Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal

Hansen Cao[1], Tzung-May Fu[1,*], Lin Zhang[1], Daven K. Henze[2], Christopher Chan Miller[3], Christophe Lerot[4], Gonzalo González Abad[3], Isabelle De Smedt[4], Qiang Zhang[5], Michel van Roozendael[4], Kelly Chance[3], Jie Li[6], Junyu Zheng[7], Yuanhong Zhao[1]

[1]Department of Atmospheric and Oceanic Sciences and Laboratory for Climate and Ocean-Atmosphere Studies, School of Physics, Peking University, Beijing,100871, China
[2]Department of Mechanical Engineering, University of Colorado, Boulder, USA
[3]Atomic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
[4]Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[5]Center for Earth System Science, Tsinghua University, Beijing, China
[6]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
[7]College of Environmental Science and Engineering, South China University of Technology, Guangzhou, China

Correspondence to: Tzung-May Fu (tmfu@pku.edu.cn)

Abstract. We used the GEOS-Chem model and its adjoint to quantify Chinese non-methane volatile organic compound (NMVOC) emissions for the year 2007, using the vertical column concentrations of formaldehyde and glyoxal observed by the Global Ozone Monitoring Experiment-2A (GOME-2A) instrument and the Ozone Monitoring Instrument (OMI) as constraints. We conducted a series of inversion experiments using different combinations of satellite observations to explore the impacts on top-down emission estimates due to different satellite retrievals. Our top-down estimates for Chinese annual total NMVOC emission was 23.4 to 35.4 (average 30.8) Tg C y^{-1}, including 13.5 to 19.7 (average 17.0) Tg C y^{-1} from anthropogenic sources, 8.9 to 14.8 (average 12.6) Tg C y^{-1} from biogenic sources, and 1.1 to 1.5 (average 1.2) Tg C y^{-1} from biomass burning. In comparison, the most widely-used bottom-up estimate for Chinese annual total NMVOC emission was 27.4 Tg C y^{-1}, including 15.5 Tg C y^{-1} from anthropogenic sources, 10.8 Tg C y^{-1} from biogenic sources, and 1.1 Tg C

y^{-1} from biomass burning. The simultaneous use of glyoxal and formaldehyde observations helped distinguish the NMVOC species from different sources and was essential in constraining anthropogenic emissions. Our four inversions consistently showed that the emissions of Chinese anthropogenic NMVOC precursors of glyoxal were larger than the *a priori* estimates. Our top-down estimates for the Chinese annual emission of anthropogenic aromatics (benzene, toluene, and xylene) ranged from 5.0 to 7.3 Tg C y^{-1}, 2% to 49% larger than the estimate of the bottom-up inventory (4.9 Tg C y^{-1}). Model simulations using the average of our top-down NMVOC emission estimates showed that surface afternoon ozone concentrations over northern and central China increased 5-12 ppb in June and decreased 5-13 ppb in December relative to the simulations using the *a priori* emissions and were in better agreement with measurements. We concluded that the satellite observations of glyoxal and formaldehyde together provided quantitative constraints on the emissions and source types of NMVOCs over China and improved our understanding on regional chemistry.

1 Introduction

Non-methane volatile organic compounds (NMVOCs) are emitted into the atmosphere from surface anthropogenic, biogenic, and biomass burning sources. NMVOCs are precursors to tropospheric ozone and secondary organic aerosols, both of which are climate forcers and major air pollutants. NMVOC also affect the oxidation capacity of the atmosphere, which in turn changes the lifetimes of greenhouse gases and other pollutants (Monks, 2005; Lelieveld et al., 2008). It is thus crucial to quantify NMVOC emissions in order to understand their impacts on atmospheric chemistry and climate on both global and regional scales. Here we used satellite observations and a chemical transport model to constrain NMVOC emissions from China and assessed their impacts on seasonal surface ozone.

Emissions of trace species are traditionally estimated in a "bottom-up" manner using activity data and emission factors, but these bottom-up estimates are sometimes susceptible to large uncertainties. This is especially true for NMVOC emissions in developing countries such as China, because (1) a wide range of species, source activities, and technologies are involved (Q. Zhang et al., 2009; Kurokawa et al., 2013; Li et al., 2014; Qiu et al., 2014), (2) locally-representative emission factors are often not measured (Wei et al., 2008; Zhao et al., 2011), and (3) reliable activity data are often incomplete,

particularly for small-scale industries, residential activities, and agricultural waste burning (Q. Zhang et al., 2009). Table 1 shows bottom-up estimates for Chinese total annual NMVOC emissions for the years 2005 to 2012, which ranged from 21.6 to 51.7 Tg y^{-1} (Guenther et al., 2006; Bolscher et al., 2007; Bo et al., 2008; Q. Zhang et al., 2009; van der Werf et al., 2010; Cao et al., 2011; Huang et al., 2012; Kurokawa et al., 2013; Li et al., 2014; Wu et al., 2016). The large uncertainties in these Chinese NMVOC emission estimates have led to great difficulty in evaluating their impacts on regional chemistry (Han et al., 2013;Wang et al., 2014).

A complementary, "top-down" approach of quantifying emissions uses observations of the targeted species or its chemical derivatives, combined with a chemical transport model acting as a transfer function, to invert for the fluxes of the targeted species. In particular, tropospheric column concentrations of formaldehyde (HCHO), retrieved from satellite UV-backscatter observations, have been used to constrain NMVOC emissions. Formaldehyde is produced at high yields during the oxidation of many NMVOC species, as well as emitted directly from anthropogenic and biomass burning activities (Akagi et al., 2011). Early applications of satellite-observed formaldehyde columns mainly focused on areas where the local NMVOC fluxes were dominated by biogenic emissions during the growing season, such as the southeast U.S. (Palmer et al., 2003, 2006; Millet et al., 2006, 2008), Europe (Dufour et al., 2009; Curci et al., 2010), the Amazon (Barkley et al., 2008, 2009, 2013), and the tropical central Africa (Marais et al., 2012, 2014a). These studies showed that the observed high concentrations of formaldehyde over densely-vegetated areas were linearly proportional to the local biogenic isoprene flux during the growing season.

Later studies constrained NMVOC emissions from multiple sources by analyzing the spatiotemporal variability of the observed formaldehyde columns (Shim et al., 2005; Fu et al., 2007; Stavrakou et al., 2009b; Curci et al., 2010; Gonzi et al., 2011; Marais et al., 2014b; Zhu et al., 2014). Fu et al. (2007), a forerunner of this study, analyzed the spatial and seasonal variation of the formaldehyde column observations from the Global Ozone Monitoring Experiment (GOME) over East and South Asia. They showed that, during the early 2000s, Chinese reactive NMVOC fluxes from biogenic, anthropogenic, and biomass burning sources were 3, 1.2, and 8.8 times their respective bottom-up estimates. In particular, Fu et al. (2007) found a large, annually-recurring NMVOC source over the North China

Plain (NCP) in June, which they attributed to crop residue burning after the local harvest of winter wheat. However, these top-down studies using only formaldehyde as constraints relied on bottom-up statistics to differentiate between NMVOC source types.

More recently, satellite measurements of tropospheric glyoxal columns emerged as an additional constraint on NMVOC emissions (Stavrakou et al., 2009a). Like formaldehyde, glyoxal is produced during the oxidation of many NMVOCs (including most importantly isoprene), as well as emitted directly from biomass burning (Fu et al., 2008; Myriokefalitakis et al., 2008). In addition, glyoxal is produced at high yields at the initial ring-cleaving stage during the oxidation of aromatics (Volkamer, 2001; Nishino et al., 2010), which are mainly anthropogenic. In contrast, formaldehyde production from the oxidation of aromatics is further downstream and thus spatially diffused (Volkamer, 2001). As such, simultaneous analyses of formaldehyde and glyoxal observations can help differentiate between biogenic and anthropogenic NMVOC emissions. Stavrakou et al. (2009a) pioneered a two-compound inversion using tropospheric glyoxal and formaldehyde column observations from the SCIAMACHY satellite instrument to constrain the global sources of glyoxal. They estimated that the anthropogenic NMVOC fluxes over East Asia for the year 2005 were a factor of 2-3 larger than the bottom-up estimates of the Emission Database for Global Atmospheric Research (EDGAR, v3.3) inventory (Olivier et al., 2001, 2002) and the REanalysis TROpospheric (RETRO) emission inventory (Schultz et al., 2007). In addition, they inferred a large missing source of glyoxal over the global continents, which they attributed to production from an unknown biogenic precursor.

Over eastern China, Liu et al. (2012) showed that the glyoxal column concentrations observed by SCIAMACHY in August 2007 was more than twice the simulated glyoxal columns using the bottom-up emission inventory developed by Q. Zhang et al. (2009). Over the Pearl River Delta area (PRD) in southern China, the discrepancy was at least a factor of three. They suggested that the missing glyoxal source over eastern China was anthropogenic, on the basis that the anonymous glyoxal columns observed by SCIAMACHY were spatially correlated with anthropogenic NO_x emissions. Their estimated Chinese anthropogenic aromatics emission was 13.4 Tg y^{-1}, which was six times the 2.4 Tg y^{-1} aromatic flux estimated by Q. Zhang et al. (2009). In contrast, Chan Miller et al. (2016) simulated the formaldehyde and glyoxal column concentrations over the Pearl River Delta area

(PRD) in southern China for the years 2006 and 2007 using the same inventory developed by Q. Zhang et al. (2009). They found that their simulated formaldehyde columns were consistent with the OMI formaldehyde observations, while their simulated glyoxal columns were lower than OMI observations by only 40%. They attributed the high anthropogenic aromatics emission estimate by Liu et al. (2012) to a regional high-bias in the SCIAMACHY data, as well as underestimated yields of glyoxal from the oxidation of aromatics.

One limitation in the use of satellite observations of formaldehyde and glyoxal for constraining NMVOC sources is their inherent uncertainty. Several studies have compared GOME-2A and OMI formaldehyde column observations against aircraft or ground-based measurements at a few locations around the world (De Smedt et al., 2015; Lee et al., 2015; Wang et al., 2017; Zhu et al., 2016). Zhu et al. (2016) compared the GOME-2A-observed formaldehyde column concentrations over the Southeast U.S. in summer 2013 against aircraft measurements and found the satellite measurements to be too low by a factor of approximately 1.7. Chan Miller et al. (2017) found that glyoxal column concentrations observed by OMI were lower than the aircraft measurements over the Southeast U.S. in summer 2013 by a factor of 1.5. Wang et al. (2017) compared the bi-monthly mean GOME-2A and OMI formaldehyde column concentrations retrieved by De Smedt et al. (2012, 2015) against ground-based multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements at a rural site in eastern China. They found that both satellite retrievals were systematically lower than the ground-based measurements by approximately 20%. These studies highlight the potential impacts on top-down NMVOC emission estimates due to uncertainty associated with satellite retrievals.

In this study, we used satellite retrievals of both formaldehyde and glyoxal, along with an updated chemical transport model and its adjoint, to constrain NMVOC emissions from China for the year 2007. We conducted sensitivity experiments to evaluate the impacts on the top-down estimates due to different satellite retrieval constraints, with the goal of obtaining a most probable range of top-down estimates. Finally, we examined the impacts of our top-down NMVOC emission estimates on surface ozone concentrations over China.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

2 Model and data

2.1 The GEOS-Chem model and its adjoint

We used the GEOS-Chem global 3D chemical transport model (version 8.2.1) to simulate the emission, transport, chemistry, and deposition of NMVOCs, as well as the resulting formaldehyde and glyoxal column concentrations for the year 2007. GEOS-Chem was driven by the assimilated meteorological data from the NASA Goddard Earth Observing System (GEOS-5) (Bey et al., 2001). To drive our simulations, the horizontal resolution of GEOS-5 data was downgraded from its native 2/3 °longitude × 1/2 °latitude to 5 °longitude × 4 °latitude. The number of vertical levels was reduced from 72 to 47 by merging layers in the stratosphere. The lower 2 km of the atmosphere was resolved by 14 levels. The temporal resolution of GEOS-5 data into GEOS-Chem is 3 h for atmospheric variables and 1 h for surface variables.

We updated the dicarbonyl chemical mechanism in GEOS-Chem developed by Fu et al. (2008), which in turn was originally adapted from the Master Chemical Mechanism (MCM) version 3.1 (Saunders et al., 2003; Bloss et al., 2005). NMVOC precursors of formaldehyde in our mechanism included ethane, propane, ≥C_4 alkanes, ethene, ≥C_3 alkenes, toluene, xylenes, isoprene, and monoterpenes. NMVOC precursors of glyoxal in our mechanism included propane, alkanes, ethene, ≥C_3 alkenes, ethyne, benzene, toluene, xylenes, isoprene, monoterpenes, glycolaldehyde, and 2-methyl-3-bute-2-nol. OH-oxidation of isoprene is a major source of both formaldehyde and glyoxal over China (Fu et al., 2007, 2008; Myriokefalitakis et al., 2008). We replaced the isoprene photochemical scheme with that used in GEOS-Chem v10.1 (Paulot et al., 2009a,b; Mao et al. 2013), where formaldehyde and glyoxal were produced from isoprene oxidation via the RO_2+NO pathway under high-NO_x conditions and via RO_2-isomerization under low-NO_x conditions. Li et al. (2016) compared the productions of formaldehyde and glyoxal from isoprene oxidation in this updated scheme with those in the MCM version 3.3.1 (Jenkin et al., 2015). They showed that the production pathways and yields of formaldehyde and glyoxal were similar in the two schemes under the high-NOx conditions typical of eastern China.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

We updated the molar yields of glyoxal from the OH oxidations of benzene (33.3%), toluene (26.2%), and xylenes (21.0%) following the latest literature (Arey et al., 2009; Nishino et al., 2010). These new molar yields were higher than those used in Fu et al. (2008) but still lower than those used by Chan Miller et al. (2016) (75%for benzene, 70%for toluene, 36%for xylenes), which were taken from MCM version 3.2 (Bloss et al., 2005). In MCM version 3.2, more than half of the glyoxal from aromatics oxidation were produced during second- and later-generation photochemistry, but such productions are still uncertain, with limited experimental support (Bloss et al., 2005).

Formaldehyde and glyoxal in our model were both removed by photolysis, as well as dry and wet deposition (Fu et al., 2008). We updated the Henry's law constant for glyoxal (Ip et al., 2009) and added the dry deposition of formaldehyde, glyoxal, methyglyoxal and glycolaldehyde on leaves (Mao et al., 2013). In addition, we assumed that glyoxal was reactively uptaken by wet aerosols and cloud droplets with an uptake coefficient $\gamma = 2.9 \times 10^{-3}$ (Liggio et al., 2005; Fu et al., 2008). All other physical and chemical processes in our forward model were as described in Fu et al. (2008).

For the forward model described above, we developed the adjoint by modifying the standard GEOS-Chem adjoint (version 34) (Henze et al., 2007). We used the Kinetic PreProcessor (KPP) (Daescu et al., 2003; Sandu et al., 2003) to construct the adjoint of the updated photochemical mechanism. Adjoint algorithms were updated to include the emission and deposition processes of formaldehyde and glyoxal precursors. The aqueous uptake rate of glyoxal by wet aerosols was a function of the ambient glyoxal concentration and the total wet aerosol surface area (Fu et al., 2008). We linearized this uptake process by archiving the wet aerosol surface areas in the forward simulations for use in the backward integrations.

We verified the adjoint model mathematically in two ways. Firstly, we used the adjoint to calculate the sensitivities of global glyoxal and formaldehyde burdens to biogenic isoprene and anthropogenic xylenes emissions, respectively, and found that the results reproduced the calculated sensitivities from the forward model (Figure S1 in Supplementary Information). Secondly, we used an *a priori* NMVOC emission inventory (Section 2.2) to drive the forward model and took the resulting global tropospheric formaldehyde and glyoxal column concentrations as pseudo observations. We then used the pseudo

observations of formaldehyde and glyoxal, respectively, to successfully optimize back to close to the *a priori* NMVOC emission estimates over high-emission areas from an initial emission guess that was five times larger (Figure S2 in Supplementary Information). These experiments demonstrated the usefulness of the adjoint model for the inversion of NMVOCs emissions.

2.2 *A priori* emission estimates of Chinese NMVOCs

As a starting point for our inversion, we used the most widely-used NMVOC emission estimates for China as the *a priori*. Table 2 summarizes the annual total of these *a priori* emission estimates and their associated uncertainties.

The *a priori* biogenic NMVOC emissions from China and from the rest of the world were calculated with the MEGAN algorithm (Guenther et al., 2006) and dependent on temperature, shortwave radiation, and monthly mean leaf area index. The annual total biogenic NMVOC emissions over China for the year 2007 was 10.8 Tg C y^{-1}, including 6.6 Tg C y^{-1} of isoprene. Previous estimates of Chinese biogenic NMVOC emissions ranged from 5.0 to 11.0Tg C y^{-1} (Guenther et al., 2006; Fu et al., 2007; Stavrakou et al., 2015). Based on this range, we estimated the uncertainty of the *a priori* biogenic emissions over China to be ±55%.

The *a priori* emissions for Chinese anthropogenic NMVOCs were from the Multi-resolution Emission Inventory for China (MEIC) inventory (Li et al., 2014), developed for the year 2010 at 0.25 º × 0.25 º resolution. The MEIC inventory, including emissions from industry, transportation, power generation and residential activities, was compiled using monthly Chinese provincial activity data and a combination of Chinese and western emission factors. The estimated annual Chinese anthropogenic NMVOC emission was 15 Tg C y^{-1}, including 64% from industries, 24% from residential activities, 10% from transportation, and 1% from power generation. The estimated annual Chinese anthropogenic emission of aromatics was 4.9 Tg C y^{-1}, including 73% from industries, 15% from residential activities, 9% from transportation, and 3% from power generation. Previous estimates of Chinese anthropogenic NMVOC emissions for the years 2005 to 2012 ranged from 10.7 to 29.8 Tg C y^{-1}, with aromatics emissions ranging from 2.1 to 11.3 Tg C y^{-1} (Bo et al., 2008; Q. Zhang et al., 2009; Cao et al., 2011;

Liu et al., 2012; Kurokawa et al., 2013; Li et al., 2014; Stavrakou et al., 2015; Wu et al., 2016). We therefore estimated the uncertainty for the *a priori* Chinese anthropogenic NMVOC emission estimates to be ±200%. Anthropogenic NMVOC emissions for the rest of the Asia were from the inventory compiled by Li et al. (2017) for the year 2010. Anthropogenic NMVOC emissions for Europe, U.S., and the rest of the world were from the European Monitoring and Evaluation Programme (EMEP) inventory (Vestreng, 2003), the U.S. EPA 2005 National Emission Inventory (NEI05) (Brioude et al., 2011; Kim et al., 2011), and the Emission Database for Global Atmospheric Research (EDGAR) inventory (version 2.0) (Olivier et al., 1999), respectively, and scaled to the year 2007 using CO_2 emissions (van Donkelaar et al., 2008).

Post-harvest, in-field burning of crop residue has been recognized as a large seasonal source of NMVOCs in China (Fu et al., 2007; Huang et al., 2012; Stavrakou et al., 2016), but this emission has been severely underestimated in inventories based on satellite burnt area observations (Liu et al., 2015). Huang et al. (2012) estimated the Chinese CO emission from crop residue burning to be 4.0 Tg y^{-1}, based on MODIS daily thermal anomalies, Chinese provincial burnt-biomass data, and emission factors from Akagi et al. (2011). We scaled this CO flux using speciated NMVOC emission factors from crop residue burning from the literature (Hays et al., 2002; Akagi et al., 2011) and then multiplied the resulting NMVOC flux estimate by two. The reason for doubling the scaled NMVOC flux was that the emission factors for many NMVOC species were not measured, such that the sum of the speciated NMVOC emission factors was only half of the measured total NMVOC emission factor (Akagi et al., 2011). This difference may partially explain why the top-down study by Stavrakou et al. (2016) using satellite observations of formaldehyde found that Huang et al. (2012) underestimated the NMVOC flux from crop fires over the North China Plain (NCP) in June by at least a factor of two.

Our resulting *a priori* estimate for Chinese annual NMVOC emissions from biomass burning was 1.1 Tg C y^{-1}, including 0.86 Tg C y^{-1} from crop residue burning (obtained by scaling Huang et al., 2012) and 0.24 Tg C y^{-1} from other types of biomass burning activities (taken from the Global Fire Emissions Database version 3, GFED3) (van der Werf et al., 2010). Previous estimates of Chinese NMVOC emissions from biomass burning for the years 1996 to 2012 ranged from 0.24 to 3.2 Tg C y^{-1} (Fu et al., 2007; van der Werf et al., 2010; Wiedinmyer et al., 2011; Huang et al., 2012; Liu et al., 2015;

Stavrakou et al., 2015, 2016). We therefore assigned an uncertainty of ±300% to the *a priori* Chinese biomass burning NMVOC flux. Biomass burning emissions from the rest of the world were taken from GFED3 (van der Werf et al., 2010).

Figure 1 (a)-(c) show the spatial distribution of the *a priori* Chinese NMVOC emissions from biomass burning, anthropogenic, biogenic, and total sources, respectively. Biomass burning emissions were highest over the NCP and southwest China, reflecting strong emissions from crop residue burning over the NCP in June and from land-clearing burning over southwest China during February to April, respectively. Chinese anthropogenic and biogenic NMVOC sources both showed a general west-to-east gradient, following population and vegetation densities. Biogenic NMVOC emissions reflected the combined modulation by vegetation densities, temperature, and sunlight. Anthropogenic NMVOC fluxes exceeded 10^3 kg C km^{-2} y^{-1} throughout the industrialized and densely populated eastern China, with the highest fluxes over the NCP and around the Yantze River Delta area.

Figure 2 shows the seasonal variation of the *a priori* Chinese NMVOC emissions. The *a priori* anthropogenic NMVOC fluxes were higher during the cold months and lower during the warm months, driven by the seasonal strengths of industrial and residential activities (Li et al., 2017). The *a priori* biogenic NMVOC fluxes showed the opposite seasonal pattern, with more than half of the total annual flux emitted in summer. The *a priori* biomass burning NMVOC source was relatively small except when it peaked in June due to the burst of post-harvest burning over the NCP and in spring due to land-clearing over southwest China. As a result, the Chinese NMVOC emissions were predominantly anthropogenic in January but mainly biogenic in June. During the transition months of April and October, the anthropogenic and biogenic contributions to the total NMVOC emissions were comparable.

2.3 Formaldehyde and glyoxal column concentrations observed by GOME-2A and OMI

We used the monthly mean tropospheric formaldehyde and glyoxal column concentrations retrieved from the Global Ozone Monitoring Experiment-2A (GOME-2A) instrument and the Ozone Monitoring Instrument (OMI) for the year 2007 to constrain Chinese NMVOC sources. The technical

297 details of these four sets of satellite retrievals are summarized in Table 3.

298

299 The native GOME-2A pixel vertical column densities (VCDs) of formaldehyde and glyoxal were
300 retrieved by De Smedt et al. (2012) and Lerot et al. (2010), respectively, using protocols briefly
301 described below. First, pixel slant column densities (SCDs) of formaldehyde and glyoxal were
302 retrieved in the 328.5-346 nm and 435-460 nm windows, respectively, using the Differential Optical
303 Absorption Spectroscopy (DOAS) technique (Platt et al., 1979). Previous glyoxal SCD retrievals often
304 showed biases over remote tropical oceans due to absorption from liquid water (Vrekoussis et al., 2010;
305 Wittrock et al., 2006; Lerot et al., 2010). This bias was corrected in Lerot et al. (2010) by explicitly
306 accounting for liquid water absorption during the DOAS fitting. Second, pixel SCDs were converted
307 into VCDs using air mass factors (AMF), which was calculated using Linearized Discrete Ordinate
308 Radiative Transfer model (LIDORT) (Spurr, 2008) and trace gas profiles simulated by the IMAGE v2
309 model (Stavrakou et al., 2009b). The native pixel VCDs were gridded to daily means at $0.25° \times 0.25°$
310 resolution (De Smedt et al., 2012; Lerot et al., 2010). We further averaged the daily means to monthly
311 means at $5°$ longitude $\times 4°$ latitude resolution. The retrieval errors of the spatially-and-temporally
312 averaged VCDs were estimated to be 30%-40% for formaldehyde and 40% for glyoxal, due to a
313 combination of errors associated with the SCD retrievals, the reference sector correction, the *a priori*
314 profile, and the AMFs (De Smedt et al., 2012; Lerot et al., 2010).

315

316 The OMI native pixel VCDs of formaldehyde and glyoxal were retrieved by González Abad et al.
317 (2015) and Chan Miller et al. (2014), respectively. Briefly, formaldehyde and glyoxal pixel SCDs were
318 retrieved by directly fitting the absorption spectra in the 328.5 – 356.5 nm (formaldehyde) and 435 –
319 461 nm (glyoxal) windows, respectively (Chance, 1998; Kurosu et al., 2004, 2007; Chan Miller et al.,
320 2014), and then converted to pixel VCDs using AMF calculated with a linearized vector discrete
321 ordinate radiative transfer model, VLIDORT (Spurr, 2006), and trace gas profiles simulated by the
322 GEOS-Chem model (González Abad et al., 2015). Liquid water absorption was also explicitly
323 calculated for the glyoxal retrieval (Chan Miller et al., 2014). The typical uncertainties of
324 OMI-observed pixel VCDs over polluted areas were estimated to be 30% to 45% for formaldehyde
325 and 104% for glyoxal (González Abad et al., 2015; Chan Miller et al., 2014). The native pixel VCDs
326 were averaged to monthly means at $5°$ longitude $\times 4°$ latitude resolution. For glyoxal, we further

removed VCDs with signal-to-uncertainty ratios less than 100%. We assumed the retrieval uncertainty of monthly mean OMI formaldehyde and glyoxal VCDs at $4^o \times 5^o$ resolution to be 40% and 100%, respectively.

To remove globally systematic biases in the satellite observations, we adjusted the global observed monthly mean VCDs by aligning the observed VCDs over remote reference areas to those simulated by the GEOS-Chem model (sampled at satellite overpass time) using the *a priori* NMVOC emissions. The remote Pacific (140^o-$160W^o$, 90^oS-90^oN) was chosen as the reference area for formaldehyde (Palmer et al., 2003, 2006; Fu et al., 2007; González Abad et al., 2015). The Sahara desert (20-30^oN, 10^oW-30^oE), where the interference from liquid water absorption was minimal, was chosen as the reference area for glyoxal (Chan Miller et al., 2014). The justification for performing the alignment was two-fold: firstly, the formaldehyde and glyoxal column concentrations over these remote reference areas were small and well simulated by the model (Fu et al., 2008; Chan Miller et al., 2014). The removed biases over the remote areas were less than 20% and 10% of the typical formaldehyde ($>8 \times 10^{15}$ molecule cm^{-2}) and glyoxal ($>4 \times 10^{14}$ molecule cm^{-2}) monthly mean VCDs observed over eastern China, respectively. More importantly, our inversion was performed over China only, assuming that the *a priori* NMVOC emissions for the rest of the world were unbiased. As will be seen in Sections 3 and 4, the optimization of NMVOC sources were predominantly driven by local formaldehyde and glyoxal enhancements produced by relatively short-lived NMVOCs.

2.4 Inversion experiments using the GEOS-Chem adjoint

We used the GEOS-Chem model to perform Bayesian inversions on Chinese NMVOC emissions, using satellite observations of formaldehyde and glyoxal over China and the *a priori* emission estimates as constraints. The inversion minimized the cost function $J(\boldsymbol{x})$ in Eq. (1) (Rodgers, 2000), which we calculated over China:

$$J(\boldsymbol{x}) = \gamma \cdot (\boldsymbol{x} - \boldsymbol{x_a}) \boldsymbol{S_a^{-1}} (\boldsymbol{x} - \boldsymbol{x_a}) + (\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{y}) \boldsymbol{S_o^{-1}} (\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{y}) \quad \text{Eq. (1)}$$

The first and second terms on the right-hand-side of Eq. (1) represented the penalty error and the prediction error, respectively. $\boldsymbol{x}$, which we sought to optimize, was the vector of scale factors (for each

NMVOC species from each emission sector) applied to the *a priori* emissions. $\boldsymbol{x_a}$ was a unit vector applied to the *a priori* NMVOC emission estimates. $\boldsymbol{y}$ was the vector of satellite-observed monthly mean VCDs of the targeted tracer (formaldehyde and/or glyoxal). $\boldsymbol{F(x)}$ was the vector of VCDs of the targeted tracer simulated by the forward model $\boldsymbol{F}$. $\boldsymbol{S_a}$ was the *a priori* emission error covariance matrix, which was a diagonal matrix with the uncertainties estimated based on ranges of previous NMVOC estimates (Section 2.2 and Table 1).

The observation error covariance matrix in Eq. (1), $\boldsymbol{S_o}$, was difficult to quantify, as it included contributions not only from the satellite retrieval, but also from the model representation of chemistry and transport. Zhu et al. (2016) and Chan Miller et al. (2017) compared vertical profiles of GEOS-Chem-simulated formaldehyde and glyoxal over the Southeast U.S. in summer against aircraft measurements. They reported that the simulated formaldehyde mixing ratios showed only a small bias (-3%±2%) in the lower troposphere but were lower than the observations by 41% in the free troposphere, likely due to insufficient deep convection in the model (Zhu et al., 2016). The simulated glyoxal mixing ratios were within 20% of the observations in the mixed layer, but they were too low in the upper troposphere by more than a factor of two, also likely due to insufficient model vertical transport (Chan Miller et al., 2017). It should be noted that these errors assessed by Zhu et al. (2016) and Chan Miller et al. (2017) likely also included the errors associated with precursor emissions. Nevertheless, based on these assessments, we roughly estimated that the model errors for formaldehyde and glyoxal VCDs to be ±80%, ±100%, respectively. Adding these estimated model errors in quadrature to the satellite retrieval errors (Section 2.3), we estimated that the observation error ($\boldsymbol{S_o}$) of formaldehyde and glyoxal to be about ±90% and ±150%, respectively.

The optimization of Eq. (1) was dependent on the relative weighting of the penalty error ($\boldsymbol{S_a}$) and the prediction error ($\boldsymbol{S_o}$). However, the errors and error correlations within $\boldsymbol{S_a}$ and $\boldsymbol{S_o}$ were often incompletely represented. In addition, we found that due to the mathematical formulation of Eq. (1), the cost function $J(\boldsymbol{x})$ was heavily weighted by grids where the *a priori* estimates were too high, such that the optimization was less effective at increasing emissions where the *a priori* emissions were too low. These issues were empirically addressed in inversion studies by the introduction of a regularization factor, γ, in Eq. (1) to adjust the relative weight of the penalty error. Henze et al. (2009)

used the L-curve method (Hansen, 1998) to find an optimized γ value, which minimized the total cost function and balanced the prediction term and the penalty term. We followed that methodology and found a γ value of 0.01 for July, which we applied for all of the warmer months (March to October). An optimized γ value of 0.1 was found for January, and we applied that value to the colder months.

Table 2 shows the setup of our inversion experiments. We experimented with four different sets of satellite retrievals as constraints, with the goal of bracketing the uncertainties of the top-down estimate of Chinese NMVOC emissions. The first two experiments (IE-1 and IE-2) constrained emissions using the formaldehyde and glyoxal VCDs observations from GOME-2A and OMI, respectively. Several studies showed that GOME-2A formaldehyde VCDs may be low by a factor of 1.3 to 1.7 (Lee et al., 2015; Zhu et al., 2016; Wang et al., 2017). As an "upper bound" constraint, we conducted a third inversion experiment (IE-3), which was constrained by 1.7 times the GOME-2A formaldehyde alone. We conducted a fourth inversion experiment (IE-4) constrained by OMI glyoxal VCDs alone to explore the impacts of glyoxal observations on the inversions of anthropogenic emissions.

Figure 3 illustrates our protocol for the inversion experiments. For each month, we began by driving the GEOS-Chem forward model with the *a priori* emissions ($\boldsymbol{x}_{i=1}$= $\boldsymbol{x}_a$=1) to simulate the monthly mean formaldehyde and glyoxal VCDs at satellite-crossing time. The simulated and satellite-observed VCDs were used to calculate the cost function, $J(\boldsymbol{x})$, and the forcing arrays ($\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{F}(\boldsymbol{x})}$). The adjoint of GEOS-Chem was then used to compute the cost function gradient ($\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}}$), and the next guess of the emission scale factor ($\boldsymbol{x}_{i+1}$) was calculated using the Quasi-Newton L-BFGS-B algorithm (Byrd et al., 1995; Zhu et al., 1997), subject to the bounds $0.32 \leq \boldsymbol{x} \leq 10$. These bounds were selected based on the largest uncertainties quoted in the literature on Chinese NMVOC emission estimates (Q. Zhang et al., 2009; Liu et al., 2012). The process was then iterated until the incremental relative reduction of the cost function ($\frac{|J(\boldsymbol{x})_{i+1} - J(\boldsymbol{x})_i|}{\max(J(\boldsymbol{x})_{i+1}, J(\boldsymbol{x})_i)}$) was less than 1% after at least six iterations. We took $\boldsymbol{x}_{i+1}$ from the last iteration as the optimized emission scale factor ($\boldsymbol{x}_p$) and applied it to calculate the top-down emission estimate.

3 Comparison of simulations using the *a priori* emissions against satellite observations and ground-based measurements

We first qualitatively compared the formaldehyde and glyoxal VCDs simulated by the GEOS-Chem model (sampled at satellite overpass times) using the *a priori* emissions against those observed by GOME-2A and OMI, as well as ground-based measurements. Figure 4 (a)-(d) show the monthly mean formaldehyde VCDs observed by GOME-2A over China for January, April, June, and October 2007. Observed formaldehyde VCDs over China showed a distinct west-to-east gradient year-round, which was driven by the higher vegetation and population densities in eastern China. Observed formaldehyde VCDs were higher during the warmer months, reflecting the stronger biogenic emissions during the growing seasons. Highest formaldehyde VCDs were observed over the NCP in June, in response to the large emissions from in-field crop residue burning. In April, high concentrations of formaldehyde were also observed near the southwestern border, reflecting the seasonal biomass burning there.

Figure 4 (e)-(h) show the simulated monthly mean formaldehyde VCDs using the *a priori* emission estimates. The model generally reproduced the observed seasonal contrast and regional patterns. The simulated formaldehyde columns were higher than the GOME-2A observations in January, implying an overestimate of anthropogenic formaldehyde precursors in the *a priori* in January. The simulated formaldehyde columns were lower than the GOME-2A observations over eastern China in June, implying an underestimation of the emissions of formaldehyde precursors in June in the *a priori*.

A few ground-based measurements of tropospheric formaldehyde VCDs have been made in China using the MAX-DOAS technique (Li et al., 2013; Stavrakou et al., 2015; Wang et al., 2017) (Table S1). Figure 4 also shows the seasonal mean of these ground-based measurements at GOME-2A overpass time. In principle, these ground-based measurements were not directly comparable to the satellite-observed and model-simulated formaldehyde columns due to the different inherent uncertainties and the coarse spatial resolution of our analyses. Nevertheless, the seasonal progression presented by these few ground-based measurements were consistent with both the GOME-2A-observed and model-simulated formaldehyde VCDs.

Figure 5 shows the monthly mean glyoxal VCDs observed by GOME-2A. Similar to the case of

formaldehyde, GOME-2A-observed glyoxal VCDs were highest over eastern China in June, reflecting large emissions of NMVOC species that are precursors to both formaldehyde and glyoxal. During January the eastern China glyoxal enhancement was more evident than formaldehyde. As biogenic emissions were small in winter, this indicated that the glyoxal VCDs were more reflective of anthropogenic source. Figure 5 (e)-(h) shows that the simulated glyoxal VCDs were higher than the GOME-2A observations in January and lower than the GOME-2A observations in June. This suggested an overestimation of anthropogenic sources in January and an underestimation of the biogenic sources in June, which was consistent with the constraints implied by the GOME-2A formaldehyde observations. During the transition months of April and October, when the anthropogenic and biogenic contributions to carbonyl productions were presumably more comparable, the simulated glyoxal VCDs were lower than the GOME-2A observations, while the simulated formaldehyde VCDs were higher than the GOME-2A observations (Figure 4 (e)-(h)). This likely indicated that the *a priori* inventory underestimated the emissions of NMVOC species (e.g. aromatics, ethyne, ethene, and glyoxal) that preferentially produced glyoxal, while it overestimated the emissions of species (e.g. $\geq C_4$ alkanes and $\geq C_3$ alkenes from anthropogenic activities) that preferentially produced formaldehyde during the transition months. Ground-based MAX-DOAS glyoxal measurements at a rural southern China site in July 2006 averaged 6.8($\pm$5.2) $\times 10^{14}$ molecules cm^{-2}, higher than both the GOME-2A-observed and simulated glyoxal VCDs. No other ground-based measurements were available to provide spatial and seasonal information.

Figure 6 (a)-(d) shows the monthly mean formaldehyde VCDs observed by the OMI instrument. Similar to the GOME-2A-observed formaldehyde VCDs, OMI formaldehyde VCDs were higher over Eastern China and enhanced during the warmer months. However, the formaldehyde VCDs observed by OMI were lower than those observed by GOME-2A by approximately 30%, likely due to the different retrieval algorithms (De Smedt et al., 2012; González Abad et al., 2015). The simulated formaldehyde VCDs over China were also lower at OMI overpass time than at GOME-2A overpass time by less than 20% in all seasons. However, the ground-based measurements at the three Chinese surface sites did not consistently show such a diurnal pattern.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Figure 7 (a)-(d) shows the monthly mean glyoxal VCDs observed by the OMI instrument. Valid OMI glyoxal VCDs observations were relatively sparse over China, especially during colder months. The seasonal and spatial patterns of the glyoxal VCDs observed by OMI were generally consistent with those observed by GOME-2A. However, the glyoxal VCDs observed by OMI were higher than those observed by GOME-2A except in January. MAX-DOAS measurements of glyoxal at a rural southern China site in July 2006 were also higher in the afternoon than in mid-morning. In contrast, the simulated glyoxal VCDs at OMI overpass time (Figure 7 (e)-(h)) were lower than those at GOME-2A overpass time. This discrepancy among the glyoxal diurnal cycles represented by the MAX-DOAS measurements and the model indicated an uncertainty in the local glyoxal budget.

Figures 6 and 7 also compare the formaldehyde and glyoxal VCDs observed by OMI to those simulated by the model using the *a priori* emission estimates over China. Formaldehyde VCDs observed by OMI were lower than those simulated by the model in all seasons, with the exception of a local hotspot over the NCP in June. However, the glyoxal VCDs observed by OMI were higher than those simulated by the model in all seasons. It thus appeared that the constraints on Chinese NMVOC emissions indicated by the OMI formaldehyde and glyoxal observations were contradictory, even during January and June when the NMVOC emissions over Eastern China were dominated by anthropogenic and biogenic sources, respectively. There may be two explanations for this apparent contradiction indicated by the OMI formaldehyde and glyoxal observations. Firstly, the simulated photochemical budgets of formaldehyde and glyoxal during the local afternoon may be in error. Errors in the model photochemical budget would also explain why the MAX-DOAS measurements of formaldehyde and glyoxal VCDs were both higher in the afternoon than in the morning, while the model showed an opposite diurnal contrast. It is also possible that there were different inherent biases in the OMI formaldehyde and glyoxal retrievals.

4 Top-down estimates of Chinese NMVOC emissions

4.1 *A posteriori* formaldehyde and glyoxal VCDs from inversion experiments

The qualitative analyses in Section 3 showed that the GOME-2A and OMI retrievals of formaldehyde and glyoxal VCDs provide disparate information on seasonal Chinese NMVOC emissions. Thus, our

four inversion experiments on monthly Chinese NMVOC emissions using different satellite observations as constraints (Table 2) represent a range of probable top-down estimates given current satellite observations. Figure 2 shows the monthly top-down Chinese NMVOC emission estimates from the four inversion experiments for January, April, June, and October and compares them against the *a priori* emission estimates. The top-down emission estimates for the full twelve months are shown in Figure S3. Figure S4 shows the changes in the normalized cost functions over China in the four inversion experiments. Relative to their respective initial cost function values, the optimized cost function values were reduced by 10%-60% for all four experiments.

Figure 4 (i-l) and Figure 5 (i-l) show the *a posteriori* monthly mean VCDs of formaldehyde and glyoxal, respectively, from the GOME-2A formaldehyde-glyoxal inversion experiment IE-1. Overall, IE-1 greatly improved the agreement between the *a posteriori* VCDs and the GOME-2A observations for both formaldehyde and glyoxal. The *a posteriori* VCDs of formaldehyde and glyoxal over eastern China both decreased in January and increased in June. During the transition months of April and October, IE-1 decreased the *a posteriori* formaldehyde VCDs while increasing the *a posteriori* glyoxal VCDs. Figure 2 illustrates how these changes in VCDs were driven by the top-down NMVOC emission estimates. For IE-1, the estimated emissions of all NMVOC species were reduced in January but enhanced in June. In April and October, however, IE-1 decreased the total NMVOC emissions while preferentially increasing the emissions of anthropogenic glyoxal precursors.

Figure 6 (i-l) and Figure 7 (i-l) show the *a posteriori* monthly mean VCDs of formaldehyde and glyoxal, respectively, from the OMI formaldehyde-glyoxal inversion experiment IE-2. IE-2 was effective in reducing the *a posteriori* formaldehyde VCDs over eastern China year-round to better agree with the OMI formaldehyde observations. However, the inversion increased the *a posteriori* glyoxal VCDs only slightly and was less effective in bringing agreement with the OMI glyoxal observations. Figure 2 shows that the *a posteriori* NMVOC emission estimates from IE-2 were lower than the *a priori* estimates for all months. This was due to a combination of factors at work in the inversion. The low formaldehyde observations from OMI in all months drove a large reduction in the emissions of NMVOCs that produced only formaldehyde ($\geq C_4$ alkanes and $\geq C_3$ alkenes from

anthropogenic activities, as well as primary formaldehyde from biomass burning). At the same time, the relatively high glyoxal observations from OMI drove an increase in the emissions of NMVOCs that produced mainly glyoxal (ethene, ethyne, and aromatics from anthropogenic activities, as well as primary glyoxal from biomass burning). For precursors that produced large amounts of both formaldehyde and glyoxal (most importantly biogenic isoprene), the inversion reduced the top-down emissions. This was because the formaldehyde observations had more leverage on the inversion due to their lower observational errors. This manifested the importance of well-characterized retrievals with reliable error estimates.

Figure 4 (q-t) shows the *a posteriori* formaldehyde VCDs from the inversion experiment IE-3, which was constrained by the GOME-2A-observed formaldehyde VCDs scaled by a factor of 1.7. The *a posteriori* formaldehyde VCDs in IE-3 were further increased during the warmer months relative to IE-1, especially over the NCP and central China in June. In January, the scaled-up GOME-2A observations were still lower than the simulated formaldehyde VCDs using the *a priori* emissions, leading to a small reduction in the *a posteriori* formaldehyde VCDs. Figure 2 shows that the top-down emission estimates of all NMVOC species were lower than the *a priori* in January and higher than the *a priori* in June. Consequently, although no observations of glyoxal were used as constraints in IE-3, the *a posteriori* glyoxal VCDs also decreased in January and increased in June (Figure 5(m) and (o)). This is in agreement with our findings in Section 3, whereby the constraints exerted by the GOME-2A formaldehyde and glyoxal observations were consistent in January and in June, when the NMVOC emissions were dominated by anthropogenic and biogenic sources, respectively. However, IE-3 had almost no effects on the simulated glyoxal VCDs and the top-down emission estimates of anthropogenic glyoxal precursors in April and October (Figure 5 (n) and (p)). This demonstrated the necessity of glyoxal observations on constraining the emissions of NMVOC species that preferentially produced glyoxal, including most importantly aromatics.

The impacts of satellite glyoxal observations on constraining Chinese NMVOC emission estimates was further demonstrated in IE-4. Figure 7 (m-p) shows the *a posteriori* glyoxal VCDs from IE-4, which used only the OMI glyoxal observations as constraints. The *a posteriori* glyoxal VCDs for all months increased, to an extent greater than those in IE-2. Figure 2 shows that this increase in the *a posteriori*

glyoxal VCDs in IE-4 was achieved mainly by increasing the emission estimates of anthropogenic glyoxal precursors for all months. In June, the emissions of biogenic isoprene (precursor to both glyoxal and formaldehyde) also increased. As a result, the *a posteriori* formaldehyde VCDs in IE-4 increased in June but remained similar to the *a priori* simulation for the other months (Figure 6 (m-p)).

4.2 Top-down estimates of Chinese NMVOC emissions from inversion experiments

Table 2 and Figure 8 shows the top-down estimates for Chinese annual total NMVOC emissions from the four inversion experiments and compare them against the *a priori*. Our top-down annual total estimates for Chinese NMVOCs ranged from 23.4 to 35.4 Tg C y^{-1}, compared to the 27.4 Tg C y^{-1} of the *a priori*. The highest top-down estimate was from IE-3, constrained by 1.7 times the GOME-2A formaldehyde VCD observations. The lowest top-down estimate was from IE-2, due to the relatively low formaldehyde observations from OMI.

Anthropogenic sources constituted 53%-57% of the total top-down NMVOC emissions. The highest top-down anthropogenic emissions estimate was from IE-4 (19.7 Tg C y^{-1}), which reflected the strong traction of the OMI glyoxal observations on constraining anthropogenic NMVOC emissions. The lowest top-down anthropogenic emission estimate was from IE-2 (13.5 Tg C y^{-1}). All four inversion experiments consistently showed larger emissions of anthropogenic glyoxal precursors than the *a priori*. In particular, our top-down estimates for anthropogenic aromatics ranged from 5.0 to 7.3 Tg C y^{-1}, consistently larger than the *a priori* of 4.9 Tg C y^{-1} (Li et al., 2014).

The top-down estimates for biogenic NMVOCs emission ranged between 8.9 and 14.8 Tg C y^{-1}. The top-down estimates for biogenic isoprene were 4.9 to 10.5 Tg C y^{-1}. The top-down estimate for biomass burning NMVOC emissions were between 1.06 to 1.47 Tg C y^{-1}, with the largest top-down estimate driven by the scaled-up GOME-2A formaldehyde VCDs (IE-3).

Figure 9 shows the spatial distribution of the scale factors for Chinese annual NMVOC emissions from each of the four inversion experiments relative to the *a priori* emission estimates. The use of GOME-2A formaldehyde and glyoxal observations as constraints in IE-1 led to a domain-wide increase

589 in biogenic NMVOC emissions, except in the northeast. IE-1 also found an increase in biomass burning
590 emissions over the NCP in June. A similar spatial distribution was found for the emission scale factors
591 of IE-3. Again, this indicated a consistency between the constraints exerted by the formaldehyde and
592 glyoxal observations from GOME-2A. The optimized emission scale factors from IE-2 and IE-4 were
593 of opposite signs. Using only OMI glyoxal observations as constraints in IE-4 led to a domain-wide
594 increase in NMVOC emissions from all sectors. However, when further constraints of the relatively
595 low OMI formaldehyde observations were added in IE-2, the top-down emission estimates decreased
596 across the domain.
597
598 As discussed previously, our four inversion experiments using different satellite retrievals as
599 constraints represent the range of probable top-down estimates given currently available satellite
600 observations. To represent the difference between these top-down estimates relative to the *a priori*, we
601 averaged the top-down estimates from the four inversion experiments. Our averaged top-down estimate
602 for Chinese total annual NMVOC emissions was 30.8 Tg C y^{-1}, including 17.0 Tg C y^{-1}, 12.6 Tg C y^{-1},
603 and 1.2 Tg C y^{-1} from anthropogenic, biogenic, and biomass burning sources, respectively. Our average
604 estimate for anthropogenic aromatic flux was 6.1 Tg C y^{-1}, which was 24% larger than the *a priori*
605 estimate of Li et al. (2014).
606
607 Figure 1 (e-l) shows the spatial distribution of annual Chinese NMVOC emission of our averaged
608 top-down estimate and the scale factors relative to the *a priori* estimates. Our averaged top-down
609 estimate of Chinese NMVOC emissions were spatially consistent with the *a priori*, but the total fluxes
610 increased by 10% to 40% throughout eastern China relative to the *a priori*. In particular, we found a 40%
611 increase in the biomass burning emissions over the NCP. We also found a 20%-40% increase in the
612 anthropogenic NMVOC emissions in coastal eastern China. Largest scale factors for biogenic NMVOC
613 were found near the northwestern border of China and along the northeast-to-southwest division line of
614 vegetation density. This potentially indicated an underestimation of biogenic NMVOC emission from
615 semi-arid ecosystems in the MEGAN inventory.
616

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

5 Comparison with previous estimates of Chinese NMVOC emissions

Table 1 compares our top-down estimates of Chinese annual NMVOC emissions for the year 2007 against estimates in the literature for the years between 2005 and 2012. It should be noted that bottom-up inventories often estimated the total NMVOCs emitted from a source sector using emission factors for total NMVOCs, then distributed the emissions using different species profile data. As a result, bottom-up estimates often included additional NMVOC species not represented here in our study.

Our top-down estimate for biogenic NMVOC emissions range from 8.9 to 14.8 Tg C y^{-1}, on average 17% larger than the flux calculated from the MEGAN inventory (Guenther et al., 2006). Our top-down estimate for isoprene emission (the single most emitted NMVOC species) ranged from 4.9 to 10.5 Tg C y^{-1}, bracketing the *a priori* of 6.6 Tg C y^{-1}. Stavrakou et al. (2015) previously used GOME-2A and OMI formaldehyde observations to derive top-down estimates of isoprene emissions over China of 5.0 Tg C y^{-1} (GOME-2A) and 5.5 Tg C y^{-1} (OMI), respectively. In comparison, our top-down isoprene emission estimates constrained by GOME-2A and OMI (both formaldehyde and glyoxal) observations, were 8.2 Tg C y^{-1} (from IE-1) and 4.9 Tg C y^{-1} (from IE-2), respectively. Our top-down estimates constrained by GOME-2A observations was larger than that of Stavrakou et al. (2015) due to the additional glyoxal constraints. Our estimate constrained by OMI observations was lower than that of Stavrakou et al. (2015) because the OMI formaldehyde VCDs over China retrieved by González Abad et al. (2015) were systematically lower than the OMI formaldehyde VCDs retrieved by De Smedt et al. (2015).

Our top-down estimates for Chinese annual biomass burning NMVOC emissions ranged from 1.06 to 1.47 Tg C y^{-1}. These numbers are in good agreement with the estimate of Huang et al. (2012). Previous bottom-up biomass burning NMVOC emission estimates by Bo et al. (2008) and Wu et al. (2016) ranged from 1.9 to 2.4 Tg C y^{-1}, but only 25% to 30% of these emissions were from open burning of crop residues; the rest were emitted from biofuel burning. The GFED3 inventory (van der Werf et al., 2010), based on satellite burnt area observations, severely underestimated biomass burning emissions over China, particularly those associated with crop residue burning. Top-down estimate of Chinese biomass burning NMVOC emissions by Stavrakou et al. (2015) was between 1.1-1.5 Tg C y^{-1}, very

close to our top-down estimate range (1.06-1.47 Tg C y^{-1}). Similar to Fu et al. (2007) and Stavrakou et al. (2016), our study also highlighted the large emissions from crop residue over the NCP in June (Figure 2).

Previous bottom-up estimates of Chinese anthropogenic NMVOC emissions ranged widely from 10.7 to 29.8 Tg C y^{-1} (Bo et al., 2008; Cao et al., 2011; Kurokawa et al., 2013; Li et al., 2014; Wu et al., 2016) due to the use of different emission factors, activity data, and statistical models. Previous top-down estimates of Chinese anthropogenic NMVOC emissions ranged from 17.3-28.7 Tg C y^{-1} (Liu et al., 2012; Stavrakou et al., 2015). Our top-down estimates had a smaller range between 13.5 to 19.7 Tg C y^{-1}. Our top-down estimates for anthropogenic aromatics (5.0 to 7.3 Tg C y^{-1}) were approximately middle-of-the-range relative to previous estimates of 2.1-11.3 Tg C y^{-1}. The large difference between previous top-down estimates and our top-down estimates of anthropogenic NMVOCs were predominantly due to the choices of satellite observation constraints, and to a lesser extent due to the choices of chemical transport model, the NMVOC species modeled, and the *a priori* emission estimates. Specifically, the much higher estimate of anthropogenic aromatic emission by Liu et al. (2012) (11.3 Tg C y^{-1}) compared to our top-down estimates (5.0-7.3 Tg C y^{-1}) was due to (1) the high glyoxal VCDs observed by the SCIAMACHY instrument compared to those observed the GOME-2A and OMI instruments over China; (2) the assumption made by Liu et al. (2012) that all anomalous glyoxal was produced by aromatics oxidation; (3) the lower yields of glyoxal from aromatics oxidation used in Liu et al. (2012) than those used in our model.

Our four inversion experiments all indicated stronger anthropogenic NMVOC emissions in summer than in winter. In contrast, the *a priori* estimates showed a slightly stronger NMVOC emission in winter than in summer, which was driven by stronger activity levels in winter in the bottom-up inventories along with seasonally-invariant emission factors (Li et al., 2017). However, studies showed that the NMVOC emission factors, in particular those for transport and industrial sectors, were strongly and positively correlated with temperature (Rubin et al., 2006; Wei et al., 2016).

6 Impacts on simulated surface ozone levels over China

We evaluated the impacts on surface ozone concentrations due to our average top-down emission estimates of NMVOCs. Figure 10 compares the monthly mean afternoon (13:00-17:00 LT) surface ozone concentrations simulated using our averaged top-down emission estimates against those simulated using the *a priori* emissions for June and December 2007. Also shown in Figure 10 are surface observations at representative regional sites (Li et al., 2007; Xu et al., 2008; J. M. Zhang et al., 2009; Zheng et al., 2010; Wang et al., 2012; Wang et al., 2015; Li and Bian, 2015; Sun et al., 2016; Xu et al., 2016).Using the *a priori* emissions, the highest simulated afternoon surface ozone concentrations were between 100-110 ppb over the NCP in June. This was lower than the observations at two sites in the NCP, including a rural site outside Beijing (>100 ppb) and Mt. Tai (108 ppb). In comparison, by using our averaged top-down NMVOC emission estimate, the simulated afternoon surface ozone increased by 5-10 ppb over the NCP in June and were in better agreement with the observations. In December, the simulated afternoon surface ozone using *a priori* emission consistently overestimated the observed concentrations in eastern China. In comparison, by using our averaged top-down NMVOC emission estimates, the simulated afternoon surface ozone over eastern China decreased by 5 to 13 ppb, again in better agreement with the observations. It thus appears that our top-down emission estimates for Chinese NMVOCs improved simulation of regional ozone.

7 Conclusions

We used the GEOS-Chem model and its adjoint, as well as satellite observations of tropospheric column concentrations of formaldehyde and glyoxal, to constrain monthly Chinese NMVOC emissions from anthropogenic, biogenic, and biomass burning sources for the year 2007. We updated the gas-phase chemistry in the GEOS-Chem model and constructed its adjoint. The *a priori* emission estimates were taken from widely-used bottom-up emission inventories. We conducted four inversion experiments, which were constrained by formaldehyde and glyoxal observations from the GOME-2A instrument (IE-1), formaldehyde and glyoxal observations from the OMI instrument (IE-2), 1.7 times the formaldehyde observations from the GOME-2A instrument (IE-3), and glyoxal observations from the OMI instrument (IE-4), respectively. The results from these experiments represented the range of probable top-down NMVOC emission estimates for China given current satellite observational

constraints.

Our top-down estimates of total annual Chinese NMVOC emission from the four inversion experiments ranged from 23.4 to 35.4 Tg C y^{-1}. Our top-down estimates of Chinese anthropogenic NMVOC emission was 13.5 to 19.7 Tg C y^{-1}. In particular, we top-down estimate of Chinese anthropogenic aromatic emissions range from 5.0 to 7.3 Tg C y^{-1}, much smaller than the top-down estimate of 11.3 Tg C y^{-1} by Liu et al. (2012). Our top-down estimate of Chinese biogenic NMVOC emission ranged from 8.9 to 14.8 Tg C y^{-1}, with 4.9 to 10.5 Tg C y^{-1} attributed to isoprene. Our top-down estimate for Chinese biomass burning NMVOC emission range from 1.1 to 1.5 Tg C y^{-1} and was mostly associated with seasonal open burning of crop residue after local harvests, such as over the NCP in June.

We evaluated the impacts on regional surface ozone concentrations from our average top-down Chinese NMVOC emission estimates. We found that the simulated monthly mean afternoon surface ozone concentrations increased by 5-12 ppb over the NCP in June, compared to the *a priori* simulation. In December, the simulated monthly mean afternoon surface ozone concentrations decreased by 5-13 ppb over northern and central China, compared to the *a priori* simulation. For both seasons, the simulation using our averaged top-down emission estimates were in better general agreement with regional surface observations.

We concluded that formaldehyde and glyoxal observations from GOME-2A and OMI provide quantitative constraints on the monthly emissions of Chinese NMVOCs. In particular, the simultaneous use of the observations of both species helps distinguish NMVOC precursors and thus provides better quantification of individual sources. However, better validation of these satellite data over China are urgently needed, particularly in terms of discrepancies between different retrievals for the same species.

Acknowledgements

This work was supported by the Ministry of Science and Technology of China (2014CB441303) and the National Natural Sciences Foundation of China (41461164007, 41222035). We thank National

732 Super Computer Center in Tianjin for supporting this work. D. K. Henze recognizes support from
733 NASA NNX17AF63G.

734 References

735 Acarreta, J. R., De Haan, J. F., and Stammes, P.: Cloud pressure retrieval using the O2-O2absorption
736 band at 477 nm, J. Geophys. Res., 109, doi: 10.1029/2003jd003915, 2004.

737 Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and
738 Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric
739 models, Atmos. Chem. Phys., 11, 4039-4072, doi: 10.5194/acp-11-4039-2011, 2011.

740 Arey, J., Obermeyer, G., Aschmann, S. M., Chattopadhyay, S., Cusick, R. D., and Atkinson, R.:
741 Dicarbonyl Products of the OH Radical-Initiated Reaction of a Series of Aromatic Hydrocarbons,
742 Environ. Sci. Technol., 43, 683-689, doi: 10.1021/es8019098, 2009.

743 Barkley, M. P., Palmer, P. I., Kuhn, U., Kesselmeier, J., Chance, K., Kurosu, T. P., Martin, R. V.,
744 Helmig, D., and Guenther, A.: Net ecosystem fluxes of isoprene over tropical South America inferred
745 from Global Ozone Monitoring Experiment (GOME) observations of HCHO columns, J. Geophys.
746 Res., 113, doi: 10.1029/2008jd009863, 2008.

747 Barkley, M. P., Palmer, P. I., De Smedt, I., Karl, T., Guenther, A., and Van Roozendael, M.: Regulated
748 large-scale annual shutdown of Amazonian isoprene emissions?, Geophys. Res. Lett., 36,
749 doi:10.1029/2008gl036843, 2009.

750 Barkley, M. P., Smedt, I. D., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D.,
751 Guenther, A., Paulot, F., Marais, E., and Mao, J.: Top-down isoprene emissions over tropical South
752 America inferred from SCIAMACHY and OMI formaldehyde columns, J. Geophys. Res. Atmos., 118,
753 6849-6868, doi:10.1002/jgrd.50552, 2013.

754 Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y.,
755 Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated
756 meteorology: Model description and evaluation, J. Geophys. Res., 106, 23073-23095, doi:
757 10.1029/2001JD000807, 2001.

758 Bloss, C., Wagner, V., Jenkin, M. E., and Volkamer, R.: Development of a detailed chemical
759 mechanism (MCMv3.1) for the atmospheric oxidation of aromatic hydrocarbons, Atmos. Chem. Phys.,
760 5, 641-664, doi:10.5194/acp-5-641-2005, 2005.

761 Bo, Y., Cai, H., and Xie, S. D.: Spatial and temporal variation of historical anthropogenic NMVOCs
762 emission inventories in China, Atmos. Chem. Phys., 8, 7297-7316, doi: 10.5194/acp-8-7297-2008,
763 2008.

764 Bolscher, M., Pulles, T., Brand, R., Pereira, J., Mota, B., Spessa, A., Dalsoren, S., Nojie, T., and Szopa,
765 S.: Emission data sets and methodologies for estimating emissions, RETRO Deliverable D1-6, 2007.

766 Brioude, J., Kim, S. W., Angevine, W. M., Frost, G. J., Lee, S. H., McKeen, S. A., Trainer, M.,
767 Fehsenfeld, F. C., Holloway, J. S., Ryerson, T. B., Williams, E. J., Petron, G., and Fast, J. D.: Top-down
768 estimate of anthropogenic emission inventories and their interannual variability in Houston using a
769 mesoscale inverse modeling technique, J. Geophys. Res. Atmos., 116, doi:10.1029/2011JD016215,
770 2011.

771 Byrd, R. H., Lu, P. H., Nocedal, J., and Zhu, C. Y.: A Limited Memory Algorithm for Bound
772 Constrained Optimization, Siam J. Sci. Comput., 16, 1190-1208, doi: 10.1137/0916069, 1995.

773 Cao, G., Zhang, X., Gong, S., An, X., and Wang, Y.: Emission inventories of primary particles and
774 pollutant gases for China, Chin. Sci. Bull., 56, 781-788, doi:10.1007/s11434-011-4373-7, 2011.

775 Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J., and Chance, K.:
776 Glyoxal retrieval from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 7, 3891-3907,
777 doi:10.5194/amt-7-3891-2014, 2014.

778 Chan Miller, C., Jacob, D. J., Abad, G. G., and Chance, K.: Hotspot of glyoxal over the Pearl River
779 delta seen from the OMI satellite instrument: implications for emissions of aromatic hydrocarbons,
780 Atmos. Chem. Phys., 16, 4631-4639, doi: 10.5194/acp-16-4631-2016, 2016.

781 Chan Miller, C., Jacob, D. J., Marais, E. A., Yu, K. R., Travis, K. R., Kim, P. S., Fisher, J. A., Zhu, L.,
782 Wolfe, G. M., Hanisco, T. F., Keutsch, F. N., Kaiser, J., Min, K. E., Brown, S. S., Washenfelder, R. A.,
783 Abad, G. G., and Chance, K.: Glyoxal yield from isoprene oxidation and relation to formaldehyde:
784 chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI
785 satellite data, Atmos. Chem. Phys., 17, 8725-8738, doi: 10.5194/acp-17-8725-2017, 2017.

786 Chance, K.: Analysis of BrO measurements from the Global Ozone Monitoring Experiment, Geophys.
787 Res. Lett., 25, 3335-3338, doi: 10.1029/98gl52359, 1998.

788 Curci, G., Palmer, P. I., Kurosu, T. P., Chance, K., and Visconti, G.: Estimating European volatile
789 organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring
790 Instrument, Atmos. Chem. Phys., 10, 11501-11517, doi: 10.5194/acp-10-11501-2010, 2010.

791 Daescu, D. N., Sandu, A., and Carmichael, G. R.: Direct and adjoint sensitivity analysis of chemical
792 kinetic systems with KPP: II—numerical validation and applications, Atmos. Environ., 37, 5097-5114,
793 doi:10.1016/j.atmosenv.2003.08.020, 2003.

794 De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J. F., Lerot, C., Theys, N., Valks, P., Hao, N.,
795 and van der A, R.: Improved retrieval of global tropospheric formaldehyde columns from
796 GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos. Meas.
797 Tech., 5, 2933-2949, doi:10.5194/amt-5-2933-2012, 2012.

798 De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C.,
799 Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J. F., and Van Roozendael, M.:
800 Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined
801 OMI and GOME-2 observations, Atmos. Chem. Phys., 15, 12519-12545, doi:
802 10.5194/acp-15-12519-2015, 2015.

803 Dufour, G., Wittrock, F., Camredon, M., Beekmann, M., Richter, A., Aumont, B., and Burrows, J. P.:
804 SCIAMACHY formaldehyde observations: constraint for isoprene emission estimates over Europe?,
805 Atmos. Chem. Phys., 9, 1647-1664, doi:10.5194/acp-9-1647-2009, 2009.

806 Fu, T.-M., Jacob, D. J., Palmer, P. I., Chance, K., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C.,
807 and Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic
808 compound emissions in east and south Asia and implications for ozone, J. Geophys. Res., 112, doi:
809 10.1029/2006jd007853, 2007.

810 Fu, T.-M., Jacob, D. J., Wittrock, F., Burrows, J. P., Vrekoussis, M., and Henze, D. K.: Global budgets
811 of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic
812 aerosols, J. Geophys. Res., 113, doi:10.1029/2007jd009505, 2008.

813 González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated
814 Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde
815 retrieval, Atmos. Meas. Tech., 8, 19-32, doi:10.5194/amt-8-19-2015, 2015.

816 Gonzi, S., Palmer, P. I., Barkley, M. P., De Smedt, I., and Van Roozendael, M.: Biomass burning
817 emission estimates inferred from satellite column measurements of HCHO: Sensitivity to co-emitted
818 aerosol and injection height, Geophys. Res. Lett., 38, doi: 10.1029/2011gl047890, 2011.

819 Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global
820 terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature),
821 Atmos. Chem. Phys., 6, 3181-3210, doi: 10.5194/acp-6-3181-2006, 2006.

822 Han, K. M., Park, R. S., Kim, H. K., Woo, J. H., Kim, J., and Song, C. H.: Uncertainty in biogenic
823 isoprene emissions and its impacts on tropospheric chemistry in East Asia, Sci. Total Environ., 463-464,
824 754-771, doi: 10.1016/j.scitotenv.2013.06.003, 2013.

825 Hansen, P. C.: Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion,
826 SIAM, Philadelphia, 1998.

827 Hays, M. D., Geron, C. D., Linna, K. J., Smith, N. D., and Schauer, J. J.: Speciation of gas-phase and
828 fine particle emissions from burning of foliar fuels, Environ. Sci. Technol., 36, 2281-2295,
829 doi:10.1021/es0111683, 2002.

830 Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos.
831 Chem. Phys., 7, 2413-2433, doi: 10.5194/acp-7-2413-2007, 2007.

832 Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality
833 influences of inorganic PM2.5 precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem.
834 Phys., 9, 5877-5903, doi:10.5194/acp-9-5877-2009, 2009.

835 Huang, X., Li, M., Li, J., and Song, Y.: A high-resolution emission inventory of crop burning in fields
836 in China based on MODIS Thermal Anomalies/Fire products, Atmos. Environ., 50, 9-15, doi:
837 10.1016/j.atmosenv.2012.01.017, 2012.

838 Ip, H. S. S., Huang, X. H. H., and Yu, J. Z.: Effective Henry's law constants of glyoxal, glyoxylic acid,
839 and glycolic acid, Geophys. Res. Lett., 36, doi: 10.1029/2008GL036212, 2009.

840 Jenkin, M. E., Young, J. C., and Rickard, A. R.: The MCM v3.3.1 degradation scheme for isoprene,
841 Atmos. Chem. Phys., 15, 11433-11459, doi: 10.5194/acp-15-11433-2015, 2015.

842 Kim, S. W., McKeen, S. A., Frost, G. J., Lee, S. H., Trainer, M., Richter, A., Angevine, W. M., Atlas, E.,
843 Bianco, L., Boersma, K. F., Brioude, J., Burrows, J. P., de Gouw, J., Fried, A., Gleason, J., Hilboll, A.,
844 Mellqvist, J., Peischl, J., Richter, D., Rivera, C., Ryerson, T., Hekkert, S. T. L., Walega, J., Warneke, C.,
845 Weibring, P., and Williams, E.: Evaluations of NOx and highly reactive VOC emission inventories in
846 Texas and their implications for ozone plume simulations during the Texas Air Quality Study 2006,
847 Atmos. Chem. Phys., 11, 11361-11386, doi: 10.5194/acp-11-11361-2011, 2011.

848 Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology
849 from 3 years of OMI data, J. Geophys. Res. Atmos., 113, doi: 10.1029/2008JD010290, 2008.

850 Kleipool, Q. L.: Transient signal flagging algorithm definition for radiance data, Tech. Rep.
851 TN-OMIE-KNMI-717 TN-OMIEKNMI-717 TN-OMIE-KNMI-717 TN-OMIE-KNMI-717
852 TNOMIE-KNMI-717, Royal Netherlands Meteorological Institute, De Bilt, the Netherlands, 2005.

853 Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima,
854 K., and Akimoto, H.: Emissions of air pollutants and greenhouse gases over Asian regions during
855 2000-2008: Regional Emission inventory in ASia (REAS) version 2, Atmos. Chem. Phys., 13,
856 11019-11058, doi:10.5194/acp-13-11019-2013, 2013.

857 Kurosu, T. P., Chance, K., and Sioris, C. E.: Preliminary results for HCHO and BrO from the
858 EOS-Aura Ozone Monitoring Instrument,Conference on Passive Optical Remote Sensing of the
859 Atmosphere and Clouds IV, 5652, 116, doi: 10.1117/12.578606, 2004.

860 Kurosu, T. P., Chance, K., Liu, X., Volkamer, R., Fu, T.-M., Millet, D., and Jacob, D. J.: Seasonally
861 resolved global distributions of glyoxal and formaldehyde observed from the Ozone Monitoring
862 Instrument on EOS Aura, Simpósio Brasileiro de Sensoriamento Remoto 13 (SBSR): 6461–6464,
863 2007.

864 Lee, H., Ryu, J., Irie, H., Jang, S.-H., Park, J., Choi, W., and Hong, H.: Investigations of the Diurnal
865 Variation of Vertical HCHO Profiles Based on MAX-DOAS Measurements in Beijing: Comparisons
866 with OMI Vertical Column Data, Atmosphere, 6, 1816-1832, doi: 10.3390/atmos6111816, 2015.

867 Lelieveld, J., Butler, T. M., Crowley, J. N., Dillon, T. J., Fischer, H., Ganzeveld, L., Harder, H.,
868 Lawrence, M. G., Martinez, M., Taraborrelli, D., and Williams, J.: Atmospheric oxidation capacity
869 sustained by a tropical forest, Nature, 452, 737-740, doi: 10.1038/nature06870, 2008.

870 Lerot, C., Stavrakou, T., De Smedt, I., Muller, J. F., and Van Roozendael, M.: Glyoxal vertical columns
871 from GOME-2 backscattered light measurements and comparisons with a global model, Atmos. Chem.
872 Phys., 10, 12059-12072, doi: 10.5194/acp-10-12059-2010, 2010.

873 Li, D., and Bian, J. C.: Observation of a Summer Tropopause Fold by Ozonesonde at Changchun,

China: Comparison with Reanalysis and Model Simulation, Adv. Atmos. Sci., 32, 1354-1364, doi: 10.1007/s00376-015-5022-x, 2015.

Li, J., Wang, Z. F., Akimoto, H., Gao, C., Pochanart, P., and Wang, X. Q.: Modeling study of ozone seasonal cycle in lower troposphere over east Asia, J. Geophys. Res.Atmos., 112, doi: 10.1029/2006JD008209, 2007.

Li, J. Y., Mao, J. Q., Min, K. E., Washenfelder, R. A., Brown, S. S., Kaiser, J., Keutsch, F. N., Volkamer, R., Wolfe, G. M., Hanisco, T. F., Pollack, I. B., Ryerson, T. B., Graus, M., Gilman, J. B., Lerner, B. M., Warneke, C., de Gouw, J. A., Middlebrook, A. M., Liao, J., Welti, A., Henderson, B. H., McNeill, V. F., Hall, S. R., Ullmann, K., Donner, L. J., Paulot, F., and Horowitz, L. W.: Observational constraints on glyoxal production from isoprene oxidation and its contribution to organic aerosol over the Southeast United States, J. Geophys. Res. Atmos., 121, 9849-9861, doi: 10.1002/2016JD025331, 2016.

Li, M., Zhang, Q., Streets, D. G., He, K. B., Cheng, Y. F., Emmons, L. K., Huo, H., Kang, S. C., Lu, Z., Shao, M., Su, H., Yu, X., and Zhang, Y.: Mapping Asian anthropogenic emissions of non-methane volatile organic compounds to multiple chemical mechanisms, Atmos. Chem. Phys., 14, 5617-5638, doi:10.5194/acp-14-5617-2014, 2014.

Li, M., Zhang, Q., Kurokawa, J. I., Woo, J. H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., and Zheng, B.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, Atmos. Chem. Phys., 17, 935-963, doi:10.5194/acp-17-935-2017, 2017.

Li, X., Brauers, T., Hofzumahaus, A., Lu, K., Li, Y. P., Shao, M., Wagner, T., and Wahner, A.: MAX-DOAS measurements of NO2, HCHO and CHOCHO at a rural site in Southern China, Atmos. Chem. Phys., 13, 2133-2151, doi: 10.5194/acp-13-2133-2013, 2013.

Liggio, J., Li, S. M., and McLaren, R.: Reactive uptake of glyoxal by particulate matter, J. Geophys. Res. Atmos., 110, doi: 10.1029/2004JD005113, 2005.

Liu, M., Song, Y., Yao, H., Kang, Y., Li, M., Huang, X., and Hu, M.: Estimating emissions from agricultural fires in the North China Plain based on MODIS fire radiative power, Atmos. Environ., 112, 326-334, doi: 10.1016/j.atmosenv.2015.04.058, 2015.

Liu, Z., Wang, Y., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., Shao, M., Chang, C.-C., Liu, S.-C., Wang, H., and Chen, C.: Exploring the missing source of glyoxal (CHOCHO) over China, Geophys. Res. Lett., 39, doi:10.1029/2012gl051645, 2012.

Mao, J. Q., Paulot, F., Jacob, D. J., Cohen, R. C., Crounse, J. D., Wennberg, P. O., Keller, C. A., Hudman, R. C., Barkley, M. P., and Horowitz, L. W.: Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry, J. Geophys. Res. Atmos., 118, 11256-11268, doi: 10.1002/jgrd.50817, 2013.

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., and Mao, J.: Isoprene emissions in Africa inferred from OMI

observations of formaldehyde columns, Atmos. Chem. Phys., 12, 6219-6235, doi: 10.5194/acp-12-6219-2012, 2012.

Marais, E. A., Jacob, D. J., Guenther, A., Chance, K., Kurosu, T. P., Murphy, J. G., Reeves, C. E., and Pye, H. O. T.: Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter, Atmos. Chem. Phys., 14, 7693-7703, doi: 10.5194/acp-14-7693-2014, 2014a.

Marais, E. A., Jacob, D. J., Wecht, K., Lerot, C., Zhang, L., Yu, K., Kurosu, T. P., Chance, K., and Sauvage, B.: Anthropogenic emissions in Nigeria and implications for atmospheric ozone pollution: A view from space, Atmos. Environ., 99, 32-40, doi: 10.1016/j.atmosenv.2014.09.055, 2014b.

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., and Clarke, A. D.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission, J. Geophys. Res., 111, doi: 10.1029/2005jd006853, 2006.

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K., Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, J. Geophys. Res., 113, doi: 10.1029/2007jd008950, 2008.

Monks, P. S.: Gas-phase radical chemistry in the troposphere, Chem. Soc. Rev., 34, 376-395, doi:10.1039/b307982c, 2005.

Myriokefalitakis, S., Vrekoussis, M., Tsigaridis, K., Wittrock, F., Richter, A., Bruehl, C., Volkamer, R., Burrows, J. P., and Kanakidou, M.: The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, Atmos. Chem. Phys., 8, 4965-4981, doi: 10.5194/acp-8-4965-2008, 2008.

Nishino, N., Arey, J., and Atkinson, R.: Formation Yields of Glyoxal and Methylglyoxal from the Gas-Phase OH Radical-Initiated Reactions of Toluene, Xylenes, and Trimethylbenzenes as a Function of NO2 Concentration, J. Phys. Chem. A, 114, 10140, doi: 10.1021/jp105112h, 2010.

Olivier, J. G. J., Bouwman, A. F., Berdowski, J. J. M., Veldt, C., Bloos, J. P. J., Visschedijk, A. J. H., van der Maas, C. W. M., and Zandveld, P. Y. J.: Sectoral emission inventories of greenhouse gases for 1990 on a per country basis as well as on 1 °×1 °, Environ. Sci. Policy, 2, 241-263, doi:10.1016/S1462-9011(99)00027-1, 1999.

Olivier, J. G. J., Berdowski, J. J. M., Peters, J. A. H. W., Bakker, J., Visschedijk, A. J. H., and Bloos, J.-P. J.: Applications of EDGAR, Including a description of EDGAR 3.0: reference database with trend data for 1970–1995, RIVM report no. 773301 001/ NOP report no. 410200 051, RIVM, Bilthoven, 2001.

Olivier, J. G. J.: Part III: Greenhouse gas emissions. 1. Shares and trends in greenhouse gas emissions; 2. Sources and methods: greenhouse gas emissions for 1990 and 1995 in "CO2 emissions from fuel combustion 1971–2000", International Energy Agency, Paris, ISBN 92-64-09794-5, 1–31, 1–31, 2002.

948 Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K., and Kurosu, T. P.: Mapping isoprene
949 emissions over North America using formaldehyde column observations from space, J. Geophys. Res.
950 Atmos., 108, doi: 10.1029/2002jd002153, 2003.

951 Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer,
952 C., Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B., and Sumner, A. L.: Quantifying the seasonal
953 and interannual variability of North American isoprene emissions using satellite observations of the
954 formaldehyde column, J. Geophys. Res., 111, doi: 10.1029/2005jd006689, 2006.

955 Paulot, F., Crounse, J. D., Kjaergaard, H. G., Kroll, J. H., Seinfeld, J. H., and Wennberg, P. O.: Isoprene
956 photooxidation: new insights into the production of acids and organic nitrates, Atmos. Chem. Phys., 9,
957 1479-1501, doi: 10.5194/acp-9-1479-2009, 2009a.

958 Paulot, F., Crounse, J. D., Kjaergaard, H. G., Kurten, A., St Clair, J. M., Seinfeld, J. H., and Wennberg,
959 P. O.: Unexpected Epoxide Formation in the Gas-Phase Photooxidation of Isoprene, Science, 325,
960 730-733, doi: 10.1126/science.1172910, 2009b.

961 Platt, U., Perner, D., auml, and tz, H. W.: Simultaneous measurement of atmospheric CH2O, O3, and
962 NO2 by differential optical absorption, J.Geophys. Res. Oceans, 84, 6329-6335,
963 doi:10.1029/JC084iC10p06329, 1979.

964 Qiu, K., Yang, L., Lin, J., Wang, P., Yang, Y., Ye, D., and Wang, L.: Historical industrial emissions of
965 non-methane volatile organic compounds in China for the period of 1980–2010, Atmos. Environ., 86,
966 102-112, do:10.1016/j.atmosenv.2013.12.026, 2014.

967 Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, World Scientific,
968 Singapore, 2000.

969 Rubin, J. I., Kean, A. J., Harley, R. A., Millet, D. B., and Goldstein, A. H.: Temperature dependence of
970 volatile organic compound evaporative emissions from motor vehicles, J. Geophys. Res. Atmos., 111,
971 doi:10.1029/2005JD006458, 2006.

972 Sandu, A., Daescu, D. N., and Carmichael, G. R.: Direct and adjoint sensitivity analysis of chemical
973 kinetic systems with KPP: Part I - theory and software tools, Atmos. Environ., 37, 5083-5096,
974 doi:10.1016/j.atmosenv.2003.08.019, 2003.

975 Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J.: Protocol for the development of the
976 Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile
977 organic compounds, Atmos. Chem. Phys., 3, 181-193, doi: 10.5194/acp-3-161-2003, 2003.

978 Schultz, M. G., Backman, L., and Balkanski, Y.: REanalysis of the TROpospheric chemical
979 composition over the past 40 years (RETRO): A long-term global modeling study of tropospheric
980 chemistry, J ¨ulich/Hamburg, Germany, 48/2007 report on Earth System Science of the Max Planck
981 Institute for Meteorology, Hamburg, http://retro.enes.org, ISSN 1614-1199, 2007.

982 Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S., and Chance, K.: Constraining global isoprene
983 emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, J.

984 Geophys. Res., 110, doi: 10.1029/2004jd005629, 2005.

985 Spurr, R.: LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate
986 radiative transfer models for use in remote sensing retrieval problems, in: Light Scattering Reviews,
987 edited by: Kokhanovsky, A., Springer, 3, 229–275, 2008.
988
989 Spurr, R. J. D.: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer
990 code for forward model and retrieval studies in multilayer multiple scattering media, J. Quant.
991 Spectrosc. Radiat. Transf., 102, 316-342, doi: 10.1016/j.jqsrt.2006.05.005, 2006.

992 Stavrakou, T., Muller, J. F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M.,
993 Wittrock, F., Richter, A., and Burrows, J. P.: The continental source of glyoxal estimated by the
994 synergistic use of spaceborne measurements and inverse modelling, Atmos. Chem. Phys., 9, 8431-8446,
995 doi: 10.5194/acp-9-8431-2009, 2009a.

996 Stavrakou, T., Muller, J. F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and
997 Guenther, A.: Global emissions of non-methane hydrocarbons deduced from SCIAMACHY
998 formaldehyde columns through 2003-2006, Atmos. Chem. Phys., 9, 3663-3679,
999 doi:10.5194/acp-9-3663-2009, 2009b.

1000 Stavrakou, T., Müller, J. F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M.,
1001 Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P. F., and Guenther, A.: How consistent
1002 are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI?,
1003 Atmos. Chem. Phys., 15, 11861-11884, doi: 10.5194/acp-15-11861-2015, 2015.

1004 Stavrakou, T., Muller, J. F., Bauwens, M., De Smedt, I., Lerot, C., Van Roozendael, M., Coheur, P. F.,
1005 Clerbaux, C., Boersma, K. F., van der, A. R., and Song, Y.: Substantial Underestimation of
1006 Post-Harvest Burning Emissions in the North China Plain Revealed by Multi-Species Space
1007 Observations, Sci. Rep., 6, 32307, doi: 10.1038/srep32307, 2016.

1008 Sun, L., Xue, L. K., Wang, T., Gao, J., Ding, A. J., Cooper, O. R., Lin, M. Y., Xu, P. J., Wang, Z., Wang,
1009 X. F., Wen, L., Zhu, Y. H., Chen, T. S., Yang, L. X., Wang, Y., Chen, J. M., and Wang, W. X.:
1010 Significant increase of summertime ozone at Mount Tai in Central Eastern China, Atmos. Chem. Phys.,
1011 16, 10637-10650, doi: 10.5194/acp-16-10637-2016, 2016.

1012 van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D.
1013 C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of
1014 deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10,
1015 11707-11735, doi:10.5194/acp-10-11707-2010, 2010.

1016 van Donkelaar, A., Martin, R. V., Leaitch, W. R., Macdonald, A. M., Walker, T. W., Streets, D. G.,
1017 Zhang, Q., Dunlea, E. J., Jimenez, J. L., Dibb, J. E., Huey, L. G., Weber, R., and Andreae, M. O.:
1018 Analysis of aircraft and satellite measurements from the Intercontinental Chemical Transport
1019 Experiment (INTEX-B) to quantify long-range transport of East Asian sulfur to Canada, Atmos. Chem.
1020 Phys., 8, 2999-3014, doi: 10.5194/acp-8-2999-2008, 2008.

1021 Vestreng, V.: Review and revision. Emission data reported to CLRTAP, Tech. rep., EMEP MSC-W,
1022 (available at: http://www.emep.int/mscw/mscw_publications.html#2003), 2003.

1023 Volkamer, R.: Primary and Secondary Glyoxal Formation from Aromatics: Experimental Evidence for
1024 the Bicycloalkyl-Radical Pathway from Benzene, Toluene, and p-Xylene, J. Phys. Chem., 105, 7865,
1025 doi:10.1021/jp010152w, 2001.

1026 Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated
1027 VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem.
1028 Phys., 10, 10145-10160, doi: 10.5194/acp-10-10145-2010, 2010.

1029 Wang, F., An, J., Li, Y., Tang, Y., Lin, J., Qu, Y., Chen, Y., Zhang, B., and Zhai, J.: Impacts of
1030 uncertainty in AVOC emissions on the summer ROx budget and ozone production rate in the three most
1031 rapidly-developing economic growth regions of China, Adv. Atmos. Sci., 31, 1331-1342, doi:
1032 10.1007/s00376-014-3251-z, 2014.

1033 Wang, H. Q., Ma, J. M., Shen, Y. J., and Wang, Y. A.: Assessment of Ozone Variations and
1034 Meteorological Influences at a Rural Site in Northern Xinjiang, Bull. Environ. Contam. Toxicol., 94,
1035 240-246, doi: 10.1007/s00128-014-1451-y, 2015.

1036 Wang, P., Stammes, P., R., v. d. A., Pinardi, G., and Roozendael, M. V.: FRESCO+: an improved O2
1037 A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8,
1038 6565-6576, doi: 10.5194/acp-8-6565-2008, 2008.

1039 Wang, Y., Konopka, P., Liu, Y., Chen, H., Muller, R., Ploger, F., Riese, M., Cai, Z., and Lu, D.:
1040 Tropospheric ozone trend over Beijing from 2002-2010: ozonesonde measurements and modeling
1041 analysis, Atmos. Chem. Phys., 12, 8389-8399, doi: 10.5194/acp-12-8389-2012, 2012.

1042 Wang, Y., Beirle, S., Lampel, J., Koukouli, M., De Smedt, I., Theys, N., Li, A., Wu, D. X., Xie, P. H.,
1043 Liu, C., Van Roozendael, M., Stavrakou, T., Muller, J. F., and Wagner, T.: Validation of OMI,
1044 GOME-2A and GOME-2B tropospheric NO2, SO2 and HCHO products using MAX-DOAS
1045 observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and
1046 aerosols on the satellite products, Atmos. Chem. Phys., 17, 5007-5033, doi: 10.5194/acp-17-5007-2017,
1047 2017.

1048 Wei, W., Lv, Z. F., Yang, G., Cheng, S. Y., Li, Y., and Wang, L. T.: VOCs emission rate estimate for
1049 complicated industrial area source using an inverse-dispersion calculation method: A case study on a
1050 petroleum refinery in Northern China, Environ. Pollut.,doi: 10.1016/j.envpol.2016.07.062, 218,
1051 681-688, 2016.

1052 Wei, W., Wang, S., Chatani, S., Klimont, Z., Cofala, J., and Hao, J.: Emission and speciation of
1053 non-methane volatile organic compounds from anthropogenic sources in China, Atmos. Environ., 42,
1054 4976-4988, doi:10.1016/j.atmosenv.2008.02.044, 2008.

1055 Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja,
1056 A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the
1057 emissions from open burning, Geosci. Model Dev., 4, 625-641, doi: 10.5194/gmd-4-625-2011, 2011.

1058 Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R.,
1059 Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde
1060 from space, Geophys. Res. Lett., 33, doi: 10.1029/2006gl026310, 2006.

1061 Wu, R., Bo, Y., Li, J., Li, L., Li, Y., and Xie, S.: Method to establish the emission inventory of
1062 anthropogenic volatile organic compounds in China and its application in the period 2008–2012, Atmos.
1063 Environ., 127, 244-254, doi:10.1016/j.atmosenv.2015.12.015, 2016.

1064 Xu, W. Y., Lin, W. L., Xu, X. B., Tang, J., Huang, J. Q., Wu, H., and Zhang, X. C.: Long-term trends of
1065 surface ozone and its influencing factors at the Mt Waliguan GAW station, China - Part 1: Overall
1066 trends and characteristics, Atmos. Chem. Phys., 16, 6191-6205, doi: 10.5194/acp-16-6191-2016, 2016.

1067 Xu, X., Lin, W., Wang, T., Yan, P., Tang, J., Meng, Z., and Wang, Y.: Long-term trend of surface ozone
1068 at a regional background station in eastern China 1991-2006: enhanced variability, Atmos. Chem. Phys.,
1069 8, 2595-2607, doi: 10.5194/acp-8-2595-2008, 2008.

1070 Zhang, J. M., Wang, T., Ding, A. J., Zhou, X. H., Xue, L. K., Poon, C. N., Wu, W. S., Gao, J., Zuo, H.
1071 C., Chen, J. M., Zhang, X. C., and Fan, S. J.: Continuous measurement of peroxyacetyl nitrate (PAN) in
1072 suburban and remote areas of western China, Atmos. Environ., 43, 228-237, doi:
1073 10.1016/j.atmosenv.2008.09.070, 2009.

1074 Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S.,
1075 Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006
1076 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131-5153, 10.5194/acp-9-5131-2009, 2009.

1077 Zhao, Y., Nielsen, C. P., Lei, Y., McElroy, M. B., and Hao, J.: Quantifying the uncertainties of a
1078 bottom-up emission inventory of anthropogenic atmospheric pollutants in China, Atmos. Chem. Phys.,
1079 11, 2295-2308, doi:10.5194/acp-11-2295-2011, 2011.

1080 Zheng, J. Y., Zhong, L. J., Wang, T., Louie, P. K. K., and Li, Z. C.: Ground-level ozone in the Pearl
1081 River Delta region: Analysis of data from a recently established regional air quality monitoring
1082 network, Atmos. Environ., 44, 814-823, doi: 10.1016/j.atmosenv.2009.11.032, 2010.

1083 Zhu, C. Y., Byrd, R. H., Lu, P. H., and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for
1084 large-scale bound-constrained optimization, ACM T. Math. Software, 23, 550-560, doi:
1085 10.1145/279232.279236, 1997.

1086 Zhu, L., Jacob, D. J., Mickley, L. J., Marais, E. A., Cohan, D. S., Yoshida, Y., Duncan, B. N., González
1087 Abad, G., and Chance, K. V.: Anthropogenic emissions of highly reactive volatile organic compounds
1088 in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns,
1089 Environ. Res. Lett., 9, 114004, doi: 10.1088/1748-9326/9/11/114004, 2014.

1090 Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M.,
1091 Sulprizio, M. P., De Smedt, I., González Abad, G., Chance, K., Li, C., Ferrare, R., Fried, A., Hair, J. W.,
1092 Hanisco, T. F., Richter, D., Jo Scarino, A., Walega, J., Weibring, P., and Wolfe, G. M.: Observing
1093 atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from
1094 four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC^{4}RS aircraft observations over the

southeast US, Atmos. Chem. Phys., 16, 13477-13490, doi: 10.5194/acp-16-13477-2016, 2016.

Table 1. Comparison of recent estimates for Chinese annual NMVOC emissions

Literature [a]	Target year	NMVOC [Tg C y^{-1}]				
		Anthropogenic		Biogenic		Biomass burning
		Total	Aromatics	Total	Isoprene	
Bottom-up						
Bo et al. (2008) [b]	2005	10.7				2.2 [e]
Zhang et al. (2009) [b]	2006	19.7 (±68%)	2.1			
Cao et al. (2011) [b]	2007	29.8				
Kurokawa et al. (2013) [b]	2008	22.8 (±46%)				
Li et al. (2014) [b]	2010	19.8	4.9			
Wu et al. (2016) [b]	2008	15.6				2.2 [e]
	2009	18.3				1.9 [e]
	2010	20.0				2.1 [e]
	2011	20.8				2.1 [e]
	2012	21.5				2.4 [e]
Huang et al. (2012) [b]	2006					1.3 (0.62-2.0)
van der Werf et al.(2010)	2007					0.24
Guenther et al.(2006)	2007			10.8 [f]	6.6 [f]	
Top-down						
Liu et al. (2012) [c]	2007	28.7	11.3			
Stavrakou et al. (2015) [d]	2010	(17.3-20.7)			(5.0-5.5)	(1.1-1.5)
This work	2007	17.0 [g] (13.5-19.7)	6.1 [g] (5.0~7.3)	12.6 [g] (8.9-14.8)	8.4 [g] (4.9-10.5)	1.22 [g] (1.06 – 1.47)

[a] Emission estimates from literature were originally in units of Tg y^{-1}. We converted the units to Tg C y^{-1} using to carbon to organic compound mass ratios (0.84 for anthropogenic VOCs, 0.57 for biomass burning VOCs, and 0.85 for biogenic VOCs based on the *a priori* emission estimates).

[b] These emission estimates included NMVOC species that were not included in this work. See color keys in Figure 2 for NMVOC species whose emissions were included in this work.

[c] Used SCIAMACHY-observed glyoxal VCDs as constraints.

[d] Used GOME-2A-observed and OMI-observed formaldehyde VCDs as constraints.

[e] Consisted of emissions from open burning of crop residues and from biofuel burning.

[f] Calculated by the GEOS-Chem model using GEOS-5 meteorological data.

[g] Average of top-down estimates from four inversion experiments.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

Table 2. Inversion experiments to constrain Chinese NMVOC emissions

Inversion experiments	Observational constraints from satellites [±uncertainties]	Annual Chinese NMVOC emission estimates [Tg C y^{-1}]			
		Anthropogenic	Biogenic	Biomass burning	Total
		A priori emission estimates [±uncertainties]			
		15.5 (4.9 for aromatics) [a] [±200%]	10.8 (6.6 for isoprene) [b] [±55%]	1.10 [±300%] [c]	27.4
		A posteriori emission estimates [range]			
IE-1	GOME-2A formaldehyde [±90%] and glyoxal [±150%]	15.7 (5.9 for aromatics)	12.5 (8.2 for isoprene)	1.13	29.3
IE-2	OMI formaldehyde [±90%] and glyoxal [±150%]	13.5 (5.0 for aromatics)	8.9 (4.9 for isoprene)	1.06	23.4
IE-3	GOME-2A formaldehyde × 170% [±90%]	19.2 (6.0 for aromatics)	14.8 (10.5 for isoprene)	1.47	35.4
IE-4	OMI glyoxal [±150%]	19.7 (7.3 for aromatics)	14.1 (9.9 for isoprene)	1.24	35.1
Our top-down estimates		17.0 [d] [13.5 - 19.7] (6.1 [d] [5.0 - 7.3] for aromatics)	12.6 [d] [8.9 - 14.8] (8.4 [d] [4.9 – 10.5] for isoprene)	1.2 [d] [1.1 - 1.5]	30.8 [d] [23.4 – 35.4]

[a] From Li et al. (2014)

[b] From Guenther et al. (2006).

[c] Compiled from the emission estimated by van der Werf et al. (2010) plus a scaling of the emission estimated by Huang et al. (2012). See text (section 2.2) for details.

[d] Average of top-down estimates from the four inversion experiments.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

Table 3. Technical details for GOME-2A and OMI formaldehyde and glyoxal retrievals used in this study

Technical details		GOME-2A		OMI	
		Formaldehyde [a]	Glyoxal [b]	Formaldehyde [c]	Glyoxal [d]
Onboard satellite		European Metop-A		NASA Aura	
Operation time		October 2006-present		July 2004-present	
Overpass time		9:30 LT		13:30 LT	
Global coverage		1.5 days [e]		1 day	
Spatial resolution		80 km × 40 km		13 km × 24 km	
Spectral window		240-790 nm		270-500 nm	
Spectral resolution		0.26-0.5 nm		0.42 nm and 0.63 nm	
Selected absorption band		328.5 - 346 nm	435 - 460 nm	328.5 - 356.5 nm	435 - 461 nm
Retrieval algorithm		DOAS fitting		Direct fitting	
Cloud parameters		FRESCO+ (Wang et al., 2008)		OMCLDO2 (Acarreta et al., 2004)	
Surface albedo		Kleipool et al. (2008)		Kleipool et al. (2008)	
Air mass factor calculation	Radiative transfer model	LIDORT (Spurr, 2008)		VLIDORT (Spurr, 2006)	
	Tracer gas profiles	IMAGE v2 (Stavrakou et al., 2009b)		GEOS-Chem (González Abad et al., 2015)	
Extinction by aerosols		Considered implicitly via cloud correction		Considered implicitly in the cloud retrieval	
Discarded pixels		Cloud fraction > 40% or zenith angles >60°		Cloud fraction > 40%	Impacted by random telegraph signals (RTS) [f]

[a] From De Smedt et al. (2012)
[b] From Lerot et al. (2010)
[c] From González Abad et al. (2015)
[d] From Chan Miller et al. (2014)
[e] Before the swath was narrowed in June 2013. After that, the global coverage is achieved every 3 days.
[f] Pixels that have been flagged as RTS in the level 1-B product (Kleipool, 2005).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

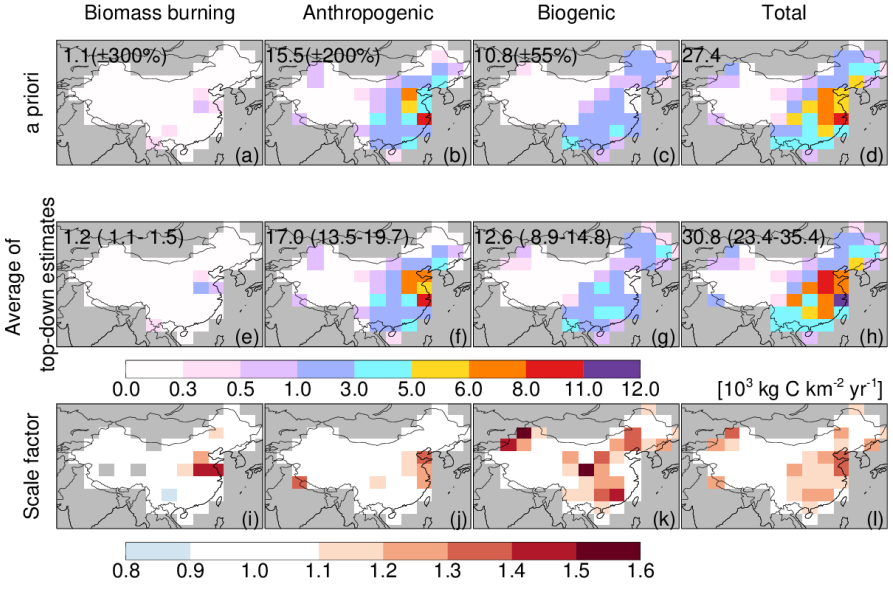

Figure 1. Spatial distributions of annual NMVOC emissions from China. (a)-(d): the *a priori* annual NMVOC emission estimates from (a) biomass burning, (b) anthropogenic, (c) biogenic, and (d) total sources. (e)-(h): averaged top-down estimates of annual NMVOC emissions. Annual Chinese total emission estimates are shown inset in units of [Tg C y^{-1}]. The uncertainties of the *a priori* emission estimates and the range of top-down emission estimates are shown in parentheses. (i)-(l): scale factors for our averaged top-down estimates relative to the *a priori* estimates.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

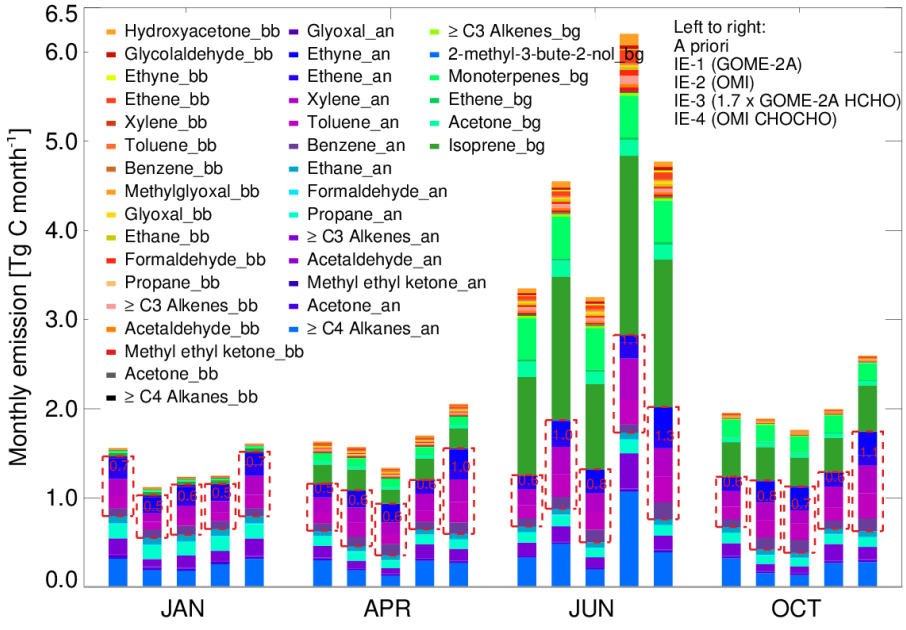

Figure 2. Estimates of monthly Chinese NMVOC emissions for January, April, June, and October 2007. For each month, the bars from left to right represent: the *a priori* emission estimates and the *a posteriori* emission estimates from IE-1, IE-2, IE-3, and IE-4. The red dashed boxes and red numbers indicate monthly emissions of anthropogenic glyoxal precursors. Color keys for NMVOC species are shown inset, with the suffixes of 'bb', 'an' and 'bg' indicating emissions from biomass burning, anthropogenic, and biogenic activities, respectively.

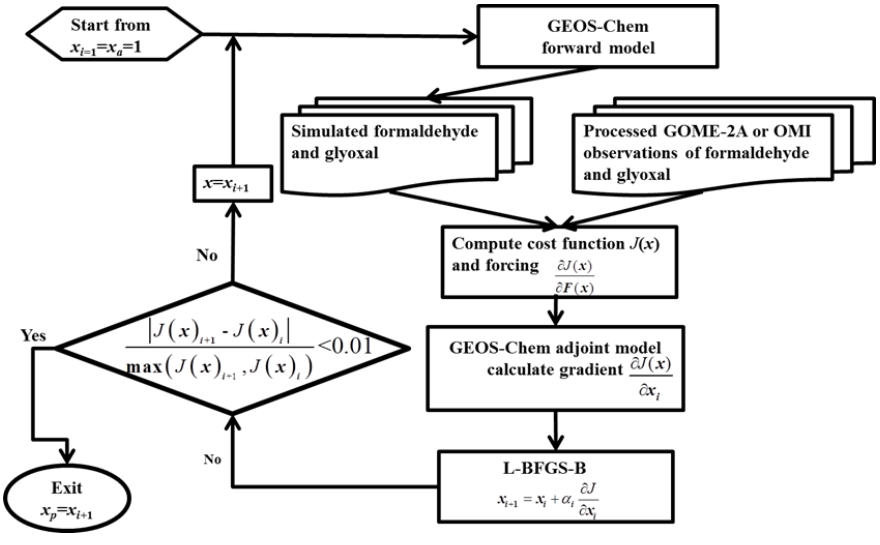

Figure 3. Protocol for the adjoint inversion experiments.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

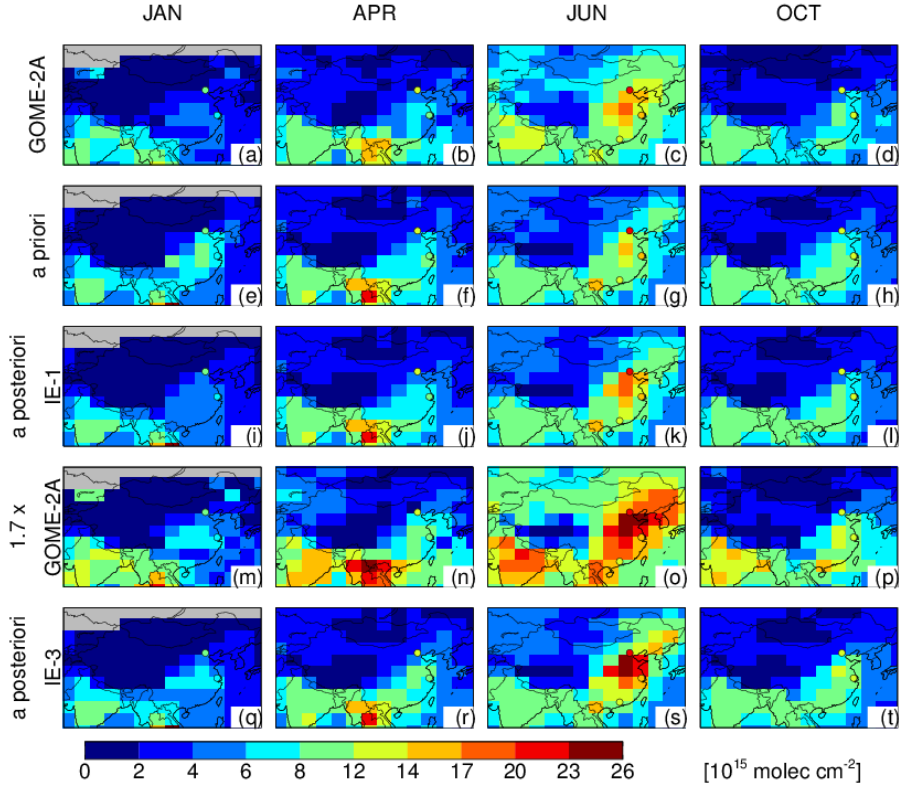

Figure 4. Monthly mean formaldehyde VCDs over China. (a-d): GOME-2A-observed formaldehyde VCDs and (m-p) GOME-2A formaldehyde VCDs scaled by a factor of 1.7. (e-h): Formaldehyde VCDs simulated by the model using *a priori* emission estimates; (i-l) the *a posteriori* formaldehyde VCDs from inversion IE-1;(q-t) the *a posteriori* formaldehyde VCDs from the inversion IE-3. Also shown are ground-based MAX-DOAS measurements at 9:30 LT (circles) at Beijing (De Smedt et al., 2015), Wuxi (Wang et al, 2017), and Back Garden (Li et al, 2013).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

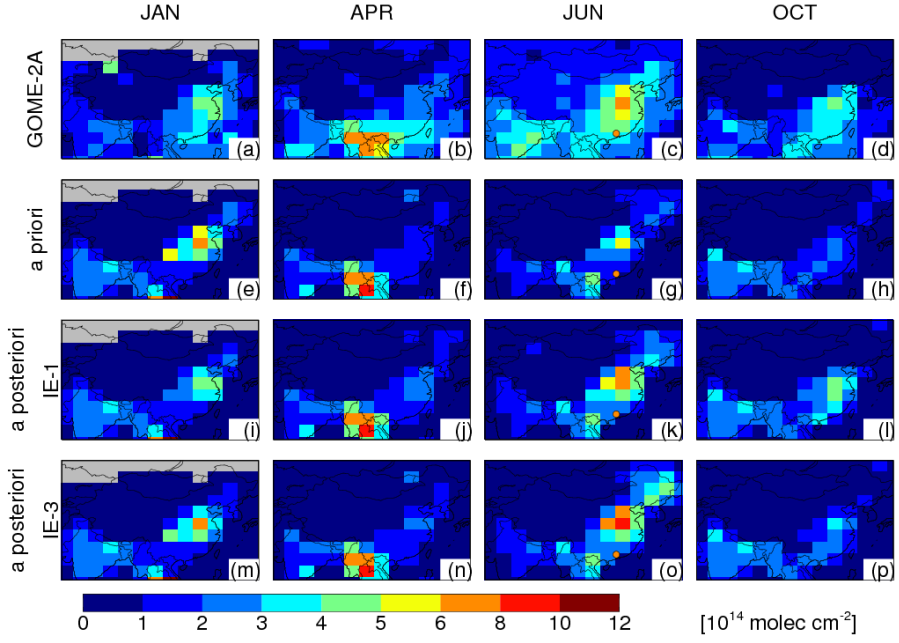

Figure 5. Monthly mean glyoxal VCDs over China (a-d) observed by the GOME-2A instrument, (e-h) simulated by the model using the *a priori* emission estimates, (i-l) obtain from inversion IE-1, and (m-p) obtain from inversion IE-3. Also shown are ground-based MAX-DOAS measurements at Back Garden in July 2006 (Li et al, 2013).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

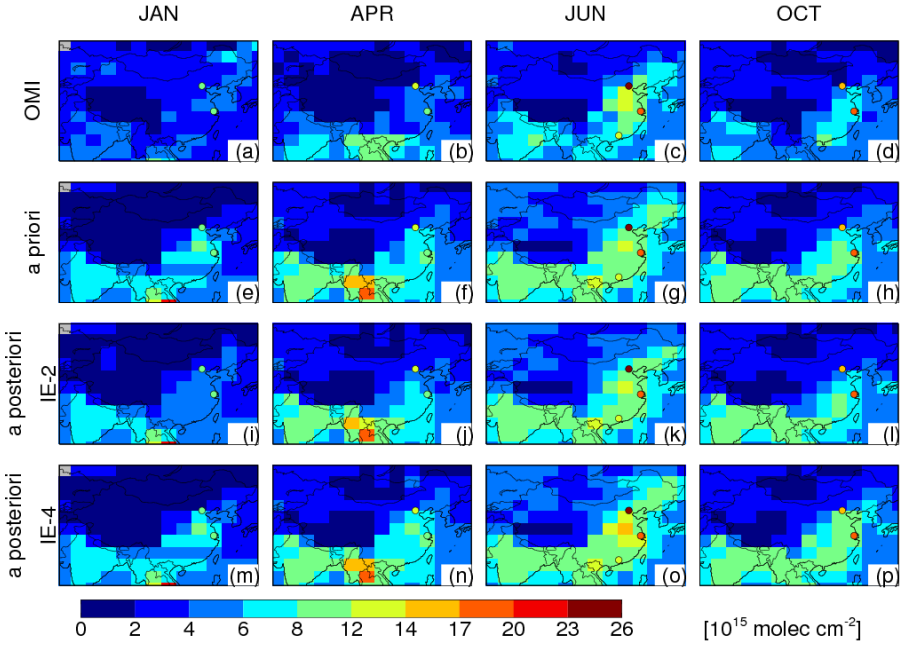

Figure 6. Monthly mean formaldehyde VCDs over China. (a-d): formaldehyde VCDs observed by the OMI instrument. (e-h): formaldehyde VCDs simulated by the model using the *a priori* emission estimates. (i-l): the *a posteriori* formaldehyde VCDs from inversion IE-2. (m-p): the *a posteriori* formaldehyde VCDs from inversion IE-4. Also shown are ground-based MAX-DOAS measurements at 13:30 LT (circles) at Beijing (De Smedt et al., 2015), Wuxi (Wang et al, 2017), and Back Garden (Li et al, 2013).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

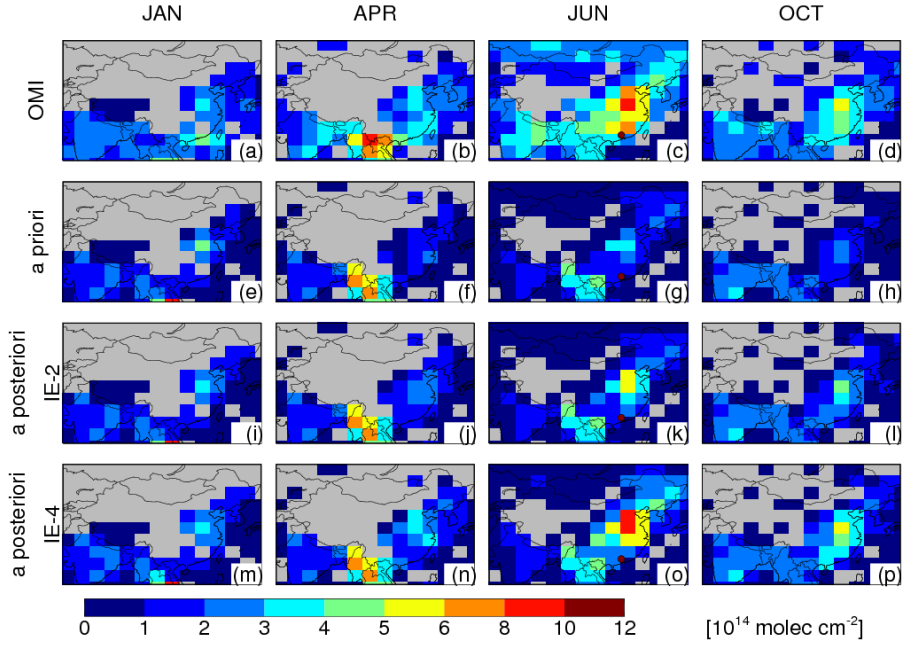

Figure 7. Monthly mean glyoxal VCDs over China (a-d) observed by the OMI instrument, (e-h) simulated by the model using the *a priori* emission estimates, (i-l) obtained from inversion IE-2, and (m-p) obtained from the inversion IE-4. Also shown are ground-based MAX-DOAS measurements at 13:30 LT (circles) at Back Garden in July 2006 (Li et al, 2013).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

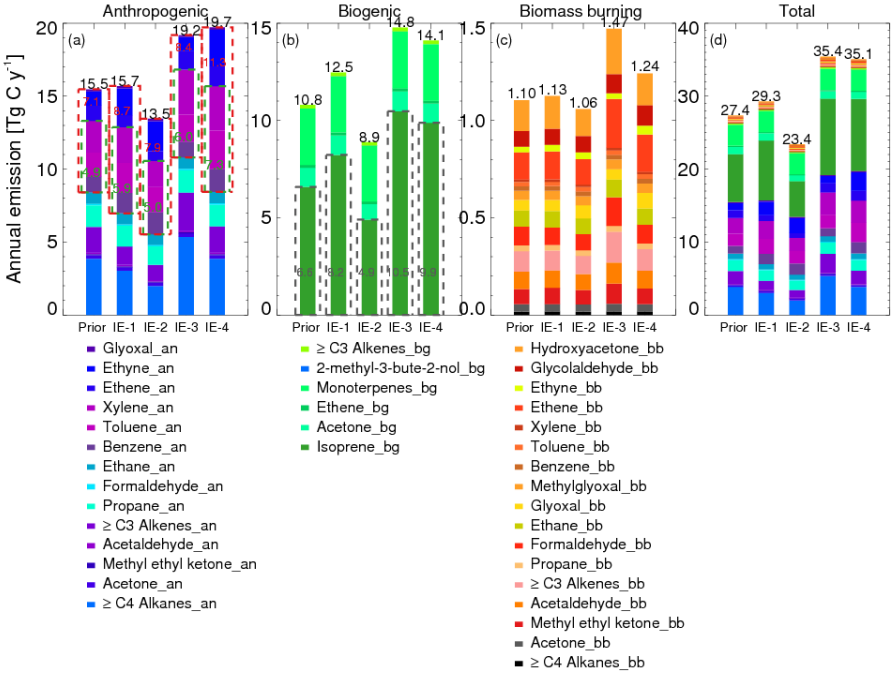

Figure 8. Comparison of estimates of annual Chinese NMVOC emissions from (a) anthropogenic, (b) biogenic, (c) biomass burning, and (d) total sources. For each subfigure, shown from left to right are the *a priori* estimates and our *a posteriori* estimates from IE-1, IE-2, IE-3, and IE-4. Annual total NMVOC emission estimates are shown in black numbers on top of each bar. The red dashed boxes and red numbers in (a) indicate annual emissions of anthropogenic glyoxal precursors. The green dashed boxes and green numbers in (a) indicate annual emissions of anthropogenic aromatics. The grey dashed boxes and grey numbers in (b) indicate annual biogenic isoprene emissions. Color keys to NMVOC species are shown at the bottom, with suffixes of 'an', 'bg', 'bb' indicating anthropogenic source, biogenic source, and biomass burning source, respectively.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1136
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 15 January 2018

Atmospheric Chemistry and Physics Discussions

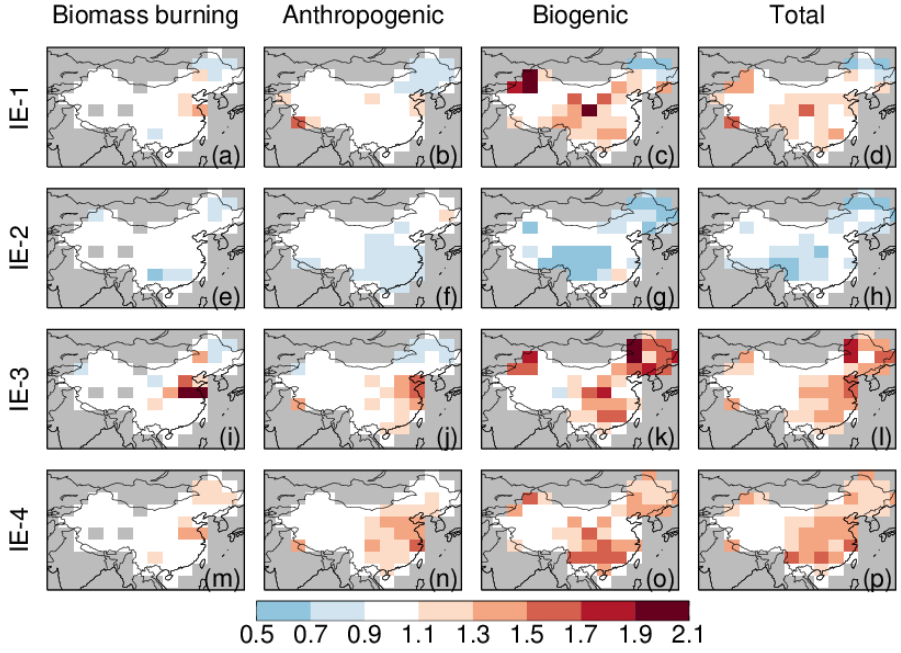

Figure 9. Spatial distributions of the optimized scale factors for Chinese annual NMVOC emissions, relative to the *a priori* emission estimates, for the four inversion experiments.

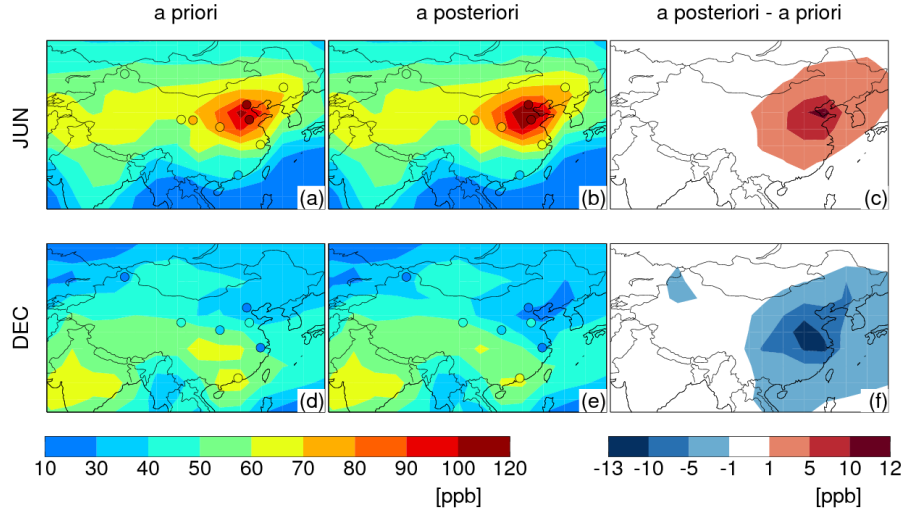

Figure 10. Simulated monthly mean afternoon (13:00-17:00 LT) surface ozone concentrations driven by the *a priori* emissions and average of our top-down emissions, respectively, as well as corresponding difference (*a posteriori–a priori*) in June and December 2007. Filled circles overlaid on the contour maps represent surface ozone observations at several sites of China (Table S2).