Dear editor,

We are very grateful to you and reviewers for the constructive comments. We respectfully submit our revised manuscript. In our revised manuscript, we addressed all comments raised by the reviewers. Our point-by-point responses to the reviewers' comments are listed below in yellow color. All changes in our manuscript are also highlighted in yellow.

This manuscript has NOT been published or is under consideration for publication elsewhere.

Thank you very much for your consideration.

Please contact myself via e-mail fangxk@mit.edu or +1-6179559144.

Sincerely,

Xuekun Fang

Center for Global Change Science, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, United States

Editor Decision: Reconsider after minor revisions (Editor review) (20 Jan 2016) by Gregory Frost

Comments to the Author:

The manuscript has improved significantly after revision. Referees #1, #3 and #4 still have a few additional comments. Comments by Referee #4 are given below. Please address all Referee comments in the next revision of the manuscript, or explain in detail why additional revisions are not needed.

———

Comments by Referee #4 on revised manuscript:

I appreciate the authors' long replies to the comments including mine (this is my second review). However, I still see the two key concerns have not been addressed. I conclude that the topic is important but the analysis method is not robust.

First, the authors did not address the potential bias in the emission estimates in the revision. They state it is out of the scope of the study. Then they should clearly state their estimates are subject to potential biases, not addressed here, so that the readers can understand the estimates with correct information. Again I argue that recent many "regional" studies evaluate meteorology, which gives the estimates more confidence. Are the authors confident of the estimates? I believe there are at least a few ground-based meteorology observation stations around the tower (if the tower itself does not have one). To the authors, it may not have been straightforward to transform the met. bias to the bias in the emission estimates although it is still doable. But at least I expected the authors to evaluate the meteorology data and discuss about it. At the minimum the authors should state this study's limitation in terms of bias analysis.

The authors did not address my previous comment about the obs-model mismatch uncertainty (covariance matrix). I see another reviewer also commented on this. The authors only say "observation–model mismatch errors were defined as RMSE value of model minus observation." This is only valid when the model does not have a mean bias compared to the observation. But this is not the case of GHG problems because the prior emission is biased (too low or too high) or imperfect (e.g., spatially). That's why we have two covariance matrices: one for the emission and another for the random error. Figuring out the random error (i.e., the RMSE in the posterior comparison) is a time-consuming process, but should be done properly as many recent regional

studies have done and developed the method. Since the treatment of random errors is “uncertain” here and I am uncertain about the posterior estimates from this study.

Referee #1

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

Figure 8. As the arrows on the figure are not determined scientifically I would like them to be removed from the figure. I do not see them as being relevant. A description of the trend is already described in the text and I do not think the arrows are necessary.

Response: Thanks for this suggestion. We have incorporated this change in our revised manuscript (also see below).

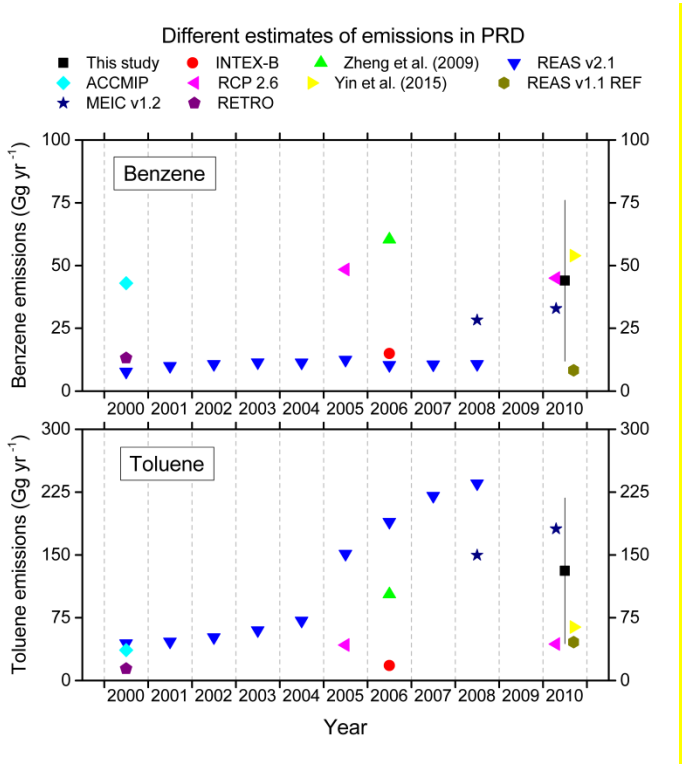

Figure 1. Estimates of benzene and toluene emissions in the PRD region for the period 2000−2010.

Referee #2

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

RE: Title: Top-down estimates of benzene and toluene emissions in Pearl River

Delta and Hong Kong, China

Author(s): X. Fang, M. Shao, A. Stohl, Q. Zhang, J. Zheng, H. Guo, C. Wang,

M. Wang, J. Ou, R.L. Thompson, and R.G. Prinn

MS No.: acp-2015-645

MS Type: Research article

Iteration: Revised Submission

Special Issue: East Asia emissions assessment (EA2)

Referee Report for a manuscript titled "Top-down estimates of benzene and toluene emissions in the Pearl River Delta and Hong Kong, China" submitted to ACP for consideration for publication.

I have reviewed all the comments, and responses to comments, and the revised manuscript. In view of the interactive public discussion on the present manuscript and taking together all the comprehensive comments provided by the four reviewers, detailed responses to comments by the authors, and the revised manuscript, I would recommend the final publication of the revised manuscript in ACP without hesitation.

Response: Thanks for the review.

Referee #3

This is my second review of this paper. The authors have addressed most of the questions that I had asked earlier but some of the responses were not included in the revised manuscript. I still have a few more concerns that need to be addressed.

My main question is whether the benzene and toluene emissions in the PRD are now better constrained or not. The authors justify the need for the top--down approach by arguing that the variation of bottom--up estimates is large. What I see in the results though, is that the uncertainty of the top--down estimate is not any smaller. A legitimate question that one might ask then, is whether these emissions are now better constrained or not. What do we learn from this study? I would like to see some more discussion devoted to the benefit of using the top--down approach in this particular case.

Response: The uncertainty of monthly total a posteriori emissions for PRD in November is indeed smaller than that of a priori emissions. Thus, uncertainty of total emissions was reduced by the inversion. Another important outcome of this study is that the spatial distributions of benzene and toluene emissions are constrained from the measurement data and inversion. Of note is that the uncertainty of "November/annual emission ratio" contributes to the uncertainty of annual total emissions presented in Table 2. If the benzene and toluene measurements span one year, the uncertainty of annual total emissions would definitely be reduced even more.

In my previous review, I asked the following question:
In terms of the prior emissions, why did the authors use the averaged emissions using MEIC v1.2 and Yin et al. (2015) for toluene for the PRD region? Did the posterior emissions change much if only MEIC v1.2 was used instead? And why was this approach not used for benzene? I feel that a similar type of sensitivity analysis using different emissions inventories might be more interesting, in addition to changing simulation length and chemical loss as currently done in the paper.

This is the response I received from the authors:
According to the comments, we tested the impact of the choice of a priori emissions. A posteriori emissions for PRD from inversions using MEIC v1.2, Yin et al. (2015) and the averaged emissions using MEIC v1.2 and Yin et al. (2015), respectively, were 13.8 Gg/month, 11.1 Gg/month and 12.0 Gg/month. Thus the difference is less than 15% which is not a big value and within the a posteriori uncertainty reported in this study. We also did the tests for benzene, and we found the difference of a posteriori emissions is about 10% using different a priori emissions. The estimated emissions for toluene by MEIC v1.2 and Yin et al. (2015) are quite different, thus we used the averaged value as the a priori emission; while the estimated emissions for benzene by MEIC v1.2 and Yin et al. (2015) are relatively close, thus we used one of them (MEIC v1.2 used in our study).

I now understand the authors' reasoning for their prior emissions but this is not reflected in the revised text. The authors are arguing that the posteriors change only by 10- -15%, "which is not a big value." If this is the case, then why don't they simply pick one of the bottom- -up estimates and state in the text that using different prior values does not change the results?
Response: Thanks for the suggestions. In our revised manuscript, we have added the explanations "Tests show that the difference between toluene a posteriori emissions for the PRD from inversions using the averaged MEIC v1.2 and Yin et al. (2015) versus the MEIC v1.2 or Yin et al. (2015), is less than 15%. The difference of benzene a posteriori emissions from inversions is about 10% using different benzene a priori emissions. Thus, the choice of priori emissions does not greatly influence the results".

I also asked the following question in the previous review:
I am also curious if the inversion is conducted for the two species together or separately. I would also like to see a better justification as to why 100 and 70% were chosen for prior emissions uncertainty for benzene and toluene, respectively. Also, it would be good to include an equation showing how the posterior uncertainty is calculated.

This is the response I received:
The inversion is conducted separately for benzene and toluene. As for the a priori uncertainty, we determined it by looking at the variations of different bottom- -up estimates for that particular species. For benzene, our priori emissions were 3.1 Gg/month, while the RCP 2.6 was 3.7 Gg/month, Yin et al. (2015) was 4.4 Gg/month, REAS v1.1 was 0.7 Gg/month. Thus, the largest deviation is about 1- -0.7/3.1=80%, and then we set the a priori uncertainty as 100%. For toluene, our priori emissions were 11.5 Gg/month, while the RCP 2.6 was 3.6 Gg/month, Yin et al. (2015) was 5.2 Gg/month, REAS v1.1 was 3.8 Gg/month and MEIC v1.2 was 14.9. Thus, the largest deviation is about 1- -3.6/11.5=69%, and then we set the a priori uncertainty as 70%.

Again, this explanation is not included in the revised manuscript. I am also puzzled why benzene uncertainty is 100% rather than 80%, as explained above. If these numbers are indeed derived from the calculations above, I do not see a reason for deviating from them. I also would like to see the equation for posterior uncertainty that I had asked earlier.
Response: Thanks for this comment. Accordingly, we did the test and found that the a posteriori emissions for the PRD from the inversion using "80%" as the benzene uncertainty are 3.85 Gg/month, which is only 2.7% smaller than the a posteriori emissions of 3.96 Gg/month from the inversion using "100%". Thus, choice of 100% and 80% does not influence the results much.

In section 2.3 in our manuscript revised this time, we have added the explanation:

"The a priori uncertainty was determined by looking at the differences among bottom-up estimates for each species. For benzene, our a priori emission was 3.1 Gg/month, compared to 3.7

Gg/month from RCP 2.6, and 4.9 Gg/month from Yin et al. (2015) for November, and 0.7 Gg/month from REAS v1.1. Thus, the largest difference with respect to our prior is 1-0.7/3.1=0.78, so we set the a priori uncertainty to be 100%. A posteriori emissions for the PRD from the inversion using 80% for the benzene uncertainty were only 2.7% smaller than those from the inversion using 100%, indicating that the choice of 80% versus 100% uncertainty does not have a significant influence on the results. For toluene, our a priori emission was 11.5 Gg/month, compared to 3.6 Gg/month from RCP 2.6, 5.6 Gg/month from Yin et al. (2015) for November, 3.8 Gg/month from REAS v1.1, and 17.4 Gg/month from MEIC v1.2. Thus, the largest deviation is 1-3.6/11.5=0.69, so we set the a priori uncertainty to be 70%."

I also have some minor comments on the figures.
For Figure 1, I suggested that the map includes the names of the cities. The authors have made changes and the maps now include city names. However, without markers pointing to the location of those cities, the readers are left unclear where these cities are located.
Response: In our revised manuscript, we have added hollow circles pointing to the location of those cities. Thanks for this suggestion. See the updated Figure 1 below.

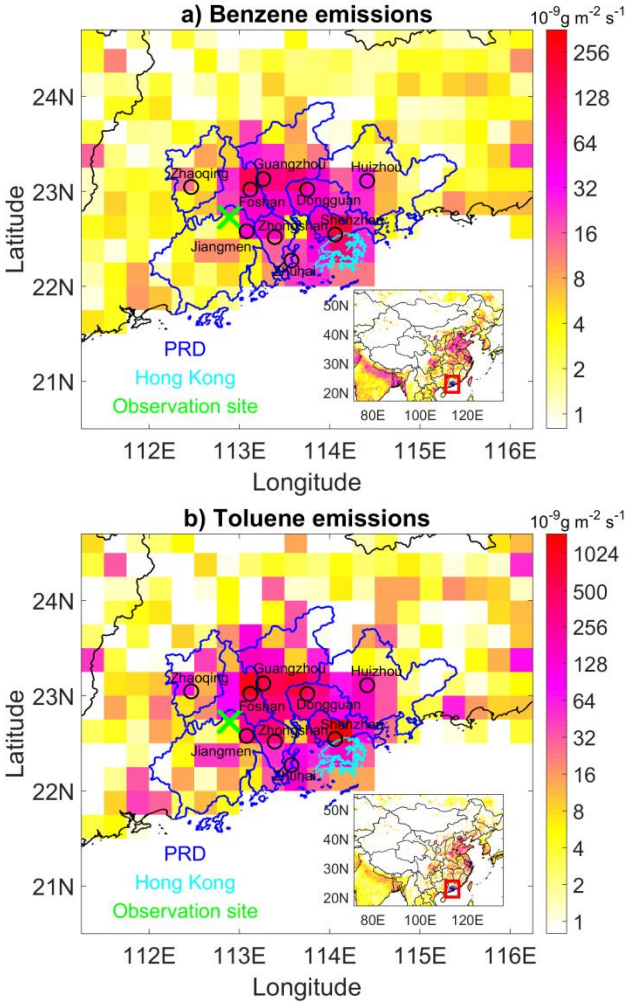

Figure 1. Map of a) benzene and b) toluene emissions from the MEIC v1.2 for China and the RCP 2.6 for outside China (inset panels), and that for the PRD and Hong Kong regions (mother panel). The PRD region is plotted with dark blue boundary lines, the Hong Kong region with cyan boundary lines. The green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

In Figure 7 all the simulated values are showing zero at the end of the time series. Is this an artifact of the plotting procedure? The same goes to Figure 3a.

Response: We have revised the Figure 3 and Figure 7. There was a minor artifact of the plotting procedure. Thank you so much! See the updated Figure 7 below.

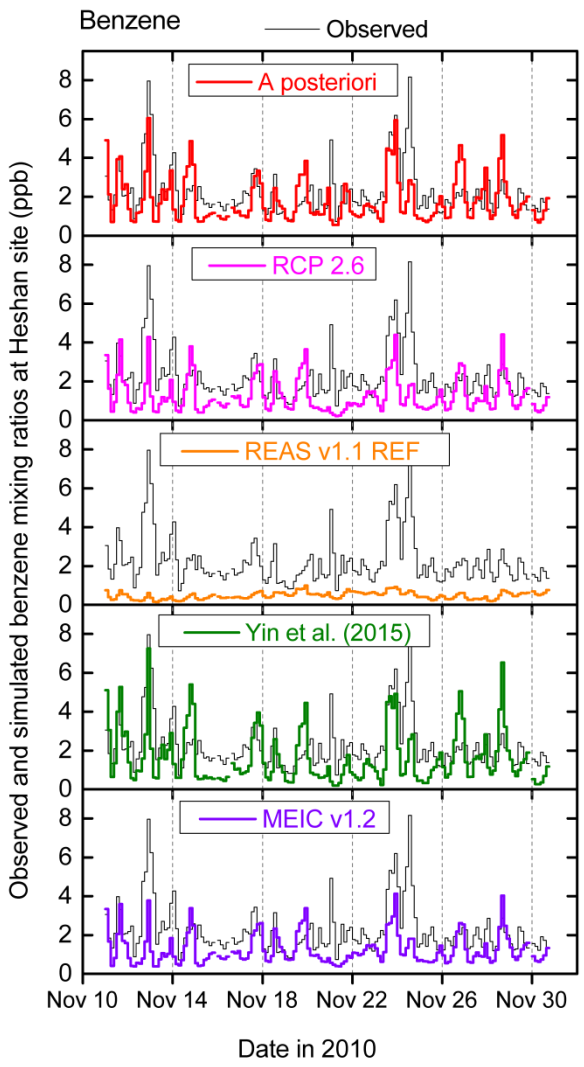

Figure 7. Time series of observed and simulated benzene mixing ratios at the Heshan site. The simulations use emission fields from inversion in this study, RCP 2.6, REAS v1.1 REF, Yin et al. (2015) and MEIC v1.2, respectively.

What do the arrows in Figure 8 represent and what does the legend "not scientifically determined" mean?

Response: We have removed these two arrows in Figure 8 in our revised manuscript, which is also suggested by another reviewer. See the updated Figure 8 below.

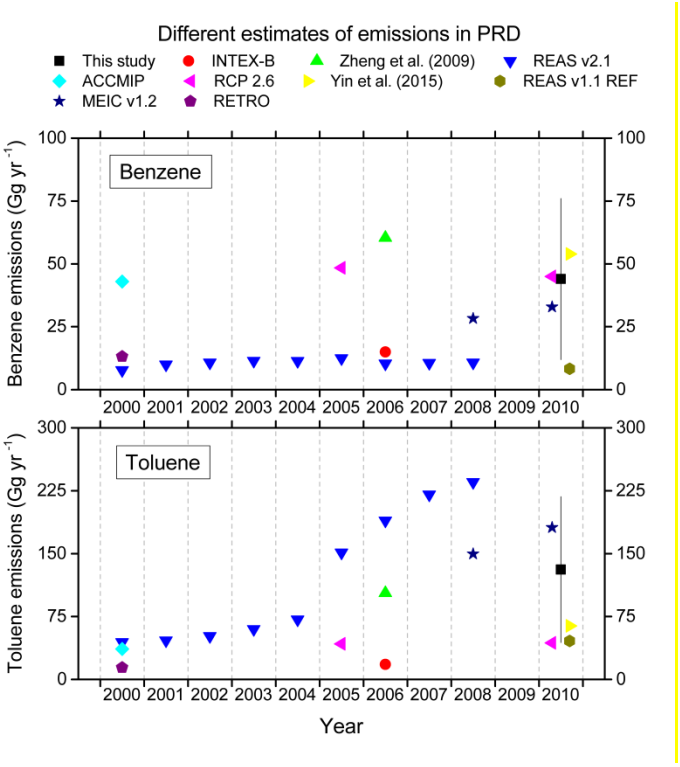

Figure 8. Estimates of benzene and toluene emissions in the PRD region for the period 2000−2010.

Referee #4

I appreciate the authors' long replies to the comments including mine (this is my second review). However, I still see the two key concerns have not been addressed. I conclude that the topic is important but the analysis method is not robust.

First, the authors did not address the potential bias in the emission estimates in the revision. They state it is out of the scope of the study. Then they should clearly state their estimates are subject to potential biases, not addressed here, so that the readers can understand the estimates with correct information. Again I argue that recent many "regional" studies evaluate meteorology, which gives the estimates more confidence. Are the authors confident of the estimates? I believe there are at least a few ground-based meteorology observation stations around the tower (if the tower itself does not have one). To the authors, it may not have been straightforward to transform the met. bias to the bias in the emission estimates although it is still doable. But at least I expected the authors to evaluate the meteorology data and discuss about it. At the minimum the authors should state this study's limitation in terms of bias analysis.

Response:

In our first response to the reviewer, we had said:

"Thus we tested the dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF). The a posteriori emissions for PRD are very close when using these two meteo. dataset, e.g. for benzene, 4.0 Gg/month from inversion using CFSR and 4.2 Gg/month from inversion using ECMWF. Although Fang et al. (2014) shows ECMWF data set performed slightly better than the NCEP CFSR dataset for the SF6 simulations in East Asia for Hateruma, Gosan and Cape Ochiishi stations, we found CFSR dataset performed slightly better than the ECMWF dataset in the benzene simulations at the Heshan site. Thus, the CFSR dataset was used in our paper".

According to the comment by the reviewer this round, we have conducted the corresponding tests. In our revised manuscript, we have added explanations in section 3.2, Table S1, Figure S1, S2 and S3. Added text in section 3.2:

"The meteorological reanalysis CFSR data were compared with measurement data from ground stations. We choose ground stations within the domain (111.45 °E−118.15 °E, 21.70 °N−27.33 °N) over which the meteorology most likely has the strongest influence on the simulation for the Heshan site. Measurement data are available at 3-hourly intervals from 34 ground stations for the

period 1 to 30 November 2010 (see the station information in Table S1 and the map of stations in Figure S1). The mean wind speed at 10 meters above ground level was 2.4 m/s in the CFSR data compared to the observed wind speed of 2.2 m/s. The mean air temperature at 2 meters above ground level was 16.1 ℃ in the CFSR data compared to the observed temperature of 17.5 ℃. Thus, the CFSR meteorological data do not differ much from the ground observations. As examples, time series of wind speed and air temperature in November 2010 at three stations are shown in Figure S2 and Figure S3, respectively."

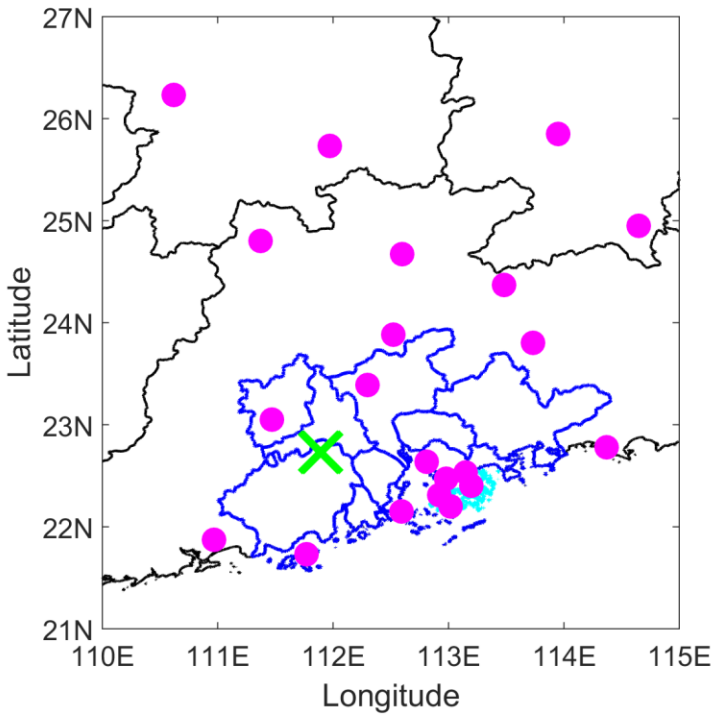

Figure S1**.** Map of meteorological measurement stations. The green cross indicates the location of the Heshan observation site (measuring benzene and toluene). The solid pink circles indicate the locations of the meteorological measurement stations. The PRD region is indicated by dark blue boundary lines and the Hong Kong region with cyan boundary lines.

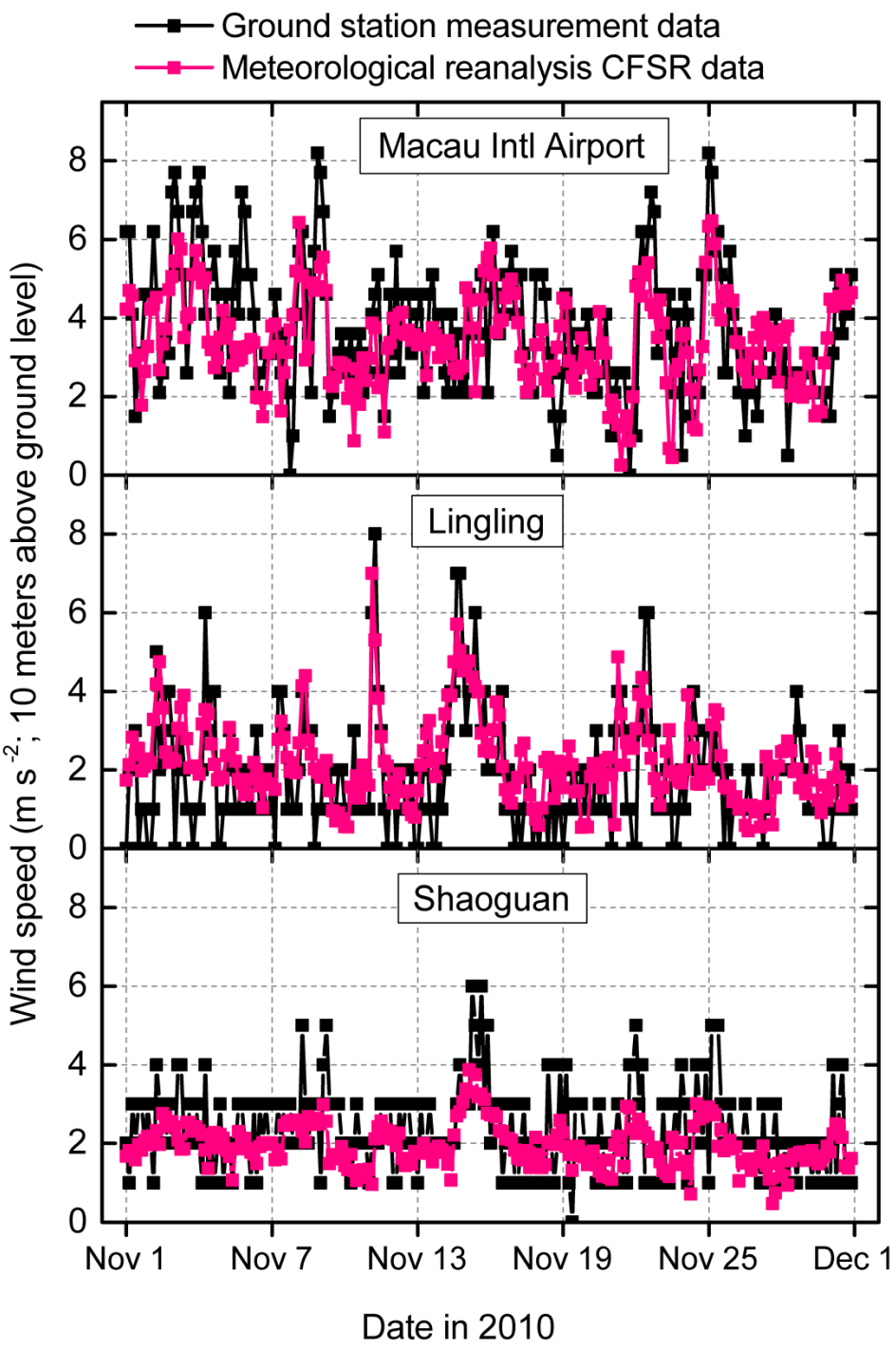

Figure S2. Wind speed at 10 meters above ground level from meteorological reanalysis CFSR data and observed at three stations (Macau Intl Airport, Lingling and Shaoguan).

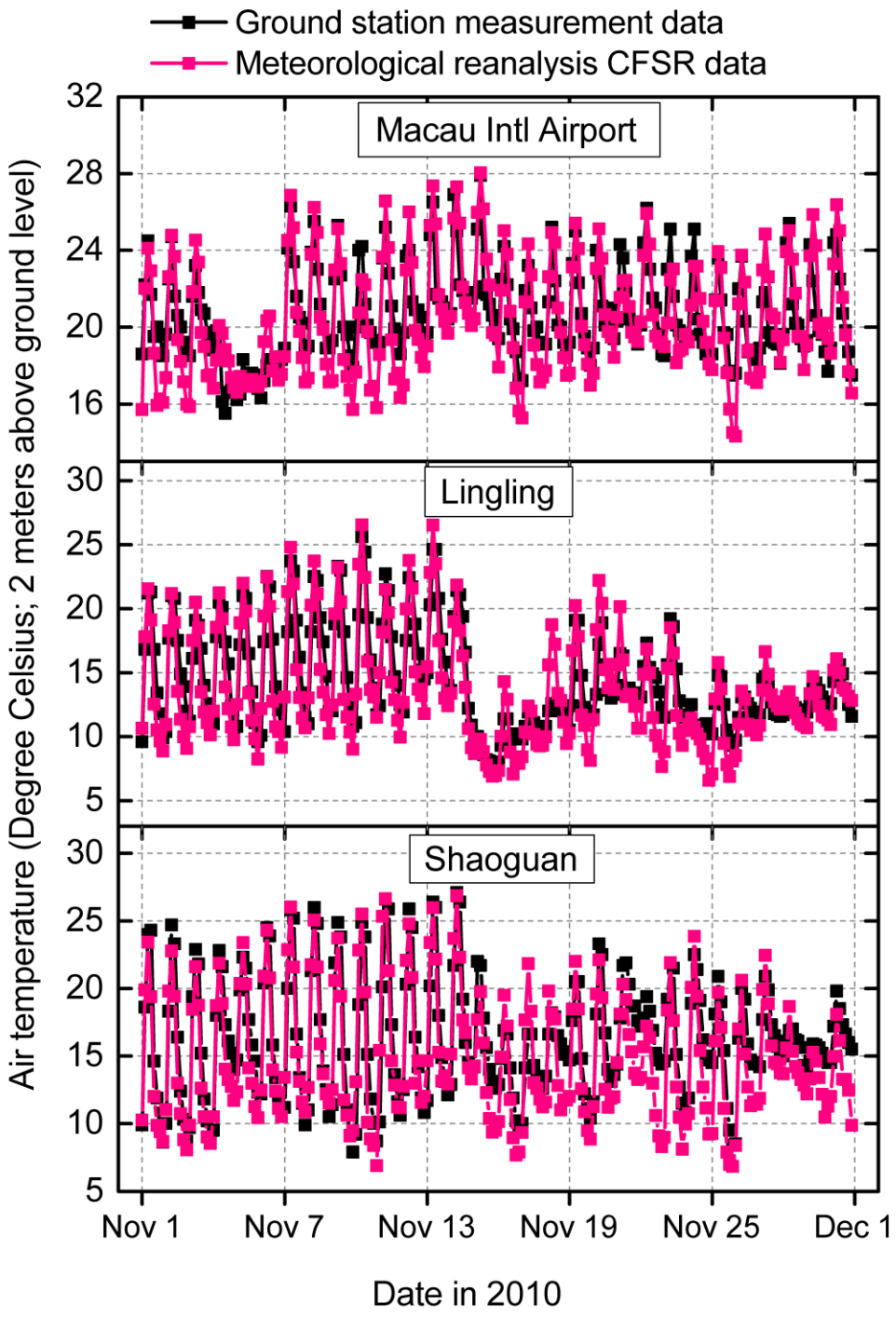

Figure S3**.** Air temperature at 2 meters above ground level from meteorological reanalysis CFSR data and observed at three stations (Macau Intl Airport, Lingling and Shaoguan).

The authors did not address my previous comment about the obs-model mismatch uncertainty (covariance matrix). I see another reviewer also commented on this. The authors only say "observation–model mismatch errors were defined as RMSE value of model minus observation." This is only valid when the model does not have a mean bias compared to the observation. But this is not the case of GHG problems because the prior emission is biased (too low or too high) or imperfect (e.g., spatially). That's why we have two covariance matrices: one for the emission and another for the random error. Figuring out the random error (i.e., the RMSE in the posterior comparison) is a time-consuming process, but should be done properly as many recent regional studies have done and developed the method. Since the treatment of random errors is "uncertain" here and I am uncertain about the posterior estimates from this study.

Response: This comment is of course right. However, it is not possible to cleanly separate bias as coming from emissions and random errors from model errors. Emissions can be biased, yes, but they can also be wrongly distributed, so this is causing both bias and random errors. And the same is true for pure model errors. E.g., systematic errors in mixing height would cause a bias, whereas errors in transport direction may be more of a random nature.

As for this study, we found that a toluene posteriori emissions for PRD from inversions using MEIC v1.2, Yin et al. (2015) and the averaged emissions using MEIC v1.2 and Yin et al. (2015), respectively, were 13.8 Gg/month, 11.1 Gg/month and 12.0 Gg/month. Thus the difference is less than 15%. The difference of benzene a posteriori emissions from inversions is about 10% using different benzene a priori emissions. Thus, that "prior emission is biased (too low or too high) or imperfect (e.g., spatially)" only affected the a posteriori emissions (and observation–model mismatch errors) to a small extent in this study.