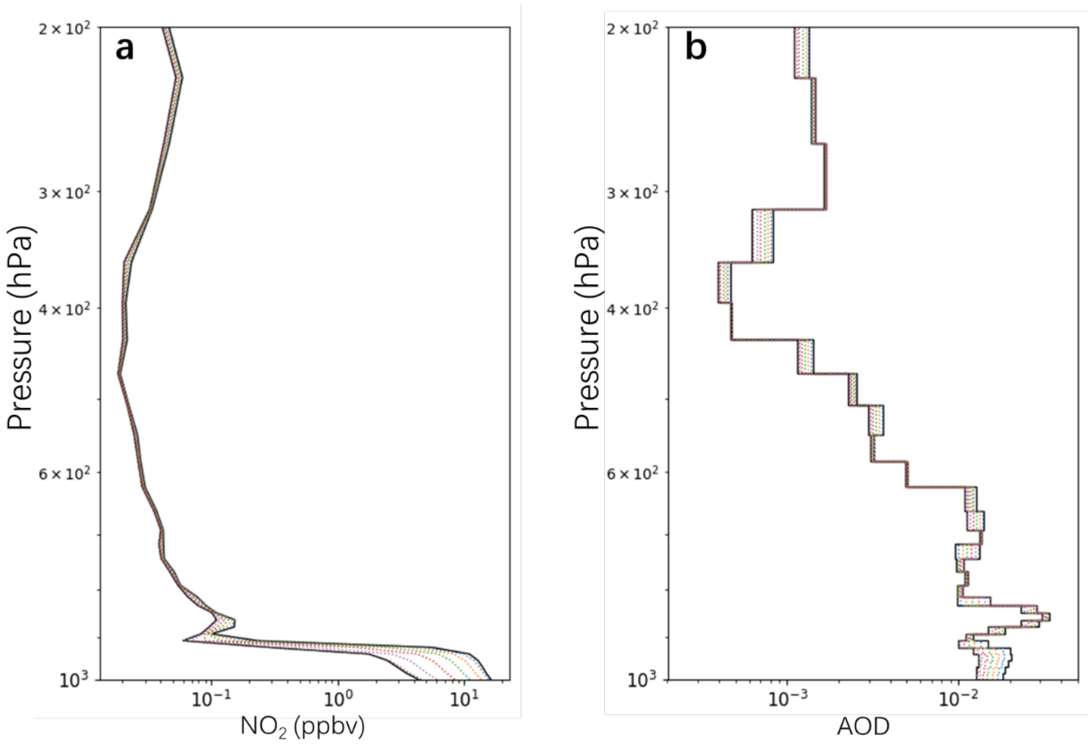


Figure S1. Vertical profiles of NO_2 (a) and sub-column AOD (b) in the troposphere. The solid black lines stand for the profiles taken from nested GEOS-Chem V9-02 simulations (as in the released version of POMINO-TROPOMI) at two adjacent grid cells centered at (121.25°E, 31°N) and (121.5625°E, 31°N), respectively. The horizontal resolution of GEOS-Chem is 0.3125° longitude × 0.25° latitude. The dotted lines in colors stand for interpolated profiles at 0.05° × 0.05° for locations in between the two adjacent GEOS-Chem grid cells.

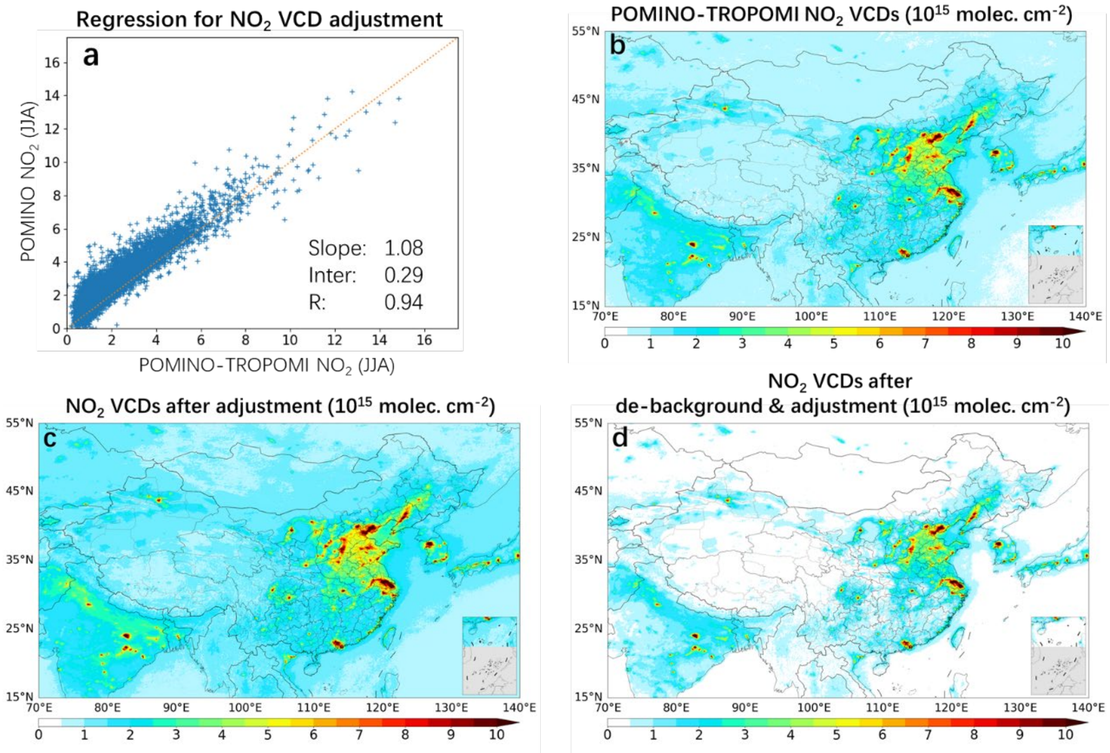


Figure S2. Regression between POMINO-TROPOMI and POMINO-OMI NO_2 VCDs sampled at 0.25° × 0.25° (a). TROPOMI NO_2 VCDs on a 0.05° × 0.05° grid re-retrieved based on interpolated a priori NO_2 vertical profiles and the POMINO algorithm (b). Similar to (b) but further adjusted based on the POMINO-OMI NO_2 (c). Similar to (c) but further de-backgrounded (d).

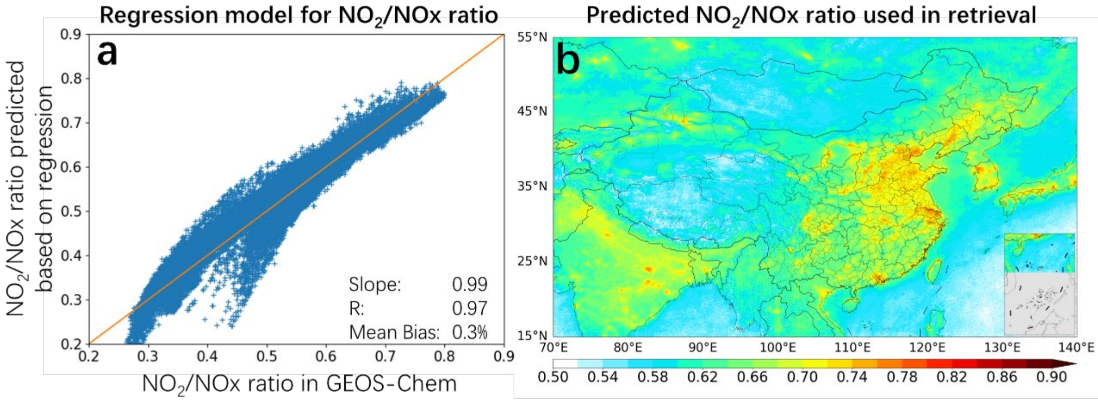


Figure S3. Scatter plot between the NO_2/NOx ratio calculated by GEOS-Chem and the ratio predicted by regression with coefficients in Table S1 (a). Spatial distribution of predicted NO_2/NOx ratio (b).

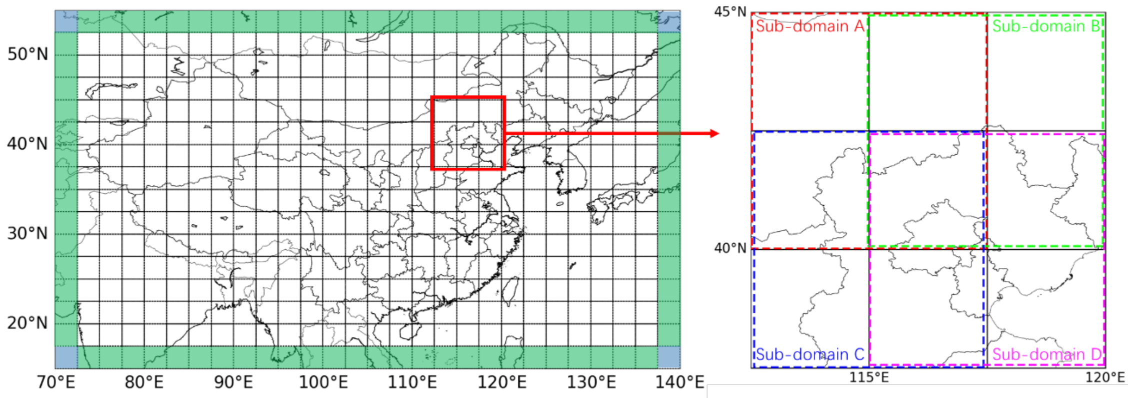


Figure S4. Illustration of how the study domain (70°-140°E, 15°-55°N) is divided into sub-domains. Each 0.05° × 0.05° grid cell is covered by 4 sub-domains inside, 2 sub-domains near the edges (denoted in green) of, or 1 sub-domain at the corners (denoted in blue) of the study domain. On the right is the four sub-domains around Beijing as an example – Sub-domains A, B, C and D are centered at (115°E, 42.5°N), (117.5°E, 42.5°N), (115°E, 40°N) and (117.5°E, 40°N), respectively.

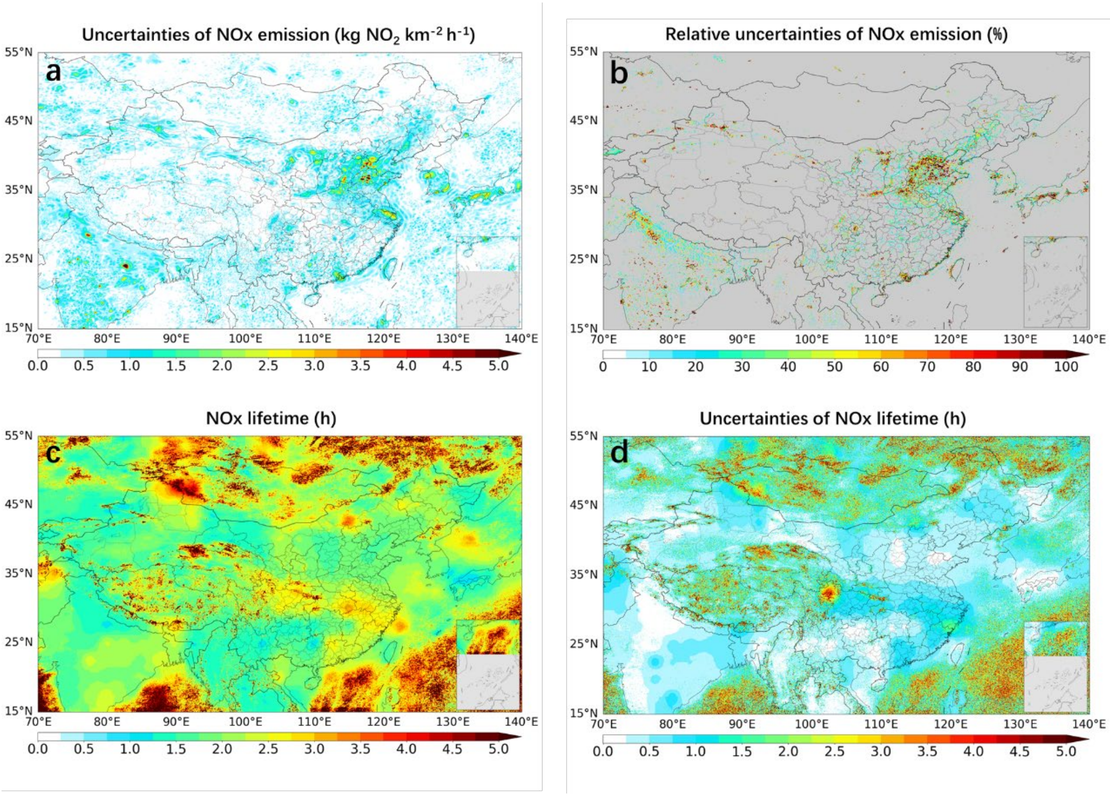


Figure S5. Uncertainty of NOx emissions (1-σ) (a), relative uncertainty of NOx emissions (1-σ) (b), NOx lifetime estimated here (c), and uncertainty of NOx lifetime (d). The grey areas in (b) stand for where NOx emissions are below 1 kg km^{-2} h^{-1}.

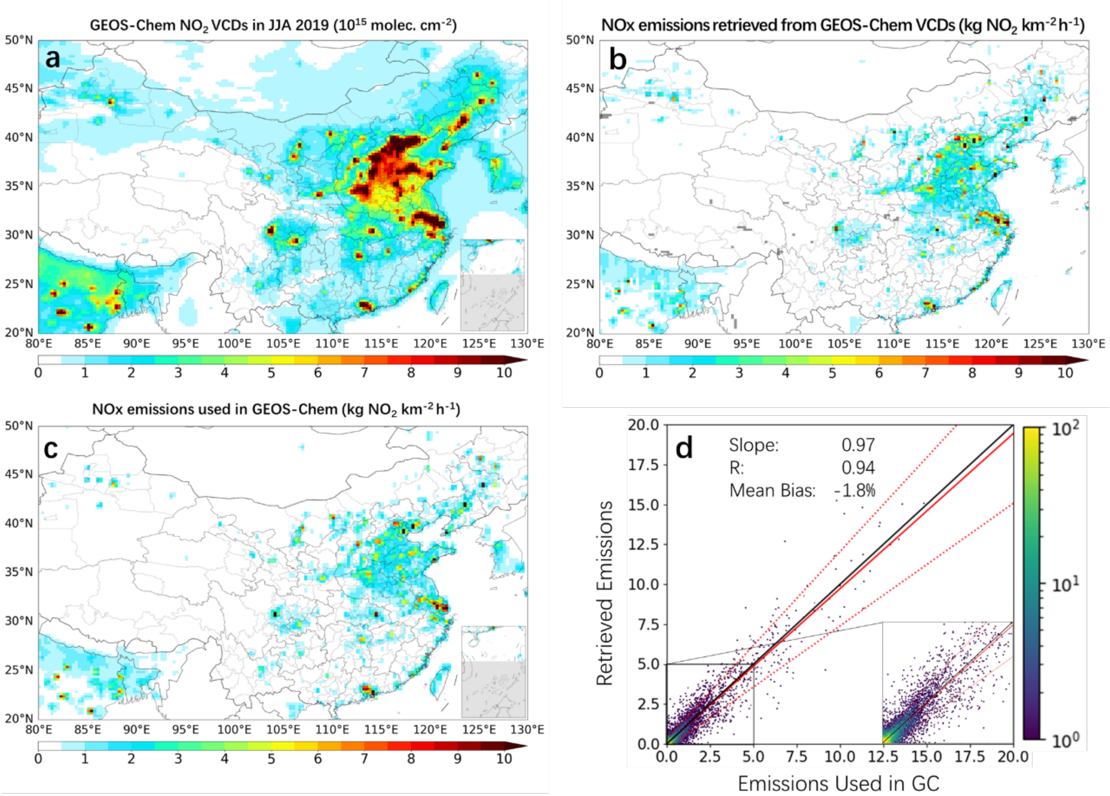


Figure S6. NO_2 VCDs simulated by GEOS-Chem (a), NOx emissions retrieved based on NO_2 VCDs simulated by GEOS-Chem (b), NOx emissions used in GEOS-Chem simulations (c), and scatter plot for (a) and (b) with colors representing data density (d).

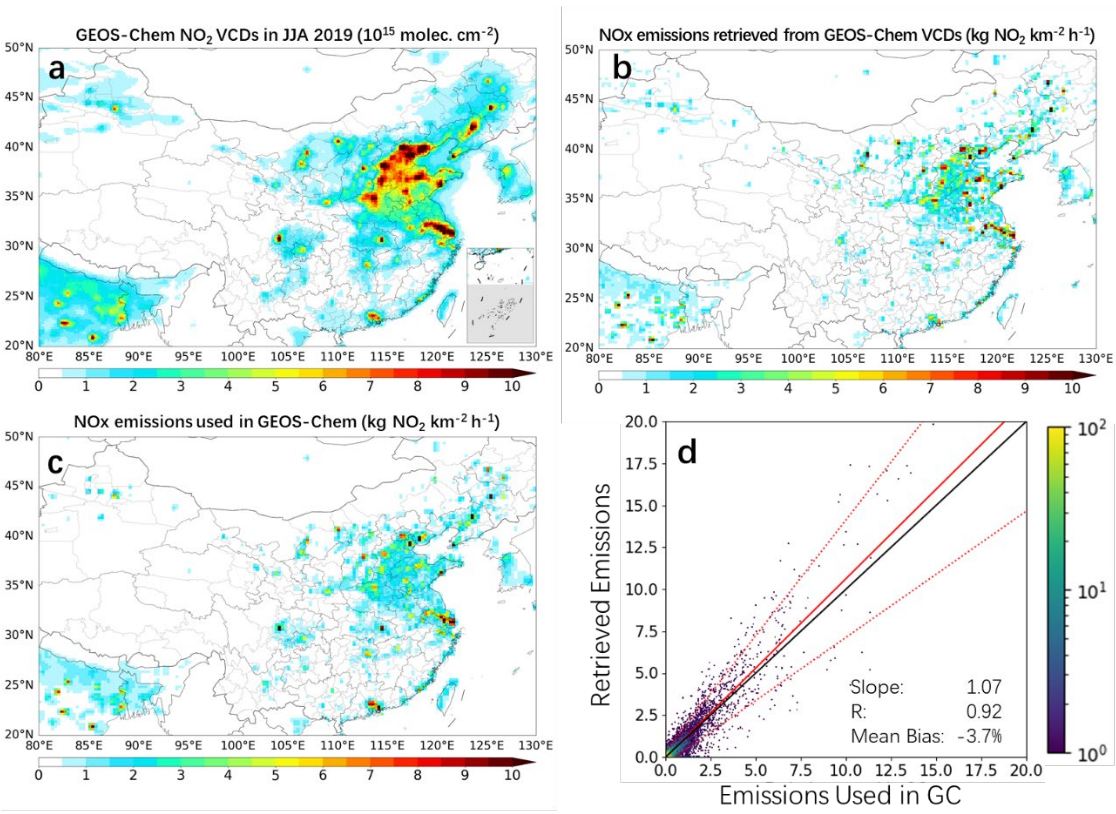


Figure S7. Similar to Fig. S6, but the GEOS-Chem NO_2 profiles are applied with the AKs from POMINO-TROPOMI.

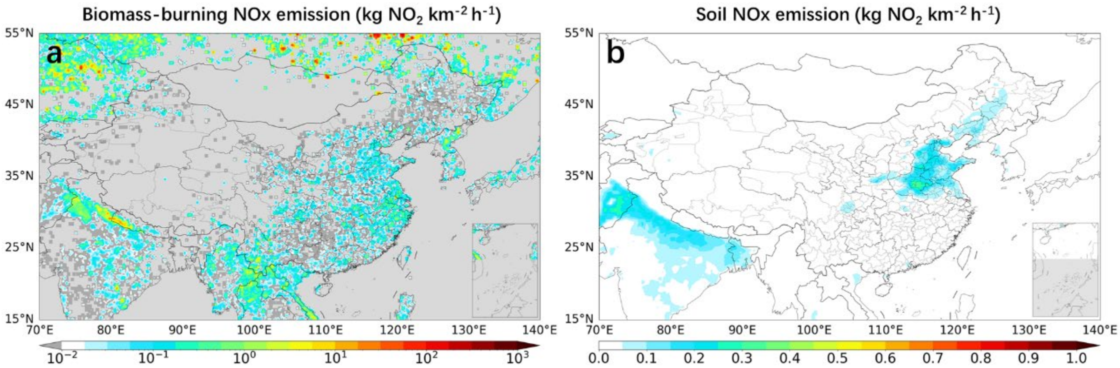


Figure S8. NOx emissions from biomass burning (a), NOx emissions from soil (b). Note the nonlinear color scales in (a).

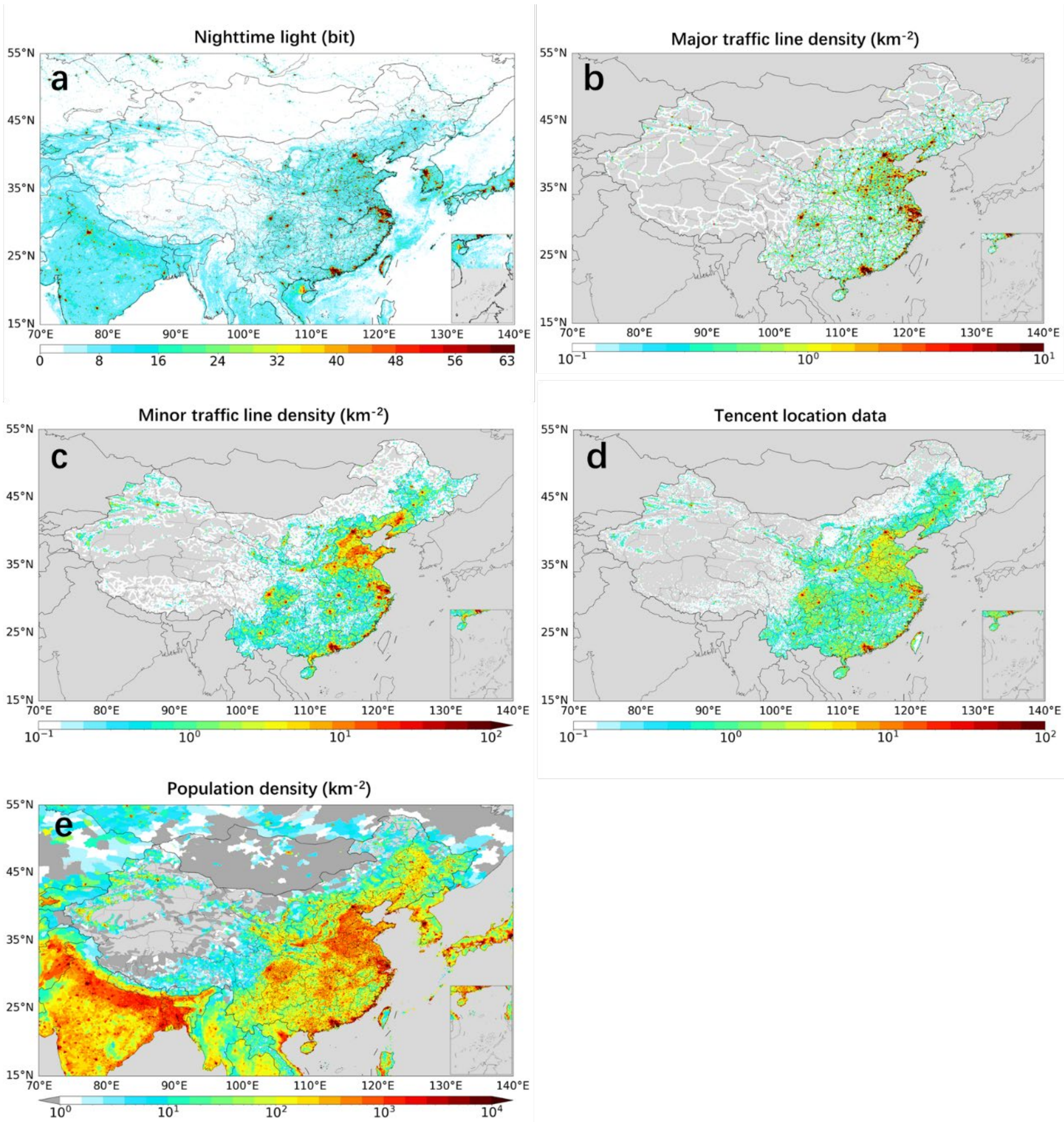


Figure S9. Proxy data at a horizontal resolution of 0.05° × 0.05°.

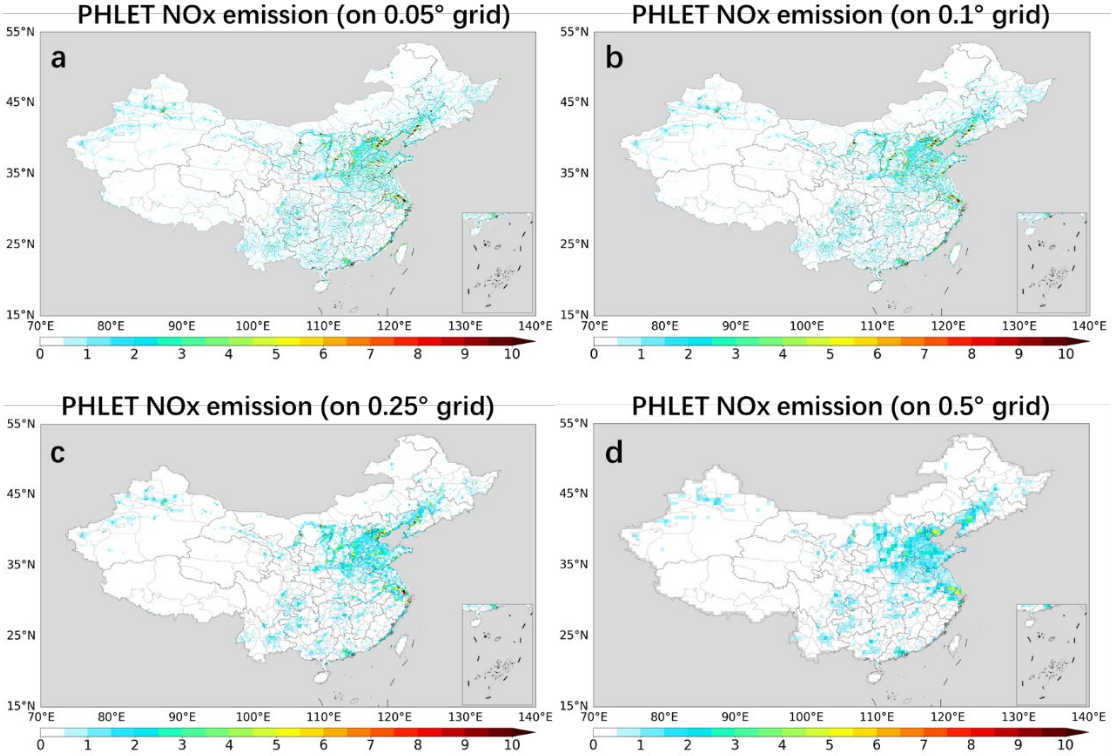


Figure S10. Horizontal distribution of NOx emissions over China from our PHLET-based emission estimate for summer 2019 at 0.05° × 0.05° (a), 0.1° × 0.1° (b), 0.25° × 0.25° (c), 0.5° × 0.5° (d).

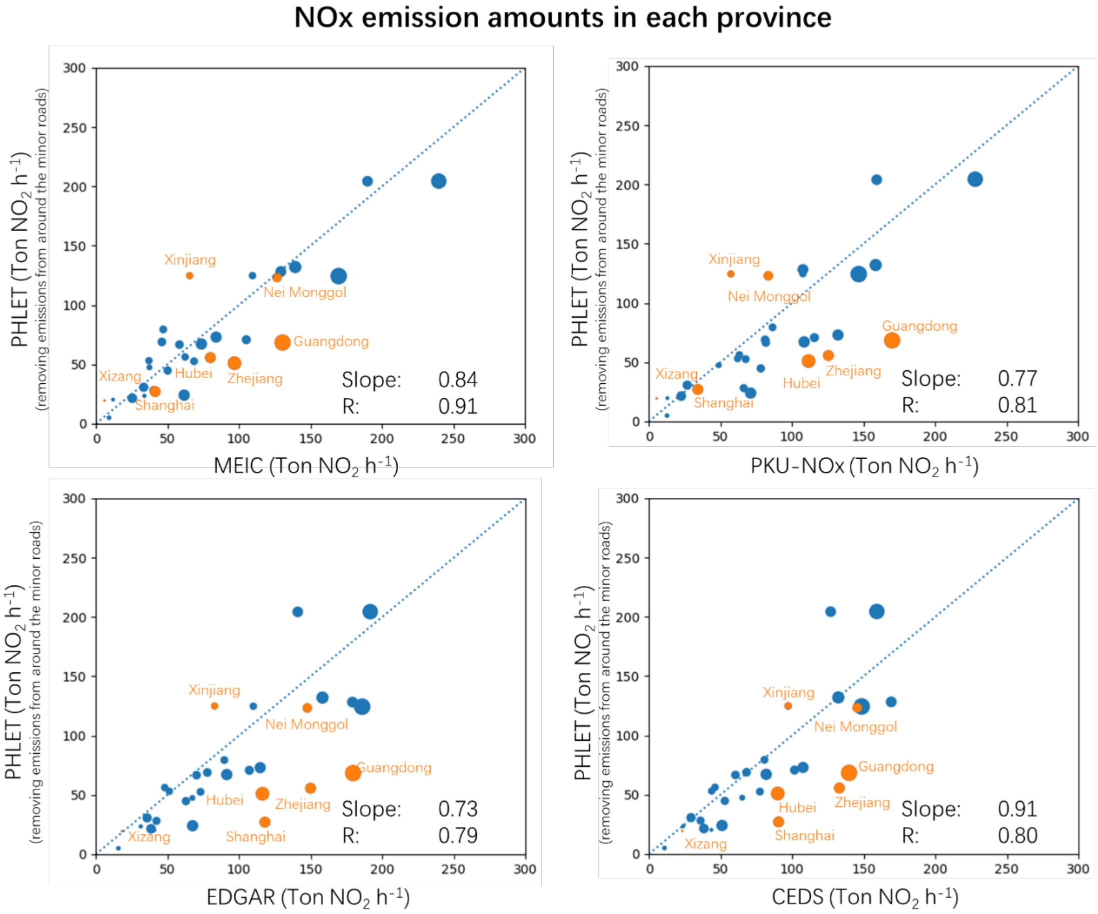


Figure S11. Similar to Fig. 3, but after removing NOx emissions at grid cells with minor roads but no major roads from the inferred emissions. The bottom-up inventories are not changed.

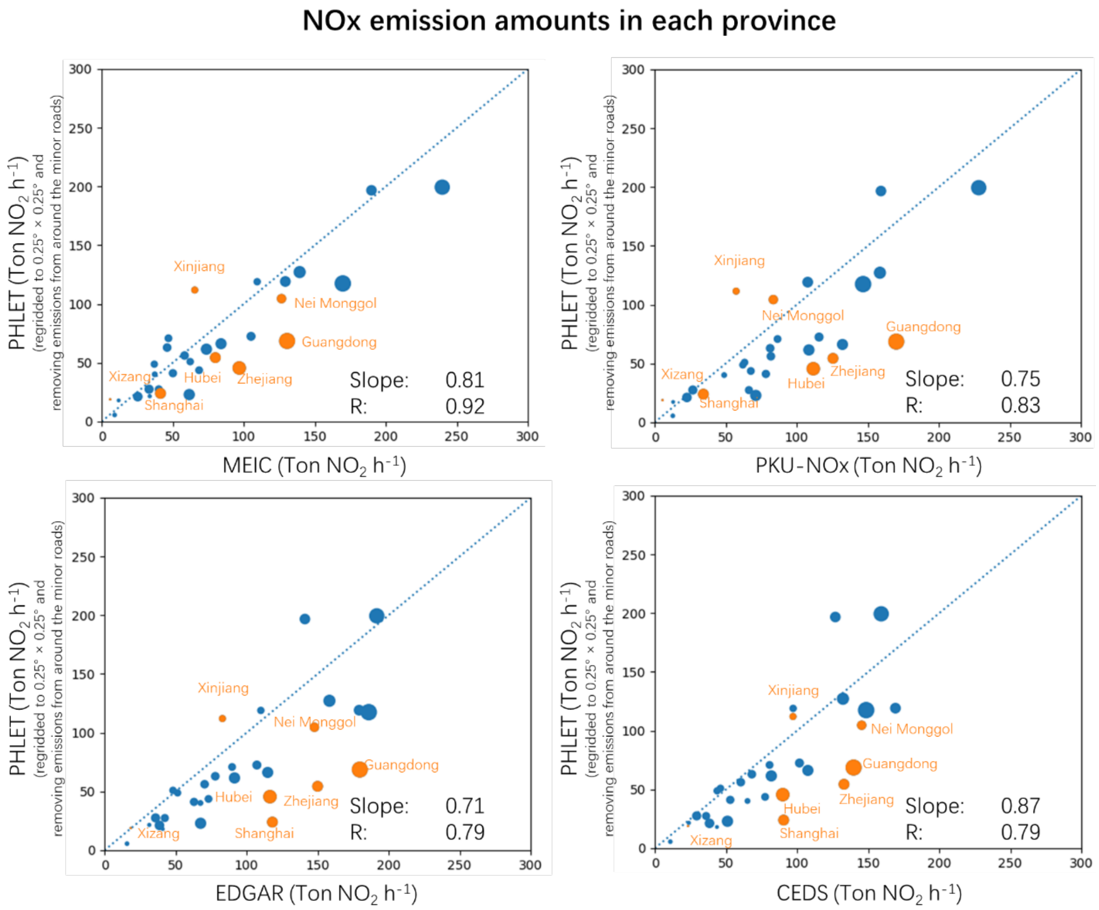


Figure S12. Similar to Fig. S11, but regridding our emissions to 0.25° × 0.25° before removing emissions at grid cells with minor roads but no major roads from the inferred emissions. The bottom-up inventories are not changed.

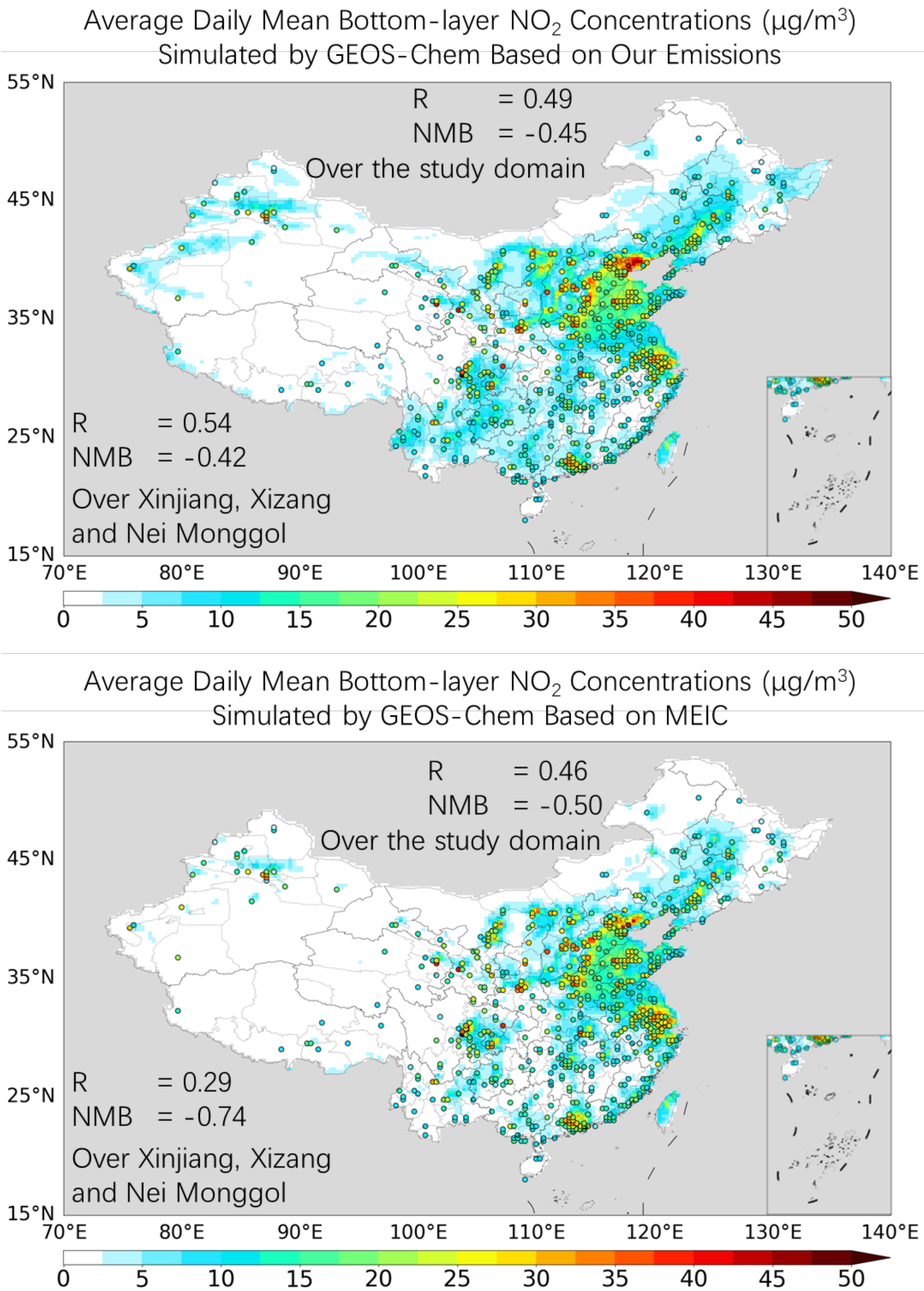
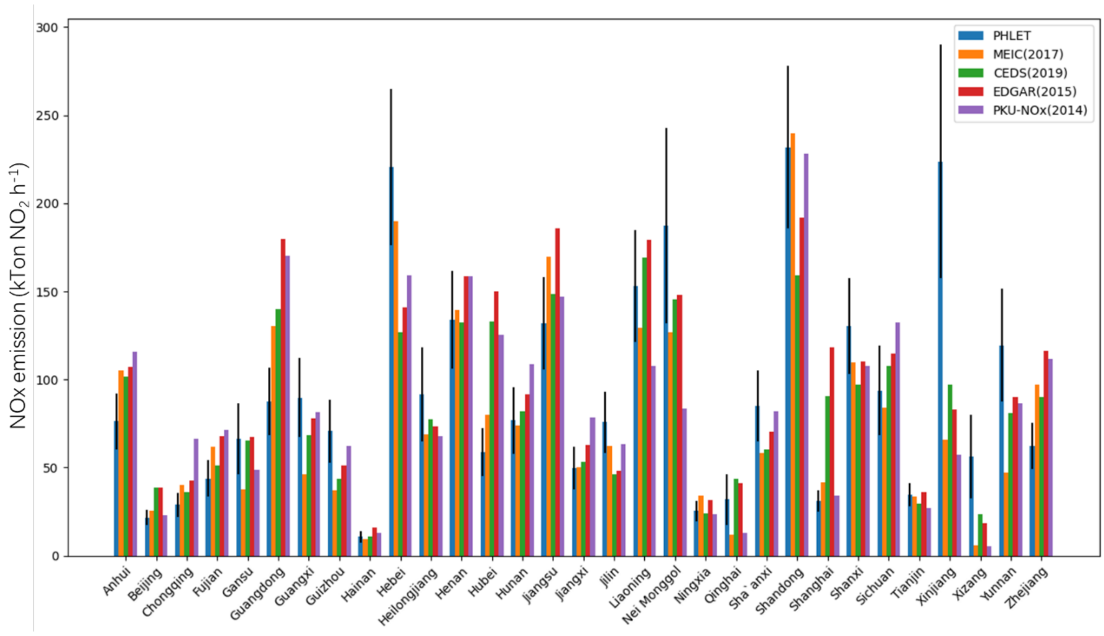


Figure S13. GEOS-Chem (v12.9.3, 0.25° × 0.3125°) simulations in July 2019, based on our emissions (top) and the MEIC inventory (bottom). The colored circles stand for MEE surface measurements.

96 

97 **Figure S14.** NOx emissions (unit: kTon NO_2 h^{-1}) in 31 provinces of mainland China.
98 Blue: our emissions (JJA 2019) with error bars (1-σ); orange: MEIC (JJA 2017); green:
99 CEDS (JJA 2019); red: EDGAR (JJA 2015); purple: PKU-NOx (JJA 2014).

Tables

Table S1. Values of coefficients of regression for the NO_2/NOx ratio (Eq. 1) derived based on least square fitting. Results with the P-value < 0.01 are marked as **.

	Best estimate	Standard deviation	Statistical significance
a (land/sea)	0.65/0.59	0.01/0.01	* * / * *
b (land/sea)	0.59/0.54	0.003/0.003	* * / * *
c	0.006	0.001	* *
d	0.0025	0.000	* *
e	0.038	0.001	* *
f (land/sea)	-19.4/-17.8	0.1/0.1	* * / * *

Table S2. Uncertainty settings (1-σ) in derivation of NOx emissions in this study and our previous work[1].

Error source	Kong et al. (2019)	This study	Notes
Errors affecting the derivation of LNS; all error sources are added in quadrature to construct the diagonal of error covariance matrix in the cost function of PHLET Adjoint.			
Satellite NO_2 VCD data	30% + 1.9×10^{15} molec. cm^{-2} for each pixel, and further reduced by a factor of $s=\sqrt{\frac{1-c}{n}+c}$ by sampling over pixels in multiple days	20% + 0.5×10^{15} molec. cm^{-2} for each pixel, and further reduced by a factor of $s=\sqrt{\frac{1-c}{n}+c}$ by sampling over pixels in multiple days	Satellite data quality is improved. c = 0.5 and n is the number of pixels with valid data during summer 2019.
Removal of NO_2 background	5% of NO_2 VCDs	0.3×10^{15} molec. cm^{-2}	Smoother background
Horizontal resolution of satellite NO_2 VCDs being lower than that of derived emissions	50% of standard deviation of surrounding grid cells	Removed	TROPOMI has a much higher horizontal resolution than OMI
Assumption of stable NO_2 (emission equals loss) at the satellite overpass time	15% of NO_2 VCDs	15% of NO_2 VCDs	
Long-term average for emission inversion	10% of NO_2 VCDs	10% of NO_2 VCDs	
2-dimesional simplification of chemistry-transport in PHLET	15% of NO_2 VCDs	15% of NO_2 VCDs	
Derivation of effective diffusion coefficients from wind field	20% of NO_2 VCDs	20% of NO_2 VCDs	

NO_2/NOx ratio	15% of NO_2 VCDs	5% of NO_2 VCDs	The value is fixed in Kong et al., but is derived here for each grid cell by regression.
Errors affecting the calculation of emissions and lifetime parameters but not LNS			
Derivation of NOx lifetime parameters	Standard deviation within multiple estimates by changing the LNS quantiles.	Root mean square difference between the emissions derived with 1%-quantile of LNS (best estimate) and those with 0.1% (upper estimate) and 2% quantiles (lower estimate).	

Table S3. Correlation coefficients between spatial proxies and NOx emissions (including our emissions denoted as PHLET, our emissions regridded on 0.25° × 0.25° grid and 0.1° × 0.1° grid, MEIC, PKU-NOx and EDGAR). In each row, the upper values are calculated for mainland China, and the lower values for three western provinces (Xinjiang, Xizang and Nei Monggol) together. Results with the P-value < 0.01 are marked as **. As emission data can be at different resolutions, they are always mapped to 0.05° × 0.05° before correlation coefficients are calculated. The resolution of CEDS is too low (0.5°) to allow a meaningful correlation analysis.

	PHLET (0.05°)	PHLET (0.1°)	PHLET (0.25°)	MEIC (0.25°)	PKU-NOx (0.1°)	EDGAR (0.1°)
Density of major road lines	0.51** 0.43**	0.53** 0.45**	0.55** 0.41**	0.52** 0.40**	0.28** 0.42**	0.17** 0.28**
Density of minor road lines	0.48** 0.46**	0.51** 0.48**	0.57** 0.45**	0.48** 0.35**	0.25** 0.37**	0.14** 0.23**
Population density	0.49** 0.41**	0.51** 0.44**	0.56** 0.45**	0.53** 0.47**	0.32** 0.49**	0.18** 0.27**
Tencent location data	0.54** 0.50**	0.55** 0.50**	0.55** 0.44**	0.52** 0.41**	0.32** 0.45**	0.18** 0.28**
Nighttime light	0.53** 0.47**	0.54** 0.50**	0.55** 0.47**	0.57** 0.44**	0.36** 0.47**	0.21** 0.29**

1. GEOS-Chem simulations

We use the nested GEO-Chem v12.9.3 to simulate NO_2 VCDs over East China (70°-140°E, 15°-55°N) at 0.3125° longitude × 0.25° latitude with 47 vertical layers, for purposes of establishing the regression model for the NO_2/NOx ratio and evaluating our emission inversion approach. The nested model is driven by the GEOS-FP assimilated meteorology in 2019 from the NASA Global Modeling and Assimilation Office and emissions in 2017 (the latest year) from MEIC. The model is run in summer (June, July and August) 2019 with the full Ox-NOx-VOC-CO-HOx gaseous chemistry and online aerosols. Model convection follows the relaxed Arakawa–Schubert scheme[2]. Vertical mixing in the planetary boundary layer employs a non-local scheme implemented by Lin et al.[3] Dry deposition follows Wesely[4], with a number of modifications, for gases[5] and Zhang et al. for aerosols[6]. Model results in a smaller domain (80°-130°N, 20°-50°E) are used here to exclude artificial noise near the lateral boundaries.

2. Proxy data of human activity

We use five gridded proxy datasets of human activity to help evaluate the derived emissions, including population density, Tencent location data, nighttime light, road line density and GPD.

We take population density data on a 2.5′ × 2.5′ grid from the Gridded Population of the World v4 database (GPWv4; https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/sets/browse; last access: 19 August 2018)[7]. GPW is a widely-used proxy in bottom-up inventories to allocate spatially aggregated emissions to individual locations[8,9]. In China, the database is based on township-level data from a census in 2010 and extrapolated to other years (2000, 2005, 2010, 2015 and 2020). Here we use the data in 2015 (Fig. S8a) since the predicted data for 2020 do not account for the effect of COVID-19. Given the relatively old base year (2010) and the large size of towns in western China, the database may contain significant errors in inferring anthropogenic activity in 2019 especially in the west.

For an up-to-date proxy, we choose Tencent user location data (https://xingyun.map.qq.com/). Tencent is a private Internet company in China providing instant messaging (text, audio and video) services with the dominant market share[10,11]. The company publishes a real-time map of the number distribution of its more than 1.2 billion users at a horizontal resolution of 0.05° × 0.05° and a temporal resolution of every 0.5 second, based on its users' location requests via WeChat and other APPs. We grip the location data from 30 August 2021 to 31 September 2021 and calculate the temporal average (Fig. S8e), since earlier data cannot be retrieved. The Tencent location data not only show intensive human activity in major city clusters, as in the GPWv4 population data, but also reveal up-to-date and subtle details of human activity especially in the western provinces, e.g., around the road lines and over the Tibetan Plateau.

We adopt nighttime light data in 2018 from version 4 of the Defense

Meteorological Satellite Program – Operational Linescan System (DMSP-OLS, last access: 31 January 2021)[12,13] in our emission filtering and analysis. Nighttime light is another widely used proxy in bottom-up inventories to spatially allocate aggregated emissions[8]. Here the data are regridded from its original resolution of 0.5′ × 0.5′ to 0.05° × 0.05° (Fig. S8b).

We take road net data for 2020 from AMAP (https://www.amap.com/; last access: 26 August 2021)[14,15], and calculate the road line densities of major and minor roads (Fig. S8c, d). The major roads are referred to here as expressways, national highways, provincial highways and first-grade roads which allow speed limits larger than 60 km h^{-1}. The minor roads include county highways, township highways and second-grade roads. For each class of roads, we calculate the number of road lines in each grid cell and refer it as road line density[14].

We also adopt gridded GDP data in 2015 at a resolution of 1 km from the Resource and Environmental Science and Data Center of the Chinese Academy of Sciences (https://www.resdc.cn/; last access: 13 August 2021)[16,17]. We calculate county-level GDP from the gridded data.

References

1. Kong, H.; Lin, J.; Zhang, R.; Liu, M.; Weng, H.; Ni, R.; Chen, L.; Jingxu, W.; Yan, Y.; Zhang, Q., High-resolution (0.05° × 0.05°) NO x emissions in the Yangtze River Delta inferred from OMI. *Atmospheric Chemistry and Physics* **2019,** *19*, 12835-12856.
2. 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.; Nielsen, J. E., *The GEOS-5 Data Assimilation System: Documentation of Versions 5.0.1, 5.1.0, and 5.2.0*. National Aeronautics and Space Administration, Goddard Space Flight Center: 2008.
3. Lin, J.-T.; McElroy, M. B., Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing. *Atmospheric Environment* **2010,** *44* (14), 1726-1739.
4. Wesely, M. L., Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmospheric Environment (1967)* **1989,** *23* (6), 1293-1304.
5. Wang, Y.; Jacob, D. J.; Logan, J. A., Global simulation of tropospheric O3-NOx-hydrocarbon chemistry: 3. Origin of tropospheric ozone and effects of nonmethane hydrocarbons. *Journal of Geophysical Research: Atmospheres* **1998,** *103* (D9), 10757-10767.
6. Zhang, L.; Gong, S.; Padro, J.; Barrie, L., A size-segregated particle dry deposition scheme for an atmospheric aerosol model. *Atmospheric Environment* **2001,** *35*, 549-560.
7. Center for International Earth Science Information Network - CIESIN - Columbia University, Gridded Population of the World, Version 4 (GPWv4): Population Count. NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, 2016.
8. Geng, G.; Zhang, Q.; Martin, R. V.; Lin, J.; Huo, H.; Zheng, B.; Wang, S.; He, K., Impact of spatial proxies on the representation of bottom-up emission inventories: A satellite-based analysis. *Atmospheric Chemistry and Physics* **2017,** *17* (6), 4131-4145.
9. Kurokawa, J.; Ohara, T.; Morikawa, T.; Hanayama, S.; Janssens-Maenhout, G.; Fukui, T.; Kawashima, K.; 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.* **2013,** *13* (21), 11019-11058.
10. Chen, B.; Song, Y.; Kwan, M.-P.; Huang, B.; Xu, B., How do people in different places experience different levels of air pollution? Using worldwide Chinese as a lens. *Environmental Pollution* **2018,** *238*, 874-883.
11. Tencent Tencent location big data. https://xingyun.map.qq.com/.
12. Elvidge, C. D.; Baugh, K. E.; Dietz, J. B.; Bland, T.; Sutton, P. C.; Kroehl, H. W., Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment* **1999,** *68(1)*, 77-88.
13. Hsu, F. C.; Baugh, K. E.; Ghosh, T.; Zhizhin, M.; Elvidge, C. D., DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Remote Sensing* **2015,** *7(2)*, 1855-1876.
14. Feng, X.; Xiu, C.; Bai, L.; Wen, Y., Urban Centrality and Influencing Factors in Jilin Province from the Perspective of Highway Traffic Flow. *Economic Geography* **2019,** *39* (1), 64-72.
15. Oda, T.; Maksyutov, S., A very high-resolution (1 km×1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.* **2011,** *11* (2), 543-556.

16. Ling, Y. I.; Xiong, L. Y.; Yang, X. H., Method of Pixelizing GDP Data Based on the GIS. *Journal of Gansu Sciences* **2006,** *18* (2), 54-58.
17. Liu, H.; Jiang, D.; Yang, X.; Luo, C., Spatialization Approach to 1km Grid GDP Supported by Remote Sensing. *Geo-information Science* **2005,** *7* (2), 120-123.