Responses to the comments of Reviewer #1:

We are truly grateful to yours' positive comments and thoughtful suggestions. Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. Based on these comments and suggestions, we have studied comments carefully and have made correction which we hope meet with approval. All changes made to the text are marked in red color. Below you will find our point-by-point responses to the reviewers' comments/ questions:

Specific Comments:

1. *L257-261. It seemed the vertical resolution of Lidar data is much finer than that of the model. Can you add a few words on the uncertainty of the Lidar AEC data? And also clarify how many data were filtered out? Thus the readers may get some more ideas why the complex data preprocess is necessary here.*

Response:

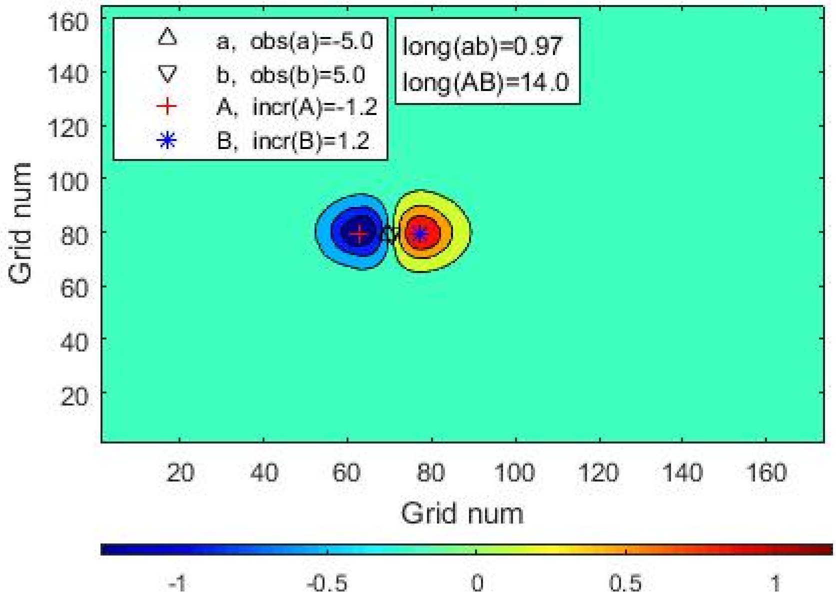
We followed the suggestion, and the following information has been added in the revised manuscript (L226-234 and L257-260). The relative standard deviation of the aerosol parameter profiles captured by the lidar over Beijing was 20.4% in the height range of 1-2 km. This lidar was calibrated via comparative observation of several lidars (Chen et al., 2019). The precision of the AEC profiles released by the other four lidars was below the quality margins (25% of the typical AEC observed in the planetary boundary layer or ±0.01km−1), as defined by Matthias et al. (2004). However, the relative standard deviation of the aerosol parameter profiles in the height range of 2-5 km released by lidar over Beijing was 35.9%.

After the quality control process, 84.32% of the original AEC data from the lidar over Beijing were accepted as valid data, and 88.75%, 54.10%, 26.74%, and 10.95% of the data from the Taiyuan, Wuhu, Shijiazhuang, and Xuzhou lidars, respectively, were valid.

2．*L285-287. It may worth trying to test the different thinning (grid-averaging) approach, from 5×5 to 1×1. As you mentioned that the spatial resolution of the model and the representativeness of Lidar AEC and surface PM data are important, since the inconsistency may cause the adjustments in two directions. It might be interesting to check if no grid-averaging is done before assimilation, but it's only a suggestion for your future study.*

Response:

We really appreciate your valuable suggestion. Actually, the scale of averaging observation data is one of the important parameters that we need to determine. However, no relevant theoretical basis has been found so far. It can only be determined roughly based on experience and a few ideal experiments. In an ideal experiment we designed, the background field is set to 0, the observation error is set to 4.6, and the two observations whose absolute value is slightly larger than the observation error a=-5.0 and b=5.0 are separated by 0.97 grid distances and are within the same grid cell. We believe that the model can only effectively simulate fluctuations with wavelengths greater than twice the grid distance. Therefore, the difference between observation a and observation b within the same grid cell represents random error, and the true value near the grid cell where the two observation points are located should be around 0. After assimilating these two observations, as showed in the following picture, the increments near observation points a and b are close to 0, which is reasonable. However, there is a negative increment center appearing at A at the 7 grid distances to the left of observation point a, and a positive increment center appearing at B at 7 grid distances to the right of observation point b, with the distance of AB reaches 14 grids distance, which is unreasonable. To avoid this unreasonable result, the simple way is averaging the two observations as one before assimilation. From the ideal experiment, we believe that the grid-averaging for observations are necessary before assimilation. As for how to choose the optimal average scale, more researches are needed in the future.



3. Section 2.3. It would be nice to add the information of observational errors for AEC and surface PM.

Response:

Thank you for your suggestion. First of all, please allow us to introduce the way of calculating observation error covariance matrix appeared in articles we have read. Following Elbern et al. [2007], Schwartz et al. [2012] and Jiang et al. [2013], the observation error covariance matrix is assumed to be diagonal, that is, the observation errors are not correlated, and the diagonal elements of R (ε_{obs}) are included contributions from measurement errors ε_m and representation errors ε_r. Elbern et al. [2007] calculated the $\varepsilon_{obs}= \varepsilon_m+ \varepsilon_r$, whereas Schwartz et al. [2012] and Jiang et al.

[2013] defined the $\varepsilon_{obs}=\sqrt{\varepsilon_m^2+\varepsilon_r^2}$. All the three articles calculated representation errors ε_r as $\varepsilon_r = \gamma\varepsilon_m\sqrt{\frac{\Delta x}{L}}$ where γ is an adjustable parameter scaling ε_m, Δx is the grid spacing and L is the radius of influence of an observation. For the ε_m of $PM_{2.5}$ or PM_{10}, Pagowski et al. [2010] used a PM2.5 measurement error of 2 μg/m3, whereas Schwartz et al. [2012] and Jiang et al. [2013] used a measurement error defined as ε_m=1.5+0.0075× Πo where Πo denotes PM observational values (units: μg/m3). For the ε_m of AEC, Yumimoto et al. [2008] introduced a minimal absolute error and defined the observation errors ε_m as ε_m=max(ε_{abs}, Πo×ε_{rel}), where ε_{abs} represents a minimal absolute error set as 0.05 km^{-1} , Πo denotes AEC observational values (units: km^{-1}) and ε_{rel} represents the relative error rate, which was assigned as 10%.

Second, please allow us to explain why the information of observational errors is not introduced in the article. The focus of this article is to accomplish the assimilation of AEC by establishing the AEC observation operator, verify the feasibility of the assimilation scheme and find some factors that may affect the assimilation effect. Because the influence of observation error on the assimilation effect is theoretically predictable, that is, the smaller the observation error, the greater the absolute value of the assimilation incremental field are, and the closer the assimilation analysis field are to the observation field deviating from the background field. In other words, no matter how large the observation error is, as long as the observation operator is correct, the assimilation analysis field will always fall between the background field and the observation field and has a positive assimilation effect, even though not the best. Because reaching the best assimilation effect through the adjustments of observation error is not the focus of this article, so in order to find factors that may affect the assimilation effect other than observation error, we set the observation error as a constant in the experiment, which is about 50% of the standard deviation of the background error of $PM_{2.5}$ (or PM_{10}, AEC). As showed in Section 2.4, the background error standard deviations of the 16 control variables have been calculated by the NMC method, and the observation operator in Section 2.5 defined the formula between the control variables and $PM_{2.5}$ (or PM_{10}, AEC), then by assuming that the background error of the control variables are uncorrelated, the background error standard deviation of $PM_{2.5}$, PM_{10} and AEC can be obtained. The observational errors of $PM_{2.5}$, PM_{10} and AEC used in this article are 5.80μg/m^3, 12.18μg/m^3 and 0.01km^{-1}, respectively.

4. *L370. Actually the application of IMPROVE algorithm is very important in this study since it simplify the complex adjoint process in the system which is innovative and interesting. However as you discussed, it may bring some uncertainties too (from observed AEC to constrain model species' concentration) since the verification of the IMPROVE parameters hadn't been thoroughly conducted for the locations where Lidar data is provided. Due to different biases between the Mie algorithm in the model and the IMPROVE algorithm in different regions, different assimilation performance may be achieved at different locations. It's suggested to clarify this point more clearly here or in the discussion.*

Response:

We really appreciated and followed the suggestion, and have added the following words in the revised manuscript (L763-769).

On the one hand, datasets from which the IMPROVE parameters were determined in previous studies were measured in specific regions and near the ground. The verification of the IMPROVE parameters had not been thoroughly conducted for the locations where lidar data were provided. Therefore, there may have been different biases between the Mie algorithm and the IMPROVE algorithm in different regions, inducing inconsistent assimilation performance.

5. L543-546. Does it also indicate different model performances for the vertical profiles at different locations? Or is it related with the different IMPROVE parametrizations for those locations? Some discussion may be nice to help the readers understand more clearly.

Response:

Thank you very much for your suggestion. We are so sorry for that the description in L543-546 is not clear enough, which increases reading difficulties for readers. What we are concerned about here is that while the lidar data are not available at surface, the DA_Ext could adjust the surface PM MCs significantly, but the adjustments could not always have positive effect. The effects of the different model performances and the different IMPROVE parametrizations at different locations are also discussed in chapter 4.

The following words have been added in the revised manuscript.

L525-536: The DA increments of AEC values from the DA_PM, that is, the AEC values obtained from the DA_PM experiment (green lines) minus those from the control experiment (blue lines), were negative for Beijing (Figure 5a), Taiyuan (Figure 5c), and Wuhu (Figure 5d) at the surface. They were also negative from the near-surface to a height of about 1000 m, although their absolute values were smaller than those at the surface. This is because the BEVCCs between each in-air layer and the surface layer were positive and decreased with height (Figure 3), so that the information contained in the surface PM MC measurements was spread to the air. However, the results of the adjustment of the AEC profiles were not always positive, because the aerosol bias of the control experiment at the surface was not always the same as it was in the atmosphere.

L546-552: In addition, although lidar data were not available at the surface, the DA_Ext adjusted of the surface PM MCs, corrected the overestimation of surface PM2.5MCs in Beijing and Wuhu, but increased the overestimation of surface PM2.5MCs in Taiyuan. This is because the information contained in the in-air AEC was spread to the surface, while the aerosol bias of the control experiment in the air did not always match that at the surface.

6． L571 Figure 6 -> 7? Please clarify.

Response:

We are so sorry for that the description legend, notes, and the description of the content shown in the figure 6 and figure 7 are not clear enough. We have revised the legend, notes, and clarified the description of the content, hoping that it will make the article clearer for readers to read.

7． L599. Actually large changes were expected to occur after sunset since PBLH and hence PM concentration change dramatically in a few hours later. For 12UTC (20LST), it's only 2-3 hours after sunset, thus continuous DA for nocturnal period should be conducted.

Response:

Thank you very much for your opinion. The characteristics of PBLH and hence PM concentration changes provide us with an important reference for design the applied assimilation scheme. The following words have been added in the revised manuscript (L603-606).

In addition, because the 1200UTC (2000LST) was only 2-3 h after sunset, so large changes of PM concentration profile may occur due to large changes in the PBLH after sunset.

We would like to express our great appreciation to you for the valuable and pertinent comment on our manuscript, which is crucial to improve the quality of our work. We hope that these revisions are satisfactory and that the revised version will be acceptable for publication in Geoscientific Model Development. Thank you very much for your work concerning my paper.

Wish you all the best!

Yours sincerely,

Yanfei Liang, Wei You and Zengliang Zang

05/10/2020
