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:

General comments:

1. *How can AOD distinguish and constrain 20 different aerosol state variables? What is the impact of using only AOD? There is no mention of other studies that assimilate more information than just AOD (e.g. AOD in other wavelengths or Angstrom Exponent, Absorption Aerosol Optical Depth or Single Scattering Albedo as well as direct radiances assimilation). Although the authors acknowledge the need for combine assimilation of various optical properties in their closing statement in conclusions (L746-751), many recent studies that are related to that are not mentioned. To name a few ones: (Chen et al., 2019; Escribano et al., 2017; Tsikerdekiset al., 2021)*

Response:

Thank you very much for your questions and suggestions. First, the forward observation operator links aerosol optical properties (including AOD, extinction coefficient, backscattering coefficient, and total attenuated backscattering coefficient) with 20 different state variables in the data assimilation system, which means that AOD observations distinguish and constrain 20 different state variables via the forward

operator. Designing and establishing the observation operator is crucial to directly assimilate optical properties in case that control or state variables are mass concentrations instead of optical properties. Fortunately, we can reduce the aerosol Optical Module within WRF-Chem to establish the forward operator, which is based on the Mie-scatter theory. Different aerosol species described by 20 aerosol state variables here make greatly different contributions to AOD, even for the same species, particles within different size bins make different contributions. The operator can quantify these contributions. Specifically, AOD can constrain particle size and number, and then adjust individual species mass concentrations denoted by 20 different aerosol state variables. Second. Only AOD observation was chosen to test the developed assimilation system, its impact may be insufficient for significantly improving aerosol forecasts. It is noted that the developed assimilation system can assimilate extinction and backscattering profiles, AOD, and attenuated backscattering at different wavelengths because the wavelength is designed as a variable parameter in the assimilation system when establishing the observation operator, but it can not assimilate other optical properties such as Angstrom Exponent, Absorption Aerosol Optical Depth or Single Scattering Albedo as well as direct radiances (Assimilating aerosol direct radiance is very challenging because it is affected by many factors). Nevertheless, we will attempt to combine assimilate more aerosol optical properties to constrain model variable more accurately in the near future work. Finally, some recent studies related to combined assimilation of various optical properties have been added in the revised version as “With the increase in aerosol observations, the simultaneous assimilation of aerosol observations from various platforms has become a trend, in particular combined assimilation of various optical properties has made great progress in recent year (Escribano et al., 2017; Chen et al., 2019; Tsikerdekiset al., 2021).”

(L399-402)

2. *The spatial aggregation of observations that the authors describe (aggregating observations in the spatial resolution of the model) is indeed often used in data assimilation studies. Although was there any consideration regarding the representation error of this aggregated observations? For example, was the observational error inflated by X amount because you were not using the original resolution of Himawari-8? (Lines 437-442)*

Response:

We really appreciate your valuable suggestion. We aggregated AOT observations in the spatial resolution of the model, which is also employed by other researchers (Yumimoto et al., 2016; Dai et al., 2019; Ha et al., 2020). The observation error plays an important role in assimilation process. In general, the observation error depends on measurement error and representation error, however, it is very difficult to accurately determine the representation error because the released AOT product gives the retrieval uncertainty rather than representation error, what is more, the retrieval uncertainty is just a reference range. Consequently, the observation error here can only be roughly determined based on experience or tuning parameter. Aggregating AOT observations by averaging them in one grid cell can not inflate observation error, conversely, this approach can smooth out much noise to improve the quality. At least, the assimilation practice has demonstrated that assimilating aggregated AOT observations is better than original observations.

3. *As a geostationary satellite, Himawari-8 is known for its high temporal frequency. Since the data assimilation cycle is in daily frequency (updating analysis once a day), are you fully exploiting this satellite capabilities or rather its strong point? I realize that the daily assimilation step was chosen for practical reasons*

(computational speed), nevertheless I would expect some discussion about it. Further related to this topic, I did not find any discussion related to temporal collocation of observation in the data assimilation system.

Response:

We really appreciate your suggestion. Himawari-8 level 3 AOT_Merged, an improved hourly product, which is derived from level 2 AOT retrievals at a 10 min interval, was employed to conduct assimilation experiments. A daily assimilation frequency seems to be an underutilization of Himawari-8 observations in comparison to its high temporal frequency. Since AOT observations are retrieved at the visible and infrared bands, observations between 03 and 08 UTC in the daytime are available for China. In fact, AOT observations are noticeably noisy, which will have a greatly negative impact on assimilation results. What is more, observations at afternoon are much noisier than those in the morning. For example, surface $PM_{2.5}$ concentration and original (not thinned) Himawari-8 AOT observations at 0300 UTC and 0600 UTC are plotted in Fig. 1 and Fig. 2, respectively. Overall, surface $PM_{2.5}$ mass concentrations change little even with a small decrease at some areas from 0300 to 0600 UTC (Fig. 1b, Fig. 2b) while there is a remarkably increase in AOTs during the same period (Fig. 1a, Fig. 2a). In terms of $PM_{2.5}$, the noticeably increase in AOT observations should not be considered as normal changes of aerosol but much noises. As a result, more frequent assimilation of AOT observations like this will certainly result in a dramatic overestimation of $PM_{2.5}$ mass concentrations. In terms of evaluation with $PM_{2.5}$ mass concentration observations, AOT observations at 0300 UTC without no temporal collocation were only assimilated in this study to test the developed assimilation system. As known, data assimilation serves only as a mathematical approach on how to introduce observations into the model, and then improves model initializations and

forecasts. Assimilation results are largely determined by the quality of observational data, as for how to deal with those with high noise and improve the quality, more researches are needed in the future. Moreover, the advanced DA system such as 4DVAR will be developed in the future that can assimilate observational data from a time window.

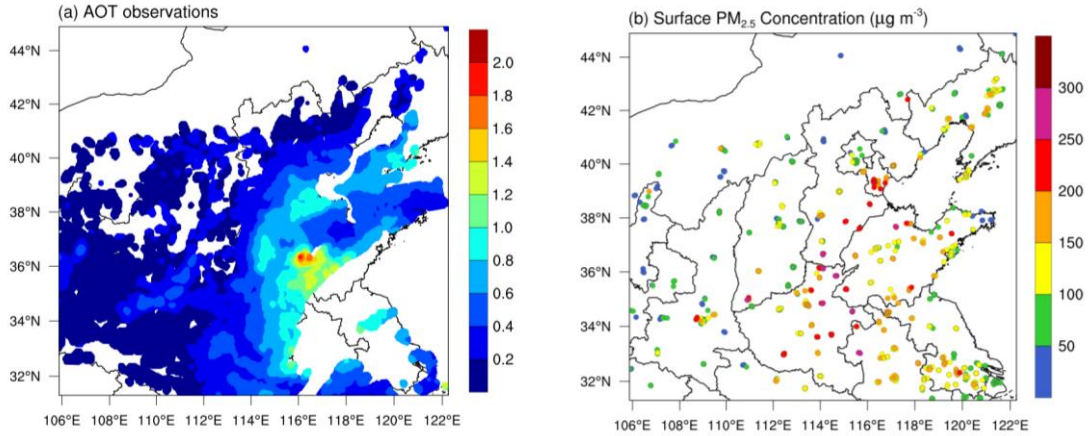


Figure 1. Observations of the original (not thinned) Himawari-8 AOTs (a) and surface $PM_{2.5}$ mass concentration (b) in D02 at 0300 UTC on 25 November 2018.

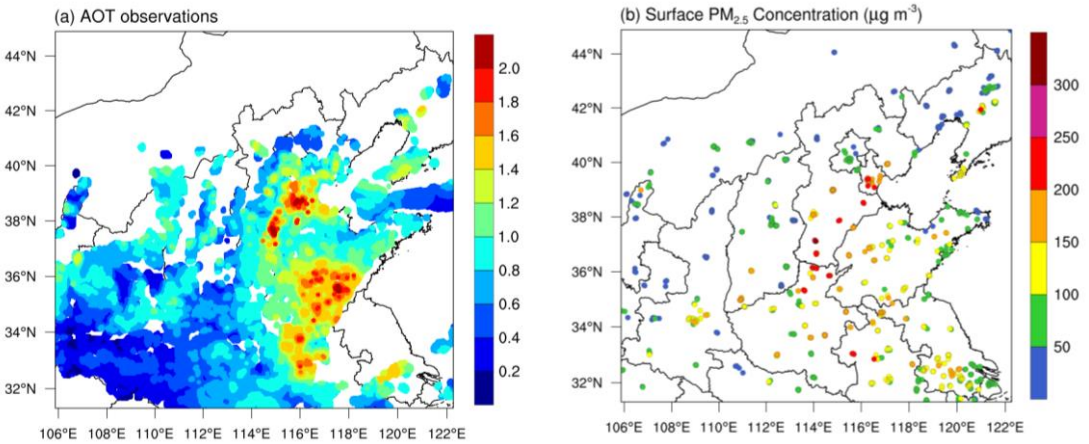


Figure 2. Same as Fig.1, but at 0600 UTC on 25 November 2018.

Specific Comments:

4. L60: Missing references.

Response:

We really appreciated the suggestion and followed it. Three references have been added here (L61).

Menon, S., Hansen, j., Nazarenko, L., and Luo, Y.: Climate Effects of Black Carbon Aerosols in China and India, Science, 297, 2250–2253. https://doi.org/10.1126/science.1075159, 2002.

Gao, M., Guttikunda, S. K., Carmichael, G. R., Wang, Y., Liu, Z., Stanier, C. O., Saide, P. E., and Yu, M.: Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area, Sci. Total. Environ., 511, 553−561, https://doi.org/10.1016/j.scitotenv.2015.01.005, 2015.

Qian, Y., Gong, D., Fan, J., Leung, L.R., Bennartz, R., Chen, D., and Wang, W.: Heavy pollution suppresses light rain in China: Observations and modeling, J. Geophys. Res., 114, D00K02, https://doi.org/10.1029/2008JD011575, 2009.

5. *L65-67: Reference, name and accessibility (or the lack of) for this dataset should be provided.*

Response:

We really appreciate your valuable suggestion. This dataset is provided by China National Environmental Monitoring Centre (CNEMC) but has no official name. This sentence has been revised as “For instance, China National Environmental Monitoring Centre (CNEMC, http://www.cnemc.cn/en/) has established a nationwide monitoring network consisting of more than 1500 stations since 2013 to provide near-time data of pollutants, including $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3.”(L66-67)

6. *L73: Probably mean “remote sensing optical properties can cover a much larger domain”. Because just optical properties can be retrieved also from AERONET stations.*

Response:

Thank you so much for your valuable suggestion. The sentence has been revised as “Remote sensing optical properties can cover a much larger domain (Kaufman et

al., 2002) and provide detailed aerosol profiles (Young and Vaughan, 2009)" (L75-76), at the same time, this reference has been added in the revised manuscript ("Young, S. A. and Vaughan, M. A.: The retrieval of profiles of particulate extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description, J. Atmos. Ocean. Tech., 26, 1105–1119, https://doi.org/10.1175/2008JTECHA1221.1, 2009.")

7. *L189-192: In principle $PM_{2.5}$ can be estimate from the modes that the MADE scheme uses, assuming you know the median and the standard deviation of the distribution for each mode. In that case MADE would be superior to MOSAIC since it will also include mixing of different species within each mode. So I would suggest to emphasize only the numerical efficiency of MOSAIC against MADE. Further, indicating how much faster it is could really promote that argument and it could be easily estimated with two forward simulations, one with MADE one with MOSAIC (no DA required).*

Response:

Thank you so much for your valuable suggestion. We agree well with you. Due to its simplicity and high numerical efficiency, the MOSAIC scheme has been chosen to develop the data assimilation system. Consequently, it seems to unnecessary to discuss how much faster is MOSAIC against MADE for aerosol simulations in the context of testing the assimilation system.

8. *L211-213: Authors could mention here that the vertical axis is on hybrid sigma-pressure levels, if that is the case.*

Response:

We followed this suggestion and this sentence has been revised as "To ensure a detailed simulation of aerosol vertical distributions, 40 vertical layers were modelled

in the simulation, and it is worth mentioning that the vertical axis is on hybrid sigma-pressure levels with a resolution decreasing with height. The lowest layer is at the surface, whereas the top reaches 50 hPa". (L214-215)

9. *L237-238: It would be really helpful to briefly mention here how Yumimoto et al. (2016) estimated this error for Himwari-8 AOD and what this error actually describes (e.g. instrument error, retrieval error, representation error) ?*

Response:

Thank you so much for your valuable suggestion. Yumimoto et al. (2016) estimated observation errors to be the retrieval uncertainty attached to the Himawari-8 AOT data plus a standard deviation calculated as the representative error in the regridding (Zhang et al., 2008, see below). The retrieval uncertainty ranged from 0.0001 to 1.04 with average of 0.013 and has larger values in the land relative to over the ocean.

The observation error plays an important role in assimilation process, however, no relevant theoretical basis has been found so far. The observation error depends on measurement error and representation error (Elbern and Schmidt, 2001; Schwartz et al., 2012; Jiang et al., 2013), nevertheless, how to determine the observation error is also a matter of assimilation practice. Because the observation error determines the weight of observation across the analysis, 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. In this study, AOT observation error was set to be a simple value which is rational only

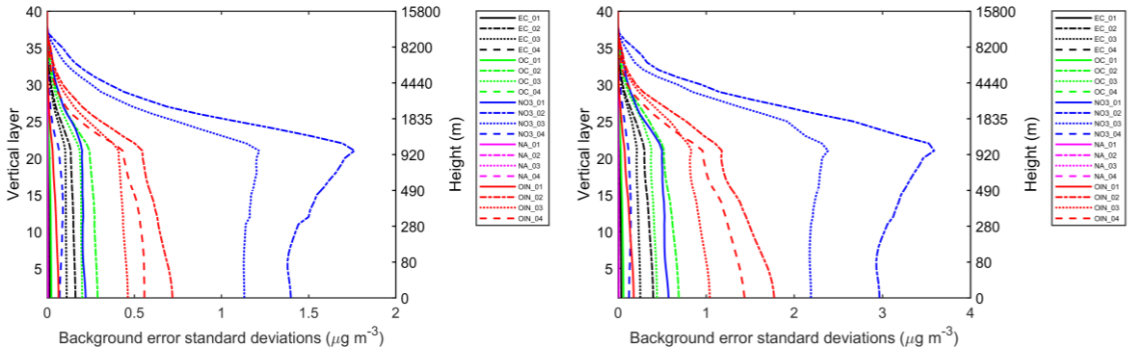
to test the developed assimilation system.

Zhang, J., Reid, J. S., Westphal, D. L., Baker, N. L., and Hyer, E. J.: A system for operational aerosol optical depth data assimilation over global oceans, J. Geophys. Res., 113, D10208, https://doi.org/10.1029/2007JD009065, 2008.

10. *L491-493: It would be interesting to compare the D02 and D01 estimated background error standard deviation. It would show how important is the model horizontal resolution for this metric. If possible an additional plot for the D01 over the domain of D02.*

Response:

We really appreciated the suggestion. Because both D01 and D01 outputs were assimilated using AOT observations in this study, background error covariance including standard derivation and correlation was estimated in D01 and D02, respectively. Only the estimated background error standard deviation in D02 was shown in manuscript, as shown in Fig .3b here, the D01 estimated background error standard deviation looks actually like D02, as shown in Fig. 3a. Obviously, the D02 estimated background error standard deviation is nearly twice than D01 estimated ones, whereas the D01 model horizontal resolution is 27km and D02 is 9km. The background error standard deviation determines the magnitude of analysis increments across aerosol control variables. As these two plots look alike, it seems unnecessary to add the plot for D01.



(a) **(b)**

Figure 3. Vertical profiles of background error standard deviation in mass concentration for aerosol control variables, (a) is for D01, and (b) is for D02.

207

208 *11. L562: I would strongly recommend to replace "improvements" with "changes" in*
209 *that sentence or rephrase. Figure 6 shows the differences of the Analysis – Control.*
210 *It is not an evaluation with observations (assimilated or independent) where we*
211 *can truly determine if there was an improvement by the data assimilation.*

212 **Response:**

213 The word "improvements" has been replaced by "changes" (L571).

214 *12. L585-587: It would be beneficial to provide how much this difference in AOD*
215 *wavelength (500nm and 550nm) is affecting your evaluation. Maybe you can use*
216 *Angstrom Exponent from AERONET to determine that and provide a number?*
217 *Usually AOD at higher wavelength (550nm) is smaller than AOD at lower*
218 *wavelength (500nm). Which means that the bias would be even more negative if*
219 *you were comparing MODIS and Model at the same wavelength at Figure 7b. I*
220 *think it is worth discussing in the manuscript (L595+) although it may enhance*
221 *the negative bias you get for both Control and Analysis.*

222 **Response:**

223 We really appreciated the suggestion and followed it, AOD simulation was
224 performed at a wavelength of 500 nm, the same as Himawari-8 retrievals, whereas

MODIS AOD is retrieved at 550 nm. Even though this difference in AOD wavelength may affect the evaluation, it is naturally convincing to evaluate AOD simulation directly employing MODIS AOD because the wavelength difference is minor.

There is no doubt that your suggestion will certainly improve the manuscript, and the following information has been added in the revised manuscript (L607-612).

Usually AOD at higher wavelength (550 nm) is smaller than AOD at lower wavelength (500 nm), so the bias would be even more negative if comparing AOD simulations with MODIS AOD for both Control and Analysis, which is demonstrated by the indicator BIAS in Fig. 7. For instance, BIAS is -0.031 when comparing with Himawari-8 AOD, while BIAS is -0.140 against MODIS AOD after assimilation.

13. L604-606: AERONET sites at Figure 1b are hardly visible (probably because 4 of them are in the Beijing area). It would be visually better to enlarge them a bit.

Response:

We really appreciated and followed the suggestion, and have added a zoomed-in map as Fig. 1c for AERONET sites in Beijing area in the revised version, which is also given as Fig. 4 below:

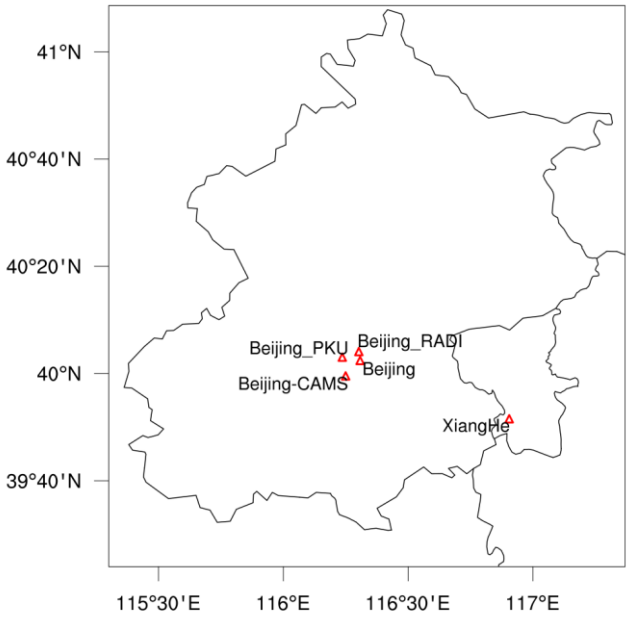


Figure 4. A zoomed-in map for AERONET sites in Beijing area, including Beijing, Beijing-CAMS, Beijing_PKU, Beijing_RADI, XiangHe.

14. L664-669: Good point, spatial availability of AOD in contrast to PM2.5 can play a role. I would also add that AOD is an atmospheric column measurement while PM2.5 is a surface measurement. Therefore, if you have an aerosol plume which is not close to the surface AOD can be increased by increasing the aerosol concentration of that plume while PM2.5 can remain almost unaffected by that change.

Response:

We really appreciated and followed the suggestion, and have added the following descriptions in the revised manuscript (L694-697).

Besides, AOD is an atmospheric column measurement while $PM_{2.5}$ is a surface measurement. Therefore, if you have an aerosol plume which is not close to the surface, AOD can be increased by increasing the aerosol concentration of that plume while $PM_{2.5}$ can remain almost unaffected by that change.

Technical Corrections:

L140: "3DAVR" to "3DVAR"

Response:

Done. (L143)

L173: "back carbon" to "black carbon"

Response:

Done. (L175-176)

L203: "/MADE/" is some kind of typo?

Response:

This sentence has been revised as "the Regional Acid Deposition Model, Version 2 (RADM2, Stockwell et al., 1990), the Modal Aerosol Dynamics Model for Europe (MADE, Ackermann et al., 1998)/Second Organic Aerosol Model (SORGAM, Schell et al., 2001) anthropogenic emissions." (L206-207)

L291: "black carton, organic carton" to "black carbon, organic carbon"

Response:

Done. (L294-295)

L609: Something is missing in the sentence. Probably "used to" to "used them to"

Response:

Done. (L622)

L1185: Figure 11: Do you mean "average over 7 analysis steps" instead of "average over 7 single experiments"?

Response:

We really appreciated and followed the suggestion. Two one-week parallel experiments have been performed to evaluate AOD assimilation effects regarding to 24 h regional $PM_{2.5}$ forecasts. For a general assessment, the statistics were averaged over 7 analysis steps. (L1213)

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,

Daichun Wang and Wei You

11/23/2021
