Dominant role of emission reduction in PM2.5 air quality improvement in Beijing during 2013-2017: a model-based decomposition analysis
by Jing Cheng et al.

General Comments

The main purpose of this paper is the discussion of the reasons of recent rapid PM2.5 decrease of Beijing, mainly by using the meteorological and emission sensitivities by chemical transport model. Based on several sensitivity simulations, this paper made a decomposition analysis framework to evaluate the impacts of local control policies, surrounding emission reductions and the meteorological changes on PM2.5 abatement in Beijing during 2013-2017 and 2016-2017. This paper made the most of important sensitivity analysis and explains the relative contribution of meteorology (12%), local emission (65%) and regional emission (23%) for the reduction of PM2.5 between 2017 and 2013. The results of detailed sensitivity analyses are useful for the understanding of PM2.5 reduction and environmental policy. Their results are very much reasonable and important. However it is difficult to find scientific uniqueness of this paper. The current version of manuscript could be published as the Technical Note. It is necessary to add the more scientific discussion in order to be accepted as a research paper.

Specific Comments

1) The paper mainly discussed the decreases of annual average PM2.5 concentration. However as shown in Figure 4, the authors are specifying the decrease of chemical compositions (i.e., SO_2, NO_x), it is necessary to compare the aerosol chemical composition changes too. The under estimation of sulfate in winter are reported in many previous papers, the comparison of sulfate between observation and model results are very much necessary.

2) Another important point is the OC (we found that model result are usually under estimated), so the detailed examinations of model reproductively of OC are necessary in order to discuss the emission change in VOC.

3) The authors show the very good agreement of PM2.5 reduction between observation and model results. It is usually difficult to have such good agreement. It is necessary to discuss the detailed reasons why model results are so good agreement from the view point of emission inventory, WRF model, model horizontal resolution, and CMAQ itself performance.

4) Zero-out emission sensitivity is used in this study by assuming the linearity. PM2.5 formation is usually nonlinear, so it is necessary why the authors are using zero-out method.

5) Section 2.3: Although we can follow by Zheng et al. (2018) to understand the emission inventory, it should be noted what is the major "updated". Especially, the inventory is named as "MEIC"; however, we noticed the emission amounts described in Zheng et al. (2018) and Li, M. (2017) are different. Does this mean "updated"? Taking into account the importance of emission inventory, more careful explanations and descriptions are needed here for the traceability of this kind of study.

6) Section 2.4.1: Model descriptions are insufficient. Model calculation was conducted after only 10 days spin-up. How does the 3 years WRF simulation perform? i.e., with or without FDDA? Need more detailed descriptions.

7) Does CMAQ study include the Asian dust? If yes, need discussion of model accuracy and problems.

8) Emission sensitivity study was conducted by each emission sector base. It is necessary to include the discussion of the accuracy (or error bars) of

emission estimate for each sector base.

9) Section 2.5 is unclear. The description of model sensitivity has to be rewrite. Equations of (1) – (3) are unclear. I think Equations of (2) and (3) are not NORMALIZED RESUTS.

10) Section 3.2 Without the model evaluation from 2013 to 2017, it is hard to discuss the source attribution results by model. From the current manuscript, we can only find time-series on 2013, 2016, and 2017 for PM2.5 and statistic evaluation only on 2017. The model evaluation is inadequate at the current manuscript. For example, Figure 3 and Figure S3 can be presented in the same form for model. See also my minor comments 4) for O3 performance.

11) The results shown here should be interpreted in depth. On 2013, especially the peak on January, model sometimes overestimated observed PM2.5. However, model simulated same level or sometimes underestimated high concentrations during winter on 2016 and 2017. Actually, the model negative-bias is larger in 2017 compared to 2013 (Table 2). Therefore, the source attribution results based on scenario analysis adopted in this study can be strongly reflected by emission variation rather than the observed facts.

12) Figures 5 and 9 are unfriendly. It need more detailed explanations or improve the presentation of figures. The basic information is same as Figure 10, so it might be better to modify Figures 5 and 9 into Figure 10 format.

13) Section 3.3: The discussion in this section needs relevant references.

Minor Comments

1) It have been reported that WRF should be updated version 3.9.0.1 or later for the upgraded NCEP dataset after 12 UTC, 19 July 2017.
   http://www2.mmm.ucar.edu/wrf/users/wpsv3.9/known-prob-3.9.html
   The used version is 3.8, but how did the authors solve this problem?

2) What is the horizontal grid resolution in second domain?

3) What is the lateral boundary condition for first domain? It will be taken from global chemical transport model, but did the global model consider year-to-year emission variation? If not, how can we conclude the importance of global-scale impacts on the air quality in China?

4) Considering the current modeling application over East Asia, the vertical 14 layers from surface to 10 km is too rough. First, the first layer is approximately 50m, but it will be usual to set 20-30m. The current model configuration is doubled thickness on first layer, and the representativeness as surface layer is ambiguous. Second, the upper model height is only 10 km. In my best knowledge, CMAQ does not support the top boundary condition. Therefore, this modeling system might have some problem to the treatment of stratospheric O3, and subsequently, to the model performance on surface level. The statistic analysis for O3 (Table S2) seems to be out of range compared to the suggested model performance (Emery et al. 2017). Furthermore, this reproducibility for O3 might lead to inaccuracy of other air pollutants.
   Reference) Emery et al. (2017, JA&WMA)
   https://www.tandfonline.com/doi/full/10.1080/10962247.2016.1265027

5) The emission inventory for Beijing is not taken from MEIC, but there is no reference and needs elevant information here. What was the difference between two inventories? Did the authors have specific reason to replace the emissions only for Beijing instead of MEIC?

6) What was the biomass burning inventory used in this study? I did not find the description.

7) Section 2.4.2: In Table S2, I can only find the statistic for the year of 2017. Why did other years not shown? This section should be clearly separated into the description and discussion. Most of this section should be moved to subsection 3.2 or 3.1.

8) Table 2: Does the parenthesis on rightmost column indicate observation? It should be clearly described.

9) Typo: Section 3.4.2 should be 3.4.3