Reviewer #1

0. *The authors provide a detailed analysis to constrain BC emissions from Jiangsu (China) using observations from two stations. They found BC emissions are significantly overestimated in the bottom-up inventories, which has important implications. However, I have some major concerns about the representation of their stations to the whole region, and the inversion methodology. I recommend the paper for publication after consideration of the points below.*

Response and revisions:

We appreciate the reviewer's remarks on the importance of the work. Regarding the limitations pointed out by the reviewer, we have improved the manuscript accordingly. The spatial representativeness of the two sites in the multiple regression model has been clearly described (please see our response to Q3). Case 2 in which observation data at only one site (NJU) were used has been further re-analysed to avoid confusion to the inversion methodology (please see our response to Q4).

1. *Abstract Lines 28-29, please confirm the same BC concentrations (i.e. 3.4 ug/m^3) at both sites. In addition, Lines 39-40 say: "the simulated annual mean was elevated to 2.6". I assume it is elevated from 3.4 to 2.6?*

Response and revisions:

We thank the reviewer's comment and reminder. We confirmed that the annual mean simulations of BC were 3.44 and 3.39 ug/m^3 at NJU and PAES, respectively. When the constrained emissions were applied, the annual mean concentration was simulated to decrease from 3.39 to 2.57 ug/m^3 at PAES, and it was indicated **in Table 2 in the revised manuscript**. We corrected the sentence **in line 41 in the revised manuscript**: "At PAES, in particular, the simulated annual mean declined to 2.6 μg/m^3 and the annual normalized mean error (NME) decreased from 72.0% to

57.6%."

2. *Line 257-258 Are 5 days long enough to minimize the influences of initial conditions? I checked the methodology of other studies and found much longer initialization periods. For example, 3 months in Wang et al. (2013) and Mao et al. (2015).*

Response and revisions:

We thank and agree with the reviewer's comment. Some studies that applied GEOS-Chem or WRF-Chem to constrain BC emissions at larger spatial scale often chose several months as spin up to minimize the influence of initial conditions (Fu et al., 2013 and studies mentioned by the reviewer). For WRF-CMAQ model, in contrast, more studies used several days as initialization periods, for example, 5 days in Chang et al. (2018) and Tran et al. (2018), and 7 days in Ran et al. (2016). The period in this study is expected to be sufficient to minimize the influence of initial condition.

3. *Table 2 As shown with the annual mean result:*

** NJU, the a priori is 0.4 lower than obs, and is reduced by 0.6 in the inversion. The a posteriori is 1.0 lower than obs.*

** PAES, the a priori is 0.9 higher than obs, and is reduced by 0.8 in the inversion. The a posteriori is 0.1 higher than obs.*

It seems that the inversion simply moves the bias from PAES to NJU by reducing the total emissions, suggesting the inversion system is dominated by PAES. Considering the inconsistency between NJU and PAES, it is hard to say whether the conclusion is reliable to provide a good representation for the whole region.

Response and revisions:

We appreciate the reviewer's important comment. As can been seen **in Table 2 and Figures 3 and 4 in the revised manuscript**, application of JS-posterior effectively reduced the large biases between simulations and observations for all seasons at PAES and for January and April at NJU, suggested by the reduced NMEs. In particular, most of the overestimations in peak concentrations were corrected at the both sites. We mentioned **in lines 489-492, 497-499 and 508 in the revised manuscript**. It should be also acknowledged that NMEs for July and October and the annual average of NME were slightly enhanced at NJU. Limitation of the multiple regression model was thus indicated that overestimation and underestimation in concentrations at different sites could hardly be corrected simultaneously without further improvement in spatial distribution of emissions, and we mentioned in details **in lines 511-516 in the revised manuscript**.

To improve the method and to quantify the effect of spatial representation of observation sites on top-down estimate, we provided Case 3 in which observation data at PAES and NJU were applied to constrain emissions from Nanjing and Suzhou-Wuxi-Changzhou-Zhenjiang city cluster, respectively, **in Section 4.1 in the revised manuscript**. The best CTM performance was obtained in Case 3, implying that inclusion of more measurement data with their spatial representativeness considered could improve the top-down method. Given the limited BC observation data in the area, therefore, more measurements with better spatiotemporal coverage were recommended for constraining BC emissions effectively, as mentioned **in lines 47-52 in the revised manuscript**.

4. *Section 4.1 As shown in Table 2, the model simulation (2.38) is already lower than obs (2.69) in April at NJU. When only NJU data is used, how could the inversion keep reducing the emissions with scaling factors, 0.42, 0.95 and 0.65? Theoretically, an inversion system should minimize the discrepancy between model and obs rather than magnifying it.*

Response and revisions:

We thank the reviewer's important comment. As can been seen **in Table 2 in the revised manuscript**, the monthly mean of simulated BC concentrations at NJU with JS-prior was 2.38 ug/m^3 for all periods in April, smaller than the observed 2.69 ug/m^3. For Case 2 in which only NJU data were applied, the scaling factors for industry, residential and transportation emissions were obtained at 0.42, 0.95, and 0.65, respectively, implying a further reduction in BC emissions. The main reason is that the data for the whole April were not fully used due to necessary data screening in the multiple regression model. We acknowledge that the data screening process was not clearly stated in the original manuscript. Before applying in the multiple regression model, we excluded the periods following the criterions: the periods lack of observation data, those for which the contribution of each emission sector (power generation, industry, residential sources and transportation) was simulated to be smaller than zero through the brute-force method, and those for which the sum of contributions of all the four sectors was larger than 100% with CTM. The data screening helped to reduce the uncertainty of CTM in the multiple regression model. We added the description of data screening **in lines 240-245 in the revised manuscript**. The number of data after screening in Case 2 was 48% of data in all periods (most data screening was due to lack of observations, accounting for 38%). We divided all the data points in April in Case 2 into two groups: those included in the multiple regression model and those excluded from the model, and analyzed the modeling performances for both groups separately. As can be seen in Table R1, the simulated concentration for periods included in the multiple regression model (2.71 μg/m^3) was larger than the observation (2.56 μg/m^3) when JS-prior was applied, different from the case without data screening (i.e., data in all periods were included). The emissions could then be reduced when the observation was applied in the constraining. As a result, application of the top-down estimate in Case 2 effectively reduced the NME for the period included in the model from 34.01% to 21.09%, and the simulated average concentration was closer to the observation. At the same time, the constrained emissions did not increase the bias for periods excluded from the multiple regression model. It thus indicated that the underestimation for periods excluded from the multiple regression model could result largely from factors other than emissions like meteorology. We added the analysis **in lines 547-559 in the revised manuscript** and included Table R1 **as Table S8 in the revised supplement**.

Table R1. Statistical indicators for observed and simulated BC concentrations for all periods, those included in the multiple regression model, and those excluded from the model in JS-prior and Case 2 for April 2015 at NJU.

Site	Parameter	JS-prior: All period	JS-prior: Included	JS-prior: Excluded	Case 2: All period	Case 2: Included	Case 2: Excluded
NJU	Average SIM (μg/m^3)	2.38	2.71	2.08	2.27	2.42	2.08
	Average OBS (μg/m^3)	2.69	2.56	2.99	2.69	2.56	2.99
	NMB (%)	-16.02	5.90	-56.48	-21.59	-5.32	-56.63
	NME (%)	42.31	34.01	57.62	32.47	21.09	57.61

5. *Table 4 More information is needed in the caption. It is really difficult to follow the discussion to distinguish the Cases (B, 1, 2, 3, 4, 5) and Cases (6, 7).*

Response and revisions:

We thank the reviewer's reminder. As suggested by the reviewer, we added the introduction of different cases **in Table 3 in the revised manuscript**.

References

Fu, T. M., Cao, J. J., Zhang, X. Y., Lee, S. C., Zhang, Q., Han, Y. M., Qu, W. J., Han, Z., Zhang, R., Wang, Y. X., Chen, D., and Henze, D. K.: Carbonaceous aerosols in China: top-down constraints on primary sources and estimation of secondary contribution, Atmospheric Chemistry and Physics, 12, 2725-2746, 10.5194/acp-12-2725-2012, 2012.

Chang, X., Wang, S., Zhao, B., Cai, S., and Hao, J.: Assessment of inter-city transport of particulate matter in the Beijing-Tianjin-Hebei region, Atmospheric Chemistry and Physics, 18, 4843-4858, 10.5194/acp-18-4843-2018, 2018.

Trang, T., Huy, T., Mansfield, M., Lyman, S., and Crosman, E.: Four dimensional data assimilation (FDDA) impacts on WRF performance in simulating inversion layer structure and distributions of CMAQ-simulated winter ozone concentrations in Uintah Basin, Atmospheric Environment, 177, 75-92, 10.1016/j.atmosenv.2018.01.012, 2018.

Ran, L., Pleim, J., Gilliam, R., Binkowski, F. S., Hogrefe, C., and Band, L.: Improved meteorology from an updated WRF/CMAQ modeling system with MODIS vegetation and albedo, Journal of Geophysical Research-Atmospheres, 121, 2393-2415, 10.1002/2015jd024406, 2016.