

# Satellite constraint for emissions of nitrogen oxides from anthropogenic, lightning and soil sources over East China on a high-resolution grid

J.-T. Lin

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

Correspondence to: J.-T. Lin ([linjt@pku.edu.cn](mailto:linjt@pku.edu.cn))

## Abstract

Vertical column densities (VCDs) of tropospheric nitrogen dioxide (NO<sub>2</sub>) retrieved from space provide valuable information to estimate emissions of nitrogen oxides (NO<sub>x</sub>) inversely. Accurate emission attribution to individual sources, important both for understanding the global biogeochemical cycling of nitrogen and for emission control, remains difficult. This study presents a regression-based multi-step inversion approach to estimate emissions of NO<sub>x</sub> from anthropogenic, lightning and soil sources individually for 2006 over East China on a 0.25 °long x 0.25 °lat grid, employing the DOMINO product version 2 retrieved from the Ozone Monitoring Instrument. The nested GEOS-Chem model for East Asia is used to simulate the seasonal variations of different emission sources and impacts on VCDs of NO<sub>2</sub> for the inversion purpose. Sensitivity tests are conducted to evaluate key assumptions embedded in the inversion process. The inverse estimate suggests annual budgets of about 7.1 TgN (± 39%), 0.22 TgN (± 47%), and 0.40 TgN (± 47%) for the a posteriori anthropogenic, lightning and soil emissions, respectively, each about 24% higher than the respective a priori values. The enhancements in anthropogenic emissions are largest in cities and areas with extensive use of coal, particularly in the north in winter, as evident on the high-resolution grid. Derived soil emissions are consistent with recent bottom-up estimates. They are less than 6% of anthropogenic emissions annually, increasing to about 13% for July. Derived lightning emissions are about 3% of anthropogenic emissions annually and about 10% in July. Overall,

anthropogenic emissions are found to be the dominant source of NO<sub>x</sub> over East China with important implications for nitrogen control.

## **1 Introduction**

Nitrogen oxides (NO<sub>x</sub> ≡ NO + NO<sub>2</sub>) are important constituents in the troposphere affecting the formation of ozone and aerosols with significant consequences on air quality, climate forcing and acid deposition. They are emitted from anthropogenic combustion sources as well as natural sources from lightning, soil and biomass burning. Understanding the individual contributions of anthropogenic and natural emissions is critical both for evaluating the effects of NO<sub>x</sub> on the global environment and for forming appropriate emission control strategies in polluted areas like East China.

Satellite remote sensing provides data for vertical column densities (VCDs) of nitrogen dioxide (NO<sub>2</sub>) in the troposphere. The data can be used to estimate emissions of NO<sub>x</sub> from the top-down perspective supplementing the bottom-up estimate (e.g., Martin et al., 2003; Jaeglé et al., 2005; Müller and Stavrou, 2005; Richter et al., 2005; Martin et al., 2006; Wang et al., 2007; Zhang et al., 2007; Stavrou et al., 2008; Zhao and Wang, 2009; Lin et al., 2010a,b; Lin and McElroy, 2010; Lin and McElroy, 2011; Wang et al., 2012). Much attention has been paid to China due to its large and rapidly growing anthropogenic emissions (e.g., Richter et al., 2005; Martin et al., 2006; Wang et al., 2007; Zhang et al., 2007; Stavrou et al., 2008; Zhao and Wang, 2009; Lin et al., 2010a,b; Lin and McElroy, 2010; Lin and McElroy, 2011; Wang et al., 2012). Retrieved VCDs, however, include contributions from all sources of NO<sub>x</sub>, requiring additional information and assumptions in distinguishing individual sources.

Several inversion studies have attempted to separate anthropogenic from other sources of NO<sub>x</sub>, especially soil sources, based on measurements from the Global Ozone Monitoring Experiment (GOME) instrument (Jaeglé et al., 2005; Müller and Stavrou, 2005; Wang et al., 2007; Stavrou et al., 2008), the Scanning Imaging Absorption Spectrometer for Atmospheric CHartography (SCIAMACHY) instrument (Müller and Stavrou, 2005; Stavrou et al., 2008), and the Ozone Monitoring Instrument (OMI) (Zhao and Wang, 2009). Jaeglé et al. (2005) and Wang et al. (2007) proposed two different methods to separate anthropogenic, soil and biomass burning emissions month by month with no attempt to constrain lightning emissions. Jaeglé et al. (2005) assumed the a posteriori non-lightning emissions to be solely anthropogenic if the a priori anthropogenic emissions exceed 90% of

the a priori total emissions or if they exceed the a posteriori non-lightning emissions. Otherwise, differences between the a posteriori and a priori emissions were attributed to soil or biomass burning sources. A similar criterion was adopted by Zhao and Wang (2009) to differentiate anthropogenic and soil emissions. Wang et al. (2007) distinguished anthropogenic and soil sources using prescribed values for errors in the a priori anthropogenic emission data. Specifically, if the a posteriori non-lightning non-biomass burning emissions exceed the a priori anthropogenic emissions plus errors (assumed to be 40-60%), the differences are attributed to soil sources. Müller and Stavrakou (2005) and Stavrakou et al. (2008) used an adjoint modeling approach for source attribution. Wang et al. (2007) suggested that soil emissions over East China amounted to 0.85 TgN per year for 1997-2000, differing significantly from other inverse estimates (Jaeglé et al., 2005; Müller and Stavrakou, 2005; Stavrakou et al., 2008; Zhao and Wang, 2009; Lyatt Jaeglé personal communication, 2011; Chun Zhao and Yuhang Wang, personal communication, 2011). Soil emissions derived from the inverse modeling also differ from the bottom-up estimates (Yienger and Levy, 1995; Yan et al., 2003; Yan et al., 2005; Hudman et al., 2012; Steinkamp and Lawrence, 2011); in particular, they are 50–300% larger than the Yienger and Levy (1995) estimate. According to Jaeglé et al. (2005) and Wang et al. (2007), soil emissions may be as large as 40-50% of anthropogenic emissions in summer for East Asia in 2000 and for East China in 1997-2000, respectively, with significant implications for the global biogeochemical cycling of nitrogen.

The magnitude of lightning emissions is difficult to estimate (Boersma et al., 2005; Schumann and Huntrieser, 2007), especially on the regional scale with significant variations in lightning occurrences from one year to another (Schumann and Huntrieser, 2007). Most inverse estimates did not attempt to constrain lightning emissions (Jaeglé et al., 2005; Wang et al., 2007; Zhao and Wang, 2009; Lin et al., 2010a,b; Lin and McElroy, 2010). The inverse modeling by Stavrakou et al. (2008) suggested lightning emissions to be 50-80% larger than their a priori values (3 TgN yr<sup>-1</sup> globally) with the largest difference over the tropics; the study did not specify the magnitude of lightning emissions over China.

This study presents a new method to inversely derive emissions of NO<sub>x</sub> for 2006 over East China (101.25 E–126.25 E, 19°–46 N; see Fig. 1) from anthropogenic, lightning and soil sources individually based on satellite retrievals of NO<sub>2</sub> columns and simulations of the global chemical transport model (CTM) GEOS-Chem. Emissions from biomass burning are not constrained since they are unimportant over East China (Wang et al., 2007; Lin et al.,

2010a). A regression-based multi-step inversion approach is used to derive emissions for all months, exploiting information on the seasonal variations of individual sources simulated by the CTM. The satellite data are taken from the DOMINO product version 2 (DOMINO-2) retrieved by the Royal Netherlands Meteorological Institute (KNMI) from the Ozone Monitoring Instrument (OMI) (Boersma et al., 2007; Boersma et al., 2011). The nested GEOS-Chem model for East Asia (Chen et al., 2009) is used to calculate VCDs of NO<sub>2</sub> in response to various emission sources for the inversion purpose. The top-down emissions are derived at a relatively fine resolution of 0.25 °long x 0.25 °lat allowing for a more detailed analysis of the spatial distribution of emissions, compared to previous studies (Jaeglé et al., 2005; Müller and Stavrakou, 2005; Wang et al., 2007; Stavrakou et al., 2008; Zhao and Wang, 2009).

The paper is organized as follows. Section 2 presents the satellite product. Section 3 describes the CTM and compares simulated VCDs with retrieved values. Section 4 describes the inversion process in detail and analyzes the resulting top-down emissions for anthropogenic, lightning and soil sources. It also evaluates the effects of key assumptions made during the inversion process. Section 5 presents the a posteriori emissions in comparison with previous inverse and bottom-up estimates. Section 6 concludes the present analysis.

## **2 VCDs of NO<sub>2</sub> retrieved from OMI**

The KNMI DOMINO-2 product offers a level-2 dataset for VCDs of NO<sub>2</sub> derived by three main steps involving the calculation of slant column densities (SCDs), tropospheric SCDs, and tropospheric VCDs (Boersma et al., 2007; Boersma et al., 2011). The derivation relies on information on air mass factors (AMFs) to convert the tropospheric SCDs to VCDs. The AMFs are interpolated from a look-up table (LUT), and are subject to errors in the predetermined information for clouds, aerosols, surface albedo, the a priori vertical profile of NO<sub>2</sub>, surface pressure, and surface height. The reader is referred to Boersma et al. (2007; 2011) for detailed derivation of the product.

Errors in retrieved VCDs are derived mainly from the calculation of SCDs and its tropospheric portion over cleaner regions, and are mainly from the calculation of AMFs for polluted regions (Boersma et al., 2007). Compared to version 1, DOMINO-2 incorporates a variety of improvements on the LUT, surface albedo, and the a priori vertical profile of NO<sub>2</sub> (Boersma et al., 2011). It also includes a cross-track stripe correction and a high-resolution

dataset for surface height. As a result, systematic biases found in version 1 are reduced significantly in DOMINO-2 (Boersma et al., 2011). The overall error for retrieved VCDs in DOMINO-2 is estimated to be about 30% (a relative error) plus  $0.7 \times 10^{15}$  molec.  $\text{cm}^{-2}$  (an absolute error), likely with a magnitude larger in winter than in summer (Boersma et al., 2007; Lin and McElroy, 2010; Lin et al., 2010a; Boersma et al., 2011; Lin and McElroy, 2011). In this study, the relative error is assumed to vary nonlinearly from 30% in summer to 50% in winter based on the following formula:  $0.3 + 0.2 * (1 - \sin(i/10 * \pi))$ , where  $i = 0, 1, 2, 3, 4, 5, 5, 4, 3, 2, 1, 0$  for months from January to December. This information will be employed for purposes of emission inversion; and the assumed seasonality will be evaluated in Sect. 4.6.

In this study, the daily level-2 data from DOMINO-2 are gridded to  $0.25^\circ \text{long} \times 0.25^\circ \text{lat}$ , which are averaged then to obtain monthly mean VCDs for subsequent emission inversion. The level-2 dataset includes measurements at 60 viewing angles corresponding to 60 ground pixels, and the pixel sizes increase nonlinearly from  $13 \times 24 \text{ km}^2$  at nadir to  $25 \times \sim 140 \text{ km}^2$  at the edges of the viewing swath. This study only uses data from the 30 pixels around the swath center (with a cross-track length less than 30 km), allowing for a better analysis of the spatial distribution of VCDs within short distances. It consequently changes the swath width in use to about 800 km so that global coverage is achieved roughly about every three days. Note that the pixel sizes here are much smaller than the GOME ( $320 \times 40 \text{ km}^2$ ) and SCIAMACHY ( $60 \times 30 \text{ km}^2$ ) instruments used in previous inverse estimates (Jaeglé et al., 2005; Müller and Stavrou, 2005; Wang et al., 2007; Stavrou et al., 2008).

### 3 Simulations of GEOS-Chem

#### 3.1 Descriptions of model simulations

This study uses the nested model of GEOS-Chem (version 08-03-02; [http://wiki.seas.harvard.edu/geos-chem/index.php/Main\\_Page](http://wiki.seas.harvard.edu/geos-chem/index.php/Main_Page)) for East Asia run at a horizontal resolution of  $0.667^\circ \text{long} \times 0.5^\circ \text{lat}$  with 47 layers vertically (Chen et al., 2009). The model is run with the full  $\text{O}_x\text{-NO}_x\text{-CO-VOC-HO}_x$  chemistry. It is driven by the assimilated meteorological fields of GEOS-5 taken from the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO). Vertical mixing in the planetary boundary layer follows the non-local scheme (Lin and McElroy, 2010) accounting for the varying magnitude of mixing from stable to unstable states of the boundary

layer. Convection is parameterized based on a modified version of the Relaxed Arakawa-Schubert scheme by Moorthi and Suarez (1992) (Rienecker et al., 2008). The lateral boundary conditions are updated every 3 hours using results from associated global simulations at  $5^{\circ}$  long x  $4^{\circ}$  lat horizontally.

Annual anthropogenic emissions of  $\text{NO}_x$ , carbon monoxide (CO) and non-methane volatile organic compounds (VOC) are taken from the INTEX-B dataset for 2006 provided by Zhang et al. (2009), including sources from power plants, industry, transportation and the residential sector. Emissions from the residential sector are further assumed to vary month to month accounting for heating related emissions that depend on ambient air temperature (Streets et al., 2003). They, however, contribute only 6% of anthropogenic sources of  $\text{NO}_x$  on the annual basis (Zhang et al., 2009). Emissions from power plants, industry and transportation are held constant across the seasons since the seasonality is relatively small and is not included in the INTEX-B dataset. The impact of such simplification is found to be small (see Sect. 5). The diurnal variations of individual sources follow Lin et al. (2010a) and Lin and McElroy (2010; 2011).

Natural sources of  $\text{NO}_x$  include lightning, soil and biomass burning. Emissions from biomass burning are taken from the year-to-year varying monthly dataset of GFED2 (van der Werf et al., 2006); their magnitudes are negligible for  $\text{NO}_x$  over China (Wang et al., 2007; Lin et al., 2010a) as the relatively low combustion temperature does not allow for significant production of  $\text{NO}_x$ . In 2006, the emission budget for East China is only about 0.013 TgN.

Production of  $\text{NO}_x$  from lightning is determined by the flash rate multiplying the yield of nitric oxide (NO) from each flash. In GEOS-Chem, the NO yield is latitude dependent: over the Asian continent, the yield is set to be 500 moles per flash north of  $35^{\circ}$  N reducing to 260 moles per flash south of  $35^{\circ}$  N (Martin et al., 2006; Hudman et al., 2007), based on previous observational constraint that the NO yield in the tropics is lower than in the midlatitude (Huntrieser et al., 2006). The total (intra-cloud and cloud-to-ground) flash rate is determined by convective cloud top height to the 4.9<sup>th</sup> power over lands free of snow and ice, as formulated by Price et al. (1997). The total amount of lightning induced emissions is distributed vertically with a backward ‘C-shape’ profile (Ott et al., 2010). Horizontally, as the flash rate depends on cloud properties that are highly parameterized, it is subject to large uncertainties particularly for individual locations. Alternate lightning parameterizations based on cloud mass flux or convective precipitation (Allen and Pickering, 2002) were found not to

1 improve the simulation of lightning distribution (Hudman et al., 2007). To improve the  
2 simulation, a horizontal adjustment is taken for each model gridbox based on the OTD/LIS  
3 satellite measurements of lightning flashes (Sauvage et al., 2007; Murray et al., 2009, 2010,  
4 2011). For each month, the mean flash rate over 2004 – 2008 is set as the monthly  
5 climatology derived from the satellite measurements from 1995 to 2005; while the interannual  
6 variability is determined by the year-to-year varying cloud heights taken from the GEOS-5  
7 meteorological fields. The constraint on the monthly climatology is meaningful since no  
8 significant trend of lightning activities is found based on the satellite measurements (Sauvage  
9 et al., 2007; Murray et al., 2009, 2010, 2011). Over East China, convection and precipitation  
10 amount are both driven by the seasonal transition of the East Asian Monsoon and thus are  
11 highly correlated month to month. Figure 3a shows that GEOS-5 captures the observed  
12 seasonal variation of precipitation and near-surface (2m) air temperature, likely indicating that  
13 the seasonality of convection is simulated reasonably well, at least on the regional mean basis.  
14 It is expected thus that the seasonality of lightning activities is likely reasonably reproduced  
15 by GEOS-Chem, although large uncertainties still exist associated with the exact magnitude  
16 and timing of lightning emissions.

17 Soil emissions are based on the Yienger and Levy (1995) scheme with the canopy reduction  
18 factors described by Wang et al. (1998). They include sources due to microbiological  
19 processes producing NO<sub>x</sub> naturally as well as those associated with use of chemical fertilizers  
20 and manure. The net emissions vary with vegetation type (Olson, 1992), temperature and  
21 precipitation. The N-pulsing is determined by the amount of precipitation over lands  
22 containing dry soils prior to the precipitation (Yienger and Levy, 1995). It is noted that  
23 GEOS-5 simulates very well the seasonal variability of precipitation and air temperature (Fig.  
24 3a). Fertilizer derived emissions are limited to agricultural lands and are distributed evenly  
25 over the growing season (May to August north of 28° N and all year long in the tropics)  
26 (Yienger and Levy, 1995; Wang et al., 1998). They are assumed to be 2.5% of the total  
27 amount of fertilizer use taken from the country based statistics of the Food and Agriculture  
28 Organization of the United Nations (FAO); for China, the fertilizer data represent the year  
29 1990 (Wang et al., 1998). As normally assumed, fertilizer associated emissions are considered  
30 to be part of natural sources in this study for comparison with anthropogenic emissions  
31 relating to combustion.

Averaged over East China, the a priori anthropogenic emissions of NO<sub>x</sub> are relatively constant across the seasons, while lightning and soil sources reach maximum values in summer and are not important in winter (Fig. 2). On the annual basis, the a priori anthropogenic emissions of NO<sub>x</sub> are about 5.8 TgN yr<sup>-1</sup> over East China; and lightning and soil emissions are only about 3-6% of anthropogenic emissions (Table 1). In July, lightning and soil emissions increase to as large as 10-13% of anthropogenic emissions (Table 1). Locally, anthropogenic emissions exhibit a seasonal pattern that is negatively or weakly correlated with the seasonality of lightning and soil emissions (Fig. 3b). The spatial correlation between anthropogenic and lightning or soil emissions is lower than 0.36 in all months, with a value larger in summer and much lower in winter.

A total of five 1-year simulations for 2006 were conducted to quantify VCDs of NO<sub>2</sub> from anthropogenic, lightning, soil and biomass burning sources, as shown in Table 2. For consistency with satellite retrievals, model VCDs in each day are obtained by regridding modeled NO<sub>2</sub> at each vertical layer to 0.25 °long x 0.25 °lat, sampled from gridboxes with valid satellite retrievals, and applied with the averaging kernel from DOMINO-2. The daily data are averaged then to obtain monthly mean values for each gridbox.

Potential sources of model errors include emissions of NO<sub>x</sub>, emissions of other pollutants affecting the chemistry of NO<sub>x</sub>, the chemical mechanism for NO<sub>x</sub>, the scheme for mixing in the boundary layer, and the meteorological fields. The total model error from factors other than emissions of NO<sub>x</sub> is estimated to be about 30-40% (Martin et al., 2003; Wang et al., 2007; Lin and McElroy, 2011); in this study, the value of 40% is taken for the inversion purpose.

## **3.2 Comparison between simulated and retrieved VCDs of NO<sub>2</sub>**

Figure 4 compares retrieved VCDs with simulated values (from all sources; Case 1 in Table 2) for January, April, July, October and annual average for 2006. Retrieved VCDs are large in regions with more advanced economic and industrial development and/or dense population, including the coastal and neighbor provinces from Beijing to Shanghai, the Pearl River Delta and the Sichuan Basin. Spike values are evident over major cities. In addition, retrieved VCDs vary across the months significantly, reaching maximum values in January and minimum values in July.

GEOS-Chem captures fairly well the spatial distributions of retrieved VCDs in different seasons (Fig. 4). The R<sup>2</sup> for spatial correlation between modeled and retrieved VCDs reaches



0.64 for January and 0.53 for July (Table 3). The smaller correlation in July is in part because the native horizontal resolution of the CTM ( $0.667^{\circ}\text{long} \times 0.5^{\circ}\text{lat}$ ) is not fine enough to capture the large spatial variation of VCDs within short distances resulting from the short lifetime of  $\text{NO}_x$  (Martin et al., 2003; Lin et al., 2010a). Spatial smoothing of 5 gridboxes by 5 gridboxes (i.e.,  $1.25^{\circ}\text{long}$  by  $1.25^{\circ}\text{lat}$ ) results in a significant enhancement of model-retrieval correlation in July with the  $R^2$  increasing to 0.67. The improvement of  $R^2$  due to the smoothing is moderate in January, compared to July, since the lifetime of  $\text{NO}_x$  is longer and the spatial variability of VCDs within short distances is smaller and better simulated by GEOS-Chem.

GEOS-Chem underestimates the magnitude of retrieved VCDs particularly over polluted regions in wintertime (Fig. 4). The range of simulated VCDs is also narrower than the retrieved range: spatially, the modeled maximum VCD is lower than the retrieved maximum with the minimum being higher (Table 3). Averaged over East China, model VCDs are about 20% lower than retrieved values in July and about 36% lower in January (Table 3).

## **4 Inversion of anthropogenic, lightning and soil emissions**

### **4.1 Method**

As discussed in Sect. 3.1, anthropogenic emissions of  $\text{NO}_x$  in East China exhibit weak seasonality and natural emissions reach maximum values in summer and minimum in winter (Fig. 2). In addition, the lifetime of  $\text{NO}_x$  is shortest in summer and longest in winter as a result of varying photochemical activity (Martin et al., 2003; Lin et al., 2010a; Lin and McElroy, 2010). This results in minimum values in summer for VCDs of  $\text{NO}_2$  of anthropogenic origin and maximum values for  $\text{NO}_2$  from natural sources, as simulated by GEOS-Chem (Fig. 5). Averaged over East China, natural sources contribute to about 30% of the total abundance of  $\text{NO}_2$  in July and August, in contrast to their negligible contributions in winter months. This characteristic is exploited here to estimate anthropogenic and natural emissions separately.

The inversion here involves a multi-step process based on a weighted multivariate linear regression analysis facilitated by several supplementary procedures. The regression is described in Sect. 4.1.1. The complete inversion process is described in Sect. 4.1.2 together with the supplementary procedures.

#### 4.1.1 A weighted multivariate linear regression analysis

Neglecting horizontal transport and assuming a linear relationship between the total VCD of NO<sub>2</sub> and VCDs from individual sources, the retrieved VCD of NO<sub>2</sub> for a given gridbox (of 0.25 °long x 0.25 °lat) in a given month can be approximated as the sum of modeled VCDs from individual emission sources, multiplied by certain scaling factors, and a random error term:

$$\Omega_r = K_a \Omega_{m,a} + K_l \Omega_{m,l} + K_s \Omega_{m,s} + K_b \Omega_{m,b} + \varepsilon \quad (1)$$

where  $\Omega_m = \Omega_{m,a} + \Omega_{m,l} + \Omega_{m,s} + \Omega_{m,b}$

Here  $\Omega_r$  denotes retrieved VCD of NO<sub>2</sub>, and  $\Omega_m$  denotes modeled VCD. The error term  $\varepsilon$  is assumed to follow a normal distribution with zero mean and standard deviation ( $\sigma$ ) equal to the sum in quadrature of errors from  $\Omega_r$  and  $\Omega_m$ . The subscripts 'a', 'l', 's', and 'b' indicate anthropogenic, lightning, soil, and biomass burning sources of NO<sub>x</sub>, respectively. The VCD that can be predicted from the inversion process is:

$$\Omega_p = k_a \Omega_{m,a} + k_l \Omega_{m,l} + k_s \Omega_{m,s} + k_b \Omega_{m,b} \quad (2)$$

where each  $k$  is an estimate of the corresponding  $K$  and is to be determined by the inverse modeling. The top-down emission is calculated as the sum of the a priori emissions from individual sources multiplied by corresponding scaling factors:

$$E_t = k_a E_{a,a} + k_l E_{a,l} + k_s E_{a,s} + k_b E_{a,b} \quad (3)$$

Here a linear relationship is assumed for each source between emissions of NO<sub>x</sub> and VCDs of NO<sub>2</sub>.  $\Omega_r$ ,  $\Omega_m$  and  $\sigma$  are known variables, and they vary from one month to the next. By minimizing the sum of  $[(\Omega_r - \Omega_p)/\sigma]^2$  in all months, Eqs. (1–2) serve as a weighted multivariate linear regression model to determine the scaling factors. Here the scaling factors are assumed implicitly to be season independent.

Over China, the contribution of biomass burning is very small for NO<sub>x</sub>, thus we do not attempt to constrain the associated emissions: the respective scaling factor is set as unity. Emissions from lightning and soil vary with seasons with similar patterns (Fig. 2), thus it is difficult to distinguish their contributions accurately based on the regression approach.

Therefore the scaling factors for the two sources are assumed to be the same in conducting the regression analysis. Under these assumptions, Eqs. (1–3) reduce to:

$$\Omega_r = K_a \Omega_{m,a} + K_l (\Omega_{m,l} + \Omega_{m,s}) + \Omega_{m,b} + \varepsilon \quad (4)$$

$$\Omega_p = k_a \Omega_{m,a} + k_l (\Omega_{m,l} + \Omega_{m,s}) + \Omega_{m,b} \quad (5)$$

$$E_t = k_a E_{a,a} + k_l (E_{a,l} + E_{a,s}) + E_{a,b} \quad (6)$$

The regression model here provides a basic statistical tool to calculate  $k_a$  and  $k_l$ . It is necessary to determine the ranges of  $k_a$  and  $k_l$  to facilitate the regression analysis; without such constraints the derived scaling factors may be unrealistic for certain gridboxes (too high, too low or even negative).  $k_a$  is set to range from 0.33 to 3 reflecting that uncertainties in anthropogenic emissions are normally moderate for individual gridboxes. Uncertainties in natural emissions are expected to be larger than those in anthropogenic emissions, therefore  $k_l$  is set to vary between 0.2 and 5. A larger range allowable for  $k_l$  (e.g., 0.1 – 10) results in more extreme values of  $k_l$  in several sparse locations with very low emissions, and has negligible impacts on the top-down emission budgets for East China.

#### 4.1.2 A multi-step inversion process beyond the regression analysis

To better estimate  $k_a$  and  $k_l$ , the gridboxes are allocated to different groups for ancillary procedures supplementing the regression analysis. The ratio  $\Omega_r / \Omega_m$  differs between winter and summer due in part to the season-dependent contribution of natural sources to the abundance of NO<sub>2</sub> and the relative magnitudes of errors in the a priori natural versus anthropogenic emissions. Thus the seasonally varying ratio is used to categorize the gridboxes, followed by a multi-step procedure to calculate  $k_a$  and  $k_l$  for all gridboxes (Fig. 6).

Gridboxes are assigned to group 1 if the ratio  $\Omega_r / \Omega_m$  for a selected winter month is smaller than the ratio for a selected summer month; otherwise they are assigned to group 2. For gridboxes in group 1, the regression analysis is performed to calculate  $k_a$  and  $k_l$  (step 1 in Fig. 6). Results are discarded if the regression is not statistically significant at a significance level of 0.1 under the Chi-square test. Spatial interpolation (step 2 in Fig. 6) is then conducted iteratively to derive  $k_l$  for all gridboxes: For a gridbox with undetermined  $k_l$ , the value of  $k_l$  is calculated as the geometric mean of values in the surrounding 24 gridboxes derived

1 previously. For gridboxes in group 2, the regression analysis usually results in unrealistically  
2 low values for  $k_l$  and thus is used to estimate  $k_a$  only (step 3 in Fig. 6); the value of  $k_l$  has  
3 been determined at previous steps. Again, the result is discarded if the regression is not  
4 statistically significant.

5 The regression approach may not be appropriate for certain gridboxes with unrealistically  
6 large  $\Omega_r/\Omega_m$  ratios in the winter month. This is because the scaling factor  $k_a$  likely has a  
7 significant seasonal dependence as a result of the seasonality in anthropogenic emissions not  
8 fully accounted for in the a priori dataset. These gridboxes are therefore reassigned to a third  
9 group, where a  $k_a$  is derived for each month as the ratio  $[\Omega_r - k_l(\Omega_{m,l} + \Omega_{m,s}) - \Omega_{m,b}]/\Omega_{m,a}$   
10 (step 4 in Fig. 6). Tentatively, a gridbox is allocated to group 3 if the ratio  $\Omega_r/\Omega_m$  exceeds 3,  
11 or if the ratio exceeds 2 with  $\Omega_r$  being higher than  $6 \times 10^{15}$  molec.  $\text{cm}^{-2}$ , in the winter month.

12 As a final step, spatial interpolation is conducted to obtain  $k_a$  for gridboxes of groups 1 and 2  
13 where the regression is not statistically significant (step 5 in Fig. 6).

14 In group allocation and subsequent inversion process, each of three winter months (December,  
15 January and February) is paired with each of two summer months (July and August) to  
16 generate a suite of six cases for winter-summer contrast. (June is not selected since the  
17 contribution of natural sources to the total abundance of  $\text{NO}_2$  is not as significant as that in  
18 July or August.) Results from the six cases are combined to obtain final values of  $k_a$  and  $k_l$ .  
19 Specifically, results at a given step available from any or all of the six cases are geometrically  
20 averaged to obtain final values of  $k_a$  and  $k_l$ , if and only if they have not been derived at  
21 earlier steps.

## 22 **4.2 Scaling factors estimated for anthropogenic, lightning and soil sources**

23 The scaling factors are determined at step 1 for most gridboxes (Figs. 7 and 8). At this step,  
24 the values of  $k_a$  are around unity in most areas, are consistently larger than 1 over the  
25 southwestern provinces, and vary between 1 and 2 in many of the remaining areas (Fig. 7). By  
26 comparison,  $k_l$  varies more significantly from one location to the next (Fig. 8). It reaches  
27 maximum values in southern Hebei Province and along the northern coasts of the Bohai Sea,  
28 which are attributable in part to an underestimate in fertilizer-associated soil emissions in the  
29 a priori dataset (see Sect. 4.4 for details). The value of  $k_l$  also has spike values at spotty

locations in other parts of East China where natural emissions are normally very low. These spikes are likely artificial results of the inversion algorithm, errors in GEOS-Chem, and/or errors in the satellite product; they however have negligible impacts on emission budgets over East China.

The special treatment at step 4 is taken mainly for gridboxes in and around Shanxi Province and in parts of Ningxia and Inner-Mongolia (Fig. 7). These places are main areas in China for coal mining and coal-fired electricity generation. The large values of  $k_a$  in these places, particularly in winter, suggest that the a priori dataset from INTEx-B likely underestimates anthropogenic sources related to use of coal, consistent with the findings of Wang et al. (2010).

The final values of  $k_a$  and  $k_l$  are determined at step 5 (see Fig. 7 for January) and step 2 (Fig. 8), respectively, for all gridboxes. They are used to calculate the top-down emissions for respective sources.

### **4.3 VCDs of NO<sub>2</sub> predicted from the inversion process**

Compared to simulation results, predicted VCDs are much closer to retrieved values (Fig. 4). The spatial correlation between predicted and retrieved VCDs is much higher than the correlation between simulated and retrieved VCDs, with the  $R^2$  increasing from 64% to 83% in January and from 53% to 86% in July (Table 3). Averaged over East China, predicted VCDs are within 15% of retrieved values across the seasons.

### **4.4 Top-down emissions**

Figure 9 compares the top-down emissions of NO<sub>x</sub> for anthropogenic, lightning and soil sources with respective a priori emissions for July. Both top-down and a priori datasets suggest significant anthropogenic sources over the coastal provinces from Shanghai to Beijing and over the Pearl River Delta, resulting in large concentrations of NO<sub>2</sub> (Fig. 10). The top-down anthropogenic emissions are much higher than the a priori emissions in cities and at locations with extensive use of coal, especially in the northern provinces.

In July, emissions from lightning are large in the north as a result of strong convection events associated with the summer monsoon (Fig. 9). Soil emissions are greatest in major agricultural areas in Hebei, Henan, Shandong and parts of neighbor provinces (Fig. 9). The

top-down estimates for lightning and soil emissions are much larger than the a priori values at many places of the respective source regions. In particular, the significant increases in soil emissions over southern Hebei and along the northern coasts of the Bohai Sea are not likely to be an artificial result of assumptions embedded in the inversion process, as they persist during various sensitivity tests (Cases 1-19 in Table 1). The increases are consistent with the new bottom-up estimates by Steinkamp and Lawrence (2011) and Hudman et al. (2012), and are contributed in part by the intensive fertilizer-derived emissions that are underestimated in GEOS-Chem (Yienger and Levy, 1995; Wang et al., 1998; Steinkamp and Lawrence, 2011). More detailed comparisons with Steinkamp and Lawrence (2011) and Hudman et al. (2012) are conducted in Sect. 5.1.2.

The contribution of anthropogenic sources to the total emissions of  $\text{NO}_x$  in July is generally consistent between the a priori and top-down datasets (Fig. 9). The anthropogenic contribution exceeds 80% over large areas of East China, but is lower than 60% over most of the northwest and Inner-Mongolia. It differs from the anthropogenic contribution to the VCD of  $\text{NO}_2$  (Fig. 10). This is in part because lightning emissions occur at higher altitudes with longer lifetime of  $\text{NO}_x$  than near the ground, compensated by a reduced fraction of  $\text{NO}_2$  in  $\text{NO}_x$ . Another factor is the use of averaging kernel in deriving model VCDs. The averaging kernel is larger for  $\text{NO}_2$  of lightning origin at higher altitudes and lower for  $\text{NO}_2$  derived from the ground. Thus the contribution is enhanced for a given amount of lightning emissions to the total abundance of  $\text{NO}_2$ , compared to the same amount of emissions from the ground. The effect is evident particularly over the southwest. The inverse modeling study by Lin et al. (2010) showed that an assumed 100% increase ( $0.57 \text{ TgN yr}^{-1}$ ) in top-down lightning emissions resulted in a 15% reduction ( $0.86 \text{ TgN yr}^{-1}$ ) in top-down anthropogenic emissions for July 2008 over East China; that is, lightning emissions are 50% more effective than anthropogenic emissions for VCD of  $\text{NO}_2$ , consistent with the analysis here. In addition, data for simulated VCDs are sampled at the time of day of satellite overpass from days with valid retrieval data; while data for monthly emissions are averaged over all time of day in all days of the month. The different sampling methods and the day-to-day and diurnal variations in natural emissions may introduce some differences between anthropogenic contributions to total emissions and to total VCDs. Another possible cause is the neglect of horizontal transport likely introducing uncertainties in source attribution for relatively clean regions downwind of polluted areas, e.g., at many places of Inner-Mongolia.

Table 1 compares the a priori and top-down emission budgets over East China for individual sources. Annually, the inversion results in budgets of 8.016 TgN, 0.228 TgN and 0.424 TgN for anthropogenic, lightning and soil emissions, respectively. These values are about 39%, 31% and 31% larger than the corresponding a priori estimates. The top-down datasets also suggest that both lightning and soil emissions are less than 6% of anthropogenic emissions. For July, the top-down budgets for lightning and soil emissions are 0.0623 TgN and 0.0810 TgN, respectively, about 10% and 13% of anthropogenic emissions estimated at 0.646 TgN.

#### **4.5 Improved GEOS-Chem simulations using the top-down emissions**

The inversion approach presented here assumes a linear relationship for a given gridbox between the total VCD of NO<sub>2</sub> and VCDs from individual sources and between emissions of NO<sub>x</sub> and VCDs of NO<sub>2</sub>. It also neglects horizontal transport, which may introduce uncertainties in deriving emissions for individual gridboxes. To evaluate the reliability of the inversion results, sensitivity simulations of GEOS-Chem were conducted for January and July 2006 by using the top-down emissions to drive the CTM. As shown in Fig. 4, VCDs resulting from the sensitivity simulations reproduce the spatial distribution of retrieved VCDs in both months. The R<sup>2</sup> for spatial correlation reaches a high level of 81% for January and 66% for July (Table 3). The smaller correlation in July is due to the short lifetime of NO<sub>x</sub> such that the CTM (at 0.667 °long x 0.5 °lat) is not able to simulate the large spatial variation of NO<sub>2</sub> within short distances, as discussed in Sect. 3.2. After (5 gridboxes by 5 gridboxes) horizontal smoothing, the R<sup>2</sup> increases to 88% for January and 81% for July (Table 3). Averaged over East China, the magnitude of model VCD is about 10% higher than the predicted value in January and about 5% lower in July (Table 3), indicating a slight nonlinear relationship between emissions and VCDs.

#### **4.6 Sensitivity of emission inversion to embedded assumptions**

This section evaluates the effects on the top-down emissions of several important assumptions taken during the inversion process, particularly for assumptions on the seasonality of various emission sources. The results are summarized in Table 1.

This study only includes 30 out of the 60 pixels from each OMI scan with smaller sizes in order to better analyze the spatial distribution of nitrogen within short distances. The top-down emission budgets for individual sources are similar to results from a sensitivity

1 calculation employing OMI data from all pixels (Case 2 in Table 1); the differences are larger  
2 for individual locations (not shown).

3 The inversion approach allocates individual gridboxes to three groups prior to the regression.  
4 The effect of group allocation was evaluated by two tests, one by re-allocating gridboxes in  
5 group 2 to group 1 and the other by re-allocating gridboxes in both group 2 and group 3 to  
6 group 1. The tests suggested that the effect of group allocation is less than 15% for top-down  
7 lightning/soil emission budgets for East China (Cases 3 and 4 in Table 1) with much larger  
8 impacts for individual locations (not shown).

9 The regression accounts for the seasonal dependence of retrieval errors. A sensitivity test  
10 assuming a time invariant relative error of 30% resulted in decreases by less than 5% in top-  
11 down lightning/soil emissions (Case 5 in Table 1). Therefore the inversion approach is not  
12 sensitive to the seasonality of retrieval errors assumed here.

13 The regression here relies on assumptions on the seasonal variations of individual emission  
14 sources. Anthropogenic emissions from power plants, industry and transportation differ  
15 slightly among seasons (Zhang et al., 2009), but are assumed here to be season independent  
16 for simulations of GEOS-Chem. The assumption was evaluated by a sensitivity analysis  
17 taking into account the seasonality estimated by Zhang et al. (2009). Specifically, modeled  
18 anthropogenic VCDs ( $\Omega_{m,a}$ ) for individual months were scaled prior to the inversion by the  
19 seasonality of total anthropogenic emissions derived from Table 9 of Zhang et al. (2009). This  
20 resulted in increases by less than 15% in top-down lightning/soil emissions (Case 6 in Table  
21 1). The effect is smaller than 5% for top-down anthropogenic emissions.

22 Another test was taken to evaluate the effect of errors in the Yienger and Levy (1995) soil  
23 emissions used as our a priori estimate. Specifically, the updated emission data by Hudman et  
24 al. (2012) (see Sect. 5.1.2 for specifications) were used to adjust modeled VCDs from soil  
25 sources ( $\Omega_{m,s}$ ) prior to the inversion, by scaling the VCDs for individual gridboxes with the  
26 ratios of Hudman et al. over Yienger and Levy soil emissions. As a result, the annual top-  
27 down lightning emissions were reduced by 11% and soil emissions enhanced by 19% (Case 7  
28 in Table 1).

29 Jaeglé et al. (2005), Wang et al. (2007) and Zhao and Wang (2009) assumed lightning  
30 emissions to be simulated well by the CTM with no attempt to constrain them inversely.  
31 Under the same assumption, a sensitivity analysis was conducted by keeping lightning



emissions unchanged during the inversion process here. This resulted in a 52% increase in the top-down soil emission budget on the annual basis and a 41% increase for July (Case 8 in Table 1).

If soil emissions are held unchanged during the inversion process, the top-down lightning emissions will be increased by less than 6% (Case 9 in Table 1).

A total of nine additional tests were performed to further analyze the sensitivity of inversion results to assumptions on the seasonality of lightning and soil emissions, including the timing and magnitude (Cases 10-18 in Table 1). To test the timing of the seasonality,  $\Omega_{m,l}$  and  $\Omega_{m,s}$  were shifted arbitrarily forward or backward by one month, separately or in combination, prior to the inversion process (Cases 10-15 in Table 1). As a result, top-down lightning and soil emissions are affected by up to 14% on the regional mean basis. Another three tests evaluate the magnitude of the seasonality, by lowering  $\Omega_{m,l}$  and  $\Omega_{m,s}$ , separately and in combination, by 20% in spring (March, April, May) and fall (September, October, November) and 40% in summer (June, July, August). These changes have significant impacts on top-down lightning and soil emissions: by reducing  $\Omega_{m,l}$  and  $\Omega_{m,s}$  simultaneously, top-down lightning and soil emissions were enhanced by about 33-34% in July.

Convection lifts pollutants in the boundary layer to the upper troposphere affecting the vertical distributions of various species. The magnitude of convection is highly parameterized in current climate models and CTMs and thus contains large uncertainties (Tost et al., 2010). The importance of model convection for simulated VCDs of  $\text{NO}_2$  originates from the altitude dependences of the lifetime of  $\text{NO}_x$ , the fraction of  $\text{NO}_2$  in  $\text{NO}_x$ , and the averaging kernel that is applied to the vertical distribution of  $\text{NO}_2$  (when comparing with satellite retrievals). Its net effect can be estimated roughly by comparing  $\text{NO}_2$  originating from a given amount of lightning emissions (i.e., locating mostly at high altitudes) and  $\text{NO}_2$  from the same amount of anthropogenic emissions (i.e., locating mostly in the boundary layer). As suggested by Lin et al. (2010) and discussed in detail in Sect. 4.4, lightning emissions are 50% more effective than anthropogenic emissions for VCD of  $\text{NO}_2$  for July 2008 over East China. Thus, doubling the magnitude of model convection and consequent changes in vertical distribution of  $\text{NO}_2$  will result in a net increase by 50% in modeled convection-associated VCD of  $\text{NO}_2$  (when the averaging kernel is taken into account). To evaluate the impact of potential errors in model convection, a sensitivity test enhanced by 50% the magnitude of convection of anthropogenic

NO<sub>2</sub> in July and August 2006, by tentatively increasing by 25% modeled concentrations of anthropogenic NO<sub>2</sub> above 500 hPa (Case 19 in Table 1). This effectively increased modeled anthropogenic VCDs of NO<sub>2</sub> over East China by about 7.5% in both months. It consequently resulted in a slight reduction in top-down anthropogenic emissions with reductions by about 4-5% in top-down lightning and soil emissions (Case 19 in Table 1).

The standard deviation (STD) of results from all inversions performed (Cases 1-19 in Table 1) is ~ 2% for anthropogenic emission budgets over East China, ~ 14% for lightning emissions, and ~ 16% for soil emissions. Locally, the STDs in most gridboxes are less than 30% for anthropogenic emissions and are within 50% for lightning and soil emissions (Fig. 11). The STDs are larger at locations containing very small emissions (Fig. 11); they however do not have a significant impact on the estimate of regional emission budgets.

Errors in the top-down emissions attributable to assumptions in the inversion process are calculated as the STD of all inversion results added in quadrature with errors associated with the nonlinearity between emissions of NO<sub>x</sub> and VCDs of NO<sub>2</sub>. The nonlinearity derived errors are taken to be about 10% for East China as a whole, based on discussions in Sect. 4.5. Thus the combined errors are about 10%, 17% and 19% for total emissions in East China from anthropogenic, lightning and soil sources, respectively (Table 1).

#### **4.7 Total errors in the top-down emission budgets over East China**

The inverse estimate here is subject to errors in retrievals, errors in model simulations, and errors in the inversion procedures as estimated from the sensitivity analyses. The total error in top-down emission budget over East China is taken to be the sum in quadrature of the three errors, amounting to about 51%, 53% and 53% for anthropogenic, lightning and soil sources, respectively (Table 1).

We do not attempt to estimate the total errors for individual gridboxes as the model and retrieval errors available here only represent bulk estimates on the regional mean basis.

### **5 A posteriori emissions**

The a posteriori emissions are estimated as the average of the a priori and top-down emissions weighted by the inverse-square of their respective errors (Martin et al., 2003). Errors in the a priori emissions are taken to be 60% for anthropogenic sources (Wang et al., 2007; Zhao and

Wang, 2009) increasing to 100% for lightning and soil sources accounting for the large range of current estimates (Boersma et al., 2005; Jaeglé et al., 2005; Schumann and Huntrieser, 2007; Wang et al., 2007; Zhao and Wang, 2009; Lin et al., 2010a). Errors in the respective top-down emissions are taken to be 51%, 53% and 53%, as derived in Sect. 4.7. Note that the error estimates here are conducted for total emissions in East China. Errors at individual locations may be much larger for both a priori and top-down datasets; the derivation however requires more detailed information that is not available currently. Therefore the error estimates for regional emission budgets are applied to individual locations, as a simplified procedure, in deriving the a posteriori emissions.

The a posteriori emissions for 2006 over East China amount to 7.071 TgN ( $\pm 39\%$ ) for anthropogenic sources, to 0.216 TgN ( $\pm 47\%$ ) for lightning sources, and to 0.402 TgN ( $\pm 47\%$ ) for soil sources (Table 1). For July, the a posteriori budgets are 0.576 TgN ( $\pm 39\%$ ), 0.0598 TgN ( $\pm 47\%$ ), and 0.0769 TgN ( $\pm 47\%$ ), respectively. The temporal and spatial distributions of the a posteriori emissions are presented in Figs. 2 and 9 for comparison with the a priori and top-down emissions.

## 5.1 Comparison with previous estimates

That anthropogenic emissions inferred from space are larger than bottom-up inventories for China is consistent with results from many previous studies (Jaeglé et al., 2005; Wang et al., 2007; Zhang et al., 2007; Lin and McElroy, 2011). Our a posteriori budget for anthropogenic emissions over East China is similar to the value of 0.565 TgN for July 2007 estimated by Zhao and Wang (2009). This study further pinpoints, at a higher resolution, cities and areas with extensive use of coal to be the main regions where bottom-up inventories likely underestimate anthropogenic emissions. Evaluation of lightning emissions is more difficult due to the large uncertainty in current research (Boersma et al., 2005; Schumann and Huntrieser, 2007) and the significant interannual variability of lightning occurrences on the regional scale (Schumann and Huntrieser, 2007). Our a posteriori emissions are within the range of previous estimates (Boersma et al., 2005; Schumann and Huntrieser, 2007; Stavrakou et al., 2008).

Soil emissions of NO<sub>x</sub> over China are of great interest concerning the extensive use of fertilizers. A detailed analysis is conducted as follows for our a posteriori estimate of soil emissions.

### **5.1.1 Comparison with previous satellite-derived soil emission estimates**

Wang et al. (2007) suggested that soil emissions are about 0.85 TgN per year for 1997–2000 over East China, amounting to 23% of anthropogenic emissions on the annual basis and to as much as 43% for summer months. In better agreement with our estimates for East China, Jaelg   et al. (2005) found an annual budget of ~ 0.40 TgN for 2000, and Zhao and Wang (2009) suggested soil emissions to be about 0.0883 TgN for July 2007 (Lyatt Jaegl   personal communication, 2011; Chun Zhao and Yuhang Wang, personal communication, 2011). The differences are derived mainly from satellite products and methods to separate anthropogenic and natural sources of NO<sub>x</sub> employed in individual studies.

### **5.1.2 Comparison with recent bottom-up estimates for soil emissions**

Two new bottom-up estimates have been conducted for soil emissions by Steinkamp and Lawrence (2011) and Hudman et al. (2012), improving upon the work of Yienger and Levy (1995) used as our a priori emissions. The new estimates employ information from more recent and complete field measurements to estimate emission factors. They use soil moisture rather than precipitation amount to separate dry and wet soil conditions. Hudman et al. (2012) also use soil moisture to calculate the duration and strength of the N-pulsing. The two new estimates include updated information on fertilizer use and vegetation map. For fertilizer use, Steinkamp and Lawrence (2011) account for its interannual variation based on the FAO statistics; while Hudman et al. (2012) use the latest available gridded dataset from Potter et al. (2010) representative of the year 2000. The two studies assume about 1.0% and 0.62%, respectively, of nitrogen in fertilizers to be emitted as NO<sub>x</sub>, for better comparisons with the observation-based estimate by Stehfest and Bouman (2006). Hudman et al. (2012) combine satellite measurements for the growing season of vegetation and attribute 75% of fertilizer derived emissions to the first month of the growing season. They also consider deposited nitrogen species as an additional fertilizer-like source of NO<sub>x</sub>. Steinkamp and Lawrence (2011) update the leaf area index (LAI) data for calculating the canopy reduction factor; while Hudman et al. (2012) assume no canopy reduction.

Our a posteriori soil emissions at 0.402 TgN (± 47%) annually is consistent with the value of 0.504 TgN calculated by Hudman et al. (2012), if canopy reduction is accounted for in their estimate (Rynda Hudman, personal communication, 2011). Our emissions are also within the range of 0.18–0.97 TgN suggested by Steinkamp and Lawrence (2011) (see their Table 7). In

addition, the horizontal distribution of soil emissions is in broad agreement between our a posteriori dataset and the two new bottom-up estimates.

## 6 Conclusions

A regression-based multi-step inversion approach is proposed to estimate emissions of  $\text{NO}_x$  from anthropogenic, lightning and soil sources for 2006 over East China on a  $0.25^\circ\text{long} \times 0.25^\circ\text{lat}$  grid. It exploits information on VCDs of tropospheric  $\text{NO}_2$  retrieved from the OMI instrument by KNMI (the DOMINO product version 2). The nested GEOS-Chem model for East Asia is used to interpret the impacts of individual sources on VCDs of  $\text{NO}_2$  to facilitate the inversion analysis.

The inversion starts by allocating the gridboxes to three groups based on analyses of the ratio of retrieved over modeled VCDs in winter and summer. A multivariate regression analysis is used then to derive emissions from individual sources for all months, taking advantage of the seasonal patterns of different sources determined by the CTM. Ancillary procedures are taken to supplement the regression analysis for gridboxes in different groups. Assumptions made during the inversion process contribute to errors in the top-down emission budgets for East China by  $\sim 10\%$  for anthropogenic sources,  $\sim 17\%$  for lightning sources and  $\sim 19\%$  for soil sources. Sensitivity simulations of GEOS-Chem driven by the top-down emission data reproduce the spatial distribution of VCDs retrieved from OMI, with the  $R^2$  for spatial correlation reaching 0.88 for January and 0.81 for July after (5 gridboxes by 5 gridboxes) horizontal smoothing.

The inversion results in an annual budget of 7.071 TgN ( $\pm 39\%$ ) for the a posteriori anthropogenic emissions of  $\text{NO}_x$  over East China, about 23% larger than the INTEx-B dataset (Zhang et al., 2009) used as our a priori emissions. On the  $0.25^\circ\text{long} \times 0.25^\circ\text{lat}$  grid, it is evident that the excess is greater over cities and areas with extensive use of coal, particularly in the north in winter.

The a posteriori budgets are 0.216 TgN ( $\pm 47\%$ ) and 0.402 TgN ( $\pm 47\%$ ) for lightning and soil emissions, respectively, for 2006 over East China. Both values are about 24% higher than the respective a priori estimates, but are each less than 6% of the a posteriori anthropogenic emissions. Even for July, the a posteriori lightning and soil emissions are only about 10% and 13% of anthropogenic emissions, respectively. Our results for soil emissions are consistent

1 with recent bottom-up estimates by Steinkamp and Lawrence (2011) and Hudman et al. (2012)  
2 and previous inverse estimates by Jaeglé et al. (2005), Stavrakou et al. (2008) and Zhao and  
3 Wang (2009). They are however about half of the inverse estimate by Wang et al. (2007) who  
4 suggested soil emissions to be more than 40% of anthropogenic emissions in summer of  
5 1997-2000.

6 In concluding, anthropogenic emissions are found to be the dominant source of NO<sub>x</sub> over East  
7 China for 2006, even in summer when natural sources reach maximum values. The  
8 contribution of anthropogenic emissions most likely has increased in more recent years due to  
9 their rapid growth (Lin and McElroy, 2011). In the coming future, the anthropogenic  
10 contribution may continue to increase along with the rapid economic and industrial  
11 development, if emission control is not taken successfully. The importance of nitrogen control  
12 has been recognized by the Chinese government, resulting in control strategies targeting the  
13 power sector. However, the successfulness of nitrogen control also depends on changes in  
14 emissions from other sectors, particularly the industrial sector for which the current  
15 inventories may be subject to much larger uncertainties than for the power sector (Zhao et al.,  
16 2011). Further research is required to evaluate the effectiveness of nitrogen control and  
17 resulting impacts on the contributions of anthropogenic versus natural sources to atmospheric  
18 nitrogen burdens.

## 20 **Acknowledgments**

21 This research is supported by the National Natural Science Foundation of China, grant  
22 41005078 and 41175127. We acknowledge the free use of tropospheric NO<sub>2</sub> column data  
23 from [www.temis.nl](http://www.temis.nl).

## 25 **References**

- 26 Allen, D. J., and Pickering, K. E.: Evaluation of lightning flash rate parameterizations for use  
27 in a global chemical transport model, *J. Geophys. Res.*, 107, 4711, doi:  
28 10.1029/2002jd002066, 2002
- 29 Boersma, K. F., Eskes, H. J., Meijer, E. W., and Kelder, H. M.: Estimates of lightning NO<sub>x</sub>  
30 production from GOME satellite observations, *Atmos. Chem. Phys.*, 5, 2311-2331, 2005.

1 Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van  
2 den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real  
3 time retrieval of tropospheric NO<sub>2</sub> from OMI, *Atmos. Chem. Phys.*, 7, 2103-2118, 2007.

4 Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P.,  
5 Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and  
6 Brunner, D.: An improved tropospheric NO<sub>2</sub> column retrieval algorithm for the Ozone  
7 Monitoring Instrument, *Atmos. Meas. Tech.*, 4, 1905-1928, doi: 10.5194/amt-4-1905-2011,  
8 2011.

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

12 Hudman, R. C., Jacob, D. J., Turquety, S., Leibensperger, E. M., Murray, L. T., Wu, S.,  
13 Gilliland, A. B., Avery, M., Bertram, T. H., Brune, W., Cohen, R. C., Dibb, J. E., Flocke, F.  
14 M., Fried, A., Holloway, J., Neuman, J. A., Orville, R., Perring, A., Ren, X., Sachse, G. W.,  
15 Singh, H. B., Swanson, A., and Wooldridge, P. J.: Surface and lightning sources of nitrogen  
16 oxides over the United States: Magnitudes, chemical evolution, and outflow, *J. Geophys. Res.*,  
17 112, D12S05, doi: 10.1029/2006jd007912, 2007

18 Hudman, R. C., Moore, N., Martin, R., Russell, A. R., Mebust, A. K., Valin, L. C., and Cohen,  
19 R. C.: A mechanistic model of global soil nitric oxide emissions: Implementation and space  
20 based-constraints, submitted to *Atmospheric Chemistry and Physics Discussions*, 2012.

21 Huntrieser, H., Schlager, H., Höller, H., Schumann, U., Betz, H. D., Boccippio, D., Brunner,  
22 D., Forster, C., and Stohl, A.: Lightning produced NO<sub>x</sub> in tropical, subtropical and  
23 midlatitude thunderstorms: New insights from airborne and lightning observations, *Geophys.*  
24 *Res. Abstr.*, 8, 03286, SRef-ID:1607-7962/gra/EGU06-A-03286, 2006.

25 Jaeglé L., Steinberger, L., Martin, R. V., and Chance, K.: Global partitioning of NO<sub>x</sub> sources  
26 using satellite observations: Relative roles of fossil fuel combustion, biomass burning and soil  
27 emissions, *Faraday Discussions*, 130, 407-423, doi: 10.1039/b502128f, 2005.

28 Lin, J.-T., and McElroy, M. B.: Impacts of boundary layer mixing on pollutant vertical  
29 profiles in the lower troposphere: Implications to satellite remote sensing, *Atmos. Environ.*,  
30 44, 1726-1739, doi: 10.1016/j.atmosenv.2010.02.009, 2010.

1 Lin, J.-T., and McElroy, M. B.: Detection from space of a reduction in anthropogenic  
2 emissions of nitrogen oxides during the Chinese economic downturn, *Atmos. Chem. Phys.*, 11,  
3 8171-8188, doi: 10.5194/acp-11-8171-2011, 2011.

4 Lin, J.-T., McElroy, M. B., and Boersma, K. F.: Constraint of anthropogenic NO<sub>x</sub> emissions  
5 in China from different sectors: a new methodology using multiple satellite retrievals, *Atmos.*  
6 *Chem. Phys.*, 10, 63-78, 2010a.

7 Lin, J.-T., Nielsen, C. P., Zhao, Y., Lei, Y., Liu, Y., and McElroy, M. B.: Recent Changes in  
8 Particulate Air Pollution over China Observed from Space and the Ground: Effectiveness of  
9 Emission Control, *Environmental Science & Technology*, 44, 7771-7776, doi:  
10 10.1021/es101094t, 2010b.

11 Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global  
12 inventory of nitrogen oxide emissions constrained by space-based observations of NO<sub>2</sub>  
13 columns, *J. Geophys. Res.*, 108, 4537, doi: 10.1029/2003JD003453, 2003.

14 Martin, R. V., Sioris, C. E., Chance, K., Ryerson, T. B., Bertram, T. H., Wooldridge, P. J.,  
15 Cohen, R. C., J. Andy Neuman, Swanson, A., and Flocke, F. M.: Evaluation of space-based  
16 constraints on global nitrogen oxide emissions with regional aircraft measurements over and  
17 downwind of eastern North America, *J. Geophys. Res.*, 111, doi: 10.1029/2005JD006680,  
18 2006.

19 Moorthi, S., and Suarez, M. J.: RELAXED ARAKAWA-SCHUBERT - A  
20 PARAMETERIZATION OF MOIST CONVECTION FOR GENERAL-CIRCULATION  
21 MODELS, *Mon. Weather Rev.*, 120, 978-1002, 10.1175/1520-0493, 1992.

22 Müller, J. F., and Stavrou, T.: Inversion of CO and NO<sub>x</sub> emissions using the adjoint of the  
23 IMAGES model, *Atmos. Chem. Phys.*, 5, 1157-1186, 2005.

24 Murray, L. T., Jacob, D. J., Logan, J. A., and Koshak, W.: Improving techniques for satellite-  
25 based constraints on the lightning parameterization in a global chemical transport model, 11th  
26 Conference on Atmospheric Chemistry, January 14, 2009.

27 Murray, L. T., Jacob, D. J., and Logan, J. A.: Investigating lightning-driven interannual  
28 variability in the oxidative capacity of the troposphere, *Geophysical Research Abstracts*, Vol.  
29 12, EGU2010-14332, 2010.



1 Murray, L. T., Logan, J. A., Jacob, D. J., and Hudman, R. C.: Spatial and interannual  
2 variability in lightning constrained by LIS/OTD satellite data for 1998-2006: implications for  
3 tropospheric ozone and OH, to be submitted to Journal of Geophysical Research, 2011.

4 Olson, J. World Ecosystems (WEI.4): Digital raster data on a 10 minute geographic 1080 x  
5 2160 grid, in Global ecosystems database, version 1.0: Disc A, edited by NOAA Natl.  
6 Geophys. Data Center, Boulder, Colorado, 1992.

7 Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley, B., Lin, R.-  
8 F., Lang, S., and Tao, W.-K.: Production of lightning NO(x) and its vertical distribution  
9 calculated from three-dimensional cloud-scale chemical transport model simulations, J.  
10 Geophys. Res.-Atmos., 115, D04301, doi: 10.1029/2009jd011880, 2010.

11 Price, C., Penner, J., and Prather, M.: NO<sub>x</sub> from lightning, 1, Global distribution based on  
12 lightning physics, J. Geophys. Res., 102, 12, 1997.

13 Richter, A., Burrows, J. P., Nüß H., Granier, C., and Niemeier, U.: Increase in tropospheric  
14 nitrogen dioxide over China observed from space, Nature, 437, 129-132, doi:  
15 10.1038/nature04092, 2005.

16 Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W.,  
17 Sienkiewicz, M., Koster, R. D., Gelaro, R., Stajner, I., and Nielsen, E.: The GEOS-5 Data  
18 Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0, NASA, 2008.

19 Sauvage, B., Martin, R. V., Donkelaar, A. v., Liu, X., Chance, K., Jaegl é, L., Palmer, P. I.,  
20 Wu, S., and Fu, T.-M.: Remote sensed and in situ constraints on processes affecting tropical  
21 tropospheric ozone, Atmos. Chem. Phys., 7, 24, 2007.

22 Schumann, U., and Huntrieser, H.: The global lightning-induced nitrogen oxides source,  
23 Atmos. Chem. Phys., 7, 3823-3907, 2007.

24 Stavrakou, T., Müller, J. F., Boersma, K. F., De Smedt, I., and van der A, R. J.: Assessing the  
25 distribution and growth rates of NO<sub>x</sub> emission sources by inverting a 10-year record of NO<sub>2</sub>  
26 satellite columns, Geophysical Research Letters, 35, L10801, doi: 10.1029/2008gl033521,  
27 2008.

28 Stehfest, E., and Bouwman, L.: N<sub>2</sub>O and NO emission from agricultural fields and soils  
29 under natural vegetation: summarizing available measurement data and modeling of global

1 annual emissions, *Nutrient Cycling in Agroecosystems*, 74, 207-228, doi: 10.1007/s10705-  
2 006-9000-7, 2006.

3 Steinkamp, J., and Lawrence, M. G.: Improvement and evaluation of simulated global  
4 biogenic soil NO emissions in an AC-GCM, *Atmos. Chem. Phys.*, 11, 6063-6082, doi:  
5 10.5194/acp-11-6063-2011, 2011.

6 Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z.,  
7 Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J. H., and Yarber, K. F.: An inventory of  
8 gaseous and primary aerosol emissions in Asia in the year 2000, *J. Geophys. Res.-Atmos.*,  
9 108, 8809, doi: 10.1029/2002jd003093, 2003.

10 Tost, H., Lawrence, M. G., Bruehl, C., Joeckel, P., Team, G., and Team, S.-O. D. A.:  
11 Uncertainties in atmospheric chemistry modelling due to convection parameterisations and  
12 subsequent scavenging, *Atmos. Chem. Phys.*, 10, 1931-1951, 2010.

13 van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and  
14 Arellano, A. F.: Interannual variability in global biomass burning emissions from 1997 to  
15 2004, *Atmos. Chem. Phys.*, 6, 3423-3441, 2006.

16 Wang, S. W., Streets, D. G., Zhang, Q. A., He, K. B., Chen, D., Kang, S. C., Lu, Z. F., and  
17 Wang, Y. X.: Satellite detection and model verification of NO(x) emissions from power  
18 plants in Northern China, *Environ. Res. Lett.*, 5, 044007, doi: 10.1088/1748-9326/5/4/044007,  
19 2010.

20 Wang, S. W., Zhang, Q., Streets, D. G., He, K. B., Martin, R. V., Lamsal, L. N., Chen, D., Lei,  
21 Y., and Lu, Z.: Growth in NO<sub>x</sub> emissions from power plants in China: Bottom-up estimates  
22 and satellite observations, *Atmos. Chem. Phys. Discuss.*, 12, 45-91, doi:10.5194/acpd-12-45-  
23 2012, 2012.

24 Wang, Y., Jacob, D. J., and Logan, J. A.: Global simulation of tropospheric O<sub>3</sub>-NO<sub>x</sub>-  
25 hydrocarbon chemistry 1. Model formulation, *J. Geophys. Res.*, 103(D9), 10713–10726, 1998.

26 Wang, Y., McElroy, M. B., Martin, R. V., Streets, D. G., Zhang, Q., and Fu, T.-M.: Seasonal  
27 variability of NO<sub>x</sub> emissions over east China constrained by satellite observations:  
28 Implications for combustion and microbial sources, *J. Geophys. Res.*, 112, doi:  
29 10.1029/2006JD007538, 2007.

- 1 Yan, X. Y., Akimoto, H., and Ohara, T.: Estimation of nitrous oxide, nitric oxide and  
2 ammonia emissions from croplands in East, Southeast and South Asia, *Global Change*  
3 *Biology*, 9, 1080-1096, doi: 10.1046/j.1365-2486.2003.00649.x, 2003.
- 4 Yan, X. Y., Ohara, T., and Akimoto, I.: Statistical modeling of global soil NO(x) emissions,  
5 *Global Biogeochemical Cycles*, 19, Gb3019, doi: 10.1029/2004gb002276, 2005.
- 6 Yienger, J. J., and Levy, H.: Empirical model of global soil-biogenic NO<sub>x</sub> emissions, *J.*  
7 *Geophys. Res.*, 100, 18, 1995.
- 8 Zhang, Q., Streets, D. G., He, K., Wang, Y., Richter, A., Burrows, J. P., Uno, I., Jang, C. J.,  
9 Chen, D., Yao, Z., and Lei, Y.: NO<sub>x</sub> emission trends for China, 1995 - 2004: The view from  
10 the ground and the view from space, *J. Geophys. Res.*, 112, doi: 10.1029/2007JD008684,  
11 2007.
- 12 Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z.,  
13 Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.:  
14 Asian emissions in 2006 for the NASA INTEx-B mission, *Atmos. Chem. Phys.*, 9, 5131-  
15 5153, 2009.
- 16 Zhao, C., and Wang, Y. H.: Assimilated inversion of NO<sub>x</sub> emissions over east Asia using  
17 OMI NO<sub>2</sub> column measurements, *Geophysical Research Letters*, 36, doi:  
18 10.1029/2008gl037123, 2009.
- 19 Zhao, Y., Nielsen, C. P., Lei, Y., McElroy, M. B., and Hao, J.: Quantifying the uncertainties  
20 of a bottom-up emission inventory of anthropogenic atmospheric pollutants in China, *Atmos.*  
21 *Chem. Phys.*, 11, 2295-2308, doi: 10.5194/acp-11-2295-2011, 2011.

1 Table 1. Emission budgets over East China (TgN) derived from various inversion calculations

| Case | Description   | Anthro. | Lightning | Soil  | Anthro. | Lightning | Soil   |
|------|---|---------|-----------|-------|---------|-----------|--------|
|      |   | Annual  |           |       | July    |           |        |
|      | a priori  | 5.763   | 0.174     | 0.324 | 0.479   | 0.0507    | 0.0627 |
|      | Total errors in a priori emissions  | 60%     | 100%      | 100%  | 60%     | 100%      | 100%   |
| 1    | Top-down (base case)  | 8.016   | 0.228     | 0.424 | 0.646   | 0.0623    | 0.0810 |
| 2    | Use level-2 retrieval data from all 60 pixels   | 7.789   | 0.223     | 0.415 | 0.633   | 0.0616    | 0.0793 |
| 3    | Re-allocate group 2 gridboxes to group 1 (and thus skip step 2 in Fig. 6)                 | 8.169   | 0.201     | 0.370 | 0.665   | 0.0529    | 0.0693 |
| 4    | Re-allocate group 2 and 3 gridboxes to group 1 (and thus skip steps 2 and 4 in Fig. 6)    | 7.720   | 0.203     | 0.373 | 0.642   | 0.0534    | 0.0698 |
| 5    | Assume constant retrieval errors across the seasons                                       | 8.306   | 0.220     | 0.400 | 0.668   | 0.0596    | 0.0754 |
| 6    | Assume seasonality of anthropogenic emissions according to Table 9 of Zhang et al. (2009) | 7.830   | 0.259     | 0.480 | 0.640   | 0.0706    | 0.0911 |
| 7    | Use soil emissions from Hudman et al. (2012)  | 7.876   | 0.202     | 0.506 | 0.636   | 0.0537    | 0.0797 |
| 8    | Keep lightning emissions unchanged  | 7.950   | 0.174     | 0.645 | 0.644   | 0.0507    | 0.1145 |
| 9    | Keep soil emissions unchanged   | 8.105   | 0.238     | 0.324 | 0.655   | 0.0657    | 0.0627 |
| 10   | Shift $\Omega_{m,l}$ and $\Omega_{m,s}$ backward by 1 month                               | 8.029   | 0.212     | 0.375 | 0.641   | 0.0564    | 0.0702 |
| 11   | Shift $\Omega_{m,l}$ and $\Omega_{m,s}$ forward by 1 month                                | 8.001   | 0.204     | 0.397 | 0.641   | 0.0542    | 0.0756 |
| 12   | Shift $\Omega_{m,l}$ backward by 1 month  | 7.991   | 0.222     | 0.401 | 0.636   | 0.0599    | 0.0757 |
| 13   | Shift $\Omega_{m,l}$ forward by 1 month   | 7.975   | 0.210     | 0.405 | 0.640   | 0.0559    | 0.0775 |
| 14   | Shift $\Omega_{m,s}$ backward by 1 month  | 7.958   | 0.224     | 0.410 | 0.634   | 0.0614    | 0.0777 |
| 15   | Shift $\Omega_{m,s}$ forward by 1 month   | 7.943   | 0.226     | 0.423 | 0.627   | 0.0620    | 0.0803 |
| 16   | Reduce magnitude of $\Omega_{m,l}$ and $\Omega_{m,s}$                                     | 8.089   | 0.309     | 0.470 | 0.660   | 0.0835    | 0.108  |
| 17   | Reduce magnitudes of $\Omega_{m,l}$   | 8.024   | 0.287     | 0.416 | 0.653   | 0.0772    | 0.0965 |
| 18   | Reduce magnitudes of $\Omega_{m,s}$   | 8.075   | 0.242     | 0.462 | 0.652   | 0.0665    | 0.0887 |
| 19   | Increase convection of anthropogenic NO <sub>2</sub> by 50%                               | 7.984   | 0.218     | 0.408 | 0.638   | 0.0595    | 0.0777 |
|      | Normalized standard deviation of Cases 1-19 relative to Case 1                            | 1.7%    | 13.7%     | 16.0% | 1.7%    | 13.6%     | 16.2%  |

|  |       |       |       |       |        |        |
|--|-------|-------|-------|-------|--------|--------|
| Errors in top-down emissions due to assumptions during the inversion process alone | 10.1% | 17.0% | 18.9% | 10.2% | 16.9%  | 19.0%  |
| Total errors in top-down emissions   | 51.0% | 52.8% | 53.4% | 51.0% | 52.8%  | 53.5%  |
| A posteriori   | 7.071 | 0.216 | 0.402 | 0.576 | 0.0598 | 0.0769 |
| Total errors in a posteriori emissions   | 38.9% | 46.7% | 47.1% | 39.9% | 46.7%  | 47.2%  |

---

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

Table 2. Descriptions of VCDs of NO<sub>2</sub> from individual sources derived from model simulations

| Case              | Description <sup>1</sup>  |
|-------------------|---|
| 1: $\Omega_m$     | Simulated by including emissions from all sources   |
| 2                 | Simulated by including all but lightning emissions  |
| 3                 | Simulated by including all but emissions from lightning and fertilizer associated soil sources, i.e., only including anthropogenic, non-fertilizer soil and biomass burning sources |
| 4                 | Simulated by including emissions from anthropogenic and non-fertilizer soil sources only  |
| 5: $\Omega_{m,a}$ | Simulated by including anthropogenic emissions only   |
| 6: $\Omega_{m,l}$ | Case 1 – Case 2   |
| 7: $\Omega_{m,s}$ | (Case 2 – Case 3) + (Case 4 – Case 5)   |
| 8: $\Omega_{m,b}$ | Case 3 – Case 4   |

1. Emissions from all sources are always included for pollutants other than NO<sub>x</sub>.

1 Table 3. Statistics for retrieved and modeled VCDs of NO<sub>2</sub> in January and July 2006

|                          | Mean                            | Median | Min  | Max   | Intercept <sup>3</sup> | Slope <sup>3</sup> | R <sup>2 3</sup> | Intercept <sup>3</sup> | Slope <sup>3</sup> | R <sup>2 3</sup> |
|--------------------------|---------------------------------|--------|------|-------|------------------------|--------------------|------------------|------------------------|--------------------|------------------|
|                          | Spatially smoothed <sup>4</sup> |        |      |       |                        |                    |                  |                        |                    |                  |
| January                  |                                 |        |      |       |                        |                    |                  |                        |                    |                  |
| Retrieved                | 5.63                            | 2.54   | 0.03 | 62.27 |                        |                    |                  |                        |                    |                  |
| Modeled                  | 3.62                            | 2.23   | 0.09 | 27.62 | 0.87                   | 0.49               | 0.64             | 0.79                   | 0.50               | 0.70             |
| Predicted <sup>1</sup>   | 4.89                            | 2.38   | 0.05 | 58.12 | 0.28                   | 0.82               | 0.83             | 0.35                   | 0.79               | 0.88             |
| Sensitivity <sup>2</sup> | 5.37                            | 2.87   | 0.10 | 45.03 | 0.53                   | 0.86               | 0.81             | 0.42                   | 0.87               | 0.88             |
| July                     |                                 |        |      |       |                        |                    |                  |                        |                    |                  |
| Retrieved                | 2.34                            | 1.67   | 0.06 | 31.16 |                        |                    |                  |                        |                    |                  |
| Modeled                  | 1.87                            | 1.47   | 0.23 | 10.93 | 0.61                   | 0.53               | 0.53             | 0.47                   | 0.59               | 0.67             |
| Predicted                | 2.39                            | 1.65   | 0.26 | 31.11 | 0.16                   | 0.95               | 0.86             | 0.06                   | 0.96               | 0.91             |
| Sensitivity              | 2.26                            | 1.65   | 0.31 | 31.32 | 0.30                   | 0.84               | 0.66             | 0.14                   | 0.89               | 0.81             |

- 2 1. Predicted from the regression-based inversion
- 3 2. Sensitivity simulation using the top-down emissions
- 4 3. With respect to retrieved VCDs
- 5 4. After 5 gridboxes by 5 gridboxes horizontal smoothing

Figure 1. Regional specifications with provincial boundaries for East China. Also presented are its subregions, provinces and province-level municipalities discussed in the text: the Yangtze River Delta (YRD), the Pearl River Delta (PRD), the Sichuan Basin, the Bohai Sea, Beijing (BJ), Shanghai, Hebei, Henan, Shandong, Shanxi, Ningxia, Inner-Mongolia, and Sichuan. On the background is the annual mean VCDs of NO<sub>2</sub> from the DOMINO-2 product.

Figure 2. Seasonal variations of the a priori, top-down and a posteriori anthropogenic, lightning, soil and total emissions of NO<sub>x</sub> ( $10^{15}$  molec. cm<sup>-2</sup> hr<sup>-1</sup>) over East China for 2006.

Figure 3. (a) Seasonal variations of monthly mean near-surface (2m) air temperature and precipitation observed from 284 meteorological stations over East China and modeled by GEOS-5. The observation data are taken from the global hourly dataset (DS3505) archived in the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NCDC) (<http://www7.ncdc.noaa.gov/CDO/cdo>; see Lin et al. (2010b)). Data from GEOS-5 are sampled at the meteorological stations for consistency with the observations. (b) Spatial distribution of the month-to-month correlation between a priori anthropogenic and natural (lightning + soil) emissions.

Figure 4. VCDs of NO<sub>2</sub> ( $10^{15}$  molec. cm<sup>-2</sup>) for January, April, July, October and annual average for 2006 retrieved from OMI, simulated by GEOS-Chem with a priori emissions, predicted by the inversion approach, and simulated by GEOS-Chem with top-down emissions. Areas outside the territory of China or without valid retrievals are shown in grey.

Figure 5. Seasonal variations of VCDs of NO<sub>2</sub> ( $10^{15}$  molec. cm<sup>-2</sup>) over East China for 2006 resulting from anthropogenic, lightning, soil and total emissions of NO<sub>x</sub> modeled by GEOS-Chem and predicted by the inversion. Also included is the seasonality of VCDs retrieved from OMI.

Figure 6. Description of the regression-based step-by-step inversion process after gridbox grouping. Values of  $k_l$  are determined at steps 1 and 2 for all gridboxes and  $k_a$  at steps 1, 3, 4, 5. See Sect. 4.1.2 for detailed analysis.

Figure 7. Scaling factors for anthropogenic emissions ( $k_a$ ) determined at steps 1, 3, 4 and 5 of the inversion process described in Fig. 6 and Sect. 4.1.2. Step 2 does not affect  $k_a$  and thus is not presented. Results at steps 4 and 5 are for January in particular. Areas outside the territory of China or with undetermined  $k_a$  are shown in grey.



Figure 8. Scaling factors for lightning/soil emissions ( $k_l$ ) determined at steps 1 and 2 of the inversion process described in Fig. 6 and Sect. 4.1.2. These two steps decide  $k_l$  for all gridboxes. Areas outside the territory of China or with undetermined  $k_l$  are shown in grey.

Figure 9. The a priori, top-down and a posteriori estimates of anthropogenic, lightning and soil emissions of  $\text{NO}_x$  ( $10^{15}$  molec.  $\text{cm}^{-2} \text{hr}^{-1}$ ) and the anthropogenic contributions (in percentage) to total emissions for July 2006 over East China. Areas outside the territory of China are shown in grey.

Figure 10. VCDs of  $\text{NO}_2$  ( $10^{15}$  molec.  $\text{cm}^{-2}$ ) in July 2006 over East China resulting from anthropogenic, lightning and soil emissions of  $\text{NO}_x$  and the anthropogenic contributions (in percentage) to the total VCDs modeled by GEOS-Chem and predicted by the inversion. Areas outside the territory of China or without valid retrievals are shown in grey.

Figure 11. (left) Ratio of the standard deviation of scaling factors in all inversions (Cases 1-19 in Table 1) over the scaling factors derived from Case 1 on the  $0.25^\circ \text{ lat} \times 0.25^\circ \text{ long}$  grid for July 2006. Areas outside the territory of China are shown in grey. (middle and right) The ratio as a function of top-down emissions for July 2006 derived from Case 1.

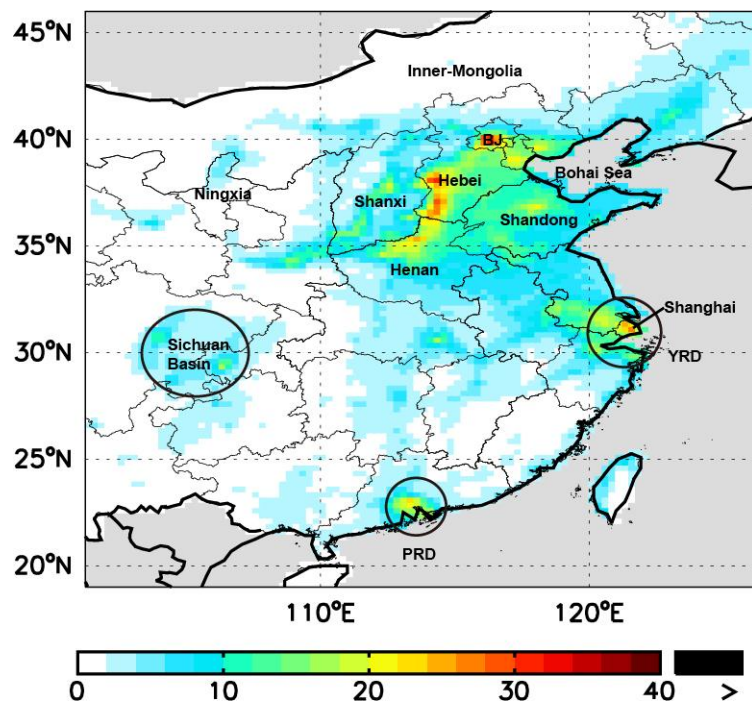


Figure 1.

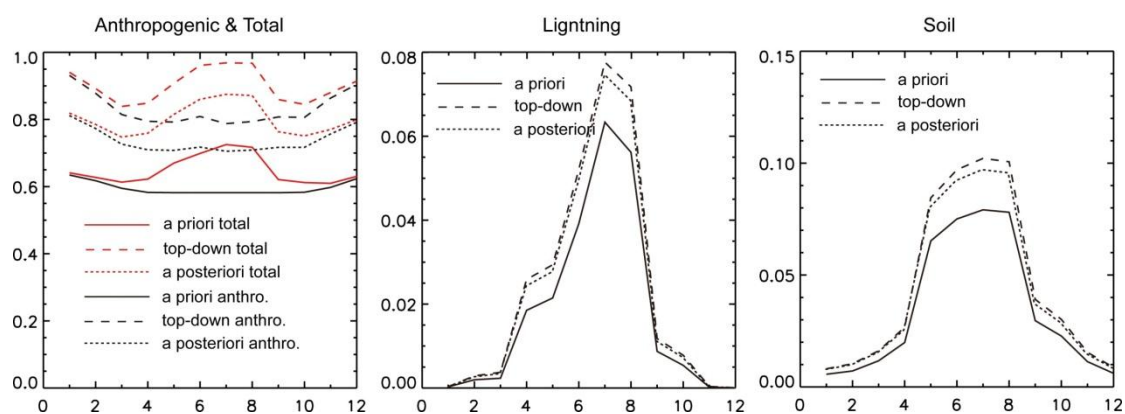
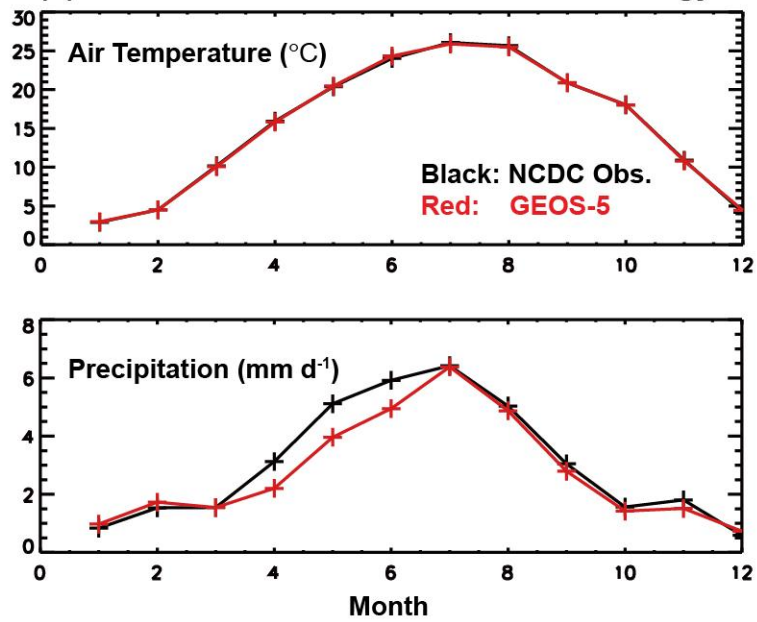
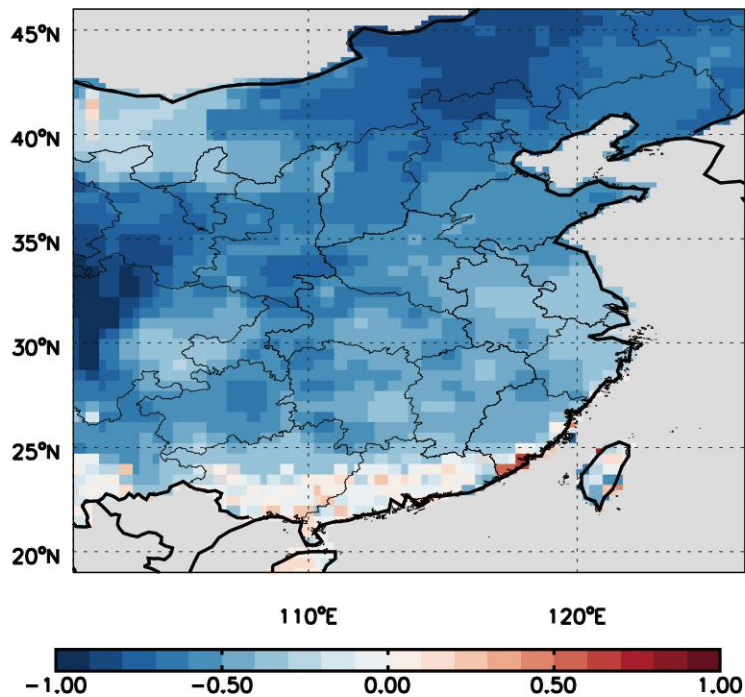


Figure 2.

### (a) Month-to-month Variation of Meteorology



### (b) Correlation of Anthro. and Lightning+Soil Emis



1

2 Figure 3.

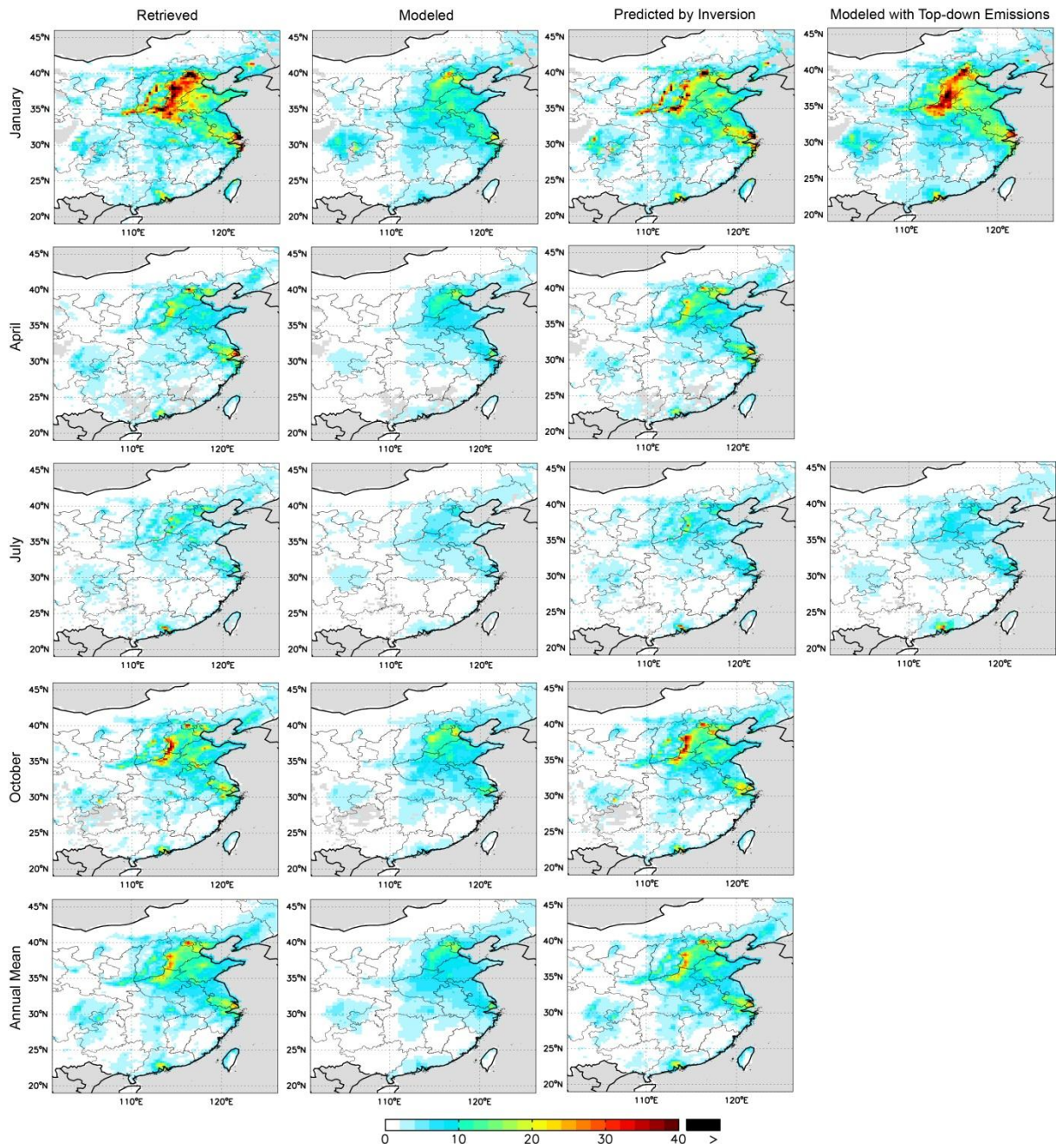


Figure 4.

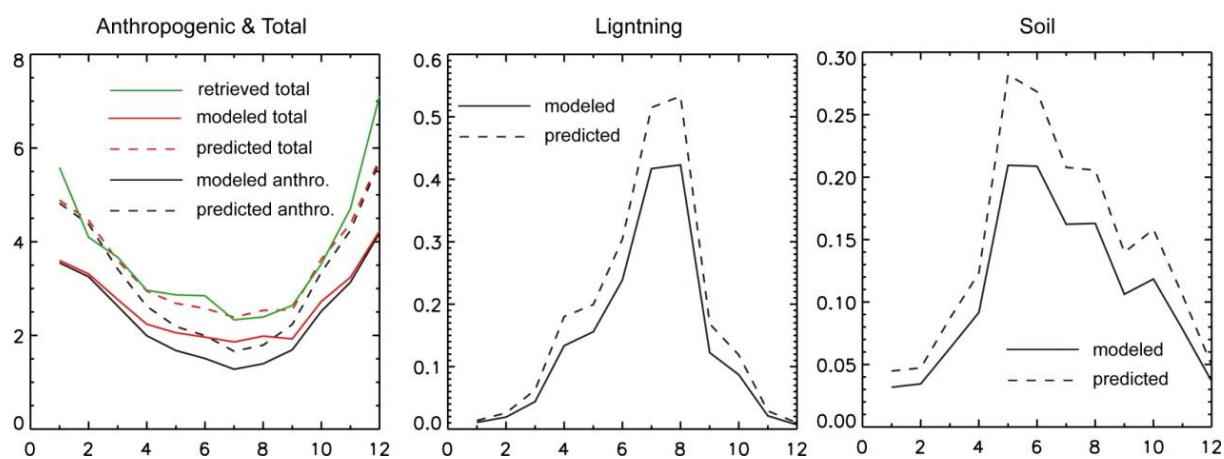


Figure 5.

# Steps of Inversion

# Gridboxes affected

(1) Regression over all months  
for  $k_a$  and  $k_i$

for

Gridboxes not evaluated at (4) , and  
 $(\Omega_r / \Omega_m)_{winter} / (\Omega_r / \Omega_m)_{summer} < 1$  , and  
regression is statistically significant



(2) Spatial interpolation for  $k_i$

for

All gridboxes not determined by (1)



(3) Regression over all months  
for  $k_a$

for

Gridboxes not done at (1) and (4) , and  
 $(\Omega_r / \Omega_m)_{winter} / (\Omega_r / \Omega_m)_{summer} \geq 1$  , and  
regression is statistically significant



(4) Special treatment:  
calculating  $k_a$  month by month

for

Gridboxes where  
 $(\Omega_r / \Omega_m)_{winter} > 3$  , or  
 $(\Omega_r / \Omega_m)_{winter} > 2$  and  $(\Omega_r)_{winter} > 6 \times 10^{15}$



(5) Spatial interpolation for  $k_a$

for

All gridboxes not determined by (1)-(4)

1

2 Figure 6.

3



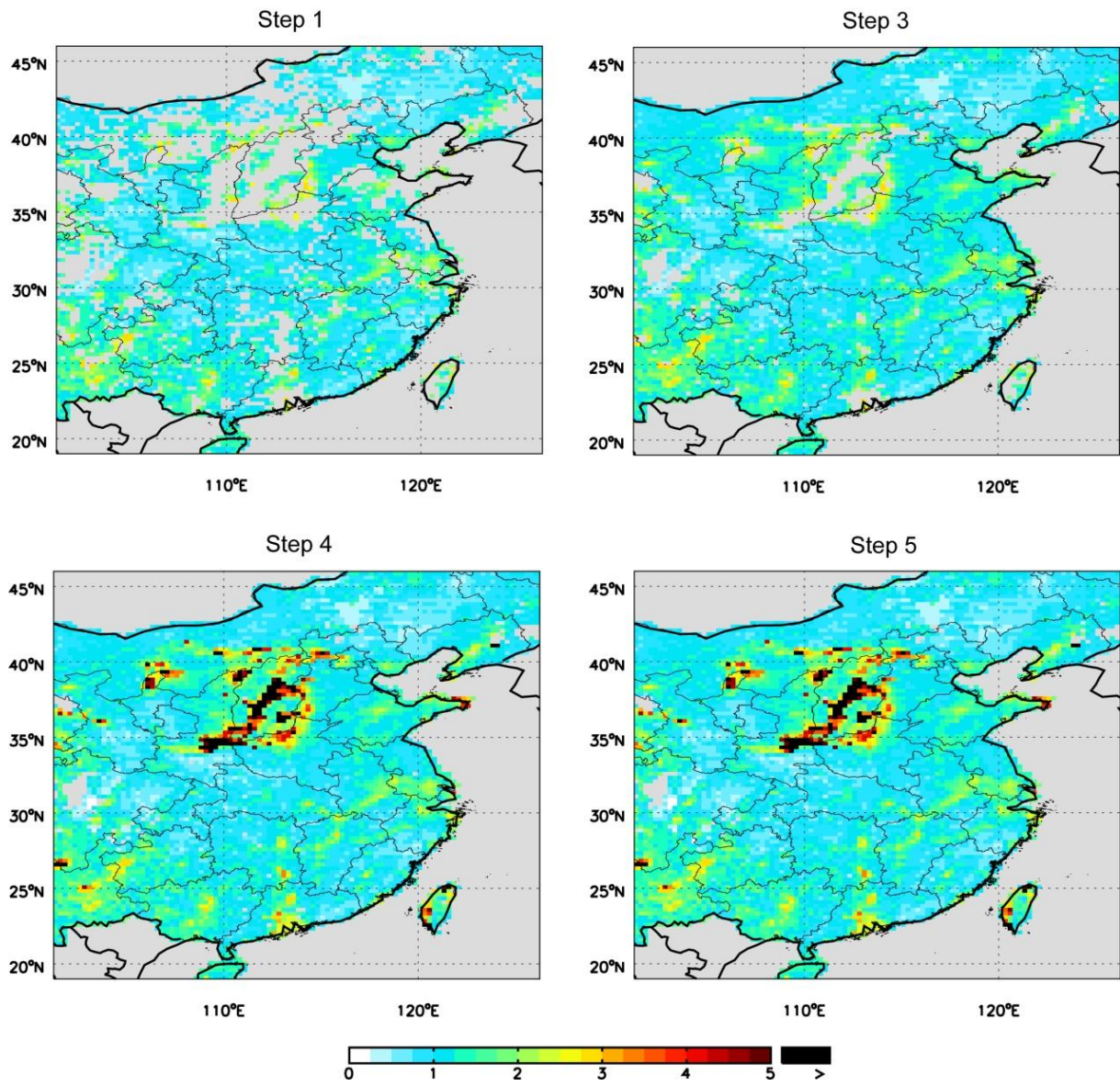


Figure 7.



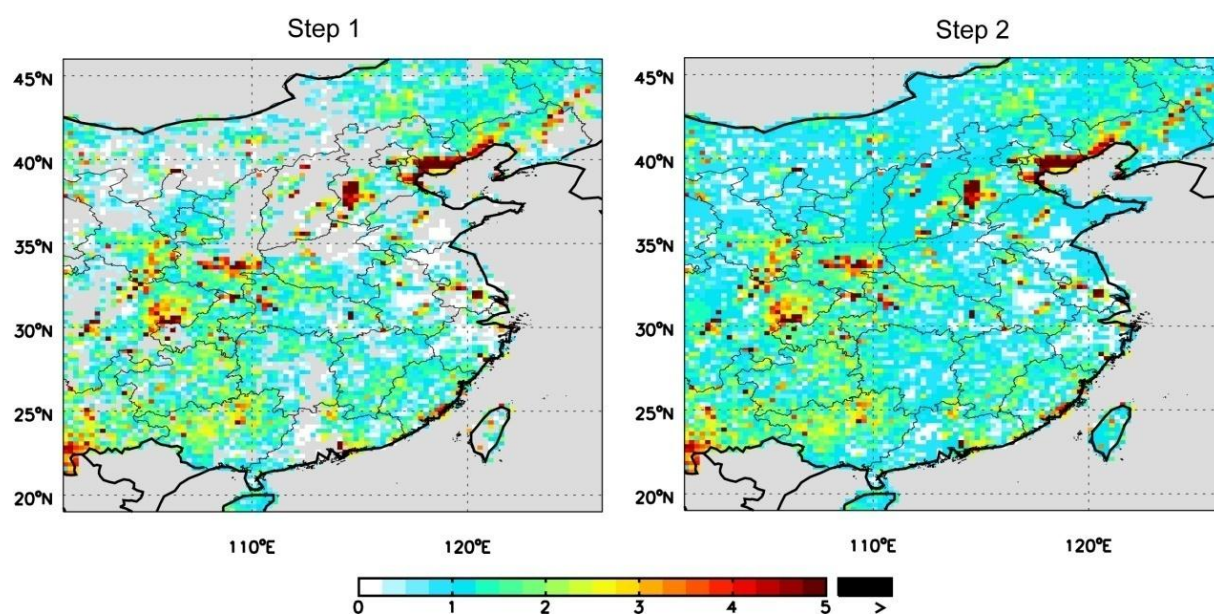


Figure 8.

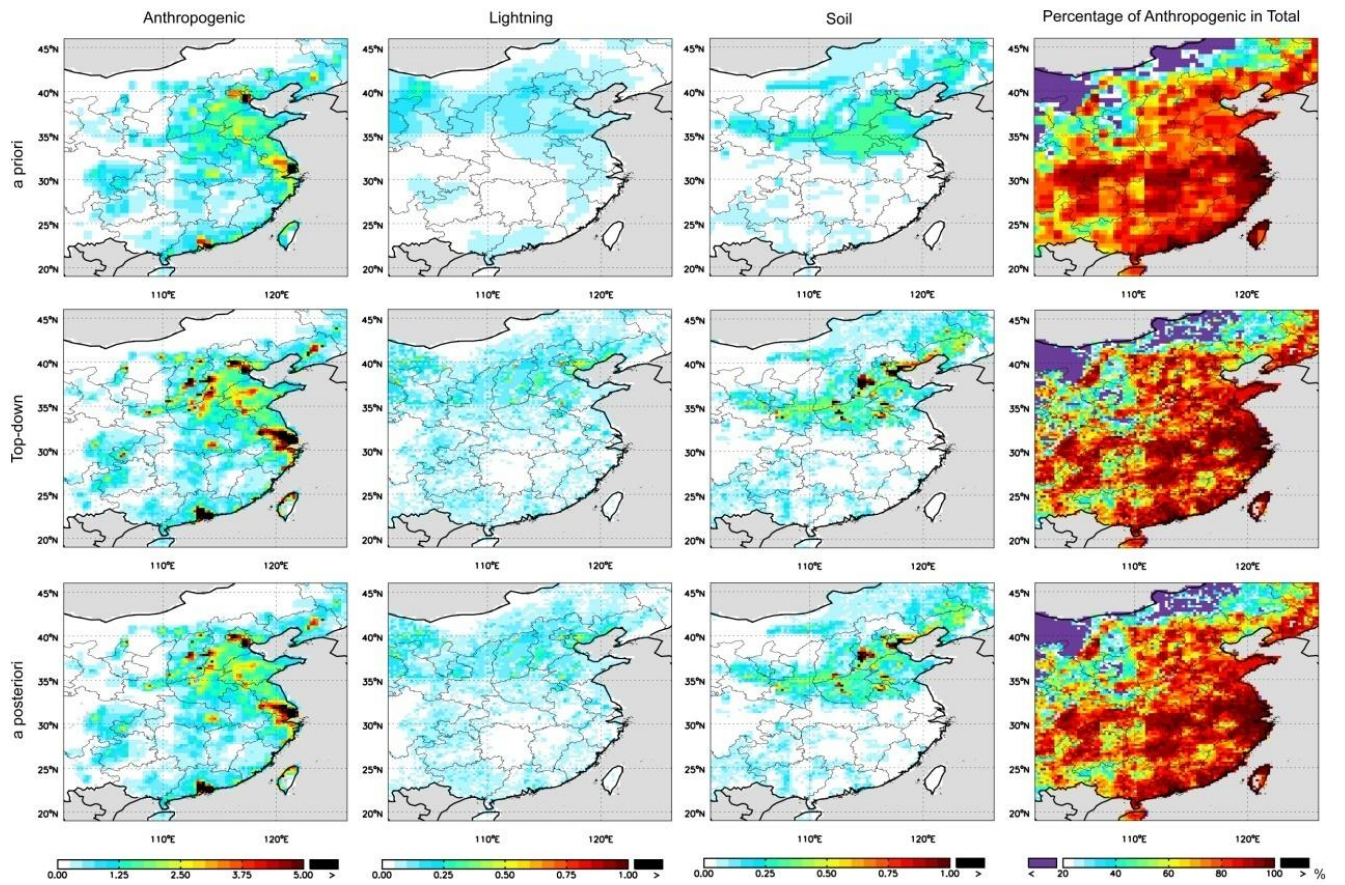


Figure 9.

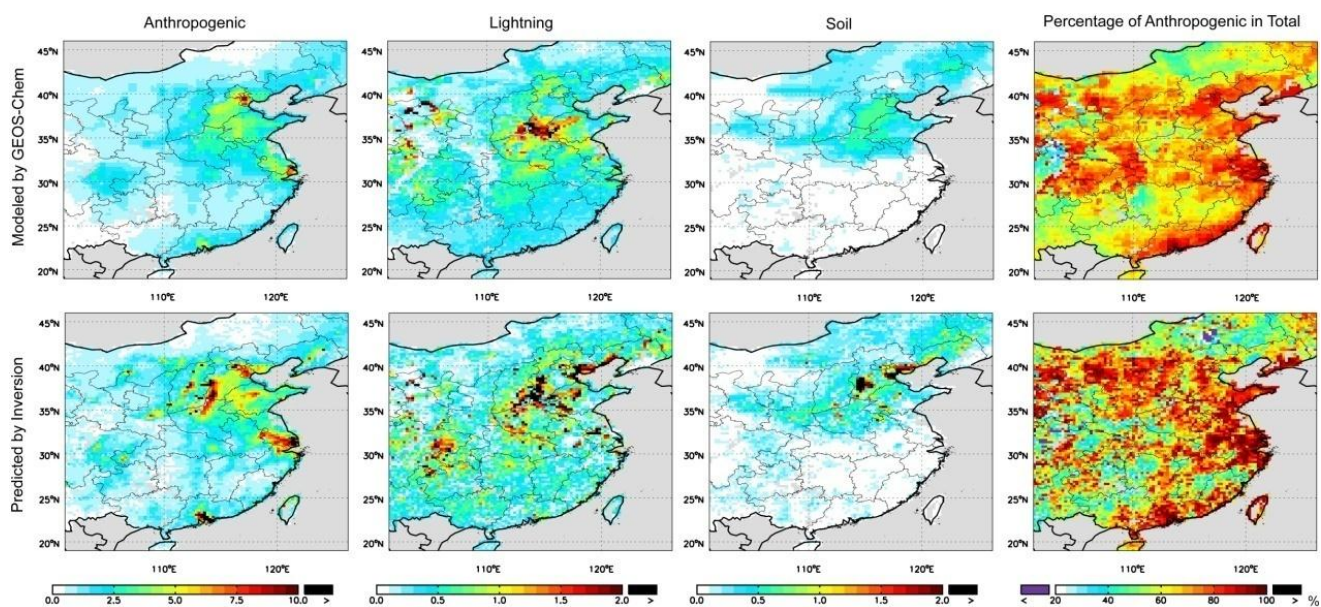


Figure 10.



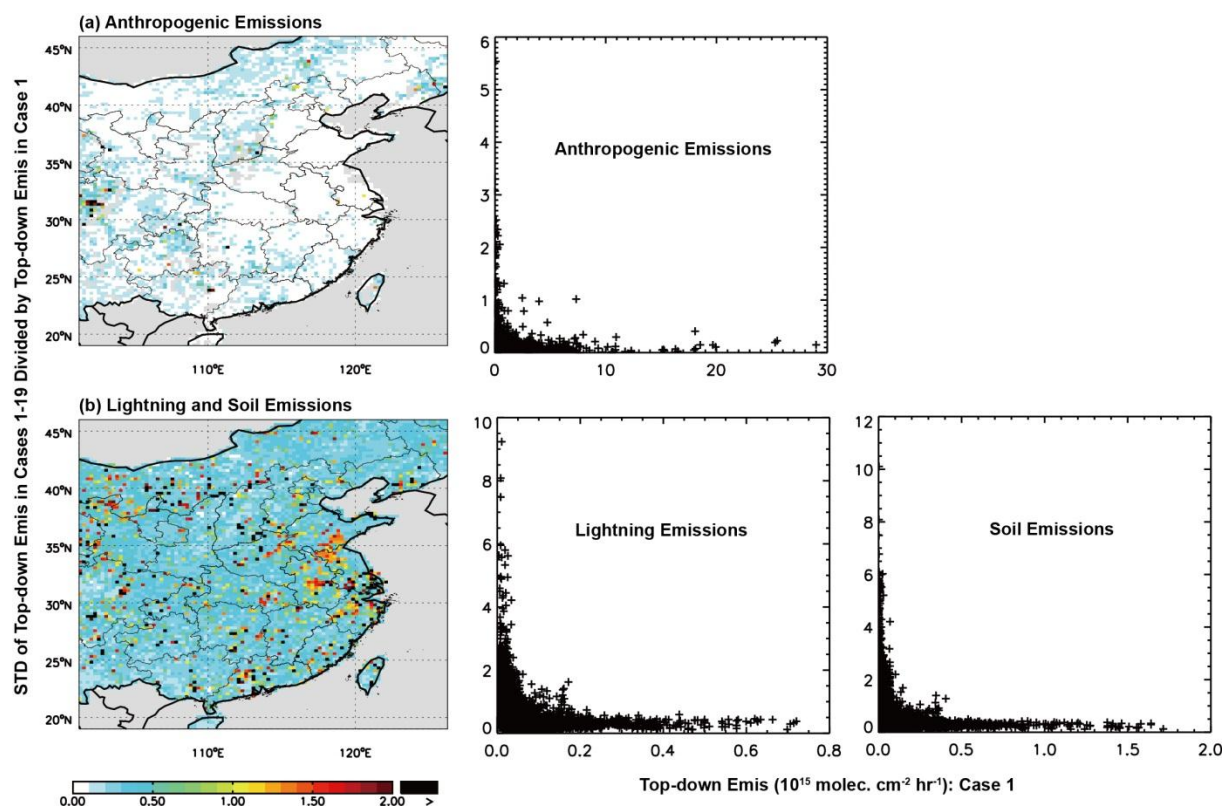


Figure 11.