Response to reviewer 3

We appreciate reviewer's thoughtful comments and suggestions, which are greatly helpful for us to improve our manuscript. The manuscript has been revised to accommodate the reviewer's comments and suggestions.

General comment As it stands, the paper needs substantial improvements to fully understand the rationales and the key findings of this work. The scientific contribution of the paper with respect to previous studies should be better emphasized. While the overall goal is to test whether accounting for emission uncertainties in BEC modelling leads to more accurate predictions of $PM_{2.5}$, two or three discussions points should be identified in Introduction and clearly addressed in a separate section. The Introduction needs to provide more detailed and accurate background on BEC modelling and emission uncertainty quantification including a consolidated review of literature on these aspects. What are the other approaches to model BEC? What are the limitations of the NMC method? An alternative approach is to rely on ensemble of analyses using random perturbations of emission. Some authors have also investigated the possibility to incorporate the emission variables in the control vector and retrieve them using data assimilation. The use of references is not accurate enough. Some definitions (e.g. data assimilation, categories of uncertainties in CTM simulations) are not accurate enough or incorrect (e.g. background error covariance matrix). The structure of Section 2 (data and methods) needs to be improved. Some aspects of the methodology are not clear or incomplete to be fully understood and be re-productible. The paper does not provide a clear understanding of the conventional NMC versus the new NMC method. The paragraph on page 6 and Fig 4. do not provide enough details to understand how the NMC method was implemented. What are the differences between Met1 and Met2? What are the actual differences between the three NMC implementations? How the emission uncertainties have been incorporated in the NMC? Part of the answers are in the text, but it requires too many efforts for the reader to find out.

I suggest having separate results and discussion section. I provide here some questions that could help to build the discussion:

- what are the limitations of the proposed NMC approach ?
- The variability in emission fluxes is accounted for using two inventories, is it enough to represent the spatiotemporal uncertainties in emission fluxes ? Would not be preferable to use spatial and temporal perturbations of emission fluxes based on a priori probability distribution function per type of emission.

- What are the benefits and limitations of assimilating $PM_{2.5}$ compared to satellite AOD? Is $PM_{2.5}$ observation more directly related to the model variable? What is the impact on the application of this method to over regions or historical periods which may not have $PM_{2.5}$ observations.

Response Thank you for your comments. We tried our best to answer your comments and suggestions. Please, refer to our responses to your comments shown below.

Comment first paragraph: I suggest to give some key references for direct and indirect radiative effects of aerosols.

Response We have added some key references in the sentence. Please, check out pp. 2:41–pp. 2:43.

Comment line 44: why limitation to GEO satellites? there are a lot of atmospheric composition observations derived from LEO satellites (e.g: MODIS/TERRA, AQUA for AOD)

Response We discussed the advantages and disadvantages of LEO and GEO satellite-retrieved products. In particular, the fraction of satellite-derived AOD available has been low due to the high cloud fractions in East Asia. In contrast, $PM_{2.5}$ has been monitored by many ground stations, regardless of the presence of clouds. This is one of the main reasons that we chose $PM_{2.5}$ as control variable. Please, refer to the discussions at pp. 2:47–pp. 2:48 and pp. 3: 82–pp. 3:90 in the revised manuscript.

Comment "The inaccuracy of CTM simulations has been associated with uncertainties in emissions of primary air pollutants and meteorological fields as well as omissions of photochemical reactions occurring in chemical mechanisms (Han et al., 2013, 2015; Kim et al., 2017a; Song et al., 2012).": The authors provide here specific examples of source of uncertainties impacting CTM simulations. I suggest to give the main categories of source of uncertainties: drivers and forcing variables (emission inventories, meteorological fields for offline simulation, land cover : : :), model structure (e.g. photochemical reactions, more or less realistic representation of atmospheric chemistry: : :), model parameters.

Response We have tried to improve the paragraph. Please, see pp. 2: 48–pp. 2:53.

Comment line 48-57: The definition of data assimilation is well known and there is no need to repeat it here. This paragraph is somehow vague. Data assimilation in NWP context has mainly two goals: provide the best estimate of initial condition and provide an estimate of the uncertainties associated with the initial state, that could include emission uncertainties.

Response Here, we intended to discuss the roles of uncertainties in observations and models in the DA process, and then discuss the importance of model uncertainty, which is related to the main purpose of our study. We thought that this description could help readers to better understand our research purpose. We have therefore decided to keep this paragraph in the manuscript with several corrections to avoid ambiguity. Please, see pp. 2:57–pp. 2:62.

Comment The list of references is sometimes too long, Authors should select two or three references for a given statement and try to be more accurate

Response We have tried to select key references directly related to the discussions. Please, check out pp. 2:55.

Comment line 58: The background error covariance matrix is a key component of both sequential and variational methods

Response Thank you for your point! We have modified the sentence. Please, check out pp. 2:63–pp. 2:64.

Comment observation errors: Observation errors include gross error (e.g. cloud detection for AOD satellite), measurement errors, representativity, observation operator uncertainties

- the background error is different than the model error.

The model error is the departure between the true atmospheric state at time k and the model prediction. The model error is represented by a dedicated model error covariance matrix. In strong constraint 4dvar the model is assumed to be perfect and the model error is neglected. In Kalman filter, the model error covariance matrix needs to be specified.

The background error is the error associated to the short-term forecast. In some assimilation system the background can be a climatology and not an output of the model. Part of the error in the background is due to the model but it can also be generated by other sources of uncertainties such as emission inventories. When the BEC is flow-dependent or in sequential assimilation scheme, the BEC is updated at each cycle and thus it is also influenced by the observation error used in the previous analysis.

I suggest to give here the main role of the BEC in terms of information spreading, information smoothing and balance properties.

Response Based upon your suggestions, we have tried to modify the sentences. Please, refer to pp. 2:65–pp. 3:66. Also, in order to avoid the confusion, we have changed the term of "model error" to "background error" throughout the entire manuscript.

Comment BEC modelling: Most methods derive the statistics of the background error from the departure between the observation and the background (expel: Hollingsworth, A., and P. Lonnberg, 1986), or using a surrogate quantity whose error statistics can be a good approximation of the unknown background errors (such as NMC). More recent approaches rely on ensembles of analyses. I suggest to provide here more background information on existing approaches to model the BEC including their advantages and limitations.

Response We focused on the BEC modeling method used in this study, so that we decided to keep the paragraph unchanged here.

Comment line 69 : "Among the greatest sources of errors in CTM simulations (e.g., Elbern et al., 2000; Wu et al., 2008) are the uncertainties of emission inventories". This sentence should be moved to the paragraph listing the sources of uncertainties affecting CTM simulations.

Response We moved this sentence into the paragraph discussing the source of uncertainties affecting CTM simulations. Please, see pp. 2:52–pp. 2:53.

Comment Paragraph on methods to account for uncertainties in emission inventories: Not enough background is given on this central aspect of this paper. There are studies that have attempted to include the emission fluxes in the control vector and estimate them using ensemble data assimilation approach.

Response We have added the sentence: "On the other hand, another recent study developed an ensemble data assimilation approach in order to constrain not only initial conditions of carbon monoxide (CO), but also surface CO emission fluxes (i.e., dual assimilation-inversion) (Barré et al., 2019). It was noted in this study that the error distributions in surface emissions play an important role in the performance of CO predictions.". Please, see pp. 3:78–pp. 3:81.

Comment need to clarify that $PM_{2.5}$ is the output variable targeted in this work. What is the rationale for choosing PM2.5 instead of AOD?

Response Again, main focus (objective) of this study was/is to improve short-term (24-hour) $PM_{2.5}$ predictions in South Korea. This was the main reason that we chose $PM_{2.5}$ as control variable, instead of AOD. We already mentioned the availability issue of AOD in East Asia. Moreover, if we selected the satellite-retrieved AOD as control variable for the purpose of short-term predictions in South Korea, we definitely need one more key technique to convert AOD into surface $PM_{2.5}$. Although this technique has been developed at authors' lab using several machine learning-based techniques, it will bring more uncertainties into our framework. All the reasons mentioned above are why we chose $PM_{2.5}$ as control variable, not AOD in this study. In order to clarify this, we have added/modified the sentences. Please, refer to the paragraph at pp. 3:82–pp. 3:90.

Comment The structure of Section 2 needs substantial revision.

Section 2.1 includes several aspects that should be included in separate subsections, I suggest the following structure:

a/Study site and observations: the second paragraph of Section 2.1 concerns the 'description of the study it is not clear how the observations used for validation and data assimilation were selected. What is the vertical footprint of the measured PM2.5? How does it compare with the modelled value?

Response We significantly modified and re-structured the paragraphs. please, refer to the pp. pp. 9:262–pp. 9:268 and pp. 11:333–pp. 11:340. We did not use observations having vertical footprint in this study. Such types of observations have not been available in East Asia.

Comment b1/model description: a short description of the CMAQ CTM is missing. Providing the version of the aerosol and chemistry module is not enough, key references are missing. I suggest to give the main characteristics of aerosol and chemistry schemes (e.g. number of species and reactions for the chemistry, list of aerosol species for the aerosol scheme) along the main characteristics of the atmospheric transport model:(which type of advection scheme is used)

b2/model configuration: it should address time and spatial resolution, coupling between WRF and CMAQ, temporal period, location, output variables.

Response We added a new table providing information on the options of WRF and CMAQ model simulations. Table will provide some details on both meteorological model and CTM simulation. Please, check out pp. 4:121–pp. 4:129 as well as Tables S1 and S2.

Comment Emission datasets: Can you justify why these two data sets have been selected for this work .

Response These two emission data sets were the most reliable in East Asia. Please, check out the details at pp. 5:161–pp. 6:167.

Comment Data assimilation and BEC modelling: Since BEC modelling is a central aspect of the methodology, a dedicated section should explain how it is parametrized and how the NMC method is used to estimate the BEC parameters in this work.

Response We seriously modified the paragraphs. Please, see pp. 7:215–pp. 7:223.

Comment Experiment design: This section should include the statements given from line 178 to 199. A table summarizing all the experiments/simulation could be helpful. The various cases of implementation of the NMC method need some clarifications.

Response We have modified Fig. 4. and added a table summarizing the experiment design. Please, check out Fig. 4. and Appendix A.

Comment Validation methodology: a clear definition of PM2.5 is missing: what is the vertical footprint of $PM_{2.5}$? What are the differences between the modelled and the observed $PM_{2.5}$?

Response Please, refer to pp. 2:38. As mentioned previously, we did not use observations related to the vertical footprint of $PM_{2.5}$ (such data was not available in East Asia!). Simulated $PM_{2.5}$ was calculated using simulated aerosol species and aerosol size fraction, and observed $PM_{2.5}$ was measured values. Please, see pp. 6:191–pp. 7:197 for the definition of modeled $PM_{2.5}$ used in this study.

Comment Some aspects of the methodology are not clear or not accurate enough
Section 2.1
line 123: which conserving method? please give a reference

Response We have used a program named "Spatial Allocator" for flux-conserving interpolation. The "Spatial Allocator" distributed by Community Modeling and Analysis System (CMAS) has been used for emission manipulations. Please, refer to the paragraphs at pp. 5:137–pp. 5:138.

Comment line 125-130: this belongs to Results and not to Methodology section. "The differences in South Korea are relatively small, except for CO in the MIX emission inventory.": Are you talking about the differences between the two databases? I do not understand "except for CO in the MIX inventory"

Response We have modified this paragraph. Please, see pp. 5:154–pp. 5:156.

Comment line 130-136 on the use of MEGAN. Why are you using LAI from MODIS and GVF from VIIRS? Are these variables required to drive MEGAN? There is a possible inconsistency between LAI from MODIS and GVF from VIIRS? Can you comment on it?

Response It would be the best option to use both LAI and GVF retrieved from the same satellite. However, both GVF from MODIS and LAI from VIIRS were not available, when our study began. MEGAN v3 requires LAIv which is the ratio of LAI to GVF. We used 8-day averaged LAI and GVF, so that we assumed that the inconsistency may not be a major issue.

Comment the description of the cost function (l14-151) is a bit confusing. x is the control vector. X and xb contain the same variables (both are of the same size). x is the analysis and xb is the background. Are you also assimilating other variables which drive the chemistry or the transport model ? line 166 redundancy with Introduction line 172, not accurate definition of S: S represents the background error and its diagonal components are the standard deviation of the error of the background. What are the differences between the measured and the simulated $PM_{2.5}$?

Response We did not assimilate other variables. We have removed the redundant sentence you pointed out. We have modified the definition of S. Please, see pp. 6:212–pp. 6:213. The difference between the measured and the simulated $PM_{2.5}$ was mentioned previously.

Comment line 211, eq 4: How a and b values have been chosen? line 213: replace 'second criterion' by 'Eq 4 criterion' line 213-215: this belong to the data assimilation section/BEC description. Some parts of the methodology are lacking such as the selection of observations for data assimilation versus validation.

Response We have added the references related to those values. Please, see pp. 9:260. We replace 'second criterion' by 'criterion of Eq 4'. We have also relocated the sentence describing BEC. Please, see pp. 6:187–pp. 6:189. For the selection of observations, please refer to pp. 9:262–pp. 9:268.

Comment Section 3 should be dedicated to the presentation of the results. A separate section should address the discussions points. I shortly review the results but further review of them should be done if the manuscript is considered for publication.

- line 221-222: "To estimate the influence of the two : : :" : this belongs to the previous section

Response In order to further clarify this point, we have changed the paragraph. Please, check out pp. 9:270–pp. 9:275.

Comment why incorporating emission uncertainties in BEC should influence the vertical distribution for $PM_{2.5}$?

Response Uncertainties in emissions are certainly related to the uncertainty in surface $PM_{2.5}$. However, because atmospheric species are transported vertically due to turbulent convection processes, the vertical distributions of $PM_{2.5}$ can also be influenced by the uncertainties in emissions.

Comment l248: "In the DA process, the horizontal length scale determines $PM_{2.5}$ increases in the horizontal spread of analysis" I do not understand this statement. The horizontal length scale refers to the horizontal correlation of $PM_{2.5}$

Response We have modified the original sentence like following: "the horizontal length scale determines the horizontal spread of $PM_{2.5}$ increments around observation locations (Descombe et al., 2016)".

Comment line 258: "The characteristics of the vertical and horizontal length scales, however, have not been fully explained in this study, thus requiring future": The authors should further discuss this aspect and provide possible explanations.

Response We intended to explain that further studies are necessary and that the studies should be carried out in the future. Regarding this point, please, check out pp. 12:391–pp. 12:399.

Comment Section 3.2 first Paragraph: This belongs to methodology and should be described in the experiment design section.

Response We have moved this paragraph into the section of experiment design.

Comment The last two paragraph should be developed in a separate discussion Section. Part of it should also be used as background information in Introduction. I can see also some redundant ideas from the Kumar et al, 2019 paper.

Response We have changed the section name from "Conclusion" to "Summary and Conclusions". Here, we intended to mention some limitations of current work.

Comment Lack of references in several part of the papers

Response We have tried to provide more relevant references in the revised papers.

Comment The use of a large number of acronyms makes the reading somehow very difficult.

Response We thought that we have used the acronyms commonly used in the atmospheric studies. To avoid the difficulty, we decided to add a list of acronyms related to the experiment design. Please, refer to Appendix A.

Comment Result description needs to be improved, some sentences are confusing.

the style is frequently not appropriate with a lot of uncertain and long sentences: for example "We found that the new approach exhibited a tendency to generate substantially increased standard deviations" , a tendency to generate : : : , I suggest using more direct sentences.

Response We have tried to modify several indirect sentences throughout the manuscript. Please, check out pp. 12:384 – pp. 12:385.