We thank both reviewers and the editor for their constructive comments, which has greatly helped us to improve the paper. Below we include a point-by point response to the referee comments, and describe the corresponding changes we have made to the manuscript

Comments from reviewer #1

General Comments

A new dataset is presented and analysed to provide a global 3-D seasonally resolved dataset of dust properties with constrained size distribution, shape and refractive index properties. The dataset (DustCOMM) provides annual and seasonal mean 3-D dust size distributions, 3-D mass extinction efficiency at 550 nm wavelength and 2-D dust column mass loading. Since models are mostly unable to reproduce these properties in faith with measurements, this is a welcome and valuable step forwards in the dust field. The authors show that the DustCOMM results perform significantly better than global model simulations when evaluated against independent measurements. This is a notable achievement considering the challenges involved in representing global dust fields. The authors outline the benefits and potential future uses of the DustCOMM dataset to the community, which are likely to be substantial.

This paper covers a considerable amount of work and therefore methodology in order to deliver the final results in the DustCOMM dataset. Although the authors do a credible job of explaining the long and complicated methodology, in places it requires further explanation and clarity, potentially with re-ordering of some sections. The results impeccably described and presented. The description of measurements used requires some corrections and clarifications. Although the authors provide the datasets online, one of the links is broken. The article fits the scope of ACP and I recommend publication subject to the corrections detailed below.

We thank the reviewer for the constructive and helpful comments that helped us to further improve the paper. We address these comments below.

General Scientific Comments

1. Overall, it is not clear why there is a need for the AOD reanalyses in this study – i.e. column 3 in Figure 1. Given the volume size distribution, and an assumed dust density, it should be possible to calculate column mass loading directly from size distribution without the requirement for AODs, and without any uncertainty involved in the mass extinction efficiency (MEE) values. Currently this is not explained, and therefore the AOD reanalysis section (fig 1 column 3 and sections 2.3) seems superfluous. More details are given below.

We thank the reviewer for this comment. We note that our dust size distributions are normalized such that the integral over the diameter range is unity, and with no information about the size-resolved volume of the particles. Therefore, the normalized dust size distribution cannot be used directly with dust density to calculate the column dust loading. We have added the following sentence in section 2.3.3 (formerly section 2.3) to clarify this point: "Since our constraints on dust size distributions are normalized to unity, and also to ensure that our estimates of dust loading produce the same extinction as those from reanalysis dataset or satellite measurements, we use this approach to estimate the atmospheric dust loading..."

2. The authors do not include any dust properties in the longwave (LW) spectrum. The impact of the coarse particles on the radiative budget is one of the motivators of this study, and indeed highlighted in the discussions and conclusion as one of the benefits of the new dataset. The LW radiative effect makes up a large part of the total change in radiative effect due to a better representation of coarse particles (Kok et al., 2017). Thus the omission of LW dust properties here detracts slightly from the novelty of this work. There may be valid reasons for excluding the LW here, such as scope of material and lack of MEE observations in the LW spectrum. However, this should be discussed, since the total radiative impact of the constrained size distribution cannot be calculated without dust LW properties. Additionally, there may be some limitations in radiative studies which could be done with the DustCOMM data due to MEE being required spectrally (at least on some spectral resolution), rather than just at 550 nm as provided.

The reviewer raises an excellent point here, as indeed a large portion of the uncertainty in dust radiative effects is due to uncertainty in LW interactions. That being said, we provide dust optical properties at the mid-visible wavelength (550 nm), rather than in the longwave spectrum, for several reasons. First, as the reviewer pointed out, measurements of dust mass extinction efficiency (MEE) in longwave are scarce and therefore it will be difficult to validate our constraints on dust MEE. Second, the estimation of dust mass loading in Eqn. 9 requires MEE value at 550 nm which is the same wavelength as the observational constraints on dust optical depth (e.g. Ridley et al., 2016). Third, the measurements of dust refractive index needed to constrain the single-particle extinction efficiency at the longwave spectrum are also scarce, potentially leading to large uncertainties in constraining the dust MEE. Despite these reasons, our future studies will focus on incorporating the available measurements of dust refraction index – such as those from Di Biagio et al. (2017, 2019) – to constrain the spectral dependence of dust optical properties for both longwave and shortwave spectra as part of DustCOMM dataset.

To clarify these reasons for not including dust properties in the LW spectrum in this paper, we have included a paragraph in section 2.3.2. "We use this constrained globally-averaged \widehat{Q}_{ext} to constrain $\widehat{\epsilon}_{\tau}$ (Eqn. 8) for every location. We thus neglect any regional variation in \widehat{Q}_{ext} because measurements of dust shapes and index of refraction are currently insufficient to constrain $\widehat{\epsilon}_{\tau}$ on a regional basis. In addition, since measurements of dust refractive index needed to constrain $\widehat{\epsilon}_{\tau}$ at other wavelengths are also scarce, we

limit our estimate here only to the 550 nm wavelength. We use 550 nm as the wavelength of choice because measurements to validate our estimate of \hat{e}_{τ} and the observational constraints to estimate the dust atmospheric loading are mostly available at mid-visible wavelength."

3. Section 2.1 (Constraining the 3-D atmospheric size distribution) is a crucial part of the paper, but difficult to follow. This section needs further explanation and clarity – see specific comments below.

Thank you for pointing this out. Based on the specific comments by the reviewer, we have reordered both sections 2 and 3 to make it easier for the reader to understand. For example, in addition to other specific changes below, we now have the description of the model simulation and reanalysis datasets before the methodology that constrains the dust size distribution, mass extinction efficiency and the atmospheric loading.

Specific Comments

1. Links to datasets – the first asset link to the DustCOMM dataset v1 (https://doi.org/10.5281/zenodo.2620475) is broken.

Thank you for pointing this out. The link has been fixed.

2. Abstract -117-21 – the use of model simulations should be mentioned here as this is a crucial part of the work.

Thank you, that's a good point. We have re-written this part to reflect the use of global model simulations in constraining the dust size distribution. The new sentence now reads as:

"This dataset leverages <u>an ensemble of global model simulations</u> with observational and experimental constraints on dust size distribution and shape to obtain more accurate constraints on three-dimensional (3-D) atmospheric dust properties than is possible from global model simulations alone."

3. It is not entirely clear what the benefits of model-constrained data presented here are over the data used in Kok et al. (2017). It would be useful to the reader to make this crystal clear, probably at the end of the introduction.

We have included a sentence at the end of the introduction to clarify the difference between the Kok et al. (2017) results and the results in the study. We added the following to the last paragraph: "DustCOMM builds on the results from Kok et al. (2017), however, unlike the globally-averaged results obtained in Kok et al. (2017), our product constrains the climatology of 3-D global atmospheric dust properties on seasonal and annual timescales".

4. All the way through the article, but particularly in the method, the authors should be absolutely clear which size distribution they are referring to - the Kok 2017 globally constrained size distribution, or the model-simulated size distribution(s) – when they state 'globally averaged size distribution,' or similar generalizations.

Thank you for the comments. We have clarified this confusion where necessary.

5. Section 5.4 – would you expect DustCOMM to show improved inter-annual variability in dust loading compared to conventional models, if more than one year of data were constructed? Or would DustCOMM's ability in this context be limited by the underlying global models' limitations? This is another significant challenge for models, e.g. Evan et al. (2014).

Thank you for the question. While we have not looked at this topic yet, we expect that the inter-annual variability in DustCOMM dust size distributions largely depends on the ensemble of model simulations used. However, the interannual variability of the dust mass loading will include not only the variability from the ensemble of model simulations but also from the observations used to constrain the dust aerosol optical depth and dust extinction. Although this study only present seasonal variability, DustCOMM inter-annual variability is expected to show some improvements when compared to individual model simulation.

Section 2.1 – p4 l21-35 – These lines cover methodology and results from Kok et al. (2017). While necessary and useful to repeat here, I suggest making it very clear that up to 1 35 this is a repeat of method and size distribution from Kok et al. (2017).

To improve clarity for this section, we have added an extra sentence before Eqn. 1, and re-write the paragraph following the same equation.

"While details can be found in Kok et al. (2017), a summary of their globally-averaged size distribution is given here as:"

"As reported in Kok et al. (2017), the constrained globally-averaged size distribution of emitted dust particles, $\left[\frac{d\hat{v}_{emit}(D)}{dD}\right]_g$, is based on an analysis of different measurements of the emitted dust size distribution, while the size-resolved globally-averaged dust lifetime, $\left[\tilde{T}(D)\right]_{\sigma}$, is based on an ensemble of global model simulations"

7. Section 2, p5, 11-17 – this section forms a crucial part of the method – the main correction/constraining of the model size distributions – yet it is difficult to follow. The authors should explain this section much more clearly and in detail, perhaps with additional figures in the supplement and/or outlining a specific example, to clearly illustrate how the PSDs are forced away from the model simulations. Some specific comments are given in the points below, but I suggest they review and rephrase these lines.

We have re-written the paragraph before Eqn. 2 to better clarify the procedure. After Equation 2, we have also made additional clarifications to the sentences, including clearly explaining the dimensions x, y, and z. For this part, we added "...;x is the dimension for longitude, y is for latitude and z is for height."

8. It is not very clear at which point annual averages vs spatially varying size distributions are used from the simulations. E.g. p5 15 – "annually-averaged, globally-averaged" size distributions are forced here – how/at which point are the spatially varying size distributions corrected? Better signposting of global averages vs spatial variations throughout the paper would be extremely useful in understanding the methodology.

We made this distinction by specifying in each formula the independent variable of each parameter. For example, $\hat{f}_{k,i}(x, y, z, D_{k,i})$ in Eqn. 2 is a spatially-varying parameter defined for x-longitude, y-latitude and z-height, while the $\left[\tilde{\tilde{f}}_{k,i}(D_{k,i})\right]_{\sigma}$ is the globally-

averaged counterpart denoting that the parameter is averaged over all space (x-y-z). To make this clearer, we have explicitly defined the independent variable following its first use in Eqn. 2., adding "x is the dimension for longitude, y is for latitude and z is for height"

9. P5 19-10 "to the global dust loading" – which global dust loading – simulated or from Kok 2017?

The sentence has been re-written for clarity. It now reads "This correction factor (α) is defined by the ratio of the Kok et al. (2017) constraint on the fractional contribution of the particle bin to the simulated fractional contribution of the particle bin per unit global dust loading."

10. P5 11-17 – the description of equation 2 is not clear enough. E.g. what is the numerator in the equation for alpha?

The numerator is the fractional contribution of the Kok et al. 2017 globally-averaged size distribution defined between diameter $D_{k,i+}$ and $D_{k,i+}$. In addition to the description of α before Eqn. 2, we have added the following statement to improve the clarity after the

equation.: "the numerator, $\int_{D_{k,i-}}^{D_{k,i+}} \left[\frac{d\hat{V}(D)}{dD}\right]_g dD$, is the constraint obtained from Kok et al. (2017)"

11. P5 116-17 – "where the discrete sum over each location and height equals unity, that is:..." – Why is the sum unity? Is this because the size distributions are normalized?

Yes, that's indeed correct. This is a key point, so we have also rewritten this for clarity.

 Fig S1 – Please include the Kok 2017 globally averaged constrained size distribution on each panel. This would enable the reader to see what the size distribution is being constrained by.

That's a good idea. Since DustCOMM is forced to the Kok et al 2017 size distribution, both DustCOMM and Kok et al 2017 are the same and thus overlap in the Figure S1. Although not always visible, we have nonetheless included the Kok et al 2017 size distribution in Fig. S1 for completeness (see below).



Figure S1 – please also include (either in this plot, or as a separate one), how the size distributions changed due to the re-binning/extension/curtailing, as described in Section 2.1.1.



We have included the figure below in the supplementary document

14. Section 2.1.1 – Is this diameter range correction performed after the size distribution correction (section 2.1)? Figure S1 suggests that first the diameter range correction is performed, and then the size distribution correction. But the ordering of the text (2.1 – size distribution correction, 2.1.1 – diameter range correction) suggests the opposite. Please clarify and order the text appropriately to follow the steps in the method.

This is indeed confusing. The order described in section 2.3.1 (formerly 2.1) does follow the steps taken in the methodology. That is, the dust mass fractions were corrected and then the size ranges were set to the common diameter limits. To make this clearer, we have added a statement that the dust mass fractions were re-normalized after the correction, such that the discrete sum still equals unity over each location. This statement at the end of section 2.1.1 reads: "After the dust mass fractions are corrected, they are renormalized such that the discrete sum between D_{min} and D_{max} equals unity over each location and height."

For the 'uncorrected' model dust mass fractions, we highlight in section 2.1 (formerly 3.1) that, for consistency, they are also set to the same diameter limits between 0.2 and $20\mu m$ following the same procedure described in section 2.3.1.1 (formerly 2.1.1).

15. P6 11-3 – Ryder et al. (2019) show that over the Sahara D>20μm contribute to at least 18% of SW extinction and 26% of LW extinction – these values are not negligible and represent aged Saharan dust.

P6 14, "dust particles with D>Dmax generally stay only for a short period in the atmosphere before they are deposited" – this is not the case in van der Does et al. (2018), as stated elsewhere in this manuscript.

That's a good point – those two recent papers indicate that the community has greatly underestimated the effect of dust with D > 20 um. In this study, we need to limit the maximum diameter to 20 µm because results from Kok et al. 2017 used to constrain our dust size distribution is limited to this diameter. But to better acknowledge that particles greater than 20 µm might be important even farther from source regions (e.g. Ryder et al. (2019)), we have re-written this part of the paragraph:

"Further, we set the upper diameter limit to $D_{max} = 20 \ \mu m$, because most global models generally do not incorporate dust particles beyond 20 μm and also because the observational constraints on the size distribution from Kok et al. (2017) is limited to this maximum diameter. Although advances in airborne observations in recent years have led to measurements of larger dust particles with $D > D_{max}$ in the atmosphere which has shown that the contribution of $D > 20 \ \mu m$ to shortwave and longwave extinction are non-negligible (e.g. Ryder et al., 2013b, 2019; Weinzierl et al., 2009, 2017), there is still a scarcity of these measurements, such that an observational constraint on dust particles with $D > D_{max}$ would be very uncertain (e.g. Mahowald et al., 2014)."

16. P5 11-6 – the authors should revisit the impact of particles d>dmax in the discussion (e.g. Section 5.4). For example, if better global constraints/observations on this size range became available, could such observations be incorporated into DustCOMM? \

Yes, that's a good point. We have added the sentence below in section 4.4 (formerly 5.4) to clarify this point.

"Given that dust particles with $D \ge 20 \ \mu m$ can contribute substantially to dust extinction both in the shortwave and longwave spectrum (Ryder et al., 2019), future versions of DustCOMM could be extended to a diameter range beyond 20 μm as more measurements of dust size distribution with $D \ge 20 \ \mu m$ become available."

17. P6 130-32 – the authors essentially extend 4 models' size distributions towards a larger size range based on the other 2, which cover a wider diameter range. Does this implicitly assume that all the models behave the same way in terms of the coarse end of the size distribution? This seems unlikely. This should be discussed more, particularly since many of the results are most sensitive to the size changes above d=10 microns

No, our constraint does not implicitly assume that the models behave the same way. Since dry depositions in these models are controlled largely by gravitational settling, the rates of deposition of the dust particles are often different. Our constraints thus account for these differences by taking into account the differences in the spatial variability of the bin that overlaps between the two models. That is, in correcting for model simulation k in Equation 4a (for example), we account for the difference between the bins $[D_{k,N_k-}, D_{k,N_k+}]$ and $[D_{r,j_r-}, D_{r,j_r+}]$ in model r, which partially overlap with each other (Eqn. 4b). Although the correction factor $\beta_r(x, y, z)$ is expected to take into account the differences between model k and model r, we however still assume that the distribution largely follow the same form controlled by the rate of dust deposition in each model. We have added the sentence below to clarify this point.

"It should be noted that the correction of Eqn. 4 takes into account the potential difference in the dust deposition between models k and r, by considering the differences in the spatial variability of dust loading between similar bins of $[D_{k,N_{k}-}, D_{k,N_{k}+}]$ and $[D_{r,j_{r}-}, D_{r,j_{r}+}]$."

18. Equation 5 – please state how/if this equation is different to that from Kok et al. (2011), and if so why.

Despite the similarity in Equation 5/6 and that from Kok 2011, an important difference is that Kok 2011 describes the size distributions at emission while Eqn. 5/6 describes the size distribution in the atmosphere. Because of that, our formulation in Eqn. 5/6 builds on the brittle fragmentation theory of Kok 2011, but adds analytical expressions of dust deposition and dust changes during transport. In addition, the generalizations of parameters in our equation also allow us to better fit different shapes of dust mass fractions over different locations, and thus able to place a better constraints on dust size distribution.

To clarify this point, we have added the following section at the end of section 2.3.1.2 (formerly 2.1.2): "Finally, we note here that although our generalized theoretical function of Eqn. 6 builds on the brittle fragmentation theory of Kok 2011, it adds analytical expressions of dust deposition and dust changes during transport that allow us to better fit different shapes of dust size distribution over different locations."

19. Equation 5 & p8 l6 – why choose D_s for the geometric median diameter by volume? S subscript typically implies with respect to surface area. D_v would be more appropriate.

Thank you for this comment. We have changed the subscript from s to v

20. P8 120-27 – This seems a great generalization. It's not clear how b is applied to specific locations as implied.

The application of all the parameters in Equation 6, including the parameter b, was done by fitting the expression using Equation 7 for each location. As a result, values for b could be different from one location to the other, based on the shape of the corrected dust mass fraction from Eqn. 2. The distributions of these parameter, including parameter b, are included in supplementary Fig. S-2. To make this clearer, we have reworded the paragraph before Eqn. 7, and also added additional sentence to the paragraph after Eqn. 7.

"To determine the parameters in Eqn. 6 for each height, horizontal location, season, and model simulation, we fit the generalized size distribution of Eqn. 6 to the corresponding corrected dust size distribution from Eqn. 2 above. To do this, we minimize the chi-squared (χ_k^2) value for each height, location, and for each model k, such that:"

"The probability distribution of these parameters for all heights, horizontal locations, and model simulations of the annually-averaged dust size distribution is shown in the supplementary Fig. S-2"

21. Section 2.1.2 – What is the reason for choosing this method of fit (eqn 6) as opposed to fitting a series of lognormal modes, as is typically done for size distribution measurements? Presumably given the simulated size distributions have been corrected based on the same function, the fitting of the corrected size distribution is more naturally aligned with eqn 6?

That's a good point. We use Eq. (6) because it includes some mechanistic understanding of what determines the functional form of dust size distributions. Furthermore, although lognormal modes are appropriate for several other aerosol species, dust size distributions generally do not follow lognormal distributions very well, in part because the emitted dust size distribution is distinctly lognormal, as detailed in Mahowald et al. (2014). We have added a sentence to clarify this point in the first paragraph of section 2.3.1.2: "Although fitting lognormal modes are appropriate for several other aerosol species, Mahowald et al., (2014) highlighted that dust size distributions are usually not lognormal and are thus better characterized by a generalized function based on mechanistic understanding of dust emission and deposition processes. "

22. P10, section 2.3, l26-30 – units for all quantities would be helpful. What do the authors mean by "mass-weighted" in "mass-weighted vertically-integrated 2-D mass extinction efficiency" and what are the units of epsilon_tau and epsilon_m?

The units of each variable have now been included. And this phrase is indeed confusing as we meant "vertically-integrated 2-D mass extinction efficiency", thus the "mass-weighted" has been removed.

23. P10, section 2.3, l26-30 – I believe this calculation is the same as first used by Kaufman et al. (2005), which should be cited.

Good point. Kaufman et al. (2005) has now been cited.

24. P10 131-32 – please list the reanalysis products (MERRA-2 etc) here to avoid confusion. Also see later comment about section ordering of 3.2. "Dataset" should be 'datasets.'

Thank you. This comment has been addressed in the paper.

25. P10 l32-p11 l2 – "This individual reanalysis dataset...." - I suggest removing this (and adding to section 3.2 if necessary). It is confusing here given that the AOD reanalyses have not yet been described.

P11 110 "the four data sets..." – this is also confusing given that the AOD datasets have not been properly introduced at this stage in the paper. See later comment about relocating section 3.2. Ordering and section 3 - I suggest the authors move sections 3.1 and 3.2 to before section 2. This would be easier to follow and understand. Section 3.3 should remain after section 2 since it follows on logically.

This is a great idea to improve the paper's clarity, thank you. As suggested by the reviewer, we have combined section 2 and 3 and reorder the subsections, such that the description of model simulations (section 2.1) and the reanalysis products (section 2.2) comes before the description of the DustCOMM products (section 2.3).

26. P16 135 – 2011-2015 is presumably limited by available years? Is there any impact of this difference in years used?

Yes indeed, the dust AOD for JRAero is only available between 2011 and 2015. Although we did not analyze the impact of the difference in DAOD climatology between JRAero and the others, we expect that the relative difference will be smaller over the dust dominated region.

27. P17 l32 – and is also a 2-D diameter project of a 3-D shape, which may introduce bias (e.g. Chou et al., 2008).

Thank you. We have reworded the relevant sentence of the 3rd paragraph of the section to better reflect this.

"during the microscopy analysis, particle diameters are usually determined as the volume-equivalent geometric diameters based on 2-dimensional images (Chou et al., 2008). Because of the asphericity of dust aerosols, this could introduce some biases (e.g., Okada et al., 2001; Huang et al., in prep.)."

28. P17136 "separate channels for different particle sizes" - this is not really relevant and could introduce confusion.

Thank you. This part of the sentence has been removed.

32. P17 125-40 – there is a 4th category, which covers imaging probes, as used in the AER-D field campaign (section 4.18 of supplement) – which are beneficial since they do not suffer from uncertainties in converting scattered light to size as OPCs do.

That's a good point. We have included two sentences to highlight this point. First sentence is added in the first paragraph stating: "Another category is the imaging probe whereby the particle image is detected by linear photodiode array providing a two-dimensional projection of the particle (Baumgardner et al., 2017; Ryder et al., 2018)."

The other is added to second paragraph of section 2.4.1 stating: "Unlike the optical particle counters that require assumption regarding dust refractive index and shape to convert scattered light intensity to particle size, the imaging probes are not subject to these uncertainties (Baumgardner et al., 2017; Ryder et al., 2018)"

33. P18 122-25 – OPCs have other sources of uncertainty – such as refractive index applied in the inversion of scattered light to size and the non-monotonic relationship between scattered light and particle size. These should also be mentioned.

We have included a sentence to clarify these points.

"In addition, optical particle counters also make assumptions about the refractive index to derive the dust size distribution, and are affected by the non-monotonic increase in the intensity of scattered light with particle size (Ryder et al., 2018; Weinzierl et al., 2011)."

34. P18 111-25 and section 3.3.2 – in-cabin measurements are also subject to uncertainties and size-bias in sampling due to aircraft inlets. As such, the MEE values from studies in table 2 are likely biased high in some cases.

Thank you for this comment. We have added a sentence in this paragraph to highlight this point. "For in-cabin measurements, studies have shown that the loss rate of coarse dust particles can be substantial due to the aircraft's instrument inlet, therefore leading to lower sampling rate and size bias (e.g. von der Weiden et al., 2009)."

35. P18 114 – although the size distribution measured does not allow aerosol type to be distinguished, various chemical composition measurements made in parallel are now mostly a matter of routine during airborne campaigns. Individual studies often use these to infer size distributions or ranges dominated by different aerosol types.

We thank the reviewer for this comments. While measurements of chemical composition helps to isolate dust from non-dust particles, there are still potential for mis-identification. We further address the discrepancy between measurements and our results in section 4.1

36. Section 3 – there is a huge variety of measurement data available, and I do not suggest the authors attempt to significantly widen their coverage. The authors should describe how and why the studies in Table 2 were selected. There also appears to be a geographical gap of sampling Arabian dust (Fig S1). Additionally, I suggest the studies of dust sampled during the AMMA airborne missions (Formenti et al., 2011) as being a very useful addition, since they provide summertime sampling in the Niger region, which is currently not covered by the studies in Table 2.

We thank the reviewer for this comment. For the most part, the studies we selected for dust size distribution are those that reported actual measurements of coarse dust particles and not log-normal fit or parameterized distributions. The dust size distribution reported in Fig. 10 of Formenti et al. (2011) is a log-normal fit to measured data, such that we cannot use this data. We have included additional sentences in the second paragraph of section 2.4.1 that clarify how and why some of the studies in Table 2 are selected.

"For the dust size distributions, our criteria for selection of studies are as follows: (1) the measured size range of the data should extend into the coarse dust (D > 5 um) size range; (2) the study should report the original in-situ measurements, instead of (lognormal) fits to the actual measurements; and (3) each study's measurements should be taken with commonly-used instrumentation in order to ensure some consistency with measurements taken by other studies. "

37. Table 2 – please indicate which studies relate to which numbers on the map on Figure S3.

We now clarify this in the caption of Fig. S-3.

38. Section 3.3.2 and Supplement section 4 – The descriptions of data taken from each measurement campaign are too vague, and occasionally in error. Often it is not enough just to reference a paper as within the measurement papers observations are collected/averaged in different ways (time periods, meteorological regimes, altitudes, etc.) and it is not clear which are being used here. The authors should state specifically which data are taken from each paper, and what the values of MEE or MSE are, preferably listing them in a table in the supplement. Specific comments about data described in the supplement are given in the Supplement section.

Thank you for the comment. We have made several additions to the supplementary section 4 to address this point. In addition, we also state in Table 2 where the data are taken from, by referencing the specific figure or table where applicable.

39. P18 l34-35 – Ryder et al. (2013b) SSA values fall well outside this SSA range. This is a fairly narrow SSA range selected. The authors should note that measured SSA is sensitive to the size range sampled in the observations, which is likely to exclude the coarse mode and often d>~2- 3 microns in many cases due to the effects of inlets. Only 3 studies are

cited, while there are a huge variety of studies in existence which have measured dust SSA.

We thank the reviewer for pointing this out. We have increased the uncertainty range to 0.03 and have also included citations to other studies.

40. P19 19-11 – "These errors include errors due to the instrument measuring the extinction coefficient" – change to 'instrumental uncertainties.' "meteorological influence" – such as? "the assumption of internal or external mixing" – how is this important?

For clarity, we have removed this sentence, and added "including instrument uncertainties" to the preceding sentence.

41. P20 113-17 – Field campaigns additionally often sample a variety of cases which are representative of the within-season variability, and also often include uncertainties/ranges to cover the variability encountered.

That is right. Thank you for pointing this out.

We have added a sentence to highlight this point: "Furthermore, most of these measurements are campaign averages often over a variety of cases that could be representative of the season-averaged size distribution."

42. P21 12-3 "(1) the ACE-2 campaign (June/July, 1997) off the west coast of Western Sahara and Morocco (Otto et al., 2007)" – would be better referred to as in the vicinity of the Canary islands. Same for caption of Figure 3.

Thank you for pointing this out. We've corrected this accordingly.

43. P21 16 "(2) the Fennec project (June 2011) between the Canary Islands and Mauritania/Mali (Ryder et al., 2013)" – If the authors refer to measurements between the Canary Islands and Morocco/Western Sahara (not Mauritania/Mali which are inland) the citation should be Ryder et al. (2013a – GRL) and the geographic references corrected. The same applies to the caption of Figure 3.

Thank you for this comment. We realize the confusion and we have clarified it in the paper. Data from the two Ryder 2013 papers are used in our study. Here, we do in fact mean the Ryder et al. (2013 - GRL) which is represented in the paper as Ryder et al. (2013a). We have adjusted the text, the geographical reference and images accordingly. The representative location is now placed at 27.65N, 14.25W.

44. Figure 3 – what is the reason for the selection of altitude choice? It seems biased very high – presence of dust at z>6km is unlikely and concentrations will be very low at

5.5km – therefore the value of such high altitude comparisons is questionable. What is the reason for the selection of these 3 studies for Figure 3? The geographic spacing is very close, with all sampling JJA SAL dust. "ACE-2" in line 7 of the caption should read "AER-D."

Thanks for the correction. ACE-2 has been changed to AER-D.

We identify these 3 studies to show that DustCOMM performs better than model simulations for a range of heights. We recognize that dust concentrations are lower for z>6km, but we follow Fig. 1 of Ryder et al. (2013 – GRL) which shows that there are still some coarse dust particles at ~6km. To address this comment, we have added additional comments in Table 2, and a justification for the height selection in the supplementary document.

45. P22 11 – 14% and 15% - in terms of which variable?

This is in terms of the dust mass fraction – i.e. the fraction of dust per unit mass of dust loading. We have clarified this in the paper. The sentence now reads "On average, simulations in our model ensemble overestimate the dust mass fraction of the fine mode by $\sim 14\%$, and underestimate that of the coarse mode by $\sim 15\%$."

46. P23 125 - and also Qinghai Province China?

Yes, good point. We have added this to the text: "... and Qinghai Province, China (Li et al., 2000)"

47. P24 117 – "weighted by the dust vertical distribution" – why is this necessary?

We agree with the reviewer that "column-integrated dust MEE" explains our point in that sentence, and "weighted by the dust vertical distribution" is indeed not necessary. As a result, we have removed it.

48. P24 131-32 – as stated earlier, it is not clear why the MEE and AODs need to be used to calculate the column mass loading, given that this is typical model output, and the size distributions are already available. It should be a direct step to calculate column mass loading from size distributions, given a dust density.

See our response to similar comments above (General Scientific Comments #1)

49. Section 5.1 – are there any impacts of uncertainties in wet deposition on the size distribution biases?

Uncertainties in dust deposition broadly affect our estimates of dust size distribution, as we discuss in the fourth paragraph of the new section 4.1 (formerly 5.1).

50. P27 130 – Fig 7a does not show MEE.

Thank you. This is noted and has been corrected.

51. P27 131 – there is no figure S7 in the supplement.

Thank you. This is noted and has been corrected. We meant to say supplementary Fig. S-6

52. Section 5.4 – the implications of dust LW properties should be reflected on here, considering the points about the LW radiative impacts of dust being crucial to the total impact on the radiation balance described above.

We thank the reviewer for this comment. We have now added a sentence in section 4.4 (formerly 5.4) that emphasizes the use of our dust size distribution for longwave radiative impacts. "With improved constraints on the dust size distribution and therefore the dust optical properties, DustCOMM could be used to determine the dust (shortwave and longwave) heating rates in the atmosphere more accurately than possible with current global model simulations. As a result, our constraints on dust size distribution could be used to better quantify radiative effects of dust, especially in the longwave spectrum which have remained very uncertain (Di Biagio et al., 2017; Dufresne et al., 2002; Kok et al., 2017; Song et al., 2018)"

53. P30 120 - should 'indirect effects' be 'semi-direct effects'?

Good point. We have included both indirect and semi-direct effects.

54. P30 128 – some reference to the SW spectrum and 550 nm should be included, since refractive index and MEE are only considered at this wavelength.

We have clarified that the DustCOMM mass extinction efficiency is at 550 nm. Since our constraints is taken from Kok et al. 2017, we have also included that reference in the sentence.

55. Section 5 – There is a general focus on in-situ observations for validation of DustCOMM. However, remote sensing observations are developing rapidly and it would be useful for the authors to consider whether lidar retrievals, for example, would be usable within the DustCOMM framework. Remote sensing observations are certainly useful as we continue to develop DustCOMM. As we stated in section 4.4, we hope to incorporate more observational constraints within DustCOMM framework. For example, to constrain the vertical distribution of the atmospheric dust loading, lidar-based retrieval of dust extinction, such as from CALIPSO, will be very useful, and it is part of future work.

We have added a sentence in section 4.4 to further highlight this point. "For instance, a next step could be to include constraints on the dust vertical concentration profile over every location, in order to more accurately estimate dust deposition, and dust concentration at the surface and in 3D. For this, lidar-based retrieval of vertical dust extinction profiles from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) can be combined with the corresponding constraints on dust mass extinction efficiency from this study to obtain constraints on the dust vertical concentration profile."

56. P31 - 125-26 - the bias across the full size range should also be stated.

We have included an additional sentence to address this point.

"Because DustCOMM underestimates the measurements for $D \le 0.5\mu m$, it shows a more negative bias (~50% more) over the full size range (between $D = 0.2 - 20\mu m$), although the error is markedly lower (~15 %), when compared to the ensemble of model simulations. Overall for $D \ge 0.5\mu m$, DustCOMM shows a bias against measured size distributions that is significantly less (about 46% less) than for an ensemble of global model simulations."

57. Are there any important dust altitude or seasonal changes in DustCOMM vs the models?

As Fig. 7 suggests, DustCOMM vertical profiles follow the form of the ensemble of global model simulations, but the fraction of dust mass in each bin is different from that of the model ensemble since the constraints adjust every location and height by the same factor. In addition, similar adjustment to the annually-average dust mass fraction between DustCOMM and model ensemble is also apparent at the seasonally-averaged timescale.

We have added a sentence in section 3.1.2.2 to clarify this: "Finally, similar changes in the spatial variability of the annually-averaged dust mass fraction are apparent in the seasonally-averaged values."

58. Discussion/Conclusion - It would be interesting if the authors could comment on bias of models vs measurements in previous studies (e.g. Hunneus et al., 2011; Evan et al., 2014), and similarities/improvements seen in those studies vs DustCOMM and the model simulations in this study.

Since models used in those studies (e.g. AeroCom in Hunneus et al., 2011) are similar to those used in our study here, they suffer from similar biases and shortcomings. That is the biases in dust properties are associated with biases in dust size distribution, it therefore suggest that better constraints on size distribution as done with DustCOMM should provide a better estimates of these dust properties. Here in section 4.4, we have added the sentences below to highlight this point.

"Furthermore, since recent studies associate much of the biases in dust properties, such as the dust aerosol optical depth, deposition fluxes and surface dust concentration, to model biases in dust size distribution (Evan et al., 2014; Huneeus et al., 2011), DustCOMM estimates can therefore serve as a better alternative. For example, DustCOMM's improved constraints on atmospheric dust loading and dust size distribution could contribute to better estimates of size-resolved dust concentration near the surface (e.g. Whicker et al., 2018). Over the ocean, such constraints on size-resolved dust concentration could potentially be used for constraints on dust deposition fluxes that are more accurate than possible from global model simulations."

59. AOD reanalyses – do the authors combine these into one single reanalysis dataset themselves? This is not really clear.

Yes, they are combined into one single data and the details of this is given in the (new) section 2.2.

Technical Comments

 P3 112 – "The resulting product constrains the climatology of 3-D global atmospheric dust properties on seasonal and annual timescales" – change to "The resulting product constrains the climatology of 3-D global atmospheric dust properties and is provided on seasonal and annual timescales" – to avoid confusion that the authors are constraining the temporal variability of dust properties.

Thanks for this comment. We have changed the sentence accordingly.

2. P3 134-35 – "After correcting..." – unclear – do you mean you combine all models into one multi-model representation?

Yes. We have added the word "multi-model" to make it clear

3. P6 128 – "globally-averaged size distribution" – Kok 2017 or the simulated one? P7 113 - "globally-averaged size distribution" – Kok 2017 or the simulated one?

We meant the Kok et al 2017 or the constrained globally-averaged dust size distribution here. We have changed them to ""constrained globally-averaged dust size distribution". We have also clarified other places where globally-averaged size distribution are mentioned.

4. P9 118 – typo – should be -10 to -4?

Thank you. It is in fact between -10 and 4, but because we realize this can be confusing, we have changed this (and others) to be -10 to 4.

5. Eqn 8 – please provide units for epsilon_tau

We have provided unit for this parameter, and others alike.

6. P10 127 – "atmospheric "column" dust loading"?

Yes. We have included "column".

7. P14 l25 – change to "...of the in-situ emission measurements.."

Thank you for the comment. The sentence now read "The dust MEE is influenced by the uncertainty in the constrained globally-averaged extinction efficiency, which in-turn is partially due to uncertainties in the in-situ emission measurements of index of refraction and dust particle shapes"

Supplement Comments

1. The supplement contains two Figure S1s. The second should be S3 (?).

This is corrected. Thank you.

2. Section 3.1 - 17-8 – mention that it is the AOD which is assimilated.

Yes. We have done that. The sentence now reads.

"For the first time, meteorological and aerosol observations (which include bias-corrected **aerosol optical depth** from MODIS, AVHRR, MISR – over desserts, and ground-based AERONET instruments) are jointly assimilated into MERRA-2..."

 Section 3.3 - "1.1ox1.1o" typo This has been corrected. Now written as 1.1°x1.1° 4. Section 4 – To make this easier to navigate, relate each observational subheading to the numbers on fig S3 (map). Also include the campaign name in the heading for each section. Take care to state for each subheading whether the campaign was ground-based or airborne. Also explain the choice of altitude selection defined in table to where relevant. Please also be aware, and state where necessary, that although a large size range may have been measured, inlet-size effects may have prevented coarser particles from being measured for some campaigns.

Thank you for the comment. We have included in each observational heading the campaign name as well as whether only PSD or MEE is taken or both. For cases where PSD are taken, we have also included sentences explaining the choice of our representative altitude.

5. Section 4 – a subsection on Kandler et al. (2009), as listed in table 2, is missing.

Thank you for pointing this out. We have included a brief description of Kandler et al (2009).

6. Section 4.1 – please make the locations listed consistent with those listed in table 2.

Thank you for the comments. We have corrected where discrepancies occur.

 Section 4.7 – note that these size distributions were not corrected for refractive index. The FSSP was *not* used as it did not operate correctly. Instead the size distribution larger than d=3 microns was taken from a sunphotometer retrieval.

Thank you for pointing this out. We have removed the mention of FSSP, and include a sentence mentioning that the they did not correct for refractive index.

"The size distributions were not corrected for refractive index because they assumed that the refractive index of latex is approximately similar to that of dust."

8. Section 4.10 – Please note that these studies operated instruments behind significant pipework and suffered loss of the majority of coarse particles (e.g. Ryder et al., 2018, Table 1).

Thank you for pointing this out. We have included a sentence in the section stating that: "Because of the aerosol inlet configuration on the aircraft, the measurement of coarse dust were particularly problematic."

9. Section 4.11 – Why is MEE only taken from DODO1 (winter time?). It appears that the MEE for DODO1 is taken from table 4 of Osborne et al. (2008), for the 'AM+CM' case (a value of 0.41). No coarse mode was measured during DODO1 (see McConnell et al., 2008).

The AM+CM DODO1 case in Osborne et al. (2008) was calculated using the coarse mode size distribution from DABEX since none was available from DODO1. This should be stated, or preferably the value from DODO2 used, where coarse mode was measured. Why is only z<1km used for the DODO2 size distributions?

Thank you for this helpful comment. We have clarified this point in the section.

10. Section 4.12 – is the campaign average size distribution used?

These data are taken from their Fig. 8 which represent the composite size distribution for L02 on flight #060519a and L07 on flight #060604a. We have included an additional sentence in this section to clarify this point.

11. Section 4.13 – and also same aircraft as SAMUM1? SAMUM1 also used a high spectral resolution lidar.

They used the Falcon aircraft, which was also used in SAMUM-1. We have added this information to the section.

12. Section 4.14 – 'used the same instrumentation...' – as which paper/campaign? Presumably the same as Kandler et al. (2009) which is missing? It is not clear which instruments the size distribution comes from – but probably because the 2009 section is missing.

Thank you for pointing this out. Yes, the instrumentations are similar to that from SAMUM-1 (Kandler et al., 2009). This detail has been clarified in the section.

13. Section 4.15 – Data from Ryder et al. (2013a – GRL, Canary Islands) is also used in the paper (Figure 3) and should be described here. Please take care to specify whether Ryder et al. (2013a or 2013b) is being cited – both are given in the references as 2013. Comparisons of both can be found in Ryder et al. (2019). "For this study, mean distribution from PCASP and CDP were selected because they were the most credible based on the authors' analysis." – change to ..."based on the authors' analysis over the size range we use here." MEE is not given in Ryder et al. (2013b - ACP), presumably this is taken from Ryder et al. (2013a) (please state). Mean values in Ryder et al. (2013a) are 0.15 for fresh dust or 0.23 for aged dust – these appear much lower than the value plotted in Figure 8 (around 0.3).

Thank you for these very helpful comments. We have re-written this section to include the descriptions of both Ryder et al. (2013a - GRL) and Ryder et al. (2013b - ACP). Indeed, data from both studies were used in this paper. For Ryder et al. (2013a), we obtained the dust size distributions as a function of heights, which were generously given to us by the first author. For Ryder et al. (2013b), we obtained the campaign averaged dust size

distribution already published. Appropriate geographical references have also been noted both in the supplementary document as well as in the main text and figures. For the MEE, we used the averaged values of 0.31+/-0.08 between the reported values for aged dust (0.23) and the SAL categories (0.39).

14. Section 4.16 "above the SAL" – I would expect a dust measurement to be taken 'in' the SAL – is this a typo?

Thank you for pointing this out. It is indeed within the SAL layer. We have made the correction accordingly.

15. Section 4.18 – "The AER-D campaign uses similar instrument as the Fennec 2011 campaign. They use wing-mounted optical particle counters and shadow probes to measure dust sizes between 0.1 and 100 μm diameter." – but additionally this AER-D used cloud imaging probes (CIP15 and 2DS) for size distributions at d>10 microns (which were used in Fennec but were not mentioned in Section 4.15 as the authors did not use the shadow probe data (d>18.5 microns) in this study).

Thank you for the comments. We have included in this section that the imaging probes are also used.

Reply to reviewer #2

This study presents a new dataset, the Dust Constraints from joint Observational- ModellingexperiMental analysis (DustCOMM), which combines in-situ measurements, reanalysis products, and an ensemble of six global model simulations. Particularly, globally-averaged dust size distribution and extinction efficiency from observational and experimental data are used to constrain the DustCOMM products. The annual and seasonal mean products of 3-dimensional (3D) dust size distribution, 3D dust mass extinction efficiency, and 2D dust loading are provided for the time period from 2004 to 2008. It is found the dataset shows a better agreement with measurements than the six-model ensemble in terms of dust size distribution and mass extinction efficiency. This dataset may be used to constrain dust simulation in global models and to study dust impacts on the earth system. The paper is generally well written. The methodology to develop the datasets is thoroughly introduced and related uncertainties are also discussed in detail. I have a few comments would like the authors to address.

We thank the reviewer for the constructive and helpful comments that helped us to further improve the paper.

Major comments:

1. Here globally-averaged dust size distribution is used to obtain 3D dust size distribution. Is it possible to demonstrate that the regional differences in dust size distribution are small? Or have you considered using different dust size distribution for different regions, e.g., by applying regional averaged values to areas where individual measurements are available and the globally-averaged value to areas where measurements are not available? This might provide better spatial constraints on the dataset. Similarly, globally-averaged dust extinction efficiency at 550 nm is used. How large are the spatial differences? Is it possible to give a rough estimation based on available data?

We thank the reviewer for this insightful comment. Assuming globally consistent size distributions and extinction efficiency is indeed one of the main assumptions in this paper.

We used globally-averaged dust size distribution and dust extinction efficiency because the measurements to constrain these parameters on a regional basis across the different dustsource regions are currently insufficient. Since North African dust dominates most of the global dust emission and many of the measurements used to constrain the globallyaveraged values are associated with the North African dust, constraining the regional dust properties with insufficient measurements will likely result in larger uncertainties than estimated in this study. Moreover, since our constraints are applied globally, regional differences in dust size distribution, for example, are assumed to follow the ensemble of the six model simulations. To better clarify this point in the manuscript, we have added additional sentences to the first paragraph of section 4.3: "We used modelling constraints in DustCOMM where observational constraints were either not available or insufficient. For example, modelling constraints are used for the regional differences in dust size distribution and extinction efficiency because the measurements to constrain these parameters on a regional basis across the different dust-source regions are currently insufficient. To further reduce the uncertainty associated with using modelling constraints, we used an ensemble of six model simulations."

 As discussed in the paper, dust aerosol optical depth from the reanalyses largely depends on the models' treatment of the dust cycle, and this adds uncertainties to the DustCOMM. I wonder if you considered using satellite products of dust optical depth, such as level 3 dust optical depth from the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP).

That's a good suggestion. Although our analysis includes both random and systematic errors in dust AOD by incorporating the satellite-based study of Ridley et al. 2016 with the ensemble of reanalysis datasets, we agree that the estimate of the dust AOD likely incurs additional uncertainties associated with model treatments of dust cycle. However, dust extinction retrieval from CALIOP also suffer from several uncertainties, such as weak signal-to-noise ratio during daytime versus nighttime retrievals (e.g. Kacenelenbogen et al., 2011; Winker et al., 2013) and erroneous assumption of aerosol extinction-to-backscatter ratio used in the extinction retrieval (e.g. Omar et al., 2009; Mamouri et al., 2013; Nisantzi et al., 2015). In addition, it is unclear how the limited spatial coverage of CALIOP AOD retrieval affects the climatological estimates. Nonetheless, our plan is that future versions of DustCOMM will incorporate extinction profiles from CALIOP in estimating the vertical distribution of dust concentration.

We have added a sentence in section 4.4 to further highlight this point. "For instance, a next step could be to include constraints on the dust vertical concentration profile over every location, in order to more accurately estimate dust deposition, and dust concentration at the surface and in 3D. For this, lidar-based retrieval of vertical dust extinction profiles from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) can be combined with the corresponding constraints on dust mass extinction efficiency from this study to obtain constraints on dust vertical concentration profile."

3. Sections 4-5 show that the new dataset has a better agreement with in-situ measurements than the multi-model mean. I think it is better to add some discussion to emphasize why this dataset is a good complement to the currently available observational data, especially individual measurements. For instance, the global coverage and vertical distribution of the

dust size distribution and mass extinction efficiency of the dataset make it easier to be adapted to global models to constrain simulations or to study global dust impacts.

We discuss these possible uses of DustCOMM in section 4.4 (formerly section 5.4). There we discuss how DustCOMM can be used to constrain dust impacts in models, and how it can also serve as alternative to global model simulations.

Minor comments:

1. Lines 21-23, page 2, this can be a bit misleading since both small are large dust particles absorb and scatter shortwave and longwave radiation.

Yes, the result of the combined absorption and scattering (i.e. extinction) is cooling for fine dust and warming for coarse dust. For clarity, we now specify that "fine dust **predominantly** cools the climate system by extinguishing shortwave (SW) radiation ..."

2. Line 6, page 3, "To address this problem", not sure the dataset would be able to address the "numerous important biases". You may want to point out a few detailed problems.

Thank you for the comment. Our dataset does address the problem showing significant improvement over model simulation, although it does not completely eliminate the biases. We have re-written this part as: "To address the problem of size and shape biases in models"

3. You may want to add the horizontal and vertical resolutions of the DustCOMM product at someplace in Section 2.

Thank you. We have added a sentence in section 2.3 stating "We estimate all DustCOMM products at 2.5° X 2.0° horizontal grid with 35 levels that is up to 100 hPa."

- 4. Line 17, page 11, what time period does the "climatology" refer to? The climatology is between 2004 and 2008. We have added this to the sentence
- 5. Section 3.1, are all the model results interpolated to the same horizontal and vertical grids? And what's the resolution?

Yes. We stated in the last paragraph of section 2.1 (formerly 3.1) that "...we interpolated seasonal and annual climatologies of these dust properties to a common resolution of approximately 2.5° by 2.0° spatial resolution, with 35 levels from the surface to 100 hPa"

- Line 35, page 16, why the JRAero in a different time period is used? It's not available from 2004 to 2008?
 Yes. JRAero is only available between 2011-2015.
- Line 16, page 21, do you refer to Fig. 4 instead of Fig. S4? Thanks for the comment. We have deleted the statement in parenthesis because it is no longer available in the supplementary document.
- 8. Line 16-18, page 21, can you please add some discussion about why the DustCOMM has a larger bias than model ensemble for $D \le 0.5 \mu m$?

Thank you for the comment. We discussed this in section 4.1 (formerly 5.1). First, we highlighted that "DustCOMM's underestimation of dust with $D \le 0.5\mu m$ may be caused by the contamination of the measured size distributions by other aerosol species for $D \le 0.5\mu m$." Second, we discussed that "the constraint on the globally-averaged dust size distribution could also underestimate the contribution from dust with $D \le 0.5\mu m$."

9. Line 19-24, page 23, "...regardless of the season and location", except Sde Boker, Israel.

Although the statement is true overall, we have removed this part of the sentence.

- Table 1, please remove "deg" in column four, since you already added a degree symbol there.
 Thank you. We have done just that.
- 11. Figs. 2-3, can you please add latitude, longitude, and location of the measurements on the top of each plot? Or you may number the measurements listed in Table 2 and then simply list the corresponding numbers in the figure.

Thank you for the comment. We have instead added the names of the campaign listed in Table 2 on each plot.

- 12. Fig. 5, is it possible to add a globally averaged PSD and its PDF to the plot? Since DustCOMM globally-averaged values are forced to the globally-averaged values from Kok et al. 2017, the globally-averaged PSD the reviewer requested is also represented by the black lines.
- 13. Fig. 6, why is dust mass fraction for $D= 0.2-2.5 \mu m$ high over the ITCZ? Is this consistent with observations?

Because dust concentration is usually low over the ITCZ region, they are dominated by the fine particles.

14. Fig. 7, it would be more interesting to show individual model results (as in Fig. 5) instead of multi-model results.

Thank you for the comment. We have included the individual model results in Figure 7.

15. Fig. 8, why do some blue dots have a light blue outline?

Those are cases with different measurement type as explained in section 2.4 (formerly 3.3). To avoid confusion, we have removed this from the figure.

References

Baumgardner, D., Abel, S. J., Axisa, D., Cotton, R., Crosier, J., Field, P., Gurganus, C., Heymsfield, A., Korolev, A., Krämer, M., Lawson, P., McFarquhar, G., Ulanowski, Z. and Um, J.: Cloud Ice Properties: In Situ Measurement Challenges, Meteorol. Monogr., 58, 9.1-9.23, doi:10.1175/AMSMONOGRAPHS-D-16-0011.1, 2017.

Di Biagio, C., Formenti, P., Balkanski, Y., Caponi, L., Cazaunau, M., Pangui, E., Journet, E., Nowak, S., Caquineau, S., Andreae, M. O., Kandler, K., Saeed, T., Piketh, S., Seibert, D., Williams, E. and Doussin, J.-F.: Global scale variability of the mineral dust long-wave refractive index: a new dataset of in situ measurements for climate modeling and remote sensing, Atmos. Chem. Phys., 17(3), 1901–1929, doi:10.5194/acp-17-1901-2017, 2017.

Di Biagio, C., Formenti, P., Balkanski, Y., Caponi, L., Cazaunau, M., Pangui, E., Journet, E., Nowak, S., Andreae, M. O., Kandler, K., Saeed, T., Piketh, S., Seibert, D., Williams, E. and Doussin, J.-F.: Complex refractive indices and single scattering albedo of global dust aerosols in the shortwave spectrum and relationship to iron content and size, Atmos. Chem. Phys. Discuss., 1–42, doi:10.5194/acp-2019-145, 2019.

Chou, C., Formenti, P., Maille, M., Ausset, P., Helas, G., Harrison, M. and Osborne, S.: Size distribution, shape, and composition of mineral dust aerosols collected during the African Monsoon Multidisciplinary Analysis Special Observation Period 0: Dust and Biomass-Burning Experiment field campaign in Niger, January 2006, J. Geophys. Res., 113(D23), D00C10, doi:10.1029/2008JD009897, 2008.

Dufresne, J.-L., Gautier, C., Ricchiazzi, P. and Fouquart, Y.: Longwave Scattering Effects of Mineral Aerosols, J. Atmos. Sci., 59(12), 1959–1966, doi:10.1175/1520-0469(2002)059<1959:LSEOMA>2.0.CO;2, 2002.

Evan, A. T., Flamant, C., Fiedler, S. and Doherty, O.: An analysis of aeolian dust in climate models, Geophys. Res. Lett., 41(16), 5996–6001, doi:10.1002/2014GL060545, 2014. Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, a., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J. J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T. and Zender, C. S.: Global dust model intercomparison in AeroCom phase i, Atmos. Chem. Phys., 11(15), 7781–7816, doi:10.5194/acp-11-7781-2011, 2011.

Kacenelenbogen, M., Vaughan, M. A., Redemann, J., Hoff, R. M., Rogers, R. R., Ferrare, R. A., Russell, P. B., Hostetler, C. A., Hair, J. W. and Holben, B. N.: An accuracy assessment of the CALIOP/CALIPSO version 2/version 3 daytime aerosol extinction product based on a detailed multi-sensor, multi-platform case study, Atmos. Chem. Phys., 11(8), 3981–4000, doi:10.5194/acp-11-3981-2011, 2011.

Kandler, K., SchüTZ, L., Deutscher, C., Ebert, M., Hofmann, H., JäCKEL, S., Jaenicke, R., Knippertz, P., Lieke, K., Massling, A., Petzold, A., Schladitz, A., Weinzierl, B., Wiedensohler, A., Zorn, S. and Weinbruch1, S.: Size distribution, mass concentration, chemical and mineralogical composition and derived optical parameters of the boundary layer aerosol at Tinfou, Morocco, during SAMUM 2006, Tellus B Chem. Phys. Meteorol., 61(1), 32–50, doi:10.1111/j.1600-0889.2008.00385.x, 2009.

Kok, J. F., Ridley, D. A., Zhou, Q., Miller, R. L., Zhao, C., Heald, C. L., Ward, D. S., Albani, S. and Haustein, K.: Smaller desert dust cooling effect estimated from analysis of dust size and abundance, Nat. Geosci., 10(4), 274–278, doi:10.1038/ngeo2912, 2017.

Li, S.-M., Tang, J., Xue, H. and Toom-Sauntry, D.: Size distribution and estimated optical properties of carbonate, water soluble organic carbon, and sulfate in aerosols at a remote high altitude site in western China, Geophys. Res. Lett., 27(8), 1107–1110,

doi:10.1029/1999GL010929, 2000.

Mahowald, N., Albani, S., Kok, J. F., Engelstaeder, S., Scanza, R., Ward, D. S. and Flanner, M. G.: The size distribution of desert dust aerosols and its impact on the Earth system, Aeolian Res., 15, 53–71 [online] Available from:

http://www.sciencedirect.com/science/article/pii/S1875963713000736 (Accessed 15 December 2017), 2014.

Mamouri, R. E., Ansmann, A., Nisantzi, A., Kokkalis, P., Schwarz, A. and Hadjimitsis, D.: Low Arabian dust extinction-to-backscatter ratio, Geophys. Res. Lett., 40(17), 4762–4766, doi:10.1002/grl.50898, 2013.

Nisantzi, A., Mamouri, R. E., Ansmann, A., Schuster, G. L. and Hadjimitsis, D. G.: Middle East versus Saharan dust extinction-to-backscatter ratios, Atmos. Chem. Phys., 15(12), 7071–7084, doi:10.5194/acp-15-7071-2015, 2015.

Omar, A. H., Winker, D. M., Kittaka, C., Vaughan, M. A., Liu, Z., Hu, Y., Trepte, C. R., Rogers, R. R., Ferrare, R. a., Lee, K. P., Kuehn, R. E. and Hostetler, C. A.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. Ocean. Technol., 26(10), 1994–2014, doi:10.1175/2009JTECHA1231.1, 2009.

Ridley, D. A., Heald, C. L., Kok, J. F. and Zhao, C.: An observationally-constrained estimate of global dust aerosol optical depth, Atmos. Chem. Phys. Discuss., 16(23), 1–31, doi:10.5194/acp-2016-385, 2016.

Ryder, C. L., Highwood, E. J., Rosenberg, P. D., Trembath, J., Brooke, J. K., Bart, M., Dean, A., Crosier, J., Dorsey, J., Brindley, H., Banks, J., Marsham, J. H., McQuaid, J. B., Sodemann, H. and Washington, R.: Optical properties of Saharan dust aerosol and contribution from the coarse mode as measured during the Fennec 2011 aircraft campaign, Atmos. Chem. Phys., 13(1), 303–325, doi:10.5194/acp-13-303-2013, 2013.

Ryder, C. L., Marenco, F., Brooke, J. K., Estelles, V., Cotton, R., Formenti, P., McQuaid, J. B., Price, H. C., Liu, D., Ausset, P., Rosenberg, P., Taylor, J. W., Choularton, T., Bower, K., Coe, H., Gallagher, M., Crosier, J., Lloyd, G., Highwood, E. J. and Murray, B. J.: Coarse mode mineral dust size distributions, composition and optical properties from AER-D aircraft measurements over the Tropical Eastern Atlantic, Atmos. Chem. Phys. Discuss., 1–49, doi:10.5194/acp-2018-739, 2018.

Ryder, C. L., Highwood, E. J., Walser, A., Seibert, P., Philipp, A. and Weinzierl, B.: Coarse and Giant Particles are Ubiquitous in Saharan Dust Export Regions and are Radiatively Significant over the Sahara, Atmos. Chem. Phys. Discuss., 1–36, doi:10.5194/acp-2019-421, 2019. Song, Q., Zhang, Z., Yu, H., Kato, S., Yang, P., Colarco, P., Remer, L. A. and Ryder, C. L.: Net radiative effects of dust in the tropical North Atlantic based on integrated satellite observations

and in situ measurements, Atmos. Chem. Phys., 18(15), 11303–11322, doi:10.5194/acp-18-11303-2018, 2018.

von der Weiden, S.-L., Drewnick, F. and Borrmann, S.: Particle Loss Calculator – a new software tool for the assessment of the performance of aerosol inlet systems, Atmos. Meas. Tech., 2(2), 479–494, doi:10.5194/amt-2-479-2009, 2009.

Weinzierl, B., Petzold, A., Esselborn, M., Wirth, M., Rasp, K., Kandler, K., SchüTZ, L., Koepke, P. and Fiebig, M.: Airborne measurements of dust layer properties, particle size distribution and mixing state of Saharan dust during SAMUM 2006, Tellus B Chem. Phys. Meteorol., 61(1), 96–117, doi:10.1111/j.1600-0889.2008.00392.x, 2009.

Weinzierl, B., Sauer, D., Esselborn, M., Petzold, A., Veira, A., Rose, M., Mund, S., Wirth, M., Ansmann, A., Tesche, M., Gross, S. and Freudenthaler, V.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Cape Verde region—an overview of the airborne in situ and lidar measurements during SAMUM-2, Tellus B Chem. Phys. Meteorol., 63(4), 589–618, doi:10.1111/j.1600-0889.2011.00566.x, 2011.

Weinzierl, B., Ansmann, A., Prospero, J. M., Althausen, D., Benker, N., Chouza, F., Dollner, M., Farrell, D., Fomba, W. K., Freudenthaler, V., Gasteiger, J., Groß, S., Haarig, M., Heinold, B., Kandler, K., Kristensen, T. B., Mayol-Bracero, O. L., Müller, T., Reitebuch, O., Sauer, D., Schäfler, A., Schepanski, K., Spanu, A., Tegen, I., Toledano, C., Walser, A., Weinzierl, B., Ansmann, A., Prospero, J. M., Althausen, D., Benker, N., Chouza, F., Dollner, M., Farrell, D., Fomba, W. K., Freudenthaler, V., Gasteiger, J., Groß, S., Haarig, M., Heinold, B., Kandler, K., Kristensen, T. B., Mayol-Bracero, O. L., Müller, T., Reitebuch, O., Sauer, D., Schäfler, A., Schepanski, K., Spanu, A., Tegen, I., Groß, S., Haarig, M., Heinold, B., Kandler, K., Kristensen, T. B., Mayol-Bracero, O. L., Müller, T., Reitebuch, O., Sauer, D., Schäfler, A., Schepanski, K., Spanu, A., Tegen, I., Toledano, C. and Walser, A.: The Saharan Aerosol Long-Range Transport and Aerosol–Cloud-Interaction Experiment: Overview and Selected Highlights, Bull. Am. Meteorol. Soc., 98(7), 1427–1451, doi:10.1175/BAMS-D-15-00142.1, 2017.
Whicker, C., Adebiyi, A. and Kok, J. F.: Improved Model Estimates of Desert Dust Aerosol Surface Concentration, Am. Geophys. Union, Fall Meet. 2018, Abstr. #A21I-2799 [online] Available from: http://adsabs.harvard.edu/abs/2018AGUFM.A21I2799W (Accessed 19 May 2019), 2018.

Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A. and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, Atmos. Chem. Phys., 13(6), 3345–3361, doi:10.5194/acp-13-3345-2013, 2013.