Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC1
https://doi.org/10.5194/acp-2022-161-RC1, 2022

Comment on acp-2022-161

Anonymous Referee #1

Referee comment on "Improving NO_x emission estimates in Beijing using network observations and a perturbed emissions ensemble" by Le Yuan et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-161-RC1, 2022

The manuscript introduces an ensemble estimate of anthropogenic emissions (including a source sector separation) by comparing simplified ensemble simulations with NO, NO2, and O3 in situ observations. The investigatino used emission data calculated for the year 2013 to analyse emission for the year 2016 in the Beijing area. The ensemble was set up by perturbing emission inventory data for NOx and CO by spatiotemporal uniform factors (parameters). The parameters have been selected by a standard multidimensional sampling technique and the valid range for each parameter was given by expert elicitation, which has been further improved. The evaluation of the ensemble members is mainly based on the mean square error (MSE) and its decomposition. The author includes a very detailed discussion on the results, which I appreciated. The manuscript is in overall good quality and it fits to ACP topics. I support publication in the ACP journal once minor revisions have been addressed.

General comments:

- Title: I am not sure what the novelty in the perturbed emission ensemble is. Thus, I suggest to remove “novel” from the title.

- The evaluation of the ensemble members is mainly based on the MSE. However, I was wondering if the ensemble shows a sign-change in the bias of NO2 concentrations, which would further support the estimation of the optimal emission data.

- I understand that the choice of spatiotemporal uniform emission perturbations suggests an evaluation of averaged concentrations over all stations. An evaluation at single stations was initiated by e. g. Fig. 5, but I would have expected a more detailed investigation of

the ensemble members in different regions. Potentially, the results allow also for a spatially heterogeneous emission correction.

- Although VOC emissions (and background concentrations) are included in the model, the impact on these emissions and potential uncertainties is not addressed adequately. Especially in high NOx concentrations, the O3 concentrations depend highly on the available VOC. The manuscript only considers NOx emissions as main source of uncertainty. A discussion on the impact of this choice is appreciated.

- A discussion on the representativity of observation sites (especially urban and traffic) is required. Is the model resolution sufficient to be compared the traffic measurement stations?

- the wording initial PEE and optimized PEE is somehow misleading. Only as I have finished section 2 I have understood that this approach is not a data assimilation or inversion method. Maybe "adjusted PEE" instead of "optimized PEE" would be clearer. In the context of observations, "optimized" always feels like there is some optimization method applied, which is certainly not the case in this manuscript.

- To the simulation setup: It is not really clear, which simulations have been done. There is a base run with additional 140 member ensemble with perturbed NOx emissions. However, the simulation episode should be state here explicitly (am I right that the full year 2016 was simulated?), also the model resolution (horizontal and vertical) is missing. A link to the discussion section, where the limitations introduced by model simplifications is discussed, would be good. It would have been easier for the understanding of the results that only the optimized PEE is used for the simulations.

Minor comments:

- line 52: A discussion on the local anthropongic and biogenic share of NOx (and esp. VOC) emissions would be appreciated. Are biogenic emissions in this regions (especially in summer) negligible compared to the anthropogenic emissions?

- Line 56-58: Citations for the different action plans required

-line 64: ...method, which… (add a comma)

-line 65/66: I guess you are talking about the amount of studies investigating emission data, please be more precise: which data? What is the large amount of the data? How can data solve the time-lag issue?

- line 68: It may be worth elaborate on emission uncertainties and their impact

- It would be worth elaborate more about the pros and cons of the different methods you are summarizing in the introduction. Why are you proposing the new method, what is the strength of your method compared to the other methods?

-line 83 – 86: This statement is not only valid for satellite data. Insufficient chemistry always influences the model results and, thus, the analysis.

-line 106 -111: I feel like this is too much detail for the manuscript. Is it necessary to follow the study to know the accuracy of the measurement instruments?

- line 116-118: Also, is this information necessary for the manuscript? I don't feel so. Are the low-cost sensors influenced by a systematic error (bias) that may have an influence on the comparison?

- Table 1: Please include the night time definition for the initial and optimized ensemble in the caption. Also, the caption states night time fractions are in %, but values show ratios, please revise.

-Line 178-179: Please include the number of experts that contributed to the poll.

-line 208/209: Although in the simulation CO is treated as inert species, in general, it does affect the NOx concentrations via O3 chemistry (Gaubert et al., 2020, Correcting model biases of CO in East Asia: impact on oxidant distributions during KORUS-AQ, Atmos. Chem. Phys., 20, 14617–14647, https://doi.org/10.5194/acp-20-14617-2020). How does this assumption influence the emission estimation? What is the benefit from adding CO perturbations to the parameter field if CO is treated as inert? Both emitted species (NOx, CO) could as well be separately optimized.

- line 210-213: Errors may not be the total emissions but the spatial distribution of the emissions, which is not addressed with the 14 parameter setup of the analysis. A

discussion is appreciated on how this influences the results (especially locally close to emissions sources).

- line 233: “high resolution” is a rather open statement. Emissions data are available at 3 km resolution. It would be good to add the exact resolution (horizontally and vertically).

- line 240-242: How is the city defined in the model? Are buildings represented as domain boundaries? If not, how is the local street canyon flow represented (e. g. channeling, overflow, small scale vorticies)? I expect from the manuscript that the model is not a LES model?

-line 325: change “length” to “number”

- line 332 – 335: I feel the reasoning in this statement is not correct. The fact that there is a large spread in the MSE depending on total NOx emissions does not necessarily mean that the emissions are higher and overestimated to a larger extent. It is rather the distribution of the MSE depending on total NOx emissions that lead to this conclusion (rapid increase in MSE for lower total NOx emissions and constant increase of MSE with increasing total NOx emissions).

- line 335 – 338: I don’t really understand this reasoning. Please rephrase.

- discussion on Fig. 4: A discussion on the fact that some stations show almost no sensitivity to the underlying NOx emissions is appreciated. How about the impact of other emissions (e. g. VOC, CO) on the O3 concentration. As stated in line 349, NOx does not seem to be the only limiting factor for O3 concentrations. A discussion on further improvements would be nice.

- line 384 – 385: I don’t really understand this. There is a change in MSE of O3 with changing total NOx emissions in Fig. 4b. Here you state, that this is associated with the mMSE. Maybe you can give examples how the mMSE is influenced (e. g. via changes in VOC concentrations by altering the NOx emissions?). Also in the discussion on Fig. 6, there is a dependence of the mMSE on the NOx emissions visible, which needs to be related to a lower correlation coefficient. Thus, in my opinion the decomposition of the MSE is mainly influenced by the changing correlation, which shifts the contribution to either the second or third term of Eq. 2 if the bias is negligible. I would like to see this discussed further.

- line 413-414: this is only valid for uniform perturbations across the domain. Please add this information to the sentence.

- line 421-422: Please add reference(s).

- line 422: An introduction to Fig. 7 is missing

- line 618-620: Comparing Fig. 8a and 8b the impact of changing the input is almost as large as the variety within the top 5 % PEE members. Thus, I feel the change of input concentrations would have also a large impact on the uncertainty of emission estimates, potentially leading to larger uncertainties in the emissions. Please add a discussion on this impact.