Top-down estimates of benzene and toluene emissions in Pearl River Delta and Hong Kong, China

X. Fang¹, M. Shao², A. Stohl³, Q. Zhang⁴, J. Zheng⁵, H. Guo⁶, C. Wang²,⁷, M. Wang², J. Ou⁸, R. L. Thompson³, and R. G. Prinn¹

¹Center for Global Change Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
²State Key Joint Laboratory of Environmental Simulation and Pollution Control, College of Environmental Sciences and Engineering, Peking University, Beijing, China
³Norwegian Institute for Air Research, Kjeller, Norway
⁴Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing, China
⁵College of Environment and Energy, South China University of Technology, University Town, Guangzhou, China
⁶Air Quality Studies, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, China
⁷College of Environmental Engineering and Science, Qilu University Of Technology, Jinan, Shandong, China
Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China

Received: 19 August 2015 – Accepted: 24 August 2015 – Published: 11 September 2015
Correspondence to: X. Fang (fangxk@mit.edu)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Benzene (C₆H₆) and toluene (C₇H₈) are toxic to humans and the environment. They are also important precursors of ground-level ozone and secondary organic aerosols and contribute substantially to severe air pollution in urban areas in China. Discrepancies exist between different bottom-up inventories for benzene and toluene emissions in Pearl River Delta (PRD) and Hong Kong (HK), which are emission hot spots in China. This study provides top-down estimates of benzene and toluene emissions in PRD and HK using atmospheric measurement data from a rural site in the area, Heshan, an atmospheric transport model and an inverse modeling method. The model simulations captured the measured mixing ratios during most pollution episodes. For PRD and HK, the benzene emissions estimated in this study for 2010 were 44 (12–75) Gg yr⁻¹ and 5 (2–7) Gg yr⁻¹ for PRD and HK, respectively, and the toluene emissions were 131 (44–218) Gg yr⁻¹ and 6 (2–9) Gg yr⁻¹, respectively. Temporal and spatial differences between the inversion estimate and four different bottom-up emission estimates are discussed, and it is proposed that more observations at different sites are urgently needed to better constrain benzene and toluene (and other air pollutants) emissions in PRD and HK in the future.

1 Introduction

Benzene and toluene, two volatile organic compounds (VOCs), are toxic to humans and the environment. For example, a sufficiently high exposure of toluene will lead to health issues like intra-uterine growth retardation, premature delivery, congenital malformations, and postnatal developmental retardation (Donald et al., 1991). VOCs, including benzene and toluene, are also important precursors of ground-level ozone that is produced from the reaction between VOCs and NOₓ in the presence of sunlight (Xue et al., 2014) and contribute to the formation of secondary organic aerosols (Henze et al., 2008). VOCs emitted from anthropogenic activities are crucial contributors to the
severe urban haze pollution in China (Guo et al., 2014). Therefore, information about the spatial and temporal distribution of benzene and toluene emissions is crucial for air quality simulations and predictions, health risk assessments and emission control policy.

The Pearl River Delta (PRD) and Hong Kong (HK) are located along the coast of southern China, which is one of the most economically developed areas in the country. Guangzhou, Shenzhen in PRD, and Hong Kong are densely populated megacities. The PRD and HK regions experience severe air pollution, namely toxic trace gases and particulates, as observed by satellites (e.g. van Donkelaar et al., 2010) and ground-based measurements (e.g. Guo et al., 2009). Toluene was the most abundant VOC and benzene was one of the most the most abundant VOCs in PRD (Chan et al., 2006). Toluene and benzene, respectively, had the largest and second largest emissions of all anthropogenic VOCs in PRD in 2010 (Ou et al., 2015), which highlights the importance of accurately quantifying these emissions. In PRD, the two major sources of benzene are industrial processes and road transport, and those of toluene are industrial solvents and road transport, while minor sources include industrial processes, gasoline evaporation, building paints, biomass burning, etc. (Ou et al., 2015).

Although some bottom-up inventories exist for benzene and toluene emissions in PRD, there are discrepancies among them. For example for benzene emissions in 2010, the Regional Emission inventory in Asia (REAS) v1.1 reference scenario (from here on referred to as REAS REF v1.1) estimates the emissions to be 8 Gg yr\(^{-1}\) (Ohara et al., 2007), while the Multi-resolution Emission Inventory (MEIC) v1.2 (available at http://www.meicmodel.org) estimate is 33 Gg yr\(^{-1}\), the Representative Concentration Pathways Scenario 3PD (RCP 3PD) estimate is 45 Gg yr\(^{-1}\) (van Vuuren et al., 2007), and the Yin et al. (2015) estimate is 54 Gg yr\(^{-1}\). Thus, estimates of the total emissions vary by a factor of approximately seven. For toluene emissions in 2010, the estimates are also quite different: the RCP 3PD and REAS v1.1 REF estimates are 44 and 46 Gg yr\(^{-1}\), respectively, the Yin et al. (2015) estimate is 64 Gg yr\(^{-1}\), and the MEIC v1.2 estimate is 181 Gg yr\(^{-1}\). Atmospheric-measurement-based estimates are needed
to validate benzene and toluene emissions estimated from bottom-up methods. However, to date no top-down estimate is available for PRD and HK.

High-frequency online measurements of VOCs (including benzene and toluene) were made during the PRIDE-PRD2010 Campaign (Program of Regional Integrated Experiments on Air Quality over Pearl River Delta) during November and December 2010. This study uses these measurement data and an inverse modeling approach to infer benzene and toluene emissions in PRD and HK. This top-down estimate is important to test and improve the existing bottom-up inventories.

2 Methodology

2.1 Measurement data

In this study, atmospheric measurements of benzene and toluene at two sites were used, the Heshan site (used for the inversion) and the Mt. Tai Mo Shan (TMS) site (used for validation). The Heshan site (112.929° E, 22.728° N) is a rural observatory located on the top of a small hill (~ 60 m above the surrounding terrain) in Jiangmen (see Fig. 1). The measurement period at Heshan site was from 11 November 2010 to 1 December 2010. Data from 1 December 2010 were not used, since we focused on mixing ratios and emissions in November. Detailed information of the measurement system and procedure can be found in Wang et al. (2014). Here we provide only a brief description. Ambient mixing ratios of VOCs were measured using an online automatic gas chromatograph system equipped with a mass spectrometer and a flame ionization detector (GC-MS/FID). Most C₂–C₅ hydrocarbons were measured by the FID Channel with a porous layer open tubular (PLOT) column, whereas other VOCs, including benzene and toluene, were measured by the MSD Channel with a DB-624 column. The time resolution of the VOC measurements was 60 min. The detection limits of this system for benzene and toluene are 0.006 and 0.015 ppb, respectively, which are much
lower values than the typical benzene and toluene mixing ratio levels of 2 and 6 ppb during the observation period at the Heshan site.

The Mt. TMS site (114.118° E, 22.405° N) was not used for the inversion but for validating the emissions derived from the inversions in this study. The sample air inlet at the TMS site was located on the rooftop of a building at Mt. TMS at an elevation of 640 m a.s.l. A total of 75 canisters of air samples were taken over different times of day and night on 1–3, 9, and 19–21, November 2010. Detailed information on the sampling time schedule can be found in Guo et al. (2013). After sampling, the VOC canister samples were sent to a laboratory at the University of California, Irvine for chemical analysis. Simpson et al. (2010) provide a full description of the analytical system, which uses a multi-column gas chromatograph (GC) with five column-detector combinations. The measurement detection limit of this system for both benzene and toluene is 0.003 ppb, which is much lower than the typical mixing ratio levels of 0.7 ppb for benzene and 1.6 ppb for toluene during the observation period at the Mt. TMS site.

### 2.2 Model simulations using FLEXPART

The source–receptor relationships (SRRs, often also called “emission sensitivities”) were calculated using the backwards in time mode of the Lagrangian particle dispersion model, FLEXPART (http://www.flextart.eu) (Stohl et al., 2005, 1998). The model was driven by hourly meteorological data of 0.5° × 0.5° horizontal resolution and 37 vertical levels from the NCEP Climate Forecast System Reanalysis (CFSR) (available at http://rda.ucar.edu/datasets/ds093.0/). During 3 hourly intervals throughout the sampling period, 80,000 virtual particles were released at the site’s location and at the height of the sampling inlet above model ground level, and followed backwards in time for 20 days. The spatial resolution of the output from the backward simulations is 0.25° × 0.25°. Loss of benzene and toluene by reaction with the hydroxyl (OH) radical in the atmosphere was considered in the backward simulation. Rate constant values...
for reaction with OH radicals were expressed for benzene as:

\[ k = 2.308 \times 10^{-12} \times \exp\left(-\frac{190}{T}\right); \]  

expressed for toluene as:

\[ k = 1.275 \times 10^{-18} \times T^2 \times \exp\left(\frac{1192}{T}\right), \]

where \( T \) is the ambient temperature (K). Gridded OH fields (hourly for the period October to December 2010, at a resolution of 0.5° \( \times \) 0.667°, 47 vertical levels) were derived from the atmospheric chemistry transport model, GEOS-Chem v5 (http://acmg.seas.harvard.edu/geos/). Compared to the reference simulation ran backwards for 20 days with atmospheric chemical loss, Sect. 3.2 shows alternative FLEXAPRT simulations that were run backwards for 10 and 40 days with atmospheric chemical loss, and 20 days without atmospheric chemical loss.

### 2.3 Inverse algorithm

The Bayesian inversion method used in this study is almost the same as described and evaluated by Stohl et al. (2009, 2010), and as used in recent studies of \( \text{SF}_6 \) emissions (Fang et al., 2014) and HFC-23 emissions (Fang et al., 2015) in East Asia. Briefly, in this study a Bayesian inversion technique is employed, based on least-squares optimization, to estimate both the spatial distribution and strength of the emissions in the domain to which the measurements are sensitive. The inversion adjusts the emissions to minimize the differences between the observed and modeled mixing ratios while also considering the deviation of the optimized emissions from an a priori emission field. Observation–model mismatch errors (including transport model errors) are determined as the root mean square error (RMSE) of the model–observation mismatch (Stohl et al., 2009, 2010). In this study, mixing ratio background values were set to 24845.
zero. This is because the backward simulations were run for 20 days and benzene and toluene in the air parcel from emissions occurring prior to this time would have been largely removed from the atmosphere by reaction with OH (typical atmospheric lifetimes of benzene and toluene are \( \sim 10 \) and \( \sim 2 \) days, respectively). Simulated benzene and toluene mixing ratios at the measurement site were obtained by integrating the gridded emission sensitivities multiplied by the gridded emissions.

For benzene, gridded a priori emission fields for mainland China were derived from MEIC v1.2 for November 2010 (\( 0.25^\circ \times 0.25^\circ \), monthly mean), and for the rest of the world the emissions were taken from RCP Scenario 3PD (\( 0.5^\circ \times 0.5^\circ \), yearly mean) (van Vuuren et al., 2007). For toluene, a priori emission fields for mainland China were derived from MEIC v1.2 for November 2010 (\( 0.25^\circ \times 0.25^\circ \), monthly mean), while for the PRD region in mainland China, a priori emissions were derived from averaging the estimates by MEIC v1.2 (\( 0.25^\circ \times 0.25^\circ \), monthly mean) and by Yin et al. (2015) (\( 0.25^\circ \times 0.25^\circ \), monthly mean) for November 2010. For the rest of the world, a priori emissions were taken from RCP 3PD inventory (\( 0.5^\circ \times 0.5^\circ \), yearly mean) (van Vuuren et al., 2007). Both monthly inventories of MEIC v1.2 and Yin et al. (2015) were obtained through personal communication with the dataset authors. A priori emission uncertainty in each grid box for benzene and toluene, respectively, was set to 100 and 70% according to the differences among the bottom-up inventories and was assumed uncorrelated in space.

3 Results and discussion

3.1 Benzene and toluene ambient mixing ratios

Table 1 shows ambient mixing ratios of benzene and toluene measured at the Heshan site and other sites all over the world. Mixing ratios of benzene at the Heshan site ranged from 0.59 to 20.23 ppb and had an average of \( 2.27 \pm 1.65 \) (mean \( \pm \) standard deviation) ppb during our observation period. Mixing ratios of toluene at the Heshan
site ranged from 0.87 to 25.05 ppb and had an average of 5.6 ± 4.15 ppb. The mixing ratios of benzene (0.67 ± 0.21 ppb) and toluene (1.58 ± 1.25 ppb) at the Mt. TMS were only ~ 30% of those at the Heshan site. In agreement with previous studies (e.g. Lau et al., 2010; Liu et al., 2008), mixing ratio levels of benzene and toluene in PRD region are overall higher than those in Hong Kong (Table 1), which is most likely due to the fact that Hong Kong often receives clean air masses from over the ocean.

Mixing ratios of benzene and toluene in some cities in Europe (e.g. Ait-Helal et al., 2014; Langford et al., 2010) and United States (e.g. USEPA, 1989; Baker et al., 2008) were about 0.5 and 1 ppb (Table 1), respectively, which is about 20% of the mean observed values in PRD in this study. Mixing ratios of benzene and toluene in Thompson Farm, United State were even 0.08 ± 0.002 and 0.09 ± 0.005 ppb, respectively, which are much lower than the lowest mixing ratios at both Heshan and Mt. TMS sites. Levels of benzene and toluene mixing ratios at different sites mainly reflect the combined influence of emission strength, seasonal changes in atmospheric OH concentration and mixing depth.

3.2 Benzene and toluene emission sensitivities

Figure 2 shows the spatial distribution of average emission sensitivity of benzene and toluene for the Heshan site for 12–30 November 2010. During the observation period, air masses transported to the Heshan site mainly came from easterly and northerly directions. Considering that the major emission sources in PRD are located to the east of the Heshan site (Fig. 1), this measurement location was ideally situated for constraining emissions from this region for this period and, as the emission sensitivities show, PRD, HK, and neighboring regions, are relatively well constrained by the observations at the Heshan site. Benzene and toluene emissions in PRD and HK are much higher than emissions in neighboring regions Fig. 1 and, consequently, the overall mixing ratio contributions (the integral of the emission sensitivities multiplied by emissions) from PRD and HK to the observation site comprise more than 80% of the total simulated mixing ratios. Note that the emission sensitivities for benzene and toluene are different
because there are differences in the chemical loss of these two compounds during atmospheric transport and in the molecular weight. Specifically, the emission sensitivities for toluene are spatially more confined because of its shorter lifetime.

As a sensitivity study, alternative simulations in which FLEXPART was run backward for 10 days were made. The derived emission sensitivities are almost identical to the reference simulations with 20 days duration (Supplement Fig. S1 for benzene and Fig. S2 for toluene), confirming that 20 day-backward simulations are sufficiently long to account for all benzene and toluene emission sources that can influence the mixing ratios at the Heshan site. Since the lifetime of benzene is ~10 days (much longer than that of toluene), we also made a 40 day-backward simulation from which the emission sensitivities for benzene are also almost identical to the reference simulation (Fig. S3). Without accounting for the loss by reaction with OH in the atmosphere, the emission sensitivities for benzene would only be a little higher (by ~10% in central PRD) (Fig. S4). On the other hand, the emission sensitivities for toluene would be much higher (by ~50% in central PRD) (Fig. S5). This indicates that accounting for chemical loss has a relatively small effect for simulating benzene mixing ratios at Heshan, whereas it has a profound effect on toluene mixing ratios. Notice that systematic errors in the simulated chemical loss would lead to smaller errors (i.e., compared to the extreme case of no loss) in the a posteriori emissions retrieved by the inversion. Thus, errors in the retrieved emissions due to errors in chemical loss are marginal for benzene but could be significant for toluene.

### 3.3 Inversion results

Figure 3 shows the observed and simulated mixing ratios at the Heshan site. The simulations captured most pollution episodes, and the inversion improved the agreement between the simulations and the observations, as expected. For benzene, root mean square errors between the observed and simulated mixing ratios decreased from 1.53 ppb, using a priori emissions, to 1.26 ppb, using a posteriori emissions, and the mean bias between the simulated mixing ratios and observations decreased...
from 0.96 ppb, using a priori emissions, to 0.41 ppb, using a posteriori emissions. For
toluene, the root mean square error between the observed and simulated mixing ratios
decreased from 4.77 ppb, using a priori emissions, to 4.30 ppb, using a posteriori emis-
sions and the mean bias between the observed and simulated mixing ratios decreased
from 2.35 ppb, using a priori emissions, to 1.99 ppb, using a posteriori emissions.

Figure 3 also shows examples of spatial distributions of toluene emission sensitivities
for two observed mixing ratios. The toluene mixing ratio at 00:00 UTC on 16 Novem-
ber 2010 was about 2 ppb and the corresponding air mass had not passed over the
strong emission sources in the central part of PRD and HK (see the backward emis-
sion sensitivities map in Fig. 3c), while the toluene mixing ratio at 00:00 UTC on 24
November 2010 was about 15 ppb and the corresponding air mass had passed over
the strong emission sources in the central part of PRD and HK (Fig. 3d).

Figure 4 shows the benzene a priori and a posteriori emission fields, their differences
and uncertainty reduction. The a priori fields show that emission hot spots are located
in Guangzhou, Shenzhen and Hong Kong megacities. Emission changes by the in-
version are positive in some grid cells and negative in some other grid cells, which
shows that the a priori emissions are not systematically lower or higher everywhere
than the a posteriori emissions. The biggest emission changes by the inversion occur
in two boxes in Guangzhou where the a priori emissions were enhanced by ~ 50 % in
one box and decreased by more than 50 % in the other box. The emission hot spot in
Shenzhen did not change much. To test the sensitivity to the a priori emission in this
grid cell, we performed an additional inversion in which the a priori emission in this grid
box was reduced, and a high a posteriori emission in this grid cell was still found, as in
the reference inversion.

Figure 5 shows the a priori and a posteriori emissions of toluene and their differ-
ence. Emission hot spots are located in Guangzhou and Shenzhen. The uncertainty
reduction map in Figs. 4d and 5d show significant error reductions, of 40 % or more,
in boxes close to the observation site, while only low emission uncertainty reductions
were achieved in boxes far from the observation site. Overall, the emission uncertain-
ties have been reduced by the inversion in PRD and HK, where the strongest emission sources are located.

The total a posteriori benzene emissions for PRD and HK, respectively, are 4.0 (1.1–6.9) Gg month\(^{-1}\) (for all cases the range represents one sigma uncertainty) and 0.4 (0.1–0.7) Gg month\(^{-1}\). A posteriori toluene emissions are 12 (4–20) Gg month\(^{-1}\) for PRD and 0.5 (0.2–0.9) Gg month\(^{-1}\) for HK. Inversion sensitivity tests, i.e., using other bottom-up emission inventories for the a priori estimate (listed in Table 2), all produce toluene emission estimates that fall within the uncertainty range of the a posteriori emissions from the reference inversion.

Benzene and toluene measurement data at the Mt. TMS site were not used in the inversion but for validating the posterior emissions. For benzene, using the a priori and a posteriori emission fields, respectively, the root mean square errors between the simulated and observed mixing ratios at Mt. TMS site are 0.367 and 0.312 ppb, and the mean bias between the simulated and observed mixing ratios are 0.314 and 0.208 ppb. For toluene, the root mean square error (1.50 ppb) between the observed and simulated mixing ratios using the a posteriori emission fields from the inversion was smaller than that (1.55 ppb) using the a priori field; the mean bias (1.06 ppb) between the observations and simulated mixing ratios using a posteriori emission fields was also smaller than that (1.12 ppb) using the a priori field. Both the root mean square error and mean bias suggest that the a posteriori emissions are more accurate than the a priori emissions.

### 3.4 Comparison with other estimates

Figures 6 and S6, respectively, show spatial distributions of benzene and toluene emissions estimated by the inversion in this study, four bottom-up inventories, and the differences among these estimates. For benzene, the spatial emission distributions in the REAS v1.1 REF have the biggest difference from our top-down emissions. Gridded emissions in the REAS v1.1 REF are always lower than the inversion emissions, while emissions in the Yin et al. (2015), MEIC v1.2 and RCP 3PD estimates are less
systematically biased. The simulated benzene mixing ratios using the REAS v1.1 inventory are much lower than the observed mixing ratios (Fig. 7). Statistics of the root mean square error, mean bias and squared Pearson correlation coefficients between the simulated and observed mixing ratios show that emission fields obtained from the inversion performed better in simulating the benzene mixing ratios than all four bottom-up inventories (See Table S1).

For toluene, in most grid cells over the PRD, emissions estimated by RCP 3PD, Yin et al. (2015) and REAS v1.1 REF are lower than the inversion estimates, while MEIC v1.2 emissions are higher than the inversion estimates (Fig. S6). Model simulations show that the simulated mixing ratios using emission estimates from RCP 3PD, Yin et al. (2015) and REAS v1.1 REF are much lower than the observed mixing ratios at the Heshan site (Fig. S7). The simulated mixing ratios using MEIC v1.2 emission fields are not consistent with some observed pollution peaks (Fig. S7). Statistics of root mean square error and squared Pearson correlation coefficients show that inversion emission fields performed better at simulating toluene mixing ratios at the Heshan site than the four bottom-up emission fields (see Table S1).

Table 2 shows five estimates of total benzene and toluene emissions in PRD and HK regions for the year 2010. The a posteriori emissions for November 2010 obtained from the inversion were extrapolated to an annual mean emission rate for the whole year 2010 by multiplying the November emissions by the ratio of emissions for the whole year 2010 to those in November 2010. For toluene, this ratio is 10.8, and was calculated from both the MEIC v1.2 and Yin et al. (2015) estimate (the November/annual emission ratio was the same in both datasets). For toluene, the factor is 10.9 (10.4–11.4), and is the average of 10.4, calculated from the MEIC v1.2 estimate, and 11.4, calculated from the Yin et al. (2015) estimate. Data in November 2010 and the whole year 2010 were obtained through personal communication with the dataset authors. Using these ratios, the benzene emissions in PRD and HK for 2010 were estimated to be 44 (12–75) and 5 (2–7) Ggyr⁻¹, respectively, and the toluene emissions were estimated to be 131 (44–218) and 6 (2–9) Ggyr⁻¹, respectively.
For benzene, emissions in PRD in 2010 calculated from the four bottom-up estimates were 45 Ggyr$^{-1}$ from RCP 3PD (van Vuuren et al., 2007), 54 Ggyr$^{-1}$ from Yin et al. (2015), 8 Ggyr$^{-1}$ from REAS v1.1 REF (Ohara et al., 2007), and 33 Ggyr$^{-1}$ from MEIC v1.2. Our inverse estimate agrees within its uncertainties with these bottom-up estimates, except that the REAS estimate is substantially lower than the other bottom-up and the top-down estimates. Emissions in HK were 5 (2–7) Ggyr$^{-1}$ estimated by this study, which agrees within uncertainties with the RCP 3PD estimate and is much higher than the REAS v1.1 REF (no estimates are available from MEIC v1.2 or Yin et al., 2015).

For toluene, emissions in PRD in 2010 calculated from the four bottom-up estimates were 44 Ggyr$^{-1}$ from RCP 3PD estimate (van Vuuren et al., 2007), 64 Ggyr$^{-1}$ from Yin et al. (2015), 46 Ggyr$^{-1}$ from REAS v1.1 REF (Ohara et al., 2007), and 181 Ggyr$^{-1}$ from MEIC v1.2. The bottom-up estimate MEIC v1.2 meets the high uncertainty range of our inversion estimates, while the other three bottom-up estimates meet the low uncertainty range of our inversion estimates. For the HK toluene emissions, estimates are not available in MEIC v1.2 or Yin et al. (2015); both RCP 3PD and REAS v1.1 REF estimates are about 4 Ggyr$^{-1}$, which agree with our inversion results within uncertainties.

### 3.5 Benzene and toluene emissions during 2000–2010

Figure 8 shows different estimates of benzene and toluene emissions in the PRD region for the period 2000–2010. For benzene, the estimate of 8 Ggyr$^{-1}$ in 2000 by REAS v2.1 (Kurokawa et al., 2013) agrees with that of 13 Ggyr$^{-1}$ by the Reanalysis of the Tropospheric chemical composition over the past 40 years project (RETO) (Schultz et al., 2007), which are substantially smaller than that of 43 Ggyr$^{-1}$ in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (Lamarque et al., 2010). For the years 2005 and 2006, different studies show substantial differences. For the year 2005, the emission estimate by RCP 3PD was $\sim$ 4 times the estimates by REAS v2.1. For the year 2006, the emission estimate by REAS v2.1 agrees with the estimate in the Intercontinental Chemical Transport Experiment-Phase
B (INTEX-B) project (Zhang et al., 2009), which were only \( \sim 1/5 \) of the estimate by Zheng et al. (2009). More studies are available for the year 2010 than for other years. For the year 2010, the estimates by RCP 3PD, MEIC v1.2 and Yin et al. (2015) agree with the inversion estimate by this study, which are higher than the estimate by REAS v1.1 REF. According to these estimates (Fig. 8), it is likely that the benzene emissions in PRD have not changed much during the 2000–2010 period, although emissions are uncertain due to limited number of estimates.

For toluene, the estimate of 45 Ggyr\(^{-1}\) in 2000 by REAS v2.1 (Kurokawa et al., 2013) agrees relatively well with the value of 36 Ggyr\(^{-1}\) by ACCMIP (Lamarque et al., 2010), but both are substantially larger than the RETRO estimate of 14 Ggyr\(^{-1}\) (Schultz et al., 2007). For the years 2005 and 2006, estimates of toluene emission are also quite different. For the year 2005, the emission estimate by REAS v2.1 was \( \sim 4 \) times the estimates by RCP 3PD. For the year 2006, the emission estimate by REAS v2.1 was \( \sim 2 \) times the estimate by Zheng et al. (2009) and even \( \sim 11 \) times the estimate by INTEX-B (Zhang et al., 2009). For the year 2010, the estimates by REAS v1.1 REF and Yin et al. (2015) meet the low end of uncertainty of inversion estimate by this study, while MEIC v1.2 estimate meets the high end. According to these estimates over 2000–2010 (Fig. 8), it is likely that the toluene emissions in PRD have increased during this period, although emissions are uncertain due to limited number of estimates.

Based on glyoxal (CHOCHO) data retrieved from satellite and inversion method, Liu et al. (2012) found their emission estimates of the lumped artificial compound ARO1 (benzene, toluene and ethylbenzene) in PRD in 2006 were \( > 10 \) times larger than the bottom-up INTEX-B estimates (also for 2006), but they did not specify which compound was responsible for the difference. As for benzene, the ratio of emissions in 2006 estimated by Zheng et al. (2009) (60 Ggyr\(^{-1}\)) to the INTEX-B estimate (15 Ggyr\(^{-1}\)) is \( \sim 4 \) times, much less than the factor \( > 10 \) discrepancy reported by Liu et al. (2012). Inversion estimate of benzene emissions (44 (12–75) Ggyr\(^{-1}\)) in 2010 is \( \sim 3 \) (1–5) times the INTEX-B emissions for 2006. Thus, we suggest that the big discrepancy is likely not due to emissions of benzene but emissions of toluene and/or ethybenzene. As
for toluene emissions, the ratios of bottom-up estimates by REAS v2.1 (190 Ggyr\(^{-1}\)) and Zheng et al. (2009) (103 Ggyr\(^{-1}\)) for 2006 to the INTEX-B bottom-up estimate (18 Ggyr\(^{-1}\)) are 11–6 times. Thus, considering the satellite-based estimate and other bottom-up estimates, the bottom-up INTEX-B estimate of toluene emissions for the PRD region for 2006 was likely too low, and estimation of toluene emissions in PRD is attributed as an important factor contributing to the big discrepancy of ARO1 emission estimates between Liu et al. (2012) and INTEX-B.

### 3.6 Suggestions for more top-down studies

To the best of our knowledge, this study provides the only available top-down estimate for toluene emissions in the PRD and HK regions. All other studies in Fig. 8 are bottom-up estimates. More top-down estimates are needed to validate the bottom-up estimates in the previous years and in the future. In this study, inversions using the Heshan measurement data reduced emission uncertainties in PRD and HK regions. However, the emission uncertainty reductions were not large because there was only one observation site suitable for the inversion (some measurements in urban environments are available but not suitable for inverse modeling) and the observation period was not long. Thus, we propose that in the future, observations with better spatial and temporal coverage are urgently needed to better constrain benzene and toluene (and other VOC) emissions in PRD and HK regions. Inversion-oriented observation sites could be situated in rural places outside of the major emission sources located in the central part of PRD and HK regions, and then the major emission sources in PRD and HK regions could be “viewed” from different angles to better constrain the benzene and toluene (and other VOC) emissions.
4 Conclusions

Using atmospheric measurements at the Heshan site, a transport model and an inversion algorithm, this study provides the first top-down estimate of benzene and toluene emissions in the Pearl River Delta (PRD) and Hong Kong (HK) regions, which are emission hot spots in China. According to the measurement data in this study and previous studies, mixing ratio levels of benzene and toluene in PRD region are overall higher than those in Hong Kong, which are much higher than those measured in the United States and Europe. Considering that air masses transported to the Heshan site mainly came from easterly and northerly directions during the observation period and that the major emissions sources in PRD are located to the east of the Heshan site, the Heshan measurement site was ideally situated for constraining emissions from these regions. Based on the measurement data, model simulations and inverse technique, the PRD and HK benzene emissions for 2010 estimated in this study were 44 (12–75) and 5 (2–7) Ggyr\(^{-1}\), respectively and the PRD and HK toluene emissions for 2010 were 131 (44–218) and 6 (2–9) Ggyr\(^{-1}\), respectively. We have discussed the spatial distributions of benzene and toluene emissions obtained by inversion in this study in the context of four different existing bottom-up inventories. The discrepancies among these bottom-up estimates for the period 2000–2010 are substantial (up to a factor of seven), while this study is the only one available top-down estimate. We propose that in the future, observations with better spatial and temporal coverage are urgently needed to constrain benzene and toluene (and other VOC) emissions in PRD and HK regions more strongly.

The Supplement related to this article is available online at doi:10.5194/acpd-15-24839-2015-supplement.

Acknowledgements. This study was partly funded by the Natural Science Foundation for Outstanding Young Scholars (Grant No. 41125018) and Key project (Grant No. 41330635). This
work was supported in part by a National Aeronautics and Space Administration (NASA) grant NNX11AF17G awarded to the Massachusetts Institute of Technology (MIT). We acknowledge Emissions of atmospheric Compounds and Compilation of Ancillary Data (http://eccad.sedoo.fr/eccad_extract_interface/JSF/page_criteres.jsf) for the archiving of the emission inventory data of ACCMIP, REAS v1.1 REF, RCP 3PD and RETRO. We thank Mingwei Li at Department of Earth, Atmospheric and Planetary Sciences, MIT for her help in GEOS-Chem model runs.

References


Chan, L.-Y., Chu, K.-W., Zou, S.-C., Chan, C.-Y., Wang, X.-M., Barletta, B., Blake, D. R., Guo, H., and Tsai, W.-Y.: Characteristics of nonmethane hydrocarbons (NMHCs) in industrial, industrial-urban, and industrial-suburban atmospheres of the Pearl River Delta (PRD)


C$_2$-C$_{10}$ volatile organic compounds (VOCs), CO$_2$, CH$_4$, CO, NO, NO$_2$, NO$_y$, O$_3$ and SO$_2$, Atmos. Chem. Phys., 10, 11931–11954, doi:10.5194/acp-10-11931-2010, 2010.


Velasco, E., Lamb, B., Westberg, H., Allwine, E., Sosa, G., Arriaga-Colina, J. L., Jobson, B. T.,
Alexander, M. L., Prazeller, P., Knighton, W. B., Rogers, T. M., Grutter, M., Herndon, S. C.,
 magnitudes, reactivities, ratios and diurnal patterns of volatile organic compounds in the
Valley of Mexico during the MCMA 2002 and 2003 field campaigns, Atmos. Chem. Phys., 7,

Zhang, Z.: Development and validation of a cryogen-free automatic gas chromatograph sy-
stem (GC-MS/FID) for online measurements of volatile organic compounds, Anal. Methods,

Wingenter, O. W., Mao, H., Talbot, R., and Sive, B. C.: Are biogenic emissions a significant
source of summertime atmospheric toluene in the rural Northeastern United States?, Atmos.

cities: precursors, regional transport and heterogeneous processes, Atmos. Chem. Phys.,

VOC emission inventory and its improvement on modeling regional ozone in the Pearl River

Yoshino, A., Nakashima, Y., Miyazaki, K., Kato, S., Suthawaree, J., Shimo, N., Matsunaga, S.,
Chatani, S., Apel, E., Greenberg, J., Guenther, A., Ueno, H., Sasaki, H., Hoshi, J.-Y.,
Yokota, H., Ishii, K., and Kajii, Y.: Air quality diagnosis from comprehensive observations
of total OH reactivity and reactive trace species in urban central Tokyo, Atmos. Environ., 49,

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z.,
emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131–5153,

Emission Inventory and Spatial Patterns of Ozone Formation Potential in the Pearl River
Table 1. Ambient mixing ratios (ppb) of benzene and toluene measured at the Heshan site and other sites all over the world (SD represents Standard Deviation; NG represents Not Given).

<table>
<thead>
<tr>
<th>Location</th>
<th>Type</th>
<th>Time</th>
<th>Sample number</th>
<th>Benzene Sample Range</th>
<th>Mean ± SD</th>
<th>Toluene Sample Range</th>
<th>Mean ± SD</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) PRD and Hong Kong regions, China</td>
<td>Rural</td>
<td>11–30 Nov 2010</td>
<td>419</td>
<td>0.59–20.23</td>
<td>2.27 ± 1.65</td>
<td>419</td>
<td>0.87–25.05</td>
<td>5.65 ± 4.15</td>
</tr>
<tr>
<td>Heshan, PRD</td>
<td>Rural</td>
<td>04 Oct to 03 Nov, 2004</td>
<td>111</td>
<td>0.66–11.35</td>
<td>2.39 ± 1.99</td>
<td>111</td>
<td>0.76–36.91</td>
<td>7.01 ± 7.33</td>
</tr>
<tr>
<td>Xikeng, PRD</td>
<td>Rural</td>
<td>04 Oct to 03 Nov, 2004</td>
<td>83</td>
<td>0.52–6.26</td>
<td>1.42 ± 0.98</td>
<td>83</td>
<td>0.54–56.41</td>
<td>8.46 ± 9.94</td>
</tr>
<tr>
<td>Dongguan, PRD</td>
<td>Urban</td>
<td>Sep 2005</td>
<td>48</td>
<td>0.27–6.45</td>
<td>1.26 ± 0.14</td>
<td>48</td>
<td>0.53–25.30</td>
<td>6.13 ± 0.81</td>
</tr>
<tr>
<td>Guangzhou, PRD</td>
<td>Urban</td>
<td>Sep 2006</td>
<td>42</td>
<td>0.65–6.80</td>
<td>2.05 ± 1.49</td>
<td>42</td>
<td>0.72–19.60</td>
<td>5.87 ± 4.11</td>
</tr>
<tr>
<td>Guangzhou, PRD</td>
<td>Industrial</td>
<td>Late summer 2000</td>
<td>15</td>
<td>2.80 ± 1.70</td>
<td>9.20 ± 4.90</td>
<td>15</td>
<td>13.5 ± 11.8</td>
<td></td>
</tr>
<tr>
<td>Area, PRD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mt. Tai Mo Shan, Hong Kong</td>
<td>Mountain</td>
<td>01–03, 09, 19–21 Nov 2010</td>
<td>75</td>
<td>0.38–1.79</td>
<td>0.67 ± 0.21</td>
<td>75</td>
<td>0.26–6.30</td>
<td>1.58 ± 1.25</td>
</tr>
<tr>
<td>Tap Mun, Hong Kong</td>
<td>Rural</td>
<td>Nov 2006 to Oct 2007</td>
<td>39</td>
<td>0.05–1.67</td>
<td>0.56 ± 0.41</td>
<td>39</td>
<td>0.15–7.12</td>
<td>1.61 ± 1.55</td>
</tr>
<tr>
<td>Central West, Hong Kong</td>
<td>Urban</td>
<td>Nov 2006 to Oct 2007</td>
<td>40</td>
<td>0.05–1.91</td>
<td>0.60 ± 0.50</td>
<td>40</td>
<td>0.28–8.81</td>
<td>2.64 ± 2.07</td>
</tr>
<tr>
<td>(2) Other sites in China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43 cities, China</td>
<td>Urban</td>
<td>Jan–Feb 2001</td>
<td>158</td>
<td>0.7–10.4</td>
<td>NG</td>
<td>158</td>
<td>0.4–11.2</td>
<td>NG</td>
</tr>
<tr>
<td>Beijing, China</td>
<td>Urban</td>
<td>Aug 2005</td>
<td>1046</td>
<td>3.03 ± 1.72</td>
<td>NG</td>
<td>1039</td>
<td>1.76 ± 0.89</td>
<td>NG</td>
</tr>
<tr>
<td>Shanghai, China</td>
<td>Urban</td>
<td>15 Jun 2006 to 14 Jun 2007</td>
<td>~ 365</td>
<td>6.07 ± 11.70</td>
<td>NG</td>
<td>~ 365</td>
<td>32.80 ± 21.60</td>
<td>NG</td>
</tr>
<tr>
<td>(3) Sites in other countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karachi, Pakistan</td>
<td>Urban</td>
<td>Winter of 1998–1999</td>
<td>78</td>
<td>0.34–19.3</td>
<td>5.20 ± 4.50</td>
<td>78</td>
<td>0.19–37.0</td>
<td>7.10 ± 7.60</td>
</tr>
<tr>
<td>Tokyo, Japan</td>
<td>Urban</td>
<td>Summer 2007</td>
<td>50</td>
<td>0.78 ± 0.61</td>
<td>NG</td>
<td>50</td>
<td>2.14 ± 0.99</td>
<td>NG</td>
</tr>
<tr>
<td>Tokyo, Japan</td>
<td>Urban</td>
<td>Winter 2007</td>
<td>16</td>
<td>0.82 ± 0.28</td>
<td>NG</td>
<td>16</td>
<td>10.10 ± 5.29</td>
<td>NG</td>
</tr>
<tr>
<td>London, UK</td>
<td>Urban</td>
<td>Oct 2010</td>
<td>601</td>
<td>0.15 ± 0.11</td>
<td>NG</td>
<td>589</td>
<td>0.68 ± 0.57</td>
<td>NG</td>
</tr>
<tr>
<td>Paris, France</td>
<td>Suburban</td>
<td>15 Jan–15 Feb 2010</td>
<td>246</td>
<td>0.32 ± 0.16</td>
<td>NG</td>
<td>246</td>
<td>0.32 ± 0.22</td>
<td>NG</td>
</tr>
<tr>
<td>Mexico City, Mexico</td>
<td>Urban</td>
<td>Feb 2002 and Apr–May 2003</td>
<td>~ 115</td>
<td>3.17 ± 1.75</td>
<td>NG</td>
<td>~ 86</td>
<td>13.5 ± 9.33</td>
<td>NG</td>
</tr>
<tr>
<td>Mexico City, Mexico</td>
<td>Rural</td>
<td>Feb 2002 and Apr–May 2003</td>
<td>~ 115</td>
<td>0.80 ± 0.91</td>
<td>NG</td>
<td>~ 86</td>
<td>1.89 ± 1.92</td>
<td>NG</td>
</tr>
<tr>
<td>39 cities, USA</td>
<td>Urban</td>
<td>Jun–Sep 1984–1986</td>
<td>835</td>
<td>0.001–0.27</td>
<td>NG</td>
<td>836</td>
<td>0.003–1.30</td>
<td>NG</td>
</tr>
<tr>
<td>28 cities, USA</td>
<td>Urban</td>
<td>Summer 1999–2005</td>
<td>530</td>
<td>(0.06 ± 0.024)</td>
<td>NG</td>
<td>530</td>
<td>(0.12 ± 0.05)</td>
<td>NG</td>
</tr>
<tr>
<td>Thompson Farm, USA</td>
<td>Rural</td>
<td>Fall 2004–2006</td>
<td>201</td>
<td>0.08 ± 0.002</td>
<td>NG</td>
<td>201</td>
<td>0.09 ± 0.005</td>
<td>White et al. (2009)</td>
</tr>
</tbody>
</table>

\* It represents the range of the minimal mean value (the corresponding standard deviation) in one of 28 cities and maximal mean value (the corresponding standard deviation) in another city.
Table 2. Benzene and toluene emissions (Ggyr$^{-1}$) in the PRD and HK regions derived from different estimates for the year 2010.

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Benzene emissions</th>
<th>Toluene emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRD</td>
<td>HK</td>
</tr>
<tr>
<td>RCP 3PD</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>Yin et al. (2015)</td>
<td>54</td>
<td>NE</td>
</tr>
<tr>
<td>REAS v1.1 REF</td>
<td>8</td>
<td>0.4</td>
</tr>
<tr>
<td>MEIC v1.2 REF</td>
<td>33</td>
<td>NE</td>
</tr>
<tr>
<td>This study</td>
<td>44 (12–75)</td>
<td>5 (2–7)</td>
</tr>
</tbody>
</table>

NE represents “Not Estimated”.
Figure 1. Map of (a) benzene and (b) toluene emissions from the MEIC v1.2 for China and the RCP 3PD for outside China (inset panels), and that for the PRD and Hong Kong regions (mother panel). The PRD region is plotted with dark blue boundary lines, the Hong Kong region with cyan boundary lines. The Green cross represents the location of the Heshan observation site.
Figure 2. Average emission sensitivities of (a) benzene and (b) toluene for the Heshan observation site for 12–30 November 2010. The green cross indicates the location of Heshan site. The blue and cyan lines represent PRD and Hong Kong boundary lines, respectively.
Figure 3. Observed and simulated (a) benzene and (b) toluene mixing ratios at the Heshan site, and two examples of spatial distributions of toluene emission sensitivities at (c) 00:00 UTC on 16 November 2010 and (d) 00:00 UTC on 24 November 2010.
Figure 4. Maps of (a) a priori benzene emissions, (b) a posteriori benzene emissions, (c) differences between (b) and (a), and (d) uncertainty reduction. The observation site is marked with a green cross.
Figure 5. Maps of (a) a priori toluene emissions, (b) a posteriori toluene emissions, (c) differences between (b) and (a), and (d) uncertainty reduction. The Heshan observation site is marked with a green cross.
Figure 6. Maps of benzene emissions for PRD, HK and surrounding regions from (a) inversion, (b) RCP 3PD, (c) Yin et al. (2015), (d) REAS v1.1 REF, (e) MEIC v1.2, and the difference between inversion results (a) and the bottom-up inventories (b, c, d, e). Note that in (c) and (g) only emissions within PRD are plotted since Yin et al. (2015) only estimated emissions within PRD, and that in (e) and (i) emissions within HK are not plotted since MEIC v1.2 has not estimated benzene emission in HK.
Figure 7. Time series of observed and simulated benzene mixing ratios at the Heshan site. The simulations use emission fields from inversion in this study, RCP 3PD, REAS v1.1 REF, Yin et al. (2015) and MEIC v1.2, respectively.
Figure 8. Estimates of benzene and toluene emissions in the PRD region for the period 2000–2010.