

The authors appreciate the reviewer's constructive and friendly comments. We have substantially revised the manuscript. New data and figures are present in the main text. A new supplementary document is included in the revision. We reply to the reviewer's comments point by point.

Anonymous Referee #1

Received and published: 19 October 2020

The manuscript presents the development of assimilation of aerosol observations into WRF-Chem using the GSI system using approaches that are different to those used in previous studies. These developments are then tested for a case of assimilating ground-based observations of particle mass concentration, scattering and absorption coefficients, and AOD, performing sensitivity simulations on assimilating datasets independently and jointly. This is done for a single site located in Kashi, representative of dust conditions. This study is within the scope of ACP and represents good contributions to the field as it develops a tool that could be used by the community and highlights shortcomings in the techniques and how they could be improved. I think the paper needs a bit more work before it's ready for publication based on the comments below.

My main comments are the following.

- While the WRF-Chem optical properties module assumes Mie theory which is based on particles being spherical, the testing of the tool is focused on dust which are mostly non-spherical particles. This is briefly mentioned in the article, but I would like to see more on the subject, including looking into literature that has explored this topic and discussion on what discrepancies obtained in this study could be explained by this issue. See more on by line comments.

Response: Yes, the spherical dust particle in WRF-Chem introduces uncertainty. We reviewed a few literatures and added a paragraph to discuss the impact of non-spherical particles in section 3.3.

- I believe that what the authors defined as Adjoint operators are really the tangent linear models, i.e., the derivative of the observables with respect to the inputs (aerosol mass). The adjoint operates on perturbations on the observables and outputs the expected perturbations on inputs. Please verify with the literature and correct accordingly.

Response: The author appreciates the reviewer's kindly comment. We have changed the misstatement of "adjoint operator" to "tangent linear operator".

- Assess representation of some intensive properties such as size (e.g., angstrom exponent, ratio of pm_{2.5} to pm₁₀), single-scattering albedo, and mass scattering efficiency to try to understand mismatches when doing assimilation. A little bit is done but it would be very helpful to expand this topic and use the nomenclature used in the literature. See more on by line comments.

Response: According to this comment, the revised manuscript shows additional assessments of angstrom exponent and SSA (Table 3). A new figure 10 shows the multi-wavelength mass scattering/absorption coefficient. Hope this additional content makes this study more convincing.

- Absorption seems completely biased even after assimilation, this points to issues probably related to underestimation of imaginary refractive index of dust. Look for literature on this depending on the deserts, I believe Chinese deserts tend to have darker (i.e., more absorbing) sands.

Response: We used the generic model value of dust refractive index in the first version manuscript. In the revision, we increased the imaginary part of the dust refractive index, which is higher than the imaginary part of the Taklimakan desert that has been retrieved by Di Biagio et al. (2019). We find that tuning dust DA is not helpful for removing the DA bias in absorption coefficient. We redid a lot of DA experiments and found a negligence of anthropogenic emission in the WRF-Chem simulation. This strong bias in absorption DA is relevant to the low concentration of black carbon (BC) and the low BC's background error. We have rewritten the relevant content in section 3.5.

Di Biagio, C., et al.: Complex refractive indices and single-scattering albedo of global dust aerosols in the shortwave spectrum and relationship to size and iron content, Atmos. Chem. Phys., doi:10.5194/acp-19-15503-2019, 2019.

Comments by line:

103-108. For completion, consider citing and discussing the study by Kumar et al (2019) that also uses GSI with CMAQ but does not use the CRTM as Tang study. This study also provides an alternative way of computing the BEC matrix (other than the NMC method) which you discuss in section 2.3

Response: We added a few words about Kumar et al. (2019) study in the revised introduction and section 2.3.

545-552. Are all of these observations in the same location? If not how far apart are they? How many PM2.5/PM10 sites are used? Also, what's the inlet cutoff size used for the scattering and absorption measurements? This is important to related mass and optical properties properly.

Response: All the observations (PM2.5, PM10, AOD, scattering/absorption coefficient) were carried out at a single site. There was no inlet cutoff for the scattering and absorption measurements. In the revised manuscript:

"The site was placed in the Kashi campus of the Aerospace Information Research Institute, Chinese Academy of Sciences (39.50°N, 75.93°E; Li et al., 2018), about 4 km in the northwest to the Kashi city. ... All the instruments were deployed at the roof of a three stories height building on the campus."

554-561. Could you add justification for the PM2.5/PM10 observation errors stated in Table 1? There is no explanation how the errors were picked. Also, why do you only use representative error for PM2.5/PM0 but not for the other observations?

Response: In the revised section 2.4:

"The observation errors of PM_x are handled in the conventional way (Schwartz et al., 2012; Chen et al., 2019), which contains the measurement error (e_1) and the representative error (e_2). The measurement error is the sum of a baseline error of $1.5 \mu\text{g m}^{-3}$ and 0.75% of the observed PM_x concentration. The representative error is the measurement error multiplied by the half-squared ratio of the grid spacing to

the scale distance. The scale distance denotes the site representation in GSI and has four default values of 2, 3, 4, and 10 km, corresponding to the urban, unknown, suburban, and rural sites. We used 3 km for the scale distance in this study. As we had a single site in Kashi, it is difficult to estimate the site representation error. Since the DA analysis was based on the child model domain with a horizontal resolution of 5 km, close to the site distance to the Kashi urban area, we assumed the aerosol optical measurement had good representativeness of the model grid covering the site. The observation error of CE318 AOD took the AERONET AOD uncertainty of 0.01 in cloud-free conditions (Holben et al., 1998).”

569. Can you add a bit more info on the vertical resolution? For instance, thickness of the 1st level and number of levels within 1km.

Response: In the revised section 2.5:

“Both domains had 41 vertical levels extending from the surface to 50 hPa. The lowest model layer at the site was approximately 25-meter height from the ground.”

594-596. Can you clarify if you did 2 simulations every cycle with and without aerosol interacting with radiation, or it was a single simulation with two calls to the radiation code within the same simulation?

Response: It was a single simulation with two calls to the radiation code. In the revised section 2.5:

“To study the impact of DA on aerosol direct radiative forcing (ADRF), we modified the WRF-Chem code to calculate the shortwave irradiance with and without aerosols at each model integration step. The modified WRF-Chem model restarted from each DA analysis and ran to the next analysis time. Each running performed the radiation transfer calculation twice, and each calculation saw the aerosols and clean air, respectively. The irradiance difference between the two pairing calls was aerosol radiative forcing.”

632-638. You are also missing some processes of potential importance such as secondary organic aerosol formation and heterogeneous sulfate formation influencing low-dust days.

Response: Yes, our simulation did not have SOA, and the heterogeneous sulfate formation in WRF-Chem may bias. Nevertheless, we accidentally lowered the anthropogenic emissions in the original WRF-Chem simulation. Because of the ambition of haze abatement in China since 2013, the anthropogenic emissions had dramatic reductions in 2013-2019. So far as we know, a timely update of emission inventories is not available, and we used the open MEIC emission inventories for the year 2010 when the anthropogenic emissions had peak values. A general way to handle this emission reduction is to scale the historical emissions, which were not appropriately handled in our first manuscript. The anthropogenic emissions that we set for Kashi in the 2019 simulation were too low. As lack of aerosol measurement at Kashi, the low bias was not identified at the first glance. In the revised manuscript, we just ignore the yearly emission differences. We redid all simulations with the MEIC emission inventories for 2010.

The revised model concentrations of PM_{2.5} and PM₁₀ are almost equivalent to the old data (Figure 1) because dust is the dominant component at Kashi. Besides, the real part of the refractive index of sulfate, nitrate, ammonium, and dust are comparable in the model. Thus, the new results do not change

the conclusion. The new advantage is that the DA bias in absorption coefficient can be somewhat attributed to black carbon when the BC's background error was amplified. We rewrote the DA of the absorption coefficient in section 3.5.

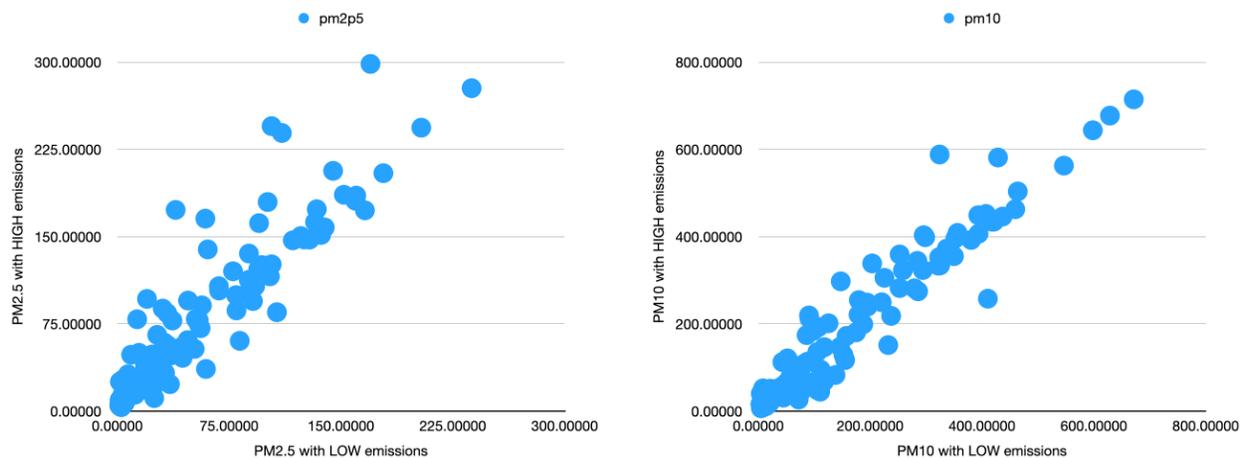


Figure 1. Comparisons of PM_{2.5} (left) and PM₁₀ (right) in the WRF-Chem simulations with high (y-axis) and low anthropogenic emissions (x-axis) at Kashi in April 2019

645-647. I think a better fit to PM_{2.5} could be achieved if you relaxed the interbin correlation. It looks like PM₁₀ is fitting pretty well but it's going a bit over the observation, so this is restricting increases in PM_{2.5} due to the correlation. Since bin 4 is 2.5-10um, in theory, if no interbin correlation was present, PM₁₀ and PM_{2.5} should be able to fit independently. For this study it would make sense to relax the interbin correlation due to the known issues in dust size distributions (see next comment)

Response: Based on lots of experiments, we find that the analyses are not sensitive to the inter-size bin correlation length in this case, though the analyses changed a lot when we turned off the inter-size bin correlation. We find that the magnitude of large background error of coarse dust is more effective in affecting the analysis of PM_{2.5}. Reducing the background error of the fourth size bin OIN (oin_a04) will increase PM_{2.5} and decrease PM₁₀. We added table S3 in the supplementary document, which shows the PM_x response to the different magnitudes of oin_a04's background error.

In the revised section 3.2:

“Applying the inter-size bin correlation length caused the interlinked analyses of PM_{2.5} and PM₁₀. In the desert area, the coarse and fine dust are readily affected by the magnitude of BEC of the fourth size-bin OIN (oin_a04). We intentionally decreased the BEC of oin_a04 by 10% each time to 30% of its original value. The magnitude of 30% of oin_a04 was comparable to the magnitude of the third size-bin (oin_a03) OIN's background error. As shown in Table S3, because the oin_a04's BEC reduction relaxes the constraint on the coarse particle, the PM₁₀ bias becomes more negative along with the decrease in oin_a04's BEC. Meanwhile, the PM_{2.5} bias becomes more positive. Correspondingly, the ratio of PM_{2.5} to

PM₁₀ was exaggerated to 0.33 with 30% of oin_a04's BEC, higher than the observed value of 0.28. Overall, the original BEC of oin_a04 is a reasonable tradeoff in our DA experiments."

647-652. Literature on dust modeling states that parameterizations tend to overpredict the fine dust and underpredict the coarse dust (see Kok et al., 2011, Adebisi and Kok 2020). So the joint assimilation of PM_{2.5} and PM₁₀ could be somewhat correcting for that, which is an additional possible explanation to the behavior explained in these sentences.

Response: Thanks for the hint. We cite the two pieces of literature in the revised section 3.2:

"As a result, the ratio of PM_{2.5} to PM₁₀ decreased from 0.39 in the background to 0.27 in DA_PMx, approaching the observed ratio of 0.28. Such improvement was consistent with the correction required to the model desert dust in literature. Kok et al. (2011) found that regional and global circulation models underestimate the fraction of emitted coast dust (>~5 μm), overestimates the fraction of fine dust (<2 μm diameter). Adebisi and Kok (2020) claimed that too rapid deposition of coarse dust out of the atmosphere accounts for the missing coarse dust in models. Similarly, WRF-Chem assimilated too much smaller dust particles than the observed. According to Kashi's AOD between 440 nm and 1020 nm, the observed Ångström exponent (AE) was 0.18 in this case, but the background value was 0.54 (Table 3). DA_PMx reduced the AE value to 0.30, a little improvement but not sufficient."

672-685. Another reason for the discrepancy is related to the size distribution. Are you assimilating multi-wavelength AOD here, right? If so, I would expect some modifications to the size distribution. It looks you are effectively modifying size distr. as the ratio of PM_{2.5} to PM₁₀ ratio is reduced from 0.31 in the background to 0.11 in the DA_AOD simulation but it might be going a bit too far as the observed ratio is 0.28. You can also check angstrom exponent. You can also explore the point you make at the end related to the dust mass extinction efficiency, you have observations to compute this at the surface. Additionally, there is also potential for your vertical distribution to be off and be generating these issues. You can diagnose this by comparing the ratio of surface extinction vs AOD. It seems the model is overpredicting this ratio, which could mean too much aerosol close to the surface.

Response: In the revised manuscript, we check the PM_{2.5}/PM₁₀ ratio, mass extinction efficiency, angstrom exponent (AE), and SSA. We do not check the ratio of surface extinction and AOD because it requires the interpolation of surface extinction and AOD to similar wavelength. The model has a large bias in AE, resulting in an unreliable interpolation. We rewrote a lot in sections 3.2 and 3.3. Please refer to the revised manuscript.

Related to this point. You are actually already computing mass scattering efficiency (2nd column in Table 3). The background already underpredicts it, and the assimilation makes it worse as you are increasing the coarser fraction. You could explore if there is an underprediction of the dust refractive index. You could look into values provided in the literature for the region studied and compare to what WRF-Chem uses.

Response: We add a supplementary document to give the complex refractive indexes for all aerosols in this study. A part of the table is shown below. We set the dust's refractive index referring to the generic

model values in literature. The imaginary part in our study is higher than the retrieved imaginary part for the Taklimakan desert dust by Di Biagio (2019).

Table S1. Multi-wavelength real and imaginary parts of refractive indexes of aerosol chemical compositions and water in this study (a part of the snapshot of table S1)

(nm)	440	450	470	520	525	550	635	660	675	870	1020
	<i>OIN, dust (Cheng et al., 2006; Zhao et al., 2010)</i>										
Real						1.53					
Imag	0.003	0.003	0.003	0.0025	0.0025	0.002	0.0015	0.0015	0.0015	0.001	0.001

678-683. I think there is no need for this very long description of the Ma paper as these results are not that relevant to the area study as RH is likely low in the desert and dust aerosols tend to be hydrophobic
 Response: We have removed the statements in the revision.

690-692. AOD to PM10 ratios depends on many variables. Since you are blaming discrepancies to issues in mass scattering/absorption efficiency it makes more sense to do direct comparisons to this variable as you have in-situ measurements of scattering and absorption

Response: In the revised table 4, we show the ratios of AOD, scattering/absorption coefficient to PM10 per DA experiment. In the revised section 3.3:

“Table 4 shows the ratios of the AOD and aerosol scattering/absorption coefficients to the surface PM₁₀ concentrations. The ratio of AOD to PM₁₀ in the background model result was one-third of the observed levels. The observed mass scattering coefficient (Esca/PM₁₀) was 1.05 Mm⁻¹ μg⁻¹ m³, while the background value was only 0.65 Mm⁻¹ μg⁻¹ m³. DA_AOD did not eliminate the low bias but enlarged the low bias to 0.51 Mm⁻¹ μg⁻¹ m³. The same thing occurred for Eabs/PM₁₀, which was 0.09 in the background and 0.05 in DA_AOD, much lower than the observed value of 0.25. Figure 10 shows these mean ratios at the other wavelengths. The low bias in AOD/PM₁₀ was comparable at each wavelength. ...”

693-696. You can assess issues with size distribution by using the angstrom exponent.

Table 2 and 3. Is there any reason behind using the lower wavelength (440-450nm) for these comparisons? Since the focus of this work is on dust, it would be preferable to compute optical properties for longer wavelengths where coarse aerosols contribute more to the scattering

Response: A new table 3 shows the angstrom exponent and SSA. The revised discussion in the main text is based on 870 nm AOD, 635 nm scattering coefficient, and 660 nm absorption coefficient.

702-706. There is extensive literature on how optical properties of dust particles deviate from Mie theory (e.g., Dubovik et al, 2006, Nousiainen et al 2015). It would be good for the authors to reference this work and attempt to explain what could be the implications of using Mie theory, and if those can explain any of the discrepancies found when assimilating multiple datasets in this study.

Response: In the revised section 3.3:

“The irregular morphology had a significant influence on the dust simulation. Okada et al. (2001) found that the aspect ratio (the ratio of the longest dimension to its orthogonal width) of the mineral dust particles (0.1-6 μm) in China arid regions exhibited a median of 1.4. Dubovik et al. (2006) suggested the aspect ratio of ~ 1.5 and higher in desert dust plumes. Kok et al. (2017) found that the dust’ sphericity assumption underestimated dust extinction efficiency by $\sim 20\text{--}60\%$ for the dust particle larger than $1\mu\text{m}$. Tian et al. (2020) found that using a dust ellipsoid model could increase the concentration of coarse dust particle (5-10 μm) by $\sim 5\%$ in eastern china and $\sim 10\%$ in the Taklimakan area because of the decrease in gravitational settling, comparing with the simulations with dust sphericity model. Nevertheless, the aspect ratio of the spheroid dust is uncertain. Even after applying the spheroidal approximation, Soorbas et al. (2015) found that the model underestimated 550 nm aerosol scattering and backscattering values by 49% and 11%, respectively, because of the uncertainties in particle axial ratio, complex refractive index, and the particle size distribution. To date, the assumption of spherical particles has been widespread in models (including WRF-Chem) for computational efficiency. Impact of dust morphology to DA deserves a further investigation.”

Figure 12. It would help to see an additional panel with these profiles being normalized, so we can more easily assess by how much the assimilation of the different datasets is changing the vertical distribution.

Response: The revised figure 12 has additional two panels showing the vertical distributions normalized to the surface PM_x concentrations.

730 You know it overestimated PM₁₀, not sure about aerosol number concentration (you would need a different observation for assessing that)

Response: We do not have the surface measurement of aerosol number concentration. The amounts of quality assured retrievals of aerosol columnar volume and effective radius by CE318 are limited in DAO-K (<9 days, 1 to 4 data samples per day). It is difficult to give a robust verification. The original statement describes that the GSI tends to increase the aerosol number in response to the high aerosol mass concentration. We changed the statement to:

“The revised GSI updates aerosol number concentration according to the analyzed aerosol mass concentration and the background ratio between mass and number concentrations. Thus, an overestimation of aerosol mass concentration inclines to raise aerosol number concentration, resulting in high scattering/absorption coefficients.”

734. Use single-scattering albedo for this

Response: We add a new table 4 to show the angstrom exponent and SSA. In the revised section 3.3:

“Additionally, we computed the surface single scattering albedo (SSA_{srf}) with the 525 nm scattering coefficient and 520 nm absorption coefficient. We did not use the Ångström exponent to interpolate the scattering/absorption coefficients to a similar wavelength because the AE itself had a large model bias even after DA (Table 3). The observed SSA_{srf} value was 0.78, indicating an emphatic absorption particle,

probably due to the mixture of anthropogenic black carbon and natural desert dust in the local air. The model background SSA_{srf} was 0.86, while the DA analyses gave even higher SSA_{srf} (0.88 to 0.9)."

738-746. This is a misconception, aerosol light extinction and AOD does not depend on sun light intensity (for instance, you can sample both at night time with different methods). What's going to change with sunlight are the radiative effects. There are likely other reasons to explain this diurnal behavior. Look into the diurnal evolution of your BEC, and also into diurnal evolution of dust reaching the city. Similar misconnection is mentioned in lines 770-771.

Response: In the revised 3.3,

"Assimilating the AOD seems to increase the diurnal variation in the DA analyses, but this variation was not conclusive since there were different amounts of AOD data for DA at 00:00, 06:00, and 12:00. The AOD data were not always available as the data quality control (i.e., cloud screening). There was a higher increase in the concentration at noon (06:00 UTC) (Figure 9b), corresponding to a few high AOD during mild dust episodes at that hour. ..."

The misstatement in the original lines 770-771 has been removed.

960-963. This is probably due to underprediction of dust imaginary refractive index

Response: We set the dust refractive index to refer to the generic model values in literature. The imaginary part in our study is higher than the imaginary part for the Taklimakan desert dust retrieved by Di Biagio (2019). The strong bias in absorption coefficient can be largely removed by tuning the background error of black carbon, though additional disadvantage is introduced. Please refer to the revised section 3.5.

Minor Edits

Fig 5 caption. It reads like a) and b) represent PM_{10} and winds, respectively, but I think that's not the case. Please revise

Response: The figure caption is changed to

"Figure 5. Monthly mean PM_{10} concentration ($\mu g m^{-3}$) and the streamlines of the 10-m wind ($m s^{-1}$) in April (a, b) and their daily mean anomalies (c, d) ..."

623. Did you mean "underestimates" instead of "lowered"?

Response: changed to "underestimates"

780. Do you mean "particles that absorb radiation" rather than "aborting particles"? Also, I would like black carbon in that list as well.

Response: Corrected.

781-791. I believe primary dust in WRF-Chem is also considered to be a bit absorbing (has a imaginary refractive index above 0). As mentioned in a previous comment, this number might be too low for dust in this region.

Response: In the first version manuscript, the imaginary part of dust was in the range of 0.002 to 0.001. In the revision, we increase the imaginary part to the range of 0.003 to 0.001. Our imaginary part is higher than the imaginary part for the Taklimakan desert dust retrieved by Di Biagio (2019). The strong bias when assimilating the absorption coefficient can be largely removed by tuning the background error of black carbon. Please refer to the revised section 3.5.

802-804. I disagree with this statement. If the model has biases that the assimilation is not able to correct (for instance, inaccurate real and imaginary refractive indexes) then assimilating multiple observation could also create unrealistic modifications to the model.

Response: The statements have been removed in the revision.

806-828. I wouldn't put DA_Esca_Eabs as an improvement over DA_Esca, they show pretty much the same results. This means that the absorption observations are not really generating any differences in the results. Also, DA_PMx_AOD matches better the assimilated variables (which off course is expected) and the better agreement with scattering you happened to underpredict it with PM assimilation, and overpredict it with AOD assimilation, so assimilating both yields you something in between.

Response: Because of the problem of individual assimilation of the absorption coefficient, we remove the DA results of DA_Esca_Eabs and DA_PMx_Esca_Eabs_AOD and just keep the result of DA_PMx_AOD.

832-834. As mentioned earlier, it would be better to check this using normalized profiles. The background profiles already had aerosols up to 4km, so is likely that the assimilation is just scaling this profile upwards rather than adding a larger fraction of the mass in these layers

Response: The revised figure 12 shows the normalized profiles. In the revised section 3.7:

"Also shown in the figure are the vertical profiles normalized to their own respective surface particulate concentrations. The assimilations not only added a larger fraction of the mass in these layers but also adjusted the shapes of the PM₁₀ profiles within 3 km above the ground (Figure 11d), following the BEC's vertical correlation length scales (Figure 3r)."

Section 4.1. I don't think this section is very relevant, the aerosols are so dominated by dust and your BEC is constructed in a way dust aerosols will be the ones largely modified. So just briefly mentioning that the composition of these other aerosols doesn't change would do.

Response: We shorten the revised section 4.1 and present the new anthropogenic aerosols' results.

Section 4.2. Might want to discuss in this section how the large underprediction of dust absorption would impact these results.

Response: It is not easy to quantify the ADRF bias due to the weak absorption with a single WRF-Chem experiment. We admit this uncertainty and add a new paragraph in section 4:

"It is noteworthy to say that the ADRF estimation remains uncertain even after DA. The AOD observation is only sporadically available because of cloud screening in retrieval data. The DA experiments cannot

eliminate the low bias in AOD in WRF-Chem. The ADRF values in the DA experiments are likely to be weaker than the plausible aerosol radiative forcing at Kashi. Neither DA experiment lowers SSAsrf to approach the observation. The observed SSAsrf (0.78) indicates likely warming forcing of aerosol at Kashi, while WRF-Chem and the DA analyses impose cooling forcing. The ADRF uncertainty is associated with the background aerosols. WRF-Chem simulates aerosol size up to 10 μm , whereas larger particles ($>10 \mu\text{m}$) exhibit substantial absorption relative to scattering in the visible wavelength (Kok et al., 2017). Anthropogenic emission inventories need an update for the year 2019, which may reduce the potential low bias in BC concentration. Additionally, the revised GSI does not concern the change in particle effective radius per size bin when calculating the aerosol number concentration in each outer loop. Low absorption cross section rises aerosol number concentration as compensation, increasing aerosol scattering coefficient too much. If our tangent operator concerns the change in particle effective radius per size bin, we can use aerosol mass and number concentration as control variables simultaneously. The DA would have a higher degree of freedom to balance the particle radius and number concentration and improve the absorption coefficient. All these need further research in the future.”