



High-resolution ($0.05^\circ \times 0.05^\circ$) NO_x emissions in the Yangtze River Delta inferred from OMI

Hao Kong¹, Jintai Lin¹, Ruixiong Zhang^{1,a}, Mengyao Liu¹, Hongjian Weng¹, Ruijing Ni¹, Lulu Chen¹, Jingxu Wang¹, Yingying Yan², and Qiang Zhang³

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

²Department of Atmospheric Science, School of Environmental Sciences, China University of Geosciences, Wuhan, China

³Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China

^anow at: School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

Correspondence: Jintai Lin (linjt@pku.edu.cn)

Received: 7 December 2018 – Discussion started: 2 January 2019

Revised: 30 August 2019 – Accepted: 3 September 2019 – Published: 15 October 2019

Abstract. Emission datasets of nitrogen oxides (NO_x) at high horizontal resolutions (e.g., $0.05^\circ \times 0.05^\circ$) are crucial for understanding human influences at fine scales, air quality studies, and pollution control. Yet high-resolution emission data are often missing or contain large uncertainties especially for the developing regions. Taking advantage of long-term satellite measurements of nitrogen dioxide (NO_2), here we develop a computationally efficient method of estimating NO_x emissions in major urban areas at the $0.05^\circ \times 0.05^\circ$ resolution. The top-down inversion method accounts for the nonlinear effects of horizontal transport, chemical loss, and deposition. We construct a two-dimensional Peking University High-resolution Lifetime-Emission-Transport model (PHLET), its adjoint model (PHLET-A), and a satellite conversion matrix approach to relate emissions, lifetimes, simulated NO_2 , and satellite NO_2 data. The inversion method is applied to the summer months of 2012–2015 in the Yangtze River Delta (YRD; $29\text{--}34^\circ \text{N}$, $118\text{--}123^\circ \text{E}$) area, a major polluted region of China, using the NO_2 vertical column density data from the Peking University Ozone Monitoring Instrument NO_2 product (POMINO). A systematic analysis of inversion errors is performed, including using an independent test based on GEOS-Chem simulations. Across the YRD area, the summer average emissions obtained in this work range from 0 to $15.3 \text{ kg km}^{-2} \text{ h}^{-1}$, and the lifetimes (due to chemical loss and deposition) range from 0.6 to 3.3 h. Our

emission dataset reveals fine-scale spatial information related to nighttime light, population density, road network, maritime shipping, and land use (from a Google Earth photo). We further compare our emissions with multiple inventories. Many of the fine-scale emission structures are not well represented or not included in the widely used Multi-scale Emissions Inventory of China (MEIC).

1 Introduction

Nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$) are a main precursor of particulate matter, ozone, and other atmospheric pollutants. NO_x strongly influences the atmospheric oxidative capacity, affects the climate, and is toxic to many organisms. NO_x is emitted from natural and anthropogenic sources (Lin, 2012). Over the past decade, China has experienced rapid growth in the gross domestic product (GDP, by $8.3\% \text{ a}^{-1}$ on average from 2008 to 2017), fossil fuel consumption (by $5.5\% \text{ a}^{-1}$ from 2007 to 2015), and urbanization (National Bureau of Statistics of China, <http://data.stats.gov.cn/>, last access: 20 August 2018). These socioeconomic changes have been accompanied by a rapid change in NO_x emissions in the urban and surrounding areas. With the large and continuously increasing urban population and motor vehicles, NO_x pollution is particularly severe in large cities such as Beijing

and Shanghai (Barnes and Rudziński, 2013; Lin et al., 2016). Many coastal cities like Shanghai have also experienced enormous growth in the shipping business. Therefore, pollution along the coastal line has become a serious problem associated with the growth of global economic trade; and emissions from seaborne transport play an increasingly important role in the global air pollution (Fu et al., 2017). Understanding the urban pollution and its environmental impacts requires accurate quantitative knowledge of NO_x emissions at a very high horizontal resolution (e.g., 0.05° × 0.05°), which is typically lacking, especially for the developing countries.

Gridded bottom-up emission inventories typically use spatial proxies (like population and GDP) to allocate provincial-level emission values, which are derived from activity statistics and emission factor data, to individual locations (Zhao et al., 2011, 2015; Janssens-Maenhout et al., 2015). Such a gridding method may lead to large uncertainties at high resolutions (Geng et al., 2017), because the mismatch between proxies and emissions becomes more significant and emitting facilities are harder to allocate accurately as the resolution increases (Zheng et al., 2017). For a small area, emission factors and activity data of the major sources can be collected by on-site surveys to allow construction of a high-resolution inventory (Zhao et al., 2015; Granier et al., 2019), such as Zhao et al. (2015) for Nanjing. However, on-site surveys are extremely time consuming and resource demanding, and therefore difficult to be applied to a large domain in a timely manner.

Top-down inversion using satellite retrieval products of tropospheric vertical column densities (VCDs) of nitrogen dioxide (NO₂) is a widely used independent estimate of NO_x emissions (Martin et al., 2003; Stavrou et al., 2008; Lin et al., 2010; Mijling and van der A, 2012; Gu et al., 2014; Beirle et al., 2015; Miyazaki et al., 2017; Ding et al., 2017b). Top-down inversion typically provides the total emission data, although emissions from individual sources can be further derived by integrating a priori data (often from bottom-up inventories) about source-specific information such as diurnal and seasonal variabilities (e.g., Lin et al., 2010; Lin, 2012) and spatial variabilities (Timmermans et al., 2016).

The traditional top-down methods based on local mass balance (LMB) or its variants assume a weak effect of horizontal transport (Martin et al., 2003; Lamsal et al., 2011; Lin, 2012; Gu et al., 2014; Boersma et al., 2015). These algorithms work relatively well at low resolutions (> 50 km) given the relatively short lifetime of NO_x (hours to 1 d) but may introduce large uncertainties when applied to higher resolutions; for example, emissions in the rural–urban fringe zone cannot be identified accurately. The adjoint model and Kalman filter methods better account for horizontal transport, although their applicability is limited by expensive computational costs. These more sophisticated methods have often been applied to relatively short time periods (e.g., Gu et al., 2016 for 1 month), small spatial domains (e.g., Tang et al., 2013 in Texas), and/or at coarse horizontal resolu-

tions (e.g., Miyazaki et al., 2012 at T41 grid, i.e., ~ 2.8°, and Stavrou et al., 2008 at 5° × 5°). Top-down estimates can be further combined with bottom-up inventories and spatial proxies to increase the spatial resolution, such as from 0.25° × 0.25° in the daily emission estimates constrained by satellite observations (DECISO)-derived emissions to 0.01° × 0.01° for 2014 during the MarcoPolo project (Hooyberghs et al., 2016; Timmermans et al., 2016) and similar inventories over Qatar and South Africa (Maiheu and Veldeman, 2013). The LMB, adjoint model, and Kalman filter approaches normally use three-dimensional chemical transport models (CTMs) to relate emissions to VCDs. CTM-based studies typically provide an estimate of the overall model error, although Lin et al. (2012) and Stavrou et al. (2013) present errors in the individual model processes (e.g., key chemical reactions and meteorological parameters). A computationally low-cost method for space-based high-resolution (0.05° × 0.05°) NO_x emission estimate will be helpful for understanding the urban pollution and its trends and variability.

This study presents a computationally low-cost space-based top-down approach to construct high-resolution NO_x emission inventories for urban and surrounding areas. The approach is applied to the Yangtze River Delta (YRD) area (29–34° N, 118–123° E), which includes Shanghai, Nanjing, Hangzhou, and 15 other cities) on a 0.05° × 0.05° grid, using the Peking University Ozone Monitoring Instrument NO₂ product (POMINO) NO₂ VCD data retrieved from the Ozone Monitoring Instrument (OMI). We derive the average NO_x emissions for the summer months (June, July, and August) of 2012–2015. We construct a model called PHLET (two-dimensional Peking University High-resolution Lifetime-Emission-Transport model) and its adjoint model (PHLET-A) to facilitate the emission estimate. The lifetimes of NO_x are estimated as well, in order to account for the nonlinear NO_x chemistry.

Section 2 presents the data and method for top-down inversion of high-resolution NO_x emissions. Inversion uncertainties are analyzed explicitly. Section 3 presents spatial distributions of NO₂ VCDs, the derived local net sources (which are used subsequently to derive NO_x emissions and lifetimes), NO_x lifetimes, and their uncertainties. Section 4 analyzes the top-down emission data estimated here, including comparisons with spatial proxies (population density, night light brightness, power plant locations, road network, marine shipping routes, and a Google Earth photo for land use indication), the Multi-scale Emissions Inventory of China (MEIC) (Zheng et al., 2014; Liu et al., 2015), the DECISO top-down emissions (Mijling et al., 2013; Ding et al., 2017a), and the MarcoPolo emissions (Hooyberghs et al., 2016; Timmermans et al., 2016). Section 5 tests our inversion method by applying it to the NO₂ VCDs simulated by the GEOS-Chem CTM. Section 6 concludes the study.

2 Data and method

2.1 A general framework to retrieve NO_x emissions at a high resolution

The high-resolution NO_x emission retrieval framework consists of multiple steps, as illustrated in the flowchart (Fig. 1). First (Sect. 2.2), the POMINO NO₂ VCD data over the summer months of 2012–2015 are averaged on a $0.05^\circ \times 0.05^\circ$ grid, using a special oversampling technique that preserves the finest spatial information possible. A satellite conversion matrix (SCM), which will be applied to PHLET-simulated NO₂ VCDs at the second step, is also calculated based on the OMI pixel parameters (i.e., corner coordinates).

Second (Sect. 2.3), the PHLET model is constructed to simulate the local net source (i.e., emission – loss) and horizontal transport of NO₂ VCDs on the $0.05^\circ \times 0.05^\circ$ grid. The SCM is then applied to PHLET-simulated VCDs to mimic how each satellite pixel averages the spatial distribution of NO₂, in order to ensure the spatial sampling consistency between PHLET and POMINO. This process is needed because satellite pixels represent the NO₂ spatial distribution at a coarser (than PHLET) resolution with irregular shapes of individual pixels.

Third (Sect. 2.4), the PHLET-A adjoint model is constructed to, together with PHLET and POMINO VCDs, derive the local net source at each $0.05^\circ \times 0.05^\circ$ grid cell. We construct a cost function to quantify the difference between the distribution of POMINO VCDs and that simulated by PHLET at the $0.05^\circ \times 0.05^\circ$ grid. The inversion process to derive the local net sources is equivalent to minimization of the cost function.

Finally (Sect. 2.5), the emission and lifetime of NO_x at each $0.05^\circ \times 0.05^\circ$ grid cell is derived from the local net source term, by fitting a formula for the nonlinear relationship between lifetimes and VCDs. The formula is assumed to be fixed; i.e., the relationship is applicable to all grid cells within the small study domain.

Furthermore, a rigorous error analysis for the framework and models is conducted (Sect. 2.6). This analysis is complemented by an test based on the GEOS-Chem-simulated distribution of NO₂ VCDs (Sect. 5).

Our inversion method explicitly accounts for horizontal transport and the nonlinear relationship between lifetimes and NO₂ VCDs. With a few reasonable assumptions, the method is computationally efficient, suitable for speedily conducting high-resolution emission estimates in multiple areas and across a long time period (2012–2015 in this study). Both PHLET and PHLET-A are numerically solved based on FEniCS, a popular open-source solver (Farrell et al., 2012; Funke and Farrell, 2013; Alnaes et al., 2015). With one computational core (Intel® Xeon® Gold 6130 CPU @ 2.10 GHz), derivation of NO_x emissions over the YRD here takes less than 1 h after necessary input data are prepared. Applying

the framework to multiple areas would take a similar amount of time by using one computational core for each area.

2.2 Tropospheric NO₂ VCDs retrieved from OMI

OMI is a UV–visible nadir solar backscatter spectrometer aboard the Aura satellite (Levelt et al., 2006). OMI provides daily global coverage. Each complete swath of OMI consists of 60 ground pixels, the sizes of which increase from 13 km × 24 km at nadir to about 40 km × 150 km at the swath edge in accordance with the view zenith angle (VZA) from 0 to 57° (de Graaf et al., 2016).

We use level-2 tropospheric NO₂ VCD data from POMINO (Lin et al., 2014, 2015). As described in detail in Lin et al. (2014, 2015), POMINO is an OMI-based regional NO₂ product that includes a number of important features. Briefly, POMINO adopts the tropospheric slant column density (SCD) data from DONIMO v2 and conducts an improved calculation of tropospheric air mass factors (AMFs) and VCDs (i.e., $VCD = SCD/AMF$) (Boersma et al., 2011). Key features of the POMINO algorithm include explicit representation of aerosol scattering and absorption (by combining aerosol data from daily nested GEOS-Chem (at 0.3125° long. × 0.25° lat.) simulations and monthly MODIS/Aqua aerosol optical depth, AOD, data), explicit representation of the angular dependence of surface reflection, high-resolution NO₂ profiles from GEOS-Chem (at 0.3125° long. × 0.25° lat.), consistent retrievals of clouds (a prerequisite for the NO₂ retrieval) and NO₂, and use of a parallelized, LIDORT-driven AMFv6 package. POMINO NO₂ VCDs are consistent with ground-based MAX-DOAS data (Liu et al., 2019b).

To better relate NO_x emissions to NO₂ VCDs at the $0.05^\circ \times 0.05^\circ$ resolution, we only employ the NO₂ data in summer (June, July, and August), in which season the lifetimes of NO₂ are the shortest (a few hours). We combine data over 2012–2015 to increase the sample size. The change in NO₂ VCD from June to August is relatively small, reducing the effect of intraseasonal variability when deriving NO_x emissions from summer mean NO₂ VCDs. We screen out the 30 outer pixels with VZA larger than 30° (cross-track width larger than 36 km) that greatly smear the spatial gradient of NO₂, pixels with cloud radiance fraction exceeding 50 %, and pixels with AOD larger than 3 (i.e., when the aerosol data used in the NO₂ retrieval are unreliable and the NO₂ retrieval is subject to an excessive error) (Lin et al., 2014, 2015; Liu et al., 2019a). We also exclude data with raw anomaly problems (<http://projects.knmi.nl/omi/research/product/rowanomaly-background.php>, last access: 30 January 2018). After data screening, we obtain valid data from 22 007 pixels. We then convert the pixel-specific level-2 data to the $0.05^\circ \times 0.05^\circ$ grid.

To convert from the satellite pixels to the $0.05^\circ \times 0.05^\circ$ grid cells, we use an oversampling method that employs satellite data on multiple days to enhance the horizontal resolution (Zhang et al., 2014). For each $0.05^\circ \times 0.05^\circ$ grid cell, we av-

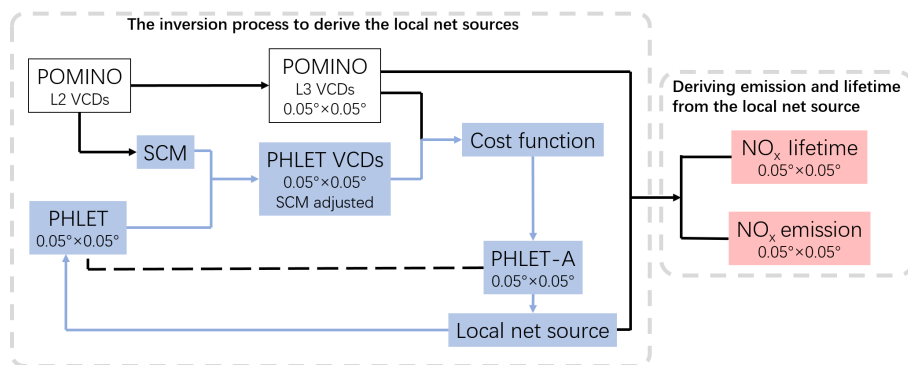


Figure 1. The flowchart of the framework of our methodology. The blue arrows show the iterative process.

erage all pixels covering the grid cell from all valid days, using area-based weighting. The oversampling approach takes advantage of the fact that the exact location of the OMI footprint slightly changes from one day to another, as does the exact location of the footprint of an OMI pixel at a given VZA. Thus, sampling from multiple days increases the horizontal resolution of data. Our oversampling approach is different from previous studies, which filled a grid cell with data from pixels within a certain distance (e.g., 30 km) and would result in spatial smoothing (Fioletov et al., 2011; Krotkov et al., 2016; Sun et al., 2018).

For the purpose of emission estimating, we assume that the error of VCD at a satellite pixel (σ_p) contains an absolute error of half of the mean VCD over the domain (i.e., 1.9×10^{-15} molec. cm⁻²) and a relative error of 30 % (Lin et al., 2010, 2015; Boersma et al., 2011; Beirle et al., 2011). We further add in quadrature an additional error (σ_g) when a satellite pixel is projected to the grid cells at a finer resolution; this error is important in the urban–rural fringe zone. For a given grid cell, σ_g is set to be 50 % of the standard deviation of VCDs at its eight surrounding grid cells. Sampling over multiple days reduces the random error by a factor of $s = (\sqrt{(1-c)/n+c})$, where c represents the fraction of systematic error (assumed to be 50 %) and n the number of days with valid data (Eskes et al., 2003; Miyazaki et al., 2012). Thus, the total error for the temporally averaged VCD at a given grid cell is $\sigma_s = \sqrt{(\sigma_p^2 + \sigma_g^2)} \cdot s$.

2.3 The PHLET model simulation

We construct the PHLET model on the $0.05^\circ \times 0.05^\circ$ grid to interpret the relationship between local net source (i.e., emission – loss), horizontal transport, and VCDs of NO₂ in a 2-D gridded space (Eq. 1) in the sense of long-time average. PHLET simulates the horizontal transport of NO_x through a time-averaged advection process and an “effective” diffusion process, which represents the residual from the temporally averaged advection. The vertical distribution is simplified as

in Sect. 2.3.2. The loss process of NO_x is represented by the lifetime.

2.3.1 Governing equation of PHLET

PHLET is an equilibrium model for the local net source, VCDs, and horizontal transport of NO₂ at each grid cell. Equation (1) presents the governing equation in PHLET:

$$\frac{\partial C(x, y)}{\partial t} = r \cdot L(x, y) - \mathbf{V}(x, y) \cdot \nabla C(x, y) + \nabla \cdot (\mathbf{K}(x, y) \cdot \nabla C(x, y)) = 0. \quad (1)$$

$C(x, y)$ represents the tropospheric NO₂ VCD (in molec. cm⁻²) due to sources over the YRD. The value of $C(x, y)$ gives the distribution of NO₂ VCDs at equilibrium ($\frac{\partial C(x, y)}{\partial t} = 0$ for every x and y). The discrete form of PHLET is set on the $0.05^\circ \times 0.05^\circ$ grid. The simulated VCDs will be applied with the SCM and compared to the gridded OMI data (after the contribution of horizontally homogeneous regional background is subtracted from the OMI data; see Sect. 2.3.2).

We assume a steady state of NO₂ in PHLET, although NO₂ observed by the satellite instrument may be in a transient state. We assign an error of 15 % to simulated $C(x, y)$ to account for the possible range of NO₂ variability at the overpass time of the instrument. Also, combining data from multiple years to derive an averaged NO₂ distribution for simulation (rather than conducting the simulations for individual years and months) leads to an additional uncertainty, which is set to be 10 % based on a comparison between the emissions estimated from multiple years together (here) and the average of emissions estimated from individual years (in a sensitivity test).

$L(x, y)$ represents the local net source term (in molec. cm⁻² s⁻¹, equivalent to 2.63×10^{-12} kg km⁻² h⁻¹), which combines the effects of ground emissions (anthropogenic + soil + biomass burning; see discussion in Sect. 2.3.2), deposition, and chemistry of NO_x. At equilibrium, the domain average of modeled $L(x, y)$ reaches zero,

because there are no horizontal fluxes into or out of the domain boundaries. $L(x, y)$ can be separated into an emission term and a loss term:

$$L(x, y) = E(x, y) - \frac{C(x, y)}{r \cdot \tau(x, y)}, \quad (2)$$

where $E(x, y)$ denotes the gridded emissions of NO_x, and $\tau(x, y)$ the lifetimes associated with deposition and chemical loss. r represents the ratio of NO₂ over NO_x concentration. The daytime NO_x chemical system reaches equilibrium rapidly and r varies little (Beirle et al., 2011; Valin et al., 2013). We set r to be 0.76 with an uncertainty of 10% (Seinfeld and Pandis, 2006; Beirle et al., 2011).

$V(x, y) \cdot \nabla C(x, y)$ represents the time-averaged advection term. $V(x, y)$ denotes the mean wind vector (in m s⁻¹) averaged over the summer months of 2012–2015. The wind data are taken from the European Centre for Medium-range Weather Forecasts (ECMWF) ERA5 dataset (see details in Sect. 2.3.3).

$\nabla \cdot (\mathbf{K}(x, y) \cdot \nabla C(x, y))$ represents the diffusion term, where $\mathbf{K}(x, y)$ denotes the “effective” diffusion coefficient tensor (in m² s⁻¹). The diffusion term accounts for transport by the residual winds deviating from the temporally averaged wind vector ($V(x, y)$). Appendix A shows how to determine the diffusion coefficient tensor.

2.3.2 Vertical shape and regional background of NO₂

PHLET assumes a horizontally homogeneous vertical shape of NO₂ concentrations, and that NO₂ is concentrated near the surface (Beirle et al., 2011). The assumption is implicitly used in many previous studies for polluted areas (Beirle et al., 2011; Liu et al., 2016, 2017). The corresponding uncertainty in the modeled NO₂ VCDs is set to 15% (Boersma et al., 2011; Lin et al., 2014).

Lightning emissions, biomass burning emissions, aircraft emissions, transport from neighboring regions, and convection can lead to NO₂ at higher altitudes over the YRD area. However, the amount of NO₂ aloft is much smaller than near-ground NO₂ due to large ground sources (Lin, 2012). Thus, we regard NO₂ aloft as the regional background and do not include it in Eq. (1). Also, for near-ground NO₂ over the YRD area, the contribution of downward vertical transport is negligible compared to the contribution of ground sources. Aircraft emissions contribute little to the total ground source, because 78% of aircraft emissions occur at the high altitudes (9–12 km) (Ma and Xiuji, 2000). Therefore, PHLET only accounts for near-ground NO₂ from ground soil, biomass burning, and anthropogenic sources (energy, industry, transportation, and residential).

To ensure the consistency between PHLET and OMI NO₂ data, we assume the background value to be half of the minimum OMI NO₂ VCD among all grid cells (i.e., 0.54×10^{15} molec. cm⁻²) and then subtract the background value from the gridded OMI data when comparing with PHLET

simulations. The corresponding uncertainty in the modeled NO₂ VCDs is set as 5%.

2.3.3 Initial conditions, lateral boundary conditions, and wind data input

To run PHLET, the NO₂ VCDs at the domain edges, as the lateral boundary conditions (LBCs), are set as the corresponding OMI NO₂ VCDs. For initial conditions, the VCD and the local net source at each grid cell inside the domain boundaries are set as zero. The horizontal distributions of modeled NO₂ VCDs and local net sources at equilibrium do not depend on the initial conditions.

For horizontal transport, we use 3-hourly wind fields from the ECMWF ERA5 dataset (<https://confluence.ecmwf.int/display/CKB/ERA5+data+documentation>; last access: 2 July 2018). The resolution of raw ERA5 data is 0.28125° on the reduced Gaussian grid, which is regridded to 0.05° × 0.05° by using the online program offered by ECMWF (see Fig. A1). We adopt the mean wind field of the lowest 14 vertical levels (out of 157 levels in total); these 14 levels represent the altitudes from surface to about 500 m (Beirle et al., 2011; Hersbach and Dee, 2016). Over the study period, the prevailing wind is northwesterly, and the wind speed is small over land (Fig. A1). For both zonal and meridional wind speeds, the uncertainty in the average wind speed is set to be 10%, which is similar to the temporal standard deviation of the wind speed and may partly account for the fact that lower-resolution wind data are used. We assess the model errors introduced by the uncertainties in the wind field and effective diffusion coefficients by Monte Carlo simulations in which the wind speeds are changed according to their uncertainties. The resulting relative uncertainty in the modeled NO₂ VCDs is about 20%.

2.3.4 Application of SCM

Remapping of PHLET-simulated NO₂ VCDs in accordance to satellite pixels is important. Given the size of OMI pixels, the OMI NO₂ data smooth to some extent the actual horizontal distribution of NO₂. To ensure consistent spatial sampling between PHLET and OMI data, for each day, we project the PHLET-modeled NO₂ VCD data (in the original 0.05° × 0.05° grid) to the satellite pixels to mimic how OMI “sees” the ground, remove the pixels with invalid OMI data, and then project the model data back to the 0.05° × 0.05° grid. The last two procedures are the same as done for OMI data. The whole process of grid conversion is done through the SCM approach (Appendix B). Although PHLET simulates summer average NO₂ VCDs (rather than daily values), we repeat the grid conversion process for as many days as there are valid OMI data.

2.3.5 Summary of model errors

The model error σ_m is set to be the sum in quadrature of errors contributed by the above mentioned steady state assumption (15 %), the time averaging over multiple years and months (10 %), the assumption of horizontally constant vertical shape of NO₂ (15 %), the NO₂/NO_x ratio (10 %), the treatment of background NO₂ concentration (5 %), and the error in the wind data and the calculation of effective diffusion coefficients (20 %).

2.4 PHLET-A: the adjoint model of PHLET

We construct the PHLET-A adjoint model to obtain an optimized horizontal distribution of the local net source term (L in Eq. 1) under the given OMI NO₂ VCDs, wind field, and other parameters. PHLET-A accounts for the complex nonlinear effects of two-dimensional transport and loss processes.

We first define a scalar cost function (Eq. 3) to quantify the difference between OMI NO₂ VCDs and PHLET-simulated (and SCM-applied) NO₂ VCDs.

$$J = \left(\mathbf{C}^{\text{OMI}} - \mathbf{C}^{\text{PHLET}} \right)^T \mathbf{S}_0^{-1} \left(\mathbf{C}^{\text{OMI}} - \mathbf{C}^{\text{PHLET}} \right) \quad (3)$$

Because PHLET does not require a priori knowledge about the local net source, the cost function does not include the a priori term either. The vector \mathbf{C} denotes gridded NO₂ VCDs. \mathbf{S}_0 denotes the observational error covariance matrix consisting of a satellite data error covariance matrix (\mathbf{S}_s) and a PHLET model error covariance matrix (\mathbf{S}_m):

$$\mathbf{S}_0 = \mathbf{S}_s + \mathbf{S}_m. \quad (4)$$

For simplicity and following previous studies (Keiya and It-sushi, 2006; Cao et al., 2018), both \mathbf{S}_s and \mathbf{S}_m are assumed to be diagonal, with the diagonal elements set to be σ_f^2 and σ_m^2 , respectively. Grid cells nearby may share the same pixels, although the area-based weights would be different. This means that nearby grid cells may not be fully independent, leading to a weakness of the diagonal assumption here. The associated uncertainty is partly accounted for by an error term based on the variability of NO₂ VCDs (i.e., 50 % of the standard deviation across the surrounding grid cells; see Sect. 2.2).

We then derive PHLET-A and its initial and lateral boundary conditions by applying Lagrange identity and integrating by parts (Marchuk, 1994; Sandu et al., 2005; Martien et

al., 2006; Hakami et al., 2007):

$$\begin{aligned} \frac{\partial \lambda(x, y)}{\partial t} + \nabla \cdot (\mathbf{V}(x, y) \cdot \lambda(x, y)) \\ + \nabla \cdot (\mathbf{K}(x, y) \cdot \nabla \lambda(x, y)) = 0 \end{aligned} \quad (5)$$

$$\lambda(x, y)|_{t=T} = \frac{\delta J}{\delta C(x, y)}|_{t=T} \quad (6)$$

$$\lambda(x, y)|_{\text{boundary}} = 0 \quad (7)$$

As shown in Eq. (5), PHLET-A represents the sensitivity of cost function (J) to local net source (\mathbf{L}) where λ stand for the adjoint variable (Marchuk, 1994; Sandu et al., 2005). T stands for the time when the domain-wide NO₂ VCDs come to equilibrium, i.e., the start time of the adjoint simulation. By discrete adjoint sensitivity analysis, the gradient of J to \mathbf{L} is obtained:

$$\frac{\delta J}{\delta L_{i,j}} = r \cdot \Delta x \cdot \Delta y \cdot \Delta t \cdot \sum_k \lambda_{i,j,k}, \quad (8)$$

where the indices i , j , and k denote zonal, meridional, and time, respectively. The gradient is then used in an iterative optimization (shown by the blue arrows in Fig. 1) to minimize the cost function J , i.e., to minimize the weighted difference between model simulated and OMI NO₂.

The numerical solution to obtain an optimized \mathbf{L} that minimizes J is as follows. Given a starting point of \mathbf{L} , we derive a search direction by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Li and Fukushima, 2001; Bousserez et al., 2015). Then, by practicing backtracking line search based on the Armijo–Goldstein condition (Armijo, 1966), we obtain a revised \mathbf{L} for the next iteration. The numerical calculation is done through FEniCS. It takes 50 iterations of PHLET and PHLET-A runs before the convergence is reached, according to the rate of reduction in J . The value of J is reduced from an initial value of 6585.2 to a stabilized value of 73.6 (Fig. 2).

The uncertainty of the optimized \mathbf{L} is given by the Hessian of the cost function, which is approximated by the BFGS method (Brasseur and Jacob, 2017):

$$\mathbf{S} = 2 \cdot \left(\nabla_{\mathbf{L}}^2 J \right)^{-1}. \quad (9)$$

2.5 Deriving emission and loss from the local net source term

The optimized local net source term combines the contributions of emission and loss (chemical loss + deposition). We further separate emission from loss by assuming a fixed formula within our small study domain for the nonlinear relationship between lifetimes and VCDs of NO₂.

In the summertime daytime, the dominant sink of NO_x is reactions with the radicals to produce nitric acid and organic nitrogen species. The NO_x chemistry quickly reaches a steady state under high solar radiation and air temperature in the early afternoon (Murphy et al., 2006; Valin et

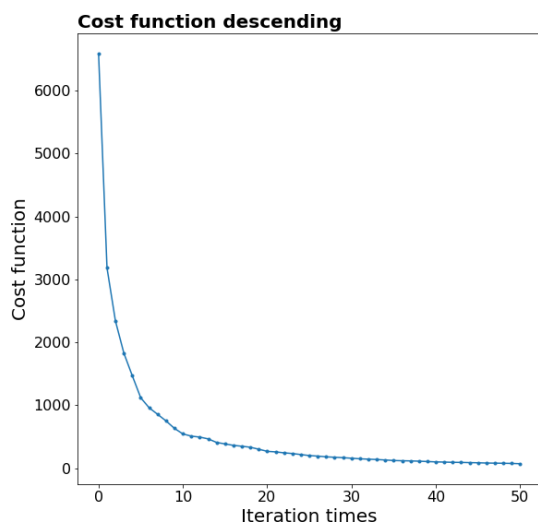


Figure 2. Cost function descending with iteration.

al., 2013) when OMI passes over the YRD. The chemical lifetime of NO_x depends on the concentrations of NO_x and non-methane volatile organic compounds (NMVOCs), radiation, temperature, and other factors. Within our small study domain, we assume the net effect of all factors except NO_x concentrations to be spatially homogeneous. As such, the chemical lifetime of NO_x at steady state is a sole function of NO_x concentration (and thus NO₂ VCD, given the constant NO₂/NO_x ratio). Appendix C shows in detail how to deduce the chemical lifetime of NO_x from NO₂ VCD, to account for the effect of dry deposition, to separate emission and lifetime from the local net source term, and to quantify the errors involved.

2.6 Uncertainty estimate for top-down emissions

For a particular grid cell, the derived emission is affected by the error involved in the estimate of L (embedded in satellite data and model simulations) and the error in the separation of emission and lifetime from L . The satellite data error (σ_s) is analyzed in Sect. 2.2. The model related error (σ_m) is analyzed in Sect. 2.3. The error of L (σ_L) is connected with σ_s and σ_m through the adjoint simulation and is given by Hessian of the cost function J (Sect. 2.4).

The error involved in the separation of emission and lifetime, σ_f , is contributed by the assumption on the NO₂/NO_x ratio r (Sect. 2.3.1), the simplified treatment of deposition and chemical processes of NO_x (Appendix C), and the assumed relationship between lifetimes and VCDs (Appendix C). σ_f is estimated by data fitting of L with different fitting parameters (Appendix C).

Thus, the error in emission, σ_e , is equal to the sum in quadrature of σ_L and σ_f , i.e., $\sigma_e = \sqrt{\sigma_L^2 + \sigma_f^2}$. The error in the lifetime is derived from the errors in NO_x loss (estimated

in Appendix C) and NO₂ VCDs, according to the common manner of error synthesis.

3 High-resolution spatial distributions of NO₂ VCDs, local net sources, and lifetimes over the YRD

Figure 3a shows the number of days with valid OMI data in the summer months of 2012–2015 over the YRD area on the $0.05^\circ \times 0.05^\circ$ grid. The number of days varies from about 12 to 101 (48 on average). There is a “band” pattern in the spatial distribution, due to the difference in the number of satellite orbits covering each grid cell (not shown). This band pattern is not obvious in the distribution of OMI NO₂ VCDs (Fig. 3b), suggesting that the temporally averaged VCD values are less sensitive to the number of days (12 or more) used for temporal averaging. There are fewer valid data in severely polluted locations. The effect of sampling size on the uncertainty in OMI NO₂ is accounted for in our study (Sect. 2.2).

Figure 3b shows the gridded horizontal distribution of OMI NO₂ VCDs. The background value (0.54×10^{15} molec. cm⁻²) has not been removed. NO₂ VCDs are high over the major urban centers along the Yangtze River and the coastal line, especially Shanghai, Nanjing (Capital of Jiangsu Province), Hangzhou (Capital of Zhejiang Province), and the Ningbo-Zhoushan area (with intensive maritime shipping activities). The maximum VCD value exceeds 16×10^{15} molec. cm⁻² in north Shanghai. NO₂ VCDs are larger than 1×10^{15} molec. cm⁻² at all grid cells, reflecting the influence of local anthropogenic sources and/or pollution transported from nearby cities (Cui et al., 2016). NO₂ VCDs are lower than 5×10^{15} molec. cm⁻² along the boundaries of our study domain.

Across the grid cells, the absolute errors in OMI NO₂ VCDs are about $1.6\text{--}4.9 \times 10^{15}$ molec. cm⁻² (Fig. 4a), and the relative errors are about 30 %–157 % (Fig. 4b). In general, the grid cells with larger NO₂ VCDs have larger absolute errors but smaller relative errors. Over the eastern sea area and the southwestern corner of the domain, NO₂ VCDs are relatively small (Fig. 3b); thus, their absolute errors are small (Fig. 4a), but their relative errors are very large (Fig. 4b).

Figure 4c shows the spatial distribution of the local net source L (emission–loss). A positive (negative) value of L indicates that the emission is larger (smaller) than the loss. The values of L are the greatest (11.7 kg km⁻² h⁻¹) over the major urban areas with high NO₂ VCDs, and are low (< -1.0 kg km⁻² h⁻¹) in many areas with low NO₂ loadings. The values of L are the lowest ($(-6.9)\text{--}(-2.0)$ kg km⁻² h⁻¹) at places in the urban-rural fringe zones with NO₂ hotspots nearby; this feature reflects that NO₂ is transported from the urban centers and destroyed in the fringe zones. The absolute errors of L vary from 0.6 to 4.5 kg km⁻² h⁻¹. The spatial correlation between the absolute errors of L (Fig. 4c) and those of the NO₂ VCDs (Fig. 4a) is about 0.5. The absolute errors

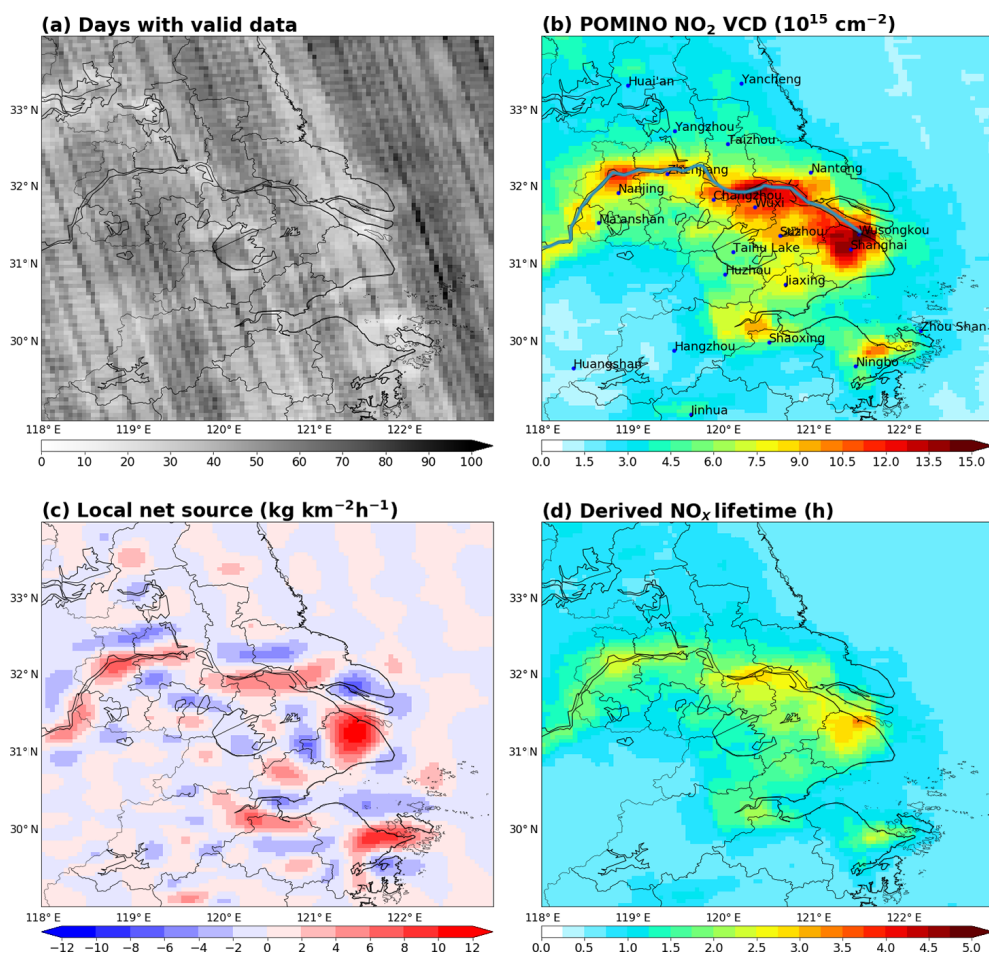


Figure 3. (a) Number of days with valid data in each grid cell over the summer months of 2012–2015. (b) POMINO NO₂ VCDs averaged over the summer months 2012–2015. Cities and locations mentioned in this paper are denoted. The Yangtze River is marked as a blue line. (c) Derived local net source. (d) Derived NO_x lifetime due to both chemical loss and deposition.

of L are notable in the urban-rural fringe zones where L is small but NO₂ VCD is high, because the deviation of L at these areas is very sensitive to errors in the assumed transport and loss process.

Figure 3d shows the derived lifetimes of NO_x on the 0.05° × 0.05° grid. The lifetimes range from 0.6 to 3.3 h across the study domain, with an average of 2.0 h. The lifetimes are about 0.6 h at grid cells with NO₂ VCDs of about 1.6 × 10¹⁵ molec. cm⁻², increasing to 0.8 h at grid cells with the lowest VCDs (around 1.0 × 10¹⁵ molec. cm⁻²), and exceeding 3 h at many polluted grid cells, i.e., the urban centers. The nonlinear dependence of lifetimes on VCDs is expected from our inversion method (Appendix C). Appendix C further shows the chemical lifetimes to be 0.6–3.8 h and the deposition lifetimes to be constantly at 30.4 h across all grid cells.

Figure 4d shows that the absolute uncertainties in the lifetimes are greater than 1.0 h at the NO₂ hotspot locations, between 0.6 and 1.0 h over the eastern sea and the southwest

of the study domain, and about 0.4 h at many other locations. The lifetimes at high-NO₂ locations (up to 3.3 h) are consistent with previous studies, e.g., a single value of 4.7 ± 1.4 h for Shanghai in the summer months of 2005–2013 by Liu et al. (2016). The short lifetimes at low-NO₂ locations over the eastern sea may be underestimated, due to the assumption that non-NO₂ factors (especially NMVOCs) are spatially homogeneous within the study domain (Appendix C). In particular, the concentrations of NMVOCs over the eastern sea may be overestimated by this assumption, based on the OMI formaldehyde data (De Smedt et al., 2015).

4 High-resolution spatial distribution of NO_x emissions over the YRD

4.1 Spatial distribution of emissions

Figure 5a shows the derived horizontal distribution of the summer months of 2012–2015 average NO_x emissions on

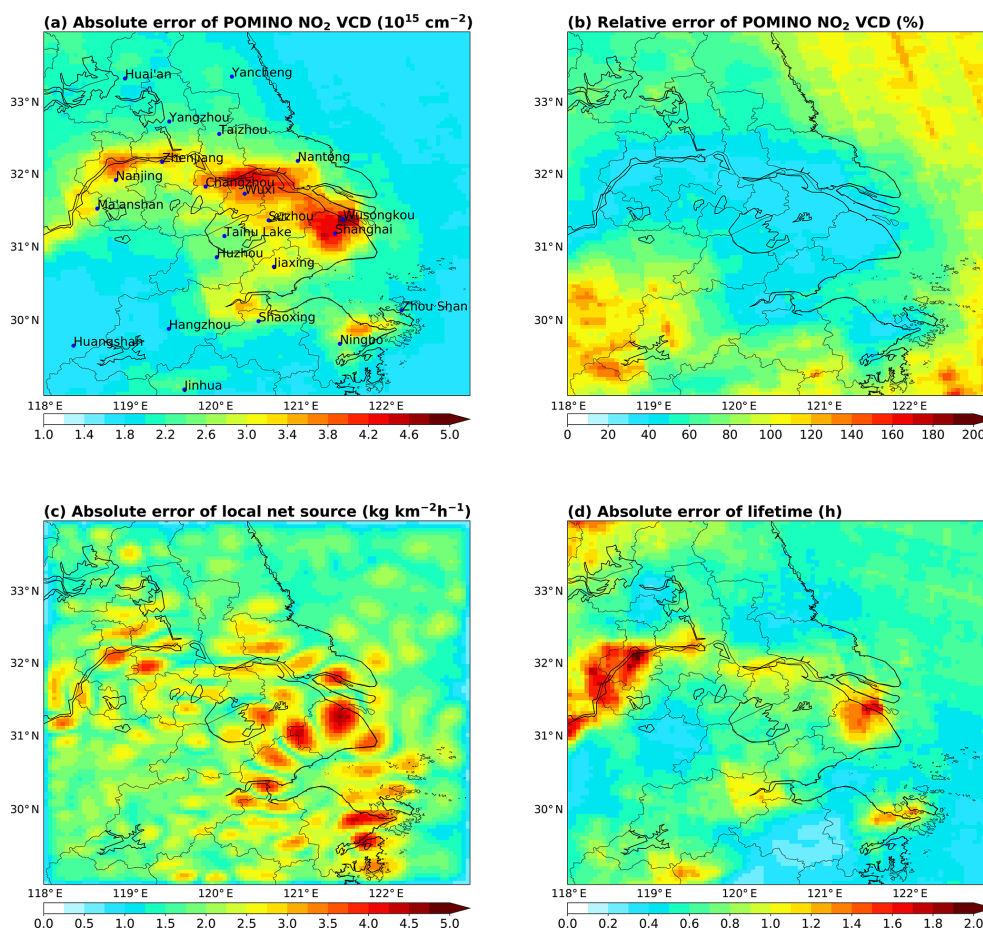


Figure 4. (a) Absolute error (1σ) of POMINO NO₂ VCD at each grid cell at the $0.05^\circ \times 0.05^\circ$ resolution. (b) Relative error (1σ) of POMINO NO₂ VCD. (c) Absolute error (1σ) of the derived local net source. (d) Absolute error (1σ) of the derived NO_x lifetime due to chemical loss and deposition.

the $0.05^\circ \times 0.05^\circ$ grid. The emissions include the contributions of ground anthropogenic (energy, industry, transportation, and residential), soil and biomass-burning sources. As discussed in Sect. 4.3, soil emissions contribute little (0.5 %) to the total emissions over the study domain, and biomass burning contributes about 5.1 %.

Figure 5a shows that NO_x emissions vary from 0 to 15.3 kg km⁻² h⁻¹ across the grid cells. The highest emission value occurs in north Shanghai, close to Wusongkou where the Port of Shanghai is located, which has become the largest container terminal in the world in 2010 (Fu et al., 2012). High emission values also occur at places along the Yangtze River and the coastal line. Along the Yangtze River, the highest emission value occurs in Nanjing city. Along the coastal line, there is an emission hotspot in the Ningbo–Zhoushan area. The general spatial distribution of NO_x emissions is consistent with that of OMI NO₂ VCDs (correlation = 0.69), reflecting the short lifetimes of NO_x and thus a modest effect of horizontal transport. Nonetheless, emissions are much more concentrated at a few sparse locations than NO₂ VCDs are,

and many locations near the emission hotspots have very low emissions but relatively large NO₂ VCDs, suggesting that the effect of horizontal transport cannot be ignored at such a high resolution.

Figure 5b shows the absolute errors of NO_x emissions at individual grid cells. The emission errors vary from 0.7 to 4.5 kg km⁻² h⁻¹ across all grid cells. The largest uncertainty occurs in north Shanghai, corresponding to the highest VCD (Fig. 3b) and emission (Fig. 5a) values. The spatial pattern of absolute emission errors (Fig. 5b) is closer to the pattern of NO₂ VCDs (Fig. 3b, correlation = 0.51) than to the pattern of emissions (Fig. 5a, correlation = 0.33). There are more hotspots in the distribution of emission errors than in the distributions of VCDs and emissions, because the emission errors can be high at locations with low VCDs and emissions. The spatial pattern of emission errors is consistent with that of L errors (Fig. 4c, correlation = 1.0), indicating that the errors in deriving emissions from the local net sources are rather homogenous. Figure 5c further shows that the relative

errors of emissions are high (> 100 %) over low-emission locations but much lower over emission hotspots.

The scatter plot in Fig. 5d shows the relationship between absolute emission errors and emissions at individual grid cells. The relationship is highly nonlinear, and there is large data spread where the emissions are low. The data spread tends to be smaller when emissions exceed $5 \text{ kg km}^{-2} \text{ h}^{-1}$. The emission errors tend to decrease as emissions increase until about $5 \text{ kg km}^{-2} \text{ h}^{-1}$, after which the emission errors tend to increase with the increasing emissions. The data points in Fig. 5d are colored to indicate the different ranges of VCDs, and they show that grid cells with higher NO₂ VCDs have larger emission errors and smaller data spread.

4.2 Comparison between our top-down emissions and spatial proxies

This section compares our NO_x emission dataset (Fig. 6b) with several spatial proxies widely used in bottom-up inventories, including nighttime light brightness (Fig. 6c), population density (Fig. 6d), road network (Fig. 6e), ship route density (Fig. 6f), power plant locations (Fig. 6g), and a satellite photo from Google Earth based on Landsat measurement that indicates the extent of land use (Fig. 6i). These proxies broadly represent the intensity of human activities and are highly related to NO_x emissions (Geng et al., 2017).

Figure 6c shows the spatial distribution of nighttime light brightness in 2012. The data are taken from version 4 of the Defense Meteorological Satellite Program – Operational Linescan System (DMSP-OLS) nighttime lights time series at a horizontal resolution of $0.5' \times 0.5'$ (<https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>; last access: 19 August 2018). The brightness is represented digitally from 0 to 63 bits. The nighttime light reflects the intensity of household activity, commercial activity, and resource consumption (Elvidge et al., 2013). When regridded to $0.05^\circ \times 0.05^\circ$, the spatial correlation between nighttime light brightness and NO_x emissions is about 0.61 over land.

Figure 6d shows the population density data, which are taken from the Gridded Population of the World v4 (GPWv4) at a horizontal resolution of $0.1^\circ \times 0.1^\circ$ (<https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/sets/browse>; last access: 19 August 2018) (Center for International Earth Science Information Network – CIESIN – Columbia University, 2016). This dataset provides population density data for every five years (2000, 2005, 2010, 2015, etc.). Data in 2012, 2013, and 2014 are estimated by fitting a natural spline to the 2000, 2005, 2010, and 2015 values. The population density varies greatly from the urban centers to the countryside. In north Shanghai, the population density exceeds $3.5 \times 10^3 \text{ km}^{-2}$. The NO_x emission hotspots match the population hotspots, and the lowest-emission locations have little population. When regridded to $0.05^\circ \times 0.05^\circ$, the spatial correlation between population densities and NO_x emissions is 0.50 over land.

Figure 6e shows the OpenStreetMap road network data (<https://download.geofabrik.de>; last access: 27 June 2018). The network includes both highways and local roads. In the southern areas (between 29° and 31° N), the spatial distribution of NO_x emissions largely coincides with the road network. The spatial coincidence is less obvious in the north because of the influence of non-mobile sources. NO_x emissions are notable along the three major national highways connecting Jinhua city (one of the largest hubs of light industry products in China), Hangzhou city, and Ningbo city. NO_x emissions are also identifiable along the national highway from Hangzhou City to Huangshan city. Pairs of NO_x emissions and traffic hubs are located to the west of Lake Taihu and in the urban centers. These results suggest the capability of our emission dataset in capturing the contribution of traffic sources.

Figure 6f shows the mean density of marine shipping routes over the eastern sea in 2016 (<https://www.marinetraffic.com>; last access: 27 June 2019). Over the northern parts of the eastern sea, the route density map shows certain north-south and northwest-southeast lines. High route densities are also evident close to the ports. These features are consistent with the distributions of NO₂ VCDs (Fig. 6a) and NO_x emissions (Fig. 6b, same as Fig. 5a).

The filled circles in Fig. 6g show the locations of coal-fired power plants in 2016 from Carbon Brief (<https://www.carbonbrief.org/>; last access: 27 June 2019). The radius of a circle denotes the power generation capacity. Figure 6h further shows the Global Power Emissions Database (GPED) v1.0 bottom-up NO_x emissions for power plants on a $0.1^\circ \times 0.1^\circ$ grid in 2016. Coal-fired power plants in the YRD are normally near the urban centers, traffic lines or other sources. Our top-down NO_x emission map shows large emission values near the power plants (Fig. 6b), although it cannot isolate the sole contribution of power plants. At the GPED power plant locations, the correlation between our and GPED emissions reaches 0.26, due to the influence by non-power plant sources; note that the correlation between GPED emissions and POMINO NO₂ VCDs are only about 0.21.

Figure 6i shows a satellite photo taken in 2018 from Google Earth (<https://www.google.com/earth/>; last access: 4 July 2019). The grey areas in the photo represent developed lands and the dark green areas indicate undeveloped places. The majority of lands over the YRD have been developed. Although the lands over the southwest are less influenced by humans than the areas like Shanghai are, many places of the southwest have been developed as cities, towns and roads. This explains the spotted emission sources (Fig. 6b) retrieved from the satellite NO₂ VCDs.

4.3 Comparison between our emission dataset and other inventories

This section compares our emission data to several inventories for the region, including the MEIC bottom-

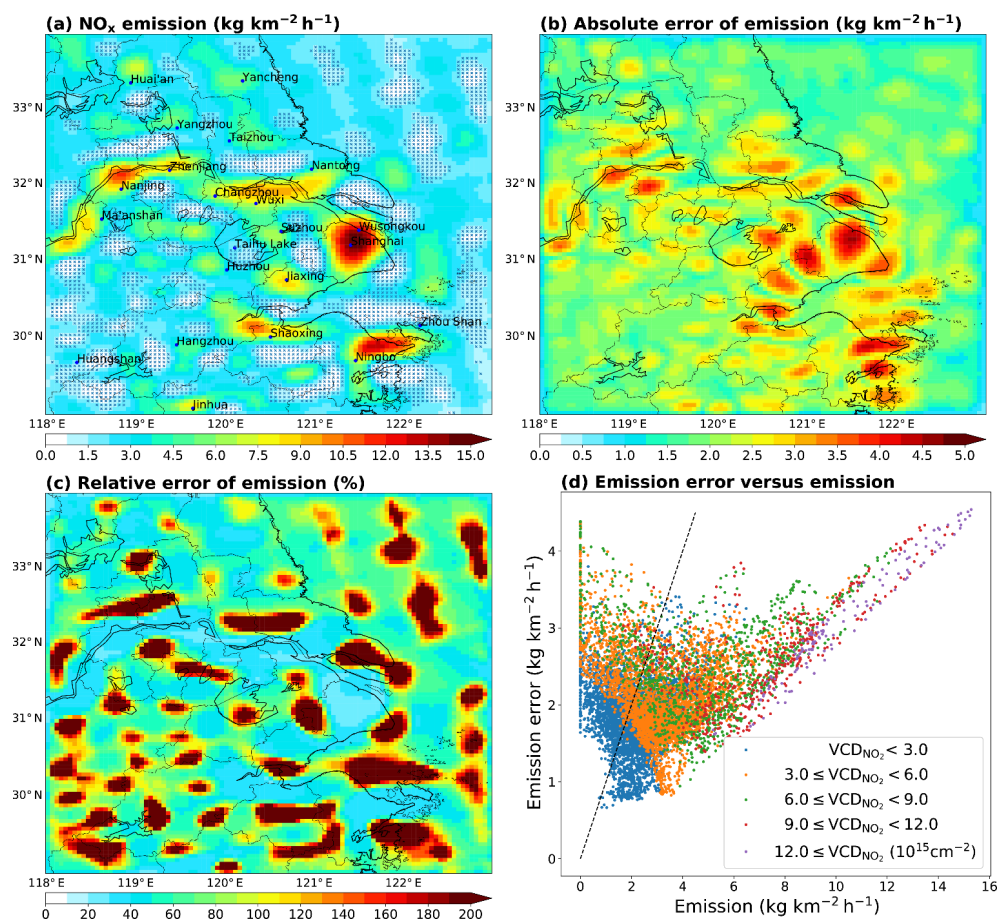


Figure 5. (a) Our NO_x emissions from anthropogenic, biomass burning and soil sources together. The blue crosses indicate where the relative errors exceed 100 %. (b) Absolute errors (1σ) of our NO_x emissions. (c) Relative errors of our emissions. (d) Absolute errors (1σ) of NO_x emissions as a function of NO_x emissions at individual grid cells. Data points are colored according to the magnitudes of POMINO NO₂ VCDs. The dashed line indicates an error of 100 %.

up anthropogenic inventory in the summer months of 2012–2015 (<http://www.meicmodel.org/>; last access: 2 July 2018) (Liu et al., 2015; Zheng et al., 2014), the MarcoPolo (bottom-up plus top-down hybrid) anthropogenic inventory in summer 2014 (<http://www.marcopolo-panda.eu/products/toolbox/emission-data/>; last access: 4 May 2019) (Hooyberghs et al., 2016), and the DECSO v5.1qa top-down emissions in the summer months of 2012–2015 (http://www.globemission.eu/region_asia/datapage.php; last access: 14 November 2018) (Mijling et al., 2013; Ding et al., 2017a). MEIC and DECSO emission data are available at the $0.25^\circ \times 0.25^\circ$ resolution, and MarcoPolo are at $0.01^\circ \times 0.01^\circ$. When compared with our emissions, these emission data are regridded to $0.05^\circ \times 0.05^\circ$.

Our emission data and the DECSO inventory are top-down estimates and include the contributions of soil and biomass-burning sources. Thus, we estimate soil and biomass-burning emissions from independent sources, and then subtract these emissions from our and DECSO emission datasets. Soil emissions are calculated by the nested GEOS-Chem

(Fig. 7c), with the uncertainties assumed to be within 50 % (Yienger and Levy, 1995; Wang et al., 1998). Biomass burning emissions (Fig. 7b) are taken from the Global Fire Emissions Database (GFED4; <https://www.globalfiredata.org>; last access: 10 July 2019) (Giglio et al., 2013), with the uncertainties estimated to be within 10 % over the YRD (Giglio et al., 2010, 2013). Summed over the study domain, the soil sources contribute about 0.5 % of our emissions, while biomass burning contribute about 5.1 %. Figure 7a shows the resulting “anthropogenic” portion of our emissions.

Compared with MEIC (Fig. 7d) and DECSO (Fig. 7e), our high-resolution anthropogenic emission dataset (Fig. 7a) provides much more detailed spatial information. Our dataset identifies the emission hotspots and their contrast with nearby low-emission areas (e.g., in the urban-rural fringe zones) better than MEIC and DECSO do. The contribution of mobile sources along the road network is clearer in our dataset. Our emission data contain sources over the nearby sea (i.e., from shipping), along the coastal line, and in the southwest of the domain, which are not included in MEIC.

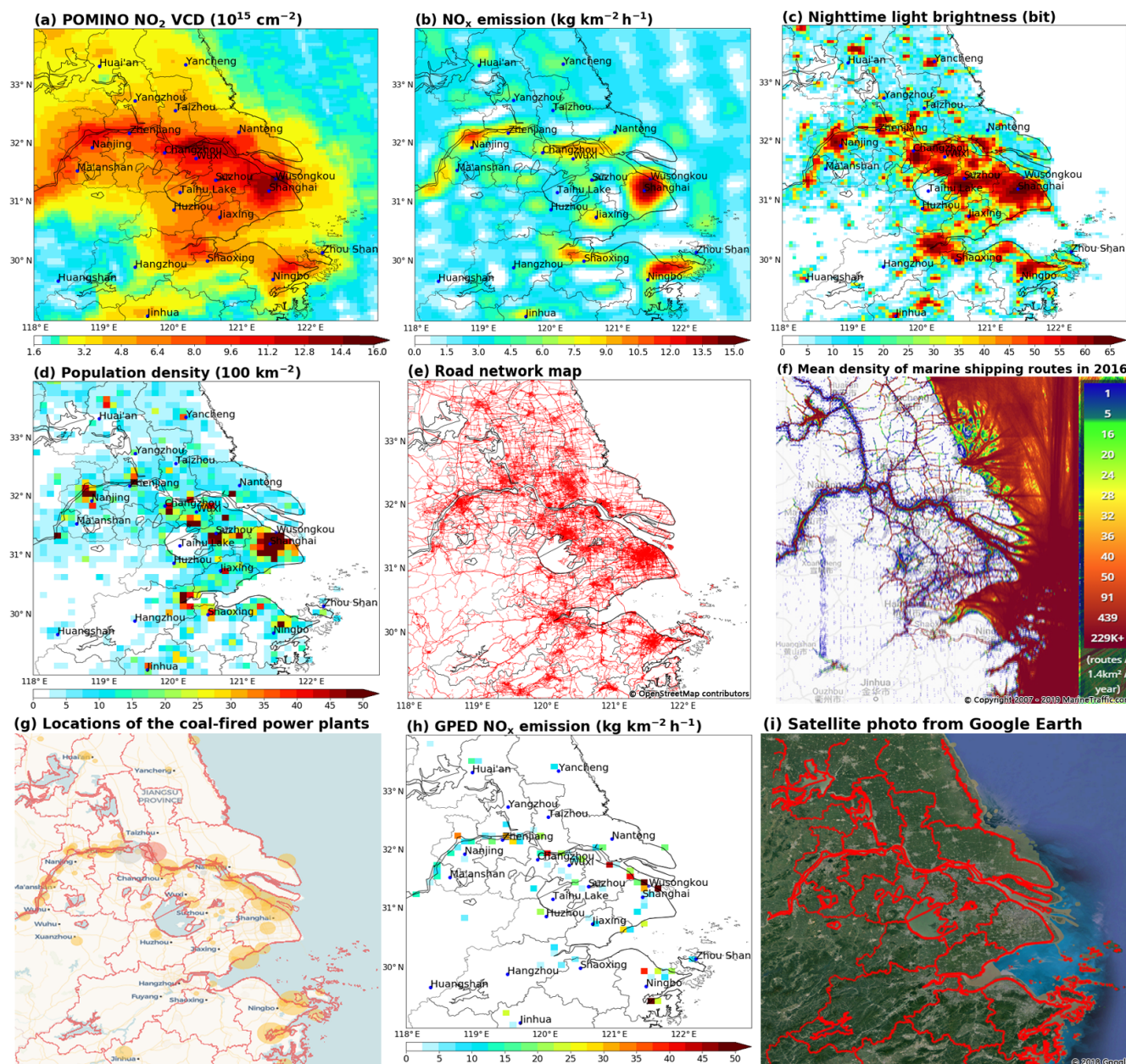


Figure 6. (a) POMINO NO_2 VCDs averaged over the summer months of 2012–2015, which is the same as Fig. 1a, but based on a different color scale. (b) Our NO_x emissions from anthropogenic, biomass-burning and soil sources together, same as Fig. 3a. (c) Nighttime light brightness in 2012. (d) Population density averaged over 2012–2015. (e) Road network (red lines). (f) Mean density of marine shipping routes in 2016 (data source: <https://www.marinetraffic.com>, last access: 27 June 2019). (g) Locations of coal-fired power plants in 2016 from Carbon Brief. (h) Global Power Emissions Database (GPED) v1.0 bottom-up NO_x emissions from coal-fired power plants. (i) A satellite photo from © Google Earth taken in 2018.

Compared with DECSO, our dataset suggests higher emissions on the northern parts of the sea, which may be due to our underestimate of NO_x lifetimes (Sect. 3) and/or errors in the DECSO estimate.

MarcoPolo emissions (regridged to $0.05^\circ \times 0.05^\circ$, Fig. 7f) show more detailed spatial information than our emission dataset (Fig. 7a) does. This is because our top-down esti-

mate is limited by the intrinsic resolution of NO_2 VCDs, i.e., our oversampling approach does not fully compensate for the large sizes of OMI pixels. Therefore, the large spatial gradient of NO_x emissions is smoothed to some extent in our dataset. On the other hand, the domain that MarcoPolo covers (29.635°N , 118.135°E – 32.625°N , 122.125°E) is much smaller than ours, and emissions of nine cities (in-

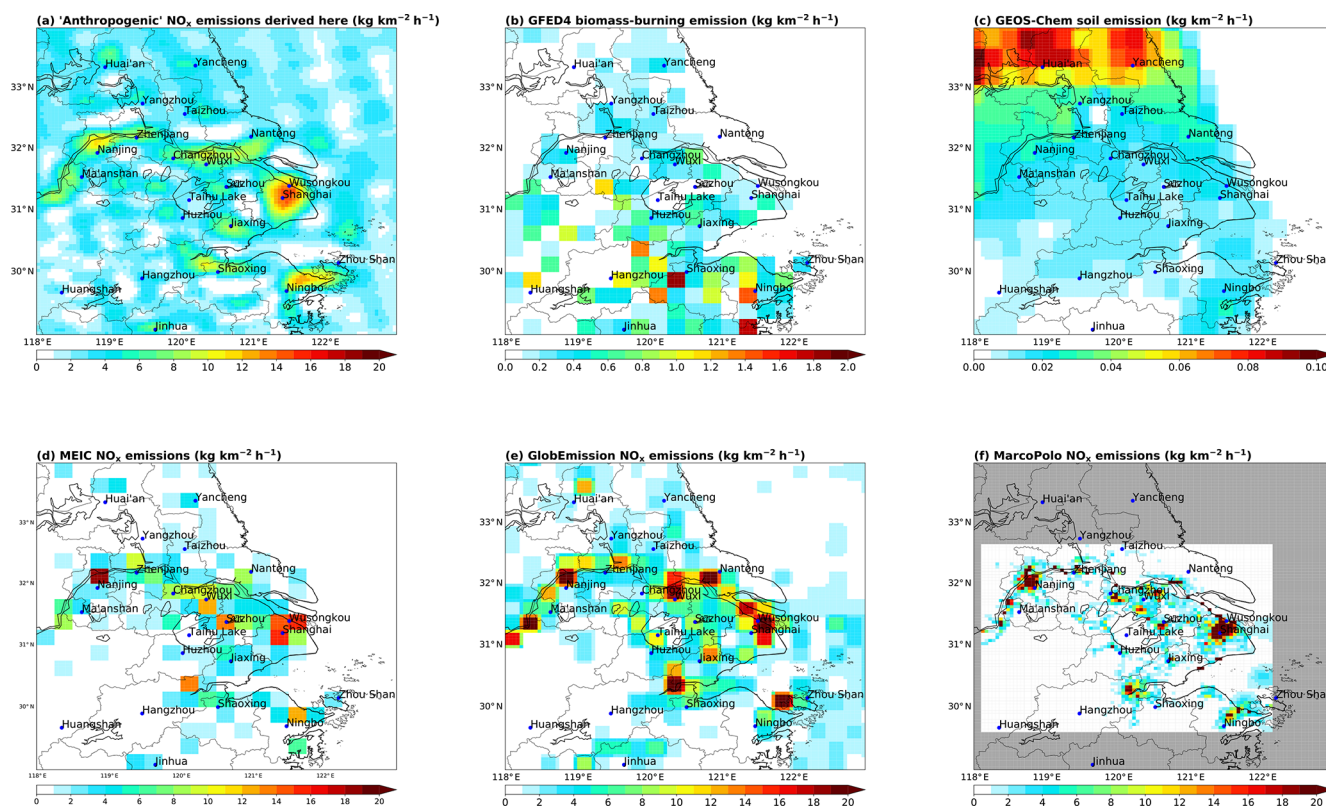


Figure 7. (a) Our “anthropogenic” NO_x emissions, by subtracting soil and biomass burning emissions from the derived emissions. (b) GFED4 biomass burning NO_x emissions. (c) Soil NO_x emissions calculated by a nested GEOS-Chem simulation. (d) MEIC NO_x emissions over the summer months of 2012–2015. (e) DECSO v5.1qa top-down emissions in the summer months of 2012–2015. (f) MarcoPolo bottom-up inventory in summer 2014; note that this inventory does not cover the grid cells shown in grey). All data are regridded to $0.05^\circ \times 0.05^\circ$.

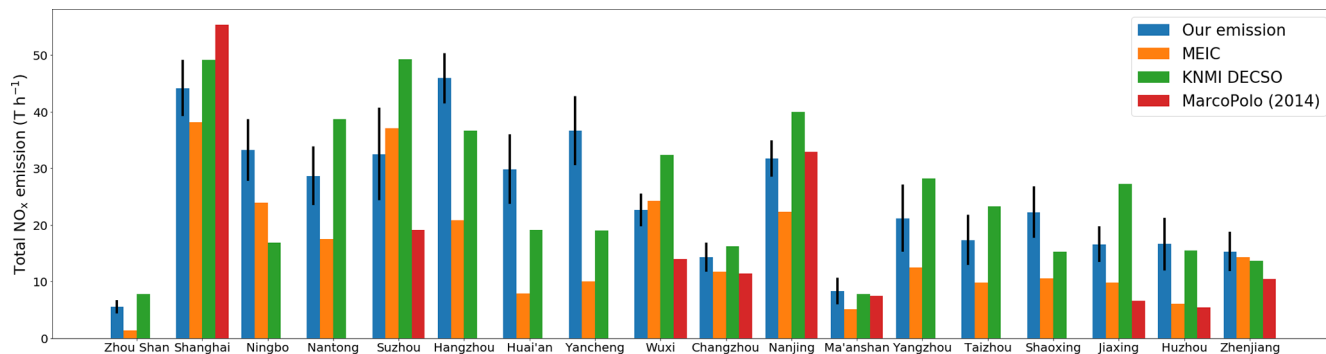


Figure 8. Total anthropogenic NO_x emission in each city for the summer months of 2012–2015 derived here, in comparison with other emission datasets. Soil NO_x emissions calculated by the nested GEOS-Chem and biomass burning NO_x emissions from GFED4 have been subtracted from our emission data and DECSO. Black vertical lines denote the uncertainty (1σ) of our emissions.

cluding Zhou Shan, Ningbo, Nantong, Hangzhou, Huai’an, Yancheng, Yangzhou, Taizhou, Shaoxing) and marine shipping emissions are not included in MarcoPolo.

At $0.05^\circ \times 0.05^\circ$, the spatial correlations between our NO_x emissions and spatial proxies are 0.61 for nighttime light brightness and 0.50 for population density (Sect. 4.2). These values can be compared to the respective results for Mar-

coPolo on the $0.05^\circ \times 0.05^\circ$ grid (0.35 and 0.55, respectively). When regridded to $0.25^\circ \times 0.25^\circ$, the correlations between our emissions and these spatial proxies become higher: 0.70 for nighttime light brightness and 0.69 for population density. The weaker correlation at a higher resolution reflects that as the spatial resolution gets finer, the chance that NO_x emissions are collocated with population or nighttime light

becomes smaller (Zheng et al., 2017), because of the influences of NO_x-emitting factories, power plants, and mobile sources. By comparison, the correlations between MEIC and these proxies on the 0.25° × 0.25° grid are 0.80 for nighttime light and 0.81 for population density. The respective correlation values for DECSO are 0.66 and 0.46. The lower correlation values for our dataset and DECSO than for MEIC partly reflect that top-down emissions better account for the influences of land transportation, which are spatially not tied closely to nighttime light and population at this resolution.

Figure 8 compares city-level emissions between our and other inventories. A total of 18 cities within the domain are selected, and for each city the NO_x emissions are summed over the grid cells within the municipal administrative boundaries. All inventories are gridded at 0.05° × 0.05° for this purpose. The MarcoPolo inventory does not include emissions for several cities; thus, the respective color bars are missing from Fig. 8. Among the cities, emission values differ from −5.8 % to +67.5 % between our emissions and the mean values of all four inventories (ours, DECSO, MEIC, and MarcoPolo, if available). For most cities, our emissions are consistent with at least one of the other three inventories, often the DECSO top-down inventory, after accounting for errors in our emission estimate. In Yancheng, Huai'an, and Ningbo, our emission values are higher than the averages of our study, DECSO, and MEIC by 67.5 %, 57.5 %, and 34.6 %, respectively. Ningbo (around 29.8° N, 121.5° E) is a coastal city with many isles and marine ports, as identifiable on the nighttime light map (Fig. 6c). The marine ports in the Ningbo–Zhoushan area contribute about 10 % of the total shipping emissions in China (Endresen et al., 2003; Fu et al., 2017). Our dataset and DECSO account for emissions from marine shipping and ports, whereas MEIC does not.

5 Test of our top-down emission derivation method by using GEOS-Chem-simulated NO₂ data

This section further presents an test to estimate the reliability of our emission derivation method, by examining to what extent the method can reproduce the emissions used in a nested GEOS-Chem simulation. Specifically, we use the nested GEOS-Chem v9-02 (Yan et al., 2016; Liu et al., 2018; Ni et al., 2018) to simulate the NO₂ VCDs in the early afternoon (around the overpass time of OMI) in summer 2014 on the 0.3125° longitude × 0.25° latitude grid. The simulated NO₂ data are shown in Fig. 9a and the emission inputs are shown in Fig. 9b. Next, we convert the GEOS-Chem NO₂ VCDs into the 0.05° × 0.05° grid, and parameterize PHLET with the wind field adopted by GEOS-Chem, following the procedures in Sect. 2.3. Then we use PHLET, PHLET-A, and the lifetime-emission separation method to estimate the NO_x emissions. Finally, we compare the derived emissions (remapped to the 0.3125° × 0.25° grid) to those used in GEOS-Chem.

Figure 9c and d shows the horizontal distributions of our “anthropogenic” emissions and emission errors, respectively. The contribution of soil and biomass burning emissions (as simulated by GEOS-Chem) are subtracted from the dataset. Figure 9e shows the differences between the derived anthropogenic emissions and those used in GEOS-Chem. The emission difference at each grid cell varies from −3.0 to 5.4 kg km^{−2} h^{−1}, which is attributed to the limitation of our inversion method. The domain average difference is 0.28 kg km^{−2} h^{−1}, or 18 % of GEOS-Chem emissions. The scatter plot in Fig. 9f suggests excellent consistency between the derived and the GEOS-Chem emissions, with a linear regression slope of 1.06 and correlation of 0.94. The emission differences for most grid cells are within the uncertainties of the derived emissions (shaded area).

6 Concluding remarks

This study presents a satellite-based top-down method to estimating NO_x emissions over urban and surrounding areas at a high horizontal resolution. As a demonstration, the method is applied to the YRD area at the 0.05° × 0.05° resolution in the summer months of 2012–2015, based on the POMINO NO₂ product. We construct a simplified, computationally efficient 2-D lifetime-emission-transport model (PHLET) and its adjoint model (PHLET-A) to, together with other procedures, facilitate the emission estimate. The reliability of our inversion method is supported by (1) a rigorous step-by-step derivation of models, assumptions, and parameters used, (2) a comprehensive uncertainty analysis, and (3) a test with GEOS-Chem-simulated NO₂ data. Our emission dataset in the YRD area on the 0.05° × 0.05° grid shows fine-scale spatial information that is tied to nighttime light, population density, road network, maritime shipping, and land use indicated from a Google Earth photo. Our dataset reveals many fine-scale spatial characteristics not well represented or not included in lower-resolution inventories such as MEIC and DECSO. Although this study derives the averaged emissions over the summer months of 2012–2015, calculations of emissions at higher temporal resolutions (e.g., every 2 years) is possible to better capture the interannual variability and trends.

Our inversion method is useful for understanding how human activities have altered the atmospheric environment at fine resolutions. Many crucial human activities, such as urbanization, are conducted at very fine spatial scales. How the resulting emissions affect air quality, public health, and geohealth are still poorly understood due to lack of high-resolution emission data. This problem is particularly severe in the developing countries, because of their rapid paces of urbanization and great inadequacies in emission-related information such as economic statistics and emission factors. This poses a grand challenge for emission control and environmental management. Thus, our inversion method

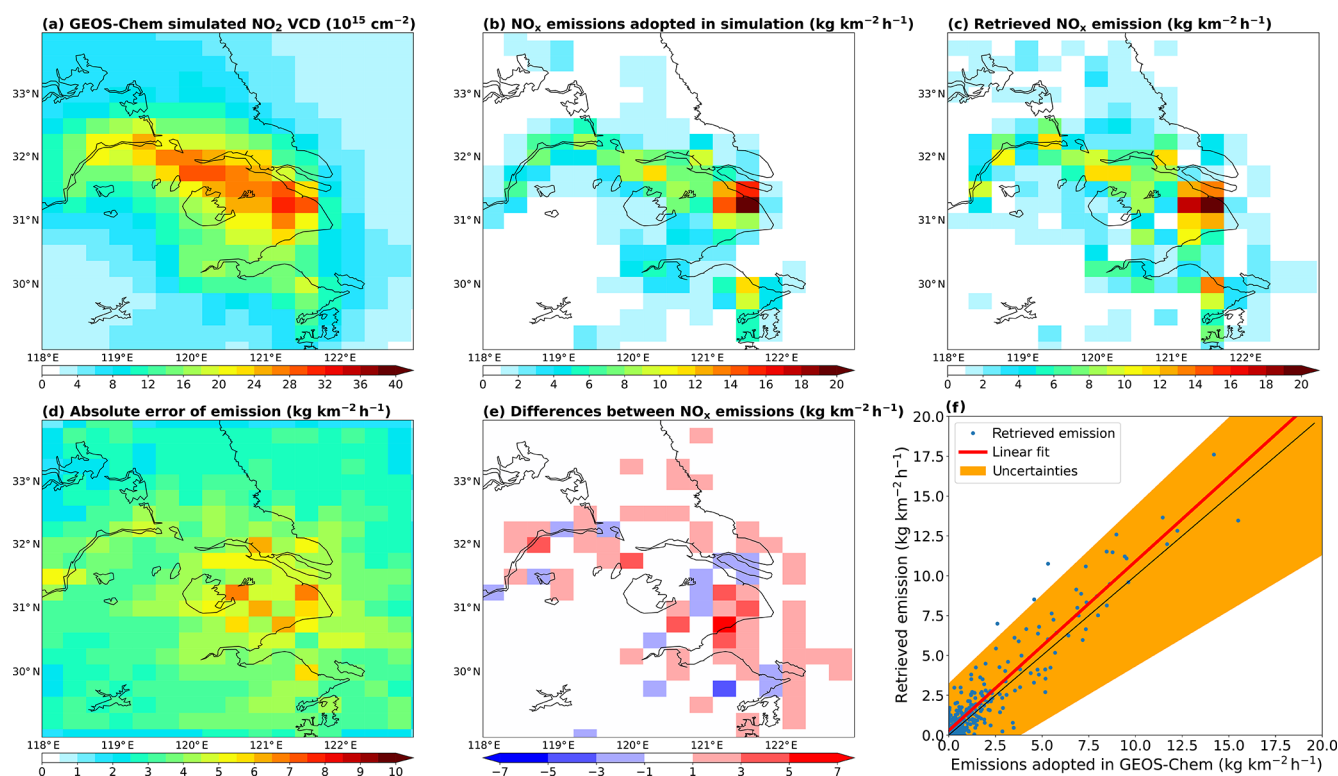


Figure 9. (a) GEOS-Chem-simulated NO_2 VCDs at the $0.3125^\circ \times 0.25^\circ$ resolution for summer 2014. (b) Anthropogenic NO_x emissions used in GEOS-Chem. (c) Anthropogenic emissions derived based on GEOS-Chem-simulated NO_2 VCDs and our inversion method. (d) Absolute errors (1σ) of derived emission data. (e) Differences between the derived emissions and GEOS-Chem emissions (derived minus GEOS-Chem). (f) Scatter plot for the derived emissions (y axis) and GEOS-Chem emissions (x axis). The red line represents least square linear fitting. The shading represents the fitting by accounting for errors in the derived emission data, i.e., derived emissions $+1\sigma$ for the upper bound, and derived emissions -1σ for the lower bound. The black dotted line denotes the 1 : 1 line.

and resulting emission data offer useful independent high-resolution information to monitor the fine-scale emission sources, to improve the bottom-up inventory, to model the urban pollution chemistry and the effect of urbanization, and to conduct spatially targeted emission control.

Our inversion method also has a few shortcomings. The derived emissions do not separate the individual contributions of anthropogenic sectors (i.e., power plants, industry, transportation, and residential). The spatial resolution of the estimated emissions is limited by that of satellite VCD data, although a special oversampling technique has been used to help achieve the highest spatial resolution possible for emissions. The PHLET model is assumed to be two-dimensional by simplifying the vertical distribution of NO_2 and not accounting for the spatial variability in the vertical shape, similar to previous studies. The adjoint model assumes the observational error covariance matrix to be diagonal, without fully considering the effect of correlations between individual grid cells. Also, we assume a spatially uniform relationship between NO_2 VCDs and NO_x lifetimes, which may lead to an underestimate in the lifetimes at low- NO_2 locations over the eastern sea.

Our emission inversion method and models have a few important features enabling their global applications. PHLET and PHLET-A are written in the Python language, which can be readily used with low financial costs. The PHLET model offers computationally efficient simulations of the NO_x chemistry, deposition, and transport. At a low computational cost, our inversion method is able to account for the nonlinear relationship between NO_x concentration, chemical loss, deposition, and transport. With the advent of TROPOMI and other satellite sensors with unprecedented spatial resolutions, our inversion approach can be applied to these measurements for continuous inference of emissions at finer and finer resolutions.

Data availability. Observational data are obtained from individual sources (see links in the text and acknowledgements). Model results are available upon request. Model codes are available on a collaborative basis.

Appendix A: Solving the diffusion process

The diffusion term can be simplified as follows:

$$\begin{aligned} & \nabla \cdot (\mathbf{K}(x, y) \cdot \nabla C(x, y)) \\ &= \frac{\partial}{\partial x} \left(K_x \cdot \frac{\partial C}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \cdot \frac{\partial C}{\partial y} \right). \end{aligned} \quad (\text{A1})$$

K_x and K_y are the diffusion coefficients in the zonal and meridional directions, respectively.

We derive the diffusion coefficients based on a random walk assumption (Schirmacher, 2015):

$$K_{x \text{ or } y} = \frac{1}{2} \overline{V_{x \text{ or } y}^2} t_0, \quad (\text{A2})$$

$\overline{V_{x \text{ or } y}^2}$ is the deviation of wind speed in the zonal or meridional direction. t_0 is 3 h, the sampling interval of ERA5 wind data. Figure A1 shows the time-averaged wind vector and the distribution of K_x and K_y . The relative uncertainty in wind speed is assumed to be 10 %, close to the temporal standard deviation of wind speed. The uncertainties of K_x and K_y are set to be 20 %, about twice of the relative uncertainty in wind speed. The calculated K_x ranges from 30 397 m² s⁻² over land to 203 783 m² s⁻² over sea. The K_y ranges from 25 811 m² s⁻² over land to 297 053 m² s⁻² over sea. These diffusion coefficients tend to be slightly underestimated, because the variabilities of wind speed at higher frequencies (than 3-hourly) are not accounted for. This means that PHLET may underestimate the horizontal transport slightly.

Appendix B: Satellite conversion matrix to account for the smoothing effect of satellite pixels

The SCM is essentially a tool to preform quick conversion between grids, regular or not. In the YRD area, there are 100 × 100 = 10 000 grid cells on the 0.05° × 0.05° grid. We use the SCM (**A** matrix: [10 000, 10 000]) to convert from its original grid (**X** vector: [10 000, 1]) to the final grid (**Y** vector: [10 000, 1]), i.e., **Y** = **A****X**. The 10 000 elements in one specific row of **A** represent the weights of the 10 000 elements of **X** to an element in **Y**. Apparently, **A** is a sparse matrix. The following description shows how **A** is constructed.

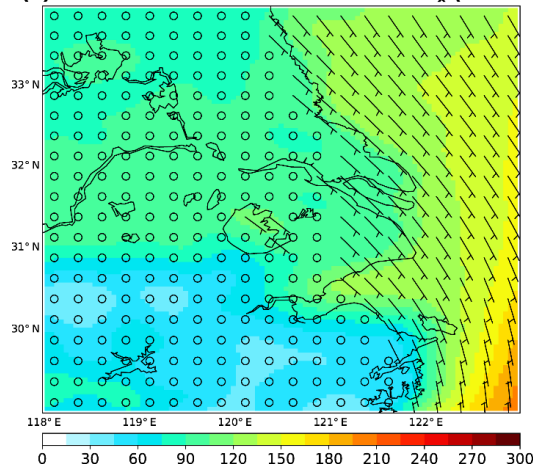
First, the VCDs specific to satellite pixels are reconstructed from the model grid cells. Each model grid cell (MGC) is divided into 10 × 10 = 100 finer grid cells (FGCs), each having the same area. Suppose the number of MGCs fully or partially covered by a given pixel p is N_c , and the number of FGCs in a given MGC i covered by p is g_i^p , then the total number of FGCs covered by p is

$$G^p = \sum_{i=1}^{N_c} g_i^p. \quad (\text{B1})$$

Thus, the average VCD for the pixel p can be reconstructed as follows:

$$\text{VCD}^p = \sum_{i=1}^{N_c} \frac{g_i^p}{G^p} \cdot \text{VCD}_i. \quad (\text{B2})$$

(a) The effective diffusion coefficient K_x (10³ ms)



(b) The effective diffusion coefficient K_y (10³ ms)

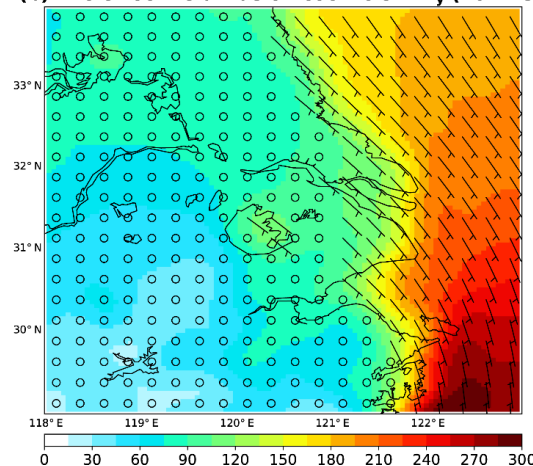


Figure A1. The effective diffusion coefficients for the summer months of 2012–2015 on a 0.05° × 0.05° grid. Overlaid is the temporal mean wind vector, which is plotted for every 5 × 5 = 25 grid cells to enhance the readability.

Equation (B2) essentially means how a satellite pixel smooths the VCD. The blue portion of Fig. B1 denotes the projection from MGC i to pixel p .

The next step represents how the oversampling approach is applied to satellite-smoothed VCD data. Suppose the number of satellite pixels fully or partially covering an MGC j is N_p , then the total number of FGCs being part of the intersection of the N_p pixels and MGC j is

$$G_j = \sum_{p=1}^{N_p} g_j^p. \quad (\text{B3})$$

Finally, the average VCD for the MGC j converted from the N_p pixels is

$$\text{VCD}'_j = \sum_{p=1}^{N_p} \frac{g_j^p}{G_j} \cdot \text{VCD}^p \quad (\text{B4})$$

The pink portion of Fig. B1 denotes the projection from pixel p to MGC j . Thus, the element of SCM converting from

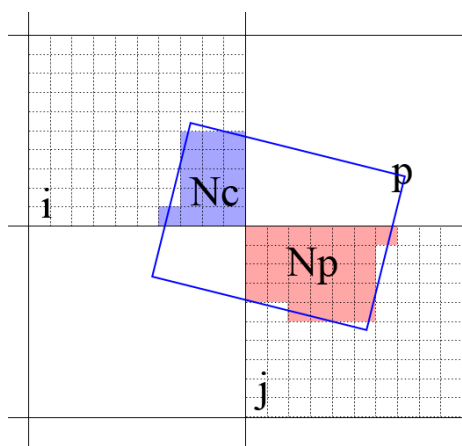


Figure B1. Illustration of how the original PHLET model grid cell i is projected to the satellite pixel p and then to the final grid cell j through the SCM approach. The size of the satellite pixel is scaled down to be comparable with the size of a model grid cell for illustration purposes.

MGC i to MGC j can be derived as follows:

$$A_{j,i} = \sum_{p=1}^{N_p} \frac{g_j^p}{G_j} \cdot \frac{g_i^p}{G^p}. \quad (\text{B5})$$

Appendix C: Deriving NO_x lifetime from VCD

We assume a steady state of radicals (HO_x), where the production rate of HO_x is equal to the loss rate through three types of termination reactions: between the hydroxyl radical (OH) and NO₂, between NO and peroxy radicals to form organic nitrates, and between peroxy radicals (Murphy et al., 2006; Valin et al., 2011):

$$P(\text{HO}_x) = k_1 C_{\text{OH}} C_{\text{NO}_2} + \alpha k_{2\text{eff}} \frac{k_1 C_{\text{NMVOC}} C_{\text{OH}}}{k_{2\text{eff}}} \\ C_{\text{NO}} C_{\text{NO}} + 6k_{3\text{eff}} \left(\frac{k_1 C_{\text{NMVOC}} C_{\text{OH}}}{k_{2\text{eff}}} C_{\text{NO}} \right)^2. \quad (\text{C1})$$

Here, $P(\text{HO}_x)$ is the production rate, and the right-hand side of Eq. (C1) is the loss rate. C_{NO_2} and C_{OH} denote the concentrations of NO₂ and OH, respectively. Since the conversion between the peroxy radicals (HO₂ + RO₂) and OH is in steady state, the term $\frac{k_1 C_{\text{NMVOC}} C_{\text{OH}}}{k_{2\text{eff}} C_{\text{NO}}}$ expresses the “effective” total concentration of peroxy radicals in terms of the concentrations of NMVOC, OH, and NO. Assuming $P(\text{HO}_x)$, C_{NMVOC} and all reaction constants to be constant (Valin et al., 2011), and given that $C_{\text{NO}} = C_{\text{NO}_2} \cdot \frac{1-r}{r}$, Eq. (C1) can be simplified as Eq. (C2):

$$a C_{\text{OH}} C_{\text{NO}_2} + b C_{\text{OH}} + c \left(\frac{C_{\text{OH}}}{C_{\text{NO}_2}} \right)^2 = 1. \quad (\text{C2})$$

Here, a , b , and c are the coefficients. Because the chemical lifetimes of NO₂ is determined by C_{OH} ($\tau_c = \frac{1}{k_1 C_{\text{OH}}}$) we can

deduce the relationship between C_{NO_2} and τ_c :

$$\frac{a'}{\tau_c} + \frac{b'}{\tau_c \cdot C_{\text{NO}_2}} + c' \left(\frac{1}{\tau_c \cdot C_{\text{NO}_2}} \right)^2 = -1. \quad (\text{C3})$$

NO₂ is lost primarily through reaction with OH and secondarily through dry deposit ($\frac{C_{\text{NO}_2}}{\tau_d}$); thus, its lifetime (τ) is also determined by these two loss processes. Therefore,

$$\tau_c = \frac{1}{\frac{1}{\tau} - \frac{1}{\tau_d}}. \quad (\text{C4})$$

In the areas of low emissions, the emission term can be neglected in Eq. (2); thus, the local net source $L = -\frac{C_{\text{NO}_2}}{r \cdot \tau}$. Therefore, Eq. (C3) becomes Eq. (C5), which connects L and C_{NO_2} .

$$a' r \frac{(L + k C_{\text{NO}_2})}{C_{\text{NO}_2}} + b' r \frac{(L + k C_{\text{NO}_2})}{C_{\text{NO}_2}^2} \\ + c' \left(\frac{r(L + k C_{\text{NO}_2})}{C_{\text{NO}_2}^3} \right)^2 = 1, \quad (\text{C5})$$

where $k = \frac{1}{\tau_d}$. We determine the coefficients a' , b' , c' , and k in Eq. (C5) by conducting nonlinear fitting of OMI NO₂ VCD data and L values in the low-emission areas (see below). This procedure establishes the nonlinear relationship between τ and VCD, which is then applied to the entire study domain.

The low-emission areas have small values of VCD and large negative values of L . Figure C1 shows a scatter plot for the derived local net source L and VCD at each individual grid cell of the study domain. The data scatter reflects the combined effect of emission, loss, and horizontal transport. We then fit the quantiles of L where the VCD is relatively low ($< 5 \times 10^{15}$ molec. cm⁻², shown as blue points in Fig. C1) into Eq. (C5) through a nonlinear quantile fitter based on Tensorflow (Abadi et al., 2016). Using the quantile fitting also means that the low-emission grid cells do not need to be explicitly identified prior to the fitting. The quantile fitting gives L as a function of VCD (when emissions are neglected), through which the relationship between lifetimes and VCDs is derived. We conduct the fitting by 50 times, each by linearly changing the assumed percentile threshold of L from 0.1 % to 5 %, to determine the fitted median value (red line in Fig. C1) and uncertainty (grey shaded areas, 95 % CI). The uncertainty is caused by the assumption on the NO₂/NO_x ratio r , the simplification of the relationship between lifetimes and VCDs, and possible misjudgment of low-emission areas.

The orange line in Fig. C1 presents the relationship between NO₂ VCDs and chemical lifetimes (τ_c) derived based on the mean value of the fitting. The value of τ_c varies from 0.6 to 3.8 h with an average of 1.2 h. The lifetimes decline rapidly with increasing VCDs from 0 to 2×10^{15} molec. cm⁻², and then grows gradually with increasing

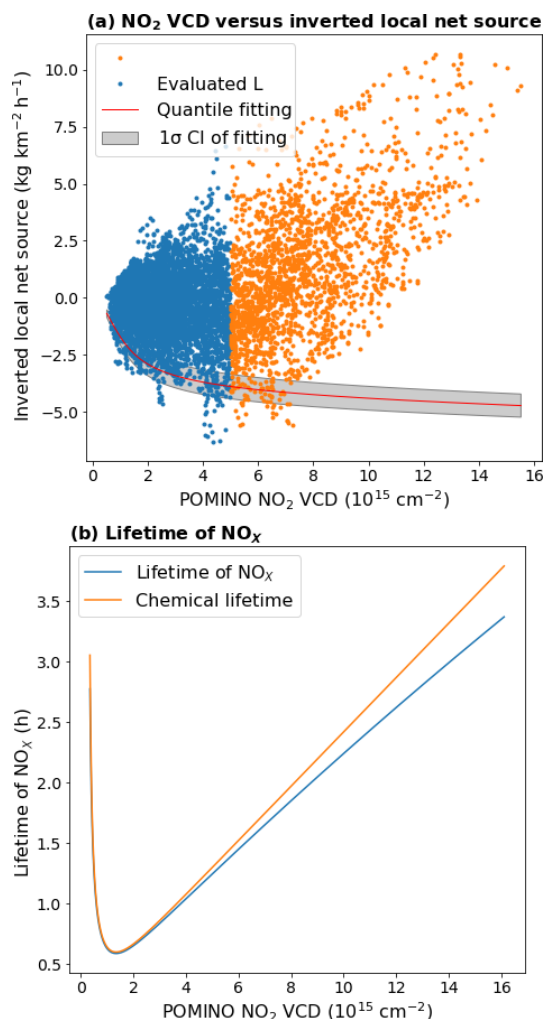


Figure C1. (a) Scatter plot for POMINO NO_2 VCDs (x axis) and derived local net sources (y axis) across individual grid cells. Grid cells with NO_2 VCDs below (above) $5 \times 10^{15} \text{ molec. cm}^{-2}$ are colored in blue (orange). The red line and shading denote the median and uncertainty (1σ CI) of the quantile fitting, respectively, to estimate the nonlinear relationship between NO_2 VCD and NO_x lifetime, based on data in the low-emission areas. (b) The derived relationship between NO_2 VCD and NO_x lifetime across the range of NO_2 VCDs in the YRD area.

VCDs. This result is consistent with Valin et al. (2011). By comparison, the value of τ_d is 30.4 h and is spatially homogeneous under the assumption here. The total lifetime (τ) varies from 0.6 to 3.3 h (Fig. C1, blue line) across the study domain.

Author contributions. JL conceived the research. HK, RZ, and JL designed the research. HK and RZ performed the data processing, model development, and simulations. ML, HW, LC, RN, JW, and YY contributed to data processing, model simulations, and data analyses. QZ provided MEIC data. HK and JL analyzed the results and wrote the paper with input from all authors.

Competing interests. The authors declare that they have no conflict of interest.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant no. 41775115) and the 973 program (grant no. 2014CB441303).

Review statement. This paper was edited by Jianping Huang and reviewed by two anonymous referees.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, arXiv:1603.04467, 2016.
- Alnaes, M. S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M. E., and Wells, G. N.: The FEniCS Project Version 1.5, Archive of Numerical Software, 3, 9–23, <https://doi.org/10.11588/ans.2015.100.20553>, 2015.
- Armijo, L.: Minimization of functions having Lipschitz continuous first partial derivatives, *Pac. J. Math.*, 16, 1–3, <https://doi.org/10.2140/pjm.1966.16.1>, 1966.
- Barnes, I. and Rudziński, K. J.: Disposal of Dangerous Chemicals in Urban Areas and Mega Cities, Springer Netherlands, 2013.
- Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, *Science*, 333, 1737–1739, <https://doi.org/10.1126/science.1207824>, 2011.
- Beirle, S., Sihler, H., and Wagner, T.: Mapping NO_x Sources on High Spatial Resolution by Combined Measurements from OMI and GOME-2 Narrow Swath Mode, ATMOS 2015, Proceedings of the Conference, Crete, 8–12 June 2015.
- 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 NO₂ column retrieval algorithm for the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 4, 1905–1928, <https://doi.org/10.5194/amt-4-1905-2011>, 2011.
- Boersma, K. F., Vinken, G., and Tournadre, J.: Ships going slow in reducing their NO_x emissions: Changes in 2005–2012 ship exhaust inferred from satellite measurements over Europe, *Environ. Res. Lett.*, 10, 074007, <https://doi.org/10.1088/1748-9326/10/7/074007>, 2015.
- Bousserez, N., Henze, D. K., Perkins, A., Bowman, K. W., Lee, M., Liu, J., Deng, F., and Jones, D. B. A.: Improved analysis-error covariance matrix for high-dimensional variational inversions: application to source estimation using a 3D atmospheric transport model, *Q. J. Roy. Meteor. Soc.*, 141, 1906–1921, <https://doi.org/10.1002/qj.2495>, 2015.
- Brasseur, G. P. and Jacob, D. J.: Modeling of atmospheric chemistry, Cambridge University Press, Cambridge, 2017.
- Cao, H., Fu, T.-M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., De Smedt, I., Zhang, Q., van Roozendaal, M., Hendrick, F., Chance, K., Li, J., Zheng, J., and Zhao, Y.: Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, *Atmos. Chem. Phys.*, 18, 15017–15046, <https://doi.org/10.5194/acp-18-15017-2018>, 2018.
- Cui, Y., Lin, J., Song, C., Liu, M., Yan, Y., Xu, Y., and Huang, B.: Rapid growth in nitrogen dioxide pollution over Western China, 2005–2013, *Atmos. Chem. Phys.*, 16, 6207–6221, <https://doi.org/10.5194/acp-16-6207-2016>, 2016.
- de Graaf, M., Sihler, H., Tilstra, L. G., and Stammes, P.: How big is an OMI pixel?, *Atmos. Meas. Tech.*, 9, 3607–3618, <https://doi.org/10.5194/amt-9-3607-2016>, 2016.
- De Smedt, I., Stavrou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendaal, M.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, *Atmos. Chem. Phys.*, 15, 12519–12545, <https://doi.org/10.5194/acp-15-12519-2015>, 2015.
- Ding, J., Miyazaki, K., van der A, R. J., Mijling, B., Kurokawa, J.-I., Cho, S., Janssens-Maenhout, G., Zhang, Q., Liu, F., and Levelt, P. F.: Intercomparison of NO_x emission inventories over East Asia, *Atmos. Chem. Phys.*, 17, 10125–10141, <https://doi.org/10.5194/acp-17-10125-2017>, 2017a.
- Ding, J., van der A, R. J., Mijling, B., and Levelt, P. F.: Space-based NO_x emission estimates over remote regions improved in DECSO, *Atmos. Meas. Tech.*, 10, 925–938, <https://doi.org/10.5194/amt-10-925-2017>, 2017b.
- Elvidge, C., Hsu, F. C., Baugh, K. E., and Ghosh, T.: National Trends in Satellite Observed Lighting: 1992–2012, in: Global Urban Monitoring and Assessment Through Earth Observation, CRC Press, Boca Raton, in press, 2013.
- Endresen, Ø., Sørsgård, E., Sundet, J. K., Dalsøren, S. B., Isaksen, I. S. A., Berglen, T. F., and Gravir, G.: Emission from international sea transportation and environmental impact, *J. Geophys. Res.*, 108, 4560, <https://doi.org/10.1029/2002jd002898>, 2003.
- Eskes, H. J., Velthoven, P. F. J. V., Valks, P. J. M., and Kelder, H. M.: Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model, *Q. J. Roy. Meteor. Soc.*, 129, 1663–1681, <https://doi.org/10.1256/qj.02.14>, 2003.
- Farrell, P. E., Ham, D. A., Funke, S. F., and Rognes, M. E.: Automated derivation of the adjoint of high-level transient finite element programs, *Comput. Sci.*, 35, C369–C393, <https://doi.org/10.1137/120873558>, 2012.
- Fioletov, V. E., McLinden, C. A., Krotkov, N., Moran, M. D., and Yang, K.: Estimation of SO₂ emissions using OMI retrievals, *Geophys. Res. Lett.*, 38, L21811, <https://doi.org/10.1029/2011gl049402>, 2011.
- Fu, M., Liu, H., Jin, X., and He, K.: National- to port-level inventories of shipping emissions in China, *Environ. Res. Lett.*, 12, 114024, <https://doi.org/10.1088/1748-9326/aa897a>, 2017.

- Fu, Q., Shen, Y., and Zhang, J.: On the ship pollutant emission inventory in Shanghai port, *Journal of Safety and Environment*, 12, 57–64, 2012.
- Funke, S. W. and Farrell, P. E.: A framework for automated PDE-constrained optimisation, *Computer Science*, arXiv:1302.3894, 2013.
- Geng, G., Zhang, Q., Martin, R. V., Lin, J., Huo, H., Zheng, B., Wang, S., and He, K.: Impact of spatial proxies on the representation of bottom-up emission inventories: A satellite-based analysis, *Atmos. Chem. Phys.*, 17, 4131–4145, <https://doi.org/10.5194/acp-17-4131-2017>, 2017.
- Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, *Biogeosciences*, 7, 1171–1186, <https://doi.org/10.5194/bg-7-1171-2010>, 2010.
- Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the Fourth-Generation Global Fire Emissions Database (GFED4), *J. Geophys. Res.-Biogeo.*, 118, 317–328, <https://doi.org/10.1002/jgrg.20042>, 2013.
- Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Lioussé, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version), Copernicus Atmosphere Monitoring Service (CAMS) report, <https://doi.org/10.24380/d0bn-kx16>, 2019.
- Gu, D., Wang, Y., Smeltzer, C., and Boersma, K. F.: Anthropogenic emissions of NO_x over China: Reconciling the difference of inverse modeling results using GOME-2 and OMI measurements, *J. Geophys. Res.-Atmos.*, 119, 7732–7740, <https://doi.org/10.1002/2014jd021644>, 2014.
- Gu, D., Wang, Y., Yin, R., Zhang, Y., and Smeltzer, C.: Inverse modelling of NO_x emissions over eastern China: uncertainties due to chemical non-linearity, *Atmos. Meas. Tech.*, 9, 5193–5201, <https://doi.org/10.5194/amt-9-5193-2016>, 2016.
- Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., and Li, Q.: The Adjoint of CMAQ, *Environ. Sci. Technol.*, 41, 7807–7817, <https://doi.org/10.1021/es070944p>, 2007.
- Hersbach, H. and Dee, D.: ERA5 reanalysis is in production, ECMWF Newsletter No. 147, 7, available at: <https://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production> (last access: 30 August 2019), 2016.
- Hooyberghs, H., Veldeman, N., and Maiheu, B.: Marco Polo Emission Inventory for EastChina: Basic Description, http://www.marcopolo-panda.eu/wp/internal/wp-content/uploads/2016/10/D4.2_DescriptionMarcoPoloInventory.pdf (last access: 30 August 2019), 2016.
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier van der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S., Koffi, B., and Li, M.: HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, *Atmos. Chem. Phys.*, 15, 11411–11432, <https://doi.org/10.5194/acp-15-11411-2015>, 2015.
- Keiyo, Y. and Itsushi, U.: Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation, *Atmos. Environ.*, 40, 6836–6845, <https://doi.org/10.1016/j.atmosenv.2006.05.042>, 2006.
- Krotkov, N. A., McLinden, C. A., Li, C., Lamsal, L. N., Celarier, E. A., Marchenko, S. V., Swartz, W. H., Bucsela, E. J., Joiner, J., Duncan, B. N., Boersma, K. F., Veefkind, J. P., Levelt, P. F., Fioletov, V. E., Dickerson, R. R., He, H., Lu, Z., and Streets, D. G.: Aura OMI observations of regional SO₂ and NO₂ pollution changes from 2005 to 2015, *Atmos. Chem. Phys.*, 16, 4605–4629, <https://doi.org/10.5194/acp-16-4605-2016>, 2016.
- Lamsal, L. N., Martin, R. V., Padmanabhan, A., van Donkelaar, A., Zhang, Q., Sioris, C. E., Chance, K., Kurosu, T. P., and Newchurch, M. J.: Application of satellite observations for timely updates to global anthropogenic NO_x emission inventories, *Geophys. Res. Lett.*, 38, L05810, <https://doi.org/10.1029/2010gl046476>, 2011.
- Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Huib, V., Johan de, V., Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, *IEEE T. Geosci. Remote*, 44, 1093–1101, <https://doi.org/10.1109/tgrs.2006.872333>, 2006.
- Li, D.-H. and Fukushima, M.: A modified BFGS method and its global convergence in nonconvex minimization, *J. Comput. Appl. Math.*, 129, 15–35, 2001.
- Lin, J., Tong, D., Davis, S., Ni, R., Tan, X., Pan, D., Zhao, H., Lu, Z., Streets, D., Feng, T., Zhang, Q., Yan, Y., Hu, Y., Li, J., Liu, Z., Jiang, X., Geng, G., He, K., Huang, Y., and Guan, D.: Global climate forcing of aerosols embodied in international trade, *Nat. Geosci.*, 9, 790–794, <https://doi.org/10.1038/ngeo2798>, 2016.
- 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, <https://doi.org/10.5194/acp-12-2881-2012>, 2012.
- 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, <https://doi.org/10.5194/acp-10-63-2010>, 2010.
- Lin, J.-T., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Van Roozendaal, M., Clémer, K., and Irie, H.: 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, <https://doi.org/10.5194/acp-14-1441-2014>, 2014.
- Lin, J.-T., Liu, M.-Y., Xin, J.-Y., Boersma, K. F., Spurr, R., Martin, R., and Zhang, Q.: Influence of aerosols and surface reflectance on satellite NO₂ retrieval: seasonal and spatial characteristics and implications for NO_x emission constraints, *Atmos. Chem. Phys.*, 15, 11217–11241, <https://doi.org/10.5194/acp-15-11217-2015>, 2015.
- 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, <https://doi.org/10.5194/acp-15-13299-2015>, 2015.
- 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, *At-*

- mos. Chem. Phys., 16, 5283–5298, <https://doi.org/10.5194/acp-16-5283-2016>, 2016.
- Liu, F., Beirle, S., Zhang, Q., van der A, R. J., Zheng, B., Tong, D., and He, K.: NO_x emission trends over Chinese cities estimated from OMI observations during 2005 to 2015, *Atmos. Chem. Phys.*, 17, 9261–9275, <https://doi.org/10.5194/acp-17-9261-2017>, 2017.
- Liu, M., Lin, J., Wang, Y., Sun, Y., Zheng, B., Shao, J., Chen, L., Zheng, Y., Chen, J., Fu, T.-M., Yan, Y., Zhang, Q., and Wu, Z.: Spatiotemporal variability of NO₂ and PM_{2.5} over Eastern China: observational and model analyses with a novel statistical method, *Atmos. Chem. Phys.*, 18, 12933–12952, <https://doi.org/10.5194/acp-18-12933-2018>, 2018.
- Liu, M., Lin, J., Boersma, K. F., Pinardi, G., Wang, Y., Chimot, J., Wagner, T., Xie, P., Eskes, H., Van Roozendaal, M., Hendrick, F., Wang, P., Wang, T., Yan, Y., Chen, L., and Ni, R.: Improved aerosol correction for OMI tropospheric NO₂ retrieval over East Asia: constraint from CALIOP aerosol vertical profile, *Atmos. Meas. Tech.*, 12, 1–21, <https://doi.org/10.5194/amt-12-1-2019>, 2019a.
- Liu, M., Lin, J., Boersma, K. F., Pinardi, G., Wang, Y., Chimot, J., Wagner, T., Xie, P., Eskes, H., Van Roozendaal, M., Hendrick, F., Wang, P., Wang, T., Yan, Y., Chen, L., and Ni, R.: Improved aerosol correction for OMI tropospheric NO₂ retrieval over East Asia: constraint from CALIOP aerosol vertical profile, *Atmospheric Measurement Techniques*, 12, 1–21, <https://doi.org/10.5194/amt-12-1-2019>, 2019b.
- Ma, J. and Xiuji, Z.: Development of a three-dimensional inventory of aircraft NO_x emissions over China, *Atmos. Environ.*, 34, 389–396, [https://doi.org/10.1016/S1352-2310\(99\)00309-X](https://doi.org/10.1016/S1352-2310(99)00309-X), 2000.
- Maiheu, B. and Veldeman, N.: NO_x Emission Downscaling for South Africa, available at: http://globemission.marvin.vito.be/southafrica/docu/GlobEmission_Downscaling_SA-v1.0.pdf (last access: 30 August 2019), 2013.
- Marchuk, G. L.: *Adjoint Equations and Analysis of Complex Systems*, Russian Academy of Sciences, Moscow, Russia, 475 pp., 1994.
- Martien, P. T., Harley, R. A., and Cacuci, D. G.: Adjoint Sensitivity Analysis for a Three-Dimensional Photochemical Model: Implementation and Method Comparison, *Environ. Sci. Technol.*, 40, 2663–2670, <https://doi.org/10.1021/es0510257>, 2006.
- 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 NO₂ columns, *J. Geophys. Res.*, 108, 4537, <https://doi.org/10.1029/2003JD003453>, 2003.
- Mijling, B. and van der A, R. J.: Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale, *J. Geophys. Res.*, 117, D17302, <https://doi.org/10.1029/2012JD017817>, 2012.
- Mijling, B., van der A, R. J., and Zhang, Q.: Regional nitrogen oxides emission trends in East Asia observed from space, *Atmos. Chem. Phys.*, 13, 12003–12012, <https://doi.org/10.5194/acp-13-12003-2013>, 2013.
- Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO₂ columns, *Atmos. Chem. Phys.*, 12, 2263–2288, <https://doi.org/10.5194/acp-12-2263-2012>, 2012.
- Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal changes in global surface NO_x emissions from multi-constituent satellite data assimilation, *Atmos. Chem. Phys.*, 17, 807–837, <https://doi.org/10.5194/acp-17-807-2017>.
- Murphy, J. G., Day, D. A., Cleary, P. A., Wooldridge, P. J., Millet, D. B., Goldstein, A. H., and Cohen, R. C.: The weekend effect within and downwind of Sacramento: Part 2. Observational evidence for chemical and dynamical contributions, *Atmos. Chem. Phys. Discuss.*, 6, 11971–12019, <https://doi.org/10.5194/acpd-6-11971-2006>, 2006.
- NBSC (National Bureau of Statistics of China): *Statistical Communiqué of the People's Republic of China on the 2009 National Economic and Social Development*, 2009.
- NBSC (National Bureau of Statistics of China): *Statistical Communiqué of the People's Republic of China on the 2013 National Economic and Social Development*, 2013.
- NBSC (National Bureau of Statistics of China): *Statistical Communiqué of the People's Republic of China on the 2017 National Economic and Social Development*, 2017.
- NBSC (National Bureau of Statistics of China): *China Statistical Yearbook 2007–2017*.
- Ni, R., Lin, J., Yan, Y., and Lin, W.: Foreign and domestic contributions to springtime ozone over China, *Atmos. Chem. Phys.*, 18, 11447–11469, <https://doi.org/10.5194/acp-18-11447-2018>, 2018.
- Sandu, A., Daescu, D. N., Carmichael, G. R., and Chai, T.: Adjoint sensitivity analysis of regional air quality models, *J. Comput. Phys.*, 204, 222–252, <https://doi.org/10.1016/j.jcp.2004.10.011>, 2005.
- Schirmacher, W.: *Theory of Liquids and Other Disordered Media*, Springer International Publishing, New York City, 2015.
- Seinfeld, J. H. and Pandis, S. N.: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 2nd Edition, John Wiley & Sons, New York, 429–443, 2006.
- Stavrakou, T., Muller, J. F., Boersma, K. F., De Smedt, I., and van der A, R. J.: Assessing the distribution and growth rates of NO_x emission sources by inverting a 10-year record of NO₂ satellite columns, *Geophys. Res. Lett.*, 35, L10801, <https://doi.org/10.1029/2008GL033521>, 2008.
- Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T., and Zhang, Q.: Key chemical NO_x sink uncertainties and how they influence top-down emissions of nitrogen oxides, *Atmos. Chem. Phys.*, 13, 9057–9082, <https://doi.org/10.5194/acp-13-9057-2013>, 2013.
- Sun, K., Zhu, L., Cady-Pereira, K., Chan Miller, C., Chance, K., Clarisse, L., Coheur, P.-F., González Abad, G., Huang, G., Liu, X., Van Damme, M., Yang, K., and Zondlo, M.: A physics-based approach to oversample multi-satellite, multispecies observations to a common grid, *Atmos. Meas. Tech.*, 11, 6679–6701, <https://doi.org/10.5194/amt-11-6679-2018>, 2018.
- 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 NO₂ observations, *Atmos. Chem. Phys.*, 13, 11005–11018, <https://doi.org/10.5194/acp-13-11005-2013>, 2013.
- Timmermans, R., Kranenburg, R., and Hooyberghs, H.: MarcoPolo project, Deliverable 4.3, available at: <http://www.atmos-chem-phys.net/19/12835/2019/>

- [//globemission.marvin.vito.be/southafrica/docu/GlobEmission_Downscaling_SA-v1.0.pdf](http://globemission.marvin.vito.be/southafrica/docu/GlobEmission_Downscaling_SA-v1.0.pdf) (last access: 30 August 2019), 2016.
- Valin, L. C., Russell, A. R., Hudman, R. C., and Cohen, R. C.: Effects of model resolution on the interpretation of satellite NO₂ observations, *Atmos. Chem. Phys.*, 11, 11647–11655, <https://doi.org/10.5194/acp-11-11647-2011>, 2011.
- Valin, L. C., Russell, A. R., and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO₂ column measurements, *Geophys. Res. Lett.*, 40, 1856–1860, <https://doi.org/10.1002/grl.50267>, 2013.
- Wang, Y., Jacob, D. J., and Logan, J. A.: Global simulation of tropospheric O₃–NO_x–hydrocarbon chemistry: 1. Model formulation, *J. Geophys. Res.*, 103, 10713–10725, <https://doi.org/10.1029/98JD00158>, 1998.
- Yan, Y.-Y., Lin, J.-T., Chen, J., and Hu, L.: Improved simulation of tropospheric ozone by a global-multi-regional two-way coupling model system, *Atmospheric Chemistry and Physics*, 16, 2381–2400, <https://doi.org/10.5194/acp-16-2381-2016>, 2016.
- Yienger, J. J. and Levy, H.: Empirical model of global soil-biogenic NO_x emissions, *J. Geophys. Res.*, 100, 11447–11464, <https://doi.org/10.1029/95JD00370>, 1995.
- Zhang, R., Lin, J., and Zhang, Q.: Satellite-based evaluation of nitrogen oxides emission inventory for Beijing at a high resolution, bachelor thesis, Peking University, China, unpublished, 19 pp., 2014.
- Zhao, Y., Nielsen, C. P., Lei, Y., McElroy, M. B., and Hao, J.: Quantifying the uncertainties of a bottom-up emission inventory of anthropogenic atmospheric pollutants in China, *Atmos. Chem. Phys.*, 11, 2295–2308, <https://doi.org/10.5194/acp-11-2295-2011>, 2011.
- Zhao, Y., Qiu, L. P., Xu, R. Y., Xie, F. J., Zhang, Q., Yu, Y. Y., Nielsen, C. P., Qin, H. X., Wang, H. K., Wu, X. C., Li, W. Q., and Zhang, J.: Advantages of a city-scale emission inventory for urban air quality research and policy: the case of Nanjing, a typical industrial city in the Yangtze River Delta, China, *Atmos. Chem. Phys.*, 15, 12623–12644, <https://doi.org/10.5194/acp-15-12623-2015>, 2015.
- 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, <https://doi.org/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, <https://doi.org/10.5194/acp-17-921-2017>, 2017, 2017.