Considerable Unaccounted Local Sources of NO\textsubscript{x} Emissions in China Revealed from Satellite

Hao Kong, Jintai Lin,\textsuperscript{*} Lulu Chen, Yuhang Zhang, Yingying Yan, Mengyao Liu, Ruijing Ni, Zehui Liu, and Hongjian Weng

ABSTRACT: High-resolution (e.g., 5 km) emission data of nitrogen oxides (NO\textsubscript{x} = NO + NO\textsubscript{2}) provide localized knowledge of pollution sources for targeted regulations, yet such data are lacking or inaccurate over most regions at present. Here we improve our PHLET-based inversion method to derive NO\textsubscript{x} emissions in China at a 5-km resolution in summer 2019, based on the TROPOMI-POMINO satellite product of nitrogen dioxide (NO\textsubscript{2}) columns. With low computational costs, our inversion explicitly accounts for the effects of horizontal transport and nonlinear chemistry. We find numerous small-to-medium sources related to minor roads and small human settlements at relatively low affinity levels, in addition to clear emission signals along major transportation lines, consistent with road line density and Tencent location data. Many small-to-medium sources and transportation emissions are unclear or missing in the spatial distributions of four widely used emission inventories. Our emissions offer a unique reference for targeted emission control.

KEYWORDS: NO\textsubscript{x} emission, high resolution, satellite, China, TROPOMI

INTRODUCTION

Tropospheric nitrogen oxides (NO\textsubscript{x} = NO + NO\textsubscript{2}) are a major short-lived air pollutant. At present, its emissions are dominated by anthropogenic sources around human settlements of various sizes.\textsuperscript{1,2} Since the anthropogenic emissions vary dramatically with intricate human activities from place to place, high-resolution (i.e., 5 km) emission data are critical for understanding the environmental effects of local human activities. Such data provide a fundamental basis for high-resolution (e.g., city- or county-level) air quality modeling and assessment and targeted emission control. However, up-to-date high-resolution NO\textsubscript{x} emissions data are still lacking over most regions of the world.

Bottom-up inventories are widely used to estimate NO\textsubscript{x} emissions, by combining regional (e.g., provincial) activity data, emission factors, and spatial proxies to downscale emissions to individual locations. These inventories are subject to large uncertainties in emission factors and activity data.\textsuperscript{3} Although much effort has been made to locate large point sources in China in a latest 1-km resolution inventory for 2013,\textsuperscript{4} about 50% of the total NO\textsubscript{x} emissions remain relying on proxy-based downscaling. At fine spatial scales, the proxy-based downscaling approach cannot work well, because emissions are related to these proxies in complex ways. Different choices of proxy can lead to large differences in the spatial distribution of NO\textsubscript{x} emissions.\textsuperscript{5} Studies have shown that common proxies such as population density tend to be decoupled with NO\textsubscript{x} emissions at high resolutions, since large emission sources tend to be away from populous city centers.\textsuperscript{6} Also, up-to-date high-resolution proxies can be difficult or impossible to obtain; for example, gridded population density data in China are developed based on the census in 2010.\textsuperscript{5,7}

Over the past two decades, top-down estimates of NO\textsubscript{x} emissions based on satellite retrievals of NO\textsubscript{2} vertical column densities (VCDs) have become popular. Traditional top-down approaches assuming local mass balance (LMB) without explicit considerations of horizontal transport and nonlinear NO\textsubscript{x} chemistry cannot work properly at high resolutions as the effect of horizontal transport becomes significant.\textsuperscript{8–13} More sophisticated methods such as Kalman Filter and Adjoint use iterative simulations of 3-dimensional chemical transport models (CTMs) to account for pollution transport and nonlinear chemistry.\textsuperscript{14–16} These methods require expensive computational resources, and are thus difficult to apply to large domains (e.g., East Asia) at high-resolutions (e.g., 5 km).

Special Issue: Urban Air Pollution and Human Health

Received: November 13, 2021
Revised: March 1, 2022
Accepted: March 4, 2022
A few low-cost inversion methods have been developed recently,17–21 often for the purpose of estimating individual point sources.17–19,21 The exponentially modified Gaussian (EMG) model is used to isolate point sources and requires a large number of samples.17 A regression model based on NO2 column—wind speed relation is developed for a particular topography,21 and it does not account for the nonlinear dependence of chemical lifetime on NO2 abundance. Another development is the divergence-based model that is dependent on prescribed, fixed NO2 lifetime and CTM-based NO2/NOx ratios.18,19 Our previous work20 develops another fast inversion approach based on a simplified 2-dimensional PHLET (Peking University High-Resolution Lifetime-Emission-Transport) model that takes into account horizontal transport and the nonlinear lifetime-abundance relationship; yet in that study, the method is only applied to a small domain (500 km of the Yangtze River Delta in China). In this study, we improve our PHLET-based emission inversion approach based on a regression model established by fitting GEOS-Chem NO2 simulations, air temperature, wind speed, and leaf area index (LAI) data. Third, regional background of NO2 VCDs is estimated and subtracted from the adjusted TROPOMI NO2 data, in order to better reveal the local hot spots of NO2. Fourth, gridded NOx emissions are derived by applying PHLET and its adjoint model to the adjusted, debackgrounded NO2 VCDs. Finally, to extract the anthropogenic sources, we filter the derived emission data using spatial proxies of human activity.

**Tropospheric NO2 VCDs Retrieved from TROPOMI.** The TROPOMI sensor is on board the sun-synchronous Copernicus Sentinel-5 Precursor satellite. With a swath of 2600 km, TROPOMI conducts space-borne remote sensing of NO2 globally on a daily basis, with an equator crossing time at 13:30 local time.22 It has an unprecedentedly high horizontal resolution, with a pixel size of 3.5 × 7 km2 at nadir (3.5 × 5.5 km2 since August 2019).23 We use the tropospheric NO2 VCD data in summer (June–August) 2019 from the POMINO-TROPOMI product.24 Compared to the KNMI official product (TMS-MP-DOMINO OFFLINE v1.1.0), POMINO-TROPOMI explicitly accounts for the effects of aerosols and surface reflectance anisotropy on NO2 retrieval, and adopts a priori NO2 vertical profiles at a much higher horizontal resolution (25 km versus 100 km). As a result, POMINO shows a smaller bias of NO2 VCDs than the official product relative to the MAX-DOAS measurements,24 and has more valid data especially in polluted situations with heavy aerosol and NO2 loadings (e.g., 60% more pixels with NO2 VCD > 5 × 1015 molecules cm−2).

Figure 1 shows the methodological flowchart. First, NO2 VCDs from TROPOMI are adjusted according to OMI NO2 data to reduce the potential systematic bias. Second, the adjusted TROPOMI NO2 VCDs are used to estimate gridded NO2/NOx ratios, based on a regression model established by fitting GEOS-Chem NO2 simulations, air temperature, wind speed, and leaf area index (LAI) data. Third, regional background of NO2 VCDs is estimated and subtracted from the adjusted TROPOMI NO2 data, in order to better reveal the local hot spots of NO2. Fourth, gridded NOx emissions are derived by applying PHLET and its adjoint model to the adjusted, debackgrounded NO2 VCDs. Finally, to extract the anthropogenic sources, we filter the derived emission data using spatial proxies of human activity.

**Materials and Methods**

Figure 1 shows the methodological flowchart. First, NO2 VCDs from TROPOMI are adjusted according to OMI NO2 data to reduce the potential systematic bias. Second, the adjusted TROPOMI NO2 VCDs are used to estimate gridded NO2/NOx ratios, based on a regression model established by fitting GEOS-Chem NO2 simulations, air temperature, wind speed, and leaf area index (LAI) data. Third, regional background of NO2 VCDs is estimated and subtracted from the adjusted TROPOMI NO2 data, in order to better reveal the local hot spots of NO2. Fourth, gridded NOx emissions are derived by applying PHLET and its adjoint model to the adjusted, debackgrounded NO2 VCDs. Finally, to extract the anthropogenic sources, we filter the derived emission data using spatial proxies of human activity.

**Figure 1.** Methodological flowchart of our emission retrieval process. Details of PHLET-based emission inversion in each subdomain and subsequent filtering to obtain “anthropogenic” emissions are shown in the bottom boxes. The orange arrows show the iterative process to retrieve the LNS (Local Net Source, emission minus loss) on a 0.25° × 0.25° grid and determine the lifetime parameters in each subdomain. The blue arrows show the iterative process to retrieve the LNS and NOx emissions on a 0.05° × 0.05° grid.
interpolate the NO\textsubscript{2} and aerosol profiles to a 5 km-resolution grid (Supporting Information (SI) Figure S1).

The official KNMI TROPMI NO\textsubscript{2} product is significantly underestimated (25%–50%) globally.\textsuperscript{25–26} Although POMINO-TROPMI reduces such underestimation, we choose to conduct further correction with the help of POMINO-OMI v2.0.1 NO\textsubscript{2} data.\textsuperscript{27–29} According to validation based on MAX-DOAS measurements, POMINO-OMI NO\textsubscript{2} data contain small biases (~3\%);\textsuperscript{28–30} and OMI NO\textsubscript{2} are often used to help evaluate TROPMI data.\textsuperscript{31} Specifically, we sample the OMI and TROPMI NO\textsubscript{2} VCD data for summer 2019 within the study domain at a horizontal resolution of 0.25° × 0.25° (given that the pixel size of OMI is 13 × 24 km\textsuperscript{2} at nadir), regress the two data sets to obtain the slope (1.08) and intercept (0.29 × 10\textsuperscript{15} molecules cm\textsuperscript{-2}) (R = 0.94, SI Figure S2a), and then use the slope and intercept to adjust TROPMI data on a 0.05° × 0.05° grid. To obtain valid data from both sensors, we remove the pixels with cloud radiance fraction exceeding 50%, with the consideration that NO\textsubscript{2} over clean sea.

We map the pixel-by-pixel (i.e., Level-2) POMINO-TROPMI NO\textsubscript{2} data to a 0.05° × 0.05° grid. Following Kong et al. (2019),\textsuperscript{32} we use a special oversampling approach to properly assign each irregularly shaped pixel to respective grid cells. We also calculate a satellite conversion matrix (SCM) to facilitate subsequent conversion from PHLET-grid cells. We also calculate a satellite conversion matrix (SCM) before and after the adjustment. The mean value of NO\textsubscript{2} VCDs is increased by 31.2% by the adjustment, and thus the underestimate is reduced.

We map the pixel-by-pixel (i.e., Level-2) POMINO-TROPMI NO\textsubscript{2} data to a 0.05° × 0.05° grid. Following Kong et al. (2019),\textsuperscript{32} we use a special oversampling approach to properly assign each irregularly shaped pixel to respective grid cells. We also calculate a satellite conversion matrix (SCM) before and after the adjustment. The mean value of NO\textsubscript{2} VCDs is increased by 31.2% by the adjustment, and thus the underestimate is reduced.

We map the pixel-by-pixel (i.e., Level-2) POMINO-TROPMI NO\textsubscript{2} data to a 0.05° × 0.05° grid. Following Kong et al. (2019),\textsuperscript{32} we use a special oversampling approach to properly assign each irregularly shaped pixel to respective grid cells. We also calculate a satellite conversion matrix (SCM) before and after the adjustment. The mean value of NO\textsubscript{2} VCDs is increased by 31.2% by the adjustment, and thus the underestimate is reduced.

We map the pixel-by-pixel (i.e., Level-2) POMINO-TROPMI NO\textsubscript{2} data to a 0.05° × 0.05° grid. Following Kong et al. (2019),\textsuperscript{32} we use a special oversampling approach to properly assign each irregularly shaped pixel to respective grid cells. We also calculate a satellite conversion matrix (SCM) before and after the adjustment. The mean value of NO\textsubscript{2} VCDs is increased by 31.2% by the adjustment, and thus the underestimate is reduced.

We map the pixel-by-pixel (i.e., Level-2) POMINO-TROPMI NO\textsubscript{2} data to a 0.05° × 0.05° grid. Following Kong et al. (2019),\textsuperscript{32} we use a special oversampling approach to properly assign each irregularly shaped pixel to respective grid cells. We also calculate a satellite conversion matrix (SCM) before and after the adjustment. The mean value of NO\textsubscript{2} VCDs is increased by 31.2% by the adjustment, and thus the underestimate is reduced.
Figure 2. Horizontal distribution of anthropogenic NO\(_x\) emissions over China from our PHLET-based emission estimate for summer 2019 at 0.05° × 0.05° (a), MEIC for summer 2017 at 0.25° × 0.25° (b), PKU-NO\(_x\) for summer 2014 at 0.1° × 0.1° (c), EDGAR for summer 2015 at 0.1° × 0.1° (d), and CEDS for summer 2019 at 0.5° × 0.5° (e). Three small areas are enlarged to highlight the local signals (I: Beijing; II: Hetao Plain; III: west Xinjiang).
of the distance between the grid cell center and each subdomain center.

**Emission Inversion in Each Subdomain.** Within each 5° × 5° subdomain, we improve the PHLET-based inversion method to quantify NOx emissions (Figure 1). The model is based on the following equations that account for emissions, nonlinear chemistry, horizontal flux, and "effective" diffusion of NOx, as detailed in Kong et al. (2019):

\[
\frac{\partial C(x,y)}{\partial t} = r(x,y)\cdot L(x,y) - \nabla \cdot (V(x,y) \cdot C(x,y)) + \nabla \cdot (\alpha K(x,y) \cdot V C(x,y)) = 0
\] (2)

\[
L(x,y) = E(x,y) - \frac{C(x,y)}{r(x,y) \cdot r(x,y)}
\] (3)

\[
-\tilde{a}'(r(x,y) \cdot L(x,y) + kC(x,y)) - b' \left( \frac{r(x,y) \cdot L(x,y) + kC(x,y)}{C(x,y)} \right)^2 = 1, \text{ where } E(x,y) \approx 0
\] (4)

\[
\frac{1}{\tau} = \frac{1}{\tau_c} + k
\] (5)

Here, eq 2 presents the governing equation for NOx; eq 3 describes the nonlinear relationship between lifetime and VCD of NOx; and eq 5 defines the lifetime. C refers to the NOx VCD, L the LNS, E the emission, r the NOx/NO2 ratio, \( r \) the lifetime, \( r_c \) the VCD-dependent chemical lifetime, and \( k \) the rate constant due to deposition. The values of \( k, a', b', \) and \( c' \) are to be fitted. V denotes the mean wind vector over the study time period (summer 2019), and \( K \) denotes the effective diffusion coefficient tensor to account for horizontal transport by the residual winds deviating from the temporally average wind field. V and \( K \) are derived from the GEOS-FP 3-hourly wind field data near the surface (0–500 m above the ground, as a common practice [17,19,20,97–99]) at a horizontal resolution of 0.3125° longitude × 0.25° latitude. As an improvement upon Kong et al. (2019), we add a scaling factor (\( \alpha = 0.5 \)) to adjust the values of diffusion coefficients, after testing different values from 0.1 to 1.0. A test applying the emission inversion procedure to GEOS-Chem simulated NOx VCDs shows that unscaled coefficients would lead to an overestimate in the horizontal transport of NOx and thus in the horizontal gradient of derived NOx emissions.

Based on the PHLET model, its adjoint and the adjusted, debackgrounded POMINO-TROPMI NO2 VCDs, we first derive the LNS on a 0.25° × 0.25° grid in order to fit the lifetime parameters \( (k_c, a', b', c' \text{ in eq 4}) \). Fitting the lifetime parameters based on the LNS data at a resolution of 0.25° × 0.25° rather than 0.05° × 0.05° reduces noise in the fitted parameters. To fit the lifetime parameters, as detailed in Kong et al. (2019), we make use of the relatively clean regions where the emission is small enough to be neglected and thus the LNS is dominated by the loss term \( L(x,y) \approx -\frac{C(x,y)}{r(x,y) \cdot r(x,y)} \). These small-emission, clean regions are selected as having low NO2 VCDs (<1.5 × 10^15 molecules cm^-2) and low LNS values (at 1% quantile). The lifetime parameters are then derived through nonlinear quantile fitting for the 1%-quantile LNS data as a function of VCDs. The fitting-related uncertainty for derived emissions is estimated by conducting the fitting procedure at other quantile thresholds (0.1% and 2%, see SI Table S3). We then apply the determined lifetime parameters to rederive LNS and estimate NOx emissions on a 0.05° × 0.05° grid.

**Uncertainty in Emission and Lifetime Estimates.** The uncertainty in NOx emissions derived in each subdomain comes from three major sources; a step-by-step uncertainty estimate follows Kong et al. (2019) with necessary updates detailed in SI Table S3. First, the uncertainty in gridded NO2 VCDs is associated with the NO2 retrieval process for each pixel, subsequent long-term averaging for NO2, and removal of the NO2 background. Since the pixel-specific NO2 data are then temporally averaged, the random component of errors is reduced. Combining data from multiple days is a common practice for top-down emission assessments to reduce the effect of random errors and missing values in satellite data sets.

Second, the PHLET-related uncertainty is derived from temporal averaging of NO2 over summer 2019, assumption of stable NO2 (emission equals loss) at the overpass time of satellite, fitting for the NO2/NOx ratio, and assumption of the 2-dimensional process with long-term average horizontal flux and effective diffusion. These two error sources affect the estimate of LNS, with individual sources of errors added in quadrature to establish the diagonal of error covariance matrix of the cost function in PHLET Adjoint.

Third, the uncertainty associated with division of the LNS into emission and loss is derived from fitting of the relationship between NO2 VCD and lifetime. A uniform relation between NO2 VCD and NOx lifetime is assumed and derived in each subdomain of 5° × 5° to retain local features of the nonlinear chemistry. In reality, the VCD-lifetime relationship within a subdomain may not exactly follow a single parametrized formula. The associated error is reflected in part in the uncertainty of fitted parameters and lifetimes. We find that the NOx lifetimes fitted from (up to) four subdomains covering each grid cell are within 15% for most grid cells, with an average of 10.8%. The respective error is reduced by averaging the results (weighted by the square of the reciprocal of the distance between each grid cell and the subdomain center) from those four subdomains.

For most grid cells, emissions estimated from four subdomains are weighted averages, so are their uncertainties. SI Figure S5a,b shows the spatial distributions of emission uncertainties (1−σ) in absolute and relative terms, respectively. As expected, the absolute uncertainties tend to be larger (by up to 7.66 kg NO2 km^-2 h^-1) over high-emission grid cells, and the relative uncertainties tend to be larger over low-emission grid cells. The spatial correlation between the derived emissions (Figure 2a) and their absolute uncertainties (SI Figure S5a) is about 0.55.

SI Figure S5c,d further shows the horizontal distributions of estimated lifetimes and their absolute uncertainties (1−σ). The lifetimes are about 1.5–3.5 h over most of the high-emission regions of East Asia and South Asia, with uncertainties of 0–1.5 h. The lifetimes are the highest (>5 h) over regions of low anthropogenic emissions such as the oceans and Mongolia, so are their uncertainties (>3 h). Especially over land, the lifetimes are smoother in space than the respective emissions.

We further use GEOS-Chem simulations to test the reliability of our emission inversion approach, following Kong et al. (2019). First, we use the nested GEOS-Chem v12.9.3 to simulate the NO2 VCDs around the overpass time...
of TROPOMI in summer 2019 on a 0.3125° longitude × 0.05° latitude grid over Asia. A brief description of model setup is presented in SI Section 1. The simulated NO2 VCDs are shown in SI Figure S6a. Then the simulated VCDs are regridd to 0.05° × 0.05° and debackgrounded, following the procedures described above for TROPOMI data. Finally, we apply our emission inversion approach to derive the emissions at 0.05° × 0.05°, regrid them to 0.3125° longitude × 0.25° latitude (SI Figure S6b), and compare them to the emissions used in GEOS-Chem (SI Figure S6c) at the same resolution. SI Figure S6d shows that the derived emissions match the GEOS-Chem emissions, with a correlation coefficient of 0.94, a slope of 0.97 and a normalized mean bias of −1.8%.

The above test with GEOS-Chem does not apply the averaging kernels (AKs) of POMINO-TROPOMI. Given that the AKs are embedded in satellite NO2 VCD data, we conduct another test in which we apply the POMINO-TROPOMI AKs to the GEOS-Chem NO2 profiles and sample the GEOS-Chem NO2 VCDs accordingly, before using them as input to retrieve NOx emissions. [Note that since there are missing values in satellite data, this unavoidably affects the subsequent test.] The resulting retrieved NOx emissions still agree well with the original emissions used in GEOS-Chem (R = 0.92 and NMB = −3.7%, SI Figure S7). There are some underestimates over parts of southern China and overestimates over parts of the North China Plain, similar to the test without applying the AKs (SI Figure S6). This result indicates a minor influence of AKs.

Extraction of Anthropogenic Contributions from Inferred Emissions. Over the study domain, the derived emissions on a 0.05° × 0.05° grid contain the contributions from anthropogenic, biomass burning (open fires) and soil sources. To further identify anthropogenic sources, we subtract from the derived emissions the contributions of biomass burning and soil (SI Figure S8). For this purpose, we adopt the soil NOx emissions from Weng et al. (2020)40 and the biomass burning emissions from GFED4;31 these emissions are interpolated to 0.05° × 0.05° by area weighting, and they together contribute about 11% of the emission total over the study domain.

Then, we examine whether the residual “anthropogenic” NOx emissions match three main proxies for human activity, including nighttime light (SI Figure S9a), traffic line data (SI Figure S9b,c), and Tencent location (SI Figure S9d). The proxy data are described in SI Section 2 and regridd to 0.05° × 0.05° by area weighting. We remove the emissions on grid cells for which none of the proxies indicate any human activities (Figure 1). For mainland China, the emissions removed amount to about 10%. Note that those emissions removed may not necessarily be of natural origin or artifact of inversion; they are excluded due to lack of ancillary data for further evaluation.

RESULTS AND DISCUSSION

Fine-Scale Characteristics of Derived Anthropogenic Emissions. Figure 2a shows the spatial distribution of derived on-land “anthropogenic” NOx emissions over China on a 0.05° × 0.05° grid for summer 2019. We focus the analysis on Chinese emissions and mask emissions outside the country and over the oceans, because the Tencent location data used to facilitate data analysis are less reliable outside China. The emissions total 24.75 ± 5.98 Tg NOx a⁻¹ (scaled based on the amounts of days in a summer and a year) for mainland China and show clear local details.

Regionally, NOx emissions are the highest over the North China Plain (including Beijing, Tianjin, Hebei, Henan, Shandong, Jiangsu and Anhui provinces), which contributes 30.1% (7.45 ± 1.71 Tg NOx a⁻¹) of the total emission in mainland China. In the northeast, emission sources gather around the G1 Expressway connecting the provincial capitals of Liaoning, Jilin and Heilongjiang provinces. The three provinces together contribute 11.4% (2.81 ± 0.69 Tg NOx a⁻¹) of the total emission in mainland China. Emissions are also large in the Yangtze River Delta (including Shanghai, Jiangsu, Anhui, and Hebei provinces) and Guangdong Province (encompassing the Pearl River Delta), contributing 10.7% (2.64 ± 0.76 Tg NOx a⁻¹) and 3.1% (0.77 ± 0.23 Tg NOx a⁻¹) of the total in mainland China, respectively. In the western regions, large emission sources are evident around the major cities (e.g., provincial capitals).

More importantly, there are clear fine-scale features at this resolution. The emission hot spots are obvious at the urban areas of each city. There are significant emissions along the traffic lines connecting major cities across the country, most apparently over the west and south, consistent with the road line data set (SI Figure S9b,c). Our emission data also reveal numerous small-to-medium sources spreading across the vast areas of the country, which are hardly seen in bottom-up inventories (see next section). These emissions are often in the suburban and rural areas and near small roads, consistent with the distributions of nighttime light (SI Figure S9a) and Tencent location data (SI Figure S9d).

Comparison with Bottom-up Anthropogenic Emission Inventories. We further compare the derived anthropogenic emissions to four bottom-up anthropogenic inventory, including the Multi-Resolution Emissions Inventory of China (MEIC v1.3, for summer 2017 at 0.25° × 0.25°),12,43 the Emission Database for Global Atmospheric Research (EDGAR v5.0, for summer 2015 at 0.1° × 0.1°),44 the Community Emissions Data System (CEDS v2021_04_21, for summer 2019 at 0.5° × 0.5°),45 and Peking University NOx (PKU-NOx v2, for summer 2014 at 0.1° × 0.1°).46 For each inventory, we use available data in the year closest to 2019. According to CEDS, the emissions decrease by 11.9% from 2014 to 2017 with high spatial correlation (R = 1.00) between the two years, and by 3.2% from 2017 to 2019 (R = 1.00). According to MEIC, the emissions in 2017 are lower than in 2014 by 14.8% (R = 0.99). The interannual variability in the spatial patterns of bottom-up emissions is negligible.

Figure 2 compares the horizontal distributions of different emission data. Our emissions regridd to lower resolutions are shown in SI Figure S10. The MEIC (21.15 Tg NOx a⁻¹), EDGAR (26.37 Tg NOx a⁻¹), CEDS (22.83 Tg NOx a⁻¹) and PKU-NOx (25.04 Tg NOx a⁻¹) emission totals for mainland China are all within 15% of ours (24.75 Tg NOx a⁻¹). The large-scale spatial distribution of our emissions is similar to the bottom-up inventories, while there are much more small-to-medium sources in our emissions.

The most important difference between our emission data and bottom-up inventories lies in the fine-scale features. To illustrate this point, Figure 2 shows enlarged emission maps for three small areas including Beijing, the Hetao Plain, and west Xinjiang. In the Beijing area (Figure 2a-d, I), our data show large emissions along the traffic lines connecting Beijing and its surrounding cities, and the emissions of Beijing concentrated within the Fifth Ring Road (denoted by the black circle in Figure 2a-I, with an average of 6.98 ± 2.07 kg NOx km⁻² h⁻¹).
There is a clear emission signal around the highway connecting the urban center and the Beijing Capital International Airport (denoted by the white cross in Figure 2a—I, with a maximum of $8.85 \pm 1.59$ kg NO$_2$ km$^{-2}$ h$^{-1}$). These fine-scale features are weak or even invisible in the bottom-up inventories, because of their lower spatial resolutions and lack of accurate spatial proxies to allocate provincial-level traffic emissions to individual locations.

In the Hetao Plain of Nei Monggol (Figure 2a—d_II), our data reveal high emissions along the G6 and G7 Highways. These signals are vague in MEIC and PKU-NO$_x$. EDGAR shows the emissions from G6 but not G7, likely because this section of G7 was not in operation in 2015 yet. The resolution of CEDS is too low to capture such traffic emissions. To the north of the highway section shared by G6 and G7, our data show emissions (with an average of $1.61 \pm 0.64$ kg NO$_2$ km$^{-2}$ h$^{-1}$) from the towns and valleys close to the Wujia River where most people in Bayannur City live. Again, these emissions are missing or not clear in the bottom-up inventories.

In west Xinjiang (Figure 2a—d_III), which covers the Aksu region and Kuqa City, our data show significant emissions around both the major roads and the minor roads; see SI Section 2 for definitions of the two road types. This information is not visible in MEIC, CEDS, and PKU-NO$_x$. EDGAR shows spotted emissions along the G3012 Highway crossing Aksu and Kuqa, and does not include emissions from the minor roads as well as nearby towns and villages. Our data also show large emissions in Baicheng County (with an average of $2.78 \pm 0.79$ kg NO$_2$ km$^{-2}$ h$^{-1}$), which is an important place for extraction of rare-earth metal minerals. The emissions resulting from industrial, transportation and residential activities around Baicheng are absent in the bottom-up inventories.

We further examine the provincial emissions. The scatter plots in Figure 3 show that in Shanghai, Hubei, Zhejiang, and

**Figure 3.** Scatter plots for provincial emissions between our and other data sets for 31 provinces in mainland China. Taiwan, Hong Kong, and Macau are not included. The sizes of the circles indicate the GDP value of each province in 2015. The yellow circles stand for the provinces explicitly mentioned in the main text.
Guangdong, our emissions are significantly lower than the inventories. Provincial total of our emissions is lower than the average of the four inventories by 56.5% for Shanghai, 51.9% for Hubei, 43.6% for Guangdong, and 39.9% for Zhejiang. These differences are attributed to the uncertainties in our inversion and the inventories. On the one hand, our NO<sub>2</sub> emissions might be underestimated in those provinces due to a possible overestimate in the NO<sub>2</sub> lifetime. A test with the lifetime shortened by 30% (according to its uncertainty in SI Figure S5d) would increase the derived emissions by 30%. On the other hand, there are substantial differences among these inventories reflecting their individual uncertainties. Besides, these provinces may have experienced larger emission reductions than other provinces over the past few years, and this difference may not be captured in the inventories. According to the POMINO-OMI data, from 2014 to 2019, the summertime mean NO<sub>2</sub> VCD decreases by 66.8% over Shanghai, 51.9% over Hubei, 43.6% over Guangdong, and 39.9% over Zhejiang. These decreases are attributed to the uncertainties in our inversion and the inventories. Over these provinces, and EDGAR and PKU-NO<sub>2</sub>x do not show such stronger-than-national-average decreases over mainland China, and leads to much improved spatial correlation with the inventories (Figure 3 and SI Figure S11). The second test regrids our emissions to 0.25°×0.25° before removing the emissions. This leads to results (SI Figure S12) similar to the first test (SI Figure S11).

To further evaluate our emission data, we conduct two additional GEOS-Chem (v12.9.3, 0.25°×0.3125°) simulations in July 2019, using our emissions and MEIC, respectively. The results are compared with the Ministry of Ecology and Environment (MEE) surface NO<sub>2</sub> measurements at 1468 sites. The simulated monthly mean daily mean surface NO<sub>2</sub> concentrations based on our NO<sub>2</sub> emissions show higher spatial correlation and a smaller bias (R = 0.49 and NMB = −45% versus R = 0.46 and NMB = −50%) (SI Figure S13). In the western, less affluent provinces such as Xinjiang, Xizang, and Nei Monggol, the simulation driven by our emissions achieves an even higher performance gain upon the simulation based on MEIC (R = 0.54 and NMB = −42% versus R = 0.29 and NMB = −74%). We note that the MEE NO<sub>2</sub> measurements are almost entirely for the urban areas and cannot be fully captured by GEOS-Chem simulations at 0.25°× 0.3125°. Also, the MEE NO<sub>2</sub> concentrations may be overestimated due to contamination from oxidized nitrogen species such as nitric acid and peroxyacyl nitrates. Nevertheless, these comparisons support the advantage of our emission data especially in the less affluent areas.
Figure 4 shows how NO\textsubscript{2} emissions cumulate with county-level GDP, after mapping the gridded data to each county and arranging the counties from the lowest to the highest affluence level (in terms of per-capita GDP). Compared to the inventories, our emission data show higher contributions (to the total in mainland China) from low-affluence counties. Specifically, counties with per-capita GDP lower than 50 thousand Yuan contribute 33% of the total GDP, but they contribute 58% of the total emission in our data, compared to 47% in MEIC, 50% in PKU-NO\textsubscript{2} and EDGAR, and 56% in CEDS. This is because NO\textsubscript{2} emissions from transportation and small-to-medium sources in these low-affluence counties are not well represented and often missing in the inventories, due to lack of detailed information (e.g., emission activities, emission factors, and locations) on these sources.

The inlet of Figure 4 highlights the contributions of most affluent counties to the total emission of mainland China. Counties with per-capita GDP larger than 150 thousand Yuan constitute 4% of the total GDP and 3% of the total emission according to our data set, in contrast to 2% for MEIC, PKU-NO\textsubscript{2} and EDGAR, and 1% for CEDS. This is likely because transportation emissions are not well captured in the inventories.

**DISCUSSION**

This study improves our PHLET-based inversion approach to estimate NO\textsubscript{2} emissions at a high horizontal resolution (5 km) for East Asia based on TROPOMI-POMINO NO\textsubscript{2} VCD data. Our inversion approach accounts for the effects of horizontal transport and nonlinear NO\textsubscript{x} chemistry for each location. A focused emission analysis is done for mainland China, facilitated by five spatial proxy data sets of human activity. Our emission data reveal fine spatial structures, transportation signals and many small-to-medium sources which are missing or poorly represented in four widely used bottom-up emission inventories.

To date, China’s effort to reduce emissions has focused on more developed regions especially in the east, particularly guided by bottom-up emission inventories. However, our emissions reveal numerous small-to-medium sources in less affluent counties with relatively low per-capita GDP, and these sources are often missing in the inventories. In the future, emissions in less affluent areas might experience further increases if the industries move to these areas, as a result of industrial restructuring and/or as a response of enterprises to strengthening emission control in more affluent regions.

Therefore, effective, targeted control measures to reduce small-to-medium emission sources in less-affluent areas are an important further step to protect China’s atmospheric environment. To this end, our high-resolution emission data offer a reference to help track and reduce emissions at places where bottom-up emission inventories are less reliable.

Our retrieval algorithm can be further improved. First, our retrieval constrains daily mean emissions using satellite NO\textsubscript{2} VCDs in the early afternoon. This is a common practice for polar orbiting satellite-based emission constraint at any horizontal resolutions. Previous studies have shown that when proper retrieval approaches are used, emissions inferred from morning satellite measurements are within 10% of those inferred from measurements in the afternoon. This offers some confidence that the intrinsic problem of polar orbiting satellite remote sensing (i.e., it lacks diurnal variability information) has a limited effect on the emission constraint.

Geostationary satellite remote sensing such as the Geostationary Environment Monitoring Spectrometer (GEMS) is expected to provide much more information on the diurnal variation of NO\textsubscript{x} useful for emission constraint. Note that the diurnal variability of emissions is also not or poorly represented in bottom-up emission inventories, including those analyzed in this study. Second, although our emissions show many high-resolution details of the NO\textsubscript{x} emission distribution, they alone cannot tell the exact source types. This is another common limitation of top-down emissions. Combining top-down emissions with sectorally detailed bottom-up data and/or local surveys will be useful to identify source types for targeted emission control.

Our emission inversion approach is a fast algorithm. After all input data are prepared, the emission inversion for the whole study domain takes about 900 core-hours. And since the calculation is highly parallelized by dividing the study domain into subdomains, it only takes several actual hours to derive the NO\textsubscript{x} emissions for the whole domain. The reliability and low cost mean that the approach can be applied to the globe and other years, allowing a low-cost, timely and reliable derivation of NO\textsubscript{x} emissions, variabilities and trends at a fine spatial scale.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.1c07723.

Examples for interpolated NO\textsubscript{2} profiles (Figure S1); NO\textsubscript{2} VCDs before/after adjustments (Figure S2); Regression model evaluation and predicted NO\textsubscript{x}/NO\textsubscript{2} ratio (Figure S3); Illustration of subdomains (Figure S4); NO\textsubscript{x} lifetime and uncertainties of NO\textsubscript{x} lifetime and emissions (Figure S5); Evaluation of emission inversion approach based on GEOS-Chem simulations (Figure S6); Evaluation of emission inversion approach based on GEOS-Chem simulations with the AKs applied (Figure S7); Natural NO\textsubscript{x} emissions (Figure S8); Proxy data of human activity (Figure S9); Our NO\textsubscript{x} emissions at lower resolutions (Figure S10); Scatter plots of provincial NO\textsubscript{x} emissions after removing minor roads from our emissions (Figure S11); Scatter plots of provincial NO\textsubscript{x} emissions after regridding our emissions to 0.25° × 0.25° and removing minor roads from our emissions (Figure S12); GEOS-Chem simulations based on our emissions and MEIC and comparison with MEE measurements (Figure S13); Provincial NO\textsubscript{x} emission amounts (Figure S14); Fitted coefficients for regression for NO\textsubscript{x}/NO\textsubscript{2} ratio (Table S1); Uncertainty description (Table S2); Correlation between proxies and emissions (Table S3); Model description of GEOS-Chem (Section 1); Description of proxy data of human activity (Section 2) (PDF).

**AUTHOR INFORMATION**

**Corresponding Author**

Jintai Lin — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China; orcid.org/0000-0002-2362-2940; Phone: +86-010-62767973; Email: linjt@pku.edu.cn

https://doi.org/10.1021/acs.est.1c07723

Environ. Sci. Technol. XXXX, XXX, XXX–XXX
Environmental Science & Technology

Authors

Hao Kong — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China; orcid.org/0000-0001-5038-2914

Lulu Chen — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China; orcid.org/0000-0002-8929-3414

Yuhang Zhang — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

Yingying Yan — Department of Atmospheric Science, School of Environmental Sciences, China University of Geosciences (Wuhan), Wuhan 430074, China

Mengyao Liu — R&D Satellite Observations Department, Royal Netherlands Meteorological Institute, De Bilt NL-3731 GA, The Netherlands

Ruijing Ni — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China; orcid.org/0000-0003-3072-6817

Zehui Liu — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

Hongjian Weng — Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.1c07723

Author Contributions

J.L. conceived the research. J.L. and H.K. designed the research. H.K. performed the data processing, model development, and simulations. Y.Z. and M.L. contributed to processing of POMINO-TROPOMI NO2 VCDs. L.C., Y.Y., R.N., and H.W. contributed to the setting of GEOS-Chem. Z.L. contributed to the model setting when revising the paper. H.W. provided soil NOx emissions. L.C., Y.Y., and R.N. contributed to the writing. H.K. and J.L. analyzed the results and wrote the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research is supported by the National Natural Science Foundation of China (Grant No. 42075175), the second Tibetan Plateau Scientific Expedition and Research Program (Grant No. 2019QZKK0604), and the National Natural Science Foundation of China (Grant No. 41775115).

REFERENCES

(14) Mijling, B.; van der A, R. J.; Zhang, Q. Regional nitrogen oxides emission trends in East Asia observed from space. Atmospheric Chemistry and Physics 2013, 13 (23), 12003–12012.
(18) Beirle, S.; Borger, C.; Dörner, S.; Eskes, H.; Kumar, V.; de Laat, A.; Wagner, T. Catalog of NOx emissions from point sources as derived from the divergence of the NOx flux for TROPOMI. Earth Syst. Sci. Data 2021, 13 (6), 2995–3012.
(20) Kong, H.; Lin, J.; Zhang, R.; Liu, M.; Weng, H.; Ni, R.; Chen, L.; Jingwu, Y.; Yan, Y.; Zhang, Q. High-resolution (0.05° × 0.05°) NOx emissions in the Yangtze River Delta inferred from OMI. Atmospheric Chemistry and Physics 2019, 19, 12835–12856.

https://doi.org/10.1021/acs.est.1c07723 Environ. Sci. Technol. XXXX, XXX, XXX–XXX
satellite NO 2 retrieval: seasonal and spatial characteristics and implications for NOx emission constraints. Environmental Science & Technology pubs.acs.org/est
