

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

3 X. Fang¹, M. Shao², A. Stohl³, Q. Zhang⁴, J. Zheng⁵, H. Guo⁶, C. Wang^{2,7}, M. Wang²,
4 J. Ou⁸, R. L. Thompson³, R. G. Prinn¹

5 ¹Center for Global Change Science, Massachusetts Institute of Technology, Cambridge,
6 Massachusetts, USA

7 ²State Key Joint Laboratory of Environmental Simulation and Pollution Control, College of
8 Environmental Sciences and Engineering, Peking University, Beijing, China

9 ³Norwegian Institute for Air Research, Kjeller, Norway

10 ⁴Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System
11 Science, Tsinghua University, Beijing, China

12 ⁵College of Environment and Energy, South China University of Technology, University
13 Town, Guangzhou, China

14 ⁶Air Quality Studies, Department of Civil and Environmental Engineering, The Hong Kong
15 Polytechnic University, Hong Kong, China

16 ⁷College of Environmental Engineering and Science, Qilu University Of Technology, Jinan,
17 Shandong, China

18 ⁸Institute of Space and Earth Information Science, The Chinese University of Hong Kong,
19 Hong Kong, China

20 Correspondence to: X. Fang (fangxk@mit.edu); M. Shao (mshao@pku.edu.cn)

21 **Abstract**

22 Benzene (C₆H₆) and toluene (C₇H₈) are toxic to humans and the environment. They
23 are also important precursors of ground-level ozone and secondary organic aerosols
24 and contribute substantially to severe air pollution in urban areas in China.
25 Discrepancies exist between different bottom-up inventories for benzene and toluene
26 emissions in the Pearl River Delta (PRD) and Hong Kong (HK), which are emission
27 hot spots in China. This study provides top-down estimates of benzene and toluene
28 emissions in the PRD and HK using atmospheric measurement data from a rural site
29 in the area, Heshan, an atmospheric transport model, and an inverse modeling method.
30 The model simulations captured the measured mixing ratios during most pollution
31 episodes. For the PRD and HK, the benzene emissions estimated in this study for
32 2010 were 44 (12–75) Gg yr⁻¹ and 5 (2–7) Gg yr⁻¹ for the PRD and HK, respectively,
33 and the toluene emissions were 131 (44–218) Gg yr⁻¹ and 6 (2–9) Gg yr⁻¹,
34 respectively. Temporal and spatial differences between the inversion estimate and four
35 different bottom-up emission estimates are discussed, and it is proposed that more
36 observations at different sites are urgently needed to better constrain benzene and
37 toluene (and other air pollutant) emissions in the PRD and HK in the future.

38 **1 Introduction**

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

51 The Pearl River Delta (PRD) and Hong Kong (HK) are located along the coast of
52 southern China, which is one of the most economically developed areas in the country.
53 It is also where the densely populated mega-cities, Guangzhou and Shenzhen (in the
54 PRD) and Hong Kong are located. The PRD and HK regions experience severe air
55 pollution, namely toxic trace gases and particulates, as observed by satellites (e.g. van
56 Donkelaar et al., 2010) and ground-based measurements (e.g. Guo et al., 2009).
57 Toluene and benzene were found to be two of the most abundant VOCs in the PRD
58 (Chan et al., 2006). Toluene and benzene, respectively, had the largest and second

59 largest emissions of all anthropogenic VOCs in the PRD in 2010 (Ou et al., 2015),
60 which highlights the importance of accurately quantifying these emissions. In the
61 PRD, the two major sources of benzene are industrial processes and road transport,
62 and those of toluene are industrial solvents and road transport, while minor sources
63 for both benzene and toluene include stationary combustion, gasoline evaporation,
64 biomass burning, etc. (Ou et al., 2015).

65 Although some bottom-up inventories exist for benzene and toluene emissions in
66 the PRD, there are discrepancies among them. For example, for benzene emissions in
67 2010, the Regional Emission inventory in Asia (REAS) v1.1 reference scenario (from
68 here on referred to as REAS REF v1.1) estimates the emissions to be 8 Gg yr⁻¹ (Ohara
69 et al., 2007), while the Multi-resolution Emission Inventory (MEIC) v1.2 (available at
70 <http://www.meicmodel.org>) estimate is 33 Gg yr⁻¹, the Representative Concentration
71 Pathways Scenario 2.6 (RCP 2.6) estimate is 45 Gg yr⁻¹ (van Vuuren et al., 2007), and
72 the Yin et al. (2015) estimate is 54 Gg yr⁻¹. Thus, estimates of the total emissions vary
73 by a factor of approximately seven. For toluene emissions in 2010, the estimates are
74 also quite different: The RCP 2.6 and REAS v1.1 REF estimates are 44 Gg yr⁻¹ and 46
75 Gg yr⁻¹, respectively, the Yin et al. (2015) estimate is 64 Gg yr⁻¹, and the MEIC v1.2
76 estimate is 181 Gg yr⁻¹. Atmospheric-measurement-based estimates are needed to
77 validate benzene and toluene emissions estimated from bottom-up methods. However,
78 to date no top-down estimate is available for PRD and HK.

79 High-frequency online measurements of VOCs (including benzene and toluene)

80 were made during the PRIDE-PRD2010 Campaign (Program of Regional Integrated
81 Experiments on Air Quality over Pearl River Delta) during November and December
82 2010. This study uses these measurement data and an inverse modeling approach to
83 infer benzene and toluene emissions in the PRD and HK. This top-down estimate is
84 important to test and improve the existing bottom-up inventories.

85 **2 Methodology**

86 **2.1 Measurement data**

87 In this study, atmospheric measurements of benzene and toluene at two sites were
88 used, the Heshan site (used for the inversion) and the Mt. Tai Mo Shan (TMS) site
89 (used for validation). The Heshan site (112.929 °E, 22.728 °N) is a rural observatory
90 located on the top of a small hill (~60 m above the surrounding terrain; ~100 m above
91 sea level) in Jiangmen (see Figure 1). The measurement period at the Heshan site was
92 from November 11, 2010 to December 1, 2010. Data from December 1, 2010 were
93 not used, since we focused on mixing ratios and emissions in November. Detailed
94 information of the measurement system and procedure can be found in Wang et al.
95 (2014). Here we provide only a brief description. Ambient mixing ratios of VOCs
96 were measured using an online automatic gas chromatograph system equipped with a
97 mass spectrometer and a flame ionization detector (GC-MS/FID). Most C2-C5
98 hydrocarbons were measured by the FID Channel with a porous layer open tubular
99 (PLOT) column, whereas other VOCs, including benzene and toluene, were measured
100 by the Mass Selective Detector (MSD) Channel with a DB-624 column. The time

101 resolution of the VOC measurements was 60 minutes. The detection limits of this
102 system for benzene and toluene are 0.006 ppb and 0.015 ppb, respectively, which are
103 much lower values than the typical benzene and toluene mixing ratio levels of 2 ppb
104 and 6 ppb during the observation period at the Heshan site.

105 The Mt. TMS site (114.118 E, 22.405 N) was not used for the inversion but for
106 validating the emissions derived from the inversions in this study. The sample air inlet
107 at the TMS site was located on the rooftop of a building at Mt. TMS at an elevation of
108 640 m above sea level. A total of 75 canisters of air samples were taken over different
109 times of day and night on November 1–3, 9, and 19–21, 2010. Detailed information on
110 the sampling time schedule can be found in Guo et al. (2013). After sampling, the
111 VOC canister samples were sent to a laboratory at the University of California, Irvine
112 for chemical analysis. Simpson et al. (2010) provide a full description of the
113 analytical system, which uses a multi-column gas chromatograph (GC) with five
114 column-detector combinations. The measurement detection limit of this system for
115 both benzene and toluene is 0.003 ppb, which is much lower than the typical mixing
116 ratio levels of 0.7 ppb for benzene and 1.6 ppb for toluene during the observation
117 period at the Mt. TMS site.

118 The TMS data were not used in the inversion because: 1) The measurements
119 performed at the two stations were calibrated according to different scales, which
120 could cause problems in the inversion (see also Weiss and Prinn (2011)). 2) The
121 number of measurement data at the TMS site (totally 75 in Nov. 2010) is much

122 smaller than that at the Heshan site (totally 419), which means that the inversion
123 results would anyway be dominated by the Heshan data. 3) TMS is relatively close to
124 central Urban Hong Kong (~7 km; Guo et al., 2013) so that the TMS site might be
125 influenced by local sources and this is not desirable for the inversion. Tests with
126 inversions including TMS data have shown that the PRD benzene emissions would be
127 only ~15% higher from those using Heshan data only, which is within the a posteriori
128 emission uncertainty.

129 **2.2 Model simulations using FLEXPART**

130 The source-receptor relationships (SRRs, often also called “emission sensitivities”, in
131 units of $\text{m}^2 \text{s}^{-1}$) were calculated using the backwards in time mode of the Lagrangian
132 particle dispersion model, FLEXPART (<http://www.flexpart.eu>) (Stohl et al., 2005;
133 Stohl et al., 1998). The model was driven by hourly meteorological data of $0.5^\circ \times 0.5^\circ$
134 horizontal resolution and 37 vertical levels from the NCEP Climate Forecast System
135 Reanalysis (CFSR) (available at <http://rda.ucar.edu/datasets/ds093.0/>). During
136 3-hourly intervals throughout the sampling period, 80,000 virtual particles were
137 released at the site’s location and at the height of the sampling inlet above model
138 ground level, and followed backwards in time for 20 days. In FLEXPART, the
139 trajectories of tracer particles are calculated using the mean winds interpolated from
140 the analysis fields plus random motions representing turbulence (Stohl and Thomson,
141 1999). The emission sensitivity value in a particular grid cell is proportional to the
142 particle residence time in that cell (Seibert and Frank, 2004). Residence time is

143 specifically for the layer from the surface up to a specified height in the planetary
144 boundary layer (100 m used by this study and previous studies, e.g., (Stohl et al.
145 (2009))). The spatial resolution of the output from the backward simulations is $0.25^\circ \times$
146 0.25° . Loss of benzene and toluene by reaction with the hydroxyl (OH) radical in the
147 atmosphere was considered in the backward simulation. Rate constant values for
148 reaction with OH radicals were expressed for benzene as:

$$149 \quad k = 2.308 \times 10^{-12} \times \exp\left(-\frac{190}{T}\right) \quad (1);$$

150 and for toluene as:

$$151 \quad k = 1.275 \times 10^{-18} \times T^2 \times \exp\left(\frac{1192}{T}\right) \quad (2),$$

152 where T is the ambient temperature (K). Gridded OH fields (hourly for the period
153 Oct to Dec 2010, at a resolution of $0.5^\circ \times 0.667^\circ$, 47 vertical levels) were derived from
154 the atmospheric chemistry transport model, GEOS-Chem v5
155 (<http://acmg.seas.harvard.edu/geos/>). A reference simulation was run backwards for
156 20 days with atmospheric chemical loss, and additional alternative FLEXAPRT
157 simulations were run backwards for 10 and 40 days with atmospheric chemical loss,
158 and 20 days without atmospheric chemical loss (see Section 3.2 shows).

159 **2.3 Inverse algorithm**

160 Simulated benzene and toluene mixing ratios at the measurement site were obtained
161 by integrating the gridded emission sensitivities ($\text{m}^2 \text{s g}^{-1}$) multiplied by the gridded
162 emissions ($\text{g m}^{-2} \text{s}^{-1}$). The Bayesian inversion method used in this study is almost the

163 same as described and evaluated by Stohl et al. (2009) and Stohl et al. (2010), and as
164 used in recent studies of SF₆ emissions (Fang et al., 2014) and HFC-23 emissions
165 (Fang et al., 2015) in East Asia. Briefly, in this study a Bayesian inversion technique
166 is employed, based on least-squares optimization, to estimate both the spatial
167 distribution and strength of the emissions in the domain to which the measurements
168 are sensitive. The inversion adjusts the emissions to minimize the differences between
169 the observed and modeled mixing ratios while also considering the deviation of the
170 optimized emissions from an a priori emission field. Observation-model mismatch
171 errors (which include transport model errors) are determined as the root mean square
172 error (RMSE) of the model-observation mismatch (Stohl et al., 2009; Stohl et al.,
173 2010). In this study, mixing ratio background values were set to zero. This is because
174 the backward simulations were run for 20 days and benzene and toluene in the air
175 parcel from emissions occurring prior to this time would have been largely removed
176 from the atmosphere by reaction with OH (typical atmospheric lifetimes of benzene
177 and toluene are ~10 days and ~2 days, respectively).

178 For benzene, gridded a priori emission fields for mainland China were derived
179 from MEIC v1.2 for November 2010 (0.25° × 0.25°, monthly mean), and for the rest
180 of the world the emissions were taken from RCP Scenario 2.6 (0.5° × 0.5°, annual
181 mean) (van Vuuren et al., 2007). For toluene, a priori emission fields for mainland
182 China were derived from MEIC v1.2 for November 2010 (0.25° × 0.25°, monthly
183 mean), while for the PRD region in mainland China, a priori emissions were derived
184 by averaging the estimates from MEIC v1.2 (0.25° × 0.25°, monthly mean) and from

185 Yin et al. (2015) ($0.25^\circ \times 0.25^\circ$, monthly mean) for November 2010. For the rest of
186 the world, a priori emissions were taken from RCP 2.6 inventory ($0.5^\circ \times 0.5^\circ$, yearly
187 mean) (van Vuuren et al., 2007). Both monthly inventories of MEIC v1.2 and Yin et
188 al. (2015) were obtained through personal communication with the dataset authors. A
189 priori emission uncertainty in each grid cell for benzene and toluene, respectively, was
190 set to 100% and 70% according to the differences among the bottom-up inventories
191 and was assumed uncorrelated in space. The a posteriori uncertainty of the emissions
192 in each grid cell was calculated as described by Seibert et al. (2011), and the
193 uncertainty reduction in each grid cell represents the difference (as a percentage)
194 between the a posteriori and priori emission uncertainties in the corresponding grid
195 cell.

196 **3 Results and Discussion**

197 **3.1 Benzene and toluene ambient mixing ratios**

198 Table 1 shows ambient mixing ratios of benzene and toluene measured at the Heshan
199 site and other sites all over the world. Mixing ratios of benzene at the Heshan site
200 ranged from 0.59 ppb to 20.23 ppb and had an average of 2.27 ± 1.65 (mean \pm standard
201 deviation) ppb during our observation period. Mixing ratios of toluene at the Heshan
202 site ranged from 0.87 ppb to 25.05 ppb and had an average of 5.6 ± 4.15 ppb. The
203 mixing ratios of benzene (0.67 ± 0.21 ppb) and toluene (1.58 ± 1.25 ppb) at the Mt.
204 TMS were only ~30% of those at the Heshan site. In agreement with previous studies
205 (e.g. Lau et al., 2010; Liu et al., 2008), mixing ratio levels of benzene and toluene in

206 the PRD region are overall higher than those in Hong Kong (Table 1), which is in part
207 due to the fact that Hong Kong often receives clean air masses from over the ocean
208 and that emissions in Hong Kong are lower than in the PRD.

209 Mixing ratios of benzene and toluene in some cities in Europe (e.g. Ait-Helal et al.,
210 2014; Langford et al., 2010) and United States (e.g. USEPA, 1989; Baker et al., 2008)
211 have been found to be approximately 0.5 ppb and 1 ppb (Table 1), respectively, which
212 is about 20% of the mean observed values in the PRD in this study. Mixing ratios of
213 benzene and toluene in Thompson Farm, United States were even 0.08 ± 0.002 ppb and
214 0.09 ± 0.005 ppb, respectively, which are much lower than the lowest mixing ratios at
215 both Heshan and Mt. TMS sites. Levels of benzene and toluene mixing ratios at
216 different sites mainly reflect the combined influence of emission strength, seasonal
217 changes in atmospheric OH concentration and mixing depth.

218 **3.2 Benzene and toluene emission sensitivities**

219 Figure 2 shows the spatial distribution of average emission sensitivity of benzene and
220 toluene for the Heshan site for November 12-November 30, 2010. During the
221 observation period, air masses transported to the Heshan site mainly came from
222 easterly and northerly directions. Considering that the major emission sources in the
223 PRD are located to the east of the Heshan site (Figure 1), this measurement location is
224 ideally situated for constraining emissions from this region for this period and, as the
225 emission sensitivities show, PRD, HK, and neighboring regions, are relatively well
226 constrained by the observations at the Heshan site. Benzene and toluene emissions in

227 the PRD and HK are much higher than emissions in neighboring regions (Figure 1)
228 and, consequently, the overall mixing ratio contributions (the integral of the emission
229 sensitivities multiplied by emissions) from PRD and HK to the observation site
230 comprise more than 80% of the total simulated mixing ratios. Note that the emission
231 sensitivities for benzene and toluene are different because there are differences in the
232 chemical loss of these two compounds during atmospheric transport and in the
233 molecular weight. Specifically, the emission sensitivities for toluene are spatially
234 more confined because of its shorter lifetime.

235 As a sensitivity study, alternative simulations in which FLEXPART was ran
236 backwards for 10 days were made. The derived emission sensitivities are almost
237 identical to the reference simulations with 20 days duration (Supporting Information
238 Figure S1 for benzene and Figure S2 for toluene), confirming that 20-day-backward
239 simulations are sufficiently long to account for all benzene and toluene emission
240 sources that can influence the mixing ratios at the Heshan site. Since the lifetime of
241 benzene is ~10 days (much longer than that of toluene), we also made a
242 40-day-backward simulation from which the emission sensitivities for benzene are
243 also almost identical to the reference simulation of 20 days (Figure S3). Without
244 accounting for the loss by reaction with OH in the atmosphere, the emission
245 sensitivities for benzene would only be a little higher (by ~10% in central PRD)
246 (Figure S4). On the other hand, the emission sensitivities for toluene would be much
247 higher (by ~50% in central PRD) (Figure S5). This indicates that accounting for
248 chemical loss has a relatively small effect for simulating benzene mixing ratios at

249 Heshan, whereas it has a profound effect on toluene mixing ratios. Thus, errors in the
250 retrieved emissions due to errors in chemical loss are marginal for benzene but could
251 be significant for toluene.

252 **3.3 Inversion results**

253 Figure 3 shows the observed and simulated mixing ratios at the Heshan site. The
254 simulations captured most pollution episodes and the inversion improved the
255 agreement between the simulations and the observations as expected (the agreement
256 between the a posteriori simulations and the observations is better than for the a priori
257 simulations and the observations). For benzene, the RMSE between the observed and
258 simulated mixing ratios decreased from 1.53 ppb, using a priori emissions, to 1.26
259 ppb, using a posteriori emissions, and the mean bias between the simulated mixing
260 ratios and observations decreased from 0.96 ppb, using a priori emissions, to 0.41 ppb,
261 using a posteriori emissions. For toluene, the RMSE between the observed and
262 simulated mixing ratios decreased from 4.77 ppb, using a priori emissions, to 4.30
263 ppb, using a posteriori emissions and the mean bias between the observed and
264 simulated mixing ratios decreased from 2.35 ppb, using a priori emissions, to 1.99
265 ppb, using a posteriori emissions.

266 Figure 3 also shows examples of spatial distributions of toluene emission
267 sensitivities for two observed mixing ratios. The toluene mixing ratio at 00:00 UTC
268 on 16 November 2010 was about 2 ppb and the corresponding air mass had not passed
269 over the strong emission sources in the central part of PRD and HK (see the backward

270 emission sensitivities map in Figure 3c), while the toluene mixing ratio at 00:00 UTC
271 on 24 November 2010 was about 15 ppb and the corresponding air mass had passed
272 over the strong emission sources in the central part of PRD and HK (Figure 3d).

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

285 Figure 5 shows the a priori and a posteriori emissions of toluene and their
286 difference. Emission hot spots are located in Guangzhou and Shenzhen. The
287 uncertainty reduction map in Figure 4d and Figure 5d shows significant error
288 reductions, of 40% or more, in boxes close to the observation site, while only low
289 emission uncertainty reductions were achieved in boxes far from the observation site.
290 Overall, the emission uncertainties have been reduced by the inversion in the PRD
291 and HK, where the strongest emission sources are located.

292 The total a posteriori benzene emissions for PRD and HK, respectively, are 4.0
293 (1.1–6.9) Gg/month (for all cases the range represents one sigma uncertainty) and 0.4
294 (0.1–0.7) Gg/month. A posteriori toluene emissions are 12 (4–20) Gg/month for PRD
295 and 0.5 (0.2–0.9) Gg/month for HK. The inversion sensitivity tests, i.e., using other
296 bottom-up emission inventories for the a priori estimate (listed in Table 2), all produce
297 toluene emission estimates that fall within the uncertainty range of the a posteriori
298 emissions from the reference inversion.

299 Benzene and toluene measurement data at the Mt. TMS site were not used in the
300 inversion but for validating the posterior emissions. For benzene, using the a priori
301 and a posteriori emission fields, respectively, the RMSE between the simulated and
302 observed mixing ratios at Mt. TMS site are 0.367 ppb and 0.312 ppb, and the mean
303 bias between the simulated and observed mixing ratios are 0.314 ppb and 0.208 ppb.
304 For toluene, the RMSE (1.50 ppb) between the observed and simulated mixing ratios
305 using the a posteriori emission fields from the inversion was smaller than that (1.55
306 ppb) using the a priori field; the mean bias (1.06 ppb) between the observations and
307 simulated mixing ratios using a posteriori emission fields was also smaller than that
308 (1.12 ppb) using the a priori field. Both the RMSEs and mean bias suggest that the a
309 posteriori emissions are more accurate than the a priori emissions.

310 We also made FLEXPART simulations driven by operational meteorological
311 analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF)
312 instead of CFSR data. The a posteriori emissions for the PRD are very similar when
313 using the emission sensitivities from the two alternative FLEXPART simulations, e.g.

314 for benzene we obtained 4.0 Gg/month from the inversion using CFSR and 4.2
315 Gg/month from the inversion using ECMWF. Although Fang et al. (2014) showed
316 that FLEXPART simulations driven with ECMWF data performed slightly better than
317 the CFSR-driven simulations for SF₆ in East Asia for Hateruma, Gosan and Cape
318 Ochiishi stations, we found that CFSR-driven FLEXPART simulations performed
319 slightly better than the ECMWF-driven simulations for the benzene simulations at the
320 Heshan site. Thus, the CFSR dataset was used in this paper.

321 **3.4 Comparison with other estimates**

322 Figure 6 and Figure S6, respectively, show spatial distributions of benzene and
323 toluene emissions estimated by the inversion in this study, four bottom-up inventories,
324 and the differences among these estimates. For benzene, the spatial emission
325 distributions in the REAS v1.1 REF have the biggest difference from our top-down
326 emissions. Gridded emissions in the REAS v1.1 REF are always lower than the
327 inversion emissions, while emissions in the Yin et al. (2015), MEIC v1.2 and RCP 2.6
328 estimates are less systematically biased. The simulated benzene mixing ratios using
329 the REAS v1.1 inventory are much lower than the observed mixing ratios (Figure 7).
330 Statistics of the RMSE, mean bias and squared Pearson correlation coefficients
331 between the simulated and observed mixing ratios show that emission fields obtained
332 from the inversion performed better in simulating the benzene mixing ratios than all
333 four bottom-up inventories (See Table S1).

334 For toluene, in most grid cells over the PRD, emissions estimated by RCP 2.6, Yin

335 et al. (2015) and REAS v1.1 REF are lower than the inversion estimates, while MEIC
336 v1.2 emissions are higher than the inversion estimates (Figure S6). Model simulations
337 show that the simulated mixing ratios using emission estimates from RCP 2.6, Yin et
338 al. (2015) and REAS v1.1 REF are much lower than the observed mixing ratios at the
339 Heshan site (Figure S7). The simulated mixing ratios using MEIC v1.2 emission
340 fields are not consistent with some observed pollution peaks (Figure S7). Statistics of
341 RMSE and squared Pearson correlation coefficients show that inversion emission
342 fields performed better at simulating toluene mixing ratios at the Heshan site than the
343 four bottom-up emission fields (see Table S1).

344 Table 2 shows five estimates of total benzene and toluene emissions in the PRD
345 and HK regions for the year 2010. The a posteriori emissions for November 2010
346 obtained from the inversion were extrapolated to an annual mean emission rate for the
347 whole year 2010 by multiplying the November emissions by the ratio of emissions for
348 the whole year 2010 to those in November 2010. For toluene, this ratio is 10.8, and
349 was calculated from both the MEIC v1.2 and Yin et al. (2015) estimate (the
350 November/annual emission ratio was the same in both datasets). For toluene, the
351 factor is 10.9 (10.4–11.4), and is the average of 10.4, calculated from the MEIC v1.2
352 estimate, and 11.4, calculated from the Yin et al. (2015) estimate. Data in November
353 2010 and the whole year 2010 were obtained through personal communication with
354 the dataset authors. Using these ratios, the benzene emissions in the PRD and HK for
355 2010 were estimated to be 44 (12–75) Gg yr⁻¹ and 5 (2–7) Gg yr⁻¹, respectively, and
356 the toluene emissions were estimated to be 131 (44–218) Gg yr⁻¹ and 6 (2–9) Gg yr⁻¹,

357 respectively.

358 For benzene, emissions in the PRD in 2010 calculated from the four bottom-up
359 estimates were 45 Gg yr⁻¹ from RCP 2.6 (van Vuuren et al., 2007), 54 Gg yr⁻¹ from
360 Yin et al. (2015), 8 Gg yr⁻¹ from REAS v1.1 REF (Ohara et al., 2007), and 33 Gg yr⁻¹
361 from MEIC v1.2. Our inverse estimate agrees within its uncertainties with these
362 bottom-up estimates, except that the REAS estimate is substantially lower than the
363 other bottom-up and the top-down estimates. Emissions in HK were 5 (2–7) Gg yr⁻¹
364 estimated by this study, which agrees within uncertainties with the RCP 2.6 estimate
365 and is much higher than the REAS v1.1 REF (no estimates are available from MEIC
366 v1.2 or Yin et al. (2015)).

367 For toluene, emissions in PDR in 2010 calculated from the four bottom-up
368 estimates were 44 Gg yr⁻¹ from RCP 2.6 estimate (van Vuuren et al., 2007), 64 Gg yr⁻¹
369 from Yin et al. (2015), 46 Gg yr⁻¹ from REAS v1.1 REF (Ohara et al., 2007), and 181
370 Gg yr⁻¹ from MEIC v1.2. The bottom-up estimate MEIC v1.2 meets the high
371 uncertainty range of our inversion estimates, while the other three bottom-up
372 estimates meet the low uncertainty range of our inversion estimates. For the HK
373 toluene emissions, estimates are not available in MEIC v1.2 or Yin et al. (2015); both
374 RCP 2.6 and REAS v1.1 REF estimates are about 4 Gg yr⁻¹, which agree with our
375 inversion results within uncertainties.

376 **3.5 Benzene and toluene emissions during 2000–2010**

377 Figure 8 shows different estimates of benzene and toluene emissions in the PRD

378 region for the period 2000–2010. For benzene, the estimate of 8 Gg yr⁻¹ in 2000 by
379 REAS v2.1 (Kurokawa et al., 2013) agrees with that of 13 Gg yr⁻¹ by the Reanalysis
380 of the Tropospheric chemical composition over the past 40 years project (RETRO)
381 (Schultz et al., 2007), which are substantially smaller than that of 43 Gg yr⁻¹ in the
382 Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP)
383 (Lamarque et al., 2010). For the years 2005 and 2006, different studies show
384 substantial differences. For the year 2005, the emission estimate by RCP 2.6 was ~4
385 times the estimates by REAS v2.1. For the year 2006, the emission estimate by REAS
386 v2.1 agrees with the estimate in the Intercontinental Chemical Transport
387 Experiment-Phase B (INTEX-B) project (Zhang et al., 2009), which were only ~20%
388 of the estimate by Zheng et al. (2009). More studies are available for the year 2010
389 than for other years. For the year 2010, the estimates by RCP 2.6, MEIC v1.2 and Yin
390 et al. (2015) agree with the inversion estimate by this study, which are higher than the
391 estimate by REAS v1.1 REF. According to these bottom-up and top-down estimates
392 (Figure 8), it is likely that the benzene emissions in the PRD have remained relatively
393 stable during the 2000–2010 period, although emissions are uncertain due to limited
394 number of estimates.

395 For toluene, the estimate of 45 Gg yr⁻¹ in 2000 by REAS v2.1 (Kurokawa et al.,
396 2013) agrees relatively well with the value of 36 Gg yr⁻¹ by ACCMIP (Lamarque et al.,
397 2010), but both are substantially larger than the RETRO estimate of 14 Gg yr⁻¹
398 (Schultz et al., 2007). For the years 2005 and 2006, estimates of toluene emission are
399 also quite different. For the year 2005, the emission estimate by REAS v2.1 was ~4

400 times the estimates by RCP 2.6. For the year 2006, the emission estimate by REAS
401 v2.1 was ~2 times the estimate by Zheng et al. (2009) and even ~11 times the estimate
402 by INTEX-B (Zhang et al., 2009). For the year 2010, the estimates by REAS v1.1
403 REF and Yin et al. (2015) meet the low end of uncertainty of inversion estimate by
404 this study, while MEIC v1.2 estimate meets the high end. According to these estimates
405 over 2000–2010 (Figure 8), it is likely that the toluene emissions in the PRD have
406 increased during this period, although emissions are uncertain due to limited number
407 of estimates.

408 Based on glyoxal (CHOCHO) data retrieved from satellite and inversion method,
409 Liu et al. (2012) found their emission estimates of the lumped artificial compound
410 ARO1 (benzene, toluene and ethylbenzene) in the PRD in 2006 were >10 times larger
411 than the bottom-up INTEX-B estimates (also for 2006), but they did not specify
412 which compound was responsible for the difference. As for benzene, the ratio of
413 emissions in 2006 estimated by Zheng et al. (2009) (60 Gg yr^{-1}) to the INTEX-B
414 estimate (15 Gg yr^{-1}) is ~4 times, much less than the factor >10 discrepancy reported
415 by Liu et al. (2012). Inversion estimate of benzene emissions ($44 (12\text{--}75) \text{ Gg yr}^{-1}$) in
416 2010 is ~3 (1–5) times the INTEX-B emissions for 2006. Thus, we suggest that the
417 big discrepancy is likely not due to emissions of benzene but emissions of toluene
418 and/or ethylbenzene. As for toluene emissions, the ratios of bottom-up estimates by
419 REAS v2.1 (190 Gg yr^{-1}) and Zheng et al. (2009) (103 Gg yr^{-1}) for 2006 to the
420 INTEX-B bottom-up estimate (18 Gg yr^{-1}) are 11–6 times. Thus, considering the
421 satellite-based estimate and other bottom-up estimates, the bottom-up INTEX-B

422 estimate of toluene emissions for the PRD region for 2006 was likely too low, and
423 estimation of toluene emissions in the PRD is attributed as an important factor
424 contributing to the big discrepancy of ARO1 emission estimates between Liu et al.
425 (2012) and INTEX-B.

426 **3.6 Suggestions for more top-down studies**

427 To the best of our knowledge, this study provides the only available top-down
428 estimate for toluene emissions in the PRD and HK regions. All other studies in Figure
429 8 are bottom-up estimates. More top-down estimates are needed to validate the
430 bottom-up estimates in previous years and in the future. In this study, inversions using
431 the Heshan measurement data reduced emission uncertainties in the PRD and HK
432 regions. However, the emission uncertainty reductions were not large because there
433 was only one observation site suitable for the inversion (some measurements in urban
434 environments are available but not suitable for inverse modeling) and the observation
435 period was not long. Thus, we propose that in the future, observations with better
436 spatial and temporal coverage are urgently needed to better constrain benzene and
437 toluene (and other VOC) emissions in the PRD and HK regions. Inversion-suited
438 observation sites could be situated in rural places outside of the major emission
439 sources located in the central part of PRD and HK regions, and then the major
440 emission sources in the PRD and HK regions could be “viewed” from different angles
441 (multiple-site inversion) to better constrain the benzene and toluene (and other VOC)
442 emissions.

443 **4 Conclusions**

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

465 **Supporting Information**

466 Supplementary material related to this article is available online at

467 <http://www.atmos-chem-phys.net/>

468 **Acknowledgement**

469 This study was partly funded by the Natural Science Foundation for Outstanding

470 Young Scholars (Grant No. 41125018) and Key project (Grant No. 41330635). This

471 work was supported in part by a National Aeronautics and Space Administration

472 (NASA) grant NNX11AF17G awarded to the Massachusetts Institute of Technology

473 (MIT). We acknowledge Emissions of atmospheric Compounds & Compilation of

474 Ancillary Data (http://eccad.sedoo.fr/eccad_extract_interface/JSF/page_critere.jsf) for

475 the archiving of the emission inventory data of ACCMIP, REAS v1.1 REF, RCP 2.6

476 and RETRO. We thank Mingwei Li at Department of Earth, Atmospheric and

477 Planetary Sciences, MIT for her help in GEOS-Chem model runs.

References

- 479 Ait-Helal, W., Borbon, A., Sauvage, S., de Gouw, J. A., Colomb, A., Gros, V., Freutel, F., Crippa, M., Afif,
480 C., Baltensperger, U., Beekmann, M., Doussin, J. F., Durand-Jolibois, R., Fronval, I., Grand, N.,
481 Leonardis, T., Lopez, M., Michoud, V., Miet, K., Perrier, S., Prévôt, A. S. H., Schneider, J., Siour,
482 G., Zapf, P., and Locoge, N.: Volatile and intermediate volatility organic compounds in
483 suburban Paris: variability, origin and importance for SOA formation, *Atmos. Chem. Phys.*, **14**,
484 10439-10464, 10.5194/acp-14-10439-2014, 2014.
- 485 Baker, A. K., Beyersdorf, A. J., Doezema, L. A., Katzenstein, A., Meinardi, S., Simpson, I. J., Blake, D. R.,
486 and Sherwood Rowland, F.: Measurements of nonmethane hydrocarbons in 28 United States
487 cities, *Atmos. Environ.*, **42**, 170-182, 2008.
- 488 Barletta, B., Meinardi, S., Simpson, I. J., Khwaja, H. A., Blake, D. R., and Rowland, F. S.: Mixing ratios of
489 volatile organic compounds (VOCs) in the atmosphere of Karachi, Pakistan, *Atmos. Environ.*,
490 **36**, 3429-3443, 2002.
- 491 Barletta, B., Meinardi, S., Rowland, F. S., Chan, C. Y., Wang, X. M., Zou, S. C., Chan, L. Y., and Blake, D. R.:
492 Volatile organic compounds in 43 Chinese cities, *Atmos. Environ.*, **39**, 5979-5990, 2005.
- 493 Barletta, B., Meinardi, S., Simpson, I. J., Zou, S., Sherwood Rowland, F., and Blake, D. R.: Ambient
494 mixing ratios of nonmethane hydrocarbons (NMHCs) in two major urban centers of the Pearl
495 River Delta (PRD) region: Guangzhou and Dongguan, *Atmos. Environ.*, **42**, 4393-4408, 2008.
- 496 Chan, L.-Y., Chu, K.-W., Zou, S.-C., Chan, C.-Y., Wang, X.-M., Barletta, B., Blake, D. R., Guo, H., and Tsai,
497 W.-Y.: Characteristics of nonmethane hydrocarbons (NMHCs) in industrial, industrial-urban,
498 and industrial-suburban atmospheres of the Pearl River Delta (PRD) region of south China, *J.*
499 *Geophys. Res. Atmos.*, **111**, D11304, 10.1029/2005JD006481, 2006.
- 500 Donald, J. M., Hooper, K., and Hopenhayn-Rich, C.: Reproductive and Developmental Toxicity of
501 Toluene: A Review, *Environ. Health Perspect.*, **94**, 237-244, 10.2307/3431317, 1991.
- 502 Fang, X., Thompson, R. L., Saito, T., Yokouchi, Y., Kim, J., Li, S., Kim, K. R., Park, S., Graziosi, F., and Stohl,
503 A.: Sulfur hexafluoride (SF₆) emissions in East Asia determined by inverse modeling, *Atmos.*
504 *Chem. Phys.*, **14**, 4779-4791, 10.5194/acp-14-4779-2014, 2014.
- 505 Fang, X., Stohl, A., Yokouchi, Y., Kim, J., Li, S., Saito, T., Park, S., and Hu, J.: Multiannual Top-Down
506 Estimate of HFC-23 Emissions in East Asia, *Environ. Sci. Technol.*, **72**, 10.1021/es505669j,
507 2015.
- 508 Guo, H., Jiang, F., Cheng, H. R., Simpson, I. J., Wang, X. M., Ding, A. J., Wang, T. J., Saunders, S. M.,
509 Wang, T., Lam, S. H. M., Blake, D. R., Zhang, Y. L., and Xie, M.: Concurrent observations of air
510 pollutants at two sites in the Pearl River Delta and the implication of regional transport,
511 *Atmos. Chem. Phys.*, **9**, 7343-7360, 10.5194/acp-9-7343-2009, 2009.
- 512 Guo, H., Ling, Z. H., Cheung, K., Jiang, F., Wang, D. W., Simpson, I. J., Barletta, B., Meinardi, S., Wang, T.
513 J., Wang, X. M., Saunders, S. M., and Blake, D. R.: Characterization of photochemical pollution
514 at different elevations in mountainous areas in Hong Kong, *Atmos. Chem. Phys.*, **13**,
515 3881-3898, 10.5194/acp-13-3881-2013, 2013.
- 516 Guo, S., Hu, M., Zamora, M. L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L., Molina,
517 M. J., and Zhang, R.: Elucidating severe urban haze formation in China, *Proc. Natl. Acad. Sci.*
518 *USA*, 10.1073/pnas.1419604111, 2014.
- 519 Henze, D. K., Seinfeld, J. H., Ng, N. L., Kroll, J. H., Fu, T. M., Jacob, D. J., and Heald, C. L.: Global

520 modeling of secondary organic aerosol formation from aromatic hydrocarbons: high- vs.
521 low-yield pathways, *Atmos. Chem. Phys.*, 8, 2405-2420, 10.5194/acp-8-2405-2008, 2008.

522 Kurokawa, J., Ohara, T., Morikawa, T., Hanayama, S., Janssens-Maenhout, G., Fukui, T., Kawashima, K.,
523 and Akimoto, H.: Emissions of air pollutants and greenhouse gases over Asian regions during
524 2000–2008: Regional Emission inventory in ASia (REAS) version 2, *Atmos. Chem. Phys.*, 13,
525 11019-11058, 10.5194/acp-13-11019-2013, 2013.

526 Lamarque, J. F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A.,
527 Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R.,
528 Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.:
529 Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive
530 gases and aerosols: methodology and application, *Atmos. Chem. Phys.*, 10, 7017-7039,
531 10.5194/acp-10-7017-2010, 2010.

532 Langford, B., Nemitz, E., House, E., Phillips, G. J., Famulari, D., Davison, B., Hopkins, J. R., Lewis, A. C.,
533 and Hewitt, C. N.: Fluxes and concentrations of volatile organic compounds above central
534 London, UK, *Atmos. Chem. Phys.*, 10, 627-645, 10.5194/acp-10-627-2010, 2010.

535 Lau, A. K. H., Yuan, Z., Yu, J. Z., and Louie, P. K. K.: Source apportionment of ambient volatile organic
536 compounds in Hong Kong, *Sci. Total Environ.*, 408, 4138-4149, 2010.

537 Liu, Y., Shao, M., Lu, S., Chang, C.-c., Wang, J.-L., and Chen, G.: Volatile Organic Compound (VOC)
538 measurements in the Pearl River Delta (PRD) region, China, *Atmos. Chem. Phys.*, 8,
539 1531-1545, 10.5194/acp-8-1531-2008, 2008.

540 Liu, Z., Wang, Y. H., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., Shao, M., Chang, C. C., Liu, S.
541 C., Wang, H. L., and Chen, C. H.: Exploring the missing source of glyoxal (CHOCHO) over China,
542 *Geophys. Res. Lett.*, 39, Art. L10812, Doi 10.1029/2012gl051645, 2012.

543 Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission
544 inventory of anthropogenic emission sources for the period 1980–2020, *Atmos. Chem.
545 Phys.*, 7, 4419-4444, 10.5194/acp-7-4419-2007, 2007.

546 Ou, J., Zheng, J., Li, R., Huang, X., Zhong, Z., Zhong, L., and Lin, H.: Speciated OVOC and VOC emission
547 inventories and their implications for reactivity-based ozone control strategy in the Pearl
548 River Delta region, China, *Sci. Total Environ.*, 530–531, 393-402, 2015.

549 Ran, L., Zhao, C., Geng, F., Tie, X., Tang, X., Peng, L., Zhou, G., Yu, Q., Xu, J., and Guenther, A.: Ozone
550 photochemical production in urban Shanghai, China: Analysis based on ground level
551 observations, *J. Geophys. Res. Atmos.*, 114, D15301, 10.1029/2008JD010752, 2009.

552 Schultz, M., Rast, S., van het Bolscher, M., Pulles, T., Brand, R., Pereira, J., Mota, B., Spessa, A.,
553 Dalsøren, S., van Noije, T., and Szopa, S.: Emission data sets and methodologies for estimating
554 emissions, RETRO project report D1-6, Hamburg, 2007,
555 http://retro.enes.org/reports/D1-6_final.pdf.

556 Seibert, P., and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion
557 model in backward mode, *Atmos. Chem. Phys.*, 4, 51-63, 10.5194/acp-4-51-2004, 2004.

558 Seibert, P., Kristiansen, N. I., Richter, A., Eckhardt, S., Prata, A. J., and Stohl, A.: Uncertainties in the
559 inverse modelling of sulphur dioxide eruption profiles, *Geomat. Nat. Hazards Risk*, 2, 201-216,
560 10.1080/19475705.2011.590533, 2011.

561 Simpson, I. J., Blake, N. J., Barletta, B., Diskin, G. S., Fuelberg, H. E., Gorham, K., Huey, L. G., Meinardi,
562 S., Rowland, F. S., Vay, S. A., Weinheimer, A. J., Yang, M., and Blake, D. R.: Characterization of
563 trace gases measured over Alberta oil sands mining operations: 76 speciated C2–C10 volatile

564 organic compounds (VOCs), CO₂, CH₄, CO, NO, NO₂, NO_y, O₃ and SO₂, *Atmos. Chem. Phys.*,
565 10, 11931-11954, 10.5194/acp-10-11931-2010, 2010.

566 Song, Y., Shao, M., Liu, Y., Lu, S., Kuster, W., Goldan, P., and Xie, S.: Source Apportionment of Ambient
567 Volatile Organic Compounds in Beijing, *Environ. Sci. Technol.*, 41, 4348-4353,
568 10.1021/es0625982, 2007.

569 Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model
570 FLEXPART against large-scale tracer experiment data, *Atmos. Environ.*, 32, 4245-4264, 1998.

571 Stohl, A., and Thomson, D.: A Density Correction for Lagrangian Particle Dispersion Models,
572 *Boundary-Layer Meteorology*, 90, 155-167, 10.1023/A:1001741110696, 1999.

573 Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle
574 dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5, 2461-2474, 2005.

575 Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Grealley, B. R., Lunder, C., Maione, M., Mühle, J.,
576 O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer,
577 M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional
578 and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons,
579 *Atmos. Chem. Phys.*, 9, 1597-1620, 2009.

580 Stohl, A., Kim, J., Li, S., O'Doherty, S., Mühle, J., Salameh, P. K., Saito, T., Vollmer, M. K., Wan, D., Weiss,
581 R. F., Yao, B., Yokouchi, Y., and Zhou, L. X.: Hydrochlorofluorocarbon and hydrofluorocarbon
582 emissions in East Asia determined by inverse modeling, *Atmos. Chem. Phys.*, 10, 3545-3560,
583 2010.

584 USEPA: Project Summary Determination of C₂ to C₁₂ Ambient Air Hydrocarbons in 39 U.S. Cities, from
585 1984 through 1986, Atmospheric Research and Exposure Assessment Laboratory, United
586 States Environmental Protection Agency, 1989.

587 van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J.: Global
588 Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol
589 Optical Depth: Development and Application, *Environ. Health Perspect.*, 118, 847-855, 2010.

590 van Vuuren, D., den Elzen, M. J., Lucas, P., Eickhout, B., Strengers, B., van Ruijven, B., Wonink, S., and
591 van Houdt, R.: Stabilizing greenhouse gas concentrations at low levels: an assessment of
592 reduction strategies and costs, *Climatic Change*, 81, 119-159, 10.1007/s10584-006-9172-9,
593 2007.

594 Velasco, E., Lamb, B., Westberg, H., Allwine, E., Sosa, G., Arriaga-Colina, J. L., Jobson, B. T., Alexander,
595 M. L., Prazeller, P., Knighton, W. B., Rogers, T. M., Grutter, M., Herndon, S. C., Kolb, C. E.,
596 Zavala, M., de Foy, B., Volkamer, R., Molina, L. T., and Molina, M. J.: Distribution, magnitudes,
597 reactivities, ratios and diurnal patterns of volatile organic compounds in the Valley of Mexico
598 during the MCMA 2002 & 2003 field campaigns, *Atmos. Chem. Phys.*, 7, 329-353,
599 10.5194/acp-7-329-2007, 2007.

600 Wang, M., Zeng, L., Lu, S., Shao, M., Liu, X., Yu, X., Chen, W., Yuan, B., Zhang, Q., Hu, M., and Zhang, Z.:
601 Development and validation of a cryogen-free automatic gas chromatograph system
602 (GC-MS/FID) for online measurements of volatile organic compounds, *Anal. Methods*, 6,
603 9424-9434, 10.1039/C4AY01855A, 2014.

604 Weiss, R. F., and Prinn, R. G.: Quantifying greenhouse-gas emissions from atmospheric measurements:
605 a critical reality check for climate legislation, *Philosophical Transactions of the Royal Society
606 a-Mathematical Physical and Engineering Sciences*, 369, 1925-1942, 10.1098/rsta.2011.0006,
607 2011.

608 White, M. L., Russo, R. S., Zhou, Y., Ambrose, J. L., Haase, K., Frinak, E. K., Varner, R. K., Wingenter, O.
609 W., Mao, H., Talbot, R., and Sive, B. C.: Are biogenic emissions a significant source of
610 summertime atmospheric toluene in the rural Northeastern United States?, *Atmos. Chem.*
611 *Phys.*, 9, 81-92, 10.5194/acp-9-81-2009, 2009.

612 Xue, L. K., Wang, T., Gao, J., Ding, A. J., Zhou, X. H., Blake, D. R., Wang, X. F., Saunders, S. M., Fan, S. J.,
613 Zuo, H. C., Zhang, Q. Z., and Wang, W. X.: Ground-level ozone in four Chinese cities:
614 precursors, regional transport and heterogeneous processes, *Atmos. Chem. Phys.*, 14,
615 13175-13188, 10.5194/acp-14-13175-2014, 2014.

616 Yin, S., Zheng, J., Lu, Q., Yuan, Z., Huang, Z., Zhong, L., and Lin, H.: A refined 2010-based VOC emission
617 inventory and its improvement on modeling regional ozone in the Pearl River Delta Region,
618 China, *Sci. Total Environ.*, 514, 426-438, 2015.

619 Yoshino, A., Nakashima, Y., Miyazaki, K., Kato, S., Suthawaree, J., Shimo, N., Matsunaga, S., Chatani, S.,
620 Apel, E., Greenberg, J., Guenther, A., Ueno, H., Sasaki, H., Hoshi, J.-y., Yokota, H., Ishii, K., and
621 Kajii, Y.: Air quality diagnosis from comprehensive observations of total OH reactivity and
622 reactive trace species in urban central Tokyo, *Atmos. Environ.*, 49, 51-59, 2012.

623 Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy,
624 S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for
625 the NASA INTEX-B mission, *Atmos. Chem. Phys.*, 9, 5131-5153, 10.5194/acp-9-5131-2009,
626 2009.

627 Zheng, J., Shao, M., Che, W., Zhang, L., Zhong, L., Zhang, Y., and Streets, D.: Speciated VOC Emission
628 Inventory and Spatial Patterns of Ozone Formation Potential in the Pearl River Delta, China,
629 *Environ. Sci. Technol.*, 43, 8580-8586, 10.1021/es901688e, 2009.

630

Tables

632 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 indicates
633 Not Given).

Location	Type	Time	Benzene			Toluene			Reference
			Sample number	Range	Mean \pm SD	Sample number	Range	Mean \pm SD	
(1) PRD and Hong Kong regions, China									
Heshan, PRD	Rural	11–30 Nov. 2010	419	0.59–20.23	2.27 \pm 1.65	419	0.87–25.05	5.65 \pm 4.15	This study
Guangzhou, PRD	Urban	4 Oct. to 3 Nov. 2004	111	0.66–11.35	2.39 \pm 1.99	111	0.76–36.91	7.01 \pm 7.33	(Liu et al., 2008)
Xinken, PRD	Rural	4 Oct. to 3 Nov. 2004	83	0.52–6.26	1.42 \pm 0.98	83	0.54–56.41	8.46 \pm 9.94	(Liu et al., 2008)
Dongguan, PRD	Urban	Sep. 2005	48	0.27–6.45	1.26 \pm 0.14	48	0.53–25.30	6.13 \pm 0.81	(Barletta et al., 2008)
Guangzhou, PRD	Urban	Sep. 2006	42	0.65–6.80	2.05 \pm 1.49	42	0.72–19.60	5.87 \pm 4.11	(Barletta et al., 2008)
Industrial Area, PRD	Industrial	Late summer 2000	15	NG	2.80 \pm 1.70	15	NG	13.5 \pm 11.8	(Chan et al., 2006)
Mt. Tai Mo Shan, Hong Kong	Mountain	1–3, 9, 19–21 Nov. 2010	75	0.38–1.79	0.67 \pm 0.21	75	0.26–6.30	1.58 \pm 1.25	This study
Tap Mun, Hong Kong	Rural	Nov. 2006 to Oct. 2007	39	0.05–1.67	0.56 \pm 0.41	39	0.15–7.12	1.61 \pm 1.55	(Lau et al., 2010)
Central West, Hong Kong	Urban	Nov. 2006 to Oct. 2007	40	0.05–1.91	0.60 \pm 0.50	40	0.28–8.81	2.64 \pm 2.07	(Lau et al., 2010)
(2) Other sites in China									
43 cities, China	Urban	Jan.–Feb. 2001	158	0.7–10.4	NG	158	0.4–11.2	NG	(Barletta et al., 2005)
Beijing, China	Urban	Aug. 2005	1046	NG	3.03 \pm 1.72	1039	NG	1.76 \pm 0.89	(Song et al., 2007)
Shanghai, China	Urban	15 Jun. 2006 to 14 Jun. 2007	~365	NG	6.07 \pm 11.70	~365	NG	32.80 \pm 21.60	(Ran et al., 2009)
(3) Sites in other countries									
Karachi, Pakistan	Urban	Winter of 1998–1999	78	0.34–19.3	5.20 \pm 4.50	78	0.19–37.0	7.10 \pm 7.60	(Barletta et al., 2002)

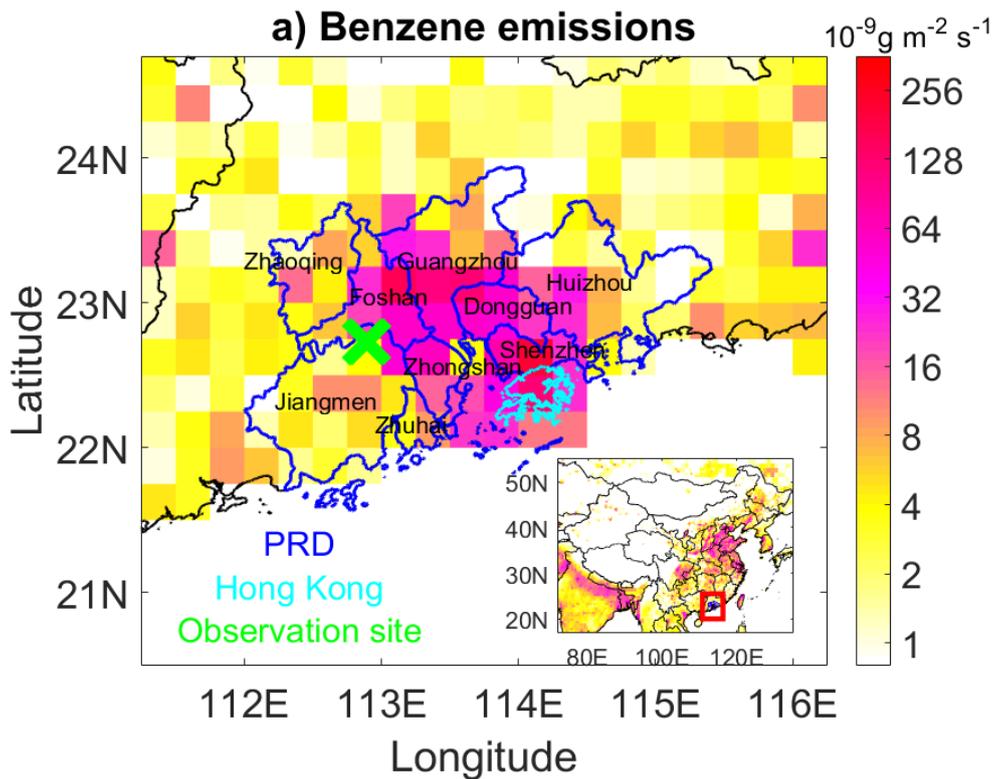
Tokyo, Japan	Urban	Summer 2007	50	NG	0.78±0.61	50	NG	2.14±0.99	(Yoshino et al., 2012)
Tokyo, Japan	Urban	Winter 2007	16	NG	0.82±0.28	16	NG	10.10±5.23	(Yoshino et al., 2012)
London, UK	Urban	Oct. 2010	601	NG	0.15±0.11	589	NG	0.68±0.57	(Langford et al., 2010)
Paris, France	Suburban	15 Jan.–15 Feb. 2010	246	NG	0.32±0.16	246	NG	0.32±0.22	(Ait-Helal et al., 2014)
Mexico City, Mexico	Urban	Feb. 2002 and Apr.–May 2003	~115	NG	3.17±1.75	~86	NG	13.5±9.33	(Velasco et al., 2007)
Mexico City, Mexico	Rural	Feb. 2002 and Apr.–May 2003	~115	NG	0.80±0.91	~86	NG	1.89±1.92	(Velasco et al., 2007)
39 cities, U.S.A.	Urban	Jun.–Sep. 1984–1986	835	0.001–0.27	NG	836	0.003–1.30	NG	(USEPA, 1989)
28 cities, U.S.A.	Urban	Summer 1999–2005	530	(0.06±0.024)– (0.48±0.24) ^a	NG	530	(0.12±0.055) –(1.54±0.88) ^a	NG	(Baker et al., 2008)
Thompson Farm, U.S.A.	Rural	Fall 2004–2006	201	NG	0.08±0.002	201	NG	0.09±0.005	(White et al., 2009)

634 ^aIt represents the range of the minimal mean value (the corresponding standard deviation) in one of 28 cities and maximal mean value (the corresponding standard
635 deviation) in another city.

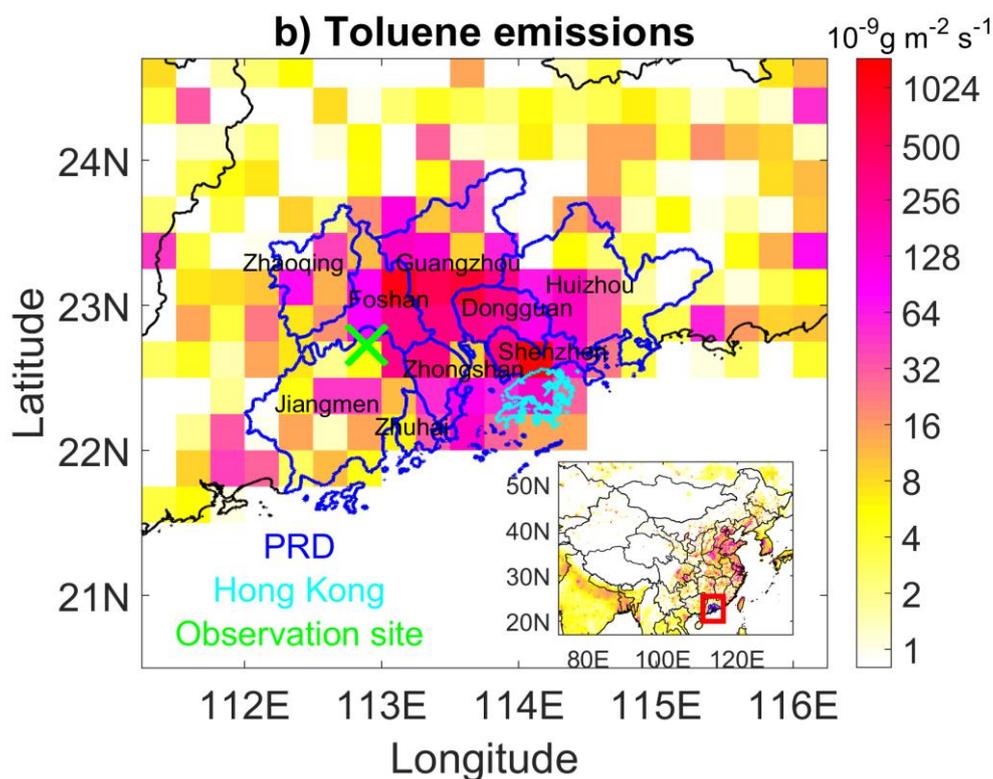
636 Table 2. Benzene and toluene emissions (Gg yr⁻¹) in the PRD and HK regions derived from
 637 different estimates for the year 2010.

Estimate	Benzene emissions		Toluene emissions	
	PRD	HK	PRD	HK
RCP 2.6	45	3	44	4
Yin et al. (2015)	54	NE	64	NE
REAS v1.1 REF	8	0.4	46	4
MEIC v1.2	33	NE	181	NE
This study	44 (12–75)	5 (2–7)	131 (44–218)	6 (2–9)

638 ^aNE indicates “Not Estimated”.

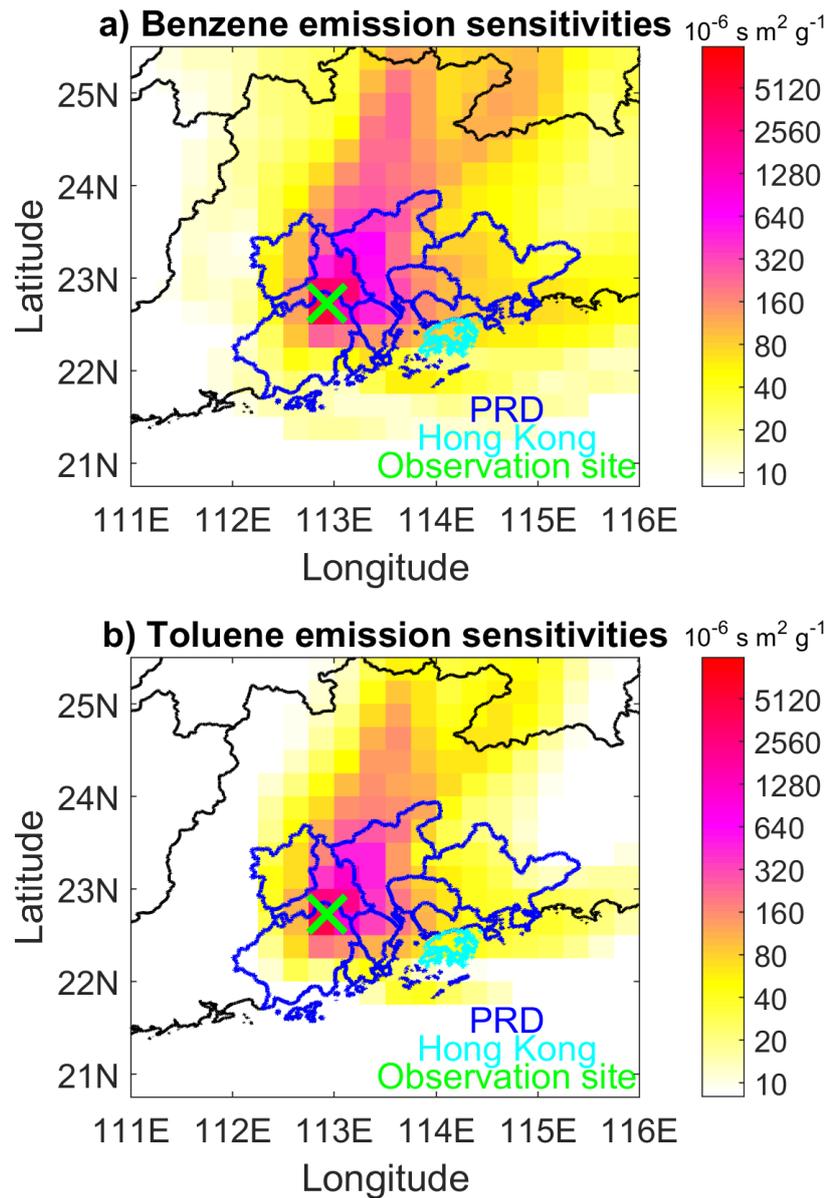


640



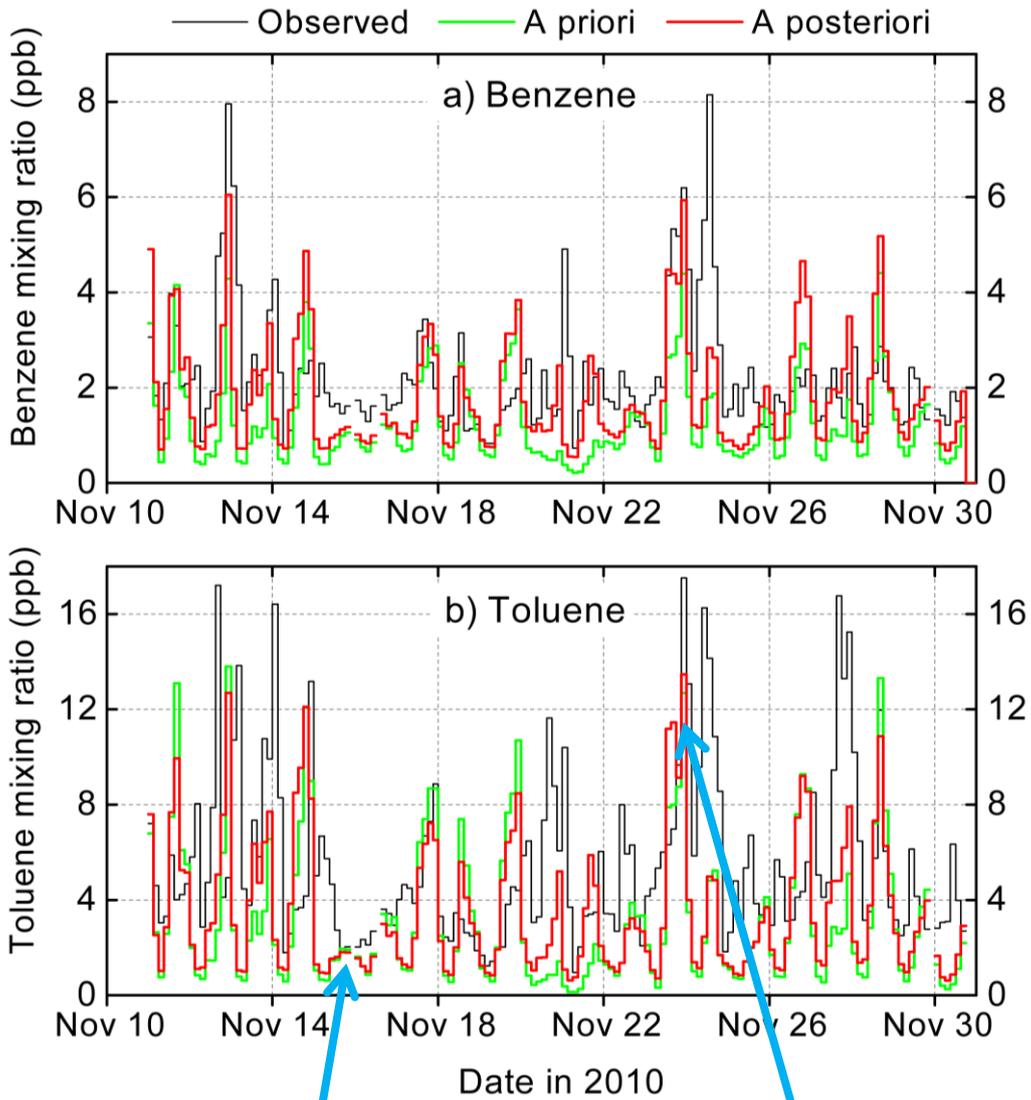
641

642 Figure 1. Map of a) benzene and b) toluene emissions from the MEIC v1.2 for China and the
 643 RCP 2.6 for outside China (inset panels), and that for the PRD and Hong Kong regions
 644 (mother panel). The PRD region is plotted with dark blue boundary lines, the Hong Kong
 645 region with cyan boundary lines. The green cross indicates the location of the Heshan
 646 observation site.

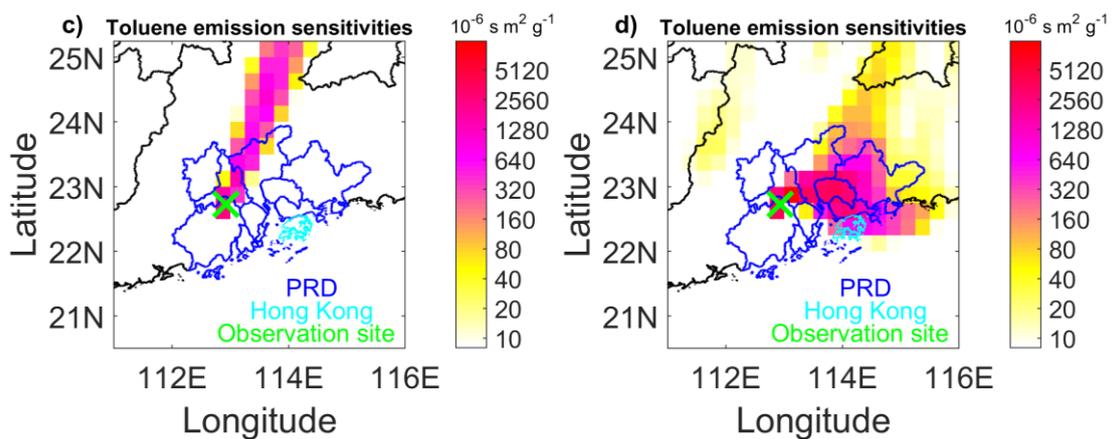


647

648 Figure 2. Average emission sensitivities of a) benzene and b) toluene for the Heshan
 649 observation site for November 12-November 31, 2010. The green cross indicates the location
 650 of Heshan site. The blue and cyan lines represent PRD and Hong Kong boundary lines,
 651 respectively.

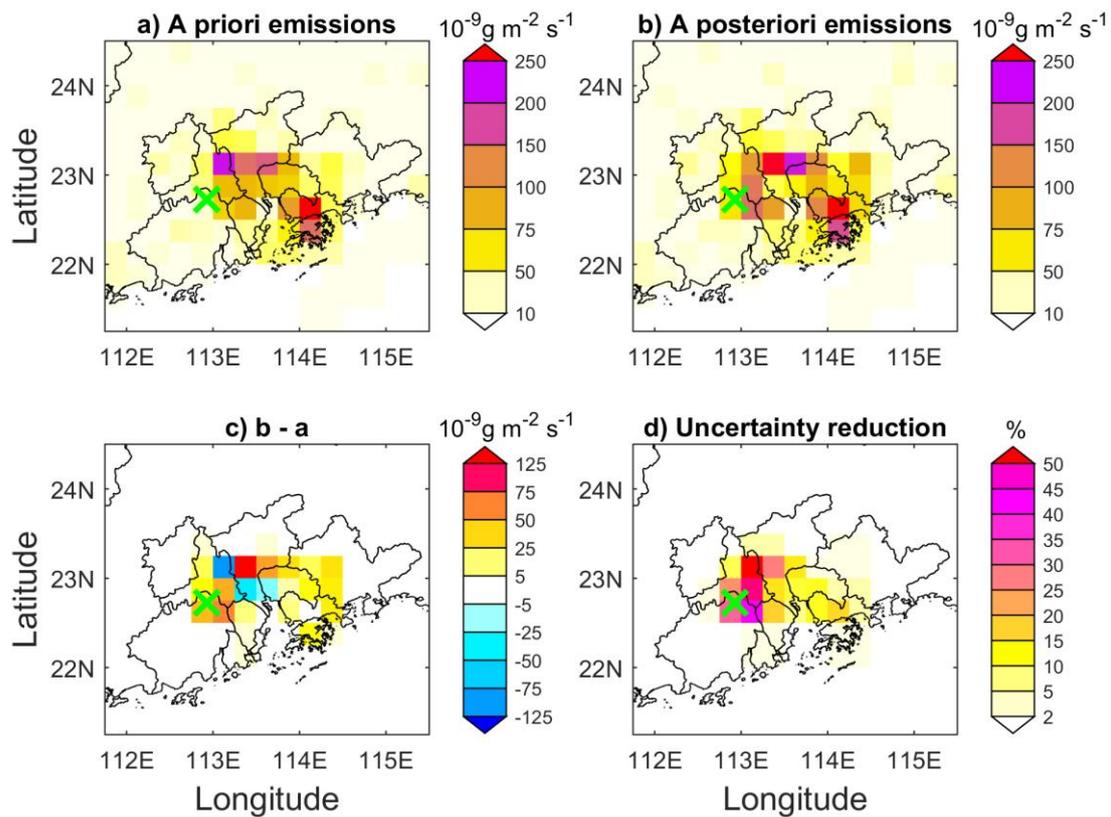


652



653

654 Figure 3. Observed and simulated a) benzene and b) toluene mixing ratios at the Heshan site,
 655 and two examples of spatial distributions of toluene emission sensitivities at c) 00:00 UTC on
 656 16 November 2010 and d) 00:00 UTC on 24 November 2010.

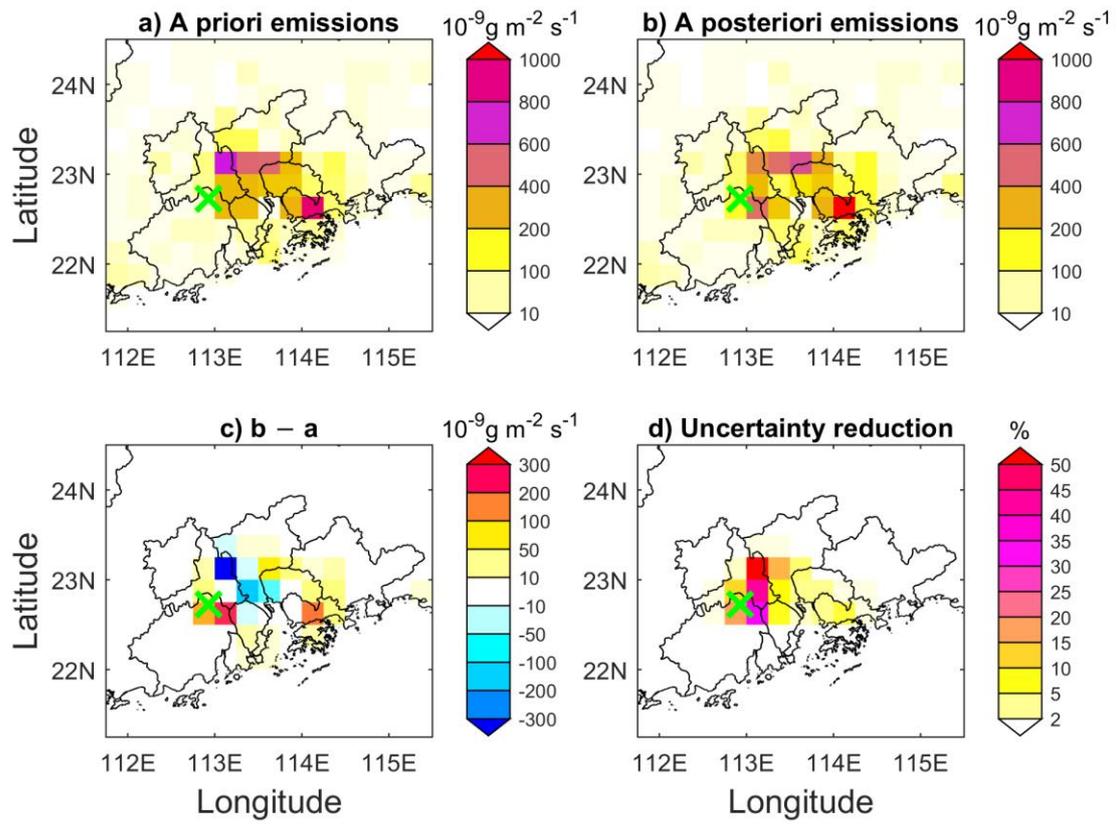


657

658 Figure 4. Maps of a) a priori benzene emissions, b) a posteriori benzene emissions, c)

659 differences between b) and a), and d) uncertainty reduction. The observation site is marked

660 with a green cross.

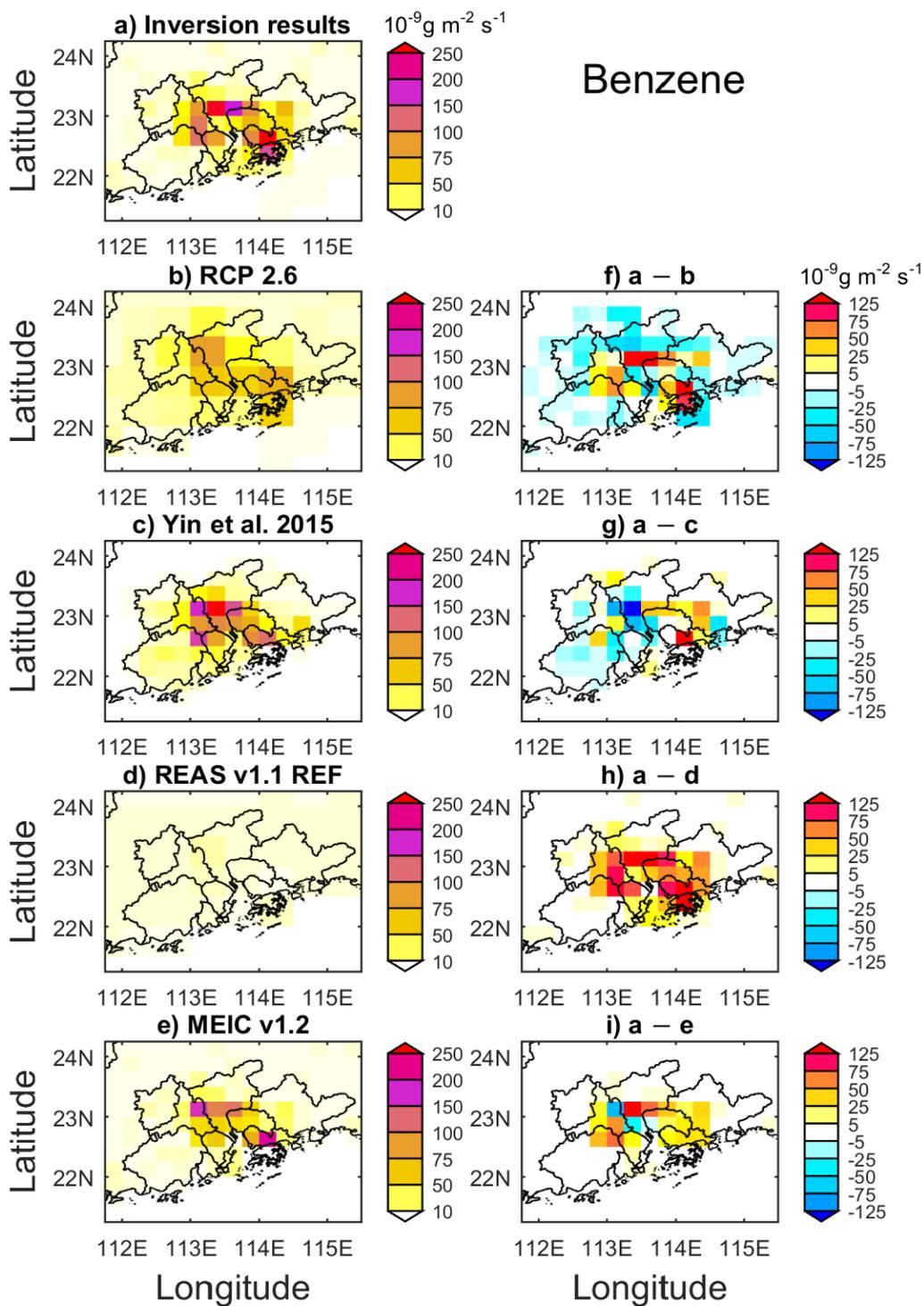


661

662 Figure 5. Maps of a) a priori toluene emissions, b) a posteriori toluene emissions, c)

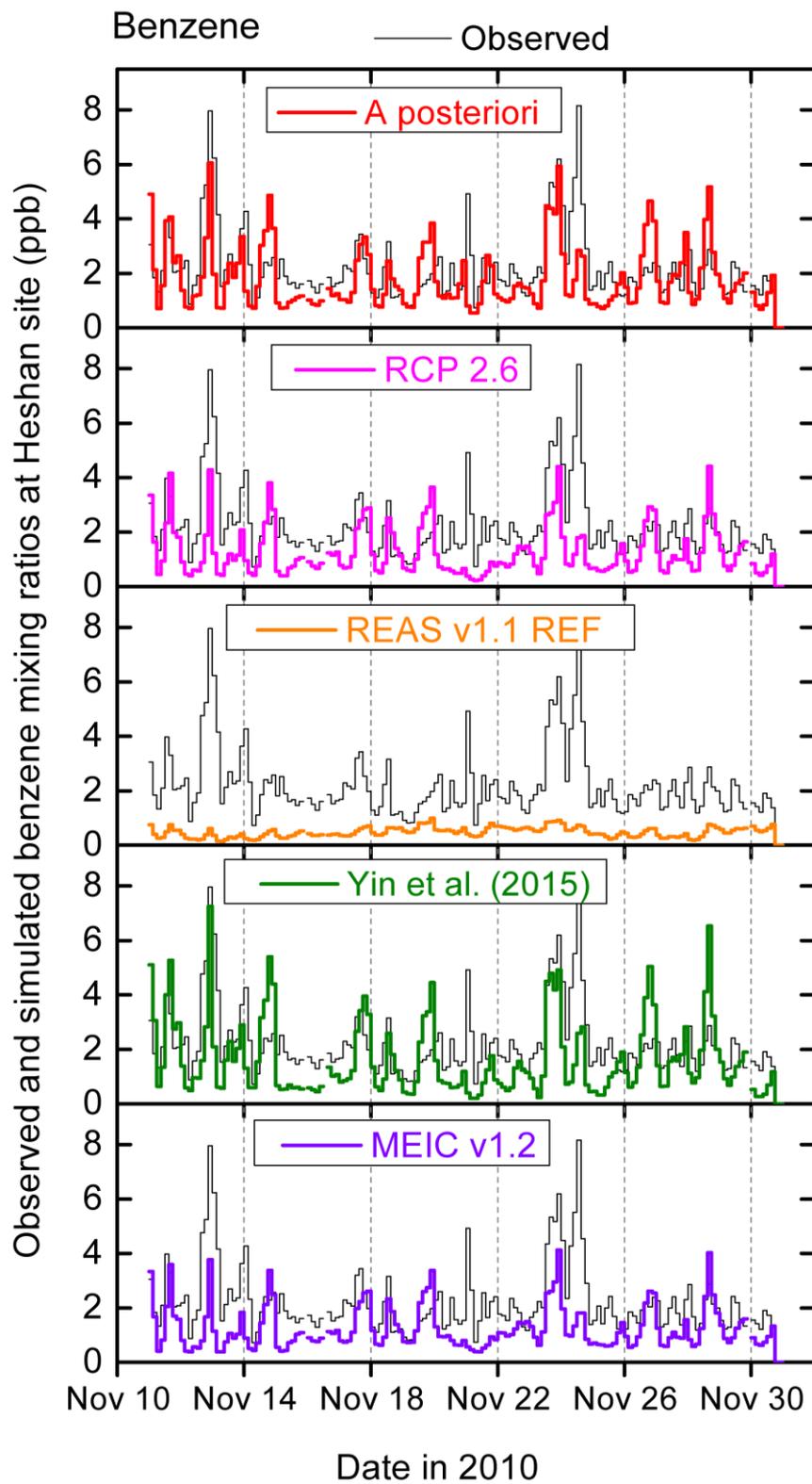
663 differences between b) and a), and d) uncertainty reduction. The Heshan observation site is

664 marked with a green cross.



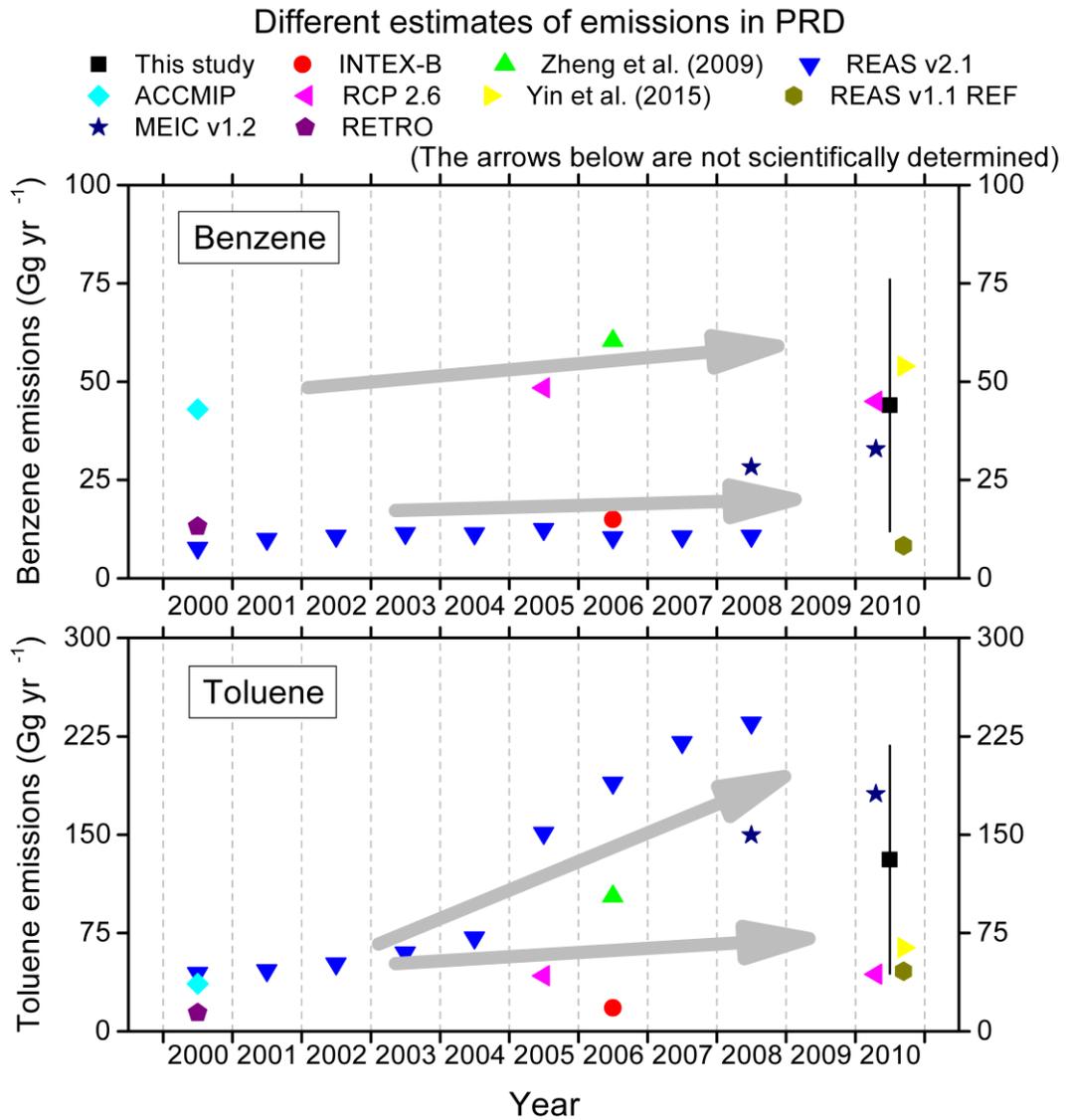
665

666 Figure 6. Maps of benzene emissions for the PRD, HK and surrounding regions from a)
 667 inversion, b) RCP 2.6, c) Yin et al. (2015), d) REAS v1.1 REF, e) MEIC v1.2, and the
 668 difference between inversion results (a) and the bottom-up inventories (b, c, d, e). Note that in
 669 c) and g) only emissions within the PRD are plotted since Yin et al. (2015) only estimated
 670 emissions within PRD, and that in e) and i) emissions within HK are not plotted since MEIC
 671 v1.2 has not estimated benzene emission in HK.



672

673 Figure 7. Time series of observed and simulated benzene mixing ratios at the Heshan site. The
 674 simulations use emission fields from inversion in this study, RCP 2.6, REAS v1.1 REF, Yin et
 675 al. (2015) and MEIC v1.2, respectively.



676

677 Figure 8. Estimates of benzene and toluene emissions in the PRD region for the period

678 2000–2010.