Responses to Referee’s Comments

We appreciate careful reading and lots of valuable comments. We wrote referee’s comments in black, our responses to comments in blue and italics, and the revised manuscript in green.

Anonymous Referee #3:

This paper addresses an important topic: which factors impact the optical absorption of modelled aerosol and how does it affect aerosol forcing calculations. Unfortunately, the paper is very limited in scope. The analysis is performed for a single model and just few factors while many others are not discussed. There is no coherent argument presented that the factors discussed in this study are the most important ones. E.g. wavelength dependence of imagninary refractive indices and wet growth are completely ignored. While the authors present this paper as a sensitivity study it feels more like a small tuning study (and perhaps should be published as a technical paper). However, the authors make use of a very interesting dataset that comprises both species concentrations and optical measurements. This dataset offers the possibility of testing modelled optical properties as a function of chemical mixing state. The authors have done this to some extent but I feel so much more is possible. Authors are also commended for linking changes in optical properties to aerosol radiative effects, although I would like to have seen an analysis of not just Direct Radiative Effect but also Present Day - PreIndustrial forcing. The absence of a proper description of the evaluation methodology for optical properties means I cannot judge whether it is appropriate or not. However, what information is presented suggests that there may be significant sampling errors in their comparison of modelled and observed SSA. On the whole, the paper provides too little substance and may have used improper methodology and I suggest rejection. I encourage the authors to consider expanding their analysis of GEOS-Chem model evaluation using species concentrations and optical properties simultaneously for a future publication.

Specific comments

One paper that deserves mention is Goto et al GRL 2011 as it seems to apply a very similar methodology. Klingmueller et al GM 2014 studies different mixing approximations and their impact on optical properties. Two other papers that deserve more attention are Schulz et al ACP 2006 and Myhre et al ACP 2013. Both describe variations in optical properties across AEROCOM models and are directly relevant to the present work. Myhre et al is quoted in the current paper but in a different context.

➔ Thank you very much for careful reading and valuable comments. There have been many studies investigating sensitivities of aerosol physical properties on AOD and SSA. The main method of our paper is extensively evaluating the model against global surface aerosol

observations and AERONET columnar aerosol optical properties, which is different from previous studies mainly focusing on specific regions or specific optical properties. Furthermore, we examined the effects of brown carbon absorption and dust size distribution, which have been ignored in previous models. Although we used a single model, we think our result can be applied to other models, because our optical calculation tool (Mie theory) has been widely used in global modeling studies and should not be importantly different from the method of other models using Mie algorithm.

Abstract p 2, l 24: please state values of DRE (before & after) as ‘increased’ is ambiguous in this context.

➔ ***We changed the text as follows***

Based on the simulations performed during this study, we found that the global aerosol direct radiative effect was increased by 10 % (from -2.62 W m^{-2} to -2.36 W m^{-2}) after the SSA bias was reduced.

p 3, l 4: Iaconno et al is a strange reference here as its main topic is long-lived greenhouse gases

➔ ***We changed the reference to Jacobson (2001).***

Jacobson, M. Z.: Global direct radiative forcing due to multicomponent anthropogenic and natural aerosols, Journal of Geophysical Research, 106, 1551-1568, 2001.

p 3, l 5: “most contributes to the uncertainty of the aerosol DRF calculations”. Please provide references.

p 3, l 9: “For example, a decrease in the SSA from 0.9 to 0.8” I would not call this a small change in SSA, given that it normally varies between 0.7 - 1.

p 3, l 10: the sign change highly depends on surface albedo which should be mentioned here

➔ ***We agreed the reviewer’s comments and changed the text as follows.***

Among aerosol optical properties, the SSA is the factor that most contributes to the uncertainty of the aerosol DRF calculations, as previous studies have reported based on the analysis of in-situ observations, chemical transport and radiative transfer model results, and satellite data (Loeb and Su, 2010; McComiskey et al., 2008; Srivastava et al., 2011). Furthermore, SSA is a critical factor to determine the sign of the DRF: DRF can be changed from negative (cooling) to positive (warming) if SSA is decreased from 0.9 to 0.8 although the sign change also highly depends on surface albedo conditions (Jethva et al., 2014).

p 4, l 5: “which can cause a wide range of calculated aerosol SSAs” I agree with the authors but it’s important to point out there are differences due to regional variations (unavoidable) and differences due to measurement uncertainties (in principle this allows improvement).

➔ ***We changed the text as follows:***

The application of Mie theory for aerosol SSA calculations requires precise information about the size distribution, refractive index, particle density, and hygroscopic growth factors of aerosols. However, the characteristics of aerosols vary widely depending on the region, source, aging and mixing state, and measurement technique, which can cause substantial uncertainties in SSA calculation. The development of measurement technique can reduce these uncertainties, but it will be still difficult to determine the best aerosol physical parameters for global model usage because aerosol physical parameters vary from region to region. As a result, the input parameters used for aerosol SSA calculations are variable and depend on the corresponding measurements, which can cause a wide range of calculated aerosol SSAs.

p 4, l 21: “variable for the mass median” → “variable in the mass median”

➔ ***We changed the text***

As a result, the physical parameters for BC used in the AeroCom models are variable in the mass median diameter

p 5, l 1: “internally mixed BC aerosol” I believe this depends on the assumed internal mixing. The authors assume either complete mixing or a solid core- liquid shell morphology. See also Klingmueller et al GMD 2014, see also Brunamonti et al GRL 2015, Liu et al Nature GeoSci 2017

➔ ***Thank you for the comment. We added the text as follows***

Both theoretical and observational studies showed that an SSA value of an internally mixed BC aerosol is lower than that of an externally mixed BC (Bond et al., 2006; Drury et al., 2010; Shiraiwa et al., 2008), which also depends on wavelength, particle morphology, and mass ratio of non-black carbon to black carbon (Klingmuller et al., 2014; Brunamonti et al., 2015; Liu et al., 2017)

p 6, l 9: “2_ x 2.5_ for this study” This is then the model resolution at which you preformed your simulations? Please state so in the text.

➔ ***We changed the text as follows***

Meteorological data including temperature, wind, humidity, planetary boundary layer

height, and other meteorological variables at 0.5° x 0.667° horizontal resolutions were re-gridded to 2° x 2.5° which is the horizontal resolution of the model used in this study.

p 6, Sect 2: Please briefly describe the aerosol model: how many tracers do you use for aerosol? Do you use a modal scheme or a bin scheme. Is the modal scheme one or two-moments? How many modes/bins? To what extent is internal mixing explicitly calculated?

➔ *We added the text as follows*

BC, OA, and inorganic aerosols (SO_4^{2-}, NO_3^-, NH_4^+) were simulated using the bulk scheme which simulated only mass concentrations. Number concentrations were calculated using the constant density of each aerosol (Table 2). Size distributions were calculated assuming log-normal distribution with geometric mean radius and standard deviation (Table 2). On the other hand, we used two bins for sea salt and four bins for dust aerosols according to their size.

p 6, l 23: It's not clear if this BrC calculation is off-line or on-line? Do you use the same GEOS-Chem simulaton but add a correction to the fields to describe BrC or is BrC an additional tracer? Or do you simply assume part of organic aerosol is absorbing in FlexAOD?

➔ ***We didn't simulate BrC in this model calculation and assumed some part of organic aerosol was BrC. Because we had 3-D BrC to OC ratio fields from Jo et al. (2016), we applied these fields to this study in order to consider BrC effects. We added the text as follows.***

For the sensitivity study described in Section 3.4, we additionally considered BrC by using the BrC/OC ratio calculated by Jo et al. (2016). Using the 3-D BrC to OC ratio fields archived from Jo et al. (2016), we applied the BrC/OC ratio to the simulated POA and SOA of each grid box in order to calculate the BrC from both primary and secondary sources. The ratio of BrC to OC burden was 0.19 across the globe, but could be regionally variable as shown in Figure 7 in Jo et al. (2016). We applied refractive indices of BrC from Jo et al. (2016) to estimate BrC absorption. The detailed BrC estimation method and the global distributions are described in Jo et al. (2016).

p 8, l 6: Does FlexAOD in principle use the same method as GEOS-Chem to calculate aerosol optical properties? I assume that GESO-Chem calculates optical properties (for radiative forcing) and wonder how the two codes relate.

➔ ***Yes, FlexAOD (if it assumes external mixing) uses the same method as GEOS-Chem does. FlexAOD aimed at extending the model capabilities concerning calculations of aerosol optical properties. We added the text as follows.***

FlexAOD uses Mie theory (Mishchenko et al., 2002) with input data including size distributions, refractive indices, particle density, and hygroscopic growth factors included for each aerosol species, which is the same method that GEOS-Chem uses for calculating aerosol optical properties.

p 8, l 14: Even though a reference is provided to Curci et al 2015, please provide more details on the main assumptions. We are talking about essential assumptions in your study so please state them explicitly. E.g. homogeneous internal mixing includes all aerosol are mixed together (some models allow external mixtures of internal mixtures, e.g. coated BC and coated dust)? Also discuss how these assumptions relate to the GEOS-Chem simulation of aerosol. E.g. a full external mixture and a full internal mixture will show different wet deposition behaviours. I suppose some of your FlexAOD modelling may be inconsistent with GEOS-Chem modelling?

➔ *FlexAOD uses the final mass concentration output from GEOS-Chem. We especially focused on sensitivities of aerosol mixing state on AOD and SSA. We agreed that different mixing state can influence the wet deposition of aerosols, but we thought the change would not be significant because aerosols were easily washed out in the atmosphere if there were rains. We added the text as follows*

In case of internal mixing, two widely used internal mixing representations were used. Homogeneous internal mixing assumed that all aerosols were well mixed, and the core-shell internal mixing assumed that an insoluble well-mixed core was coated by a concentric well-mixed soluble shell. In both assumptions, the refractive index was calculated as the volume-weighted average of the components. Calculating optical properties approximated the integrals for the Mie efficiencies by dividing the size range into 100 geometrically spaced bins, and then calculated the wet volume (hygroscopic growth is considered) concentration of each species in the well-mixed particle, or the well-mixed core and the shell, in each size bin from the sum of all log-normal modes.

In case of homogeneous internal mixing, the Mie efficiencies (extinction, absorption, and scattering) were calculated in each bin for a monodisperse aerosol of radius r using the Mishchenko et al. (1999) code. Then extinction, absorption, and scattering coefficients were calculated summed over the whole size distribution.

In case of core-shell internal mixing, the same averaging procedure was applied separately to the core and the shell. The Mie efficiencies are calculated in each bin for a monodisperse aerosol of radius r and the calculated core-to-shell volume ratio using the Toon and Ackerman (1981) code for stratified spheres. Hydrophilic OA constitutes shell and hydrophobic OA is treated as the core for the internal mixing calculation. BC is assumed as core regardless of its hygroscopicity. Other aerosols including inorganic, dust, and sea salt constituted the shell. A more detailed description of the aerosol optical property calculation with different mixing states can be found in Section 2 of Curci et al. (2015).

p 8, l 23: “updates to the method of Martin et al.” Does Martin et al describe GEOS-Chem or FlexAOD?

➔ Martin et al. (2003) described aerosol physical and optical properties of GEOS-Chem. We added the text as follows

Herein, we briefly discuss the significant updates to the method of Martin et al. (2003), which was the basis of GEOS-Chem aerosol optical property calculations.

p 9, l 15: “which makes it difficult to choose the appropriate BC radius in the Mie calculation” So how did you resolve this problem?

➔ We examined BC absorption versus mean BC radius (Figure 1) and conducted sensitivity studies by varying mean radiuses of BC as shown in Table 3. We found that the mean radius of GEOS_BR_DI_E case showed relatively good results compared to other cases especially for 870 nm - RMSE of GEOS_BR_DI_E case (0.024) was less than that of other cases (0.033-0.041).

p 10, l 15: As far as I understand you do not really consider different morphologies (apart from core-shell) as all optical properties are calculated with Mie theory. Please state this clearly.

➔ We agreed with the reviewer’s comment and removed the sentence about morphologies in this section as follows.

Here we focused on BC because of its high contribution to light absorption.

p 10, l 19: Please define BB, BBR and BBHR.

➔ We defined these cases in Table 3. We changed the text for the clarity.

Refractive index of 1.95 – 0.79i were used for BB, BBR, and BBHR cases with different geometric radiuses of BC as described in Table 3.

p 11, l 6: “We also implemented a new dust size distribution in the model as suggested” Implemented in FlexAOD or in GEOS-Chem? If in GEOS-Chem then please move this to Sect 2. If in FlexAOD, then wouldn’t this imply an inconsitency between GEOS-Chem and FlexAOD?

➔ We agreed with the reviewer’s comment that different size distribution could affect the

burden in GEOS-Chem. However, the mass distribution in submicron bins does not affect the dust mass concentrations in GEOS-Chem (Ridley et al., 2012; Zhang et al., 2013). The mass concentration in supermicron dust bins could be changed, but the submicron dust is more important to AOD because of its high extinction efficiency (Ridley et al., 2012). Therefore, we think this doesn't change the dust burden and AOD significantly, and this change would be a lot less than the uncertainty of dust emissions (Huneeus et al., 2011). We changed the text to clarify this part as follows.

We also conducted the sensitivity test by changing the dust size distribution as suggested by Zhang et al. (2013). Zhang et al. (2013), developed the dust size distribution with constraints from in-situ measurements and the work of Kok (2011), who suggested new dust size distributions based on the physics of the scale-invariant fragmentation of brittle materials. We noted that the change in size distribution only applied to the calculation of dust optical properties, and there was no change in the dust aerosol mass within the GEOS-Chem simulation as per Ridley et al. (2012). The mass distribution in submicron bins does not affect the dust mass concentrations in GEOS-Chem (Ridley et al., 2012; Zhang et al., 2013). The mass concentration in supermicron dust bins could be changed by dry and wet deposition, but the submicron dust is more important to AOD because of its high extinction efficiency (Ridley et al., 2012).

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M. and Dentener, F.: Global dust model intercomparison in AeroCom phase I, Atmos. Chem. Phys., 11(15), 7781–7816, 2011.

p 12, l 18: "was degraded" Degraded compared to what? Earlier studies? Maybe AMS? In that case, say something like "Compared to AMS, the model performs worse with respect to SPARTAN".

➔ ***We changed the text as follows.***

Compared to AMS, the model showed worse results for inorganic aerosols with respect to SPARTAN, especially for nitrate aerosols.

p 13, l 7: "nitrate overestimation" See also recent AEROCOM comparison of nitrate simulation (Bian et al ACP 2017) which shows _ 10x difference amongst models

➔ ***Thank you for the comment. We added the text as follows.***

Recent AeroCom phase Ⅲ study showed that nitrate simulations have the largest diversity among inorganic aerosols (Bian et al., 2017).

p 13, Sect 4.1: Fig 2 & 3 show evaluation of individual species. Can something be said about how well the model simulates chemical mixing? If it overestimates sulfate when it underestimates BC, chemical mixing states are very poorly represented. But if it overestimates both, SSA may be quite accurate.

➔ We agreed that BC underestimation (or sulfate overestimation) leads to uncertainty of SSA calculation because of limited core aerosol amount in internal mixing assumptions although the total aerosol mass concentration is the same. In order to remove this uncertainty, we filtered out observation sites based on mass concentration criteria of each aerosol.

p 14, l 5: "The performance of the model was comparable to that of global PM2.5 estimates constrained by using satellite AOD observations" This sentence is not that clear, can you rephrase? van Donkelaar's PM2.5 evaluation yielded similar results to yours but they adjusted PM2.5 based on AOD?

➔ We removed this part and changed the text as follows.

Simulated $PM_{2.5}$ concentrations were generally in good agreement with the observations (R = 0.76) with a slight low bias observed for the mean concentrations (-24 %). However, the model significantly overestimates or underestimates individual aerosol concentrations in some observation sites. We screened out some poor representation sites based on criteria, which was described in next section.

p 14, Sect 4.2: Please describe how you compare model data with observations. I believe you only have monthly model data so you need to construct monthly AERONET data somehow. According to the caption of Fig 5, monthly AERONET averages were only calculated when there were more than 10 days of observations with AOT > 0.4? In that case monthly model data and monthly observations are based on very different temporal sampling. This will lead to large errors in the comparison, see Schutgens et al ACP 2016 (on temporal collocation) and Schutgens et al ACP 2017. Note that such errors need not be random but can be biases, especially because observations use AOT>0.4 but monthly model data do not. Please discuss.

➔ Thank you for the comment. We think the evaluation results will not be significantly changed (within 0.01 for mean SSA). We added the text as follows.

Here we used monthly mean AERONET and model outputs for 2008 – 2010. We noted

that temporal averaging without collocation could cause significant errors up to 0.05 for SSA in specific region and time (Schutgens et al., 2016; Schutgens et al., 2017). However, mean temporal sampling error caused by using monthly averaged SSA is less than 0.01 (Figure 16 in Schutgens et al., 2016), which is less than the mean model error in this section (~0.05).

p 14, l 17: can something be said about AOD errors as agreement with AMS becomes better and better? Or conversely, can you imagine a plot of AOD error vs AMS error? I wonder if AOT errors are (relatively) smaller or larger than AMS errors. Smaller AOT errors suggest balancing errors, larger errors suggest incorrect assumption on morphology, mixing state etc. I realise this is very hand-wavey but can the authors say something about this?

➔ ***We calculated the mean bias of SPARTAN $PM_{2.5}$ for selected observation sites based on the criteria in Section 4.2. The model underestimated the observed $PM_{2.5}$ by 19%. On the other hand, the model underestimated the observed AOD by 10%. However, it is difficult to make a direct comparison between $PM_{2.5}$ and AOD because AOD is affected by not only surface concentrations but also vertical distributions. Different observed years between AOD (2008-2010) and SPARTAN (2013-2015) also make the comparison difficult.***

p 15, l 1: "From this result, we concluded that the aerosol optical property calculation in FlexAOD was properly performed". I am not sure what the authors mean by this. Clearly there are still errors. There is no proof that your calculation is bug-free. I admit it's a sanity test of your data. But the best you can conclude is that there are no major errors in either your GEOS_Chem calculation, your FLexAOD calculation or your observations.

➔ ***We changed the text as follows***

From this result, we concluded that there were no significant errors in our simulated aerosol concentration (GEOS-Chem) and aerosol optical property calculation (FlexAOD).

p 15, l 8: In addition to Dubovik et al. 2002, see also: E. Andrews at al ACP 2017

➔ ***Thank you for the comment. Andrews et al. (2017) reported that most AERONET SSA values are within reported uncertainty bounds for 440 nm AOD > 0.2. However, we decided to use SSA values for 440 nm AOD > 0.4 for the conservative comparison because SSA uncertainty was further reduced as AOD increased as shown in Figure 3 of Andrews et al. (2017).***

p 15, l 19: “The model showed an improved result especially for the SSA at 440 nm”. For 870nm, I see that R decreases a little bit. There is no info on RMS errors. I suggest removing ‘especially’ and explicitly describe the results for 870nm.

p 15, l 21: “we concluded that the model simulation capability of the AOD and SSA was improved when we selected AERONET sites where the model showed good results against the AMS and SPARTAN network” I think this conclusion is too strong, as there is no improvement for 870nm.

➔ *We changed the text as follows*

The model showed an improved result especially for the SSA at 440 nm. The correlation was increased to 0.5 from 0.34, which was analogous to the evaluated result of the AOD at 500 nm. On the other hand, the model didn’t show improvement for 870 nm SSA. Although there was no improvement for 870 nm SSA, the model showed improved results for AOD and 440 nm SSA when we selected AERONET sites. Hereafter, we used these selected AERONET sites for the sensitivity studies below.

p 15: You have a very interesting dataset of both species concentrations and optical measurements. It is worthwhile to dig deeper into this dataset. How do the aerosol change when you filter the AERONET sites in Fig 5? It appears filtered AERONET sites have less absorbing aerosol. Do they also have less BC? Does SSA agreement improve when chemical mixing is better modelled? You filter by requiring species concentrations to be within 2x of observations but that does not necessarily mean mixing is better. You can also adress questions like: are dust-heavy cases better modelled than BC-heavy cases, etc.

➔ *Observed mean BC concentration was decreased by 10% after filtering observations (2.58 to 2.32 μgC m^{-3}). The normalized mean bias of BC was changed to 3% from -37% (without filtering). SSA values were still overestimated although bias of BC was reduced, which indicated other factors affecting SSA rather than BC concentration.*

p 15, Sect 4.2: One thing I miss in this Section is the realization that you are talking about fundamentally different observations: optical column measurements and surface measurements. This creates uncertainty whenever you interprete SSA model results and needs to be acknowledged.

➔ *Thank you for the comment. Because of this reason, we re-compared AOD after we filtered out AERONET sites. We added the text as follows*

Although we selected column observations (AERONET) based on surface observations (AMS and SPARTAN), the model results were improved for AOD at 500 nm and SSA at 440 nm as discussed below.

p 15, discussion of Fig 5: The most obvious reason why SSA might agree at 870nm but is overestimated at 440nm is of course wavelength dependency of imaginary refractive index. However, the authors do not really discuss this. Note that due to the temporal sampling issue I discussed earlier, I’m not convinced that the agreement at 870nm is undisputed.

➔ The mean biases of SSA at 870 nm were 0.005 and -0.010 for GEOS_E and BB_E cases, respectively. The only difference between two cases was imaginary refractive index. GEOS_E case used -0.44i, and BB_E case used -0.79i. We think this imaginary refractive index difference (0.44i vs. 0.79i) is higher than the refractive index change by wavelength dependency (Kirchstetter et al., 2004), and the SSA bias at 870 nm is less than the bias for SSA at 440 nm. The mean temporal sampling error caused by using monthly averaged SSA is less than 0.01, which is less than the difference between 440 nm SSA bias and 870 nm SSA bias.

p 16, Sect 5: There is a difference between a sensitivity study (as the Section’s title inidcates) and a parameter estimation (as the first paragraph indicates). Please decide which one you want to do and adjust the text accordingly.

➔ The main content of Section 5 was to determine best parameters for simulating global SSA by conducting sensitivity cases. We changed the section title according to this.

Section 5. Parameter estimates for global SSA simulation

p 16, l 18: I still don’t understand how BrC ‘works’. GEOS-Chem simulates organic carbon but assumes zero imaginary refractive index? So in FlexAOD you now assume that part of the organics are actually absorbing? Please describe this earlier in the model or FlexAOD Sections.

➔ Yes, we estimate some part of absorbing organics in FlexAOD calculation. We added the text in the model description section

For the sensitivity study described in Section 3.4, we additionally considered BrC by using the BrC/OC ratio calculated by Jo et al. (2016). Using the 3-D BrC to OC ratio fields archived from Jo et al. (2016), we applied the BrC/OC ratio to the simulated POA and SOA of each grid box in order to calculate the BrC from both primary and secondary sources. The ratio of BrC to OC burden was 0.19 across the globe, but could be regionally variable as shown in Figure 7 in Jo et al. (2016). We applied refractive indices of BrC from Jo et al. (2016) to estimate BrC absorption. The detailed BrC estimation method

and the global distributions are described in Jo et al. (2016).

p 17, l 5: “this remaining gap could be reduced when we consider strongly absorbing BrC” Many reasons can be used to explain this. Why do the authors ignore e.g. wavelength dependence of BC? Your dataset of both species conentrations and optical properties can give you hints as to what is the problem?

➔ We thought the wavelength dependency of the imaginary refractive index of BC was not significant based on Table 4 of Kirchstetter et al. (2004). It changed from 0.71 to 0.77 in the range of 350 nm to 750 nm.

p 17, l 11: The authors never really discuss how they define the radius of BC. Clearly the Mie sphere is a bad approximation for the strands of black carbon observed in nature.

➔ We agreed with the reviewer’s comment that Mie theory is a bad approximation for the BC aggregates. However, in this study, we focused on finding best parameter using Mie algorithm, which has been widely used in global models. Therefore, we tried to improve the model simulation ability using Mie algorithm for the global model use. We added the text as follows.

On the other hand, we found that the correlation coefficients and regression slopes decreased as the geometric mean radius of BC increased. We surmised that the use of a spherule size radius (~0.02 µm) seemed more appropriate than the aggregate size radius (~0.1 µm) for the calculations using Mie theory. It should be noted that the spherical Mie theory is not applicable to BC optical property calculation because of its non-spherical and agglomerates characteristics. However, the use of 0.02 µm could result in the better simulation of SSA when Mie theory is used in global models.

p 17, l 17: Why is BC only contributor? I cannot find information on the wavelength dependence of BrC absorption in the paper.

➔ Yang et al. (2009) reported mass absorption efficiency of BC at 880 nm as 5.9 m^2 g^{-1}, which was higher than that of dust (0.001 m^2 g^{-1}) and brown carbon (0.02 m^2 g^{-1}). We changed the text as follows

Aerosol absorption at 870 nm is mainly affected by BC rather than BrC and dust. Yang et al. (2009) reported mass absorption efficiency of BC at 880 nm as 5.9 m^2 g^{-1}, which was higher than that of dust (0.001 m^2 g^{-1}) and brown carbon (0.02 m^2 g^{-1}).

p 18, Sect 6: It’s a pity the authors do not provide a similar analysis for direct aerosol forcing (DRE Present Day - PreIndustrial). I surmise in that case, dust would not be the main cause of sensitivity but perhaps BrC.

➔ ***We agreed with the reviewer’s comment. However, we mainly focused on present DRE change by SSA correction. Because DRF includes not only the DRE change from present day minus preindustrial condition but also the DRE change by SSA correction, we think the DRE analysis can easily reflect SSA correction effect.***

p 19, Sect 7: I suggest the authors start the Conclusions with a brief recap of their study. It also is better to prevent use of abbreviations like GEOS_BR_DI_E, OP_E, BB_E which will only be obvious to people who have read the full paper.

p 19, l 18, Finding 1: Actually what is critical to SSA calculation is the Mass Absorption Coefficient (MAC), which is the ratio of AAOD by column load (and so includes density). This MAC is still uncertain, see Boucher et al PNAS 2016

➔ ***We agreed with the reviewer’s comment that MAC is directly comparable to AAOD. However, we divided it into imaginary refractive index and density because they are used as input parameters of practical aerosol optical calculations. We changed the conclusion as follows.***

In this study, we used the GEOS-Chem model for aerosol mass concentrations and FlexAOD for aerosol optical property calculation. We heavily evaluated the model with the surface networks (AMS and SPARTAN) and AERONET AOD, and then we conducted sensitivity tests for global SSA calculation at 440 nm and 870 nm. We found that the model with external mixing assumption, brown carbon, and observationally constrained dust size distribution showed the best performance against the AERONET AOD and SSA. However, it is difficult to confirm whether this result can be applied to other studies. There were many factors simultaneously affecting the AOD and SSA calculation such as the hygroscopic growth factor, which was not discussed in detail. However, we were able to identify some notable findings that could apply to future modeling studies when using Mie theory for the calculation of aerosol optical properties:

1. There were no significant differences between the calculated BC absorptions using input parameters from the OPAC database and data from Bond and Bergstrom (2006). The refractive indices used by Bond and Bergstrom (2006) (1.95 – 0.79i) were higher than those used by OPAC (1.74 – 0.44i), and the resulting BC absorption of Bond and Bergstrom (2006) was 49 % higher than that of OPAC. However, the low BC particle density of OPAC (1.0 g cm^{-3}) increased the BC number concentration and the AOD, which enhanced the BC absorption. We found that the mean SSA using OPAC refractive index (0.927) was slightly lower than the mean SSA using refractive index from Bond and Bergstrom (2006) (0.934) (Table 5). Although many previous studies did not provide the

BC density they used for the Mie calculation (Feng et al., 2013; Jo et al., 2016; Lin et al., 2014), an assumed BC density was critical to the SSA calculation and should be provided in future inter-comparison studies.

2. The model using the external mixing assumption showed better performance than the model using internal mixing assumptions in a global modeling sense. Drury et al. (2010), also showed similar results against aircraft observations. The internal mixing assumptions caused very high absorption and an underestimation of the AERONET SSA. The effect of internal mixing on absorption could be overestimated in the model calculations; Cappa et al. (2012) reported that the observed ambient BC absorption enhancements obtained by internal mixing were small (~6 %), which is less than predicted from theoretical calculations and observed from laboratory measurements. Furthermore, BC absorption enhancement by internal mixing could vary by region (Liu et al., 2015). Coating thickness could also be different due to aging timescales and the distance of source regions. Further observational and modeling studies are needed especially for considering region-specific absorption enhancement.

3. The modelled SSA calculations from previous studies were mostly evaluated between 440 – 550 nm (Dai et al., 2015; Jo et al., 2016; Lin et al., 2014). However, SSAs at both shorter and longer wavelengths should be evaluated together for the model evaluation. The calculated SSA could show different performances at different wavelengths because they are affected by different aerosols and wavelength-dependent physical characteristics.

4. BrC absorption and observationally constrained dust size distributions should be considered for the SSA calculation in the model, especially for the shorter wavelengths. These factors reduced the SSA at shorter wavelengths, resulting in a decrease of the positive SSA bias at 440 nm generally found in global models.

5. The changes of dust size distribution led to a significant increase of the global DRE, whose value (0.17 W m-2) was more than two times higher than the global DRE increase by BrC absorption (0.07 W m-2). Global models likely underestimate the global dust DRE because of the overestimation of fine mode dust and the underestimation of the coarse mode dust as shown by Kok et al. (2017), who constrained the global dust size distribution using global observations and models.

p 20, l 24: “resulting in a decrease of the positive SSA bias at 440 nm generally found in global models.” The authors have only shown this to be the case for one model (GEOS-Chem)

➔ ***We changed the text and added references as follows.***

These factors reduced the SSA at shorter wavelengths, resulting in a decrease of the positive SSA bias at 440 nm found in some global models (Lin et al., 2014; Jo et al., 2016).