Dear anonymous referee #1,

We sincerely appreciate your time on review RC2 and your insightful comments on the important points, including careless mistakes and typos. Thanks to your review, our manuscript has been substantially improved, especially for methodology of emission flux optimization. We have carefully addressed all your comments in the revised version.

The region name “Delhi NCR” includes surrounding districts of Delhi. Thus, to be precise, the category name for CPCB and the region name in the maps are changed from “Delhi NCR” to “Delhi”, throughout the revised manuscript.

According to referee #2’s comment, Figs. 5 and 6 are removed. Figs. 7 to 12 are changed to Figs. 5 to 10, respectively, in the revised manuscript.

Point-by-point responses to your comments are written in blue in this letter.

With best regards,
Mizuo Kajino and Kentaro Ishijima

---

[1] Authors have used chemical transport model and observations from dense network of low cost sensor to derive emission fluxes at resolution of 12 hours over a region known for crop residue burning (CRB) to find their impact on downwind mega city Delhi. The analysis is very important for the policy makers and air pollution researchers who try to see impact of such activity on air quality, health and cost-benefit analysis. Overall paper is well written, I suggest few minor correction which I believe will make manuscript better and are necessary.
Thank you very much for your evaluation.

[2] Authors use emissions from inventories to calculate non-CRB contribution. One can see that simulation and observation for non-CRB period compare well from Fig. 3. However, authors have used selected stations to prepare those plots. Authors do mention criteria for selection and list of stations selected vis-a-vis not selected in main text and supplementary but I think more details are necessary to replicate their finding by others as well as to understand subjectivity vis-a-vis objectivity in the criteria for

selection of stations. As authors indicate in the absence of these selection, the comparison may not be as good, and in that case CRB fluxes which are interpreted based on difference between observed and simulated concentration will also be different.

Thank you for your comments. There are no objective thresholds of data selection. First, we simply excluded unreliable data, which have sudden gaps or zero drift. For $PM_{2.5}$, most data can be used, more than 70% of station data (78 stations out of 107 stations) are included for the analysis, and there are not many differences between $PM_{2.5}$ for selected sites and all sites in both CPCB and CUPI-G (please compare left and right panels in Fig. S2). Thus, we presume that the optimized emissions may not be so much different either using all stations or using only selected stations. However, because the purpose of our study is to estimate the CRB emission flux, we were especially careful on the selection of data during the non CRB period, October 15-28. If there are big gaps in the simulated and observed data during the non CRB period, our optimization will be misled to fill the gap by altering CRB emissions, which may not be realistic. We inserted the following statement in the 5th paragraph of Sect. 2.3 as:

"In other words, if there are big gaps in the simulated and observed data during the non CRB period, our optimization would be misled to fill the gaps by altering CRB emissions."

Also, I visited not all but some of CPCB stations in Delhi, where wind speed data are very low, or solar radiation data are always much lower than the clear sky conditions. Most stations are well situated in open places, but some are situated in locations surrounded by tall buildings. Those stations are suitable for human exposure studies, but not for the NHM-Chem study, because our model does not simulate air quality inside the urban canopy. This is also the reason why we excluded such stations for the emission optimization.

Still, it is not our motivation to show which stations are suitable or not suitable for the emission optimization, we didn't show time series of meteorological and air quality data of individual stations but only showed the list of stations we used as in Table S1.

re

Thank you very much for your understanding.

[3] Emission flux estimates in top down approach are highly sensitive to simulation of vertical distribution of species which in turn is highly sensitive to simulation of boundary layer dynamics. Authors should show either using previously published studies for their model or from this study, how good were boundary layer simulation over South Asia and preferably over North India.

Thank you for your suggestion. As shown in Fig. S9, we compared the simulated vertical profiles of CO and O_3 against observation data made by IAGOS, the commercial aircraft monitoring campaign. Based on the comparison, we made a new section 3.6 to discuss uncertainty in top-down emission flux due to boundary layer simulation. Please see Sect. 3.6, “Uncertainties in our top-down emission optimization approach”.

[4] Top down approach for flux estimate is also sensitive error in observation. Why authors have not mentioned errors in observations explicitly, they imply temporal standard deviation of 3 hours exceeds error in observation. This may be true for random error but if the observation had systematic error then it will still affect the calculation of emission flux. Authors should provide information on error in observation and discussion on how they would affect emission flux estimate.

Honestly speaking, this is my first paper to conduct emission optimization, and so I did not realize showing observation errors are critically important.

This observational data uncertainty generally consists of observational errors and spatial representativeness errors. In the analysis of the global to subcontinental scale carbon dioxide flux (Maki et al., 2010), observational errors are clearly smaller compared to spatial representativeness errors. Therefore, they mainly estimated spatial representativeness errors for each observation site based on the difference from the fitting curve of the monthly average carbon dioxide observations. In the analysis of dust emissions in the Gobi Desert (Maki et al., 2011), they used PM_{10} observations over a wide range from China to Japan (considering observations above a certain value as dust). Although they assumed that there were no significant differences in instrument errors spatially, they considered spatial representativeness errors to vary by location and set the observational data uncertainty to 10% of the observed PM_{10} concentrations. This paper, since we are using an observation network in a relatively small area in northern India, we believe it is reasonable to treat observation uncertainty (spatial representativeness errors and observational errors) uniformly.

We elaborated on this in the 4th paragraph of Sect. 2.3 by inviting an expert of inverse

modeling Dr. Takashi Maki to the coauthor. Uncertainty in emission flux estimates are discussed in Sect. 3.6.

References:

Maki, T., Ikegami, M., Fujita, T., Hirahara, T., Yamada, K., Mori, K., Takeuchi, A., Tsutsumi, Y., Suda, K. and Conway, T. J.: New technique to analyse global distributions of CO2 concentrations and fluxes from non-processed observation data, Tellus, 62B, 797-809, doi:10.1111/j.1600-0889.2010.00488.x, 2010.

Maki, T., Tanaka, T. Y., Sekiyama, T. T. and Mikami, M.: The impact of ground-based observations on the inverse technique of aeolian dust aerosol, SOLA, 7A, 21-24, doi:10.2151/sola.7A-006, 2011.

[5] The discussion regarding emission optimization is not very clear -- at least to me.

The point-by-point responses are written as below from [5a] to [5c]. We also had similar comments by Referee 2. Please refer to the RC1 comments #6 together with our corresponding replies.

[5a] Based on the discussion that follows in and after section 2.3, x_m simply appears to be ratio of observed concentration to simulated concentration or a multiplier when multiplied to simulated concentration excluding non-CRB concentration, it matches simulated concentration to observed concentration. I might be wrong to interpret x_m in this manner but If I am correct about x_m as ratio, then authors should explain

x_m is the scaling parameter for each CRB emission flux tag. S_m and S_0 are the $PM_{2.5}$ simulation data with and without the CRB emission tags, so that S_m minus S_0 is considered to be $PM_{2.5}$ from CRB emission tag (this is so-called the brute-force concept with a zeroed-out test to derive source contributions, as explained in our reply to your Comment #7).

[5b] Are optimized emission fluxes are simply scaled up emission fluxes based on that ratio?

That's why, yes.

[5c] How would second term in the equation 1 would imply importance of apriori in the cost function?

This can be altered by setting the upper and lower limits (u_m, l_m). As the limits are broader, the relative importance of the second term is smaller so that the algorithm

tried to minimize deviations in the first term. Anyway, relative importance of the first and second terms in the cost function depends on the relative magnitudes of σ_n and (u_m, l_m). The sensitivities of those parameters to emission optimization were tested by altering (u_m, l_m) as (0.5, 2), (0.1, 10), and (0.01, 100), as presented in the new section, Sect. 3.6.

[6] Authors write that the denominator in the second term of Eq 1 (u_m, l_m) were taken as (2.0 and 0.5) and in an multistep process calculations were repeated until x_m values were between 0.009 and 1.001. Later-on in the result section authors show values of x_m ranging from 0 to 69 with several of them more than 1. The text is not clear enough to describe how the values of x_m get these high values. Authors should clarify x_m calculations and how they can have such high value in spite of their set criteria.

x_m was obtained after multiple steps of optimization. Suppose x_m obtained after the i-th step is defined as $x_{m,i}$, optimization was repeated until k-th step when $x_{m,n}$ get enough close to 1, and $x_m = x_{m,1}\, x_{m,2}\, x_{m,3} \ldots x_{m,k}$. When $x_{m,i}$ (i = 1, 2, …) are (2.0, 2.0, 2.0, …, 1.5, 1.05, 1.005, 1.001, 1.0001, 1.0001, …), x_m larger than 2.0 is obtained.

We inserted the following description in the 3rd paragraph of Sect. 2.3:

"Let xm,i be the value of xm obtained in the i-th step. This multi-step optimization was repeated until 0.009 < xm,i < 1.001 was obtained for all m valuesat the k-th step, so that the final xm was obtained as:

$$x_m = \prod_{i=1}^{k} x_{m,i}. \qquad (2)$$"

[7] Page 19 lines 5-10: The description of calculation of anthropogenic contribution is not clear. Why the 20% emission reduction, why not 25% or 40%? Which emissions -- All the emissions or the emission that are newly calculated over and above non emissions? What does multiply by 5 implies? Is the simulated concentrations multiplied by 5? Why would difference between control and reduced/multiplied concentrations would imply anthropogenic contribution?

This is so-called a brute-force method to estimate source-receptor relationship. If zeroed-out sensitivity test (anthropogenic NO_x, SO_x, NH_x, and VOC concentrations are set all zero) was conducted, the oxidant fields such as O_3 and OH to produce secondary components of $PM_{2.5}$ (such as SO_4^{2-}, NH_4^+, and SO_4^{2-}) are significantly altered (i.e.,

significantly underestimated). Thus 20% reduction run was conducted, so as not to change the oxidant fields much, and 5 times the deviations between the control run and the 20% reduction run are defined as the source contributions (anthropogenic contributions) (20% times 5 = 100%). For the sensitivity runs of biomass burning sources, zeroed-out runs were conducted because it may not alter the oxidant fields significantly, because the contributions from other sources such as anthropogenic sources may be more to the oxidant fields.

Yes, in fact, the brute-force method with a 25% reduction run then 4 times the deviation (Lin et al., 2008) or a 10% reduction then 10 times the deviation (Bartinicki, 1999), existed. As the reduction rate is smaller, the artifact due to altering oxidant fields get smaller, but the numerical errors become larger (for example, the deviation between 1% reduction run and control run is so small that the S/N ratio is small. Then 100 times the deviation may contain larger noise in their source-receptor estimation). Thus, we reached a consensus that 20% is a suitable value, after our extensive activities for the source-receptor relationship studies in East Asia conducted by the Long-range Transboundary Air Pollutants in Northeast Asia (LTP project) among China, Korea, and Japan (e.g., NIER, 2010, Kajino et al., 2013).

We modified the description of the brute-force method (2nd paragraph of Sect. 3.2) as follows:

“The anthropogenic contribution was calculated using the brute-force method, derived from the difference between the control run and the 20% emission reduction run multiplied by 5 to reduce the effect of nonlinearity in chemical reactions (Kajino et al., 2013): The brute-force method with a zeroed-out simulation (no anthropogenic NO_x and NMVOC) affects oxidant concentrations substantially and thus chemical production rates of secondary $PM_{2.5}$, which deteriorates the calculation of source contribution estimations.”

Refs.:

Bartinicki, J., 1999. Computing Source-receptor Matrices with the EMEP Eulerian Acid Deposition Model. EMEP/MSC-W, Note 5/99. Norwegian Meteorological Institute, Oslo, Norway, p. 37.

Kajino, M., Sato, K., Inomata, Y. and Ueda, H., 2013. Source-receptor relationships of nitrate in Norheast Asia and influence of sea salt on the long-range transport of

nitrate, Atmos. Environ., 79, 67-78, doi:10.1016/j.atmosenv.2013.06.024.
Lin, M., Oki, T., Bengtsson, M., Kanae, S., Holloway, T., Streets, D.G., 2008. Long-range transport of acidifying substances in East Asia — Part II source-receptor relationships. Atmospheric Environment 42, 5956-5967.
National Institute of Environmental Research (NIER), 2010. The 10th Year's Joint Research on Long-range Transboundary Air Pollutants in Northeast Asia. Annual Report of LTP Project 2009. NIER, Korea.

Minor Issues:

[8] Throughout the text and in abstract and in conclusion authors mention their model as NHM-Chem. However, in the methodology section, authors mention that only the Chem part is used from NHM-Chem whereas meteorology is simulated using WRF. Since, the WRF model is also available with its own Chem version known as WRF-Chem. It is imperative that authors should briefly describe why instead of WRF's Chem they chose the different Chem model and instead of NHM's meteorology, they chose WRF's meteorology. Also they need to come-up with better nomanclature then NHM-Chem since WRF-Chem would be misunderstood as default WRF-Chem model and NHM-Chem would be misunderstood as default NHM meteorology.

I agree with you. The offline coupling between WRF and the CTM part of NHM-Chem was referred to as NHM(WRF)-Chem, as defined in the 1st paragraph of Sect. 2.1. The reason why we selected NHM(WRF)-Chem is because it was performed slightly better than NHM-Chem, in terms of surface $PM_{2.5}$. The reason why we used NHM-Chem and did not use WRF-Chem is because NHM-Chem is our model. The reason why we selected NHM(WRF)-Chem was explained in the last sentence of 1st paragraph of Sect. 2.1 as "because NHM(WRF)-Chem exhibited a slightly superior performance compared to NHM-Chem when evaluated against observed time series of surface $PM_{2.5}$ in North India" We modified the nomenclature elsewhere in the entire manuscript.

[9] Page 3 Line 29: Authors write "polar orbiting satellites travel once during the day ..". Most polar orbiting satellites including Terra and Aqua travel twice over a place in 24 hours, once during daytime and once during nighttime. Satellites use thermal channels to find fire hot-spot which in principle should work better in night. Should authors clarify this point in the text and explain why there are no nighttime data for the fire hotspots?

Thank you for the correction. The statement was totally wrong. Burned area (BA) product is obtained only in the daytime, but the polar orbiting satellites travel twice per day. The statement was based on my presumption that the biomass burning emissions

are primarily derived using BA (such as GFED and FINN), but they are also derived by other parameters such as fire radiative power (FRP), that are efficiently measured during the night. In fact, GFAS emission we mainly used in this study was based FRP and their daily emission is the combination of daytime and nighttime emissions (Kaiser et al., 2012). Thus, the sentence was changed to “polar-orbiting satellites travel twice during the day and night, usually around noon and midnight”, in the 3rd paragraph of Sect. 1. The similar statement in the same paragraph, “Nevertheless, nighttime observational data are unavailable even though some farmers ignite fires after sunset (Liu et al., 2020)” was deleted. Similarly, there was another sentence in the 2nd paragraph of Sect. 4: “Satellite-derived fire emissions are measured in the daytime on daily basis by polar-orbiting satellites; thus, emissions at nighttime have been previously unavailable.”, which was changed to “Diurnal variations of satellite-derived fire emissions have been previously unavailable”.

[10] Page 9 Eq 1: There is an extra plus sign in the equation. Also, the symbol σ_0 should be σ_n if separate sigma was used for each observational data. Or otherwise mention how the σ_0 is calculated.

Thank you for finding the error. I deleted the extra plus sign. As for σ_0, it is not sigma_zero, but sigma_O (O of Observation). Anyway, since it was misleading, I changed it from sigma_zero to sigma_n. Relating to your comment #4, how sigma_O was calculated was explained in the 4th paragraph of Sect. 2.3.

[11] Fig 7e: GFAS is written as GFSS

Thank you for finding the typo. I modified it.